Deterministic Apple Tasting

Zachary Chase ZACHMAN99323@GMAIL.COM
University of California San Diego

Idan Mehalel IDANMEHALEL@GMAIL.COM
The Hebrew University

Editors: Nika Haghtalab and Ankur Moitra

Abstract

In binary $(0/1)$ online classification with apple tasting feedback, the learner receives feedback only when predicting 1. Besides some degenerate learning tasks, all previously known learning algorithms for this model are randomized. Consequently, prior to this work it was unknown whether deterministic apple tasting is generally feasible. In this work, we provide the first widely-applicable deterministic apple tasting learner, and show that in the realizable case, a hypothesis class is learnable if and only if it is deterministically learnable, confirming a conjecture of Raman et al. (2024). Quantitatively, we show that every class $\mathcal{H}$ is learnable with mistake bound $O\left(\sqrt{\mathtt{L}(\mathcal{H})T\log T}\right)$ (where $\mathtt{L}(\mathcal{H})$ is the Littlestone dimension of $\mathcal{H}$), and that this is tight for some classes. This demonstrates a separation between a deterministic and randomized learner, where the latter can learn every class with mistake bound $O\left(\sqrt{\mathtt{L}(\mathcal{H})T}\right)$, as shown in Raman et al. (2024).

We further study the agnostic case, in which the best hypothesis makes at most k many mistakes, and prove a trichotomy stating that every class $\mathcal{H}$ must be either easy, hard, or unlearnable. Easy classes have (both randomized and deterministic) mistake bound $\Theta_{\mathcal{H}}(k)$. Hard classes have randomized mistake bound $\tilde{\Theta}_{\mathcal{H}}\left(k+\sqrt{T}\right)$, and deterministic mistake bound $\tilde{\Theta}_{\mathcal{H}}\left(\sqrt{k\cdot T}\right)$, where T is the time horizon. Unlearnable classes have (both randomized and deterministic) mistake bound $\Theta(T)$.

Our upper bound is based on a deterministic algorithm for learning from expert advice with apple tasting feedback, a problem interesting in its own right. For this problem, we show that the optimal deterministic mistake bound is $\Theta\left(\sqrt{T(k+\log n)}\right)$ for all k and $T \leq n \leq 2^T$, where n is the number of experts. Our algorithm is an optimal (up to constant factors), natural and computationally efficient variation of the well-known exponential weights forecaster.

Keywords: Online learning, partial feedback, deterministic algorithms

1. Introduction

We study online classification under the apple tasting feedback model presented in Helmbold et al. (2000). In this problem, the learner only observes the correct label when predicting 1, in contrast to standard online learning where the learner observes the correct label after every prediction. In more detail, in the problem of online classification with apple tasting feedback, an *adversary* and a *learner* are rivals in a repeated game played for T many rounds. In each round t, the adversary provides an instance x_t from a domain $\mathcal{X}$ and chooses a label $y_t \in \{0,1\}$, the learner then decides (possibly at random) on a prediction $\hat{y}_t \in \{0,1\}$ and suffers the loss $\mathbb{1}[\hat{y}_t \neq y_t]$. If $\hat{y}_t = 1$, the adversary needs to send back the correct label $y_t \in \{0,1\}$ to the learner.

There are many examples for online learning scenarios in which the provided feedback follows the apple tasting model. In many of them, the goal is to separate between good and bad instances, where the bad instances are trashed and their classification cannot be verified. A common example is the problem of *spam filtering*. When a spam-filtering algorithm marks an e-mail as spam, it is moved to the spam folder and usually not presented to the user, which thus cannot verify that the e-mail is indeed spam. On the other hand, an e-mail marked by the algorithm as legitimate will go straight to the inbox, and the user can then indicate to the algorithm in case of a mistake, when the e-mail is spam.

Online classification with apple tasting feedback was previously studied by Helmbold et al. (2000); Raman et al. (2024) However, the algorithms proposed in those works are randomized (except for some specific degenerate tasks). Consequently, prior to this work it was unknown if deterministic apple tasting is even possible. This is in contrast with other online learning settings, in which deterministic algorithms are usually studied, and considered as fundamental. Notable examples are the most basic SOA of Littlestone (1988) and the *Weighted Majority* algorithm of Littlestone and Warmuth (1994) for standard online classification, as well as the *binoimal weights* algorithm for prediction with expert advice Cesa-Bianchi et al. (1996). For online classification with bandit feedback, deterministic algorithms were presented in Auer and Long (1999); Daniely et al. (2015); Long (2020).

It is not a coincidence that in other online learning settings deterministic algorithms were designed and analyzed. There are many reasons to prefer deterministic (instead of randomized) learners, when possible:

1. Mistake bounds of deterministic learners hold with probability 1 and not in expectation.
2. Randomness is a resource. Like any other resource, it could be missing or expensive.
3. We do not have to worry about measurability constraints on the concept class we wish to learn.
4. The predictions of deterministic learners can be simulated in advance, so every deterministic learner has the same guarantees for either an oblivious or an adaptive adversary.
5. We do not need to worry that the adversary might have knowledge on the internal mechanism generating the algorithm's randomness.
6. The introduction of randomized algorithms can be seen as a necessity to overcome the impossibility of achieving sublinear regret by deterministic algorithms, as shown in Cover (1967) by a simple argument. It is more natural to consider deterministic algorithms in realizable settings, where the argument of Cover (1967) does not hold.

This discussion raises the following natural main question, raised also by Raman et al. (2024), and guiding our work:

> Is it possible to learn deterministically under apple tasting feedback?

We answer this question positively: any class that can be learned in the easiest standard online learning setting, can be also learned deterministically, when only apple tasting feedback is provided. This affirmative answer raises the problem of finding the best possible deterministic mistake bounds achievable under apple tasting feedback. In Section 1.1 we present mistake bounds that are optimal in the sense that for some classes, they are the best possible (up to a constant factor). We also provide mistake bounds that are universally tight for every class, but only up to logarithmic factors of T and constants depending on the class.

1.1. Main results

To state the results, we present some of the notation used in the paper. All notations are formally defined in Section A. Let $\mathcal{H} \in \{0,1\}^{\mathcal{X}}$ be a *concept class*, or a *hypothesis class*, where $\mathcal{X}$ is a *domain* of instances. A pair $(x, y) \in \mathcal{X} \times \{0,1\}$ is called an *example*.

1.1.1. LEARNABILITY

We say that $\mathcal{H}$ is *deterministically learnable under apple tasting feedback* (or *deterministically learnable*, in short), if there exists a sub-linear function $M : \mathbb{N} \to \mathbb{N}$ and a deterministic learner that, for any $T \in \mathbb{N}$, $h^\star \in \mathcal{H}$, and a *realizable* input sequence of examples $S = (x_1, h^\star(x_1)), \ldots, (x_T, h^\star(x_T))$, makes at most $M(T)$ prediction mistakes on S. The sequence S is called realizable since its labels are realized by $h^\star \in \mathcal{H}$. We also use the analogue natural definitions for randomized learnability with apple tasting feedback, and for learnability with full information feedback.[1]

We first wish to understand which classes are deterministically learnable. We show that even when receiving only apple tasting feedback and forcing the learner to be deterministic, the exact same classes which are learnable in the easiest standard online learning setting, are learnable in this difficult setting as well.

Theorem 1 (Qualitative characterization) *The following conditions are equivalent for every class $\mathcal{H}$.*

1. *$\mathcal{H}$ is learnable with full-information feedback.*
2. *$\mathcal{H}$ is learnable with apple tasting feedback.*
3. *$\mathcal{H}$ is deterministically learnable with apple tasting feedback.*

This characterization follows from our most basic quantitative upper bound on the mistake bound for learning a class $\mathcal{H}$, which we will now discuss. For any class $\mathcal{H}$ and time horizon T, let $\mathtt{M}^\star(\mathcal{H}, T)$ be the number of mistakes made by an optimal deterministic learner for $\mathcal{H}$, on the worst-case realizable sequence S of examples of length T. We sometimes refer to this quantity as the deterministic *mistake bound* (or just mistake bound, when the context is clear) of $\mathcal{H}$ with time horizon T. In his seminal work Littlestone (1988), Littlestone proved that the optimal deterministic mistake bound of any class $\mathcal{H}$, and for any time horizon T in standard online learning (with full-information feedback) equals exactly to the *Littlestone dimension* of $\mathcal{H}$, denoted by $\mathtt{L}(\mathcal{H})$. We prove the following mistake bounds in terms of $\mathtt{L}(\mathcal{H})$ and T.

Theorem 2 (Quantitative bounds - realizable case) *For every class $\mathcal{H}$ and time horizon T:*

$$\mathtt{M}^\star(\mathcal{H}, T) = O\Big(\sqrt{\mathtt{L}(\mathcal{H}) T \log T}\Big).$$

1. In the case of full-information feedback, the learnability definition is stronger and requires that there exists $M^\star \in \mathbb{N}$ so that $M(T) \le M^\star$ for all T.

Furthermore, for every natural $d \geq 1$ there exists a class $\mathcal{H}$ and $T_0(d)$ such that $\mathtt{L}(\mathcal{H}) = O(d)$ and

$$\mathtt{M}^\star(\mathcal{H}, T) = \Omega\left(\sqrt{dT \log T}\right)$$

for all $T \geq T_0(d)$.

All of our upper bounds do not require that T is given to the learner. The lower bound implies a strict separation between randomized and deterministic learners, since Raman et al. (2024) showed that every class can be learned with randomized mistake bound $O\left(\sqrt{\mathtt{L}(\mathcal{H})T}\right)$.

1.1.2. AGNOSTIC LEARNING

It is interesting to prove a version of Theorem 2 for the *agnostic* setting, in which the input sequence S can be any sequence of examples $(x_1, y_1), \ldots, (x_T, y_T)$, and not necessarily choose the labels $\{y_t\}_{t=1}^T$ to be $\{h^\star(x_t)\}_{t=1}^T$ for some $h^\star \in \mathcal{H}$. In this setting, we measure $\mathtt{M}^\star(\mathcal{H}, T, k)$, which is the number of mistakes made by an optimal learner on the worst-case k-*realizable* sequence S of length T. The sequence $S = (x_1, y_1), \ldots, (x_T, y_T)$ is k-realizable if there exists $h^\star \in \mathcal{H}$ so that $y_t \neq h^\star(x_t)$ for at most k many indices t. We often call k the *realizability* parameter, as it indicates how far is S from being realizable. All of our bounds do not require that k is given to the learner. Theorem 2 generalizes to the agnostic setting as follows.

Theorem 3 (Quantitative bounds - agnostic case) *For every class $\mathcal{H}$, time horizon T, and realizability parameter k:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O\left(\sqrt{T(k + \mathtt{L}(\mathcal{H}) \log T)}\right).$$

Furthermore, for every natural $d \geq 1$ there exists a class $\mathcal{H}$ and $T_0(d)$ such that $\mathtt{L}(\mathcal{H}) = O(d)$ and for every realizability parameter k and $T \geq T_0(d)$:

$$\mathtt{M}^\star(\mathcal{H}, T, k) = \Omega\left(\sqrt{T(k + d \log T)}\right).$$

Theorem 3 is implied by combining Theorem 14, Theorem 25, and Theorem 23.

1.1.3. NEARLY-TIGHT UNIVERSAL MISTAKE BOUNDS

It is interesting to obtain mistake bounds that are nearly tight for all classes, and not just for the specific classes mentioned in the lower bound of Theorem 3. When fixing the class $\mathcal{H}$ and taking $T \to \infty$, we are able to obtain bounds which are tight up to logarithmic factors for every class. In Raman et al. (2024), a new combinatorial dimension of hypothesis classes coined *effective width* is defined and proved to establish a trichotomy in the randomized mistake bounds of hypothesis classes in the realizable setting. We recall the definition of effective width from Raman et al. (2024) in the beginning of Section C.

The following bounds are proved in Raman et al. (2024). Let $\mathtt{w}(\mathcal{H})$ be the effective width of a class $\mathcal{H}$, and let $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T)$, $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k)$ be the analogue definitions of $\mathtt{M}^\star(\mathcal{H}, T)$, $\mathtt{M}^\star(\mathcal{H}, T, k)$ when randomized learners are allowed. If $\mathtt{w}(\mathcal{H}) = 1$ then $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T) = \Theta_{\mathcal{H}}(1)$, where the notation $\Theta_{\mathcal{H}}(\cdot)$ hides constants depending on $\mathcal{H}$. We call these *easy* classes. Othrewise, if $1 < \mathtt{w}(\mathcal{H}) < \infty$ then $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T) = \Theta_{\mathcal{H}}\left(\sqrt{T}\right)$. We call these *hard* classes. Finally, if $\mathtt{w}(\mathcal{H}) = \infty$ then $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T) =$

$\Theta(T)$. We call these *unlearnable* classes. We prove that the same partition of all classes to easy, hard, and unlearnable classes by the value of their effective width (1, finite larger than 1, or infinite) provides a trichotomy also for deterministic mistake bounds, and also for the agnostic setting, both for randomized and deterministic learners. However, our results demonstrate a large separation between deterministic and randomized mistake bounds of hard classes in the agnostic setting, as described in the following theorem.

Theorem 4 (Mistake bound trichotomy) *Fix a class $\mathcal{H}$. Let $k \in \mathbb{N}$ be the realizability parameter, and let $T \geq k$ be the time horizon. Then:*

1. *If $\mathcal{H}$ is easy ($\mathtt{w}(\mathcal{H}) = 1$) then:*

$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k), \mathtt{M}^\star(\mathcal{H}, T, k) = \Theta_{\mathcal{H}}(k).$$

2. *If $\mathcal{H}$ is hard ($1 < \mathtt{w}(\mathcal{H}) < \infty$) then:*

$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k) = \Theta(k) + \tilde{\Theta}_{\mathcal{H}}\left(\sqrt{T}\right), \quad \mathtt{M}^\star(\mathcal{H}, T, k) = \Theta\left(\sqrt{Tk}\right) + \tilde{\Theta}_{\mathcal{H}}\left(\sqrt{T}\right).$$

3. *If $\mathcal{H}$ is unlearnable ($\mathtt{w}(\mathcal{H}) = \infty$) then:*

$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k), \mathtt{M}^\star(\mathcal{H}, T, k) = \Theta(T).$$

We prove Theorem 4 in Section C.3. Theorem 4 implies a large seperation between randomized and deterministic learning of hard classes: for any hard class $\mathcal{H}$, if k grows for example as $\sqrt{T}$, then the randomized mistake bound grows at most as $\sqrt{T \log T}$, while the deterministic mistake bound grows as $T^{3/4}$.

1.1.4. PREDICTION WITH EXPERT ADVICE

One of the main building blocks in this paper, which is interesting on its own right, is a deterministic algorithm for *prediction with expert advice* with apple tasting feedback. Our algorithm is the first deterministic learner for this problem, and it is optimal (up to constant factors) for reasonably large enough number of experts n. It uses a novel variation of the well-known exponential weights forecaster (see (Cesa-Bianchi and Lugosi, 2006, Section 2.2)). This problem is similar to the problem of learning hypothesis classes, only that instead of hypotheses, having predictions that are determined by the provided instance x_t, we have experts that decide on their predictions in every round, possibly in an adversarial manner. For this problem, we prove the following bounds. Let $\mathtt{M}^\star(n, T, k)$ be the optimal mistake bound in the instance of the problem where there are n experts, the time horizon is T and the best expert makes at most k many mistakes.

Theorem 5 (Optimal bound for experts) *For every n, T, k so that $k \leq T \leq n < 2^{T/2}$ and $T \geq 2$:*

$$\mathtt{M}^\star(n, T, k) = \Theta\left(\sqrt{T(k + \log n)}\right).$$

We prove Theorem 5 in Section B.1 (upper bound) and Section B.2 (lower bound). To provide some understanding of the techniques used in this paper in the main text, we provide a complete proof of the upper bound of Theorem 5 in the realizable case ($k = 0$) in Section 3.

Similar results for randomized learners were proved in Helmbold et al. (2000); Raman et al. (2024). Raman et al. (2024) proved a randomized upper bound of $k + O\big(\sqrt{T \log n}\big)$. Helmbold et al. (2000) studied the realizable case with an oblivious adversary (all expert's predictions are fixed in advance) and showed that the optimal randomized mistake bound in this setting is $\Theta\left(\sqrt{\frac{T \log n}{\log \frac{T}{\log n}}}\right)$ in the range of parameters of Theorem 5 (in their work, they proved tight bounds for all regimes of T, n).

Following these results, it is interesting to observe that in the experts' setting, a strict separation between randomized and deterministic learners in the agnostic case appears via a simple argument, given that n is sufficiently large. Consider an adversary that operates in $\sqrt{T/k}$ many phases, each of $\sqrt{Tk}$ many rounds, where in phase i the only expert predicting 1 is i, and the correct label reported by the adversary is always 0. It is not difficult to prove that $\sqrt{Tk}$ many mistakes can be forced on a deterministic learner regardless of its predictions, which is impossible for an optimal randomized learner. A formal statement and proof are given in Theorem 13.

1.2. Additional related work

In this section, we mention some additional related work not mentioned in Section 1.1.

Apple tasting feedback, like other partial feedback models, can be seen as a special case of the more general graph-feedback model Mannor and Shamir (2011); Alon et al. (2015), which in turn is a special case of partial monitoring Bartók et al. (2014). The more famous bandit-feedback model Daniely et al. (2015); Daniely and Helbertal (2013) is also a special case of the graph-feedback model.

This work aims to understand how the restriction of using only deterministic learners affects the mistake bounds in the apple-tasting feedback model. Problems in the same vein were studied for other feedback models: full-information feedback in Abernethy et al. (2006); Filmus et al. (2023), and bandit-feedback in Filmus et al. (2024).

Most of our upper bounds are based on a novel algorithm we present for the problem of *prediction with expert advice*, which is a variation of the *exponentially weighted average forecaster* of Cesa-Bianchi and Lugosi (2006). Using this algorithm, we prove tight bounds that hold under the restriction that both the experts' and learner's predictions are deterministic, and only apple-tasting feedback is provided. Many variations of this problem were extensively studied in the past decades. The most basic setting of binary online learning with full-information feedback was studied, e.g, in Vovk (1990); Littlestone and Warmuth (1994); Cesa-Bianchi et al. (1997); Abernethy et al. (2006); Filmus et al. (2023). The work of Cesa-Bianchi et al. (1996) studied this setting with deterministic predictions. In Brânzei and Peres (2019), the multiclass version of the problem is studied. The seminal work of Auer et al. (2002) studied the bandit-feedback variation of the problem.

2. Technical overview

In this section, we informally describe the ideas behind the main novel techniques used to prove our results. We focus on the realizable setting, which already captures many of the ideas we find intereseting.

Section organization. In Section 2.1 we overview our algorithm for the problem of prediction with expert advice, which attains the upper bound of Theorem 5 in the realizable case $(k = 0)$. This result is a main building block of the results proved in this paper. We then explain in Section 2.2 how to use it to prove the upper bound of Theorem 2. Finally, in Section 2.3 we explain how to construct the classes mentioned in the lower bound of Theorem 2.

2.1. Technical overview: upper bound for prediction with expert advice

Recall that in the realizable setting of prediction with expert advice with apple tasting feedback, we have n many experts, where each expert i provides a prediction $\hat{y}_t^{(i)} \in \{0,1\}$ on each round t, such that there exists $i^\star \in [n]$ for which $\hat{y}_t^{(i^\star)} = y_t$ for all t (where y_t is the correct label). The learner's goal, as usual, is to minimize the number of rounds where it errs. The realizability assumption implies that whenever the learner predicts $\hat{y}_t = 1$ but $y_t = 0$, it can eliminate all experts predicting 1 in this round. Note that these rounds are exactly the false positives, and denote their number with M_+. Let $M_- = |t \in [T] : \hat{y}_t = 0, \hat{y}_t^{(i^\star)} = 1|$, which is exactly the number of false negatives. The total number of mistake M made by the learner is thus exactly $M = M_+ + M_-$. Therefore, the learner wishes to balance M_+ and M_-, such that both are not too high.

Consider the following two extreme scenarios, which are very easily handled by the learner:

1. In every round, the adversary let exactly a experts to predict 1, where $a \geq 1$ is some constant.
2. In every round, the adversary let exactly $b \cdot n$ experts to predict 1, where $b < 1$ is some constant.

We call the experts that are not yet eliminated the *living* experts. In the first scenario, the learner can predict 1 only when some living expert i has $|t \in [T] : \hat{y}_t^{(i)} = 1, \hat{y}_t = 0| \geq \sqrt{T}$. This strategy is similar to the deterministic learner of Helmbold et al. (2000) for the class of singletons, and produces $\mathtt{M}_+, M_- = O\left(\sqrt{T}\right)$. In the second scenario, the learner can predict 1 in every round, which produce $M_- = 0$ and $M_+ = O(\log n)$.

Note that in the first scenario, the adversary always let very few experts to predict 1. It thus wastes many rounds on increasing the value of $|t \in [T] : \hat{y}_t^{(i)} = 1, \hat{y}_t = 0|$, which reflects M_- in case that i realizes the input sequence, for just a small number of experts. Therefore, in this scenario the learner benefits from making most of its predictions 0. The second scenario is the other extreme, in which the adversary allows the learner to eliminate many experts in every round in which it predicts 1 and make a mistake, and thus the learner benefits from always predicting 1. The same idea of course applies even if the adversary follows the first scenario in some rounds, and the second scenario in other rounds. Unfortunately, the adversary is not limited to those scenarios, and can follow many other intermediate scenarios, which are harder to handle. Roughly speaking, to handle it we need to "smooth" the binary treatment of both extreme scenarios to some unified strategy. This is done by using *exponential weights*. For every expert i, we denote $d_t^{(i)} = |t \in [T] : \hat{y}_t^{(i)} = 1, \hat{y}_t = 0|$, and call $d_t^{(i)}$ the *distance* of expert i in round t. The distance of expert i reflects the number of false negatives to be made by the learner in case that the input sequence is realized by expert i. We define $2^{\eta d_t^{(i)}}$ to be the *weight* of expert i in round t, where η is a parameter to be chosen optimally. In every round t, our algorithm predicts 1 if and only if the sum of weights of living experts predicting 1 is larger than n. Namely, the algorithm predicts 1 if and only if:

$$\sum_{i \in V_t} \mathbb{1}[\hat{y}_t^{(i)} = 1] 2^{\eta d_t^{(i)}} \geq n, \tag{1}$$

where V_t is the set of living experts in round t. Note that this strategy handles the two extreme scenarios well: If in round t just a few living experts predict 1 and the learner predicts 1 as well, then those experts must have already travelled a long distance, which means that the adversary has wasted a lot of rounds on increasing their distance. On the other hand, if many experts predict 1, then even if all of them have low distances, condition (1) is satisfied. Using an exponential value of the distance as weight in condition (1) produces a balanced strategy that can handle all adversarial strategies. We formalize this in the proof the upper bound of Theorem 5 for the realizable case ($k = 0$), in Section 3. In a nutshell, note that $M_- \leq \frac{\log n}{\eta}$, since any expert can reach distance at most $\frac{\log n}{\eta}$ by condition (1). To bound M_+, we show that in order to make the learner predict 1, the adversary must either waste many rounds, or many experts to be eliminated if the learner predicts 1. The bound we get is $M_+ = O(\eta T)$. Thus, we get in total

$$M = O\left(\frac{\log n}{\eta} + \eta T\right). \tag{2}$$

Choosing $\eta = \sqrt{\frac{\log n}{T}}$ gives the upper bound $O\big(\sqrt{T \log n}\big)$ of Theorem 5 when $k = 0$. Remarkably, the bound (2) is exactly the regret bound for agnostic prediction with expert advice under full information feedback appearing in (Cesa-Bianchi and Lugosi, 2006, Corollary 2.2). This does not seem like a coincidence: our algorithm is a variation of their *exponentially weighted average forecaster* achieving the regret bound of their Corollary 2.2. However, we do not know if there is a direct rigorous connection between detereministic realizable prediction with expert advice under apple tasting feedback to agnostic prediction with expert advice under full-information feedback (for example, is there a reduction between the problems in some direction?).

2.2. Technical overview: upper bound for learning hypothesis classes

To prove the upper bound of Theorem 2, we use a reduction from learning hypothesis classes to the problem of prediction with expert advice, proved and used in Ben-David et al. (2009) for agnostic learning with full-information feedback. They showed that for every class $\mathcal{H}$ and every T, there exists a single class of $n = T^{\mathtt{L}(\mathcal{H})}$ many experts that *covers* any sequence of instances $\{x_t\}_{t=1}^T$ of length T with respect to $\mathcal{H}$, where "covers" means that for every $h \in \mathcal{H}$, there exists $i \in [n]$ so that $\hat{y}_t^{(i)} = h(x_t)$ for every $t \in [T]$. This means that there is essentially no difference between learning $\mathcal{H}$, and predicting using the advice of the covering experts. Applying the bound of Theorem 5 with $n = T^{\mathtt{L}(\mathcal{H})}$ gives the upper bound in Theorem 2.

2.3. Technical overview: lower bound for specific hypothesis classes

The lower bound of Theorem 2 claims that for every natural d larger than some universal constant there exists $T_0(d)$ and a class $\mathcal{H}$ such that $\mathtt{L}(\mathcal{H}) = O(d)$ and $\mathtt{M}^\star(\mathcal{H}, T) = \Omega\big(\sqrt{dT \log T}\big)$ for all $T \geq T_0(d)$. The class of n experts can be seen as a concept class $\mathcal{H}$ of size n, and it has $\mathtt{L}(\mathcal{H}) = \lfloor \log n \rfloor$. Therefore, using this class directly with the lower bound of Theorem 5 will only give a lower bound of $\Omega\left(\sqrt{\mathtt{L}(\mathcal{H})T}\right)$. However, looking into the proof of the lower bound in Theorem 5 reveals that in contrast with full-information lower bounds, the "hard" experts' predictions used by the adversary are very unbalanced, in the sense that many experts predict 0 and only few predict 1. Since $\mathtt{L}(\mathcal{H}) = \Omega(\log n)$ comes from choosing experts' predictions which are as balanced as possible, this intuitively means that we can restrict the adversary such that only unbalanced experts' predictions are available, in a way that decreases the Littlestone dimension from $\Omega(\log n)$ to $O(\log_T n)$, but maintains the $\Omega\big(\sqrt{T\log n}\big)$ lower bound from Theorem 5. Choosing $n = T^d$ results in a class $\mathcal{H}$ with $\mathtt{L}(\mathcal{H}) = O(d)$ that maintains a lower bound of $\Omega(T\log T^d)$. In Section C.4, we provide a non-constructive proof of existence of $\mathcal{H}$. Finding an explicit class $\mathcal{H}$ realizing this lower bound remains open.

The technique explained above inherently assumes that T is given in advance, before choosing the concept class used for the lower bound. As this is not the case, we still need to remove this assumption. This is done by "gluing" together instances of the class described above for all $T \geq T_0(d)$, in a way that keeps the Littlestone dimension $O(d)$. The details of how to glue the classes can be found in Section C.4.

3. Prediction with expert advice: realizable case

To provide some understanding of the techniques used in this paper, we incldue here a complete proof of the upper bound of Theorem 5 in the realizable case ($k = 0$). That is, we will prove that

$$\mathtt{M}^\star(n, T) = O\Big(\sqrt{T\log n}\Big),$$

where $\mathtt{M}^\star(n,T) = \mathtt{M}^\star(n,T,0)$. The general case $k \geq 0$ is proved in Section B.1.

Concretely, we will study the optimal relizable mistake bound of a specific class called the *universal class* (of size n). For every $n \geq 1$, the universal class $\mathcal{U}_n = \{h_1,\dots,h_n\}$ is defined as follows. Let $\mathcal{X}_n = \{0,1\}^n$ be the instance domain. For every $i \in [n]$, the hypothesis h_i is defined to be $h_i(x) = x_i$ for every $x \in \mathcal{X}_n$, where x_i denotes the i'th entry of x.

Note that every partition of $\mathcal{U}_n$ to hypotheses predicting 0 and hypotheses predicting 1 has an appropriate $x \in \mathcal{X}_n$ inducing it. Therefore, learning $\mathcal{U}_n$ is equivalent to the game of *prediction with expert advice with apple tasting feedback*, when the experts provide deterministic predictions. Formally, each round t of the game proceeds as follows:

(i) The n experts present predictions $\hat{y}_t^{(1)},\dots,\hat{y}_t^{(n)} \in \{0,1\}$.

(ii) The learner predicts a value $\hat{y}_t \in \{0,1\}$ and suffers the *loss* $\mathbb{1}[\hat{y}_t \neq y_t]$.

(iii) If $\hat{y}_t = 1$, the adversary reveals y_t to the learner.

In the realizable setting, there is an expert who never errs. The upshot of viewing $\mathcal{U}_n$ as a class of experts is twofold: first, it makes the problem easier to describe and analyze. Second, it allows to pick the experts to be learning algorithms and extend the solution to all concept classes, as done in Section C.

For brevity, we will usually refer to the experts simply by their indices $[n]$.

Our upper bound on $\mathtt{M}^\star(n,T)$ is attained by the RealizableExpAT learner presented in Figure 1.

Theorem 6 *For every time horizon $T \geq 2$, number of experts $n \geq 2$, and a realizable input sequence S of experts' predictions and true labels of length T:*

$$\mathtt{M}(\mathsf{RealizableExpAT}, S) = O\Big(\sqrt{T\log n}\Big).$$

Furthermore, RealizableExpAT *is computationally efficient.*

In section B.2 we show that the upper bound in Theorem 6 is tight for $n \geq T$. Our algorithm is a variation of the known exponential weights forecaster (Cesa-Bianchi and Lugosi, 2006, Section 2.2). Let us briefly describe RealizableExpAT. We maintain a version space V_t containing of all experts that have been consistent with the feedback gathered until round $t-1$. Initially, V_1 contains all the experts. For each expert $j \in V_t$, we maintain a weight which will be a function of its *distance* $d_t^{(j)}$. The distance of an expert in round t is defined to be the number of rounds in which it has predicted 1 while the learner predicted 0, until round $t-1$. We define the *weight* of expert j in round t to be $2^{\eta d_t^{(j)}}$, where η is a parameter to be optimized. The algorithm's decision making mechanism is very simple and efficient: it predicts 1 if and only if the total sum of weights of all experts in V_t predicting 1 exceeds some threshold L, where L is a parameter to be optimized. If the learner has made a mistake when predicting 1, it removes all experts predicting 1 from the version space.

RealizableExpAT

Input: Set of experts indexed by $[n]$, learning parameters $\eta \in (0,1), L > 1$.
Initialize: Let $V^{(1)} = [n]$, let $d_1^{(j)} = 0$ for all $j \in [n]$.

for $t = 1, \ldots, T$**:**

1. Receive experts predictions $\hat{y}_t^{(1)}, \ldots, \hat{y}_t^{(n)}$.
2. Predict $\hat{y}_t = 1$ if and only if
$$\sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{\eta d_t^{(j)}} \geq L.$$
3. If $\hat{y}_t = 1$:
   (a) Set $d_{t+1}^{(j)} = d_t^{(j)}$ for every expert $j \in V_t$.
   (b) If $y_t = 0$: set $V_{t+1} = V_t \backslash \{j \in V_t : \hat{y}_t^{(j)} = 1\}$.
4. If $\hat{y}_t = 0$: set $d_{t+1}^{(j)} = d_t^{(j)} + \mathbb{1}[\hat{y}_t^{(j)} = 1]$ for every expert $j \in V_t$.

Figure 1: A deterministic apple tasting learner for realizable prediction with expert advice.

We will now analyze the algorithm's mistake bound. Towards this end, we introduce some notation. Fix the number of experts $n \geq 2$, the horizon $T \geq 2$, and an execution of RealizableExpAT on a sequence S of experts' predictions and true labels. A mistake in which the algorithm predicts 1 is called a *false positive*, and a mistake in which the algorithm predicts 0 is called *false negative*. For $r \in \{0,1\}$, let T_r be the set of rounds in which RealizableExpAT predicts r. Let $T_- = \{t \in T_0 : y_t = 1\}$ be the set of rounds in which RealizableExpAT makes a false negative mistake, and let $T_+ = \{t \in T_1 : y_t = 0\}$ be the set of rounds in which RealizableExpAT makes a false positive mistake. Denote $M_- = |T_-|$, $M_+ = |T_+|$. For $d \in \mathbb{N}$, let $I_d = [d/\eta, (d+1)/\eta]$. For every expert j, let D_j be it's distance in the last round where it is still in the version space, or at round T if $j \in V_T$. Let $N_d = \{j : D_j \in I_d\}$ and $n_d = |N_d|$.

Lemma 7 *For any learning parameters $\eta \in (0,1), L > 1$ satisfying $n \leq \eta LT$, we have:*

$$M_- \leq \frac{\log L}{\eta} + 1, \quad \textit{and} \quad M_+ \leq 6\eta T.$$

The challenging part in proving Lemma 7 is to upper bound M_+. The key ingredient of the proof of this bound is an inequality proved in Lemma 8 by forming lower and upper bounds on the *total weighted distance gained* by experts in the version space. Formally, we will lower and upper bound the quantity

$$\texttt{TWDG} := \sum_{t \in T_0} \sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{\eta d_t^{(j)}}.$$

Indeed, by multiplying the weight of expert j in round $t \in T_0$ by $\mathbb{1}[\hat{y}_t^{(j)} = 1]$, we make sure to take into account its current weight only in rounds where it gains distance. We sum this quantity over all rounds and all experts to get TWDG.

Lemma 8 *Suppose that $n \leq \eta T L$. Then:*

$$\sum_{d \in \mathbb{N}} n_d 2^d \leq 3\eta T L.$$

Proof The main part of the proof is to lower and upper bound TWDG. We start with the upper bound. By definition of the algorithm, for every $t \in T_0$ it holds that

$$\sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{\eta d_t^{(j)}} < L,$$

which implies $\texttt{TWDG} \leq LT_0 \leq LT$. For the lower bound, note that for every $d \geq 1$, every expert j with $D_j \in I_d$ must have gained distance for $1/\eta$ many times while having weight at least $2^{\eta(d-1)/\eta} = 2^{d-1}$. Therefore:

$$\texttt{TWDG} \geq \sum_{d \geq 1} n_d \frac{1}{\eta} 2^{d-1}.$$

Combined with the upper bound $\texttt{TWDG} \leq LT$, this implies

$$\sum_{d \geq 1} n_d 2^d \leq 2\eta T L.$$

Applying the assumption $n \leq \eta T L$ gives the statement of the lemma. ■

We can now prove Lemma 7.

Proof [Proof of Lemma 7] Let's first bound M_-. Let $j^\star$ be an expert who makes no mistakes. In each of the rounds of T_-, the distance of $j^\star$ is increased by 1. Therefore, if $|T_-| > \frac{\log L}{\eta} + 1$, then $d_t^{(j^\star)} > \frac{\log L}{\eta} + 1$ for some round t. However, by the algorithm's definition, this means that there was some round $t > t' \in T_0$ for which $d_{t'}^{(j^\star)} > \log L/\eta$ and $\hat{y}_{t'}^{(j^\star)} = 1$. This is not possible, since it would have imply $2^{\eta d_{t'}^{(j^\star)}} > L$ while $j^\star$ predicts 1, which implies $\hat{y}_{t'} = 1$.

Let's now bound M_+. Denote $m = M_+$. Denote $T_+ = \{t_1, \dots, t_m\}$. For every round $t_i \in T_+$, the set A_i of experts predicting 1 must have sum of weights at least L. Therefore, the set A_i will be removed from the version space at the end of round t_i, which implies that all A_i are disjoint. Therefore, we have:

$$\sum_{i\in[m]} \sum_{j\in A_i} 2^{\eta D_j} \geq mL.$$

On the other hand, the same quantity is upper bounded by:

$$\sum_{i\in[m]} \sum_{j\in A_i} 2^{\eta D_j} = \sum_{d\in\mathbb{N}} \sum_{i\in[m]} \sum_{A_i\cap\{j:D_j\in I_d\}} 2^{\eta D_j} \leq \sum_{d\in\mathbb{N}} n_d 2^{d+1} \leq 6\eta T L,$$

where the last inequality is due to Lemma 8. Combining both equations above gives $m \leq 6\eta T$ as required. ■

It is now straightforward to infer Theorem 6.

Proof [Proof of Theorem 6] We run RealizableExpAT with learning parameters $L = n$ and $\eta = \sqrt{\frac{\log n}{T}}$. The parameters are valid since $L > 1$ and $\eta \in (0, 1)$. Further, we have $\eta L T = \sqrt{\frac{\log n}{T}} n T \geq n$, since $\sqrt{\frac{\log n}{T}} T = \sqrt{T \log n} \geq 1$. Therefore Lemma 7 holds and we have

$$\mathtt{M}(\mathsf{RealizableExpAT}, S) = M_- + M_+ \leq \frac{\log L}{\eta} + 1 + 6\eta T = \sqrt{T\log n} + 1 + 6\sqrt{T\log n}.$$

Furthermore, it is clear that RealizableExpAT is computationally efficient by its definition. ■

4. Future work

Explicit construction for Lemma 26. In Lemma 26, we give a non-constructive proof of existence of the class used for the lower bound of Theorem 25. It will be interesting to construct this class explicitly. Note that the lower bound for prediction with expert advice (Lemma 12) is explicit and uses a similar technique. However, in this bound we do not require that the Littlestone dimension of the class will be only $O(\log_T n)$, so it is simpler.

Relation to standard agnostic online learning. The techniques used to prove the upper bound for prediction with expert advice in the realizable case are similar to techniques used for proving upper bounds on the regret in the problem of agnostic prediction with expert advice under full-information feedback. Remarkably, the optimal bounds are also identical, up to constant factors. Are the problems related? For example, is there a reduction between them, at least in one direction?

Complete landscape of mistake bounds. Given a fixed hard class $\mathcal{H}$, the bounds established in Theorem 24 when $k = 0$ are tight up to a $\sqrt{\log T}$ factor. Furthermore, in Section C.4 we show that both sides of Inequality (6) are attained by some hard classes. Is it true that every mistake bound between the lower and upper bound of (6) can be attained? For example, is there a hard class $\mathcal{H}$ that has mistake bound $\Theta_{\mathcal{H}}\left(\sqrt{T\log\log T}\right)$ in the realizable setting?

Prediction with slightly more than $\sqrt{T}$ many experts. We prove an upper bound of $O\big(\sqrt{T \log n}\big)$ for prediction with expert advice in the realizable setting. We also prove a matching lower bound of $\Omega\big(\sqrt{T \log n}\big)$, but it holds only for $n \geq T^{1/2+\epsilon}$, where $\epsilon > 0$ is a universal constant. On the other hand, it is not hard to see that if $n \leq \sqrt{T}$ then the optimal mistake bound is $\Theta(n)$. The regime where $n = T^{1/2+\epsilon}$ and $\epsilon = o(1)$ thus remains open.

Acknowledgments

We would like to thank Vinod Raman and Unique Subedi for introducing the problem of deterministic apple tasting to us.

References

Jacob Abernethy, John Langford, and Manfred K. Warmuth. Continuous experts and the binning algorithm. In *International Conference on Computational Learning Theory*, pages 544–558. Springer, 2006.

Noga Alon, Nicolo Cesa-Bianchi, Ofer Dekel, and Tomer Koren. Online learning with feedback graphs: Beyond bandits. In *Conference on Learning Theory*, pages 23–35. PMLR, 2015.

Peter Auer and Philip M. Long. Structural results about on-line learning models with and without queries. *Machine Learning*, 36:147–181, 1999.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multi-armed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

Gábor Bartók, Dean P Foster, Dávid Pál, Alexander Rakhlin, and Csaba Szepesvári. Partial monitoring—classification, regret bounds, and algorithms. *Mathematics of Operations Research*, 39(4): 967–997, 2014.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *COLT*, 2009.

Simina Brânzei and Yuval Peres. Online learning with an almost perfect expert. *Proceedings of the National Academy of Sciences*, 116(13):5949–5954, 2019.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Nicolo Cesa-Bianchi, Yoav Freund, David P. Helmbold, and Manfred K. Warmuth. On-line prediction and conversion strategies. *Machine Learning*, 25(1):71–110, 1996.

Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P. Helmbold, Robert E. Schapire, and Manfred K. Warmuth. How to use expert advice. *Journal of the ACM (JACM)*, 44(3):427–485, 1997.

Thomas M. Cover. Behavior of sequential predictors of binary sequences. In *Proceedings of the 4th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 263–272, Prague, 1967. Publishing House of the Czechoslovak Academy of Sciences.

Amit Daniely and Tom Helbertal. The price of bandit information in multiclass online classification. In *Conference on Learning Theory*, pages 93–104. PMLR, 2013.

Amit Daniely, Sivan Sabato, Shai Ben-David, and Shai Shalev-Shwartz. Multiclass learnability and the ERM principle. *J. Mach. Learn. Res.*, 16(1):2377–2404, 2015.

Yuval Filmus, Steve Hanneke, Idan Mehalel, and Shay Moran. Optimal prediction using expert advice and randomized littlestone dimension. In *COLT*, volume 195 of *Proceedings of Machine Learning Research*, pages 773–836. PMLR, 2023.

Yuval Filmus, Steve Hanneke, Idan Mehalel, and Shay Moran. Bandit-feedback online multiclass classification: Variants and tradeoffs. *arXiv preprint arXiv:2402.07453*, 2024.

David P Helmbold, Nicholas Littlestone, and Philip M Long. Apple tasting. *Information and Computation*, 161(2):85–139, 2000.

Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine learning*, 2(4):285–318, 1988.

Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

Philip M. Long. New bounds on the price of bandit feedback for mistake-bounded online multiclass learning. *Theoretical Computer Science*, 808:159–163, 2020.

Shie Mannor and Ohad Shamir. From bandits to experts: On the value of side-observations. *Advances in Neural Information Processing Systems*, 24, 2011.

Vinod Raman, Unique Subedi, Ananth Raman, and Ambuj Tewari. Apple tasting: Combinatorial dimensions and minimax rates. In *COLT*, volume 247 of *Proceedings of Machine Learning Research*, pages 4358–4380. PMLR, 2024.

Volodimir G. Vovk. Aggregating strategies. *Proc. of Computational Learning Theory, 1990*, 1990.

Appendix Table of Contents

Appendix A. Definitions

A.1. Realizable learning definitions

Let $\mathcal{X}$ be a (possibly infinite) *domain*. A pair $(x, y) \in \mathcal{X} \times \{0, 1\}$ is called an *example*, and an element $x \in \mathcal{X}$ is called an *instance*. A function $h \colon \mathcal{X} \to \{0, 1\}$ is called a *hypothesis* or a *concept*. A *hypothesis class*, or *concept class*, is a non-empty set $\mathcal{H} \subset \{0, 1\}^{\mathcal{X}}$. A sequence of examples $S = \{(x_i, y_i)\}_{t=1}^{T}$ is said to be *realizable* by $\mathcal{H}$ if there exists $h \in \mathcal{H}$ such that $h(x_t) = y_t$ for all $1 \leq i \leq T$. We say that such h is *consistent* with S, or *realizes* it. For simplicity of results statements, we will always assume that $n, T \geq 2$. The notation $\mathbb{1}[\cdot]$ denotes an indicator function. Online learning with apple tasting feedback Helmbold et al. (2000) is a repeated game between a learner and an adversary. Each round t, in the game proceeds as follows:

(i) The adversary picks an example $(x_t, y_t) \in \mathcal{X} \times \{0, 1\}$, and reveals only x_t to the learner.

(ii) The learner predicts a value $\hat{y}_t \in \{0, 1\}$ and suffers the *loss* $\mathbb{1}[\hat{y}_t \neq y_t]$.

(iii) If $\hat{y}_t = 1$, the adversary reveals y_t to the learner.

In this work, unless stated otherwise, we study the case where the predictions of the learner are deterministic. We model apple tasting learners as functions $\mathsf{Lrn} \colon (\mathcal{X} \times \{0, 1, \star\})^* \times \mathcal{X} \to \{0, 1\}$. The input of the learner has two parts: a *feedback sequence* $F \in (\mathcal{X} \times \{0, 1, \star\})^*$, and the current instance $x \in \mathcal{X}$. The feedback sequence is naturally constructed throughout the game: if in round t the prediction is 0, then the learner appends $(x_t, \star)$ to the feedback sequence, to indicate that no feedback was given for x_t. If, on the other hand, the prediction is 1, then the learner appends (x_t, y_t) to the feedback sequence. In round t, the prediction of the learner is $\mathsf{Lrn}(F, x_t)$, where F is the feedback sequence gathered by the learner in rounds $1, \ldots, t-1$.

Given a learning rule Lrn and an input sequence of examples $S = (x_1, y_1), \ldots, (x_T, y_T)$ such that (x_t, y_t) is the example picked by the adversary in round t, we denote the number of mistakes that Lrn makes on S by

$$\mathtt{M}(\mathsf{Lrn}; S) = \sum_{i=1}^{T} \mathbb{1}[\hat{y}_t \neq y_t].$$

It is worth noting that fixing S beforehand is usually linked with an *oblivious* adversary setting, in which the adversary cannot pick the examples on the fly. However, when the learner is deterministic, the adversary can simulate the entire game on its own, since we assume that the learning algorithm is known to all. Thus, oblivious and adaptive adversaries are in fact equivalent, ans we will refer to the adversary as being either adaptive or oblivious, depending on whichever is more convenient in the given context.

Let $\mathbb{R}^+$ be the set of non-negative real numbers. An hypothesis class $\mathcal{H}$ is *learnable with apple tasting feedback*, if there exists a function $M_{\mathcal{H}} \colon \mathbb{R}^+ \to \mathbb{R}^+$ satisfying $\lim_{T\to\infty} \frac{M_{\mathcal{H}}(T)}{T} = 0$, and a learning rule Lrn, such that for any T and for any input sequence S of length T which is realizable by $\mathcal{H}$, it holds that $\mathtt{M}(\mathsf{Lrn}; S) \leq M_{\mathcal{H}}(T)$.

We define the *optimal* mistake bound of $\mathcal{H}$ with horizon T to be

$$\mathtt{M}^\star(\mathcal{H}, T) = \inf_{\mathsf{Lrn}} \sup_{S} \mathtt{M}(\mathsf{Lrn}; S), \tag{3}$$

where the infimum is taken over all deterministic learning rules, and the supremum is taken over all realizable input sequences S of length T. It is convenient to assume that T is given to the learner. However, in Section D we show how to remove this assumption using standard doubling tricks.

A.2. Agnostic learning definitions

We also study the *agnostic* setting, in which even the best hypothesis might be inconsistent with the input sequence. Towards this end, we use the k*-realizable* framework, in which it is assumed that an upper bound k on the number of mistakes made by the best hypothesis is given to the learner. This assumption is convenient but not necessary, and we show in Section D how to remove it. Formally, an input sequence $S = (x_1, y_1), \ldots, (x_T, y_t)$ is k-realizable by a class $\mathcal{H}$ if there exists $h \in \mathcal{H}$ such that $h(x_t) \neq y_t$ for at most k indices $t \in [T]$. We say that such a hypothesis is k*-consistent* with S, or k*-realizes* it. For a given k, we denote the corresponding optimal mistake bound of $\mathcal{H}$ with horizon T by

$$\mathtt{M}^\star(\mathcal{H}, T, k) = \inf_{\mathsf{Lrn}} \sup_{S} \mathtt{M}(\mathsf{Lrn}; S), \tag{4}$$

where the infimum is taken over all deterministic learning rules, and the supremum is taken over all k-realizable input sequences S of length T. Note that the realizable setting is a special case of the k-realizable setting, attained when $k = 0$. We define and sometimes analyze the realizable case separately for didactic reasons: it is often significantly easier to handle.

A.3. Decision trees and the Littlestone dimension.

In this paper, a *tree* $\mathbf{T}$ refers to a finite full rooted ordered binary tree (that is, a rooted binary tree where each node which is not a leaf has a left child and a right child), equipped with the following information:

1. Each internal node v is associated with an instance $x \in \mathcal{X}$.
2. For every internal node v, the left outgoing edge is associated with the label 0, and the right outgoing edge is associated with the label 1.

A *prefix* of the tree T is any path that starts at the root. In this paper, a path is defined by a sequence of consecutive vertices. If a path is not empty, we may refer it by the sequence of consecutive edges corresponding with the sequence of consecutive vertices defining it. A prefix $v_0, v_1, \ldots, v_t$ defines a sequence of examples $(x_1, y_1), \ldots, (x_t, y_t)$ in a natural way: for every $i \in [t]$, x_i is the instance corresponding to the node v_{i-1}, and y_i is the label corresponding to the edge $v_{i-1} \to v_i$. A prefix is called *maximal* if it is maximal with respect to containment, that is, there is no prefix in the tree that strictly contains it. This is equivalent to requiring that v_t be a leaf. A maximal prefix is called a *branch*. The length of a prefix is the number of edges in it (so, the length is equal to the size of the corresponding sequence of examples). The length of a maximal branch in a tree is referred to as the tree's *depth*. We sometimes also refer to the length of a prefix as its depth.

A prefix in the tree is said to be k*-realizable* by $\mathcal{H}$ if the corresponding sequence of examples is k-realizable by $\mathcal{H}$. A tree $\mathbf{T}$ is k*-shattered* by $\mathcal{H}$ if all branches in $\mathbf{T}$ are k-realizable by $\mathcal{H}$. For every branch, it is convenient to relate to some hypothesis k-realizing it as the hypothesis that labels the leaf at the end of this branch. The *Littlestone dimension* of an hypothesis class $\mathcal{H}$, denoted by $\mathtt{L}(\mathcal{H})$, is the maximal depth of a *perfect* tree (that is, a tree in which all branches have the same depth) shattered by $\mathcal{H}$. If $\mathcal{H}$ shatters trees of arbitrarily large depth, then $\mathtt{L}(\mathcal{H}) = \infty$.

Appendix B. Prediction with expert advice

In this section we study the optimal mistake bound of a specific class called the *universal class* (of size n). For every $n \geq 1$, the universal class $\mathcal{U}_n = \{h_1, \ldots, h_n\}$ is defined as follows. Let $\mathcal{X}_n = \{0,1\}^n$ be the instance domain. For every $i \in [n]$, the hypothesis h_i is defined to be $h_i(x) = x_i$ for every $x \in \mathcal{X}_n$, where x_i denotes the i'th entry of x.

Note that every partition of $\mathcal{U}_n$ to hypotheses predicting 0 and hypotheses predicting 1 has an appropriate $x \in \mathcal{X}_n$ inducing it. Therefore, learning $\mathcal{U}_n$ is equivalent to the game of *prediction with expert advice with apple tasting feedback*, when the experts provide deterministic predictions. Formally, each round t of the game proceeds as follows:

(i) The n experts present predictions $\hat{y}_t^{(1)}, \ldots, \hat{y}_t^{(n)} \in \{0,1\}$.

(ii) The learner predicts a value $\hat{y}_t \in \{0,1\}$ and suffers the *loss* $\mathbb{1}[\hat{y}_t \neq y_t]$.

(iii) If $\hat{y}_t = 1$, the adversary reveals y_t to the learner.

In the realizable setting, there is an expert who never errs, and in the k-realizable setting there is an expert who errs for at most $k \in \mathbb{N}$ many times. Denote

$$\mathtt{M}^\star(n, T) = \mathtt{M}^\star(\mathcal{U}_n, T), \quad \text{and} \quad \mathtt{M}^\star(n, T, k) = \mathtt{M}^\star(\mathcal{U}_n, T, k).$$

For brevity, we will usually refer to the experts simply by their indices $[n]$. The main goal of this section is to prove an upper bound on $\mathtt{M}^\star(n, T, k)$ for all non-trivial n, T, k.

B.1. Prediction with expert advice: agnostic case

In this section we extend Theorem 6 to the k-realizable setting, in which an upper bound $k \in \mathbb{N}$ on the number of mistakes of the best expert is given. This assumption is not necessary, and in Section D we show how to remove it using a variation of the standard "doubling trick" of, e.g., Cesa-Bianchi et al. (1997). This section follows lines similar to Section 3 that handles the realizable case. We prove the following upper bound, attained by the ExpAT learner presented in Figure 2.

Theorem 9 *For every number of experts $n \geq 2$, horizon $T \geq 2$, realizability parameter k, and a k-realizable input sequence S of experts' predictions and true labels of length T:*

$$\mathtt{M}(\mathsf{ExpAT}, S) = O\left(\sqrt{T(k + \log n)}\right).$$

Furthermore, ExpAT *is computationally efficient.*

In section B.2 we show that the upper bound in Theorem 9 is tight for $n \geq T$. ExpAT is almost identical to RealizableExpAT from Section 3. The difference is that in RealizableExpAT the weight of an expert j in round t is defined to be $2^{\eta d_t^{(j)}}$, and in ExpAT it is defined to be $2^{k_t^{(j)} + \eta d_t^{(j)}}$, where $k_t^{(j)}$ is the number of future *allowed false positives* for expert j in round t. The number of future allowed false positives of expert j is the number of false positives that j can make in the future and still being kept in the version space. Namely, it is k minus the number of false positives that j have already made. As usual, a false positive mistake of an expert j refers to the scenario $\hat{y}_t^{(j)} = 1$ and $y_t = 0$. Other than that, all notation appearing in ExpAT has the same meaning as in RealizableExpAT.

ExpAT

Input: Set of experts indexed by $[n]$, realizability parameter $k \in \mathbb{N}$, learning parameters $\eta \in (0,1), L > 1$.
Initialize: Let $V_1 = [n]$, let $d_1^{(j)} = 0$ and $k_1^{(j)} = k$ for all $j \in [n]$.

for $t = 1, \ldots, T$**:**

1. Receive experts predictions $\hat{y}_t^{(1)}, \ldots, \hat{y}_t^{(n)}$.
2. Predict $\hat{y}_t = 1$ if and only if
$$\sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{k_t^{(j)} + \eta d_t^{(j)}} \geq L.$$
3. If $\hat{y}_t = 1$:
   (a) Set $d_{t+1}^{(j)} = d_t^{(j)}$ for every expert $j \in V_t$.
   (b) If $y_t = 0$:
      i. Set $V_{t+1} = V_t \backslash \{j \in V_t : \hat{y}_t^{(j)} = 1, k_t^{(j)} = 0\}$.
      ii. Set $k_{t+1}^{(j)} = k_t^{(j)} - \mathbb{1}[\hat{y}_t^{(j)} = 1]$ for every expert $j \in V_{t+1}$.
4. If $\hat{y}_t = 0$: set $d_{t+1}^{(j)} = d_t^{(j)} + \mathbb{1}[\hat{y}_t^{(j)} = 1]$ and $k_{t+1}^{(j)} = k_t^{(j)}$ for every expert $j \in V_t$.

Figure 2: A deterministic apple tasting learner for prediction with expert advice.

We will now analyze the algorithm's mistake bound. Towards this end, we introduce some additional notation. Fix T, k, and a k-realizable sequence S of experts' predictions and true labels of length T that ExpAT is executed on. It is convinient to assume without loss of generality, that all experts are eventually removed from the version space. To make sure that this assumption does not weaken the adversary, we allow it to continue after the T legitimate rounds, only if it causes the learner to make a false positive in each one of the rounds to follow the first T rounds. For every expert j and $\ell \in \{0, \ldots, k\}$, let $D_j^{(\ell)}$ be the distance of expert j when it makes its $(\ell+1)$'th false positive. Let $N_d^{(\ell)} = \{j \in [n] : D_j^{(\ell)} \in I_d\}$. Namely, $N_d^{(\ell)}$ is the set of experts having distance in $I_d = [d/\eta, (d+1)/\eta]$ when they make their $(\ell+1)$'th false positive. Let $n_d^{(\ell)} = |N_d^{(\ell)}|$.

Lemma 10 *For any learning parameters $\eta \in (0,1), L > 1$ satisfying $n \leq \eta L T / 2^{k+1}$, we have:*

$$M_- \leq \frac{\log L}{\eta} + k + 1, \quad \textit{and} \quad M_+ \leq 200 \eta T.$$

As in the realizable case, the challenging part in proving Lemma 10, is to upper bound M_+, and the key ingredient in this bound is an inequality proved in Lemma 11 by forming lower and upper bounds on the *total weighted distance gained* by experts in the version space. Formally, we will lower and upper bound the quantity

$$\mathtt{TW^2DG} := \sum_{t \in T_0} \sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{k_t^{(j)} + \eta d_t^{(j)}}.$$

We replace the notation from $\mathtt{TWDG}$ in the realizable case to $\mathtt{TW^2DG}$ in the agnostic case, because in the agnostic case the weight of an expert j is also controlled by its number of allowed false positives $k_t^{(j)}$ in addition to its distance $d_t^{(j)}$. The following decomposition of $\mathtt{TW^2DG}$ will be helpful. For every $\ell \in \{0, \dots, k\}$, let

$$\mathtt{TW^2DG}^{(\ell)} = \sum_{t \in T_0} \sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1 \wedge k_t^{(j)} = k - \ell] 2^{k_t^{(j)} + \eta d_t^{(j)}}.$$

That is, $\mathtt{TW^2DG}^{(\ell)}$ is the total weighted distance gained by the experts in the portion of the game in which they have made exactly ℓ false positives.

Lemma 11 *Suppose that* $\eta \in (0, 1), L > 1$ *and* $n \leq \eta LT/2^{k+1}$. *Then:*

$$\sum_{\ell=0}^{k} \sum_{d \in \mathbb{N}} n_d^{(\ell)} 2^{k-\ell+d} \leq 200 \eta LT.$$

Proof The main part of the proof is to lower and upper bound $\mathtt{TW^2DG}$. The upper bound is just as in the realizable case: by definition of the algorithm, for every $t \in T_0$ it holds that

$$\sum_{j \in V_t} \mathbb{1}[\hat{y}_t^{(j)} = 1] 2^{k_t^{(j)} + \eta d_t^{(j)}} < L,$$

which implies $\mathtt{TW^2DG} \leq LT_0 \leq LT$. Let's prove the lower bound, which is more challenging. Fix $\ell \geq 0$. Decompose $[n]$ into two sets $G^{(\ell)}, B^{(\ell)}$, where $G^{(\ell)}$ is the set of all experts that gained at least $\frac{1}{2\eta}$ distance while having made exactly ℓ false positives. Therefore:

$$\begin{aligned}
\mathtt{TW^2DG}^{(\ell)} &\geq \sum_{j \in G^{(\ell)}} \sum_{D \in \{D_j^{(\ell-1)}, \dots, D_j^{(\ell)}\}} 2^{k-\ell+D} \\
&\geq \sum_{j \in G^{(\ell)}} \frac{1}{2\eta} 2^{k-\ell+\eta\left(D_j^{(\ell)} - \frac{1}{2\eta}\right)} \\
&= 2^{k-\ell-3/2} \sum_{j \in G^{(\ell)}} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}}.
\end{aligned}$$

For the experts in $B^{(\ell)}$ we have:

$$\sum_{j \in B^{(\ell)}} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \leq \sum_{j \in B^{(\ell)}} \frac{1}{\eta} 2^{\eta\left(D_j^{(\ell-1)} + \frac{1}{2\eta}\right)} = \sqrt{2} \sum_{j \in B^{(\ell)}} \frac{1}{\eta} 2^{\eta D_j^{(\ell-1)}} \leq \sqrt{2} \sum_{j \in [n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell-1)}}.$$

Summing both inequalities above gives:

$$\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \leq \mathtt{TW}^2\mathtt{DG}^{(\ell)}/2^{k-\ell-3/2} + \sqrt{2} \sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell-1)}}. \tag{5}$$

Summing over ℓ and incorporating the number of future allowed false positives, we have:

$$\sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) = 2^k \sum_{j\in N_0^{(0)}} \frac{1}{\eta} 2^{\eta D_j^{(0)}} + 2^k \sum_{d\geq 1} \sum_{j\in N_d^{(0)}} \frac{1}{\eta} 2^{\eta D_j^{(0)}} + \sum_{\ell=1}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right).$$

The first two summands are a decomposition of the case $\ell = 0$ to experts with distance in I_0 and to experts with distance in I_d for $d > 0$. The third summand handles $\ell \geq 1$. We prove an upper bound for each summand. We can upper bound the first summand as

$$2^k \sum_{j\in N_0^{(0)}} \frac{1}{\eta} 2^{\eta D_j^{(0)}} \leq 2^{k+1} n \frac{1}{\eta} \leq LT,$$

by the assumption $n \leq \eta LT/2^{k+1}$. We can upper bound the second summand as follows:

$$2^k \sum_{d\geq 1} \sum_{j\in N_d^{(0)}} \frac{1}{\eta} 2^{\eta D_j^{(0)}} \leq 4 \cdot \sum_{d\geq 1} \sum_{j\in N_d^{(0)}} \frac{1}{\eta} 2^{k+d-1} \leq 4\mathtt{TW}^2\mathtt{DG}^{(0)} \leq 4\mathtt{TW}^2\mathtt{DG}.$$

By (5), we can bound the third summand by:

$$\begin{aligned}
\sum_{\ell=1}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) &\leq \sum_{\ell=0}^{k-1} 2^{k-(\ell+1)} \left(\mathtt{TW}^2\mathtt{DG}^{(\ell+1)}/2^{k-(\ell+1)-3/2} + \sqrt{2} \sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) \\
&\leq \sum_{\ell=0}^{k-1} 2^{3/2}\mathtt{TW}^2\mathtt{DG}^{(\ell+1)} + \sqrt{2} \sum_{\ell=0}^{k-1} 2^{k-(\ell+1)} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right). \\
&\leq 4\mathtt{TW}^2\mathtt{DG} + \frac{\sqrt{2}}{2} \sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right).
\end{aligned}$$

Summing the three bounds above together, we have:

$$\begin{aligned}
\sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) &\leq LT + 4\mathtt{TW}^2\mathtt{DG} + 4\mathtt{TW}^2\mathtt{DG} + \frac{\sqrt{2}}{2} \sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) \\
&\leq 9LT + \frac{\sqrt{2}}{2} \sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right).
\end{aligned}$$

Rearranging the inequality above gives:

$$\frac{9}{1-\frac{\sqrt{2}}{2}} LT \geq \sum_{\ell=0}^{k} 2^{k-\ell} \left(\sum_{j\in[n]} \frac{1}{\eta} 2^{\eta D_j^{(\ell)}} \right) \geq \sum_{\ell=0}^{k} \frac{1}{2} 2^{k-\ell} \left(\sum_{d\geq 1} n_d^{(\ell)} 2^d \right).$$

By assumption:

$$\sum_{\ell=0}^{k} 2^{k-\ell} n_0^{(\ell)} \leq 2^{k+1} n \leq LT.$$

Summing the two inequalities above gives the statement of the lemma. ■

We may now prove Lemma 10.

Proof [Proof of Lemma 10] Let's first bound M_-. Let $j^\star$ be an expert who makes at most k mistakes. In each of the rounds of T_- except for at most k, the distance of $j^\star$ is increased by 1. Therefore, if $|T_-| > \frac{\log L}{\eta} + k + 1$, then $d_t^{(j^\star)} > \frac{\log L}{\eta} + k + 1$ for some round t. However, by the algorithm's definition, this means that there was some round $t > t' \in T_0$ for which $d_{t'}^{(j^\star)} > \log L/\eta$ and $\hat{y}_{t'}^{(j^\star)} = 1$. This is not possible, since it would have imply $2^{\eta d_{t'}^{(j^\star)}} > L$ while $j^\star$ predicts 1, which implies $\hat{y}_{t'} = 1$.

Let's now bound M_+. Denote $m = M_+$ and $T_+ = \{t_1, \ldots, t_m\}$. For every round $t_i \in T_+$, the set A_i of experts predicting 1 must have sum of weights at least L. Therefore, we have:

$$\sum_{i \in [m]} \sum_{j \in A_i} 2^{k_{t_i}^{(j)} + \eta D_{t_i}^{(j)}} \geq mL.$$

On the other hand, the same quantity is upper bounded by:

$$\begin{aligned}
\sum_{i \in [m]} \sum_{j \in A_i} 2^{k_{t_i}^{(j)} + \eta D_{t_i}^{(j)}} &= \sum_{\ell=0}^{k} \sum_{j \in [n]} 2^{k-\ell+\eta D^{(\ell)} j} \\
&\leq \sum_{\ell=0}^{k} \sum_{d \in \mathbb{N}} n_d^{(\ell)} 2^{k-\ell+d+1} \\
&\leq 200 \eta LT. && \text{(Lemma 11)}
\end{aligned}$$

Combining both equations above gives $m \leq 200\eta T$ as required. ■

It is now straightforward to infer Theorem 9.

Proof [Proof of Theorem 9] We run ExpAT with learning parameters $L = n \cdot 2^{k+1}$ and $\eta = \sqrt{\frac{k+\log n}{T}}$. The parameters are valid since $L > 1$ and $\eta \in (0, 1)$. Further, we have $\eta LT/2^{k+1} = \sqrt{\frac{k+\log n}{T}} nT \geq n$, since $\sqrt{\frac{k+\log n}{T}} T = \sqrt{T(k + \log n)} \geq 1$. Therefore Lemma 10 holds and we have

$$\mathtt{M}(\mathsf{ExpAT}, S) = M_- + M_+ \leq \frac{\log L}{\eta} + k + 1 + 200\eta T = \sqrt{T(k + \log n)} + k + 1 + 200\sqrt{T(k + \log n)}.$$

Furthermore, it is clear that ExpAT is computationally efficient by its definition. ■

B.2. Prediction with expert advice: lower bounds

In this section we prove lower bounds for the problem of prediction with expert advice, which will affirm that our algorithms are optimal (up to a constant multiplicative factor) for $n \geq T$. We start with a bound that holds already in the realizable case.

Lemma 12 *For every* $2 \leq T \leq n \leq 2^T$*:*

$$\mathtt{M}^\star(n,T) = \Omega\Big(\sqrt{T \log n}\Big).$$

Proof If $\sqrt{2}^T \leq n \leq 2^T$ the lower bound is trivial. Suppose that $n < \sqrt{2}^T$. The adversary's strategy is to operate in phases, as explained below. Let n_i be the number of consistent experts in the beginning of phase i (initially we have $n_1 = n$). In phase i, as long as $n_i \geq \left\lceil \sqrt{T/\log n} \right\rceil$, the adversary splits the n_i experts into $\left\lceil \sqrt{T/\log n} \right\rceil$ equal as possible blocks, each containing at most a $\sqrt{\log n/T}$-fraction of the n_i experts (this is indeed a fraction since we assume $n < \sqrt{2}^T$). Denote the blocks by $b_1, \ldots, b_{\left\lceil \sqrt{T/\log n} \right\rceil}$. The phase continues as long as the learner predicts 0, and operates as follows. Starting with $j = 1$, in each round of the phase the experts in b_j predict 1, and the rest predict 0. At the end of the round, increase j by 1 if $j < \left\lceil \sqrt{T/\log n} \right\rceil$, and set $j = 1$ otherwise. When the learner predicts 1, the true label reported is 0, and the adversary moves on to the next phase. Let T_i be the number of rounds in phase i. Denote $T_i = t_i \left\lceil \sqrt{T/\log n} \right\rceil + r_i$ where $0 \leq r_i < \left\lceil \sqrt{T/\log n} \right\rceil$. In words, t_i counts the minimal number of times that some expert predicts 1 during phase i. If all experts but at most $\left\lceil \sqrt{T/\log n} \right\rceil$ are inconsistent before T rounds have passed, it must hold that:

$$e^{-2\sqrt{\log n/T}P} n \leq e^{-\frac{\sqrt{\log n/T}}{1-\sqrt{\log n/T}}P} n \leq (1 - \sqrt{\log n/T})^P n \leq 2\sqrt{T/\log n},$$

where the first and last inequalities are by the assumption $n < \sqrt{2}^T$, and the second inequality uses $1 + x \geq e^{\frac{x}{1+x}}$ for $x > -1$. After rearranging and using $n \geq T$, the inequality above implies $P > \sqrt{T/\log n}/8$. Therefore we are done, since the learner makes a mistake every time that a phase ends. So, we may now assume that all T rounds of the game are played. Thus:

$$T = \sum_{i=1}^{P} T_i = \sum_{i=1}^{P} \Big(t_i \left\lceil \sqrt{T/\log n} \right\rceil + r_i \Big).$$

If $P \geq \sqrt{T \log n}/4$ we are done, so assume that $P < \sqrt{T \log n}/4$. Therefore:

$$\sum_{i=1}^{P} r_i < \frac{\sqrt{T \log n}}{4} \cdot \left\lceil \sqrt{T/\log n} \right\rceil < T/2,$$

where the second inequality is due to $n < \sqrt{2}^T$. By the two equations above, we have:

$$\sum_{i=1}^{P} t_i \left\lceil \sqrt{T/\log n} \right\rceil \geq T/2.$$

Rearranging and using $n < 2^{\sqrt{T}}$ again gives:

$$\sum_{i=1}^{P} t_i > \frac{T/2}{\left\lceil \sqrt{T/\log n} \right\rceil /2} \geq \sqrt{T \log n}/4.$$

Every consistent expert predicted 1 when the learner predicted 0 for at least $\sum_{i=1}^{P} t_i$ many times. This finishes the proof by letting some consistent expert to determine the true labels. ■

We now prove a lower bound for the agnostic case.

Theorem 13 *For every $2 \leq T \leq n < \sqrt{2}^T$ and $k \geq 0$:*

$$\mathtt{M}^\star(n, T, k) = \Omega(\sqrt{T(k + \log n)}).$$

Proof The $\sqrt{T \log n}$ term is due to Lemma 12. It remains to prove a lower bound of $\Omega(\sqrt{kT})$ when assuming $k \geq 1$. It is convenient (but not necessary) to assume that $\sqrt{T/k} \in \mathbb{N}$. The adversary's strategy proceeds in $\sqrt{T/k}$ phases, each of length $\sqrt{Tk}$ rounds. In phase i, expert i predicts 1 and all other experts predict 0. The adversary reports that the true label is 0 every time the learner predicts 1. If in some phase i the learner predicted 1 for less than k many times, then we set the true label in all rounds of phase i in which the learner predicted 0 to be 1. Therefore expert i is k-consistent with the input, and the learner has made $\sqrt{Tk}$ many mistakes. Otherwise, the learner predicts 1 for at least k many times in all $\sqrt{T/k}$ phases, which results in $\sqrt{Tk}$ many mistakes. Since there are more than $\sqrt{T/k}$ experts, there is an expert who is consistent with letting all true labels to be 0. ■

Appendix C. Learning hypothesis classes

In this section we prove our results for learning hypothesis classes. Towards this end, we informally recall the definition of the *effective width* of a class $\mathcal{H}$, formally defined in Raman et al. (2024), and denoted by $\mathtt{w}(\mathcal{H})$. We assume that $\mathtt{L}(\mathcal{H}) \geq 1$, as otherwise $|\mathcal{H}| \leq 1$. For any natural w, d so that $d \geq w$, define a tree of width w and depth d as follows. Start from a perfect tree of depth d. Traverse every branch in the tree, while counting right edges. Once the w'th right edge of a branch is reached, remove the entire subtree beneath it, except for its root, which now becomes a leaf. For every $w \geq 1$, $D_w(\mathcal{H})$ is defined to be the maximal $d \geq w$ such that there exists a tree of width w and depth d which is shattered by $\mathcal{H}$. If there is no such tree even for $d = w$, then $D_w(\mathcal{H}) = 0$. If there are such trees with arbitrarily large depth then $D_w(\mathcal{H}) = \infty$. The effective width $\mathtt{w}(\mathcal{H})$ is defined to be the minimal w so that $D_w(\mathcal{H}) < \infty$. If there is no such minimal value then $\mathtt{w}(\mathcal{H}) = \infty$.

It was proved in Raman et al. (2024) that $\mathtt{w}(\mathcal{H})$ controls the randomized mistake bound of $\mathcal{H}$ under apple tasting feedback in the realizable setting. In this section we prove that it controls both its deterministic and randomized mistake bounds, even in the agnostic setting.

C.1. Learning hypothesis classes: Upper bounds

In this section, we prove the following upper bounds.

Theorem 14 *Let $\mathcal{H}$ be a class, k be a realizability parameter and T be the time horizon.*

1. *If* $\mathtt{w}(\mathcal{H}) = 1$*, then:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O(D_1(\mathcal{H})(k+1)).$$

2. *If* $1 < \mathtt{w}(\mathcal{H}) < \infty$*, then:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O\Big(\sqrt{T(k + \mathtt{L}(\mathcal{H}) \log T)}\Big).$$

Lemmas 15,20, to be proved in the following subsections, imply Theorem 14. In Section C.2 we prove that for every class, the first item is tight as long as T is sufficiently large, and the second item is tight up to a logarithmic factor of T and constants depending on the class.

C.1.1. EASY CLASSES

In this section we handle the case $\mathtt{w}(\mathcal{H}) = 1$. Towards this end, we provide an extension of the deterministic algorithm of Raman et al. (2024) for the case $\mathtt{w}(\mathcal{H}) = 1$ to the k-realizable setting.

Lemma 15 *Let* $\mathcal{H}$ *be a class with* $\mathtt{w}(\mathcal{H}) = 1$*. Then for all* k, T*:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O(D_1(\mathcal{H})(k+1)).$$

In Section C.2 we show that this mistake bound is tight for all easy classes, as long as T is sufficiently large. The idea of the upper bound is to use a variation of the budgeted concept classes technique of Filmus et al. (2023).

Definition 16 (false positive budgeted concept class) *A* false positive budgeted concept class *(or* budgeted class*, for short) is a collection* $\mathcal{B}$ *of pairs* (h, b) *where* $h \colon \mathcal{X} \to \{0, 1\}$ *is a hypothesis and* $b \in \mathbb{N}$ *is the allowed number of false positives. Furthermore, all hypotheses are distinct.*

An input sequence $S = (x_1, y_1), \ldots, (x_T, y_T)$ is realizable by a budgeted class $\mathcal{B}$ if there exists $(h, b) \in \mathcal{B}$ so that $h(x_t) = 1$ for for all $y_t = 1$, and $h(x_t) = 1$ for at most b many indices t with $y_t = 0$. A tree is shattered by $\mathcal{B}$ if every branch in it is realizable by $\mathcal{B}$. Given a concept class $\mathcal{H} \subset \{0, 1\}^{\mathcal{X}}$, $x \in \mathcal{X}$ and $y \in \{0, 1\}$, we defined $\mathcal{H}^{(x \to y)} = \{h \in \mathcal{H} : h(x) = y\}$. For budgeted classes, we define

$$\mathcal{B}^{(x \to 1)} = \{(h, b) \in \mathcal{B} : h(x) = 1\},$$
$$\mathcal{B}^{(x \to 0)} = \{(h, b) \in \mathcal{B} : h(x) = 0\} \cup \{(h, b-1) : (h, b) \in \mathcal{B}, b \geq 1, h(x) = 1\}.$$

In words, $\mathcal{B}^{(x \to 1)}$ removes from $\mathcal{B}$ all pairs with $h(x) = 0$. $\mathcal{B}^{(x \to 0)}$ decreases the budget of every hypothesis h with $h(x) = 1$, or removes it from $\mathcal{B}$ if it is out of budget. We now define an appropriate notion of k-shattering. A tree is k-shattered by a class $\mathcal{H}$ if for every branch, there exists a hypothesis that agrees with the branch on all of its right edges, and on all but at most k of its left edges. We can naturally simulate k-shattering using budgeted classes. For every hypothesis class $\mathcal{H}$ and $k \in \mathbb{N}$, define $\mathcal{B}_{\mathcal{H},k} = \{(h, k) : h \in \mathcal{H}\}$.

Observation 17 *Let* $\mathcal{H}$ *be a hypothesis class. A tree is* k*-shattered by* $\mathcal{H}$ *if and only if it is shattered by* $\mathcal{B}_{\mathcal{H},k}$*.*

NarrowConceptAT

Input: Class $\mathcal{H}$ with $\mathtt{w}(\mathcal{H}) = 1$, realizability parameter $k \in \mathbb{N}$.
Initialize: Let $V_1 = \mathcal{B}_{\mathcal{H},k}$.

for $t = 1, \ldots, T$**:**

1. Receive instance x_t.
2. If there exists $h \in V_t$ so that $h(x_t) = 1$: predict $\hat{y}_t = 1$. Else, predict $\hat{y}_t = 0$.
3. If $\hat{y}_t = 1$: Update $V_{t+1} = V_t^{(x_t \to y_t)}$

Figure 3: A deterministic apple tasting learner for concept classes with $\mathtt{w}(\mathcal{H}) = 1$.

For any class $\mathcal{H}$, define $D_w^{(k)}(\mathcal{H})$ to be the maximal d such that there exists a tree of width w and depth d that is k-shattered by $\mathcal{H}$, or, equivalently, shattered by $\mathcal{B}_{\mathcal{H},k}$. For budgeted classes in general, we use the standard notation $D_w(\mathcal{B})$ for the maximal d such that there exists a tree of width w and depth d that is shattered by $\mathcal{B}$. We may use the notation $h \in \mathcal{B}$ to indicate that $(h, b) \in \mathcal{B}$ for some $b \geq 0$. Our algorithm makes at most $D_1^{(k)}(\mathcal{H})$ false positives and at most k false negatives. It is presented in Figure 3.

Lemma 18 *For every class $\mathcal{H}$, $k \in \mathbb{N}$, and a k-realizable input sequence S:*

$$\mathtt{M}^\star(\mathsf{NarrowConceptAT}, S) \leq D_1^{(k)}(\mathcal{H}) + k$$

Proof We will show that NarrowConceptAT makes at most k false negatives and at most $D_1^{(k)}(\mathcal{H})$ false positives. Let us start with the false negatives. NarrowConceptAT makes a false negative exactly in rounds t where all hypotheses in V_t agree on the label 0, but the correct label is 1. For every round t, the hypotheses in V_t are exactly those who have made at most k false positives. By the k-realizability guarantee, at all times there must be some $h^\star \in V_t$ that made at most k many mistakes, and thus there could be at most k many rounds in which all hypotheses in V_t agree on the label 0, and yet the correct label is 1.

Let us now show that NarrowConceptAT makes at most $D_1^{(k)}(\mathcal{H})$ false positives. Let t be a round in which NarrowConceptAT makes a false positive. Suppose that $D_1(V_t) = D_1(V_{t+1})$. Then we can reach a contradiction by constructing a tree of depth $D_1(V_t) + 1$ and width 1 that is shattered by V_t, as follows. Let r be the root of this tree, labeled with x_t. By the algorithm's definition, there must be $h \in \mathcal{B}$ so that $h(x_t) = 1$, so we add a right edge emanating from r, and a leaf labeled by h beneath this edge. We also add a left edge emanating from r. Beneath this edge, we attach a tree of width 1 and depth $D_1(V_t) = D_1(V_{t+1})$ which is shattered by $V_{t+1} = V_t^{(x_t \to 0)}$, by assumption. We conclude that $D_1(V_t)$ drops by at least 1 after each false positive. Note that $D_1(V_t) = 0$ implies that for every x, all hypotheses in V_t agree on the same label. If that label is 1, then all hypotheses have made at least k many false positives, thus 1 must be the correct label. If this label is 0, a false positive cannot be made, by the algorithm's definition. Therefore, at most $D_1^{(k)}(\mathcal{H})$ false positives are made. ■

In order to prove Lemma 15, it remains to upper bound $D_1^{(k)}(\mathcal{H})$.

Lemma 19 *Let $\mathcal{H}$ be a class with $\mathtt{w}(\mathcal{H}) = 1$. We have*

$$D_1^{(k)}(\mathcal{H}) = O(D_1(\mathcal{H})(k+1)).$$

Proof If $k = 0$ then the lemma is trivial, so suppose that $k > 0$ and that $D_1^{(k)}(\mathcal{H}) \geq 4D_1(\mathcal{H}) \cdot k$. Let $\mathbf{T}$ be a tree witnessing $D_1^{(k)}(\mathcal{H})$. Let ℓ_1 be the lowest leaf that is right beneath a right edge in $\mathbf{T}$, and let h be the hypothesis labeling it. Let r_1 be the parent vertex of ℓ_1. For every instance x labeling an internal node v that is above r_1 in $\mathbf{T}$, and such that $h(x) = 1$, we conduct the following process. Remove from $\mathbf{T}$ the node v labeled by x, it's right outgoing edge and leaf beneath it, and it's left outgoing edge e. Let v' be the node beneath the left edge of v, and let e' be the ingoing edge of v, if exists. We naturally reconnect the tree by connecting e' as an ingoing edge to v'. Note that every hypothesis k-realizing a branch in $\mathbf{T}$ has at most k disagreements on left edges of the branch, and therefore we overall decreased the depth of $\mathbf{T}$ by at most k. Now, if exists, let r_2 be the node above r_1 (after all manipulations made), and let ℓ_2 be the right child of r_2. We now perform exactly the same process made before for ℓ_1, on ℓ_2. We repeat this process also for $r_3, r_4 \ldots, r_D$, as many times as possible, until we reach to r_D, which is the root of the tree (after the manipulations made). We argue that $D \geq 2D_1(\mathcal{H})$. Indeed, every treatment of r_i decreases the depth of $\mathbf{T}$ by at most k. By assumption, $D_1^{(k)}(\mathcal{H}) \geq 4D_1(\mathcal{H}) \cdot k$, and therefore after decreasing it by $2kD_1(\mathcal{H})$, there are still internal vertices above $r_{2D_1(\mathcal{H})}$ to handle.

When the process is finished, we have a tree of width 1 and depth $D \geq 2D_1(\mathcal{H})$ which is shattered by $\mathcal{H}$, and that is a contradiction. ■

Lemma 20 follows from plugging the bound of Lemma 19 to the bound of Lemma 18.

C.1.2. HARD CLASSES

Lemma 20 *Let $\mathcal{H}$ be a class with $1 < \mathtt{w}(\mathcal{H}) < \infty$. Then for all k, T:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O\Big(\sqrt{T(k + \mathtt{L}(\mathcal{H})\log T)}\Big).$$

The idea is to use the covering technique from Ben-David et al. (2009) with our bounds for *prediction with expert advice*. Formally, we say that an expert (or a learning algorithm) indexed by i covers a hypothesis h with respect to a sequence of instances $S = x_1, \ldots, x_T \in \mathcal{X}$ if $\hat{y}_t^{(i)} = h(x_t)$ for all $t \in [T]$. We say that a class $\mathcal{H}$ is covered by a set of experts with respect to S if every $h \in \mathcal{H}$ has a covering expert with respect to S within the set. The work of Ben-David et al. (2009) proves the following.

Theorem 21 (Ben-David et al. (2009)) *Let $\mathcal{H}$ be a class with $\mathtt{L}(\mathcal{H}) < \infty$, and let $T \in \mathbb{N}$. There exists a class of $n \leq T^{\mathtt{L}(\mathcal{H})}$ experts that covers $\mathcal{H}$ with respect to any sequence of instances of length at most T.*

We may now prove the upper bound.

Proof [Proof of Lemma 20] As argued in Raman et al. (2024), $\mathtt{w}(\mathcal{H}) < \infty$ implies $\mathtt{L}(\mathcal{H}) < \infty$. By Theorem 21, we can cover $\mathcal{H}$ with respect to any sequence of length at most T with at most $T^{\mathtt{L}(\mathcal{H})}$ experts. Since S is k-realizable by $\mathcal{H}$, the best expert makes at most k many mistakes on S. Thus, Theorem 9 implies that

$$\mathtt{M}^\star(\mathcal{H}, T, k) = O\left(\sqrt{T\left(k + \log T^{\mathtt{L}(\mathcal{H})}\right)}\right),$$

as required. ■

C.2. Learning hypothesis classes: universal lower bounds

In this brief section, we prove simple lower bounds that apply to all classes, and showing that our upper bounds are nearly tight. The unified adversarial strategy is given in the following lemma.

Lemma 22 *Let $\mathcal{H}$ be a class, and let $k \leq T$. Let $D \in \mathbb{N}$ so that $D \leq \min\{D_1(\mathcal{H}), \sqrt{T/(k+1)}\}$. Then*

$$\mathtt{M}^\star(\mathcal{H}, T, k) \geq D(k+1).$$

Proof Let $\mathbf{T}$ be a shattered tree of width 1 and depth D, which exists by assumption. The adversary's strategy is simple. Starting from the root, ask on any instance labeling the vertices of the branch consists of only left edges, for $D(k+1)$ many times each. Always respond with 0 as the true label. This amounts to a total of $D^2(k+1) \leq T$ many rounds by assumption, thus the strategy is implementable when there are T many rounds. If for some instance x the learner predicts 1 for less than $k+1$ many times, stop. We determine the true labels by the branch that ends at the right edge of the vertex labelled by x. Therefore, the adversary adheres to the k-realizability constraint and the learner has made at least $D(k+1)$ many mistakes. If, on the other hand, the learner predicts 1 for more than k many times on all instances, then we determine all true labels to be 0. This choice is realized by the branch consists of only left edges. In this case, the learner has made at least $k+1$ many mistakes on every instance, and overall at least $D(k+1)$ many mistakes. ■

The following theorem is an immediate corollary.

Theorem 23 *Let $\mathcal{H}$ be a class, and let $k \leq T$. Then:*

1. *If $\mathtt{w}(\mathcal{H}) = 1$ and $T \geq (D_1(\mathcal{H}))^2(k+1)$ then:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) \geq D_1(\mathcal{H})(k+1).$$

2. *If $\mathtt{w}(\mathcal{H}) > 1$ then:*

$$\mathtt{M}^\star(\mathcal{H}, T, k) = \Omega\left(\sqrt{T(k+1)}\right).$$

Proof The first item follows from applying Lemma 22 with $D = D_1(\mathcal{H})$. The second item follows from applying Lemma 22 with $D = \left\lfloor \sqrt{T/(k+1)} \right\rfloor$. ■

C.3. Trichotomy of mistake bounds

In this section we prove that up to logarithmic factors and constants depending on the class, there is a trichotomy of mistake bounds for learning concept classes. Towards this end, we define $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k)$ to be as $\mathtt{M}^\star(\mathcal{H}, T, k)$, only that we also allow randomized learners. Respectively, $\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k)$ measures the optimal *expected* number of mistakes, and not the number of mistakes, as $\mathtt{M}^\star(\mathcal{H}, T, k)$.

Theorem 24 *Fix a class $\mathcal{H}$. Let $k \in \mathbb{N}$, and let $T \geq k$ be the time horizon. Then:*

1. *If $\mathtt{w}(\mathcal{H}) = 1$:*
$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k), \mathtt{M}^\star(\mathcal{H}, T, k) = \Theta_{\mathcal{H}}(k).$$

2. *If $1 < \mathtt{w}(\mathcal{H}) < \infty$:*
$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k) = \Theta(k) + \tilde{\Theta}_{\mathcal{H}}\left(\sqrt{T}\right), \quad \mathtt{M}^\star(\mathcal{H}, T, k) = \Theta\left(\sqrt{Tk}\right) + \tilde{\Theta}_{\mathcal{H}}\left(\sqrt{T}\right).$$

3. *If $\mathtt{w}(\mathcal{H}) = \infty$:*
$$\mathtt{M}^\star_{\text{rand}}(\mathcal{H}, T, k) = T/2, \quad \mathtt{M}^\star(\mathcal{H}, T, k) = T.$$

The notation $\Theta_{\mathcal{H}}(\cdot)$ hides constants depending on the class $\mathcal{H}$.

Proof We prove every item separately, starting from the first item. The lower bound holds even for randomized learners with full information feedback. The deterministic upper bound is given in Lemma 15.

Let us prove the second item. The randomized upper bound is due to the agnostic regret bound given in Raman et al. (2024). In the randomized lower bound, the $\sqrt{T}$ term is due to the realizable lower bound of Raman et al. (2024), and the k term is a lower bound even for $\mathtt{w}(\mathcal{H}) = 1$. The deterministic upper bound is due to Lemma 20, and the deterministic lower bound is due to Theorem 23.

In the third item, the upper bounds are trivial. The lower bounds are since $\mathtt{w}(\mathcal{H}) = \infty$ implies $\mathtt{L}(\mathcal{H}) = \infty$, as noted in Raman et al. (2024). When $\mathtt{L}(\mathcal{H}) = \infty$, those lower bounds hold even with full-information feedback. ■

Note that Theorem 24 implies a large separation between deterministic and randomized learners for every hard hypothesis class. For example, for any hard class $\mathcal{H}$, if k grows as $\sqrt{T}$, then the randomized mistake bound grows at most as $\sqrt{T \log T}$ and the deterministic mistake bound grows as $T^{3/4}$.

C.4. Lower bounds for specific classes

When considering logarithmic factors of T and constants depending on the class, the deterministic bounds stated in Theorem 24 are not tight for hard classes already in the realizable case. In fact, the accurate best known bounds for any hard class $\mathcal{H}$ are:

$$\Omega\left(\sqrt{\mathtt{w}(\mathcal{H})T}\right) \leq \mathtt{M}^\star(\mathcal{H}, T) \leq O\left(\sqrt{\mathtt{L}(\mathcal{H})T \log T}\right), \tag{6}$$

where the lower bound is due to Raman et al. (2024) and the upper bound is by Theorem 14.

The goal in this section is to show that both inequalities in (6) can be attained as equalities for infinitely many $\mathtt{w}(\mathcal{H}), \mathtt{L}(\mathcal{H})$ values, and therefore the bounds of (6) are in fact the best possible bounds that hold for any hard class $\mathcal{H}$, even when the effective width/Littlestone dimension is very large. The work of Helmbold et al. (2000) describes a deterministic learner for the class of singletones over $\mathbb{N}$ that has a mistake bound of $O\left(\sqrt{T}\right)$. This is rather straightforward to extend this result to the class $\mathcal{H}_d$ of d-hamming balls over $\mathbb{N}$, for which the mistake bound will be $O\left(\sqrt{dT}\right)$. It is also not hard to see that $\mathtt{w}(\mathcal{H}_d) = d + 1 = \mathtt{L}(\mathcal{H}) + 1$, and thus for every value of $\mathtt{w}(\mathcal{H})$ the left inequality can be attained as equality. In this section, we will prove that the right inequality can also be attained for any $\mathtt{L}(\mathcal{H})$ larger than some constant.

Theorem 25 *For any natural d larger than some universal constant there exists a class $\mathcal{H}$ with $\mathtt{L}(\mathcal{H}) = O(d)$ and $T_0 = T_0(d)$, such that for every $T \geq T_0$:*

$$\mathtt{M}^\star(\mathcal{H}, T) = \Omega\left(\sqrt{dT \log T}\right).$$

To prove Theorem 25, we first assume that T is given and fixed, even before choosing the class $\mathcal{H}$. We will later show how to remove this assumption by a simple "gluing" technique that takes into account all values of T. The class $\mathcal{U}_n$ has $\mathtt{L}(\mathcal{U}_n) = \lfloor \log n \rfloor$. Therefore, using this class directly with the lower bound of Theorem 5 will only give a lower bound of $\Omega\left(\sqrt{\mathtt{L}(\mathcal{U}_n)T}\right)$. However, looking into the proof of the lower bound in Theorem 5 reveals that in contrast with full-information lower bounds, the "hard" experts' predictions used by the adversary are very unbalanced, in the sense that many experts predict 0 and only few predict 1. Since $\mathtt{L}(\mathcal{U}_n) = \Omega(\log n)$ comes from choosing experts' predictions which are as balanced as possible, this intuitively means that we can remove instances from $\mathcal{X}_n$ such that only unbalanced experts' predictions are available, in a way that decreases the Littlestone dimension from $\Omega(\log n)$ to $O(\log_T n)$, but maintains the $\Omega\left(\sqrt{T \log n}\right)$ lower bound from Theorem 5. Choosing $n = T^d$ results in a class $\mathcal{H}$ with $\mathtt{L}(\mathcal{H}) = O(d)$ that maintains a lower bound of $\Omega(T \log T^d)$.

We will now establish the existence of this class. Towards this end, we define a (p, T, n)-random class as follows. The domain is $\mathcal{X} = \{x_1, \ldots, x_T\}$. The class consists of n many hypotheses $\mathcal{H} = \{h_1, \ldots, h_n\}$. Every hypothesis h_i independently predicts 1 on every instance x_j with probability p.

Lemma 26 *For every natural d larger than some universal constant there exists $T_0(d)$ such that for every $T \geq T_0(d)$, there exists a hypothesis class $\mathcal{H}(d, T) = \{h_1, \ldots, h_n\}$ of size $n = T^d$ over the domain $\mathcal{X} = \{x_1, \ldots, x_T\}$ satisfying:*

1. *Every hypothesis in $\mathcal{H}$ predicts 1 on at least $\sqrt{dT \log T}$ instances.*

2. *For a subset $X \subset \mathcal{X}$, denote $H_{X \to 0} = \{h \in \mathcal{H} : \forall x \in X, h(x) = 0\}$. For every subset $X \subset \mathcal{X}$, if $|H_{X \to 0}| \geq T^{d/2}$ then for every $x \in \mathcal{X}$:*

$$\left| H_{X \to 0}^{(x \to 0)} \right| \geq \left(1 - 1000 \sqrt{\frac{d \log T}{T}} \right) |H_{X \to 0}|. \tag{7}$$

3. $\mathtt{L}(\mathcal{H}) < 10d$.

Proof We prove the existence of $\mathcal{H} := \mathcal{H}(d,T)$ by drawing a (p,T,n)-random class with $p = 100\sqrt{\frac{d\log T}{T}}$ and $n = T^d$ and prove that with positive probability, all items hold. Denote the event where Item i holds by E_i. Let us start with the first item. Fix a hypothesis h. By Chernoff's bound the probability that h predicts 1 for less than $\sqrt{dT\log T}$ many instances is at most

$$e^{-100\sqrt{dT\log T}/8} \leq e^{-10\sqrt{T}}.$$

A union bound now gives

$$\Pr\big[\bar{E}_1\big] \leq T^d/e^{10\sqrt{T}}. \tag{8}$$

Let us handle the second item. Fix $X \subset \mathcal{X}$ and $x \in \mathcal{X}$. If $x \in X$ then $H^{(x\to 0)}_{X\to 0} = H_{X\to 0}$, so in such a case (7) does not hold with probability 0. If $x \notin X$, then since the predictions of hypotheses for different instances are independent, the probability that (7) does not hold is at most

$$e^{-T^{d/2}100\sqrt{\frac{d\log T}{T}}} \leq e^{-T^{(d-1)/2}}$$

by Chernoff's bound. By a union bound:

$$\Pr\big[\bar{E}_2\big] \leq \frac{2^{T+\log T}}{e^{T^{(d-1)/2}}}. \tag{9}$$

Finally, we handle the third item. Let $\mathbf{T}$ be a perfect tree of depth $10d$, with vertices labelled by instances from $\mathcal{X}$. Fix a branch b in $\mathbf{T}$ with exactly $5d$ right edges. The probability that b is realized by a fixed $h \in \mathcal{H}$ is at most $\sqrt{\frac{d\log T}{T}}^{5d}$. Therefore the probability that b is realized by $\mathcal{H}$ is at most

$$1 - \left[1 - \sqrt{\frac{d\log T}{T}}^{5d}\right]^{T^d} \leq 2T^d\sqrt{\frac{d\log T}{T}}^{5d} = 2\frac{(d\log T)^{2.5d}}{T^{1.5d}} \leq 1/T^{1.4d}. \tag{10}$$

The first inequality is due to the Taylor series of $(1-\alpha)^\beta$ at $\alpha = 0$ implying $(1-\alpha)^\beta \leq 1 - \alpha\beta + O\big((\alpha\beta)^2\big)$ for $\alpha\beta < 1$. The second inequality holds for sufficiently large T_0 since $T \geq T_0$.

We now upper bound the probability that all b are realized by $\mathcal{H}$. Towards this end, we first show that the events where different branches in $\mathbf{T}$ are shattered by $\mathcal{H}$ are negatively correlated. Let $\mathcal{B} := \mathcal{B}(\mathbf{T})$ be the set of all branches in $\mathbf{T}$. For a class $\mathcal{H}'$, let $\mathcal{S}(\mathbf{T},\mathcal{H}')$ be the set of all branches in $\mathbf{T}$ which are realized by $\mathcal{H}'$. Since $\mathcal{H}$ is random and the predictions of its functions are chosen independently, the probability that a branch b is realized by $\mathcal{H}$ depends only on the size of $\mathcal{H}$. Therefore, there exists an increasing function $f\colon \mathbb{N} \to [0,1]$ such that $\Pr[b \in \mathcal{S}(\mathbf{T},\mathcal{H})] = f(|\mathcal{H}|)$. Since every two branches disagree on at least one instance in $\mathcal{X}$, it holds that if b is realized by some function $h \in \mathcal{H}$, then any other branch $b' \in \mathcal{B}$ cannot be realized by h. Thus, for any non-empty subset of branches $\mathcal{B}' \subset \mathcal{B}$ and a branch $b \notin B$, we have

$$\Pr\big[b \in \mathcal{S}(\mathbf{T},\mathcal{H}) | \mathcal{B}' \subset \mathcal{S}(\mathbf{T},\mathcal{H})\big] \leq f(|\mathcal{H}| - |\mathcal{B}'|).$$

Therefore, by the conditional probability formula, for any subset $\mathcal{B}' \subset \mathcal{B}$:

$$\Pr[\mathcal{B}' \subset \mathcal{S}(\mathbf{T},\mathcal{H})] \leq \prod_{b\in\mathcal{B}'} \Pr[b \in \mathcal{S}(\mathbf{T},\mathcal{H})]. \tag{11}$$

Let $\mathcal{B}'$ be the set of branches with precisely $5d$ right edges. By (10), (11) and $\binom{2n}{n} \geq 2^n/\sqrt{4n}$ for any n larger than some universal constant:

$$\Pr[\mathcal{B}' \subset \mathcal{S}(\mathbf{T}, \mathcal{H})] \leq \left(1/T^{1.4d}\right)^{\binom{10d}{5d}} \leq 1/T^{0.5\sqrt{d}\cdot 2^{10d}}.$$

On the other hand, there are at most $T^{2^{10d}}$ trees with vertices labeled by instances from $\mathcal{X}$ of depth $10d$. A union bound gives:

$$\Pr[\bar{E}_3] \leq \frac{T^{2^{10d}}}{T^{0.5\sqrt{d}\cdot 2^{10d}}} = 1/\sqrt{T}^{\sqrt{d}}. \tag{12}$$

Finally, we deduce:

$$\Pr[E_1 \cap E_2 \cap E_3] = 1 - \Pr[\bar{E}_1 \cup \bar{E}_2 \cup \bar{E}_3] \geq 1 - \sum_{i=1}^{3} \Pr\left[\bar{E}_i\right] > 0,$$

where the first inequality is by a union bound, and the second inequality is by summing 8, 9, 12 when d is larger than some universal constant and $T \geq T_0$ where T_0 is sufficiently large. ■

We may now prove Theorem 25 under the assumption that T is fixed.

Lemma 27 *Fix d larger than some universal constant, and $T \geq T_0$, where T_0 is as in Lemma 26. There exists a class $\mathcal{H}$ with $\mathtt{L}(\mathcal{H}) = O(d)$, such that:*

$$\mathtt{M}^\star(\mathcal{H}, T) = \Omega\left(\sqrt{dT\log T)}\right).$$

Proof We choose the class $\mathcal{H} := \mathcal{H}(d, T)$ guaranteed by Lemma 26. The third item of Lemma 26 states that $\mathtt{L}(\mathcal{H}) \leq 10d$. It remain to prove that $\mathtt{M}^\star(\mathcal{H}, T) = \Omega\left(\sqrt{dT\log T)}\right)$. The adversary's strategy is very simple. It maintains a version space V_t of hypotheses consistent with the feedback provided until (and include) round $t-1$. As long as $|V_t| \geq T^{d/2}$, for every round t where $\hat{y}_t = 1$, the true label is $y_t = 0$. Once $|V_t| < T^{d/2}$, the adversary chooses an arbitrary hypothesis $h^\star \in V_t$ and picks the true label to be $h^\star(x'_t)$ for every round t' larger than the first round t in which $|V_t| < T^{d/2}$. This strategy guarantees that in any case, there exists $h^\star \in \mathcal{H}$ which is consistent with the feedback.

It remains to analyze the number of mistakes made by the learner. We have two cases. If $|V_T| \geq T^{d/2}$, we choose an arbitrary $h^\star \in V_T$ and determine $y_t = h^\star(x_t)$ for every round t. By definition of V_T, this choice of true labels is consistent with the feedback. By the first item of Lemma 26, there are at least $\sqrt{dT\log T}$ many rounds t so that $h^\star(x_t) = 1$ and $\hat{y}_t = 0$, and therefore the learner makes at least $\sqrt{dT\log T}$ many mistakes. Otherwise, there exists some $t \in [T]$ such that $|V_t| \geq T^{d/2}$ and $|V_{t+1}| < T^{d/2}$. Denote the set of rounds in $\{1, \ldots, t\}$ where the learner predicted 1 by $M = \{i_1, \ldots, i_m\}$ (where $i_m = t$). For every $j \in [m]$, denote $X_j = \{x_{i_1}, \ldots, x_{i_j}\}$. Note that by the adversary's strategy, for every $i_j \in M$, we have $V_{i_j+1} = \mathcal{H}_{X_j \to 0}$. The second item of Lemma 26 thus implies

$$\left(1 - 1000\sqrt{\frac{d\log T}{T}}\right)^m T^d \leq T^{d/2}.$$

Rearranging and using $1-x > e^{\frac{-x}{1-x}}$ for $x < 1$ gives $e^{1000\sqrt{\frac{d \log T}{T}}m} \geq T^{d/2}$ since d, T are sufficiently large. After taking $\log$ of both sides, we obtain $1000\sqrt{\frac{d \log T}{T}}m \geq \frac{d}{2} \log T$, which finally gives:

$$m \geq \sqrt{dT \log T}/2000.$$

Since $m = |M|$ and the learner makes a mistake in every round from M, this finishes the proof. ■

The final step remained for proving Theorem 25 is to remove the assumption that we know T before choosing the class. This assumption can be removed by a rather simple technique, which we call *gluing* concept classes. To the best of our knowledge, it was not defined in existing literature.

We define how to glue together two concept classes, and the definition is easily extendable to any countable collection of concept classes, by gluing two classes from the collection, and then gluing the obtained class again with another class, and so on.

Definition 28 (Glued concept class) *Let $\mathcal{H}_1$, $\mathcal{H}_2$ be concept classes defined on finite, disjoint domains $\mathcal{X}_1, \mathcal{X}_2$. We call two classes satisfying these properties* glueable classes. *Define the glued domain of $\mathcal{X}_1, \mathcal{X}_2$ to be $G(\mathcal{X}_1, \mathcal{X}_2) = \mathcal{X}_1 \cup \mathcal{X}_2$. Define the glued concept class of $\mathcal{H}_1, \mathcal{H}_2$, denoted by $G(\mathcal{H}_1, \mathcal{H}_2)$ as follows. Let $r \in \{1, 2\}$. For every $h \in \mathcal{H}_r$, the hypothesis $h^{(r)}$ belongs to $G(\mathcal{H}_1, \mathcal{H}_2)$, where:*

$$h^{(r)}(x) = \begin{cases} h(x) & x \in \mathcal{X}_r, \\ 0 & x \notin \mathcal{X}_r. \end{cases}$$

We say that the source of $h^{(r)}$ is $\mathcal{H}_r$.

For any countable collection of hypotheses classes $\{\mathcal{H}_i\}_{i=1}^{\infty}$, if the domains of all classes are finite and pairwise disjoint, we say that the collection $\{\mathcal{H}_i\}_{i=1}^{\infty}$ is glueable. The glued class $\mathcal{H} = G(\{\mathcal{H}_i\}_{i=1}^{\infty})$ is obtained by gluing $\mathcal{H}_1$ and $\mathcal{H}_2$, and the gluing $G(\mathcal{H}_1, \mathcal{H}_2)$ with $\mathcal{H}_3$, and so on.

Proposition 29 *Let $\{\mathcal{H}_i\}_{i=1}^{\infty}$ be a countable collection of glueable classes, and let $\mathcal{H} = G(\{\mathcal{H}_i\}_{i=1}^{\infty})$. Then $\mathtt{L}(\mathcal{H}) \leq \max_i \mathtt{L}(\mathcal{H}_i) + 1$.*

Proof We design an explicit learner for learning $\mathcal{H}$ in the realizable, full-information feedback setting. The learner always predicts 0 until it makes a mistake. Once that happens, the learner knows the source $\mathcal{H}_i$ of the the target concept. Therefore it will make at most $\mathtt{L}(\mathcal{H}_i)$ more mistakes. ■

We can now fully prove Theorem 25.

Proof [Proof of Theorem 25] Consider the collection $\{\mathcal{H}(d, T)\}_{T \geq T_0}$, where T_0 is the same as in Lemma 27. For every T, we differentiate the instances of $\mathcal{H}(d, T)$ from the instances of other classes by relating to them as $x_1(T), \ldots, x_T(T)$, thus making the collection $\{\mathcal{H}(d, T)\}_{T \geq T_0}$ glueable. We can now define $\mathcal{H} = G(\{\mathcal{H}(d, T)\}_{T \geq T_0})$. By proposition 29 and Lemma 26, $\mathtt{L}(\mathcal{H}) \leq 10d$. Now, given T, we use the adversary's strategy defined in Lemma 27 for the class $\mathcal{H}(d, T)$, which is consistent with $\mathcal{H}$. Therefore $\mathtt{M}^{\star}(\mathcal{H}, T) = \Omega\Big(\sqrt{dT \log T)}\Big)$ as required. ■

DTExpAT

Input: Class of n experts indexed by $[n]$.
Initialize: Let $g_k = g_T = 1$, $T_0 = 0$.
Denote: $k = g_k \log n, T = g_T \log n$.

for $t = 1, \ldots, T^\star$**:**

1. If $t > T - T_0$, or, all experts have made more than k false positives:
   (a) If $t > T - T_0$: Double g_T.
   (b) If all experts have made more than k false positives: Double g_k.
   (c) Restart ExpAT: Set $T_0 = t$, zero the false positives count of all experts, remove all information gathered in previous rounds from its memory.
2. Predict as ExpAT predicts under the assumption that the best expert makes at most k many mistakes and that there are T many rounds, given all information gathered by ExpAT in previous rounds. Receive feedback when the prediction is 1.

Figure 4: A deterministic apple tasting learner for prediction with expert advice without prior knowledge.

Appendix D. Prediction without prior knowledge

In this section, we remove the assumption that the number of rounds T and the realizability parameter k are given to the learner. We denote the unknown number of rounds and realizability parameter by $T^\star, k^\star$. The idea is similar to standard doubling tricks, such as the doubling trick of Cesa-Bianchi et al. (1997).

D.1. Prediction with expert advice without prior knowledge

We present the doubling trick algorithm DTExpAT for prediction with expert advice in Figure 4.

Theorem 30 *Theorem 9 holds even without assuming that the number of rounds $T^\star$ and the realizability parameter $k^\star$ are given to the learner.*

Proof Let S be the input sequence. Let $i^\star$ be a target expert who makes at most $k^\star$ many mistakes. It is convenient to assume that $i^\star$ makes no false negative mistakes. If a mistake bound M holds under this assumption, then a mistake bound of $M + k^\star$ holds without this assumption: the adversary may choose at most $k^\star$ rounds in which DTExpAT and $i^\star$ predict 0, and let the true label in those rounds to be 1. Since the mistake bound proved in Theorem 9 (in which we "compete") is larger than $\sqrt{T^\star k^\star} \geq k^\star$, adding $k^\star$ to M does not change the bound in more than a constant factor.

We use the learner DTExpAT described in Figure 4. Let $g_T^\star, g_k^\star$ such that $T^\star = g_T^\star \log n$ and $k^\star = g_k^\star \log n$. By the assumptions in Thoeorem 9, $g_T^\star \geq 1$. We consider two cases. In the first case, $g_k^\star < 1$. This means that g_k is never doubled, and that the mistake bound given in Thoeorem 9 is $O\big(\sqrt{T \log n}\big)$, which is the same as $O\big(\sqrt{g_T^\star} \log n\big)$. Consider the following partition of $[T^\star]$ to intervals. Every interval in the partition ends at the round before the round where g_T is doubled, and then a new interval begins. Let $\mathcal{I}_T = \{I_0, \dots, I_{T'}\}$ be the set of intervals in the partition, Where I_i is the interval in which the value of g_T is 2^i. For every $I \in \mathcal{I}_T$, let $\mathtt{M}^\star(I)$ be the number of mistakes made by $\mathsf{DTExpAT}$ in all rounds of interval I. Thus:

$$\mathtt{M}^\star(\mathsf{DTExpAT}, S) = \sum_{I_i \in \mathcal{I}_T} \mathtt{M}^\star(I_i) \leq \sum_{I_i \in \mathcal{I}_T} \sqrt{2^i 2 \log^2 n} = \sqrt{2} \log n \sum_{I_i \in \mathcal{I}_T} \sqrt{2^i} \leq 10 \sqrt{g_T^\star} \log n.$$

The first inequality is by the mistake bound of Theorem 9. The second inequality is since $\sum_{i=0}^{T'} \sqrt{2^i} \leq 4\sqrt{2^{T'}}$ and because at all times, $g_T \leq 2 g_T^\star$.

We now handle the second case in which $g_k^\star \geq 1$. In this case, the mistake bound given in Theorem 9 is $O\Big(\sqrt{Tk}\Big)$, which is the same as $O\big(\sqrt{g_T^\star g_k^\star} \log n\big)$. In this case as well, we consider a partition of $[T^\star]$ to intervals. Every interval in the partition ends one round before at least one of g_k, g_T is doubled, and then a new interval begins. Let $\mathcal{I}_k = \{I_1, \dots, I_{k'}\}$ be the set of intervals initiated by doubling g_k, where I_i is the interval in which the value of g_k is 2^i. For every $I \in \mathcal{I}_k$, let $\mathtt{M}^\star(I)$ be the number of mistakes made by $\mathsf{DTExpAT}$ in all rounds of interval I. We have:

$$\sum_{I_i \in \mathcal{I}_k} \mathtt{M}^\star(I) \leq \sum_{I_i \in \mathcal{I}_k} \sqrt{2 g_T^\star \log n (2^i \log n + \log n)} \leq 2\sqrt{g_T^\star} \log n \sum_{I_i \in \mathcal{I}_k} \sqrt{2^i} \leq 20 \sqrt{g_T^\star g_k^\star} \log n.$$

The first inequality is since $g_T \leq 2 g_T^\star$ at all times, and by the guarantees of Theorem 9. The last inequality is since $\sum_{i=1}^{k'} \sqrt{2^i} \leq 4\sqrt{2^{k'}}$. We conduct roughly the same analysis for $\mathcal{I}_T = \{I_1, \dots, I_{T'}\}$ which is the set of intervals initiated by doubling g_T, Where I_i is the interval in which the value of g_T is 2^i. This gives:

$$\sum_{I_i \in \mathcal{I}_T} \mathtt{M}^\star(I) \leq \sum_{I_i \in \mathcal{I}_T} \sqrt{2^i \log n (2 g_k^\star \log n + \log n)} \leq 2\sqrt{g_k^\star} \log n \sum_{I_i \in \mathcal{I}_T} \sqrt{2^i} \leq 20 \sqrt{g_T^\star g_k^\star} \log n.$$

The first inequality is since $g_k \leq 2 g_k^\star$ at all times. For the first interval before either g_k or g_T is doubled, $O(\log n)$ mistakes are made. Summing the total number of mistakes made in all rounds, we get a mistake bound of $\mathtt{M}^\star(\mathsf{DTExpAT}, S) \leq O\big(\sqrt{g_T^\star g_k^\star} \log n\big)$, as required. ■

D.2. Learning hypothesis classes without prior knowledge

Theorem 31 *Theorem 14 holds even without assuming that the number of rounds $T^\star$ and the realizability parameter $k^\star$ are given to the learner.*

Proof We first prove the bound for hard classes, and then for easy classes.

Our bound for a hard class $\mathcal{H}$ is obtained via a reduction to prediction with expert advice, thus it is almost completely handled by Theorem 30. The only difference is that in the reduction, the number of experts depends on T. Thus, when $\mathsf{DTExpAT}$ doubles the guess of T to $2T$, it also needs to enlarge the number of experts from $T^{\mathtt{L}(\mathcal{H})}$ to $(2T)^{\mathtt{L}(\mathcal{H})}$. Since the final guess of T is at most $2T^\star$, we obtain the upper bound $O\Big(\sqrt{T^\star\big(k^\star + \log(2T^\star)^{\mathtt{L}(\mathcal{H})}\big)}\Big)$, which is $O\Big(\sqrt{T^\star(k^\star + \mathtt{L}(\mathcal{H})\log T^\star)}\Big)$, as required.

For easy classes, the mistake bound achieved by NarrowConceptAT does not depend on $T^\star$, but it does use knowledge of $k^\star$. To remove the assumption that $k^\star$ is given, it is convenient to assume, as we did for prediction with expert advice, that the best hypothesis makes no false negative mistakes. Removing this assumption will add a $k^\star$ term to the mistake bound, which does not increase it by more than a constant factor. Now, we initialize $k = 1$ and run NarrowConceptAT with k as the realizability parameter. If at some point the version space becomes empty, then the guess of k is too small, and we double it. The final guess of k will be at most $2k^\star$. Therefore the total number of mistakes will be at most

$$\sum_{i=1}^{\lceil \log 2k^\star \rceil} D_1^{(2^i)}(\mathcal{H}) \leq \sum_{i=1}^{\lceil \log 2k^\star \rceil} O(D_1(\mathcal{H}) + 1)2^i = O((D_1(\mathcal{H}) + 1)(k^\star + 1)),$$

where the inequality is by Lemma 19, and the last equality is since $\sum_{i=1}^{n} 2^i = O(2^n)$. ■