Corruption Robust Active Learning

Yifang Chen, Simon S. Du, Kevin Jamieson
Paul G. Allen School of Computer Science & Engineering
University of Washington, Seattle, WA
{yifangc, ssdu, jamieson}@cs.washington.edu

Abstract

We conduct theoretical studies on streaming-based active learning for binary classification under unknown adversarial label corruptions. In this setting, every time before the learner observes a sample, the adversary decides whether to corrupt the label or not. First, we show that, in a benign corruption setting (which includes the misspecification setting as a special case), with a slight enlargement on the hypothesis elimination threshold, the classical RobustCAL framework can (surprisingly) achieve nearly the same label complexity guarantee as in the non-corrupted setting. However, this algorithm can fail in the general corruption setting. To resolve this drawback, we propose a new algorithm which is provably correct without any assumptions on the presence of corruptions. Furthermore, this algorithm enjoys the minimax label complexity in the non-corrupted setting (which is achieved by RobustCAL) and only requires $\tilde{O}(C_{\text{total}})$ additional labels in the corrupted setting to achieve $O(\varepsilon + \frac{C_{\text{total}}}{n})$, where $\varepsilon$ is the target accuracy, $C_{\text{total}}$ is the total number of corruptions and $n$ is the total number of unlabeled samples.

1 Introduction

An active learning algorithm for binary classification aims to obtain the best hypothesis (classifier) from some given hypothesis set while requesting as few labels as possible. Under some favorable conditions, active learning algorithms can require exponentially fewer labels than passive, random sampling [Hanneke, 2014]. Active learning is ideally suited for applications where large datasets are required for accurate inference, but the cost of paying human annotators to label a dataset is prohibitively large [Joshi et al., 2009, Yang et al., 2015, Beluch et al., 2018].

A bit more formally, for an example space $\mathcal{X}$ (such as a set of images) and label space $\{0, 1\}$ (like whether the image contains a human-made object or not), let $\mathcal{H}$ be a hypothesis class such that for each $h \in \mathcal{H}$, we have $h : \mathcal{X} \rightarrow \{0, 1\}$. After a certain number of labels are requested, the learner will output a target hypothesis $h_{\text{out}} \in \mathcal{H}$. In this paper, we consider the streaming setting where at each time $t$ nature reveals $x_t \sim \mathcal{D}_X$ and an active learning algorithm must make the real-time decision on whether to request the corresponding label $y_t$ or not. Such a streaming setting of active learning is frequently encountered in online environments such as learning a spam filter or fraud detection (i.e., mark as spam/fraudulent and do not request the label, or send to inbox/expert to obtain a label).

This paper is interested in a setting when the requested label $y_t$ is potentially corrupted by an adversary. That is, when requesting the label for some example $x_t \in \mathcal{X}$, if uncorrupted the learner will receive a label drawn according to the “true” conditional label distribution, but if corrupted, the learner will receive a label drawn from an arbitrary distribution decided by an adversary. This setting is challenging because the learner has no a priori knowledge of when or how many corruptions will occur. And if the learner is collecting data adaptively, he may easily be misled into becoming confident in an incorrect belief, collect data based on that belief, and never recover to output an accurate classifier even if the adversary eventually stops serving corrupted labels later. This greatly contrasts with the passive setting (when all labels are observed) where as long as the number of...
corrupts grows sub-linearly over time, the effect of the corruptions will fade and the empirical risk minimizer will converge to an accurate classifier with respect to the uncorrupted labels.

The source of corruptions can come from automatic labeling, non-expert labeling, and, mostly severely, adaptive data poisoning adversaries. Particularly, with the rise of crowdsourcing, it is increasingly feasible for such malicious labelers to enter the system [Miller et al. 2014, Awasthi et al. 2014]. There have been many prior works that consider robust offline training using corrupted labels (i.e., the passive setting) [Hendrycks et al. 2018, Yu et al. 2019]. Correspondingly, related corruption settings have also been considered in online learning [Gupta et al. 2019, Zimmert and Seldin 2019, Wei et al. 2020] and reinforcement learning [Lykouris et al. 2020, Chen et al. 2021a, Wei et al. 2021]. However, there is a striking lack of such literature in the active learning area. Existing disagreement-based active learning algorithms nearly achieve the minimax label complexity for a given target accuracy when labels are trusted [Hanneke 2014], but they fail to deal with the case where the labels are potentially corrupted.

**Our contributions:** In this paper, we study active learning in the agnostic, streaming setting where an unknown number of labels are potentially corrupted by an adversary. We begin with the performance of existing baseline algorithms.

- Firstly, we analyze the performance of empirical risk minimization (ERM) for the passive setting where all labels are observed, which will output an \( \left( \varepsilon + \frac{R^*C_{\text{total}}}{n} \right) \)-optimal hypothesis as long as \( n \gtrapprox \frac{1}{\varepsilon} + \frac{R^*}{\varepsilon} \), where \( R^* \) is the risk of best hypothesis. This result serves as a benchmark for the following active learning results (Section 3).

If we assume that the disagreement coefficient, a quantity that characterizes the sample complexity of active learning algorithms [Hanneke 2014], is some constant, then we obtain the following results for active learning.

- Secondly, we analyze the performance of a standard active learning algorithm called RobustCAL [Balcan et al. 2009, Dasgupta et al. 2007, Hanneke 2014] under a benign assumption on the corruptions (misspecification model is a special case under that assumption). We show that, by slightly enlarging the hypothesis elimination threshold, this algorithm can achieve almost the same label complexity as in the non-corrupted setting. That is, the algorithm will output an \( \tilde{O}(\varepsilon + \frac{R^*C_{\text{total}}}{n}) \)-optimal hypothesis as long as \( n \gtrapprox \frac{R^*}{\varepsilon^2} + \frac{1}{\varepsilon} \) with at most \( \tilde{O}(R^*n + \log(n)) \) number of labels (Section 4).

- Finally and most importantly, in the general corruption case without any assumptions on how corruptions allocated, we propose a new algorithm that matches RobustCAL in the non-corrupted case and only requires \( \tilde{O}(C_{\text{total}}) \) additional labels in the corrupted setting. That is, the algorithm will output an \( \left( \varepsilon + \frac{C_{\text{total}}}{n} \right) \)-optimal hypothesis as long as \( n \gtrapprox \frac{1}{C_{\text{total}}} \) with at most \( \tilde{O} ((R^*)^2 n + \log(n) + C_{\text{total}}) \) number of labels. Besides, this algorithm also enjoys an improved bound under a benign assumption on the corruptions. That is, the algorithm will output an \( \left( \varepsilon + \frac{R^*C_{\text{total}}}{n} \right) \)-optimal hypothesis with at most \( \tilde{O} ((R^*)^2 n + \log(n) + R^*C_{\text{total}}) \) number of labels (Section 5).

Note that \( C_{\text{total}} \) can be regarded as a fixed budget or can be increasing with incoming samples. In the latter case, \( C_{\text{total}} \) in the second and third result can be different since they may require different \( n \). Detailed comparison between these two results will be discussed in corresponding sections.

**Related work:** For nearly as long as researchers have studied how well classifiers generalize beyond their performance on a finite labelled dataset, they have also been trying to understand how to minimize the potentially expensive labeling burden. Consequently, the field of active learning that aims to learn a classifier using as few annotated labels as possible by selecting examples sequentially is also somewhat mature [Settles 2011, Hanneke 2014]. Here, we focus on just the agnostic, streaming setting where there is no relationship assumed a priori between the hypothesis class and the example-label pairs provided by Nature. More than a decade ago, a landmark algorithm we call RobustCAL was developed for the agnostic, streaming setting and analyzed by a number of authors that obtains nearly minimax performance [Balcan et al. 2009, Dasgupta et al. 2007, Hanneke 2014].
The performance of RobustCAL is characterized by a quantity known as the disagreement coefficient that can be large, but in many favorable situations can be bounded by a constant $\theta^*$ which we assume is the case here. In particular, for any $\epsilon > 0$, once Nature has offered RobustCAL $n$ unlabeled samples, RobustCAL promises to return a classifier with error at most $\sqrt{R^* \log(|\mathcal{H}|)/n + \log(|\mathcal{H}|)/n}$ and requests just $nR^* + \theta^*(\sqrt{nR^* \log(|\mathcal{H}|)} + \log(|\mathcal{H}|))$ labels with high probability. Said another way, RobustCAL returns an $\epsilon$-good classifier after requesting just $\theta^*(nR^*/\epsilon^2 + \log(1/\epsilon))$ labels. If $\theta$ is treated as an absolute constant, this label complexity is minimax optimal [Hanneke, 2014]. While there exist algorithms with other favorable properties and superior performance under special distributional assumptions (c.f., Zhang and Chaudhuri [2014], Koltchinskii [2010], Balcan et al. [2007], Balcan and Long [2013], Huang et al. [2015]), we use RobustCAL as our benchmark in the uncorrupted setting. We note that since RobustCAL is computationally inefficient for many classifier classes of interest, a number of works have addressed the issue at the cost of a higher label sample complexity [Beygelzimer et al., 2009, 2010, Hsu, 2010, Krishnamurthy et al., 2017] or higher unlabeled sample complexity [Huang et al., 2015]. Our own work, like RobustCAL, is also not computationally efficient but could benefit from the ideas in these works as well.

To the best of our knowledge, there are few works that address the non-IID active learning setting, such as the corrupted setting of this paper. Nevertheless, [Miller et al., 2014] describes the need for robust active learning algorithms and the many potential attack models. While some applied works have proposed heuristics for active learning algorithms that are robust to an adversary [Deng et al., 2018, Pi et al., 2016], we are not aware of any that are provably robust in the sense defined in this paper. Active learning in crowd-sourcing settings where labels are provided by a pool of a varying quality of annotators, some active learning algorithms have attempted to avoid and down-weight poorly performing annotators, but these models are more stochastic than adversarial [Khetan and Oh, 2016]. The problem of selective sampling or online domain adaptation studies the setting where $P(Y_t = 1|X_t = x)$ remains fixed, but $P(X_t = x)$ drifts and the active learner aims to compete with the best online predictor that observes all labels [Yang, 2011, Dekel et al., 2012, Hanneke and Yang, 2021, Chen et al., 2021b]. Another relevant line of work considers the case where the distribution of the examples drifts over time (i.e., $P(X_t = x)$) [Rai et al., 2010] or the label proportions have changed (i.e., $P(Y_t = 1)$) [Zhao et al., 2021], but the learner is aware of the time when the change has occurred and needs to adapt. These setting are incomparable to our own.

Despite the limited literature in active learning, there have been many existing corruption-related works in the related problem areas of multi-arm bandits (MAB), linear bandits and episodic reinforcement learning. To be specific, for MAB, Gupta et al. [2019] achieves $\tilde{O}(\sum_{a \neq a^*} \frac{1}{n^2} + KC)$ by adopting a sampling strategy based on the estimated gap instead of eliminating arms permanently. Our proposed algorithm is inspired by this “soft elimination” technique and requests labels based on the estimated gap of each hypothesis. Later, Zimmert and Seldin [2019] achieves a near-optimal result $\tilde{O} \left( \frac{1}{\sqrt{\sum_{a \neq a^*} \frac{C}{\theta_{a^*}^2}}} + \sqrt{\sum_{a \neq a^*} \frac{C}{\theta_{a^*}}^2} \right)$ in MAB by using Follow-the-Regularized Leader (FTRL) with Tsallis Entropy. How to apply the FTRL technique in active learning, however, remains an open problem. Besides MAB, Lee et al. [2021] achieves $\tilde{O} \left( \text{GapComplexity} + C \right)$ in stochastic linear bandits. We note that the linear bandits papers of Lee et al. [2021] and Camilleri et al. [2021] both leverage the Catoni estimator that we have repurposed for robust gap estimation in our algorithm. Finally, in the episodic reinforcement learning, Lykouris et al. [2020] achieves $\tilde{O} \left( C \cdot \text{GapComplexity} + C^2 \right)$ in non-tabular RL and Chen et al. [2021a] achieves $\tilde{O} \left( \text{PolicyGapComplexity} + C^2 \right)$ in tabular RL. Very recently, Wei et al. [2021] obtains an $\text{GapComplexity} + C$ bounds for linear MDP.

## 2 Preliminaries

**General protocol:** A hypothesis class $\mathcal{H}$ is given to the learner such that for each $h \in \mathcal{H}$ we have $h : \mathcal{X} \to \{0, 1\}$. Before the start of the game, Nature will draw $n$ unlabeled samples in total. At each time $t \in \{1, \ldots, n\}$, nature draws $(x_t, y_t) \in \mathcal{X} \times \{0, 1\}$ independently from a joint distribution $D_t$, the learner observes just $x_t$ and chooses whether to request $y_t$ or not. Note that in this paper, we assume $\mathcal{X}$ is countable, but it can be directly extended to uncountable case. Next, We denote the expected risk of a classifier $h \in \mathcal{H}$ under any distribution $D$ as $R_D(h) = \mathbb{E}_{x,y \sim D} \left( 1 \{ h(x) \neq y \} \right)$, the marginalized distribution of $x$ as $\nu$ and probability of $y = 1$ given $x$ and $D$ as $\nu^y$. Finally we define $\rho_D(h, h') = \mathbb{E}_{x \sim \nu} 1 \{ h(x) \neq h'(x) \}$.  


**Uncorrupted model:** In the traditional uncorrupted setting, there exists a fixed underlying distribution $D_*$ where each $(x_t, y_t)$ is drawn from this i.i.d distribution. Correspondingly, we define the marginalized distribution of $x$ as $\nu_*$ and probability of $y = 1$ given $x$ and $D_*$ as $\eta_*^x$.

**Oblivious and non-oblivious adversary model:** In the corrupted setting, the label at time $t$ is corrupted if $(x_t, y_t)$ is drawn from some corrupted distribution $D_t$ that differs from the base $D_*$. At the start of the game, an oblivious adversary will choose a sequence of functions $\eta_t^x : \mathcal{X} \to [0, 1]$ for all $t \in \{1, \ldots, n\}$. The corruption level at time $t$ is measured as

$$c_t = \max_{x \in \mathcal{X}} |\eta_t^x - \eta_*^x|,$$

and the amount of corruptions during any time interval $\mathcal{I}$ as $C_\mathcal{I} = \sum_{t \in \mathcal{I}} c_t$. Correspondingly, we define $C_{\text{total}} = C_{\{0,n\}}$. Then, Nature draws $x_t \sim \nu_*$ for each $t \in \{1,\ldots,n\}$ so that each $x_t$ is independent of whether $y_t$ was potentially corrupted or not.

One notable case case of the oblivious model is the $\gamma$-misspecification model. In the binary classification setting, it is equivalent to

$$\eta_t^x = (1 - \gamma)\eta_*^x + \gamma\tilde{\eta}_t^x, \forall x, t,$$

where $\tilde{\eta}_t^x$ can be any arbitrary probability. Such label contamination model can be regarded a special case of corruption where for each $t$,

$$c_t = \max_x |\eta_t^x - \eta_*^x| = \gamma \max_x |\eta_*^x - \tilde{\eta}_t^x|$$

Moreover, our main algorithm actually works for the non-oblivious adversary. In this more challenging case, each time $t$, the adversary adaptively decides $\eta_t^x$ before seeing actual $x_t$, based on all the previous history.

**Other notations:** For convenience, we denote $R_{D_t}(h)$ as $R_t(h)$, $R_{D_*}(h)$ as $R_* (h)$, $\rho_{D_t}(h, h') = \rho_t(h, h')$ and $\rho_{D_*}(h, h') = \rho_*(h, h')$. We also define an average expected risk that will be used a lot in our analysis, $R_{\mathcal{I}}(h) = \frac{1}{|\mathcal{I}|} \sum_{t \in \mathcal{I}} R_t(h)$. In addition, we define $h^* = \arg \min R_* (h)$, $R^* = R_*(h^*)$ and the gap of the suboptimal classifier $h$ as $\Delta_h = R_*(h) - R^*$.

**Disagreement coefficient:** For some hypothesis class $\mathcal{H}$ and subset $V \subset \mathcal{H}$, the region of disagreement is defined as $\text{Dis}(V) = \{ x \in \mathcal{X} : \exists h, h' \in V \text{ s.t. } h(x) \neq h'(x) \}$, which is the set of unlabeled examples $x$ for which there are hypotheses in $V$ that disagree on how to label $x$. Correspondingly, the disagreement coefficient of $h^* \in \mathcal{H}$ with respect to a hypothesis class $\mathcal{H}$ and distribution $\nu_*$ is defined as

$$\theta^*(r) = \sup_{r \geq r_0} \frac{\mathbb{P}_{x \sim \nu_*} (X \in \text{Dis}(B(h^*, r)))}{r}.$$
We restate the classical RobustCAL [Balcan et al., 2009, Dasgupta et al., 2007, Hanneke, 2014] in Algorithm 1 RobustCAL (modified the elimination condition)

**Theorem 4.1.** Suppose the additional term \( \frac{1}{n} \hat{\rho}_t(h, h_t) \) ensures robustness because each \( (R_t(h) - R_t(h')) \) will be corrupted at most \( 2\rho_* (h, h')c_t \). In the theorem below we show that, it can achieve the similar label complexity result as in the non-corrupted setting as long as the growth rate of corruptions is at most in a certain fraction of number of unlabeled samples.

**Theorem 4.1.** Suppose the \( C_{[t],\epsilon} \leq \frac{1}{\delta} \) for all \( t \in \{ \log(t) = \mathbb{N} \} \), for example, the (1/8)-mis specification model. Then with high probability at least \( 1 - \delta \), for any \( n \geq \frac{(2R_h^* + \frac{22}{\epsilon}) \log((\log(n))|\mathcal{H}|^2/\delta)}{2} \), we have \( R_{h_{\text{out}}} - R^* \leq \epsilon + \mathcal{O}(\frac{R_{h_{\text{out}}}^*}{n}) \) with label complexity at most

\[
\mathcal{O} \left( \theta^* (14R^* + 120 \frac{\log((\log(n))|\mathcal{H}|^2/\delta)}{n}) \log((\log(n))|\mathcal{H}|^2/\delta) (R^*n + \log(n)) \right)
\]

**Remark 4.1.** In Appendix [C.2] we show the necessity of enlarging the threshold in line [10] from the original

\[
V_{t+1} = \left\{ h \in V_{\log(t)} : \hat{L}_t(h) - \hat{L}_t(h_t) \leq \frac{\sqrt{2\beta_t \hat{\rho}_t(h, h_t) \hat{\beta}_t}}{t} \right\}
\]

by giving an counter-example. The counter-example shows that, when \( R^* \gg 0 \), the best hypothesis will be eliminated under the original condition even the “\( C_{[0,t]} \leq \frac{1}{\delta} \) for all \( t \in \{ \log(t) = \mathbb{N} \} \)” assumption is satisfied.

**Proof Sketch** For correctness, it is easy to show by Bernstein inequality. For the sample complexity, Theorem [3.1] implies that, for any interval \([0,t]\), as long as \( C_{[0,t]} \leq \frac{1}{\delta} \), the learner can always identify
We adopt the idea from the BARBAR algorithm proposed by Gupta et al. [2019] which was originally
true care is the gap between any hypothesis
This aligns with its MAB feedback structure, where only the information of the pulled arm will be
\[ \tilde{\rho}_t(h, \hat{h}_t) \]
Then by standard analysis we can connect this disagreement probability with the disagreement
\[ \delta(\hat{h}) \]
By Camilleri et al. [2021], we construct such estimation by using
\[ \delta(h) \]
Comparison between the modified RobustCal and passive learning: Assume disagreement
coefficient is a constant. In the non-corrupted case, the algorithm achieves the same performance
guarantee as the vanilla Robust CAL. In the corrupted case, we still get the same accuracy as in
Theorem 3.1 with at most \( \tilde{O}(R^* n + \log(n)) \) number of labels, which is the same as the non-corrupted
case.
Discussion on the “\( C_{[0,t]} \leq \frac{t}{8} \) for all the \{t \in \mathbb{N} \}” condition: This condition can be
reduced to the \((1/8)\)-mis specification model as defined in Section 2 since \( C_I \leq \frac{|I|}{8} \) for any \( I \). But
this condition does not contain the case where an adaptive poisoning adversary corrupts all the labels
at the earlier stage and stop corrupting later, which still ensures the small total amount of corruptions,
but will clearly mislead the algorithm to delete a true best hypothesis \( h^* \). In Section 5 we will show
a more general result that applies to scenarios beyond \( C_{[0,t]} \leq \frac{t}{8} \).

5 Main algorithm - CALruption

5.1 Algorithm

In this section we describe our new algorithm, CALruption. The pseudo-code is listed in Algorithm 2.
Our previous analysis showed that in the agnostic setting the classical RobustCAL may permanently
eliminate the best hypothesis due to the presence of corruptions. To fix this problem, in our CAL-
ruption algorithm, the learner never makes a “hard” decision to eliminate any hypothesis. Instead, it
assigns different query probability to each hypothesis, as shown in line 9 to 11.

We adopt the idea from the BARBAR algorithm proposed by Gupta et al. [2019] which was originally
designed for multi-armed bandits (MAB). Instead of permanently eliminating a hypothesis, the learner
will continue pulling each arm with a certain probability defined by its estimated gap. However, the
original BARBAR algorithm is mainly focused on estimating the reward of each individual arm.
This aligns with its MAB feedback structure, where only the information of the pulled arm will be
gained at each time. In the active learning setting, we instead focus on the
In line 7, we estimate the disagreement probability for each hypothesis pair \( (h, h') \) with an empirical
quantity that upper bounds the expectation. In line 8 instead of estimating the value of each
hypothesis, we estimate the gap between each hypothesis pair \( (h, h') \), denoted as \( W_{h,h'}^t \), by any
\( \delta \)-robust estimator that satisfies eq. 1. One example of \( \delta \)-robust estimator is Catoni estimator [Lugosi
and Mendelson, 2019]. Note that simple empirical estimator will lead to potentially rare but large
variance, which has been discussed in Stochastic rounding section in Camilleri et al. [2021]. But what
we truly care is the gap between any hypothesis \( h \) and the best hypothesis \( h^* \). Therefore, inspired
by Camilleri et al. [2021], we construct such estimation by using \( W_{h^*,h'}^t \) as shown in line 9 to 11.
Finally, we divide the hypothesis set into several layers based on the estimated gap and set the query
probability for each \( x \) based on the hypothesis layers, as shown in line 12 and 13. For more detailed
explanation on line 9-13, please refer to Appendix D.
Algorithm 2 CALruption

1: Initialize: $\beta_3 = 2 \log(\frac{1}{2} |\log(n)| |\mathcal{H}|^2 / \delta), \beta_1 = 32 \times 640 \beta_3, \beta_2 = \frac{5}{32}, \epsilon_i = 2^{-i}, N_l = \beta_1 \epsilon_i^{-2}$, $\hat{\Delta}_h^0 = 0, V_i^0 = \mathcal{Z}$ and $\tau_1 = 1, q^F = 1$ for all $x \in \mathcal{X}$
2: for $l = 1, 2, \ldots, n$ do
3: Nature reveals unlabeled data point $x_l$
4: Set $Q_l \sim \text{Ber}(q^F)$ and request $y_l$ if $Q_l = 1$.
5: Set estimated loss for all $h \in \mathcal{H}$ as $\hat{\ell}_t(h) = \frac{1}{q^F} \frac{|h(x_l) \neq y_l|}{Q_l}$
6: if $t = \tau_l + N_l - 1$ then
7: Set $\hat{\rho}_l(h, h') = \frac{1}{N_l} \sum_{x \in I_t} 1\{h(x) \neq h'(x)\}$ for all $h, h' \in \mathcal{H}$
8: For each $(h, h')$, set $W_i^{h, h'} = \text{RobustEstimator}\left(\{\hat{\ell}_t(h) - \hat{\ell}_t(h')\}_{t \in I_t}\right)$, which satisfies that, with probability at least $1 - \delta$, 
\[ |(\hat{R}_l(h) - \hat{R}_l(h')) - W_i^{h, h'}| \leq \sqrt{\frac{10 \beta_3 \hat{\rho}_l(h, h')}{N_l \min_{x \in \text{Dis}(h, h')} q^F}}, \]
\[ \text{(1)} \]
where $\hat{R}_l(h) = \frac{1}{|I_t|} \sum_{x \in I_t} \mathbb{E}[y \sim \text{Ber}(\eta^F)] 1\{h(x) \neq y\}$.
9: Set $D_l = \arg\min_h \max_{h' \in \mathcal{H}} (R_D(h) - R_D(h') - W_i^{h, h'}) \sqrt{\frac{\min_{x \in \text{Dis}(h, h')} q^F}{\hat{\rho}_l(h, h')}}$
10: Set $\hat{h}_l^i = \arg\min_h (R_{D_l}(h) + \beta_2 \hat{\Delta}_h^{l-1})$
11: Set $\hat{\Delta}_h^l = \max\{\epsilon_i, R_{D_l}(h) - (R_{D_l}(\hat{h}_l^i) + \beta_2 \hat{\Delta}_h^{l-1})\}$
12: Construct $V_{l+1}^i$ for all $i = 0, 1, 2, \ldots, l$, such that, 
\[ \hat{\Delta}_h^l \leq \epsilon_i, \forall h \in V_{l+1}^i \quad \text{and} \quad \hat{\Delta}_h^l > \epsilon_i, \forall h \notin V_{l+1}^i \]

Therefore, $V_{l+1}^0 \subset V_{l+1}^1 \subset \ldots \subset V_{l+1}^0$
13: Calculate the query probability $q^F$ for each $x$ as follows 
\[ \mathcal{Z}(x) = \{(h, h') \in \mathcal{H} \mid x \in \text{Dis}(\{h, h'\})\} \]
\[ k(h, h', l + 1) = \max\{i \mid h, h' \in V_{l+1}^i\} \]
\[ q^F = \max_{(h, h') \in \mathcal{Z}(x)} \frac{\beta_1 \hat{\rho}_l(h, h')}{N_{l+1}^i} \epsilon_i^{-2} k(h, h', l + 1) \]

14: Set $\tau_{l+1} = \tau_l + N_l$ and denote the epoch $l$ as $\mathcal{I} = [\tau_l, \tau_{l+1} - 1]$. Set $l \leftarrow l + 1$, go to the next epoch
15: end if
16: end for
17: Output: $h \in V_{l+1}^0$

Remark 5.1. In Line 5, instead of estimating over all possible distribution $D$, we actually just need to estimate $\eta^F$ for all $x \in \{x_l\}_{l \in I_l}$ and set the corresponding $x$ distribution of $D$ as the empirical distribution of $x$ inside $I_l$.

Theorem 5.1 (CALruption). With $n \geq 72 \varepsilon^{-2} \beta_1$ number of unlabeled samples, with probability at least $1 - \delta$ we can get an $h_{out}$ satisfying 
\[ R_{\star}(h_{out}) - R_{\star} \leq \varepsilon + 24 \frac{C_{\text{total}}}{n}, \]
with label complexity as most 
\[ \mathcal{O}\left(\theta^* (R_{\star}^* + 3 \sqrt{\frac{\beta_1}{n} + \frac{64 C_{\text{total}}}{n}}) \log(\log(n)|\mathcal{H}|^2 / \delta) ((R_{\star}^*)^2 n + \log(n)(1 + C_{\text{total}}))\right) \]
where $C_{\text{total}} = \sum_{l \in I} \mathbb{E}_{n, \beta} \mathbb{E}_y l \left( R_{\star}^* \mathbb{1}\left\{ \frac{C_{\text{epoch}} l}{N_l} \leq \frac{1}{32\beta_1} \right\} + \mathbb{1}\left\{ \frac{C_{\text{epoch}} l}{N_l} > \frac{1}{32\beta_1} \right\} \right)$ and $\beta_1 = 16 \times 640 \log(\frac{3}{2} |\log(n)| |\mathcal{H}|^2 / \delta)$. Note that epoch $l$ is prescheduled and not algorithm-dependent.
Corollary 5.1. Suppose the corruptions satisfy \( \frac{C_{\text{miss}_{l}}}{N_{l}} \leq \frac{1}{32} \) for all epochs, for example, the (1/32)-misspecification case, then for any \( n \geq 72\varepsilon^{-2}\beta_{1} \) number of unlabeled samples, with probability at least \( 1 - \delta \) we can get a \( h_{\text{out}} \) satisfying

\[
R_{*}(h_{\text{out}}) - R^{*} \leq \varepsilon + 24R^{*}\frac{C_{\text{total}}}{n},
\]

with label complexity as most

\[
\mathcal{O}\left( \theta^{*}(R^{*} + 3\sqrt{\frac{R^{*}\beta_{1}}{n}} + \frac{64R^{*}C_{\text{total}}}{n}) \log(\log(n)|\mathcal{H}|/\delta) \left( (R^{*})^{2}n + (R^{*}C_{\text{total}} + 1) \log(n) \right) \right)
\]

Comparison with passive learning and the Calruption: Consider the case where \( \theta^{*}(\cdot) \) is of lower order like a constant. The Corollary 5.1 shows that, when \( \frac{C_{\text{miss}_{l}}}{N_{l}} \leq \frac{1}{32} \) for all epochs, our algorithm achieves a similar accuracy \( \mathcal{O}\left( \varepsilon + \frac{C_{\text{total}}}{n} \right) \) as in the passive learning case, while only requiring \( \mathcal{O}\left( (R^{*})^{2}n + \log(n)(1 + R^{*}C_{\text{total}}) \right) \) number of labels, for \( n \geq \frac{1}{\varepsilon^{2}} \). So if we set \( n = \mathcal{O}\left( \frac{1}{\varepsilon^{2}} \right) \), then the label complexity becomes \( \mathcal{O}\left( \frac{R^{*}C_{\text{total}}}{n} + \log(1/\varepsilon)(1 + R^{*}C_{\text{total}}) \right) \), which matches the minimax label complexity in the non-corrupted case.

Going beyond the \( \frac{C_{\text{miss}_{l}}}{N_{l}} \leq \frac{1}{32} \) constraint, the general Theorem 5.1 shows that, for \( n \geq \frac{1}{\varepsilon^{2}} \), our algorithm achieves an accuracy \( \mathcal{O}\left( \varepsilon + \frac{C_{\text{total}}}{n} \right) \) while only requiring \( \mathcal{O}\left( (R^{*})^{2}n + \log(n) + C_{\text{total}} \right) \) number of labels no matter how corruptions are allocated. When \( R^{*} \) is some constant, this result becomes similar to the Corollary 5.1. Moreover, we will argue that upper bound \( C_{\text{total}} \) by \( C_{\text{total}} \) is loose and in many case \( C_{\text{total}} \) will be close to \( R^{*}C_{\text{total}} \) instead of \( C_{\text{total}} \). We show one example in the paragraph below.

When is Calruption better than modified Robust CAL? Consider the case where the adversary fully corrupts some early epoch and then performs corruptions satisfying \( \frac{C_{\text{miss}_{l}}}{N_{l}} \leq \frac{1}{32} \) for rest epochs. Then the modified Robust CAL will mistakenly eliminate \( h^{*} \) so it can never achieve target result when \( \varepsilon < \min_{\Delta_{h}} \Delta_{h} \) while Calruption can surely output the correct hypothesis. Moreover, according to Theorem 5.1, since the total amount of early stage corruptions are small, so here \( C_{\text{total}} \) is close to \( R^{*}C_{\text{total}} \), which implies a similar result as in Corollary 5.1.

When is Calruption worse then modified Robust CAL? Consider the case where the total amount of corruption is, instead of fixed, increasing with incoming unlabeled samples, for example, the misspecification case. Then \( C_{\text{total}} \) in modified Robust CAL can be \( \mathcal{O}\left( \frac{C_{\text{total}}}{N_{l}} + \frac{1}{l} \right) \) while \( C_{\text{total}} \) in CALruption can goes to \( \mathcal{O}\left( \frac{1}{l} \right) \). Such gap comes from the extra unlabeled sample complexity, which we discuss in the paragraph below.

Discussion on the extra unlabeled samples complexity: We note that we require a larger number of unlabeled data than ERM in the passive learning setting. Here we explain the reason. Consider the version spaces \( V_{l}^{l-1} \) for any fixed epoch \( l \). In the non-corrupted setting, this version space serves the similar purpose as the active hypothesis set in Robust CAL. In Robust CAL, its elimination threshold is about \( \mathcal{O}\left( \sqrt{\frac{\rho_{*}(h, h')}{l}} + \frac{1}{l^{2}} \right) \) (or \( \mathcal{O}\left( \rho_{*}(h, h') + \frac{1}{l} \right) \) in our modified version) while in our CALruption, the threshold is about \( \mathcal{O}\left( \sqrt{\frac{1}{l}} \right) \), which is more conservative than the Robust CAL and leads to the extra unlabeled sample complexity. The reason about being conservative here is that we need more samples to weaken the effects of corruptions on our estimation. Whether such extra unlabeled samples complexity is unavoidable remains an open problem.

5.2 Proof sketch for Theorem 5.1

Here we provide main steps of the proof and postpone details in Appendix E.

First we show a key lemma which guarantees the closeness between \( \Delta_{l}^{h} \) and \( \Delta_{h} \) for all \( l \) and \( h \).
Lemma 5.1 (Upper bound and lower bound for all estimation). With probability at least $1 - \delta$, for all epoch $l$ and all $h \in \mathcal{H}$,

$$\hat{\Delta}_h^l \leq 2(\Delta_h + \epsilon_l + g_l), \quad \Delta_h \leq \frac{3}{2} \Delta_h^l + \frac{3}{2} \epsilon_l + 3g_l,$$

where $g_l = \frac{2}{\beta_1} \epsilon_l^2 \sum_{s=1}^{l} C_s \left(2R^* \mathbf{1} \left\{ \frac{2C_s}{N} \leq \frac{1}{10} \right\} + 1 \left\{ \frac{2C_s}{N} > \frac{1}{10} \right\} \right)$.

Here the $g_l$ term implies that, as long as the total corruption is sublinear in $n$, the misleading effects on the gap estimations will fade when the number of unlabeled samples increasing.

Based on this lemma, we can directly get another useful lemma as follows.

Lemma 5.2. For all epoch $l$ and layer $j$, we have $\max_{h \in V_j^l} \rho_j(h, h^*) \leq 2R^* + 3\epsilon_j + 3g_{l-1}$

In the following we first deal with the correctness then sample complexity.

Correctness. By Lemma 5.1, we have

$$\Delta_{h_{\text{out}}} \leq \frac{3}{2} \hat{\Delta}_{h_{\text{out}}} + \frac{3}{2} \epsilon_{L-1} + 3g_{L-1} \leq 6 \sqrt{\frac{2\beta_1}{n} + 24 \frac{C_{\text{total}}}{n}}.$$  

Sample complexity. For any $t \in I_l$, recall that $q_l^t = \max_{(h, h') \in \mathcal{Z}(x)} \frac{\beta_l \rho_{l,h,h'} \epsilon_l^2}{N_t} \epsilon_{k,h',t}^l$, the probability of $x_t$ being queried ($Q_t = 1$) is

$$\mathbb{E}[Q_t] \leq 10 \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \max_{h \in V_j^l} \rho_j(h, h^*) \epsilon_l^{-2} + 8 \frac{\beta_1}{N_l} \leq 10 \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \left(2R^* \epsilon_l^{-2} + 3\epsilon_l^{-2} + 3g_{l-1} \epsilon_l^{-2}\right) + 8 \frac{\beta_1}{N_l} \leq 10 \frac{\beta_1}{N_l} \sum_{t=0}^{l-1} \left(2R^* \epsilon_l^{-2} + 3\epsilon_l^{-2} + 3g_{l-1} \epsilon_l^{-2}\right) \mathbb{P}(x \in \text{Dis}(V_j^l)) + 8 \frac{\beta_1}{N_l} \text{Prob}$$

Here $j^*$ is some arbitrary mapping from $\mathcal{X}$ to $[l]$, which is formally defined in detailed version in Appendix E.6. The first inequality comes from the closeness of estimated $\hat{\rho}_{l,h,h'}(h, h')$ and the true $\rho_j(h, h')$, as well as some careful relaxation. The second inequality comes from Lemma 5.2.

Now we can use the standard techniques to upper bound $\mathbb{P}(x \in \text{Dis}(V_j^l))$ as follows,

$$\mathbb{P} \left( \exists h \in V_j^l : h(x) \neq h^*(x) \right) \leq \mathbb{P} \left( \exists h \in \mathcal{H} : h(x) \neq h^*(x), \rho_j(h, h^*) \leq 2R^* + 3\epsilon_l + 3g_{l-1} \right) \leq \theta^* (2R^* + 3\epsilon_l + g_{l-1}) (2R^* + 3\epsilon_l + g_{l-1})$$

where again the first inequality comes from Lemma 5.2. Again we postpone the full version into Appendix E.6.

Combining the above results with the fact that $g_l = \frac{2}{\beta_1} \epsilon_l^2 \bar{C}_{l-1}$ and $\bar{C}_{l-1} \leq \sum_{s=1}^{l-1} C_s \leq 2\beta_1 \epsilon_l^{-2}$, we get the expected number of queries inside a complete epoch $l$ as,

$$\sum_{t \in Z_l} \mathbb{E}[Q_t] \leq 20 \beta_1 \theta^* (2R^* + 3\epsilon_l + g_{l-1}) * \left(4(R^*)^2 \epsilon_l^{-2} + 12R^* \epsilon_l^{-1} + \frac{132}{\beta_1} \bar{C}_{l-1} + 10 \right)$$

Finally, summing over all $L = \lceil \frac{1}{2} \log(n/\beta_1) \rceil$ number of epochs, for any $n$, we can get the target label complexity.

6 Conclusion and future works

In this work we analyzed an existing active learning algorithm in the corruption setting, showed when it fails, and designed a new algorithm that resolve the drawback. Relative to RobustCAL,
our algorithm requires a larger number of unlabeled data. One natural question is to design a corruption robust algorithm which requires the same number of unlabeled data as RobustCAL in the non-corrupted setting. Another potential question is that, the $O\left(\epsilon + \frac{C_{\text{total}}}{n}\right)$ accuracy from Algo. 2 in the general corruption case is generally worse than $O\left(\epsilon + \frac{R^* C_{\text{total}}}{n}\right)$ accuracy of passive learning. Although we state that these two bounds are close in many cases, a question is if there exists an alternative algorithm or analysis that will result a more smooth final bound to interpolate between different corruptions cases.

Finally, we believe it is possible to replace the Catoni estimator with naive importance sampling estimator and even further simplify the layered active hypothesis sets construction step by adopting Beygelzimer et al. [2010]’s idea. We plan to work on this direction in the future.

Acknowledgements

SSD gratefully acknowledges funding from NSF Award’s IIS-2110170 and DMS-2134106

References


Chen-Yu Wei, Christoph Dann, and Julian Zimmert. A model selection approach for corruption robust reinforcement learning, 2021.


Burr Settles. From theories to queries: Active learning in practice. In Active Learning and Experimental Design workshop In conjunction with AISTATS 2010, pages 1–18, 2011.


A Lemmas related to corruption effects

Here we states some basic lemmas that will be used all over the proofs.

Lemma A.1 (Corruption effects 1). For any interval $I$ and hypothesis $h$, we have

$$\frac{1}{|I|} \sum_{t \in I} (R_t(h) - R^*_t(h)) \leq \frac{C_I}{|I|}$$
Proof.

\[
\frac{1}{|I|} \sum_{t \in I} (R_t(h) - R_t(h'))
\]

\[
= \mathbb{E}_{x \sim \nu_t} \frac{1}{|I|} \sum_{t \in I} (\mathbb{E}_{y \sim \eta_t^x} [1\{h(x) \neq y\}] - \mathbb{E}_{y \sim \eta_t^x} [1\{h(x) \neq y\}])
\]

\[
\leq \frac{1}{|I|} \sum_{t \in I} \max_{x \in X} (\mathbb{E}_{y \sim \eta_t^x} [1\{h(x) \neq y\}] - \mathbb{E}_{y \sim \eta_t^x} [1\{h(x) \neq y\}])
\]

\[
\leq \frac{1}{|I|} \sum_{t \in I} \max_{x \in X} |\eta_t^x - \eta_t^x| \leq C_I \frac{1}{|I|}
\]

\[\square\]

Lemma A.2 (Corruption effects 2). For any interval \(I\) and hypothesis pair \(h, h'\), we have

\[
\frac{1}{|I|} \sum_{t \in I} (R_t(h) - R_t(h')) - (R_\star(h) - R_\star(h')) \leq 2\rho_\star(h, h') C_I \frac{1}{|I|}
\]

Proof.

\[
\frac{1}{|I|} \sum_{t \in I} (R_t(h) - R_t(h')) - (R_\star(h) - R_\star(h'))
\]

\[
= \mathbb{E}_x \left[ \frac{1}{|I|} \sum_{t \in I} (\mathbb{E}_{x_t \sim \eta_t^x} [1\{h(x_t) \neq y\}] - \mathbb{E}_{x_t \sim \eta_t^x} [1\{h(x_t) \neq y\}]) \right]
\]

\[
= \mathbb{E}_x \left[ \frac{1}{|I|} \sum_{t \in I} (\mathbb{E}_{x_t \sim \eta_t^x} [1\{h(x_t) \neq y\}] - \mathbb{E}_{x_t \sim \eta_t^x} [1\{h(x_t) \neq y\}]) \right]
\]

\[
\leq \rho_\star(h, h') \left( \frac{2}{|I|} \sum_{h \in H} \max_{h \in H} (R_t(h) - R_\star(h)) \right)
\]

\[
\leq 2\rho_\star(h, h') C_I \frac{1}{|I|}
\]

\[\square\]

B Analysis for Passive Learning: Proof of Theorem [3.1]

With probability at least \(1 - \delta\), we have for any \(n\) samples,

\[
R_\star(h_{\text{out}}) - R^\star
\]

\[
\leq (R_\star(h_{\text{out}}) - R^\star) - \left( \tilde{R}_{[1,n]}(h_{\text{out}}) - \tilde{R}_{[1,n]}(h^\star) \right) + \left( \tilde{R}_{[1,n]}(h_{\text{out}}) - \tilde{R}_{[1,n]}(h^\star) \right)
\]

\[
\leq 2 \frac{C_{\text{total}}}{n} \rho_\star(h_{\text{out}}, h^\star) + \left( \tilde{R}_{[1,n]}(h_{\text{out}}) - \tilde{R}_{[1,n]}(h^\star) \right)
\]

\[
\leq 2 \frac{C_{\text{total}}}{n} \rho_\star(h_{\text{out}}, h^\star) + \left( \tilde{R}_{[1,n]}(h_{\text{out}}) - \tilde{R}_{[1,n]}(h^\star) \right) + \sqrt{\rho_\star(h_{\text{out}}, h^\star) \frac{4 \log(|H|/\delta)}{n}} + \frac{\log(|H|/\delta)}{n}
\]

\[
\leq 2 \frac{C_{\text{total}}}{n} \max \{ R_\star(h_{\text{out}}) - R^\star, 2R^\star \} + \sqrt{\max \{ R_\star(h_{\text{out}}) - R^\star, 2R^\star \} \frac{4 \log(|H|/\delta)}{n}} + \frac{\log(|H|/\delta)}{n}
\]

where the second step can from our definition of corruptions and fact that \(\nu_\star\) is not corrupted (see Lemma A.2 for details), third inequality comes from the Bernstein inequality and the last inequality comes from the definition of \(h_{\text{out}}\) and the fact \(\rho_\star(h, h') \leq 2 \max \{ R_\star(h) - R^\star, 2R^\star \} \). Now if \(2R^\star \geq R_\star(h_{\text{out}}) - R^\star\), then we directly get the target result. Otherwise, by solving the quadratic inequality, we have

\[
R_\star(h_{\text{out}}) - R^\star \leq \frac{5 \log(|H|/\delta)}{n} \frac{1}{(1 - \frac{4C_{\text{total}}}{n})^2}
\]
**C Analysis for Robust CAL**

**C.1 Proof of Theorem 4.1**

For convenient, for all subscripts \([0, t]\), we simply write as subscript \(t\).

We first state a key lemma that is directly inspired by Theorem 3.1.

**Lemma C.1.** For any \(t\) that \(\log(t) = N\), under the assumption of this theorem, as long as \(h^* \in V_t\), we have

\[
R_*(\hat{h}_t) - R^* \leq \frac{22 \log(|\mathcal{H}|/\delta)}{t} + \frac{4}{n} R^* + \sqrt{\frac{8 \log(|\mathcal{H}|/\delta)}{t}} 
\]

\[
\leq \frac{22 \log(|\mathcal{H}|/\delta)}{t} + \frac{R^*}{2} + \sqrt{\frac{8 \log(|\mathcal{H}|/\delta)}{t}} \quad \text{(By assumption on } C_t) 
\]

\[
\leq \frac{26 \log(|\mathcal{H}|/\delta)}{t} + R^* \quad \text{(By the fact } \sqrt{AB} \leq A + B \text{)}
\]

**Proof.** With probability at least \(1 - \delta\), by combine the same proof steps as in Theorem 3.1 and the fact that \(\hat{R}_{[1,t]}(\hat{h}_t) - \hat{R}_{[1,t]}(h^*) = \hat{L}_t(\hat{h}_t) - \hat{L}_t(h^*) \leq 0\), we can get the similar inequality as follows

\[
R_*(\hat{h}_t) \leq \frac{4 C_t}{n} \max\{R_*(\hat{h}_t) - R^*, 2R^*\} + \sqrt{\max\{R_*(\hat{h}_t) - R^*, 2R^*\}} \frac{4 \log(|\mathcal{H}|/\delta)}{t} + \frac{\log(|\mathcal{H}|/\delta)}{t}
\]

Then again by quadratic inequality and the assumption that \(C_t \leq \frac{1}{8} t\), we have

\[
R_*(\hat{h}_t) \leq \frac{22 \log(|\mathcal{H}|/\delta)}{t} + \frac{4}{n} R^* + \sqrt{\frac{8 \log(|\mathcal{H}|/\delta)}{t}}
\]

This lemma suggests that, as long as the corruptions are not significantly large. For example, in this theorem, \(C_t \leq \frac{1}{8} t\). Then the learner can still easily identify the \(\tilde{O}(\frac{1}{\gamma} + R^*)\)-optimal hypothesis even in the presence of corruptions. Therefore, we can guarantee that the best hypothesis always stay in active set \(V_t\) after elimination. We show the detailed as follows.

Define \(E_1, E_2\) as

\[
E_1 := \{\forall t \text{ that } \log(t) = N, (\overline{R}_t(h) - \overline{R}_t(h')) - (\hat{R}_t(h) - \hat{R}_t(h')) \leq \frac{2\beta_t \hat{\rho}_t(h, h')}{t} + \frac{\beta_t}{t}\}
\]

\[
E_2 := \{\forall t \text{ that } \log(t) = N, (\overline{R}_t(h) - \overline{R}_t(h')) - (\hat{R}_t(h) - \hat{R}_t(h')) \leq \frac{2\beta_t \rho_*(h, h')}{t} + \frac{\beta_t}{t}\}
\]

\[
E_3 := \{\forall t \text{ that } \log(t) = N, |\rho_*(h, h') - \hat{\rho}_t(h, h')| \leq \frac{2\beta_t \hat{\rho}_t(h, h')}{t} + \frac{\beta_t}{t}\}
\]

By (empirical) Bernstein inequality plus union bound, it is easy to see \(P(E_1 \cap E_2 \cap E_3) \geq 1 - \delta\).

**First we show the correctness.**
For any $t$ that $\log(t) = \mathbb{N}$, assume that $h^* \in V_t$, then we have

$$
\tilde{L}_t(h^*) - \tilde{L}_t(h_t) = R_t(h^*) - R_t(h_t)
\leq R_t(h^*) - \tilde{R}_t(h_t) + \sqrt{\frac{\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{\beta_t}{2t}
\leq R^* - R_t(h_t) + \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{\beta_t}{t} + \rho_*(h^*, h_t) 2C_t/t
\leq \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{\beta_t}{t} + \rho_*(h^*, h_t) 2C_t/t
\leq \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{\beta_t}{t} + \left(\hat{\rho}_t(h^*, h_t) + \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{\beta_t}{t}\right) 2C_t/t
\leq \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{3\beta_t}{2t} + \frac{1}{2} \hat{\rho}_t(h^*, h_t)
$$

where the first and forth inequality comes from the event $E_1$ and $E_3$, the second inequality comes from Lemma A.2, the third inequality comes from the definition of $R^*$ and last inequality comes from $\sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} \leq \frac{\hat{\rho}_t(h^*, h_t)}{2t} + \frac{\beta_t}{t}$ and the assumption that $\frac{\beta_t}{t} \leq \frac{1}{8}$.

According to the elimination condition [10] in Algo. [1] this implies that $h^* \in V_{t+1}$. Therefore, by induction, we get that $h^* \in V_n$. By again using Lemma C.1 we can guarantee that

$$
R_*(h_{out}) - R^* \leq \frac{22 \log(|H|/\delta)}{n} + \frac{4R^* C_{total}}{n} + \sqrt{R^* \frac{8 \log(|H|/\delta)}{n}}
$$

**Next we show the sample complexity.** For any $t$ that $\log(t) = \mathbb{N}$ and any $h \in V_t$, we have

$$
\Delta_h = (\Delta_h - (\bar{R}_t(h) - \hat{R}_t(h))) + (\hat{R}_t(h) - \bar{R}_t(h)) + (\hat{R}_t(h) - \bar{R}_t(h^*))
\leq \frac{2C_t}{t} \rho_*(h, h^*) + \sqrt{\frac{2\beta_t \rho_*(h^*, h_t)}{t}} + \frac{\beta_t}{t} + \hat{R}_t(h) - \hat{R}_t(h_t)
\leq \frac{1}{4} \rho_*(h, h^*) + \sqrt{\frac{2\beta_t \rho_*(h^*, h_t)}{t}} + \frac{\beta_t}{t} + \frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t} + \frac{3\beta_t}{2t} + \frac{1}{2} \hat{\rho}_t(h^*, h_t)
\leq \frac{19}{24} \rho_*(h, h^*) + \sqrt{\frac{2\beta_t \rho_*(h^*, h_t)}{t}} + \sqrt{\frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t}} + \frac{2\beta_t \hat{\rho}_t(h^*, h_t)}{t} + \frac{6\beta_t}{t}
\leq \left(\frac{19}{24} + \frac{25}{24\beta_4}\right) \rho_*(h, h^*) + \frac{13}{24\beta_4} \rho_*(h_t, h^*) + \left(\frac{2\beta_4 + 6}{2}\right) \frac{19}{24} + \frac{25}{24\beta_4} + \frac{13}{24\beta_4}
\leq \left(\frac{19}{24} + \frac{25}{24\beta_4}\right) \Delta_h + \frac{13}{24\beta_4} \Delta_{h_t} + \left(\frac{2\beta_4 + 6}{2}\right) \frac{19}{24} + \frac{25}{24\beta_4} + \frac{13}{24\beta_4}
\leq \left(\frac{19}{24} + \frac{25}{24\beta_4}\right) \Delta_h + \left(\frac{2\beta_4 + 6}{2}\right) \frac{19}{24} + \frac{25}{24\beta_4} + \frac{13}{24\beta_4}
\leq \frac{12\beta_4}{t} + 12 R^*
$$

where the first inequality comes from the event $E_2$ and the definition of $h_t$, the second inequality comes from the elimination condition [10] in Algo. [1]. For the third and forth inequality, we use the fact $\sqrt{AB} \leq \frac{A^2 + B^2}{2}$ multiple times and the last inequality comes from Lemma C.1.

Finally, choose $\beta_4 = 25$ and solve this inequality, we get $\Delta_h \leq \frac{12\beta_4}{t} + 12 R^*$
Therefore, we get the probability of query as
\[
\mathbb{P}(x_{t+1} \in \text{Dis}(V_{t+1})) \leq \mathbb{P}\left( \exists h \in V_{t+1} : h(x_t) \neq h^*(x_t), \Delta_h \leq \frac{120\beta_t}{t} + 12R^* \right)
\]
\[
\leq \mathbb{P}\left( \exists h \in V_{t+1} : h(x_t) \neq h^*(x_t), \rho_*(h, h^*) \leq 14R^* + \frac{120\beta_t}{t} \right)
\]
\[
\leq \theta^*(14R^* + \frac{120\beta_t}{t}) \left(14R^* + \frac{120\beta_t}{t}\right)
\]
Therefore, we get the final prove by summing this probability over all the time.

C.2 Why vanilla Robust CAL does not work?

Proposition C.1. When \( R^* \gg 0 \) and the corruptions are unknown to the learner, there exists an instance and an adversary such that the vanilla Robust CAL can never output the target hypothesis.

Proof. Suppose \( \mathcal{X} = \{x_1, x_2, x_3\} \) where \( \nu_*(x_1) = \xi_1 \gg 0, \nu_*(x_2) = \xi_2 \leq \frac{\xi_1}{64} \) and \( \nu_*(x_3) = 1 - \xi_1 - \xi_2 \). Here we further assume that \( \nu \) is given to the learner. For labels, we set \( \eta_{x_1}^s = \frac{1}{2}, \eta_{x_2}^s = \eta_{x_2}^s = 1 \). Now consider \( h_1 : h_1(x_1) = h_1(x_2) = h_1(x_3) = 1 \) and \( h_2 : h_2(x_1) = h_2(x_2) = 0, h_2(x_3) = 1 \).

With some routine calculations, we can obtain that:
\[
R^* = R_*(h_1) = \frac{1}{2}\xi_1, \quad R_*(h_2) = \frac{1}{2}\xi_1 + \xi_2, \quad \rho_*(h_1, h_2) = \xi_1 + \xi_2
\]
Now suppose the adversary corrupts \( \eta_{x_1}^s \) from \( \frac{1}{2} \) to \( \eta_{x_1}^s = \frac{15}{64} \) for all \( s \leq \tau \) and will stop corrupting at certain time \( \tau \). Consider this case \( C_t \leq \frac{1}{32}t \), which satisfies our corruption assumption.

With such corruptions, we have that for any \( t \leq \tau \),
\[
\hat{R}_t(h_1) = \frac{17}{32}\xi_1, \quad \hat{R}_t(h_2) = \frac{15}{32}\xi_1 + \xi_2,
\]
Since \( \hat{R}_t(h_2) \geq \hat{R}_t(h_1) \), so \( h_2 \) will never be eliminated before \( \tau \). Next we show that \( h_1 \) can be eliminated before \( \tau \). Note that, when \( \tau \geq O\left(\frac{1}{t^2}\right) \), we can always find a proper \( t \leq \tau \) such that
\[
\hat{R}_t(h_1) - \hat{R}_t(h_2) \geq \frac{1}{16}\xi_1 - \xi_2 - \tilde{O}\left(\sqrt{\frac{\xi_1 + \xi_2}{t}} + \frac{1}{t}\right)
\]
In the non-corrupted setting, the confidence threshold of vanilla Robust CAL is always \( \tilde{O}\left(\sqrt{\frac{\xi_1 + \xi_2}{t}} + \frac{1}{t}\right) \), which can be smaller than \( \frac{1}{16} \xi_1 - \xi_2 - \tilde{O}\left(\sqrt{\frac{\xi_1 + \xi_2}{t}} + \frac{1}{t}\right) \) for large enough \( t \), so the above inequality shows that \( h_1 \) can be eliminated before \( \tau \). This implies that, if our target accuracy \( \varepsilon < \xi_2 \), then the vanilla Robust CAL will never able to output the correct answer no matter how many unlabeled samples are given. On the other hand, in the passive learning, one can still output the target \( h_1 \) as long as \( n \gg \tau \).

D More detailed explanation for CALRuption for line 9 to 13

Here we provide a more detailed explanation on line 9 to 13

- In Line 9, we are going to estimate the underlying distribution of samples based on the collected samples. To be specific, we have the estimated gap between each pair of \( h \) and \( h' \), so the initial desire is to find a proper distribution that induces all gaps uniformly close to all the estimated gaps. But this is impossible, so we instead choose the distribution that minimizes the worst-case pairs scaled with its variance. With such an estimated distribution, we can naturally get the estimated error of each hypothesis \( h \) denoted as \( R_{\hat{P}}(h) \).
- In Line 10, recall that we already have the \( R_{\hat{P}}(h) \), and the previously estimated gap between any hypothesis \( h \) and the previous estimated best hypothesis \( h^{l-1}_s \), denoted as \( \Delta^{l-1}_h \).
So based on these two terms, we can have a pessimistic estimation of the current best hypothesis \( \hat{h}^l_{\ast} \).

- Then in Line 11, based on the estimated best hypothesis \( \hat{h}^l_{\ast} \), we can further have a new estimated gap \( \hat{\Delta}^l_\ast \).

Up to this point, we have an estimate of the performance of each hypothesis (\( \hat{\Delta}^l_\ast \)). Now recall that in the traditional elimination-style algorithms like Robust CAL, we will permanently eliminate all the hypotheses for which \( \hat{\Delta}^l_\ast \) is larger than some threshold and then do a disagreement-based query on the remaining hypothesis set. But here, the learner never makes a “hard” decision to eliminate any hypothesis. Instead, it assigns different query probability to each based on the estimated gap \( \hat{\Delta}^l_\ast \) for each hypothesis, That is what Line 12 and Line 13 are doing. To be specific:

- In Line 12, we divide the hypothesis into \( l + 1 \) sets based on \( \hat{\Delta}^l_\ast \). Again in the traditional elimination-style algorithm, the only remaining active hypothesis set is \( V_{l+1}^l \).
- In Line 13, based on these layered hypothesis sets, we are going to assign the query probability on the incoming \( x \). Intuitively, for each \( x \), we want to find the lowest policy set it belongs to, among all those layered sets. Then, because the lower the set is, the smaller its corresponding estimated gap is, so intuitively, we want to assign a higher query probability to those that have a lower corresponding hypothesis set.

### E Analysis for CALRuption

#### E.1 Notations

Let \( \mathcal{I}_l \) denotes the epoch \( l \), \( C_l \) denotes \( C_{\mathcal{I}_l} \).

#### E.2 Concentration guarantees on \( \delta \)-robust estimator

In this section, we show the analysis by using the Catoni’s estimator which is described in detail as below. Note that the same estimator has been used in previous works including [Wei et al. 2020], [Camilleri et al. 2021], [Lee et al. 2021].

**Lemma E.1.** (Concentration inequality for Catoni’s estimator [Wei et al. 2020]) Let \( F_0 \subset \cdots \subset F_n \) be a filtration, and \( X_1, \ldots, X_n \) be real random variables such that \( X_i \) is \( F_i \)-measurable. Let \( \mu_i \) be some fixed \( \mu_i \), and let \( \hat{\mu}_{n,\alpha} \) be the Catoni’s robust mean estimator of \( X_1, \ldots, X_n \) with a fixed parameter \( \alpha > 0 \), that is, \( \hat{\mu}_{n,\alpha} \) is the unique root of the function

\[
f(z) = \sum_{i=1}^{n} \psi(\alpha (X_i - z))
\]

where

\[
\psi(y) = \begin{cases} 
\ln(1 + y + y^2/2), & \text{if } y \geq 0 \\
-\ln(1 - y + y^2/2), & \text{else}
\end{cases}
\]

Then for any \( \delta \in (0, 1) \), as long as \( n \) is large enough such that \( n \geq \alpha^2 \left( V + \sum_{i=1}^{n} (\mu_i - \mu)^2 \right) + 2\log(1/\delta) \), we have with probability at least \( 1 - 2\delta \),

\[
|\hat{\mu}_{n,\alpha} - \mu| \leq \alpha \left( V + \sum_{i=1}^{n} (\mu_i - \mu)^2 \right) + \frac{2\log(1/\delta)}{\alpha n}
\]

\[
\leq \alpha (V + \sum_{i=1}^{n} \mu_i^2) + \frac{2\log(1/\delta)}{\alpha n}.
\]

**Lemma E.2** (Concentration inequality in our case). For any fixed epoch \( l \) and any pair of classifier \( h, h' \in \mathcal{H} \), as long as \( N_l \geq 4 \log(1/\delta) \), with probability at least \( 1 - \delta \), we have

\[
|\langle \hat{R}_l(h) - \hat{R}_l(h') \rangle - W_l^{h, h'}| \leq \sqrt{\frac{10 \log(1/\delta)}{N_l \min_{x \in \mathcal{D}(h, h')} Q_l^2}}
\]
where \( \hat{R}_t(h) = \frac{1}{|T|} \sum_{t \in T} \mathbb{E}_{y \sim \text{Ber}(\eta_t^*)} \left[ 1\{ h(x_t) \neq y \} \right] \) (restate).

**Proof.** First we calculate the expectation and variance of \( (\hat{l}_i(h) - \hat{l}_i(h')) \) for each \( t \in T \),
\[
\mathbb{E}_{y \sim \text{Ber}(\eta_t^*)} \mathbb{E}_{Q_t} \left[ \hat{l}_i(h) - \hat{l}_i(h') \right] = \mathbb{E}_{y \sim \text{Ber}(\eta_t^*)} \left[ 1\{ h(x_t) \neq y \} - 1\{ h'(x_t) \neq y \} \right]
\leq 1\{ h(x_t) \neq h'(x_t) \}
\]
and,
\[
\text{Var}_t \left( \hat{l}_i(h) - \hat{l}_i(h') \right) \leq \mathbb{E}_{y \sim \text{Ber}(\eta_t^*)} \mathbb{E}_{Q_t} \left[ (\hat{l}_i(h) - \hat{l}_i(h'))^2 \right]
= \mathbb{E}_{y \sim \text{Ber}(\eta_t^*)} \mathbb{E}_{Q_t} \left[ 1\{ h(x_t) \neq h'(x_t) \} \right]
\leq \frac{1\{ h(x_t) \neq h'(x_t) \}}{q_t^2}
\]

Then according to the Lemma E.1, we have
\[
\| (\hat{R}_t(h) - \hat{R}_t(h')) - W_t^{h,h'} \|
\leq 2 \alpha_t^{h,h'} \left( \frac{\sum \frac{1\{ h(x_t) \neq h'^*(x_t) \}}{\min_{x \in \text{Dis}(h,h')} q_t^2}}{N_t} + \frac{2 \log(1/\delta)}{\alpha_t^{h,h'} N_t} \right)
\leq 2 \alpha_t^{h,h'} \hat{\rho}_t(h,h') + \frac{2 \log(1/\delta)}{\alpha_t^{h,h'} N_t}
\leq \sqrt{\frac{10 \log(1/\delta) \hat{\rho}_t(h,h')}{N_t \min_{x \in \text{Dis}(h,h')} q_t^2}}
\]
The last one comes from choosing \( \alpha_t^{h,h'} = \sqrt{\frac{2 \log(1/\delta) \min_{x \in \text{Dis}(h,h')} q_t^2}{5 N_t \hat{\rho}_t(h,h')}} \) and also it is easy to verify that
\[
(\alpha_t^{h,h'})^2 \left( \frac{N_t \hat{\rho}_t(h,h')}{\min_{x \in \text{Dis}(h,h')} q_t^2} + \sum_t ((R_t(h) - R_t(h')) - (R_t(h) - R_t(h')))^2 \right) + 2 \log(1/\delta)
\leq 4 \log(1/\delta) \leq N_t.
\]

**E.3 High probability events**

Define the event \( \mathcal{E}_{\text{gap}} \) as
\[
\mathcal{E}_{\text{gap}} := \left\{ \forall i, \forall h, h' \in \mathcal{H}, |(\hat{R}_t(h) - \hat{R}_t(h')) - W_t^{h,h'}| \leq \sqrt{\frac{10 \beta_3 \hat{\rho}(h,h')}{N_t \min_{x \in \text{Dis}(h,h')} q_t^2}} \right\},
\]
and event \( \mathcal{E}_{\text{dis1}}, \mathcal{E}_{\text{dis2}} \) as
\[
\mathcal{E}_{\text{dis1}} := \left\{ \forall i, \forall h, h' \in \mathcal{H}, |\hat{\rho}(h,h') - \rho_*(h,h')| \leq \sqrt{\frac{\beta_3 \hat{\rho}(h,h')}{N_t}} + \frac{\beta_3}{N_t} \right\}
\]
\[
\mathcal{E}_{\text{dis2}} := \left\{ \forall i, \forall h, h' \in \mathcal{H}, |\hat{\rho}(h,h') - \rho_*(h,h')| \leq \sqrt{\frac{\beta_3 \rho_*(h,h')}{N_t}} + \frac{\beta_3}{N_t} \right\}.
\]

By condition \([\delta]\) of \( \delta \)-robust estimator in Algo 2, the (empirical) Bernstein inequality and the union bounds, we have easily get \( \mathbb{P}(\mathcal{E}_{\text{gap}} \cap \mathcal{E}_{\text{dis1}} \cap \mathcal{E}_{\text{dis2}}) \geq 1 - \delta \) as shown in the following lemmas.
Lemma E.3. $\mathbb{P}(\mathcal{E}_{\text{est}}) \geq 1 - \delta/3$

Proof. We prove this by condition 1 in Algo 2 and the union bound over $|\mathcal{H}|^2$ number of hypothesis pairs and $\frac{1}{2}|\log(n)|$ number of epochs.

Lemma E.4. $\mathbb{P}(\mathcal{E}_{\text{gap1}}) \geq 1 - \delta/3, \mathbb{P}(\mathcal{E}_{\text{gap2}}) \geq 1 - \delta/3$

Proof. We prove this by (empirical) Bernstein inequality in Algo 2 and the union bound over $|\mathcal{H}|^2$ number of hypothesis pairs and $\frac{1}{2}|\log(n)|$ number of epochs.

E.4 Gap estimation accuracy

In this section, we show that $\Delta^l_h$ is close to $\Delta_h$ for all $l, h$. To prove this, we first show some auxiliary lemmas as follows.

Lemma E.5 (Estimation accuracy for $\hat{D}_l$). On event $\mathcal{E}_{\text{gap}}$, for any fixed epoch $l$, for any fixed pair $h, h' \in \mathcal{H}$, suppose $j = \max\{i \mid h, h' \in V_i^l\}$, we have

$$
| (R_{\hat{D}_l}(h) - R_{\hat{D}_l}(h')) - (R_s(h) - R_s(h')) |
\leq \frac{1}{16} \left( \max\{\Delta^l_{h-1}, \Delta^l_{h'}-1\} + \epsilon_l \right) + \frac{4c_l}{N_l} R^* + \frac{2c_l}{N_l} \max\{\Delta_h, \Delta_{h'}\}
$$

Proof. Firstly we show that, for any pair $h, h' \in \mathcal{H}$, we have

$$
| (R_{\hat{D}_l}(h) - R_{\hat{D}_l}(h')) - (R_s(h) - R_s(h')) |
\leq \max_{h_1, h_2 \in \mathcal{H}} \left| (R_{\hat{D}_l}(h_1) - R_{\hat{D}_l}(h_2)) - W^h_{l, h_1, h_2} \sqrt{\min_{x \in \text{Dis}(h_1, h_2)} q_i^x} \frac{\hat{\rho}(h, h')}{\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x} \right|
+ |W^h_{l, h_1, h_2} - (R_t(h) - R_t(h'))| + |(R_t(h) - R_t(h')) - (R_s(h) - R_s(h'))|
\leq 2 \max_{h_1, h_2 \in \mathcal{H}} \left| (R_{\hat{D}_l}(h_1) - R_{\hat{D}_l}(h_2)) - W^h_{l, h_1, h_2} \sqrt{\min_{x \in \text{Dis}(h_1, h_2)} q_i^x} \frac{\hat{\rho}(h, h')}{\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x} \right|
+ |(R_t(h) - R_t(h')) - (R_s(h) - R_s(h'))|
\leq 2 \sqrt{\frac{10\beta_3}{N_l}} \frac{\hat{\rho}(h, h')}{\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x} + |(R_t(h) - R_t(h')) - (R_s(h) - R_s(h'))|
$$

The third inequality comes from the definition of $\hat{D}_l$ and the last inequality comes from the Condition 1 of $\delta$-robust estimator in Algo 2.

For the first term, for any $x \in \text{Dis}(h, h')$, by the definition of $q_i^x$ in line 13 and the fact that $(h, h') \in \mathcal{Z}(x)$, we have that

$$
q_i^x \geq \frac{\beta_1 \hat{\rho}(h, h')}{N_l} \epsilon_j^2, \text{ where } j = \max\{i \mid l - 1 \mid h, h' \in V_i^l\}
$$

So we can further lower bound the $\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x$ by

$$
\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x \geq \frac{\beta_1 \hat{\rho}(h, h')}{N_l} \epsilon_j^2, \text{ where } j = \max\{i \mid l - 1 \mid h, h' \in V_i^l\}
$$

and therefore upper bound the first term as

$$
2 \sqrt{\frac{10\beta_3}{N_l}} \frac{\hat{\rho}(h, h')}{\min_{x \in \text{Dis}(h, h', h') \cap \mathbb{Z}} q_i^x} \leq 2 \sqrt{\frac{10\beta_3}{\beta_1}} \epsilon_j.
$$
For the second term, by the definition of corruptions, we have
\[ |(\hat{R}_l(h) - \hat{R}_l(h')) - (R_*(h) - R_*(h'))| \]
\[ \leq |(\hat{R}_l(h) - \hat{R}_l(h')) - (\mathcal{R}_l(h) - \mathcal{R}_l(h'))| + |(\mathcal{R}_l(h) - \mathcal{R}_l(h')) - (R_*(h) - R_*(h'))| \]
\[ \leq 2\sqrt{\frac{2s}{N_1}} + 2C_l \rho_*(h, h') \]
\[ \leq 2\sqrt{\frac{2s}{N_1} \beta_1} + \frac{2C_l}{N_1} (\rho_*(h, h^*) + \rho_*(h', h^*)) \]
\[ \leq 2\sqrt{\frac{2s}{N_1} \beta_1} + 4C_l R^* + \frac{2C_l}{N_1} \max\{\Delta_h, \Delta_{h'}\} \]

where the second inequality comes from Bernstein inequality and Lemma \[A.2\]

Finally, we are going to make the connection between \( \epsilon_j \) and the \( \hat{\Delta}_{h_{l-1}}, \hat{\Delta}_{h'_{l-1}} \). Note that if \( j < l - 1 \), by definition of \( j \), we must have \( h, h' \notin V_{l+1}^1 \). By the definition that \( \forall h \notin V_{l+1}^1, \hat{\Delta}_h \geq \epsilon_i \), we have
\[ \max\{\hat{\Delta}_{h_{l-1}}, \hat{\Delta}_{h'_{l-1}}\} > \epsilon_{j+1} = \frac{\epsilon_j}{2} \]

and if \( j = l - 1 \), we directly have \( \frac{\epsilon_j}{2} \leq \epsilon_i \). Therefore, we have \( \frac{\epsilon_j}{2} \leq \max\{\hat{\Delta}_{h_{l-1}}, \hat{\Delta}_{h'_{l-1}}\} + \epsilon_i \).

\[ \square \]

**Lemma E.6** (Upper bound of the estimated gap). On event \( \mathcal{E}_{\text{gap}} \), for any fixed epoch \( l \), suppose its previous epoch satisfies that, for all \( h \in \mathcal{H} \),
\[ \Delta_h \leq \frac{3}{2} \hat{\Delta}_{h_{l-1}} + \frac{3}{2} \epsilon_{l-1} + 3g_{l-1}, \quad (2) \]
\[ \hat{\Delta}_{h_{l-1}} \leq 2(\Delta_h + \epsilon_{l-1} + g_{l-1}), \quad (3) \]

then we have,
\[ \hat{\Delta}_h \leq 2(\Delta_h + \epsilon_l + g_l) \]

where \( g_l = 2\frac{C_l}{N_1} \sum_{s=1}^{l} C_s \left( 2R^* \left\{ \frac{2C_s}{N_s} \leq \frac{1}{16} \right\} + 1 \left\{ \frac{2C_s}{N_s} > \frac{1}{16} \right\} \right) \).

**Proof.** According to the definition of \( \hat{\Delta}_h \), If \( \left( h - \hat{h}_l, \hat{\theta}_l \right) - \beta_2 \hat{\Delta}_{h_{l-1}} \leq \epsilon_i \), then the above trivially holds. Otherwise, we have
\[ \hat{\Delta}_h = R_{\mathcal{D}_l}(h) - (R_{\mathcal{D}_l}(\hat{h}_l) + \beta_2 \hat{\Delta}_{h_{l-1}}) \]
\[ = \left( (R_{\mathcal{D}_l}(h) - R_{\mathcal{D}_l}(\hat{h}_l)) - (R_*(h) - R_*(\hat{h}_l)) \right) + (R_*(h) - R_*(\hat{h}_l)) - \beta_2 \hat{\Delta}_{h_{l-1}} \]
\[ \leq \left( (R_{\mathcal{D}_l}(h) - R_{\mathcal{D}_l}(\hat{h}_l)) - (R_*(h) - R_*(\hat{h}_l)) \right) + \Delta_h - \beta_2 \hat{\Delta}_{h_{l-1}} \]
\[ \leq \frac{1}{16} \left( \max\{\hat{\Delta}_{h_{l-1}}, \hat{\Delta}_{h'_{l-1}}\} + \epsilon_l \right) + \frac{1}{16} \max\{\Delta_h, \Delta_{h_l}\} + \Delta_h - \beta_2 \hat{\Delta}_{h_{l-1}} \]
\[ + \frac{4C_l}{N_1} R^* \left\{ \frac{2C_l}{N_1} \leq \frac{1}{16} \right\} + \frac{2C_l}{N_1} \left\{ \frac{2C_l}{N_1} > \frac{1}{16} \right\} \]

Corruption Term
\[ = \frac{1}{16}(\hat{\Delta}_{h_{l-1}} + \epsilon_l) + \frac{1}{16} \Delta_h + \frac{1}{16} \Delta_{h_l} + \frac{1}{16} \Delta_{h'_{l-1}} + \frac{1}{16} \Delta_{h'_{l-1}} - \beta_2 \hat{\Delta}_{h_{l-1}} + \Delta_h + \text{Corruption Term} \]
\[ \leq \left( \frac{1}{16}(\hat{\Delta}_{h_{l-1}} + \epsilon_l) + \frac{1}{16} \Delta_h + \Delta_{h_l} \right) + \left( \frac{1}{16} \Delta_{h_{l-1}} + \frac{3}{32} \hat{\Delta}_{h_{l-1}} - \beta_2 \hat{\Delta}_{h_{l-1}} \right) + \frac{3}{32} (\epsilon_{l-1} + 2g_{l-1}) + \text{Corruption Term} \]
\[ \leq \left( \frac{1}{16}(\hat{\Delta}_{h_{l-1}} + \epsilon_l) + \frac{1}{16} \Delta_h + \Delta_{h_l} \right) + \frac{3}{32} (\epsilon_{l-1} + 2g_{l-1}) + \text{Corruption Term} \]
\[ = \frac{1}{16} \hat{\Delta}_{h_{l-1}} + \left( 1 + \frac{1}{16} \right) \Delta_h + \frac{1}{4} \epsilon_l + 4R^* \frac{C_l}{N_1} + \frac{3}{16} g_{l-1} \]
\[ \leq 2(\Delta_h + \epsilon_i + g_l) \]
Here the first inequality comes from the definition of $h^*$, the second inequality comes from Lemma [E.5], the third inequality comes from the assumption (1) and the penultimate inequality comes from the fact that $\beta_2 \geq \frac{1}{16}$. Finally, the last inequality comes from assumption (2).

\[ \Delta_h^I \leq 2 (\Delta_h^I + \epsilon_{l-1} + g_{l-1}) \]

Proof. We prove this by induction.

For the base case where $l = 1$, we can easily have the following

\[ \Delta_h^I \leq 1 \leq 2\Delta_h + 2\epsilon_1 + 2g_l \]

and also, by using Lemma [E.7] and the fact that $\Delta_h^I \leq 2(\Delta_h + \epsilon_0 + g_0)$, it is easy to get

\[ \Delta_h \leq \frac{3}{2} \Delta_h^I + \frac{3}{2} \epsilon_1 + 3g_l \]

So the target inequality holds for $l = 1$.

Suppose the target inequality holds for $l' - 1$ where $l' \geq 2$, then by Lemma [E.6] we show that the first target inequality holds for $l'$. Also by Lemma [E.7] we show that the second target inequality holds for $l'$. Therefore, we finish the proof.
E.5 Auxiliary lemmas

Lemma E.9. For any epoch $l$ and layer $j$, we have

$$\max_{h \in V_l^j} \rho_\ast(h, h^\ast) \leq 2R^* + 3\epsilon_j + 3g_{l-1}$$

Proof.

$$\max_{h \in V_l^j} \rho_\ast(h, h^\ast) \leq 2R^* + \max_{h \in V_l^j} \Delta_h$$

$$\leq 2R^* + \max_{h \in V_l^j} \left( \frac{3}{2} \hat{\Delta}_h^{l-1} \right) + \frac{3}{2} \epsilon_{l-1} + 3g_{l-1}$$

The first inequality comes from the fact that $\rho_\ast(h, h^\ast) \leq R_\ast(h) + R^* = 2R^* + \Delta_h$, the second inequality comes from the lower bound in Lemma E.8, and the last inequality is by the definition of $V_l^j$. \hfill \square

E.6 Main proof for Theorem 5.1

Here we assume $\log_4 \left( \frac{N}{n} \right) \notin \mathbb{N}$ and there are no corruptions in the last unfinished epoch $[\log_4 \left( \frac{N}{n} \right)]$. This will not affect the result but will make the proof easier. Given that events $\mathcal{E}_{\text{gap}}, \mathcal{E}_{\text{dis}1}$ and $\mathcal{E}_{\text{dis}2}$, then we have the following proofs.

First we deal with the sample complexity. For any $t \in \mathcal{I}_t$, the probability of $x_t$ being queried $(Q_t)$ is

$$\mathbb{E}[Q_t] = \sum_{x \in \mathcal{X}} P(x_t = x) q_t^x$$

$$= \sum_{x \in \mathcal{X}} P(x_t = x) \max_{(h, h') \in Z(x)} \frac{\beta_1 p_{l-1}(h, h')}{N_l} \epsilon_k(h, h', t)$$

$$\leq \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \max_{(h, h') \in Z(x)} \rho_\ast(h, h') \epsilon_k(h, h', t)$$

$$+ 4\frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \sqrt{\rho_\ast(h, h') \epsilon_k(h, h', t)} + 4\frac{\beta_1}{N_l}$$

$$\leq \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \max_{(h, h') \in Z(x)} \rho_\ast(h, h') \epsilon_k(h, h', t) + 8\frac{\beta_1}{N_l}$$

$$= \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \rho_\ast(h_1^t, h_z^t) \epsilon_j^{x^2} + 8\frac{\beta_1}{N_l}$$

$$\leq \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \max_{h_3, h_4 \in V_{l-1}^j} \rho_\ast(h_3, h_4) \epsilon_j^{x^2} + 8\frac{\beta_1}{N_l}$$

$$\leq \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \max_{h \in V_{l-1}^j} \rho_\ast(h, h^\ast) \epsilon_j^{x^2} + 8\frac{\beta_1}{N_l}$$

$$\leq \frac{\beta_1}{N_l} \sum_{x \in \mathcal{X}} P(x_t = x) \left( 2R^* \epsilon_j^{x^2} + 3\epsilon_j^{x_1} + 3g_{l-1} \epsilon_j^{x_2} \right) + 8\frac{\beta_1}{N_l}$$

$$= \frac{\beta_1}{N_l} \sum_{i=1}^{l-1} \left( 2R^* \epsilon_i^{x^2} + 3\epsilon_i^{x_1} + 3g_{l-1} \epsilon_i^{x_2} \right) \sum_{x \in \mathcal{X}} P(x_t = x) 1\{j^z = i\} + 8\frac{\beta_1}{N_l}$$

$$\leq \frac{\beta_1}{N_l} \sum_{i=0}^{l-1} \left( 2R^* \epsilon_i^{x^2} + 3\epsilon_i^{x_1} + 3g_{l-1} \epsilon_i^{x_2} \right) \mathbb{P}(x \in \text{Dis}(V_l^j)) + 8\frac{\beta_1}{N_l}$$
Here \((h^*_t, h^*_t) = \arg \max_{(h, h') \in Z(x)} \rho_s(h, h') \epsilon_{k(h, h', t)}^{-2}\) and \(j^* = k(h^*_t, h^*_t, l)\). The first inequality comes from the event \(\mathcal{E}_{dis_2}\), the second inequality comes from the fact that \(\sqrt{\rho_s(h, h') \epsilon_{k(h, h', t)}^{-2}} \leq \rho_s(h, h') \epsilon_{k(h, h', t)}^{-2} + 1\) and penultimate inequality comes from the Lemma \ref{lem:E9}

Now we can use the standard techniques to bound \(\mathbb{P}(x \in \text{Dis}(V^*_j))\) as follows

\[
\mathbb{P}(x \in \text{Dis}(V^*_j)) = \mathbb{P}(\exists h, h' \in V^*_i : h(x) \neq h'(x)) \\
\leq \mathbb{P}(\exists h \in V^*_i : h(x) \neq h^*(x)) \\
\leq \mathbb{P}(\exists h \in \mathcal{H} : h(x) \neq h^*(x), \rho_s(h, h^*) \leq 2R^* + 3\epsilon_i + 3g_{l-1}) \\
\leq \theta^* (2R^* + 3\epsilon_i + g_{l-1}) (2R^* + 3\epsilon_i + 3g_{l-1})
\]

where again the first inequality comes from Lemma \ref{lem:E9}

Combine with the above result, we get the expected number of queries inside a complete epoch \(l\) as,

\[
\sum_{i \in I^*_l} \mathbb{E}[Q_i] = 10\beta_1 \sum_{i=0}^{l-1} \theta^* (2R^* + 3\epsilon_i + g_{l-1}) \\
* (4(R^*)^2 \epsilon_i^{-2} + 12R^* \epsilon_i^{-1} + 12R^* g_{l-1} \epsilon_i^{-2} + 18g_{l-1} \epsilon_i^{-1} + 9g_{l-1}^2 \epsilon_i^{-2} + 9) \\
\leq 20\beta_1 \theta^* (2R^* + 3\epsilon_i + g_{l-1}) \\
* \left(4(R^*)^2 \epsilon_i^{-2} + 12R^* \epsilon_i^{-1} + \frac{24}{\beta_1} R^* C_{l-1} + \frac{36}{\beta_1} C_{l-1} \epsilon_i^{-1} + \frac{36}{\beta_1} \bar{C}^2_{l-1} \epsilon_i^{-1} + 9\right) \\
\leq 20\beta_1 \theta^* (2R^* + 3\epsilon_i + g_{l-1}) * \left(4(R^*)^2 \epsilon_i^{-2} + 12R^* \epsilon_i^{-1} + \frac{132}{\beta_1} \bar{C}_{l-1} + 10\right)
\]

where the second inequality comes from the fact that \(g_i = \frac{2}{\beta_1} \epsilon_i^2 \bar{C}_i\) and the third inequality comes from that fact that \(C_{l-1} \leq \sum_{s=1}^{l-1} C_s \leq 2\beta_1 \epsilon_i^{-2}\).

Summing over all \(L = \lceil \frac{1}{\theta} \log(n/\beta_1) \rceil\) number of epochs, we have that, for any \(n\),

\[
\begin{align*}
\text{Query complexity} \\
&\leq \sum_{l=1}^{L} \sum_{i \in I^*_l} \mathbb{E}[Q_i] \\
&\leq 40\beta_1 \theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) \left(4(R^*)^2 \epsilon_L^{-2} + 12R^* \epsilon_L^{-1}\right) \\
&\quad + 40\beta_1 \theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) L \left(\frac{132}{\beta_1} \bar{C}_{total} + 10\right) \\
&= 40\beta_1 \theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) \left(4(R^*)^2 \frac{n}{\beta_1} + 12R^* \sqrt{\frac{n}{\beta_1}} + 5\log(n/\beta_1)\right) \\
&\quad + 2450\theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) (\log(n/\beta_1)) \bar{C}_{total} \\
&= \theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) \left(160(R^*)^2 n + 480R^* \sqrt{n/\beta_1} + 200\beta_1 \log(n/\beta_1)\right) \\
&\quad + 2450\theta^* (2R^* + 3\epsilon_{L-1} + g_{L-1}) (\log(n/\beta_1)) \bar{C}_{total} \\
&\leq O \left(\theta^* (2R^* + 3\sqrt{\frac{\beta_1}{n} + \bar{C}_{total}}) ((R^*)^2 n + \log(n/\beta_1)) \beta_1\right) \\
&\quad + O \left(\theta^* (2R^* + 3\sqrt{\frac{\beta_1}{n} + \bar{C}_{total}}) \log(n/\beta_1) \bar{C}_{total}\right)
\end{align*}
\]

where the last inequality comes from the following lower bound,

\[
3\epsilon_{L-1} + g_{L-1} = 3\epsilon_{L-1} + \frac{2}{\beta_1} \bar{C}_{total} \epsilon_{L-1}^2 \geq 3\sqrt{\frac{\beta_1}{n} + \frac{2\bar{C}_{total}}{n}}
\]
Now we will deal with the correctness. By Lemma E.8 we have
\[
\Delta h_{\text{out}} \leq \frac{3}{2} \Delta h_{\text{out}}^{L-1} + \frac{3}{2} \epsilon_{L-1} + 3g_{L-1}
\]
\[
\leq 3\epsilon_{L-1} + 3g_{L-1}
\]
\[
\leq 6\sqrt{\frac{2\beta_1}{n}} + 3g_{L-1}
\]
\[
\leq 6\sqrt{\frac{2\beta_1}{n}} + 24 \frac{C_{\text{total}}}{n}
\]
where the second inequality comes from the definition of $h_{\text{out}}$ and $V_{L}^{L-1}$ and the third and last inequality is just by replacing the value of $\epsilon_{L-1}$ and $g_{L-1}$. Finally, we can write this result in the $\varepsilon$-accuracy form. Set $6\sqrt{\frac{2\beta_1}{n}} := \varepsilon$, we have $n = \frac{72\beta_1}{\varepsilon^2}$.

25