A. Omitted proofs - Reduction and a lower bound

We now present the reduction that we will use for our QC lower bounds later on.

Theorem 1. ***[Reduction.]*** *Let $\epsilon \in \mathbb{R}_{\geq 0}$ and let T be a binary classification task on $\mathbb{R}^d$ with separable classes. Let ALG be a randomized learning algorithm for T that uses m samples. Then for every $\kappa \in [0,1]$ the following holds:*

$$QC(ALG, T, m, \epsilon, 1/2, \kappa) \geq \log\left(\frac{1-\kappa}{\sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{S\sim\mathcal{D}^m, B\sim\mathcal{B}}\left[\mathcal{E}(S,B,p)\right]}\right),$$

where the event $\mathcal{E}(S,B,p)$ is defined as:

$$\mu(p^{-1}(E(ALG(S,B)))) \geq \frac{AR(ALG(S,B),\epsilon)}{2}.$$

Proof. We first prove the Theorem when ALG is deterministic. Let $\mathcal{A}$ be a q-bounded adversary that performs a successful attack on ALG with respect to $(T, m, \epsilon, 1/2, \kappa)$ (as per Definition 5). We will show that q is lower-bounded by the value from the statement of the Theorem.

The behavior of $\mathcal{A}$ can be represented as a binary tree $\mathcal{T}$ where each non-leaf vertex $v \in \mathcal{T}$ contains a query point $x_v \in \mathbb{R}^d$ and each leaf $l \in \mathcal{T}$ contains an ϵ-perturbation $p_l : \mathbb{R}^d \to \mathbb{R}^d$. Then $\mathcal{A}$ works as follows: it starts in the root r of $\mathcal{T}$ and queries the vertex x_r. Depending on $f(x_r) \stackrel{?}{=} 1$ it proceeds left or right. It continues in this manner, querying the points stored in the visited vertices until it reaches a leaf l. At the leaf it outputs the perturbation function p_l.

Let us partition all possible data sets $S \in (\mathbb{R}^d)^m$ depending on which leaf is reached by $\mathcal{A}$ when interacting with $ALG(S)$. Let $l_1, \ldots, l_n$ be the leaves of $\mathcal{T}$ and $C_1, \ldots, C_n \subseteq (\mathbb{R}^d)^m$ be the respective families of data sets that end up in the corresponding leaves. Let $Z := \{S \in (\mathbb{R}^d)^m : \mathcal{A} \text{ succeeds on } S\}$. By assumption $\mathcal{A}$ is guaranteed to succeed with probability $1-\kappa$, so

$$\mathbb{P}_{S\sim\mathcal{D}^m}[S \in Z] \geq 1-\kappa. \tag{2}$$

Now observe that for every $i \in [n]$ and $S \in C_i \cap Z$

$$\mu(p_{l_i}^{-1}(E(ALG(S)))) \geq \frac{AR(ALG(S),\epsilon)}{2}.$$

In words, for every $S \in C_i \cap Z$ the adversary $\mathcal{A}$ succeeds if at least $\frac{AR(ALG(S),\epsilon)}{2}$ of the probability mass of $\mathcal{D}$ is moved by p_{l_i} into the error set of $ALG(S)$. Thus we get that for every $i \in [n]$:

$$\mathbb{P}_{S\sim\mathcal{D}^m}[S \in C_i \cap Z] \leq \sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{S\sim\mathcal{D}^m}\left[\mu(p^{-1}(E(ALG(S)))) \geq \frac{AR(ALG(S),\epsilon)}{2}\right]. \tag{3}$$

By standard properties of entropy we know that for a discrete random variable W any protocol asking yes-no questions that finds the value of W must on average ask at least $H(W)$ many questions. Let W be a random variable that takes values in $\{1,2,\ldots,n\}$, where for every $i \in [n]$ we have $\mathbb{P}[W=i] := \mathbb{P}_{S\sim\mathcal{D}^m}[S \in C_i \cap Z]/\mathbb{P}_{S\sim\mathcal{D}^m}[S \in Z]$. Note that $\mathcal{A}$'s protocol can be directly used to find a protocol that asks yes-no questions and finds the value of W with at most q queries. It is enough to prove a lower-bound on $H(W)$ as the expected number of questions can only be lower than the maximum number.

Note that by (2) and (3) we get that for every $i \in [n]$:

$$\mathbb{P}[W=i] \leq \frac{1}{1-\kappa} \cdot \sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{S\sim\mathcal{D}^m}\left[\mu(p^{-1}(E(ALG(S)))) \geq \frac{AR(ALG(S),\epsilon)}{2}\right].$$

By properties of entropy we know that $H(W) \geq \log(1/\max_{i\in[n]} \mathbb{P}[W=i])$, so in the end we get that:

$$H(W) \geq \log\left(\frac{1-\kappa}{\sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{S\sim\mathcal{D}^m}\left[\mu(p^{-1}(E(ALG(S)))) \geq \frac{AR(ALG(S),\epsilon)}{2}\right]}\right).$$

The proof for the case when ALG is randomized in analogous. The only difference is that instead of partitioning the space $(\mathbb{R}^d)^m$ we partition the space $(\mathbb{R}^d)^m \times \text{supp}(\mathcal{B})$. □

Remark 2. *For the sake of clarity and consistency with the standard setup we fixed the approximation constant to be equal* $1/2$ *and the data generation process to be* $S \sim \mathcal{D}^m$*. We note however, that Theorem 1 (and its proof with minor changes) is also true for all approximation constants and for general data generation processes. By different generation process we mean anything different from* $S \sim \mathcal{D}^m$*, for instance a case where samples are dependent or where the number of samples* m *is itself a random variable. This distinction will become important in the proof of Theorem 2.*

The following theorem states that if an algorithm ALG applied to a learning task satisfies the following: ALG learns low-risk classifier with constant probability, the adversarial risk is high with constant probability and every point from the support of the distribution is misclassified with small probability then the QC of ALG is high. The core of the proof is the reduction from Theorem 1.

Theorem 4. *For every* $\epsilon \in \mathbb{R}_{\geq 0}, C, \delta, \eta \in \mathbb{R}_+$ *and* T *a binary classification task on* $\mathbb{R}^d$ *with separable classes the following conditions hold. If ALG is a learning algorithm for* T *and satisfies the following properties:*

1. $\forall x \in \mathit{supp}(\mathcal{D}) + B_\epsilon$,
   $\mathbb{P}_{S\sim\mathcal{D}^m}[\mathit{ALG}(S)(x) \neq h(x)] \leq C \cdot \delta$,
2. $\mathbb{P}_{S\sim\mathcal{D}^m}[AR(\mathit{ALG}(S), \epsilon) \geq \eta] \geq 0.99$,
3. $\mathbb{P}_{S\sim\mathcal{D}^m}[R(\mathit{ALG}(S)) \leq \delta] \geq 0.99$,

then:

$$QC(\mathit{ALG}, T, m, \epsilon) \geq \log\left(\frac{\eta}{3 \cdot C \cdot \delta}\right).$$

Proof. Let $p : \mathbb{R}^d \to \mathbb{R}^d$ be an ϵ-perturbation. For simplicity we introduce the notation $\rho := \mathbb{P}_{S\sim\mathcal{D}^m}[AR(ALG(S), \epsilon) \geq \eta \wedge R(ALG(S), \epsilon) \leq \delta]$. We define two new data distributions:

$$\mathcal{D}_1 := \mathcal{D}^m | \left(AR(ALG(S), \epsilon) \geq \eta \wedge R(ALG(S), \epsilon) \leq \delta\right),$$
$$\mathcal{D}_2 := \mathcal{D}^m | \left(AR(ALG(S), \epsilon) < \eta \vee R(ALG(S), \epsilon) > \delta\right).$$

Observe that $\text{supp}(\mathcal{D}_1) \cap \text{supp}(\mathcal{D}_2) = \emptyset$ and:

$$\mathcal{D}^m = \rho \cdot \mathcal{D}_1 + (1 - \rho) \cdot \mathcal{D}_2. \tag{4}$$

Let $\mathcal{A}$ be an adversary that succeeds on $\mathcal{D}^m$ with probability 0.99. By (4) and the union bound $\mathcal{A}$ has to succeed on $\mathcal{D}_1$ with probability of success s that satisfies:

$$\rho \cdot s + (1 - \rho) \geq 0.99,$$

or, equivalently,

$$s \geq \frac{1}{\rho}\left(0.99 - (1 - \rho)\right).$$

By Assumption 2 and 3, this implies

$$s \geq 0.97. \tag{5}$$

Now observe:

$$\begin{aligned}
&\mathbb{E}_{S\sim\mathcal{D}_1}[\mu(p^{-1}(E(ALG(S))))] \\
&= \int_{\text{supp}(\mathcal{D})} \mathbb{P}_{S\sim\mathcal{D}_1}[p(x) \in E(ALG(S))]\, d\mu \\
&= \int_{\text{supp}(\mathcal{D})} \mathbb{P}_{S\sim\mathcal{D}^m}\left[p(x) \in E(ALG(S)) | \left(AR(ALG(S), \epsilon) \geq \eta \wedge R(ALG(S), \epsilon) \leq \delta\right)\right]\, d\mu \\
&= \int_{\text{supp}(\mathcal{D})} \frac{\mathbb{P}_{S\sim\mathcal{D}^m}\left[p(x) \in E(ALG(S)) \cap AR(ALG(S), \epsilon) \geq \eta \cap R(ALG(S), \epsilon) \leq \delta\right]}{\mathbb{P}_{S\sim\mathcal{D}^m}[AR(ALG(S), \epsilon) \geq \eta \wedge R(ALG(S), \epsilon) \leq \delta]}\, d\mu \\
&\leq \int_{\text{supp}(\mathcal{D})} \frac{\mathbb{P}_{S\sim\mathcal{D}^m}\left[p(x) \in E(ALG(S))\right]}{\mathbb{P}_{S\sim\mathcal{D}^m}[AR(ALG(S), \epsilon) \geq \eta \wedge R(ALG(S), \epsilon) \leq \delta]}\, d\mu \\
&\leq (C \cdot \delta)/\rho \\
&\leq \frac{1}{0.98} \cdot C \cdot \delta,
\end{aligned} \tag{6}$$

where the second equality follows from the definition of $\mathcal{D}_1$, third equality follows from the definition of conditioning, first inequality follows from the fact that intersection decreases probability, second inequality is a result of Assumption 1 (which can be applied as $p(x) \in \text{supp}(\mathcal{D}) + B_\epsilon$) and the last inequality is obtained by Assumptions 2, 3 and the union bound. Using (6) we get:

$$\begin{aligned}
&\mathbb{P}_{S\sim\mathcal{D}_1}\left[\mu(p^{-1}(E(ALG(S)))) \geq \frac{AR(ALG(S),\epsilon)}{2}\right] \\
&\leq \frac{2\cdot\mathbb{E}_{S\sim\mathcal{D}_1}[\mu(p^{-1}(E(ALG(S))))]}{AR(ALG(S),\epsilon)} && \text{by Markov inequality} \\
&\leq \frac{2\cdot\frac{1}{0.98}\cdot C\cdot\delta}{\eta} && \text{by (6) and definition of } \mathcal{D}_1 \qquad (7)
\end{aligned}$$

Applying Theorem 1 to (5) and (7) we get that:

$$QC(ALG,T,m,\epsilon) \geq \log\left(\frac{0.97\cdot 0.98\cdot\eta}{2\cdot C\cdot\delta}\right) \geq \log\left(\frac{\eta}{3\cdot C\cdot\delta}\right).$$

□

B. Omitted Proofs - K-NN

Theorem 2. *There exists a function $\lambda : \mathbb{R}^+ \to (0,1)$ such that the 1-Nearest Neighbor (1-NN) algorithm applied to the learning task $T_{intervals}(z)$ satisfies:*

$$QC(\text{1-NN}, T_{intervals}(z), 2m, z/10, 1-\lambda(z), 0.1) \geq \Theta(m),$$

provided that $z = \Omega(1)$.

Proof. For $x \in L_- \cup L_+$ and $\rho \in \mathbb{R}$ we will use $x + \rho$ to denote $x + (\rho, 0)$. Finally, for $x \in L_- \cup L_+$ we will use $g(x)$ to denote the closest point to x in the other interval.

Data generation process. Instead of letting $S \sim \mathcal{D}^{2m}$ we will use a standard trick and employ a Poisson sampling scheme. This will simplify our proof considerably. Specifically, we think of the samples as being generated by two Poisson processes: Let N_- be a homogeneous Poisson process on the line defined by the extension of L_- and N_+ be a independent of N_- homogeneous Poisson process on the line defined by the extension of L_+, both of rate $\lambda = 1$. Then we define $A_- := ([0,m) \times \{0\}) \cap N_-, A_+ := ([0,m) \times \{z\}) \cap N_+$ and finally:

$$S := \{(x,-1) : x \in A_-\} \cup \{(x,+1) : x \in A_+\} \text{ and}$$

$$\widetilde{S} := \{(x,-1) : x \in N_-\} \cup \{(x,+1) : x \in N_+\}.$$

By design we have $\mathbb{E}[|S|] = 2m$ as $|S|$ is distributed according to $\text{Pois}(2m)$. Moreover, using a standard tail bound for a Poisson random variable, we get that for every $t > 0$:

$$\mathbb{P}[||S| - 2m| \geq t] \leq 2e^{-\frac{t^2}{2(2m+t)}}. \tag{8}$$

This means that the size of the dataset generated with the new process is concentrated around $2m$ (with likely deviations of order $\sqrt{m}$). Let $\{x_1^-, x_2^-, \dots\}$ be the points from N_- with non-negative first coordinate ordered in the increasing order and similarly let $\{x_1^+, x_2^+, \dots\}$. Then note that $A_- = \{x_1^-, \dots, x_{|A_-|}^-\}$ and $A_+ = \{x_1^+, \dots, x_{|A_+|}^+\}$. To simplify notation we let $E(S) := E(\text{1-Nearest Neighbor}(S)), E(\widetilde{S}) := E(\text{1-Nearest Neighbor}(\widetilde{S}))$, where we recall that E denotes the error set. Moreover let:

$$x_0^- := \max_{x \in N_-, x<0} x,\ x_0^+ := \max_{x \in N_+, x<0} x$$

We also define the corresponding random variables $\{X_0^-, X_1^-, \dots\}$ and $\{X_0^+, X_1^+, \dots\}$, where for every i we have $x_i^- \sim X_i^-$ and $x_i^+ \sim X_i^+$.

Upper-bounding $\mu(p^{-1}(E(\text{1-Nearest Neighbor}(S))))$. Let p be a $z/10$-perturbation. We analyze only one of the intervals, namely L_+, as the situation for L_- is symmetric. For $i \in \mathbb{N}_+ \cup \{0\}$ let $\widetilde{Z}_i$ be a non-negative random variable defined as:

$$\widetilde{Z}_i := \nu\left(p^{-1}\left(\overline{P}_{x_i^+} \cup \overline{P}_{x_{i+1}^+}\right)\right),$$

where we define for every $(x,y) \in L_+$:

$$\overline{P}_{(x,y)} := \left\{(x',y') \in \mathbb{R}^2 : y' < \frac{1}{2z}(x'-x)^2 + \frac{z}{2}\right\}$$

Note that by construction:

$$\sum_{i=0}^{|A_-|} \widetilde{Z}_i \geq \nu(p^{-1}(E(S)) \cap L_+). \tag{9}$$

We divide $\widetilde{Z}_i$'s into k groups, where k will be chosen later. For $i \in \mathbb{N}_+ \cup \{0\}$ we define:

$$\widetilde{Z}_{i/k}^{i \bmod k} := \widetilde{Z}_i.$$

Let $g \in \{0, \ldots, k-1\}$. We will upper-bound the probability:

$$\mathbb{P}\left[\sum_{i=0}^{\lceil (1+c(z))m/k \rceil} \widetilde{Z}_i^g \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right],$$

where the function $c : \mathbb{R}^+ \to \mathbb{R}^+$ will be defined later.

Let $i \in [\lceil (1 + c(z))m/k \rceil]$ and $x_0^+, x_1^+, \ldots, x_{(i-1)k+g+1}^+ \in \mathbb{R}$ be an increasing sequence such that $x_0^+ < 0 < x_1^+$. Assume that p maximizes $\mathbb{E}\left[\widetilde{Z}_i^g \middle| X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right]$. Note that by construction of $\overline{P}$'s we have the following property. For every $i \in [|A_+|]$ and every $(x, y) \in \overline{P}_{x_i^+}$, $y \geq 1/2$ we have that for every $y' \in [1/2, y]$ $(x, y') \in \overline{P}_{x_i^+}$. Using this fact we can assume without loss of generality that for every $t \in L_+$ we have $p(t) = (x, y), y \leq z$ and $\|p(t) - t\|_2 = z/10$. The reason is that if $p(t)$ is above L_+ we can flip it with respect to L_+ and preserve the distance to t and if $\|p(t) - t\|_2 < z/10$ we can create a new p' that moves t to $p'(t) := (x, y')$, where $(x, y) = p(t), y' < y$ and $\|p'(t) - t\|_2 = z/10$.

For every $t \in L_+$ let:

$$\alpha(t) := \sphericalangle((-1, 0), p(t) - t).$$

Now observe that $p(t) \in \overline{P}_{x_i^+} \cup \overline{P}_{x_{i+1}^+}$ iff $x_i^+ \leq \tau_1$ and $\tau_2 \leq x_{i+1}^+$, where the two threshold can be computed from $p(t)$ or equivalently from t and $\alpha(t)$. We get the following:

$$\begin{aligned}
&\mathbb{P}\left[p(t) \in \overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+} \middle| X_{(i-1)\cdot k+g+1} = x_{(i-1)k+g+1}^+\right] \\
&= \int_{x_{(i-1)k+g+1}^+}^{\tau_1} f_{X_{i\cdot k+g}^+ - X_{(i-1)k+g+1}^+}(x' - x_{(i-1)k+g+1}^+) \cdot e^{-(\tau_2 - x')} dx' \\
&= e^{-\frac{2\sqrt{5}z}{5}\sqrt{5-\cos(\alpha(t))}} \cdot \frac{\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5 - \cos(\alpha(t))})\right]^k}{k!} \cdot e^{-\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))}\right]} \\
&\leq e^{-\frac{4\sqrt{5}z}{5}} \cdot \frac{\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5 - \cos(\alpha(t))})\right]^k}{k!} \cdot e^{-\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))}\right]} \qquad (10)
\end{aligned}$$

The first equality follows from the fact that inter-arrival times are independent on L_+. To see the second observe that $f_{X_{i\cdot k+g}^+ - X_{(i-1)k+g+1}^+}$ is the density of Erlang distribution with parameters $(k-1, 1)$ and the formula $t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5 - \cos(\alpha(t))}$ gives the expression for τ_1 and $\frac{2\sqrt{5}z}{5}\sqrt{5 - \cos(\alpha(t))} + \tau_1$ gives the expression for τ_2.

Then we have:

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{Z}_i^g \middle| X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] \\
&= \int_{x_{(i-1)k+g+1}^+}^{\infty} \mathbb{P}_S\left[p(t) \in \overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+} \middle| X_0^+ = x_0^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] dt \\
&= \int_{x_{(i-1)k+g+1}^+}^{\infty} \mathbb{P}_S\left[p(t) \in \overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+} \middle| X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] dt \\
&\leq e^{-\frac{4\sqrt{5}z}{5}} \int_0^{\infty} \frac{\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5 - \cos(\alpha(t))})\right]^k}{k!} \cdot e^{-\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))}\right]} \quad \text{By (10)} \quad (11)
\end{aligned}$$

Now we bound the expression from (11). Note that the range of $\sin$ and $\cos$ is $[-1, 1]$ so:

$$\left|\frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5 - \cos(\alpha(t))})\right| \leq \frac{11z}{10}$$

Function $e^{-x} \cdot \frac{x^k}{k!}$ is increasing on $[-\infty, k]$ and decreasing on $[k, \infty]$ thus

$$\begin{aligned}
&\int_0^\infty \frac{\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))})\right]^k}{k!} \cdot e^{-\left[t - \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))})\right]} \\
&\leq \int_0^{k-\frac{11z}{10}} e^{-(t'+\frac{11z}{10})} \cdot \frac{(t'+\frac{11z}{10})^k}{k!}\, dt' + \int_{k-\frac{11z}{10}}^{k+\frac{11z}{10}} e^{-k} \cdot \frac{k^k}{k!}\, dt' + \int_{k+\frac{11z}{10}} e^{-(t'-\frac{11z}{10})} \cdot \frac{(t'-\frac{11z}{10})^k}{k!}\, dt' \\
&\leq \int_0^\infty e^{-t'} \cdot \frac{t'^k}{k!}\, dt' + \frac{22z}{10} e^{-k} \cdot \frac{k^k}{k!} \\
&\leq 1 + \frac{22z}{10\sqrt{2\pi k}} \qquad (12)
\end{aligned}$$

where the last inequality follows from the fact that the function $e^{-t'} \cdot \frac{t'^k}{k!}$ is the density function of the Erlang distribution with parameters $(k, 1)$ and Stirling factorial bounds.

Combining (11) and (12) we get that:

$$\mathbb{E}\left[\widetilde{Z}_i^g \middle| X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] \leq \left(1 + \frac{z}{\sqrt{k}}\right) e^{-\frac{4\sqrt{5}z}{5}}. \qquad (13)$$

Note that in order for $\widetilde{Z}_i^g \geq 0$ one needs $x_{i\cdot k+g+1}^+ - x_{i\cdot k+g}^+ \geq 2 \cdot \frac{z}{10}(-\sin(\alpha(t)) + 2\sqrt{5}\sqrt{5-\cos(\alpha(t))})$, for $\alpha(t) = 0$. Simplifying this is equivalent to $x_{i\cdot k+g+1}^+ - x_{i\cdot k+g}^+ \geq \frac{4\sqrt{5}z}{5}$. As the lengths of intervals are independent we get that for every $i \in \mathbb{N}_+ \cup \{0\}$:

$$\mathbb{P}\left[\widetilde{Z}_i^g = 0 | X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] \geq 1 - e^{-\frac{4\sqrt{5}z}{5}} \qquad (14)$$

From definition of $\widetilde{Z}_i^g$ we have that $\widetilde{Z}_i^g \leq \nu\left(\left(\overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+}\right) + B_{z/10}\right)$. We will give an upper bound on $\nu\left(\left(\overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+}\right) + B_{z/10}\right)$ depending on $x_{i\cdot k+g+1}^+ - x_{i\cdot k+g}^+$. For simplicity let $l := x_{i\cdot k+g+1}^+ - x_{i\cdot k+g}^+$. Let α^* be the minimizer of $\frac{z}{10}(-\sin(\alpha) + 2\sqrt{5}\sqrt{5-\cos(\alpha)})$ and $x^* := \frac{\sqrt{5}\sqrt{5-\cos(\alpha^*)}z}{5}$. Then for $l \in \left[\frac{4\sqrt{5}z}{5}, 2x^*\right]$ we have that:

$$\nu\left(\left(\overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+}\right) + B_{z/10}\right) = 2\sqrt{\frac{z^2}{100} - \left(\frac{z}{2} - \frac{1}{2z}(l/2)^2\right)^2}. \qquad (15)$$

For $l \in (2x^*, \infty)$ we have:

$$\nu\left(\left(\overline{P}_{x_{i\cdot k+g}^+} \cup \overline{P}_{x_{i\cdot k+g+1}^+}\right) + B_{z/10}\right) = l - 2x^* + \frac{2z}{10}\sin(\alpha^*). \qquad (16)$$

Thus as the length of the intervals are distributed according to the exponential distribution we get that for $l \in \left[\frac{4\sqrt{5}z}{5}, 2x^*\right]$:

$$\mathbb{P}\left[\widetilde{Z}_i^g \geq 2\sqrt{\frac{z^2}{100} - \left(\frac{z}{2} - \frac{1}{2z}(l/2)^2\right)^2} \middle| X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] \leq e^{-l}, \qquad (17)$$

and similarly for $l \in (2x^*, \infty)$:

$$\mathbb{P}\left[\widetilde{Z}_i^g \geq l - 2x^* + \frac{2z}{10}\sin(\alpha^*) \middle| X_0^+ = x_0^+, X_1^+ = x_1^+, \ldots, X_{(i-1)k+g+1}^+ = x_{(i-1)k+g+1}^+\right] \leq e^{-l}, \qquad (18)$$

Now we bound the probability that sum of variables from the g-th group deviates considerably from its expectation. The

idea is to use a method similar to the proof of the Chernoff bound.

$$\begin{aligned}
&\mathbb{P}\left[\sum_{i=0}^{\lceil(1+c(z))m/k\rceil}\widetilde{Z}_i^g \geq \frac{1}{k}\left(1+\frac{\epsilon(z)}{2}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] \\
&\leq \mathbb{P}\left[\widetilde{Z}_0^g \geq \frac{1}{k}\cdot\frac{\epsilon(z)}{4}\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] + \mathbb{P}\left[\sum_{i=1}^{\lceil(1+c(z))m/k\rceil}\widetilde{Z}_i^g \geq \frac{1}{k}\left(1+\frac{\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] \quad \text{By the union bound}
\end{aligned} \tag{19}$$

We bound the two terms from (19) separately. Using (17) we get that for $m \geq \frac{2z}{10}\sin(\alpha^*)\cdot k\cdot\frac{4}{\epsilon(z)}\cdot e^{\frac{4\sqrt{5}z}{5}}$:

$$\mathbb{P}\left[\widetilde{Z}_0^g \geq \frac{1}{k}\cdot\frac{\epsilon(z)}{4}\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] \leq \exp\left(-\frac{1}{k}\cdot\frac{\epsilon(z)}{4}\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m + 2x^* - \frac{2z}{10}\sin(\alpha^*)\right), \tag{20}$$

which implies:

$$\mathbb{P}\left[\widetilde{Z}_0^g \geq \frac{1}{k}\cdot\frac{\epsilon(z)}{4}\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] \leq O_z(1)e^{-\Omega_z(m)} \tag{21}$$

Now we bound the second term from (19). For every $s>0$::

$$\begin{aligned}
&\mathbb{P}\left[\sum_{i=1}^{\lceil(1+c(z))m/k\rceil}\widetilde{Z}_i^g \geq \frac{1}{k}\left(1+\frac{\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right] \\
&\leq \mathbb{P}\left[\exp\left(s\sum_{i=1}^{\lceil(1+c(z))m/k\rceil}\widetilde{Z}_i^g\right) \geq \exp\left(s\cdot\frac{1}{k}\left(1+\frac{\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(s\sum_{i=1}^{\lceil(1+c(z))m/k\rceil}\widetilde{Z}_i^g\right)\right]\cdot\exp\left(-s\cdot\frac{1}{k}\left(1+\frac{\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right) \qquad \text{By Markov inequality} \\
&\leq \mathbb{E}\left[\prod_{i=1}^{\lceil(1+c(z))m/k\rceil}\left[\exp\left(s\cdot\left(\widetilde{Z}_i^g - \frac{1}{1+c}\left(1+\frac{\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]\right]
\end{aligned} \tag{22}$$

Set $c(z) := \frac{1+\frac{\epsilon(z)}{4}}{1+\frac{\epsilon(z)}{8}} - 1$. Using the chain rule we obtain:

$$\begin{aligned}
&\mathbb{E}\left[\prod_{i=1}^{\lceil(1+c(z))m/k\rceil}\left[\exp\left(s\cdot\left(\widetilde{Z}_i^g - \left(1+\frac{\epsilon(z)}{8}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]\right] \\
&\mathbb{E}\left[\prod_{i=1}^{\lceil(1+c(z))m/k\rceil-1}\left[\exp\left(s\cdot\left(\widetilde{Z}_i^g - \left(1+\frac{\epsilon(z)}{8}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]\cdot\mathbb{E}\left[\exp\left(s\cdot\left(\widetilde{Z}_i^g - \left(1+\frac{\epsilon(z)}{8}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\middle|\{\widetilde{Z}_i^g\}_{i=1}^{\lceil 2m/k\rceil-1}\right]\right]
\end{aligned} \tag{23}$$

Using the fact that variables $X_0^+,\ldots,X_{(i-1)k+g+1}^+$ determine values of $\widetilde{Z}_0^g\ldots,\widetilde{Z}_{i-1}^g$ and the bound from (13) holds for all possible realizations of $X_0^+,\ldots,X_{(i-1)k+g+1}^+$ if we maximize the inner conditional expectation of (23) over variables $\widetilde{Z}_i^g$ satisfying property (13) we can get an upper bound on $\mathbb{P}\left[\sum_{i=0}^{\lceil(1+c)m/k\rceil}\widetilde{Z}_i^g \geq \frac{1}{k}\left(1+\frac{\epsilon(z)}{2}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right]$ via (19), (20) and (22). More formally let's consider a family of random variables Z satisfying:

1. $Z \geq 0$
2. $\mathbb{E}[Z] \leq (1+\frac{z}{\sqrt{k}})\cdot e^{-\frac{4\sqrt{5}z}{5}}$,

3. For $l \in \left[\frac{4\sqrt{5}z}{5}, 2x^*\right]$: $\mathbb{P}\left[Z \geq 2\sqrt{\frac{z^2}{100} - \left(\frac{z}{2} - \frac{1}{2z}(l/2)^2\right)^2}\right] \leq e^{-l}$,
4. For $l \in (2x^*, \infty)$: $\mathbb{P}\left[Z \geq l - 2x^* + \frac{2z}{10}\sin(\alpha^*)\right] \leq e^{-l}$.

Consider the following optimization problem.

$$\sup_{Z:Z \text{ satisfies 1, 2, 3 and 4}} \mathbb{E}\left[\exp\left(s \cdot \left(Z - \left(1 + \frac{\epsilon(z)}{8}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]. \tag{24}$$

The supremum of this problem is attained for some Z^* from the family. This is the case as Properties 3 and 4 guarantee that the objective function is bounded. Set $k := \frac{256z^2}{\epsilon(z)^2}$. Observe then that because of Property 2 we have that $\mathbb{E}\left[Z^* - \left(1 + \frac{\epsilon(z)}{8}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right] < 0$. Taylor expanding the function e^{sX} we get that:

$$\mathbb{E}\left[\exp\left(s \cdot \left(Z^* - \left(1 + \frac{\epsilon(z)}{8}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right] = 1 + s \cdot \mathbb{E}\left[Z^* - \left(1 + \frac{\epsilon(z)}{8}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right] + o(s^2).$$

Thus we get that there exists $s^* > 0$ such that:

$$\sup_{Z:Z \text{ satisfies 1, 2, 3 and 4}} \mathbb{E}\left[\exp\left(s^* \cdot \left(Z - \left(1 + \frac{\epsilon(z)}{8}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right] < e^{-\Omega_z(1)}. \tag{25}$$

So combining (21), (23) and (25) we get that :

$$\begin{aligned}
&\mathbb{P}\left[\sum_{i=0}^{\lceil (1+c)m/k \rceil} \widetilde{Z}_i^g \geq \frac{1}{k}\left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] \\
&\leq O_z(1)e^{-\Omega_z(m)} + e^{-\Omega_z((1+c(z))m/k)} \\
&\leq O_z(1)e^{-\Omega_z(m)} && \text{As } k \text{ is a function of } z
\end{aligned} \tag{26}$$

Thus we get that:

$$\begin{aligned}
&\mathbb{P}\left[\nu(p^{-1}(E(S)) \cap L_-) \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] \\
&\leq \mathbb{P}\left[\sum_{i=0}^{|A_-|} \widetilde{Z}_i \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] && \text{By (9)} \\
&\leq \mathbb{P}\left[\left(\sum_{i=0}^{\lceil (1+c(z))m \rceil} \widetilde{Z}_i \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right) \vee (|A_-| > (1+c(z))m)\right] \\
&\leq \mathbb{P}\left[\sum_{i=0}^{\lceil (1+c(z))m \rceil} \widetilde{Z}_i \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] + 2e^{-\frac{(1+c(z))^2m^2}{2(m+(1+c(z))m)}} && \text{Union bound and (8)} \\
&\leq k \cdot O_z(1)e^{-\Omega_z(m)} + 2e^{-\Omega_z(m)} && \text{By (26) and union bound over groups} \\
&\leq \frac{256z^2}{\epsilon(z)^2} \cdot O_z(1)e^{-\Omega_z(m)} + 2e^{-\Omega_z(m)} && \text{By setting of } k \\
&\leq O_z(1)e^{-\Omega_z(m)}
\end{aligned}$$

The above fact together with the union bound over L_- and L_+ gives:

$$\mathbb{P}[\mu(p^{-1}(E(\text{1-Nearest Neighbor}(S)))) \geq \left(1 + \frac{\epsilon(z)}{2}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}] \leq O_z(1)e^{-\Omega_z(m)} \tag{27}$$

Lower-bounding $AR(1$-Nearest Neighbor$(S), z/10)$. We will focus on L_+ as the argument for L_- is analogous. Let $a = (a_1, z), b = (b_1, z) \in A_+$ be two consecutive points from A_+. Note that a parabola defined by:

$$P_a := \left\{ \left(t + a_1, \frac{1}{2z} t^2 + \frac{z}{2} \right) : t \in \mathbb{R} \right\},$$

is exactly the set of points that are equally distant to a and L_-. An analogous parabola can be defined for the point b. Let $P_a^{\rightarrow}$ be the parabola P_a shifted to the right by ρ (to be fixed later). Formally:

$$P_a^{\rightarrow} := \left\{ \left(t + a_1, \frac{1}{2z} (t - \rho)^2 + \frac{z}{2} \right) : t \in \mathbb{R} \right\}.$$

Similarly let:

$$P_b^{\leftarrow} := \left\{ \left(t + b_1, \frac{1}{2z} (t + \rho)^2 + \frac{z}{2} \right) : t \in \mathbb{R} \right\}.$$

We will show if a point $(x, y) \in \mathbb{R}^2$ is below $P_a^{\rightarrow}, P_b^{\leftarrow}$ and $y \le 2z, x \in (a_1, b_1)$ then $(x, y) \in E(S)$ with high probability. More precisely let $(x, y) \in \mathbb{R}^2$ be such that: $x \in (a_1, b_1), y \le \frac{1}{2z}(x - a_1 - \rho)^2 + \frac{z}{2}, y \le \frac{1}{2z}(x - b_1 + \rho)^2 + \frac{z}{2}, y \in \left[\frac{9z}{10}, 2z\right]$. By construction $d((x, y), A_+)$ is obtained at a or b. We have that:

$$\begin{aligned} d(a, (x, y))^2 &\ge (x - a_1)^2 + \left(\frac{z}{2} - \frac{(x - a_1 - \rho)^2}{2z} \right)^2 \\ &= (x - a_1)^2 + \frac{(x - a_1 - \rho)^4}{4z^2} - \frac{(x - a_1 - \rho)^2}{2} + \frac{z^2}{4} \end{aligned} \tag{28}$$

$$\begin{aligned} d(L_-, (x, y))^2 &\le \left(\frac{(x - a_1 - \rho)^2}{2z} + \frac{z}{2} \right)^2 \\ &= \frac{(x - a_1 - \rho)^4}{4z^2} + \frac{(x - a_1 - \rho)^2}{2} + \frac{z^2}{4} \end{aligned} \tag{29}$$

Now if $d(a, (x, y)) > 3z$ then $d(a, (x, y)) - d(L_-, (x, y)) \ge z$ by assumption that $y \le 2z$. Otherwise we have:

$$\begin{aligned} d(a, (x, y)) - d(L_-, (x, y)) &= \frac{d(a, (x, y))^2 - d(L_-, (x, y))^2}{d(a, (x, y)) + d(L_-, (x, y))} \\ &\ge \frac{\rho(2x - 2a_1 - \rho)}{3z + 2z} && \text{By (28), (29)} \\ &\ge \frac{\rho\left(\frac{4z}{\sqrt{5}} - \rho\right)}{5z} && \text{As } y \ge 0.9z \\ &\ge 0.3\rho && \text{As } z > 10\rho \end{aligned} \tag{30}$$

By symmetry an analogous bound holds for $d(b, (x, y)) - d(L_-, (x, y))$.

Observe that if there exist a point $c \in A_-$ such that $c \in [(x - \sqrt{0.1\rho z}, 0), (x + \sqrt{0.1\rho z}, 0)]$ then $d(a, (x, y)) > d(L_-, (x, y))$. That's true because:

$$\begin{aligned} d(c, (x, y)) &\le \sqrt{y^2 + 0.1\rho z} \\ &\le y\sqrt{1 + \frac{0.1\rho z}{y^2}} \\ &\le y\sqrt{1 + \frac{0.13\rho}{z}} && \text{As } y \ge 0.9z \\ &\le y\left(1 + \frac{0.07\rho}{z}\right) \\ &\le y + 0.14\rho && \text{As } y \le 2z, \end{aligned} \tag{31}$$

Noticing that $y = d(L_-, (x, y))$ we get:

$$\begin{aligned} d(a,(x,y)) - d(c,(x,y)) &\geq d(a,(x,y)) - d(L_-,(x,y)) - 0.14\rho && \text{By (31)} \\ &\geq 0.3\rho - 0.14\rho \\ &> 0 \end{aligned} \tag{32}$$

By symmetry we also get that $d(b,(x,y)) > d(L_-,(x,y))$, which also implies that $(x,y) \in E(S)$. Note that:

$$|N_- \cap [(x - \sqrt{0.1\rho z}, 0), (x + \sqrt{0.1\rho z}, 0)]| \sim \text{Pois}(2\sqrt{0.1\rho z}),$$

so $\mathbb{P}[N_- \cap [(x - \sqrt{0.1\rho z}, 0), (x + \sqrt{0.1\rho z}, 0)] \neq \emptyset] = 1 - e^{-2\sqrt{0.1\rho z}} \geq 1 - e^{-0.6\sqrt{\rho z}}$, which gives:

$$\mathbb{P}[(x,y) \in E(S)] \geq 1 - e^{-0.6\sqrt{\rho z}} \tag{33}$$

For $i \in \mathbb{N}_+$ let $\widetilde{Y}_i$ be the random variable defined as:

$$\widetilde{Y}_i := \nu((E(\widetilde{S}) + B_{z/10}) \cap [x_i^+, x_{i+1}^+)),$$

where ν is one dimensional Lebesgue measure on L_+. In words, $\widetilde{Y}_i$ is the random variable that is equal to how much the interval $[x_i^+, x_{i+1}^+)$ contributes to $AR(\text{1-Nearest Neighbor}(\widetilde{S}), z/10)$. Observe that $\widetilde{Y}_i$ is primarily determined by the length of $[x_i^+, x_{i+1}^+)$ as well as where the points of N_- are located with respect to $[x_i^+, x_{i+1}^+)$.

$\widetilde{Y}_i$'s satisfy the following properties:

1. $\widetilde{Y}_i$ is non-negative,
2. For all $l \in \left[\frac{4\sqrt{5}z}{5}, 2x^*\right]$: $\mathbb{P}\left[\widetilde{Y}_i \geq 2\sqrt{\frac{z^2}{100} - \left(\frac{z}{2} - \frac{1}{2z}(l/2)^2\right)^2}\right] \geq e^{-l-2\rho} \cdot (1 - 2e^{-0.6\sqrt{\rho z}})$,
3. For all $l \in (2x^*, \infty)$: $\mathbb{P}\left[\widetilde{Y}_i \geq l - 2x^* + \frac{2z}{10}\sin(\alpha^*)\right] \geq e^{-l-2\rho} \cdot (1 - 2e^{-0.6\sqrt{\rho z}})$,
4. $\widetilde{Y}_i$'s are i.i.d. .

The first property (non-negativity) is true by definition. To see the second and the third (observe similarity to (17) and (18)) consider $P^{\rightarrow}_{x_i^+}, P^{\leftarrow}_{x_{i+1}^+}$ and define $\overline{P}^{\rightarrow}_{x_i^+}$ to be all the points below $P^{\rightarrow}_{x_i^+}$ and $\overline{P}^{\leftarrow}_{x_{i+1}^+}$ analogously. Note that:

$$\nu\left(\left(\left(\overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}\right) \cap E(S)\right) + B_{z/10}\right) \leq \widetilde{Y}_i$$

Moreover:

$$\nu\left(\left(\left(\overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}\right) \cap E(S)\right) + B_{z/10}\right) = \nu\left(\left(\overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}\right) + B_{z/10}\right),$$

as for the equality to hold it is enough for $E(S)$ to contain an interval $[(x,y),(x',y)] \subseteq \overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}$ that certifies $\nu\left(\left(\overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}\right) + B_{z/10}\right)$. This happens if $(x,y),(x',y) \in E(S)$ as then $[(x,y),(x',y)]$ by construction. Finally $(x,y),(x',y) \in E(S)$ with probability at least $1 - 2e^{-0.6\sqrt{\rho z}}$ by (33) and the union bound. Properties two and three follow by observing that $\nu\left(\left(\overline{P}^{\rightarrow}_{x_i^+} \cup \overline{P}^{\leftarrow}_{x_{i+1}^+}\right) + B_{z/10}\right)$ was already computed in (15) and (16). The last property is in turn a consequence of the fact that the inter-arrival times of a Poisson process are i.i.d. and that the points on the "other" line are Poisson as well and independent of the first line.

Using these properties we have that for every $i \in \mathbb{N}_+$:

$$\begin{aligned} \mathbb{E}[\widetilde{Y}_i] &= \int_0^\infty \mathbb{P}[\widetilde{Y}_i > t]dt \\ &\geq (1 - 2e^{-0.6\sqrt{\rho z}}) \cdot \left(\int_0^{\frac{2z}{10}\sin(\alpha^*)} e^{-\frac{2\sqrt{5z^2 - \sqrt{z^4 - 25z^2t^2}}}{\sqrt{5}} - 2\rho} dt + \int_{\frac{2z}{10}\sin(\alpha^*)}^\infty e^{-t - 2x^* + \frac{2z}{10}\sin(\alpha^*) - 2\rho} dt\right) \\ &= (1 - 2e^{-0.6\sqrt{\rho z}}) \cdot e^{-2\rho} \cdot \left(\int_0^{\frac{2z}{10}\sin(\alpha^*)} e^{-\frac{2\sqrt{5z^2 - \sqrt{z^4 - 25z^2t^2}}}{\sqrt{5}}} dt + \int_{\frac{2z}{10}\sin(\alpha^*)}^\infty e^{-t - 2x^* + \frac{2z}{10}\sin(\alpha^*)} dt\right) \end{aligned}$$

Our goal now is to show that $\sum_{i=1}^{(1-c'(z))m}\widetilde{Y}_i \geq \left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m$ with high probability, where the function $c':\mathbb{R}^+\to\mathbb{R}^+$ will be defined later. Similarly to the standard proof of the Chernoff bound, for every $s>0$:

$$\begin{aligned}
&\mathbb{P}\left[\sum_{i=1}^{(1-c'(z))m}\widetilde{Y}_i \leq \left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right]\\
&=\mathbb{P}\left[\exp\left(-s\sum_{i=1}^{(1-c'(z))m}\widetilde{Y}_i\right)\geq \exp\left(-s\cdot\left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right)\right]\\
&\leq \mathbb{E}\left[\exp\left(-s\sum_{i=1}^{(1-c'(z))m}\widetilde{Y}_i\right)\right]\cdot\exp\left(s\cdot\left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right) && \text{by Markov inequality}\\
&=\left(\mathbb{E}\left[\exp\left(-s\widetilde{Y}_1\right)\right]\cdot\exp\left(s\cdot\frac{1}{1-c'(z)}\left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right)\right)^{(1-c'(z))m} && \text{as } \widetilde{Y}_i\text{'s are i.i.d.} \qquad (34)
\end{aligned}$$

Set $c'(z):=1-\frac{1+\frac{2\epsilon(z)}{3}}{1+\frac{3\epsilon(z)}{4}}$. Then the above becomes:

$$\left(\mathbb{E}\left[\exp\left(s\left(-\widetilde{Y}_1+\left(1+\frac{3\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]\right)^{(1-c'(z))m}$$

Taylor expanding the function e^{sX} we get that:

$$\mathbb{E}\left[\exp\left(s\left(-\widetilde{Y}_1+\left(1+\frac{3\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]=1+s\cdot\mathbb{E}\left[-\widetilde{Y}_1+\left(1+\frac{3\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right]+o(s^2). \qquad (35)$$

We will show now that there exists a function $\epsilon:\mathbb{R}^+\to\mathbb{R}^+$ such that for z bigger than a constant:

$$\mathbb{E}[\widetilde{Y}_i]\geq(1+\epsilon(z))\cdot e^{-\frac{4\sqrt{5}z}{5}}.$$

Note that it is equivalent to showing that for z bigger than a constant

$$\max_{z/10>\rho>0}(1-2e^{-0.6\sqrt{\rho z}})\cdot e^{-2\rho}\cdot\left(\int_0^{\frac{2z}{10}\sin(\alpha^*)}e^{-\frac{2\sqrt{5z^2-\sqrt{z^4-25z^2t^2}}}{\sqrt{5}}}dt+\int_{\frac{2z}{10}\sin(\alpha^*)}^{\infty}e^{-t-2x^*+\frac{2z}{10}\sin(\alpha^*)}dt\right)>e^{-\frac{4\sqrt{5}z}{5}}.$$

First observe that $(1-2e^{-0.6\sqrt{\rho z}})\cdot e^{-2\rho}$ can be made arbitrarily close to 1 by setting $\rho:=z^{-1/2}$. Next we lower bound the first integral:

$$\begin{aligned}
&\int_0^{\frac{2z}{10}\sin(\alpha^*)}e^{-\frac{2\sqrt{5z^2-\sqrt{z^4-25z^2t^2}}}{\sqrt{5}}}dt\\
&\geq\int_0^{\sqrt{\frac{z}{5\sqrt{5}}}}e^{-\frac{2\sqrt{5z^2-\sqrt{z^4-25z^2t^2}}}{\sqrt{5}}}dt && \text{For } z \text{ big enough, as } \sin(\alpha^*) \text{ is a constant}\\
&\geq e^{-\frac{4\sqrt{5}z}{5}}\cdot\int_0^{\sqrt{\frac{z}{5\sqrt{5}}}}e^{-\frac{10\sqrt{5}}{4z}t^2}dt && \text{As } \frac{2\sqrt{5z^2-\sqrt{z^4-25z^2t^2}}}{\sqrt{5}}\leq\frac{4\sqrt{5}z}{5}+\frac{10\sqrt{5}}{4z}t^2 \text{ for } t\leq z/5\\
&\geq e^{-\frac{4\sqrt{5}z}{5}}\cdot\sqrt{\frac{z}{5\sqrt{5}}}\cdot\frac{\sqrt{\pi}}{2}\text{erf}(1/\sqrt{2}),
\end{aligned}$$

thus for z big enough we get that $\mathbb{E}[\widetilde{Y}_i]\geq(1+\epsilon(z))\cdot e^{-\frac{4\sqrt{5}z}{5}}$. Using (35) this implies that there exists $s^*>0$ such that $\mathbb{E}\left[\exp\left(s^*\cdot\left(-\widetilde{Y}_1+\left(1+\frac{3\epsilon(z)}{4}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\right)\right)\right]<1$. Using (34) we get then that:

$$\mathbb{P}\left[\sum_{i=1}^{(1-c'(z))m}\widetilde{Y}_i\leq\left(1+\frac{2\epsilon(z)}{3}\right)\cdot e^{-\frac{4\sqrt{5}z}{5}}\cdot m\right]\leq e^{-\Omega_z(m)}. \qquad (36)$$

Now for $i \in [|A_+| - 1]$ let Y_i be the random variable defined as:

$$Y_i := \nu(((E(S) \cap [x_i^+, x_{i+1}^+)) + B_{z/10}) \cap L_-).$$

Notice that for all $i \in [|A_+| - 1]$ we have $Y_i = \widetilde{Y}_i$. Note that by Poisson tail bound we have:

$$\mathbb{P}\left[|A_+| \leq (1 - c'(z))m\right] \leq 2e^{-\frac{(1+c'(z))^2 m^2}{2(m+(1+c'(z))m)}} \leq 2e^{-\Omega_z(m)}. \tag{37}$$

Combining (37) and (36) and the union bound we get that:

$$\begin{aligned}
\mathbb{P}\left[\sum_{i\in[|A_+|-1]} Y_i \leq \left(1 + \frac{2\epsilon(z)}{3}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] &\leq \mathbb{P}\left[(|A_+| \leq (1 - c'(z))m) \vee \left(\sum_{i=1}^{(1-c'(z))m} \widetilde{Y}_i \leq \left(1 + \frac{2\epsilon(z)}{3}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right)\right] \\
&\leq \mathbb{P}[|A_-| \leq m/2] + \mathbb{P}\left[\sum_{i=1}^{(1-c'(z))m} \widetilde{Y}_i \leq \left(1 + \frac{2\epsilon(z)}{3}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}} \cdot m\right] \\
&\leq 2e^{-\Omega_z(m)} + e^{-\Omega_z(m)} \\
&\leq O_z(1)e^{-\Omega_z(m)}.
\end{aligned}$$

Note that we omitted the first and the last interval $\left(\left[0, x_1^-\right) \text{ and } \left[x_{|A_-|}^-, m\right)\right)$. Omitting these intervals is valid as we are deriving a lower bound for $AR(\text{1-Nearest Neighbor}(S), z/10)$. We conclude using the union bound over two intervals L_- and L_+ to obtain:

$$\mathbb{P}\left[AR(\text{1-Nearest Neighbor}(S), z) \leq \left(1 + \frac{2\epsilon(z)}{3}\right) \cdot e^{-\frac{4\sqrt{5}z}{5}}\right] \leq O_z(1)e^{-\Omega_z(m)}. \tag{38}$$

Lower-bounding QC. To prove a lower-bound on the QC of 1-NN applied to this task we will use Theorem 1. This means that we need to upper-bound:

$$\sup_{p:\, z/10\text{-perturbation}} \mathbb{P}_S\left[\mu(p^{-1}(E(\text{1-Nearest Neighbor}(S)))) \geq (1 - \lambda) \cdot AR(\text{1-Nearest Neighbor}(S), z/10)\right],$$

where S is generated from the two independent Poisson processes as described at the beginning of the proof. By Remark 2 we can use Theorem 1 in this case.

Combining (27) and (38) we get that there exists $\lambda : \mathbb{R}^+ \to \mathbb{R}^+$, which can be set to $\lambda(z) = \frac{1+\frac{\epsilon(z)}{2}}{1+\frac{2\epsilon(z)}{3}}$, so that:

$$\begin{aligned}
&\sup_{p:\, z/10\text{-perturbation}} \mathbb{P}_S\left[\mu(p^{-1}(E(S))) \geq (1 - \lambda(z)) \cdot AR(\text{1-Nearest Neighbor}(S), z/10)\right] \\
&\leq O_z(1)e^{-\Omega_z(m)} + O_z(1) \cdot e^{-\Omega_z(m)} \\
&\leq O_z(1)e^{-\Omega_z(m)}.
\end{aligned}$$

This, by Theorem 1, means that:

$$\begin{aligned}
QC(\text{1-Nearest Neighbor}, T_{\text{intervals}}, 2m, z) &\geq \Theta\left(\log\left(\frac{0.1}{O_z(1)e^{-\Omega_z(m)}}\right)\right) \\
&\geq \Theta_z(m).
\end{aligned}$$

□

C. Omitted Proofs - Quadratic Neural Network

We now present proofs of claims from Section 5. Recall that this section deals with quadratic neural nets applied to the concentric spheres dataset.

C.1. QC lower bounds for exponentially small risk

We first use the results from Section 7 to argue that increased accuracy leads to an improved guarantee for robustness. We analyze the QC of QNN for $\epsilon = 0.1$, that is ϵ which is comparable with the separation between the classes. It was experimentally shown in Gilmer et al. (2018) that increasing the sample size for QNN leads to a higher accuracy on the CS dataset. Thus we assume that for some $m \in \mathbb{N}$ the following holds:

$$\mathbb{P}_{S\sim\mathcal{D}^m}[R(\text{QNN}(S)) \in [\delta/2, 2\delta]] \geq 1-\delta. \tag{39}$$

Let $S \in (\mathbb{R}^d)^m$ be such that $R(QNN(S)) \geq \delta/2$. We have that $AR(QNN(S),\epsilon) = \mu(E(QNN(S)) + B_\epsilon) \geq \mu((E(QNN(S)) \cap (S_1^{d-1} \cup S_{1.3}^{d-1})) + B_\epsilon)$. Isoperimetric inequality for spheres states that $\mu((E(QNN(S)) \cap (S_1^{d-1} \cup S_{1.3}^{d-1})) + B_\epsilon)$ is maximized when $(E(QNN(S)) \cap (S_1^{d-1} \cup S_{1.3}^{d-1})$ is a spherical cap of $S_{1.3}^{d-1}$. Using the standard bounds on volumes of spherical caps we get that there exists a universal constant $c > 0$ such that if $\delta \geq 2^{-c\cdot d}$ then $\mu((E(QNN(S)) \cap (S_1^{d-1} \cup S_{1.3}^{d-1})) + B_\epsilon) \geq 1/5$. By (39) it implies that:

$$\mathbb{P}_{S\sim\mathcal{D}^m}[AR(QNN(S),\epsilon) \geq 1/5] \geq 0.99. \tag{40}$$

Moreover note that $\mathcal{D}$ is symmetric and thus it is natural to assume that $\mathbb{P}_{S\sim\mathcal{D}^m}[\text{ALG}(S)(x) \neq h(x)]$ only depends on $\|x\|_2$. Using (39) we bound:

$$\begin{aligned}&\int_{\text{supp}(\mathcal{D})} \mathbb{P}_{S\sim\mathcal{D}^m}[\text{ALG}(S)(x) \neq h(x)]\, d\mu = \mathbb{E}_{S\sim\mathcal{D}^m}[R(\text{QNN})]\\ &\leq 2\delta \cdot \mathbb{P}_{S\sim\mathcal{D}^m}[R(\text{QNN}(S)) \leq 2\delta] + 1\cdot(1 - \mathbb{P}_{S\sim\mathcal{D}^m}[R(\text{QNN}(S)) \leq 2\delta]) \leq 3\delta,\end{aligned}$$

which, assuming that points from S_1^{d-1} are misclassified equally likely as points from $S_{1.3}^{d-1}$ gives that $\forall x \in \text{supp}(\mathcal{D})$, $\mathbb{P}_{S\sim\mathcal{D}^m}[\text{ALG}(S)(x) \neq h(x)] \leq 3\delta$. Finally we assume that there exists a universal constant $C \in \mathbb{R}$ such that $\forall x \in \text{supp}(\mathcal{D}) + B_\epsilon$, $\mathbb{P}_{S\sim\mathcal{D}^m}[\text{ALG}(S)(x) \neq h(x)] \leq 3\cdot C\cdot\delta$. This assumption is consistent with the experimental results from Gilmer et al. (2018) and intuitively it states that points that are ϵ close to the $\text{supp}(\mathcal{D})$ are at most C times more likely to be misclassified as points from the $\text{supp}(\mathcal{D})$.

Combining the properties and applying Theorem 4 we get: $QC(\text{QNN}, \text{CS}, m, 0.1) \geq \log\left(\frac{1/5}{9\cdot C\cdot\delta}\right)$, provided that $\delta \geq 2^{-c\cdot d}$. In words, if QNN has a risk of $2^{-\Omega(k)}$ then it is secure against $\Theta(k)$-bounded adversaries for $\epsilon = 0.1$.

C.2. QC lower bounds for constant risk

We give QC lower bounds for the case where the risk achieved by the network is as large as a constant. To get started, let us formally define the distributions and error sets that we will be concerned with. Recall that for $y \in S_1^{d-1}$ we define $\text{cap}(y, r, \tau) := B_r \cap \{x \in \mathbb{R}^d : \langle x, y\rangle \geq \tau\}$. Let $\tau : [0,1] \to [0,1]$ be such that for every $\delta \in [0,1]$ we have $\nu(\text{cap}(\cdot, 1, \tau(\delta)))/\nu(S_1^{d-1}) = \delta$, where ν is a $d-1$ dimensional measure on the sphere S_1^{d-1}. Recall that for $k \in \mathbb{N}_+$:

$$E_-(k) = \text{cap}(e_1, 1.15, \tau(\delta/k)) \setminus B_{1.15/1.3}$$

$$E_+(k) = \text{cap}(e_1, 1.495, 1.3\tau(\delta/k)) \setminus B_{1.15},$$

Definition 6. ***(Distributions on Spherical Caps)***

- ***Cap***. *Let $\delta \in (0,1)$. We define $Cap(\delta)$ as a distribution on subsets of $B_{1.15}$ defined by the following process: generate $y \sim U[S_1^{d-1}], b \sim U\{-1,1\}$. Return: $cap(y_i, 1.15, \tau(\delta/k))\setminus B_{1.15/1.3}$ if $b = -1$ and $cap(y, 1.495, 1.3\tau(\delta/k))\setminus B_{1.15}$ otherwise.*

- ***Caps***$_k^{i.i.d.}$. *Let* $k \in N_+, \delta \in (0,1)$. *We define* $Caps_k^{i.i.d.}(\delta)$ *as a distribution on subsets of* $\mathbb{R}^d$ *defined by the following process: generate a sequence of random bits* $b_1, \ldots, b_k \sim U\{-1,1\}$, *generate a sequence of random vectors* $y_1, \ldots, y_k \sim U[S_1^{d-1}]$. *Return:*

$$\bigcup_{i:b_i=-1} [cap(y_i, 1.15, \tau(\delta/k)) \setminus B_{1.15/1.3}] \cup \bigcup_{i:b_i=+1} [cap(y_i, 1.495, 1.3\tau(\delta/k)) \setminus B_{1.15}]$$

  In words $Caps_k^{i.i.d.}(\delta)$ *generates* k *i.i.d. randomly rotated sets, each either* $E_-(k)$ *or* $E_+(k)$.

- ***Caps***$_k^{\mathcal{G}}$. *Let* $k \in N_+, \delta \in (0,1)$, $\mathcal{G}$ *be a distribution on* $(S_1^{d-1})^k$. *We define* $Caps_k^{\mathcal{G}}(\delta)$ *as a distribution on subsets of* $\mathbb{R}^d$ *defined by the following process: generate a sequence of random bits* $b_1, \ldots, b_k \sim U\{-1,1\}$, *generate* $y_1, \ldots, y_k \sim \mathcal{G}$, *generate an orthonormal matrix* $M \sim O(d)$. *Return:*

$$\bigcup_{i:b_i=-1} M(cap(y_i, 1.15, \tau(\delta/k)) \setminus B_{1.15/1.3}) \cup \bigcup_{i:b_i=+1} M(cap(y_i, 1.495, 1.3\tau(\delta/k)) \setminus B_{1.15})$$

  In words $Caps_k^{\mathcal{G}}(\delta)$ *generates* k *randomly rotated sets, each either* $E_-(k)$ *or* $E_+(k)$, *where relative positions of normal vectors of the sets are defined by* $\mathcal{G}$.

Note that definitions of Cap, $\text{Caps}_k^{i.i.d.}$ and $\text{Caps}_k^{\mathcal{G}}$ are compatible in the following sense:

Observation 1. *For every* $k \in \mathbb{N}_+, \delta \in (0,1)$*:*

- $Caps_1^{i.i.d.}(\delta) = Cap(\delta)$,
- $Caps_k^{i.i.d.}(\delta) = Caps_k^{U[(S_1^{d-1})^k]}(\delta)$.

In the following lemma we show a reduction from $\text{Cap}_k^{\text{i.i.d}}$ to Cap. This means that we show that if there is an adversary that uses q queries and succeeds on $\text{Cap}_k^{\text{i.i.d}}$ then there exists an adversary that succeeds on Cap and also asks at most q queries. The takeaway from this lemma is that the QC of $\text{Cap}_k^{\text{i.i.d}}$ is no smaller than the QC of Cap. Formally:

Lemma 2 (Reduction from $\text{Caps}_k^{\text{i.i.d.}}$ to Cap). *Let* $k \in \mathbb{N}_+$. *If there exists a* q*-bounded adversary* $\mathcal{A}$ *that succeeds on* $Caps_k^{i.i.d.}(0.01)$ *with approximation constant* $1/2$, *error probability* 0.01 *and* $\epsilon = \tau(0.01/k)$ *then there exists a* q*-bounded adversary* $\mathcal{A}'$ *that succeeds on* $Cap(0.01/k)$ *with approximation constant* $\frac{1}{2k}$, *error probability of* $1 - \frac{1}{3k}$ *and the same* ϵ.

Proof. Algorithm 1 invoked with $\delta = 0.01$ defines the protocol for $\mathcal{A}'$. We will show that this protocol satisfies the statement of the Lemma.

Algorithm 1 EMULATEIID$(f, \mathcal{A}, \delta, k)$ ▷ f is the attacked classifier ▷ $\mathcal{A}$ is an adversary for distribution $\text{Cap}_k^{\text{i.i.d.}}(\delta)$

$y_1, \ldots, y_{k-1} \sim U[S_1^{d-1}]$
$b, b_1, \ldots, b_{k-1} \sim U\{-1,1\}$
for i = 1, . . . , k-1 **do**
  $C_i := \begin{cases} \text{cap}(y_i, 1.15, \tau(\delta/k)) \setminus B_{1.15/1.3} & \text{if } b_i = -1 \\ \text{cap}(y_i, 1.4, 1.3\tau(\delta/k)) \setminus B_{1.15} & \text{if } b_i = +1 \end{cases}$
end for
$p :=$ Simulate $\mathcal{A}$, to query x answer $\begin{cases} f(x) & \text{if } (x \in C_1) \vee \cdots \vee (x \in C_{k-1}) = \text{False} \\ +1 & \text{if } (x \in C_1) \vee \cdots \vee (x \in C_{k-1}) = \text{True and } \|x\|_2 \leq 1.15 \\ -1 & \text{if } (x \in C_1) \vee \cdots \vee (x \in C_{k-1}) = \text{True and } \|x\|_2 > 1.15 \end{cases}$
Return p

At the first sight it might seem that the protocol for $\mathcal{A}'$ uses kq queries. But due to the fact that $k-1$ caps were added artificially the answer to $(k-1)q$ of those queries is known to $\mathcal{A}'$ beforehand. This gives us that $\mathcal{A}'$ is q-bounded as every query of $\mathcal{A}'$ corresponds to a query of $\mathcal{A}$.

For simplicity we will refer to C_i's and C as caps even though they formally are caps with a ball carved out of them. Let $C \subseteq \mathbb{R}^d$ be the hidden cap that was generated from Cap. Observe that:

$$C \cup \bigcup_{i=1}^{k-1} C_i$$

is distributed according to $\text{Cap}_k^{\text{i.i.d.}}$, as $C_1, \ldots, C_{k-1}$ are i.i.d. uniformly random spherical caps, C is a random spherical cap. Thus by the guarantee for $\mathcal{A}$ we know that with probability at least 0.99:

$$\mu\left(p^{-1}\left(C \cup \bigcup_{i=1}^{k-1} C_i\right)\right) \geq \frac{1}{2} \cdot \mu\left(\left(C \cup \bigcup_{i=1}^{k-1} C_i\right) + B_\epsilon\right)$$

As $C, C_1, \ldots, C_k$ are indistinguishable from the point of view of $\mathcal{A}$ we get that with probability at least $0.99/k$:

$$\mu\left(p^{-1}(C)\right) \geq \frac{1}{2k} \cdot \mu\left(\left(C \cup \bigcup_{i=1}^{k-1} C_i\right) + B_\epsilon\right),$$

where $\mu\left(\left(C \cup \bigcup_{i=1}^{k-1} C_i\right) + B_\epsilon\right) \geq 1/4$ with probability $1 - 2^{-k}$ as with this probability there is a cap among $C, C_1, \ldots, C_{k-1}$ that is of the form $textcap(y_i, 1.15, \tau(\delta/k))$. Then $textcap(y_i, 1.15, \tau(\delta/k)) + B_e$ covers $1/4$ of the mass of μ. Thus by the union bound we get that with probability at least $0.99/k - 2^{-k} \geq 1/3k$:

$$\mu\left(p^{-1}(C)\right) \geq \frac{1}{8k},$$

which is equivalent to $\mathcal{A}'$ succeeding on $\text{Cap}(0.01/k)$ with approximation constant of $\frac{1}{2k}$, error probability of at most $1 - \frac{1}{3k}$ for the same ϵ. □

In the next lemma we generalize Lemma 2 to more complex distributions. More formally we show that if there is an adversary that uses q queries and succeeds on $\text{Cap}_k^{\mathcal{G}}$ then there exists an adversary that succeeds on Cap and asks at most kq queries. Formally:

Lemma 3 (Reduction from $\text{Caps}_k^{\mathcal{G}}$ to Cap). *Let $k \in \mathbb{N}_+$ and let $\mathcal{G}$ be any distribution on $(S_1^{d-1})^k$. If there exists a q-bounded adversary $\mathcal{A}$ that succeeds on $\text{Caps}_k^{\mathcal{G}}(0.01)$ with approximation constant $1/2$, error probability 0.01 and $\epsilon = \tau(0.01/k)$ then there exists a kq-bounded adversary $\mathcal{A}'$ that succeeds on $\text{Cap}(0.01/k)$ with approximation constant $\frac{1}{2k}$, error probability 0.76 and the same ϵ.*

Proof. Algorithm 2 defines the protocol for $\mathcal{A}'$. We will show that this protocol satisfies the statement of the lemma.

Algorithm 2 EMULATEGENERAL($f, \mathcal{A}, \mathcal{G}, k$) ▷ f is the attacked classifier
▷ $\mathcal{A}$ is an adversary for distribution $\text{Cap}_k^{\mathcal{G}}(0.01)$

1: $T(x) := \begin{cases} 1.3 \cdot x & \text{if } \|x\|_2 \leq 1.15 \\ x/1.3 & \text{if } \|x\|_2 > 1.15 \end{cases}$
2: $(y_1, \ldots, y_k) \sim \mathcal{G}$
3: **for** i = 1, . . . , k **do**
4: $R_i :=$ rotation such that $R_i(e_1) = y_i$ ▷ Any rotation satisfying the condition is valid
5: **end for**
6: $M \sim O(d)$
7: $b_1, \ldots, b_k \sim U\{-1, 1\}$
8: **for** i = 1, . . . , k **do**
9: $T_i := \begin{cases} T & \text{if } b_i = -1 \\ \text{Id} & \text{if } b_i = +1 \end{cases}$
10: **end for**
11: **for** i = 1, . . . , k **do**
12: $err_i := (f(M(R_i(T_i(x))) = -1 \wedge \|T_i(x)\|_2 > 1.15) \vee (f(M(R_i(T_i(x))) = +1 \wedge \|T_i(x)\|_2 \leq 1.15)$
13: **end for**
14: $err := \bigvee_{i \in [k]} err_i$
15: $p :=$ Simulate $\mathcal{A}$, to x answer $\begin{cases} +1 & \text{if } \|x\|_2 \leq 1.15 \wedge (err = \text{True}) \\ -1 & \text{if } \|x\|_2 \leq 1.15 \wedge (err = \text{False}) \\ -1 & \text{if } \|x\|_2 > 1.15 \wedge (err = \text{True}) \\ +1 & \text{if } \|x\|_2 > 1.15 \wedge (err = \text{False}) \end{cases}$
16: **for** i = 1, . . . , k **do**
17: $p_i := T_i^{-1} \circ R_i^{-1} \circ M^{-1} \circ p \circ M \circ R_i \circ T_i$
18: **end for**
19: **Return** $p' := \frac{1}{k} \sum_{i=1}^k p_i \Big|_{S_1^{d-1}} + \text{Id} \Big|_{S_{1.3}^{d-1}}$ ▷ understood as a linear combination of transport maps

First observe that $\mathcal{A}'$ asks at most kq queries as every query of $\mathcal{A}$ is multiplied k times (see line 2 of Algorithm 2). Observe that p' is a well defined ϵ-perturbation as all p_i's are ϵ-perturbations when restricted to S_1^{d-1}. It follows from the fact that all p_i's are of the form $F^{-1} \circ p \circ F$ where F is a composition of an isometry and either T or the identity. This implies that for all $x \in S_1^{d-1}$ we have $\|x - F^{-1} \circ p \circ F(x)\|_2 \leq \epsilon$. Let C be the hidden spherical cap. Observe that:

$$\bigcup_{i=1}^k M(R_i(T_i(C)))$$

is distributed according to $\text{Cap}_k^{\mathcal{G}}$, as the relative positions of normal vectors of $M(R_1(C)), M(R_2(C)), \ldots, M(R_k(C))$ are distributed according to the process: generate $(y_1', \ldots, y_k') \sim \mathcal{G}$, $M' \sim O(d)$, return $M'((y_1', \ldots, y_k'))$. Thus by the fact that $\mathcal{A}$ succeeds with $\alpha = 1/2$ we know that with probability at least 0.99:

$$\mu\left(p^{-1}\left(\bigcup_{i=1}^k M(R_i(T_i(C)))\right)\right) \geq \frac{1}{2} \cdot \mu\left(\left(\bigcup_{i=1}^k M(R_i(T_i(C)))\right) + B_\epsilon\right).$$

If $C \subseteq B_{1.5}$ then:

$$\mu\left(\left(\frac{1}{k}\sum_{i=1}^k p_i\right)^{-1}(C)\right) = \frac{1}{k}\sum_{i=1}^k \mu\left(p^{-1}(M(R_i(T_i(C))))\right) \geq \frac{1}{k} \cdot \mu\left(p^{-1}\left(\bigcup_{i=1}^k M(R_i(T_i(C)))\right)\right)$$

Combining the two bounds we get that if $C \subseteq B_{1.5}$ then with probability at least 0.99:

$$\mu\left(p'^{-1}(C)\right) \geq \frac{1}{2k} \cdot \mu\left(\left(\bigcup_{i=1}^k M(R_i(T_i(C)))\right) + B_\epsilon\right) \tag{41}$$

We note that with probability at least $(1-2^{-k})\cdot 1/2$ we have that $C \subseteq B_{1.5}$ and there exists $i_0 \in [k]$ such that $T_{i_0}(C) \subseteq B_{1.5}$ as the two events are independent. This means that with probability at least $1/4$:

$$\mu\left(\left(\bigcup_{i=1}^{k} M(R_i(T_i(C)))\right) + B_\epsilon\right) \geq 1/4, \tag{42}$$

as $\mu(M(R_{i_0}(T_{i_0}(C))) + B_\epsilon) = \mu(S_1^{d-1})/2$. Combining (41) and (42) and using the union bound we get that with probability of at least 0.24:

$$\mu(p'^{-1}(C)) \geq \frac{1}{8k},$$

which is equivalent to $\mathcal{A}'$ succeeding on $\mathrm{Cap}(0.01/k)$ with approximation constant of at least $\frac{1}{2k}$, error probability of at most 0.76 for the same ϵ.

□

The following tail bound will be useful.

Lemma 4. *Let X be a zero-mean Gaussian with variance σ^2. Then for every $t \geq 0$:*

$$\frac{1}{\sqrt{2\pi}} \cdot \left(\frac{1}{t} - \frac{1}{t^3}\right) \cdot e^{-t^2/2} \leq \mathbb{P}_{X\sim\mathcal{N}(0,\sigma^2)}[X \geq \sigma \cdot t] \leq \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{t} \cdot e^{-t^2/2}$$

In Lemma 2 and Lemma 3 we showed that the QC of $\mathrm{Cap}_k^{\text{i.i.d.}}$ and $\mathrm{Cap}_k^{\mathcal{G}}$ can be lower-bounded in terms of the QC of Cap. We now show an upper bound $\Theta(d)$ for the the QC of Cap. Further, we give the proof for a lower bound of $\Theta(d)$ for $\mathrm{Cap}(0.01)$. The summary of these results is presented in Table 1.

The upper-bound for Cap, that we are going to show, holds even if we restrict the adversary to be non-adaptive. I.e., the bound holds even if we require the adversary to declare the set of queries up front.

Definition 7 (Non-adaptive query-bounded adversary). *For $\epsilon \in \mathbb{R}_{\geq 0}$ and $f : \mathbb{R}^d \to \{-1, 1\}$ a q-bounded adversary with parameter ϵ is a deterministic algorithm $\mathcal{A}$ that asks at most $q \in \mathbb{N}$ **non-adaptive** queries of the form $f(x) \stackrel{?}{=} 1$ and outputs an ϵ-perturbation $\mathcal{A}(f) : \mathbb{R}^d \to \mathbb{R}^d$.*

Lemma 5 (Upper bound for Cap). *For every d bigger than an absolute constant there exists a non-adaptive $\Theta(d)$-bounded adversary $\mathcal{A}$ that succeeds on $\mathrm{Cap}(0.01)$ with approximation constant $1/2$, error probability 0.01 for $\epsilon = \tau(0.01)$. Moreover $\mathcal{A}$ can be implemented in $O(d^2)$ time.*

Proof. We will first show that there exists a randomized $\mathcal{A}$ that satisfies the statement of the Lemma. This adversary uses Algorithm 3 invoked with $s = \Theta(d)$ as its protocol. Later we will show how to derandomize the protocol. The adversary we design is more produces adversarial examples only on the support of $\mathcal{D}$. This makes the goal of the adversary harder to achieve.

Algorithm 3 CAPADVERSARYRANDOMIZED(f, s, ϵ) ▷ f is the classifier, s is the number of sampled points per sphere ▷ ϵ is the bound on allowed perturbations

1: $Q^- := \{x_1^-, \ldots, x_s^-\}$, where x_i^-'s are i.i.d. $\sim U[S_1^{d-1}]$
2: $Q^+ := \{x_1^+, \ldots, x_s^+\}$, where x_i^+'s are i.i.d. $\sim U[S_{1.3}^{d-1}]$
3: $R := \{x \in Q^- : f(x) = +1\} \cup \{x \in Q^+ : f(x) = -1\}$
4: $v := 1/|R| \cdot \sum_{x\in R} x$
5: $p(x) := \begin{cases} \mathrm{argsup}_{x' \in S_1^{d-1}, \|x-x'\|_2 \leq \epsilon} \langle x' - x, v\rangle & \text{if } x \in S_1^{d-1} \\ \mathrm{argsup}_{x' \in S_{1.3}^{d}, \|x-x'\|_2 \leq \epsilon} \langle x' - x, v\rangle & \text{if } x \in S_{1.3}^{d-1} \end{cases}$
6: **Return** p

Randomized algorithm. First notice that $\mathcal{A}$ is non-adaptive. The queries asked by $\mathcal{A}$ are from $Q^- \cup Q^+$ which were generated (see lines 3 and 3) before any queries were asked and, hence, answered were received. Note further that $\mathcal{A}$ is $\Theta(d)$ bounded as she asks $2 \cdot s = \Theta(d)$ queries.

Run time. We first remark that $\mathcal{A}$ can be implemented in $O(d^2)$ time as the run time is dominated by asking $\Theta(d)$ queries and each vector is in $\mathbb{R}^d$. Formally, p is not returned explicitly but one can imagine that $\mathcal{A}$, after preprocessing that takes $O(d^2)$ time, provides oracle access to p where each evaluation takes time $O(d)$.

Now we prove that $\mathcal{A}$ succeeds with probability 0.99 with approximation constant $1/2$. Let E be the hidden spherical cap that contains all errors of 1-NN and let $u \in S_1^{d-1}$ be its normal vector. First assume that $E \subseteq S_1^{d-1}$. We start by lower-bounding $|R|$. For every $i \in [s]$ let Y_i^- be a random variable which is equal to 1 if $x_i^- \in E$ and 0 otherwise. Then, by the Chernoff bound, we have that for every $\delta < 1$:

$$\mathbb{P}\left[\left|\sum_{i=1}^{s} Y_i^- - \mathbb{E}\left[\sum_{i=1}^{s} Y_i^-\right]\right| > \delta \cdot \mathbb{E}\left[\sum_{i=1}^{s} Y_i^-\right]\right] \leq 2e^{-\frac{\delta^2}{3}\mathbb{E}\left[\sum_{i=1}^{s} Y_i^-\right]}, \tag{43}$$

Noticing that $\mathbb{E}\left[\sum_{i=1}^{s} Y_i^-\right] = s \cdot 0.01$ if we set $\delta = 1/2$ we get that:

$$\mathbb{P}\left[\left|\sum_{i=1}^{s} Y_i^- - s/100\right| > s/200\right] \leq 2e^{-\frac{\delta^2}{3}\cdot s/100} = 2e^{-s/1200}.$$

So with probability at least $1 - 2e^{-s/1200}$ we have that:

$$|R| \geq s/200. \tag{44}$$

Now assume $R = \{z_1, \ldots, z_k\}$ and observe that for every $z \in R$ we have $\langle z, u\rangle \geq \tau(0.01)$ and note that z_i's are i.i.d. uniformly distributed on $\mathrm{cap}(u, 1, \tau(0.01))$. We will model $U[S_1^{d-1}]$ as $\mathcal{N}(0, 1/d)^d$. Then we have that:

$$\begin{aligned}
\langle u, v\rangle &= \frac{1}{k}\left\langle u, \sum_{i=1}^{k} z_i\right\rangle \\
&= \frac{1}{k}\sum_{i=1}^{k}\langle u, z_i\rangle \\
&\geq \tau(0.01) && \text{as } z_i \in R \\
&\geq 2.2/\sqrt{d} && \text{by Lemma 4}
\end{aligned} \tag{45}$$

Moreover if Π is the orthogonal projection onto $u^\perp$ then $\Pi(k \cdot v) \sim \mathcal{N}(0, k/d)^{d-1}$ and $\Pi(v) \sim \mathcal{N}(0, 1/(dk))^{d-1}$ thus:

$$\|\Pi(v)\|_2^2 \sim \frac{1}{dk} \cdot \chi^2(d-1)$$

So, using standard tail bounds for χ^2 distribution, we get that for all $t \in (0, 1)$:

$$\mathbb{P}\left[\left|\frac{dk}{d-1} \cdot \|\Pi(v)\|_2^2 - 1\right| \geq t\right] \leq 2e^{-(d-1)t^2/8} \tag{46}$$

Moreover observe:

$$\begin{aligned}
\left\langle u, \frac{v}{\|v\|_2}\right\rangle &= \frac{\langle u, v\rangle}{\sqrt{\langle u, v\rangle^2 + \|\Pi(v)\|_2^2}} \\
&= \frac{1}{\sqrt{1 + \|\Pi(v)\|_2^2/\langle u, v\rangle^2}} \\
&\geq \frac{1}{\sqrt{1 + \frac{d}{2.2^2} \cdot \|\Pi(v)\|_2^2}} && \text{by (45)}
\end{aligned} \tag{47}$$

Observe that if $\sphericalangle(u, v/\|v\|_2) = 0$ then $\mu(p^{-1}(\mathrm{cap}(u, 1, \tau(0.01)))) \geq 1/5$, as the preimage is exactly $(\mathrm{cap}(u, 1, \tau(0.01)) + B_e) \cap S_1^{d-1}$. Moreover $\mu(p^{-1}(\mathrm{cap}(u, 1, \tau(0.01))))$ is a continuous function of $\sphericalangle(u, v/\|v\|_2)$. Observe that in a coordinate

system where the first basis vector is $v/\|v\|_2$ we have $p(\mu|_{S_1^{d-1}}) \approx (\mathcal{N}(\tau(0.01), 1/d), \mathcal{N}(0, 1/d), \ldots, \mathcal{N}(0, 1/d))$. Assume $\sphericalangle(u, v/\|v\|_2) = \alpha$. We bound:

$$\begin{aligned}
&\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01))))\\
&= \int_{-\infty}^{+\infty} \int_{\frac{\tau(0.01)-x_1\cos(\alpha)}{\sin(\alpha)}}^{+\infty} d/2\pi \cdot e^{-\frac{d}{2}((x_1-\tau(0.01))^2+x_2^2)}\, dx_2\, dx_1\\
&\geq \int_{-\infty}^{+\infty} \int_{\frac{2.4/\sqrt{d}-x_1\cos(\alpha)}{\sin(\alpha)}}^{+\infty} d/2\pi \cdot e^{-\frac{d}{2}((x_1-2.2/\sqrt{d})^2+x_2^2)}\, dx_2\, dx_1 && \text{by Lemma 4}\\
&= \int_{-\infty}^{+\infty} \int_{\frac{2.4-x_1'\cos(\alpha)}{\sin(\alpha)}}^{+\infty} 1/2\pi \cdot e^{-\frac{1}{2}((x_1'-2.2)^2+x_2'^2)}\, dx_2'\, dx_1' && x_1' = x_1\cdot\sqrt{d}, x_2' = x_2\cdot\sqrt{d}
\end{aligned}$$

This means that there exists $\alpha \in (0, \pi/2]$ (independent of d) such that for all v such that $\sphericalangle(u, v/\|v\|_2) \leq \alpha$ we have $\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01)))) \geq 1/6$. Thus by (47) we get that there exists $\xi > 0$ such that if $\|\Pi(v)\|_2^2 \leq \xi/d$ then $\sphericalangle(u, v/\|v\|_2) \leq \alpha$ and in turn $\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01)))) \geq 1/6$.

Setting $k := \frac{2d}{\xi}, t := 1/2$ in (46) we get that with probability at least $1 - e^{-(d-1)/32}$ we have:

$$\|\Pi(v)\|_2^2 \leq \xi/d,$$

which in turn means that with probability at least $1 - e^{-(d-1)/32}$:

$$\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01)))) \geq 1/6. \tag{48}$$

Now combining (44), (48) and the union bound we get that if we set $s := \frac{400d}{\xi}$ then with probability at least $1 - 2e^{-s/200} - e^{-(d-1)/32} = 1 - 2e^{-2d/\xi} - e^{-(d-1)/32}$ we have:

$$\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01)))) \geq 1/6. \tag{49}$$

This probability is bigger than 0.99 if d is bigger than an absolute constant that depends on ξ. Observing that $\mu(E + B_\epsilon) \geq 1/5$ we conclude that if $E \subseteq S_1^{d-1}$ then if $s = \Theta(d)$ then with probability 0.99 $\mathcal{A}$ succeeds on Cap with approximation constant at least $1/2$. To finish the proof one notices that the case $E \subseteq S_{1.3}^{d-1}$ is analogous. The final constant hidden under $\Theta(d)$ for the number of samples is a maximum of constants for S_1^{d-1} and $S_{1.3}^{d-1}$.

Deterministic algorithm. We know show how to derandomize Algorithm 3 to design an adversary $\mathcal{A}_{\text{det}}$. We observe that in Algorithm 3 randomness was used only to generate query points Q^-, Q^+. Instead of generating the query points randomly we use fixed sets. We define the deterministic adversary, $\mathcal{A}_{\text{det}}$, as:

$$\mathcal{A}_{\text{det}} := \textsc{CapAdversaryDeterministic}(\cdot, Q^-, Q^+),$$

for fixed (for a given d) sets Q^-, Q^+ that we define next.

Algorithm 4 $\textsc{CapAdversaryDeterministic}(f, Q^-, Q^+, \epsilon)$

▷ f is the classifier, Q^-, Q^+ are query points on $S_1^{d-1}, S_{1.3}^{d-1}$ respectively

▷ ϵ is the bound on allowed perturbations

1: $R := \{x \in Q^- : f(x) = +1\} \cup \{x \in Q^+ : f(x) = -1\}$

2: $v := 1/|R| \cdot \sum_{x \in R} x$

3: $p(x) := \begin{cases} \text{argsup}_{x' \in S_1^{d-1}, \|x-x'\|_2 \leq \epsilon} \langle x' - x, v\rangle & \text{if } x \in S_1^{d-1} \\ \text{argsup}_{x' \in S_{1.3}^{d}, \|x-x'\|_2 \leq \epsilon} \langle x' - x, v\rangle & \text{if } x \in S_{1.3}^{d-1} \end{cases}$

4: **Return** p

For $u \in S_1^{d-1}$ let:

$$f_u(x) := \begin{cases} -1 & \text{if } x \in S_1^{d-1} \setminus \text{cap}(u, 1, \tau(0.01) \\ +1 & \text{otherwise} \end{cases}.$$

We say that an adversary succeeds on f_u if she, run for f_u, returns p such that $\mu(p^{-1}(\text{cap}(u, 1, \tau(0.01)))) \geq 1/8$. From (49) we know that for every $d \in \mathbb{N}_+$, for every $u \in S_1^{d-1}$:

$$\mathbb{P}_{x_1^-,\ldots,x_{400d/\xi}^- \sim U\left[S_1^{d-1}\right]}\left[\mathcal{A}(f_u, 400d/\xi, \epsilon) \text{ succeeds}\right] \geq 1 - 2e^{-2d/\xi} - e^{-(d-1)/32}$$

Thus we get that for every $d \in \mathbb{N}_+$ that:

$$\mathbb{P}_{u,x_1^-,\ldots,x_{400d/\xi}^- \sim U\left[S_1^{d-1}\right]}\left[\mathcal{A}(f_u, 400d/\xi, \epsilon) \text{ succeeds}\right] \geq 1 - 2e^{-2d/\xi} - e^{-(d-1)/32}$$

And finally, this means that for every $d \in \mathbb{N}_+$ there exists $Q_d^- \subseteq S_1^{d-1}, |Q_d^-| = 400d/\xi$ such that:

$$\mathbb{P}_{u \sim U\left[S_1^{d-1}\right]}\left[\mathcal{A}_{\text{det}}(f_u, Q_d^-, \emptyset, \epsilon) \text{ succeeds}\right] \geq 1 - 2e^{-2d/\xi} - e^{-(d-1)/32}$$

Thus, for d bigger than an absolute constant we get that conditioned on $E \subseteq S_1^{d-1}$ $\mathcal{A}_{\text{det}}$ run with $Q^- = Q_d^-$ succeeds with probability at least 0.99 and asks $|Q_d^-| = \Theta(d)$ queries. Analogous argument shows that for every $d \in \mathbb{N}_+$ there exists $Q_d^+ \subseteq S_{1.3}^{d-1}, |Q_d^+| = \Theta(d)$ such that the following holds. For every d bigger than an absolute constant conditioned on $E \subseteq S_{1.3}^{d-1}$ $\mathcal{A}_{\text{det}}$ run with $Q^+ = Q_d^+$ succeeds with probability at least 0.99. Combining these two results we get that $\mathcal{A}_{\text{det}}$ satisfies statement of the lemma. □

Remark 3. *As we have seen in the proof of Lemma 5 it was more natural to design an adversary that was randomized. We believe that allowing the adversary to use randomness would not change the results in a fundamental way.*

Lemma 1 (**Lower bound for Cap**). *There exists $\lambda > 0$ such that if a q-bounded adversary $\mathcal{A}$ succeeds on Cap(0.01) with approximation constant $\geq 1 - \lambda$, error probability $2/3$ for $\epsilon = \tau(0.01)$. Then*

$$q \geq \Theta(d).$$

Proof. To simplify computations we will sometimes approximate the uniform distribution on S_1^{d-1} as a d-dimensional normal distribution: $\mathcal{N}\left(0, \frac{1}{d}\right)$. This change is valid as the norm of $\mathcal{N}\left(0, \frac{1}{d}\right)^d$ is closely concentrated around 1.

Lower-bounding QC. $\mathcal{A}$ succeeds on Cap(0.01) with probability at least $1/3$ this means that if succeeds with probability at least $1/3$ on either Cap(0.01) conditioned on the error set intersecting S_1^{d-1} or Cap(0.01) conditioned on the error set intersecting $S_{1.3}^{d-1}$. We first prove the result in the first case.

To use Theorem 1 we think that there is an algorithm ALG for which the distribution of errors coincides with Cap(0.01) conditioned on the error set intersecting S_1^{d-1}. Let's call this distribution $Cap'(0.01)$. Note that by definition $AR(ALG(S), \epsilon) = 1/2$. Thus we analyze:

$$\begin{aligned}
&\sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{S \sim \mathcal{D}^m}\left[\mu(p^{-1}(E(ALG(S)))) \geq (1-\delta) \cdot AR(ALG(S), \epsilon)\right] \\
&= \sup_{p\text{: }\epsilon\text{-perturbation}} \mathbb{P}_{E \sim \text{Cap}'(0.01)}\left[\mu(p^{-1}(E)) \geq \frac{1-\delta}{2}\right], \qquad (50)
\end{aligned}$$

for a constant δ that will be fixed later. Let p be an ϵ-perturbation and $y \in S_1^{d-1}$ be such that $\mu(p^{-1}(\text{cap}(y, 1.15, \tau(0.01)) \setminus B_{1.15/1.3})) \geq \frac{1-\delta}{2}$. We will show that for every $x \in S_1^{d-1}$ if $\sphericalangle(y, x) \in \left[\frac{49\pi}{100}, \frac{51\pi}{100}\right]$ then $\mu(p^{-1}(\text{cap}(x, 1.15, \tau(0.01)) \setminus B_{1.15/1.3})) < \frac{1-\delta}{2}$. This will conclude the proof as then:

$$\begin{aligned}
\mathbb{P}_{E \sim \text{Cap}'(0.01)}\left[\mu(p^{-1}(E)) \geq \frac{1-\delta}{2}\right] &\leq 2 \cdot \mu\left(\text{cap}\left(\cdot, 1, \arccos\left(\frac{49\pi}{100}\right)\right)\right) \\
&\leq 2^{-\Omega(d)} \qquad \text{By Lemma 4}
\end{aligned}$$

combined with (50) and Theorem 1 gives the result.

Now let $x \in S_1^{d-1}$ be such that $\sphericalangle(y, x) \in \left[\frac{49\pi}{100}, \frac{51\pi}{100}\right]$. To simplify notation let $C_x := \text{cap}(x, 1.15, \tau(0.01)) \setminus B_{1.15/1.3}, C_y := \text{cap}(y, 1.15, \tau(0.01)) \setminus B_{1.15/1.3}$. Now define:

$$I := \left\{z \in S_1^{d-1} \;\middle|\; d(z, C_y) \leq \epsilon \wedge d(z, C_x) \leq \epsilon \wedge d(z, C_x \cap C_y) > \epsilon\right\},$$

where d denotes the ℓ_2 distance between sets. By Lemma 4 we have:

$$2.2/\sqrt{d} \leq \tau(0.01) \leq 2.4/\sqrt{d} \tag{51}$$

Now observe that:

$$\begin{aligned} I &\supseteq \left\{ z \in S_1^{d-1} \;\middle|\; \langle z, y\rangle \geq 0 \wedge \langle z, x\rangle \geq 0 \wedge \left\langle z, \frac{x+y}{\|x+y\|_2} \right\rangle < \frac{2.2}{\sqrt{d}\cdot \cos(\sphericalangle(y,x)/2)} - \frac{2.4}{\sqrt{d}} \right\} \\ &\supseteq \left\{ z \in S_1^{d-1} \;\middle|\; \langle z, y\rangle \geq 0 \wedge \langle z, x\rangle \geq 0 \wedge \left\langle z, \frac{x+y}{\|x+y\|_2} \right\rangle < \frac{1}{20\sqrt{d}} \right\} =: \widehat{I} \end{aligned} \tag{52}$$

where in the first transition we used (51) and in the second transition we used that $\sphericalangle(y,x) \in \left[\frac{49\pi}{100}, \frac{51\pi}{100}\right]$. Note that $\mu\left(\widehat{I}\right)$ is minimized for $\sphericalangle(y,x) = \frac{51\pi}{100}$. Thus:

$$\begin{aligned} &\mu\left(\widehat{I}\right) \\ &\geq \int_0^\infty \int_{\tan(\pi/100)\cdot x_1}^\infty d/2\pi \cdot e^{-\frac{d}{2}(x_1^2+x_2^2)} \cdot \mathbb{1}\left[x_1 \cos\left(\frac{51\pi}{200}\right) + x_2 \sin\left(\frac{51\pi}{200}\right) < \frac{1}{20\sqrt{d}}\right] dx_2\, dx_1 \\ &= \int_0^\infty \int_{\tan(\pi/100)\cdot x_1}^\infty 1/2\pi \cdot e^{-\frac{1}{2}(x_1^2+x_2^2)} \cdot \mathbb{1}\left[x_1 \cos\left(\frac{51\pi}{200}\right) + x_2 \sin\left(\frac{51\pi}{200}\right) < \frac{1}{20}\right] dx_2\, dx_1, \end{aligned} \tag{53}$$

where the first equality comes from integration by substitution. The integral from (53) is positive, which means that there exists $\delta > 0$ such that $\mu(\widehat{I}) > \delta$. Combining that with (52) we get that $\mu(I) > \delta$. Observe that by definition of I for every $z \in I$ we have that at most one of $p(z) \in C_x$, $p(z) \in C_y$ can be true. Thus, using the fact that $\mu(C_x + B_\epsilon) = \mu(C_y + B_\epsilon) = 1/2$, we get that:

$$\min(\mu(p^{-1}(C_x)), \mu(p^{-1}(C_y))) < 1/2 - \delta/2. \tag{54}$$

This ends the proof as by assumption we know that $\mu(p^{-1}(C_y)) \geq 1/2 - \delta/2$, so by (54) we get that $\mu(p^{-1}(C_x)) < 1/2 - \delta/2$. The proof for the other case is analogous. □

Note that Lemma 1 is equivalent to the statement of Conjecture 1 for $k = 1$.

Conjecture 1 (Cap conjecture). *For every $k \in [d]$ if a q-bounded adversary $\mathcal{A}$ succeeds on Cap$(0.01/k)$ with approximation constant $\geq \frac{1}{2k}$, error probability $\leq 1 - \frac{1}{3k}$ for ϵ such that cap$(\cdot, 1, \tau(0.01/k)) + B_\epsilon =$ cap$(\cdot, 1, 0)$. Then*

$$q \geq \Theta(d).$$

D. Figures

In Figure 2, similar to Figure 1, we present visualizations of decision boundaries for 1-NN. Each subfigure represents a random decision boundary for a different sample $S \sim \mathcal{D}^m$. The aim of these visualizations is to give an intuition for why Theorem 2 is true.

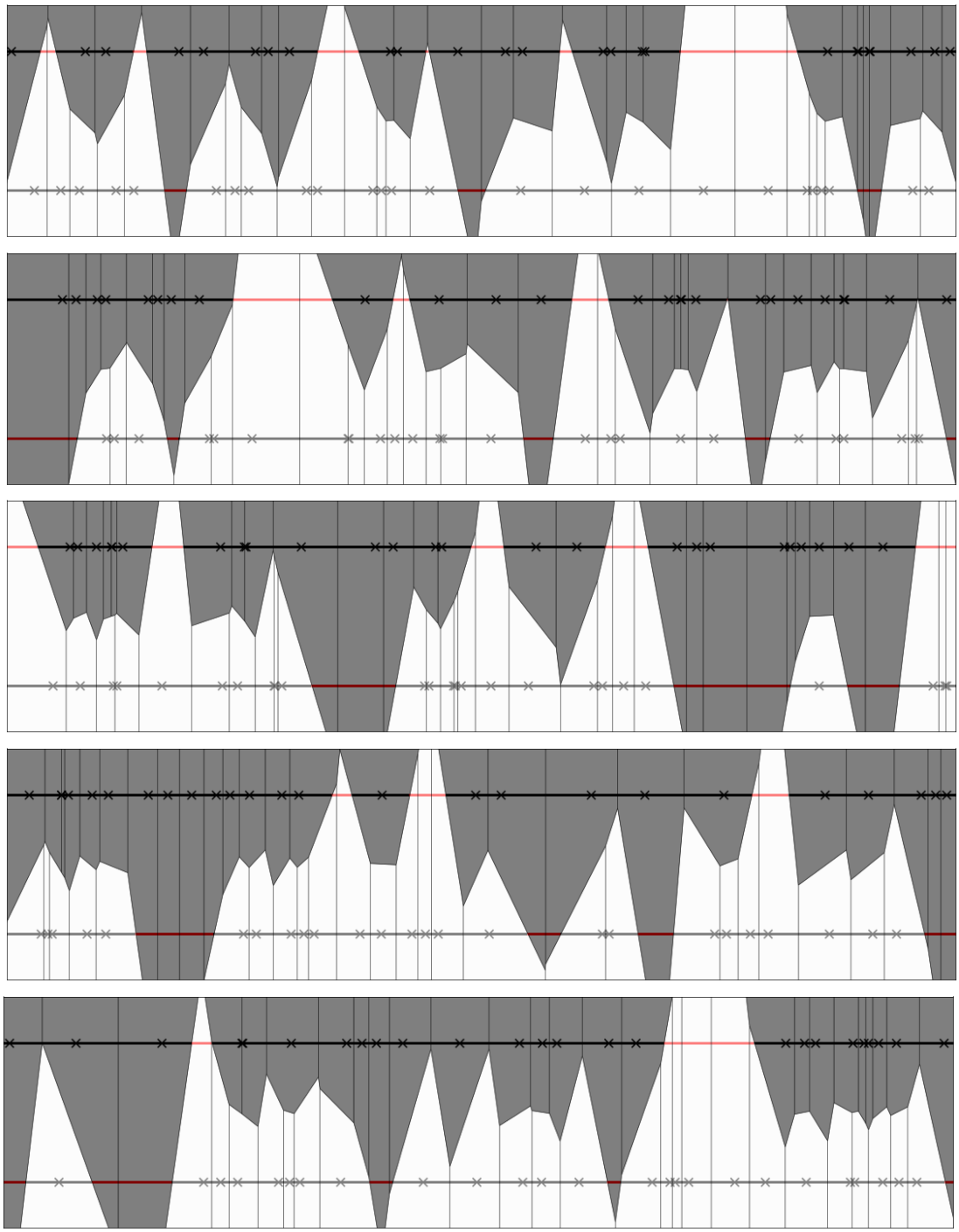

Figure 2. Random decision boundaries of 1-NN for $T_{\text{intervals}}$.