A New Perspective on Pool-Based Active Classification and False-Discovery Control

Anonymous Author(s)
Affiliation
Address
email

Abstract

In many scientific settings there is a need for adaptive experimental design to guide the process of identifying regions of the search space that contain as many true positives as possible subject to a low rate of false discoveries (i.e. false alarms). Such regions of the search space could differ drastically from a predicted set that minimizes 0/1 error and accurate identification could require very different sampling strategies. Like active learning for binary classification, this experimental design cannot be optimally chosen a priori, but rather the data must be taken sequentially and adaptively. However, unlike classification with 0/1 error, collecting data adaptively to find a set with high true positive rate and low false discovery rate (FDR) is not as well understood. In this paper we provide the first provably sample efficient adaptive algorithm for this problem. Along the way we highlight connections between classification, combinatorial bandits, and FDR control making contributions to each.

1 Introduction

As machine learning has become ubiquitous in the biological, chemical, and material sciences, it has become irresistible to use these techniques not only for making inferences about *previously* collected data, but also for guiding the data collection process, closing the loop on inference and data collection [9, 37, 39, 38, 32, 30]. However, though collecting data randomly or non-adaptively can be inefficient, ill-informed ways of collecting data adaptively can be catastrophic: a procedure could collect some data, adopt an incorrect belief, collect more data based on this belief, and leave the practitioner with insufficient data in the right places to infer anything with confidence.

In a recent high-throughput protein synthesis experiment [32], thousands of short amino acid sequences (length less than 60) were evaluated with the goal of identifying and characterizing a subset of the pool of all possible sequences ($\approx 10^{80}$) containing many sequences that will fold into stable proteins. That is, given an evaluation budget that is just a minuscule proportion of the total number of sequences, the researchers sought to make predictions about individual sequences that would never be evaluated. An initial first round of sequences uniformly sampled from a predefined subset were synthesized to observe whether each sequence was in the set of sequences that will fold, $\mathcal{H}_1$, or in $\mathcal{H}_0 = \mathcal{H}_1^c$. Treating this as a classification problem, a linear logistic regression classifier was trained, using these labels and physics based features. Then a set of sequences to test in the next round were chosen to maximize the probability of folding according to this empirical model - a procedure repeated twice more. This strategy suffers two flaws. First, selecting a set to maximize the likelihood of hits given past rounds' data is effectively using logistic regression to perform *optimization* similar to follow-the-leader strategies [13]. While more of the sequences *evaluated* may fold, these observations may provide little information about whether sequences that were *not* evaluated will fold or not. Second, while it is natural to employ logistic regression or the SVM

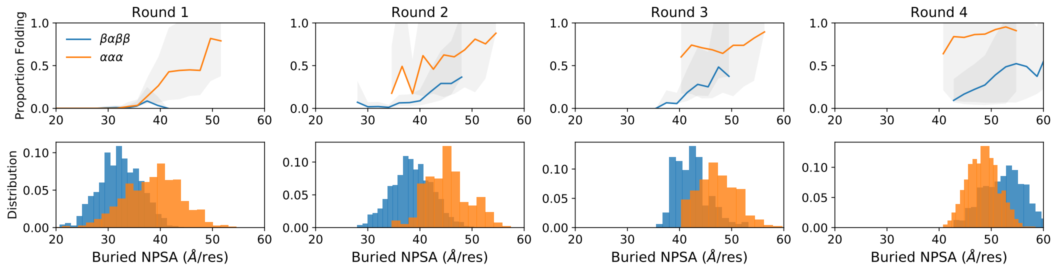

Figure 1: The distribution of a feature that is highly correlated with the fitted logistic model (bottom plot) and the proportion of sequences that fold (top plot). The distribution of this feature for the sequences drifts right.

to discriminate between binary outcomes (e.g., fold/not-fold), in many scientific applications the property of interest is incredibly rare and an optimal classifier will just predict a single class e.g. not fold. This is not only an undesirable inference for prediction, but a useless signal for collecting data to identify those regions with higher, but still unlikely, probabilities of folding. Consider the data of [32] reproduced in Figure 1, where the proportion of sequences that fold along with their distributions for a particularly informative feature are shown in each round for two different protein topologies. In the last column of Figure 1, even though most of the sequences evaluated are likely to fold, we are sampling in a small part of the overall search space. This limits our overall ability to identify under-explored regions that could potentially contain many sequences that fold, even though the logistic model does not achieve its maximum there. On the other hand, in the top plot of Figure 1, sequences with topology $\beta\alpha\beta\beta$ (shown in blue) so rarely folded that a near-optimal classifier would predict "not fold" for every sequence.

Instead of using a procedure that seeks to maximize the probability of folding or classifying sequences as fold or not-fold, a more natural objective is to predict a set of sequences π in such a way as to *maximize the true positive rate (TPR)* $|\mathcal{H}_1 \cap \pi|/|\mathcal{H}_1|$ while *minimizing the false discovery rate (FDR)* i.e. $|\mathcal{H}_0 \cap \pi|/|\pi|$. That is, π is chosen to contain a large number of sequences that fold while the proportion of false-alarms among those predicted is relatively small. For example, if a set π for $\beta\alpha\beta\beta$ was found that maximized TPR subject to FDR being less than $9/10$ then π would be non-empty with the guarantee that at least one in every 10 suggestions was a true-positive; not ideal, but making the best of a bad situation. In some settings, such as for topology $\alpha\alpha\alpha$ (shown in orange), training a classifier to minimize 0/1 loss may be reasonable. Of course, before seeing any data we would not know whether classification is a good objective so it is far more conservative to optimize for maximizing the number of discoveries.

Contributions. We propose the first provably sample-efficient adaptive sampling algorithm for maximizing TPR subject to an FDR constraint. This problem has deep connections to active binary classification (e.g., active learning) and pure-exploration for combinatorial bandits that are necessary steps towards motivating our algorithm. We make the following contributions:

1. We improve upon state of the art sample complexity for pool-based active classification in the agnostic setting providing novel sample complexity bounds that do not depend on the disagreement-coefficient for sampling with or without replacement. Our bounds are more granular than previous results as they describe the contribution of a single example to the overall sample complexity.
2. We highlight an important connection between active classification and combinatorial bandits. Our results follow directly from our improvements to the state of the art in combinatorial bandits, extending methods to be near-optimal for classes that go beyond matroids where one need not sample every arm at least once.
3. Our main contribution is the development and analysis of an adaptive sampling algorithm that minimizes the number of samples to identify the set that maximizes the true positive rate subject to a false discovery constraint. To the best of our knowledge, this is the first work to demonstrate a sample complexity for this problem that is provably better than non-adaptive sampling.

1.1 Pool Based Classification and FDR Control

Here we describe what is known as the pool-based setting for active learning with stochastic labels. Throughout the following we assume access to a finite set of items $[n] = \{1, \cdots, n\}$ with an associated label space $\{0, 1\}$. The items can be fixed vectors $\{x_i\}_{i=1}^n \in \mathbb{R}^d$ but we do not restrict to this case. Associated to each $i \in [n]$ there is a Bernoulli distribution $\text{Ber}(\eta_i)$ with $\eta_i \in [0,1]$. We imagine a setting where in each round a player chooses $I_t \in [n]$ and observes an i.i.d $Y_{I_t,t}$ where where $Y_{I_t,t} \sim \text{Ber}(\eta_{I_t})$. Borrowing from the multi-armed bandit literature, we may also refer to the items as *arms*, and *pulling an arm* is receiving a sample from its corresponding label distribution. We will refer to this level of generality as the **stochastic noise** setting. The case when $\eta_i \in \{0,1\}$, i.e. each point $i \in [n]$ has a deterministic label $Y_{i,j} = \eta_i$ for all $j \geq 1$, will be referred to as the **persistent noise** setting. In this setting we can define $\mathcal{H}_1 = \{i : \eta_i = 1\}, \mathcal{H}_0 = [n] \setminus \mathcal{H}_1$. This is a natural setting if the experimental noise is negligible so that performing the same measurement multiple times gives the same result. A classifier is a decision rule $f : [n] \to \{0,1\}$ that assigns each item $i \in [n]$ a fixed label. We can identify any such decision rule with the set of items it maps to 1, i.e. the set $\pi = \{i : i \in [n], f(i) = 1\}$. Instead of considering all possible sets $\pi \subset [n]$, we will restrict ourselves to a finite class $\Pi \subset 2^{[n]}$. With this interpretation, one can imagine Π being a combinatorial class, such as the collection of all subsets of $[n]$ of size k, or if we have features, Π could be the sets induced by the set of all linear separators over $\{x_i\}$.

The *classification error*, or *risk* of a classifier is given by the expected number of incorrect labels, i.e.

$$R(\pi) = \mathbb{P}_{i \sim \text{Unif}([n]), Y_i \sim \text{Ber}(\eta_i)} \left(\pi(i) \neq Y_i\right) = \frac{1}{n}\left(\sum_{i \notin \pi} \eta_i + \sum_{i \in \pi}(1 - \eta_i)\right)$$

for any $\pi \in \Pi$. In the case of persistent noise the above reduces to $R(\pi) = \frac{|\pi \cap \mathcal{H}_0| + |\pi^c \cap \mathcal{H}_1|}{n} = \frac{|\mathcal{H}_1 \Delta \pi|}{n}$ where $A \Delta B = (A \cup B) - (A \cap B)$ for any sets A, B.

Problem 1:(Classification) Given a hypothesis class $\Pi \subseteq 2^{[n]}$ identify $\pi^* := \underset{\pi \in \Pi}{\text{argmin}} R(\pi)$ by requesting as few labels as possible.

As described in the introduction, in many situations we are not interested in finding the lowest risk classifier, but instead returning $\pi \in \Pi$ that contains many *discoveries* $\pi \cap \mathcal{H}_1$ without too many false alarms $\pi \cap \mathcal{H}_0$. Define $\eta_\pi := \sum_{i \in \pi} \eta_x$. The *false discovery rate (FDR)* and *true positive rate (TPR)* of a set π in the stochastic noise setting are given by

$$FDR(\pi) := 1 - \frac{\eta_\pi}{|\pi|} \quad \text{and} \quad TPR(\pi) := \frac{\eta_\pi}{\eta_{[n]}}$$

In the case of persistent noise, $FDR(\pi) = \frac{|\mathcal{H}_0 \cap \pi|}{|\pi|} = 1 - \frac{|\mathcal{H}_1 \cap \pi|}{|\pi|}$ and $TPR(\pi) = \frac{|\mathcal{H}_1 \cap \pi|}{|\mathcal{H}_1|}$. A convenient quantity that we can use to reparametrize these quantities is the *true positives:* $TP(\pi) := \sum_{i \in \pi} \eta_i$. Throughout the following we let $\Pi_\alpha = \{\pi \in \Pi : FDR(\pi) \leq \alpha\}$.

Problem 2:(Combinatorial FDR Control) Given an $\alpha \in (0,1)$ and hypothesis class $\Pi \subseteq 2^{[n]}$ identify $\pi_\alpha^* = \underset{\pi \in \Pi, FDR(\pi) \leq \alpha}{\text{argmax}} TPR(\pi)$ by requesting as few labels as possible.

In this work we are *agnostic* about how η relates to Π, ala [2, 19]. For instance we do *not* assume the Bayes classifier, $\text{argmin}_{B \in \{0,1\}^n} R(B)$ is contained in Π.

2 Related Work

Active Classification. Active learning for binary classification is a mature field (see surveys [35, 24] and references therein). The major theoretical results of the field can coarsely be partitioned into the streaming setting [2, 5, 19, 25] and the pool-based setting [18, 23, 31], noting that algorithms for the former can be used for the latter, [2], an inspiration for our algorithm, is such an example. These results rely on different complexity measures known as the splitting index, the teaching dimension, and (arguably the most popular) the disagreement coefficient.

Computational Considerations. While there have been remarkable efforts to make some of these methods more computationally efficient [5, 25], we believe even given infinite computation, many of these previous works are fundamentally inefficient from a sample complexity perspective. This stems from the fact that when applied to common combinatorial classes (for example the collection of all subsets of size k), these algorithms have sample complexities that are off by at least $\log(n)$ factors from the best algorithms for these classes. Consequently, in our work we focus on sample complexity alone, and leave matters of computational efficiency for future work.

Other Measures. Given a static dataset, the problem of finding a set or classifier that maximizes TPR subject to FDR-control in the information retrieval community is also known as finding a binary classifier that maximizes recall for a given precision level. There is extensive work on the non-adaptive sample complexity of computing measures related to precision and recall such as AUC, and F-scores [34, 8, 1]. However, there have been just a few works that consider adaptively collecting data with the goal of maximizing recall with precision constraints [33, 4], with the latter work being the most related. We will discuss it further after the statement of our main result. In [33], the problem of adaptively estimating the whole ROC curve for a threshold class is considered under a monotonicity assumption on the true positives; our algorithm is agnostic to this assumption.

Combinatorial Bandits: The pure-exploration combinatorial bandit game has been studied for the case of all subsets of $[n]$ of size k known as the Top-K problem [21, 28, 29, 27, 36, 16], the bases of a rank-k matroid (for which Top-K is a particular instance) [17, 22, 14], and in the general case [10, 15]. The combinatorial bandit component of our work (see Section A.1) is closest to [10]. The algorithm of [10] uses a disagreement-based algorithm in the spirit of Successive Elimination for bandits [21], or the A^2 for binary classification [2]. Exploring precisely what $\log$ factors are necessary has been an active area. [15] demonstrates a family of instances in which they show in the worst-case, the sample complexity must scale with $\log(|\Pi|)$. However, there are many classes like best-arm identification and matroids where sample complexity does *not* scale with $\log(|\Pi|)$ (see references above). Our own work provides some insight into what $\log$ factors are necessary by presenting our results in terms of VC dimension. In addition, we discuss situtations when a $\log(n)$ could potentially be avoided by appealing to Sauer's lemma in the supplementary material.

Multiple Hypothesis Testing. Finally, though this work shares language with the adaptive multiple-hypothesis testing literature [11, 26, 40], the goals are different. In that setting, there is a set of n hypothesis tests, where the null is that the mean of each distribution is zero and the alternative is that it is nonzero. [26] designs a procedure that adaptively allocates samples and uses the Benjamini-Hochberg procedure on p-values to return an FDR-controlled set. We are not generally interested in finding which individual arms have means that are above a fixed threshold, but instead, given a hypothesis class we want to return an FDR controlled set in the hypothesis class with high TPR. This is the situation in many structured problems in scientific discovery where the set of arms corresponds to an extremely large set of experiments and we have feature vector associated with each arm. We can't run each one but we may have some hope of identifying a region of the search space which contains many discoveries. In summary, unlike the setting of [26], Π encodes structure among the sets, we do not insist each item is sampled, and we are allowing for persistent labels - overall we are solving a different and novel problem.

3 Pool Based Active Classification

We first establish a pool based active classification algorithm that motivates our development of an adaptive algorithm for FDR-control. For each i define $\mu_i := 2\eta_i - 1 \in [-1, 1]$ so $\eta_i = \frac{1+\mu_i}{2}$. By a simple manipulation of the definition of $R(\pi)$ above we have

$$R(\pi) = \frac{1}{n}\sum_{i=1}^{n}\eta_i + \frac{1}{n}\sum_{i\in\pi}(2\eta_i - 1) = \frac{1}{n}\sum_{i=1}^{n}\eta_i + \frac{1}{n}\sum_{i\in\pi}\mu_i$$

so that $\underset{\pi\in\Pi}{\text{argmin}} R(\pi) = \underset{\pi\in\Pi}{\text{argmax}} \sum_{i\in\pi}\mu_i$. Define $\mu_\pi := \sum_{i\in\pi}\mu_i$. If for some $i \in [n]$ we map the jth draw of its label $Y_{i,j} \mapsto 2Y_{i,j} - 1$, then $\mathbb{E}[2Y_{i,j} - 1] = \mu_i$ and returning an optimal classifier in the set is equivalent to returning $\pi \in \Pi$ with the largest μ_π. Algorithm 1 exploits this. The algorithm maintains a collection of active sets $\mathcal{A}_k \subseteq \Pi$ and an active set of items $T_k \subseteq [n]$ which is the symmetric difference of all sets in $\mathcal{A}_k$. To see why we only sample in T_k, if $i \in \cap_{\pi\in\mathcal{A}_k}\pi$ then π and π' agree on the label of item i, and any contribution of arm i is canceled in each difference $\widehat{\mu}_\pi - \widehat{\mu}_{\pi'} = \widehat{\mu}_{\pi\setminus\pi'} - \widehat{\mu}_{\pi'\setminus\pi}$ for all $\pi, \pi' \in \mathcal{A}_k$ so we should not pay to sample it. In each round sets π with lower empirical means that fall outside of the confidence interval of sets with higher empirical means are removed. There may be some concern that samples from previous rounds are reused. Since our sampling strategy is uniformly drawing from $[n]$ in each round, but only paying to see a label if $I_t \in T_k$, the underlying sampling distribution is still uniform regardless of the round and so the estimate of $\widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k}$ is unbiased. In practice, since the number of samples that land in T_k follow a geometric distribution, instead of using rejection sampling we could instead have drawn a single sample from a geometric distribution and sampled that many uniformly at random from T_k.

Input: $\delta, \Pi \subset 2^{[n]}$, Confidence bound $C(\pi', \pi, t, \delta)$.
Let $\mathcal{A}_1 = \Pi, T_1 = (\cup_{\pi \in \mathcal{A}_1} \pi) - (\cap_{\pi \in \mathcal{A}_1} \pi)$, $k = 1$, $\mathcal{A}_k$ will be the active sets in round k
for $t = 1, 2, \cdots$
  if $t == 2^k$**:**
    Set $\delta_k = .5\delta/k^2$. For each π, π' let
    $\widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} = \frac{n}{t}(\sum_{s=1}^{t} R_{I_s,s}\mathbf{1}\{I_s \in \pi' \setminus \pi\} - \sum_{s=1}^{t} R_{I_s,s}\mathbf{1}\{I_s \in \pi \setminus \pi'\})$
    Set $\mathcal{A}_{k+1} = \mathcal{A}_k - \big\{\pi \in \mathcal{A}_k : \exists \pi' \in \mathcal{A}_k \text{with } \widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} > C(\pi', \pi, t, \delta_k)\big\}$.
    Set $T_{k+1} = \big(\cup_{\pi \in \mathcal{A}_{k+1}} \pi\big) - \big(\cap_{\pi \in \mathcal{A}_{k+1}} \pi\big)$.
    $k \leftarrow k + 1$
  endif
  Stochastic Noise:
    If $T_k = \emptyset$, **Break**. Otherwise, draw I_t uniformly at random from $[n]$ and if $I_t \in T_k$ receive an associated reward $R_{I_t,t} = 2Y_{I_t,t} - 1, Y_{I_t,t} \overset{iid}{\sim} \text{Ber}(\eta_{I_t})$.
  Persistent Noise:
    If $T_k = \emptyset$ or $t > n$, **Break**. Otherwise, draw I_t uniformly at random from $[n] \setminus \{I_s : 1 \le s < t\}$ and if $I_t \in T_k$ receive associated reward $R_{I_t,t} = 2Y_{I_t,t} - 1$, $Y_{I_t,t} = \eta_{I_t}$.
Output: $\pi' \in \mathcal{A}_k$ such that $\widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} \ge 0$ for all $\pi \in \mathcal{A}_k \setminus \pi'$

Algorithm 1: Action Elimination for Active Classification

For any $\mathcal{A} \subseteq 2^{[n]}$ define $V(\mathcal{A})$ as the VC-dimension of a collection of sets $\mathcal{A}$. Given a family of sets, $\Pi \subseteq 2^{[n]}$, define $B_1(k) := \{\pi \in \Pi : |\pi| = k\}$, $B_2(k, \pi') := \{\pi \in \Pi : |\pi \Delta \pi'| = k\}$. Also define the following complexity measures:

$$V_\pi := V(B_1(|\pi|)) \wedge |\pi| \text{ and } V_{\pi,\pi'} := \max\{V(B_2(|\pi\Delta\pi'|, \pi), V(B_2(|\pi\Delta\pi'|, \pi'))\} \wedge |\pi\Delta\pi'|$$

In general $V_\pi, V_{\pi,\pi'} \le V(\Pi)$. A contribution of our work is the development of confidence intervals that do not depend on a union bound over the class but instead on local VC dimensions. These are described carefully in Lemma 1 in the supplementary materials.

Theorem 1 *For each $i \in [n]$ let $\mu_i \in [-1, 1]$ be fixed but unknown and assume $\{R_{i,j}\}_{j=1}^{\infty}$ is an i.i.d sequence of random variables such that $\mathbb{E}[R_{i,j}] = \mu_i$ and $R_{i,j} \in [-1, 1]$. Define $\widetilde{\Delta}_\pi = |\mu_\pi - \mu_{\pi^*}|/|\pi\Delta\pi^*|$, and*

$$\tau_\pi = \frac{V_{\pi,\pi^*}}{|\pi^*\Delta\pi|} \frac{1}{\widetilde{\Delta}_\pi^2} \log\Big(n \log(\widetilde{\Delta}_\pi^{-2})/\delta\Big).$$

Using $C(\pi, \pi', t, \delta) := \sqrt{\frac{8|\pi\Delta\pi'|nV_{\pi,\pi'}\log\left(\frac{n}{\delta}\right)}{t}} + \frac{4nV_{\pi,\pi'}\log\left(\frac{n}{\delta}\right)}{3t}$ for a fixed constant c, with probability greater than $1 - \delta$, in the ***stochastic noise*** *setting Algorithm 1 returns π_* after a number of samples no more than $c\sum_{i=1}^{n} \max_{\pi\in\Pi: i\in\pi\Delta\pi^*} \tau_\pi$ and in the* ***persistent noise*** *setting the number of samples needed is no more than $c\sum_{i=1}^{n} \min\{1, \max_{\pi\in\Pi: i\in\pi\Delta\pi^*} \tau_\pi\}$*

One always has $1/|\pi^*\Delta\pi| \le V_{\pi,\pi^*}/|\pi^*\Delta\pi| \le 1$ and both bounds are achievable by different classes Π. In addition, in terms of risk $\widetilde{\Delta}_\pi = |\mu_\pi - \mu_{\pi^*}|/|\pi\Delta\pi^*| = n|R(\pi) - R(\pi^*)|/|\pi\Delta\pi^*|$. Since sampling is done without replacement for persistent noise, there are improved confidence intervals that one can use in that setting described in Lemma 1 in the supplementary materials. Finally, if we had sampled non-adaptively, i.e. without rejection sampling, we would have had a sample complexity of $O(n \max_{i\in[n]} \max_{\pi:\Pi: i\in\pi\Delta\pi^*} \tau_\pi)$.

Remark: Our rewards could be drawn from arbitrary distributions, not just Bernoulli label distributions. In fact if we allow $R_{I_t,t} \sim^{i.i.d.} \nu_i$, where ν_i is a distribution supported on $[-1, 1]$ with $\mathbb{E}[\nu_i] = \mu_i$, then Algorithm 1 gives state of the art results for the more general *pure exploration combinatorial bandit* problem and furthermore Theorem 1 holds verbatim. Algorithm 1 is similar to previous action elimination algorithms for combinatorial bandits in the literature, e.g. Algorithm 4 in [10]. However, unlike previous algorithms, we do not insist on sampling each item once, an unrealistic requirement for classification settings. We discuss this connection further in Section A.1 in the supplementary materials and prove Theorem 1 in this more general setting.

3.1 Comparison with previous Active Classification results.

One Dimensional Thresholds: In the bound of Theorem 1, a natural question to ask is whether the $\log(n)$ dependence can be improved. In the case of nested classes, such as thresholds on a line, we can replace the $\log(n)$ with a $\log\log(n)$ using empirical process theory. This leads to confidence intervals dependent on $\log\log(n)$ that can be used in place of $C(\pi', \pi, t, \delta)$ in Algorithm 1 (see sections C for the confidence intervals and A.1 for a longer discussion). Under specific specific noise models we can give a more interpretable sample complexity. Let $h \in (0, 1]$, $\alpha \geq 0$, $z \in [0, 1]$ for some $i \in [n-1]$ and assume that $\eta_i = \frac{1}{2} + \frac{\text{sign}(z-i/n)}{2} h|z - i/n|^\alpha$ so that $\mu_i = h|z - i/n|^\alpha \text{sign}(z - i/n)$ (this would be a reasonable noise model for topology $\alpha\alpha\alpha$ in the introduction). Let $\Pi = \{[k] : k \leq n\}$. In this case, inspecting the dominating term of Theorem 1 for $i \in \pi^*$ we have $\arg\max_{\pi \in \Pi : i \in \pi \Delta \pi^*} \frac{V_{\pi,\pi^*}}{|\pi \Delta \pi^*|} \frac{1}{\tilde{\Delta}_\pi^2} = [i]$ and takes a value of $\left(\frac{1+\alpha}{h}\right)^2 n^{-1}(z - i/n)^{-2\alpha-1}$. Upper bounding the other terms and summing, the sample complexities can be calculated to be $O(\log(n)\log(\log(n)/\delta)/h^2)$ if $\alpha = 0$, and $O(n^{2\alpha}\log(\log(n)/\delta)/h^2)$ if $\alpha > 0$. These rates match the minimax lower bound rates given in [12] up to $\log\log$ factors. Unlike the algorithms given there, our algorithm works in the *agnostic* setting, i.e. it is making no assumptions about whether the Bayes classifier is in the class. In the case of non-adaptive sampling, the sum is replaced with the max times n yielding $n^{2\alpha+1}\log(\log(n)/\delta)/h^2$ which is substantially worse than adaptive sampling.

Comparison to previous algorithms: One of the foundational works on active learning is the DHM algorithm of [19] and the A^2 algorithm that preceded it [2]. Similar in spirit to our algorithm, DHM requests a label only when it is uncertain how π^* would label the current point. In general the analysis of the DHM algorithm can not characterize the contribution of each arm to the overall sample complexity leading to sub-optimal sample complexity for combinatorial classes. For example in the the case when $\Pi = \{[i]\}_{i=1}^n$, with $i^* = \arg\max_{i \in [n]} \mu_i$, ignoring logarithmic factors, one can show for this problem the bound of Theorem 1 of [19] scales like $n^2 \max_{i \neq i_*}(\mu_{i^*} - \mu_i^{-2})$ which is substantially worse than our bound for this problem which scales like $\sum_{i \neq i_*} \Delta_i^{-2}$. Similar arguments can be made for other combinatorial classes such as all subsets of size k. While we are not particularly interested in applying algorithms like DHM to this specific problem, we note that the style of its analysis exposes such a gross inconsistency with past analyses of the best known algorithms that the approach leaves much to be desired. For more details, please see A.3 in the supplementary materials.

4 Combinatorial FDR Control

Algorithm 2 provides an active sampling method for determining $\pi \in \Pi$ with $FDR(\pi) \leq \alpha$ and maximal TPR, which we denote as π_α^*. Since $TPR(\pi) = TP(\pi)/\eta_{[n]}$, we can ignore the denominator and so maximizing the TPR is the same as maximizing TP. The algorithm proceeds in epochs. At all times a collection $\mathcal{A}_k \subseteq \Pi$ of active sets is maintained along with a collection of FDR-controlled sets $\mathcal{C}_k \subseteq \mathcal{A}_k$. In each time step, random indexes I_t and J_t are sampled from the union $S_k = \cup_{\pi \in \mathcal{A}_k \setminus \mathcal{C}_k} \pi$ and the symmetric difference $T_k = \cup_{\pi \in \mathcal{A}_k} \pi - \cap_{\pi \in \mathcal{A}_k} \pi$ respectively. Associated random labels $Y_{I_t,t}, Y_{J_t,t} \in \{0, 1\}$ are then obtained from the underlying label distributions $\text{Ber}(\eta_{I_t})$ and $\text{Ber}(\eta_{J_t})$. At the start of each epoch, any set with a FDR that is statistically known to be under α is added to $\mathcal{C}_k$, and any sets whose FDR are greater than α are removed from $\mathcal{A}_k$ in condition 1. Similar to the active classification algorithm of Figure 1, a set $\pi \in \mathcal{A}_k$ is removed in condition 2 if $TP(\pi)$ is shown to be statistically less than $TP(\pi')$ for some $\pi' \in \mathcal{C}_k$ that, crucially, is FDR controlled. In general there may be many sets $\pi \in \Pi$ such that $TP(\pi) > TP(\pi_\alpha^*)$ that are not FDR-controlled. Finally in condition 3, we exploit the positivity of the η_i's: if $\pi \subset \pi'$ then deterministically $TP(\pi) \leq TP(\pi')$, so if π' is FDR controlled it can be used to eliminate π. The choice of T_k is motivated by active classification: we only need to sample in the symmetric difference. To determine which sets are FDR-controlled it is important that we sample in the entirety of the union of all $\pi \in \mathcal{A}_k \setminus \mathcal{C}_k$, not just the symmetric difference of the $\mathcal{A}_k$, which motivates the choice of S_k. In practical experiments persistent noise is not uncommon and avoids the potential for unbounded sample complexities that potentially occur when $FDR(\pi) \approx \alpha$. Figure 2 demonstrates a model run of the algorithm in the case of five sets $\Pi = \{\pi_1, \ldots, \pi_5\}$.

Recall that Π_α is the subset of Π that is FDR-controlled so that $\pi_\alpha^* = \arg\max_{\pi \in \Pi_\alpha} TP(\pi)$. The following gives a sample complexity result for the number of rounds before the algorithm terminates.

Input: Confidence bounds $C_1(\pi, t, \delta), C_2(\pi, \pi', t, \delta)$
$\mathcal{A}_k \subset \Pi$ will be the set of active sets in round k. $\mathcal{C}_k \subset \Pi$ is the set of FDR-controlled policies in round k.
$\mathcal{A}_1 = \Pi,\ \mathcal{C}_1 = \emptyset,\ S_1 = \cup_{\pi\in\Pi}\pi,\ T_1 = \bigcup_{\pi\in\Pi}\pi - \bigcap_{\pi\in\Pi}\pi, k = 1.$
for $t = 1, 2, \cdots$
 if $t = 2^k$**:**
  Let $\delta_k = .25\delta/k^2$
  For each set $\pi \in \mathcal{A}_k$, and each pair $\pi', \pi \in \mathcal{A}_k$ update the estimates:
  $\widehat{FDR}(\pi) := 1 - \frac{n}{|\pi|t}\sum_{s=1}^{t} Y_{I_s,s}\mathbf{1}\{I_s \in \pi\}$
  $\widehat{TP}(\pi') - \widehat{TP}(\pi) := \frac{n}{t}\left(\sum_{s=1}^{t} Y'_{J_s,s}\mathbf{1}\{J_s \in \pi'\backslash\pi\} - \sum_{s=1}^{t} Y'_{J_s,s}\mathbf{1}\{J_s \in \pi\backslash\pi'\}\right)$
  Set $\mathcal{C}_{k+1} = \mathcal{C}_k \cup \{\pi \in \mathcal{A}_k \setminus \mathcal{C}_k : \widehat{FDR}(\pi) + C_1(\pi, t, \delta_k)/|\pi| \le \alpha\}$
  Set $\mathcal{A}_{k+1} = \mathcal{A}_k$
  Remove any π from $\mathcal{A}_{k+1}$ and $\mathcal{C}_{k+1}$ such that one of the conditions is true:
   1. $\widehat{FDR}(\pi) - C_1(\pi, t, \delta_k)/|\pi| > \alpha$
   2. $\exists\pi' \in \mathcal{C}_{k+1}$ with $\widehat{TP}(\pi') - \widehat{TP}(\pi) > C_2(\pi, \pi', t, \delta_k)$ and add π to a set R
  Remove any π from $\mathcal{A}_{k+1}$ and $\mathcal{C}_{k+1}$ such that:
   3. $\exists\pi' \in \mathcal{C}_{k+1} \cup R$, such that $\pi \subset \pi'$.
  Set $S_{k+1} := \bigcup_{\pi\in\mathcal{A}_{k+1}\setminus\mathcal{C}_{k+1}}\pi$, and $T_{k+1} = \bigcup_{\pi\in\mathcal{A}_{k+1}}\pi - \bigcap_{\pi\in\mathcal{A}_{k+1}}\pi$.
  $k \leftarrow k + 1$
 endif

 Stochastic Noise:
  if $|\mathcal{A}_k| = 1$, **Break**. Otherwise:
  Sample $I_t \sim \text{Unif}([n])$. If $I_t \in S_k$, then receive a label $Y_{I_t,t} \sim \text{Ber}(\eta_{I_t})$.
  Sample $J_t \sim \text{Unif}([n])$. If $J_t \in T_k$, then receive a label $Y'_{J_t,t} \sim \text{Ber}(\eta_{J_t})$.
 Persistent Noise:
  If $|\mathcal{A}_k| = 1$ or $t > n$, **Break.** Otherwise:
  Sample $I_t \sim [n]\backslash\{I_s : 1 \le s < t\}$. If $I_t \in S_k$, then receive a label $Y_{I_t,t} = \eta_{I_t}$.
  Sample $J_t \sim [n]\backslash\{J_s : 1 \le s < t\}$. If $J_t \in T_k$, then receive a label $Y'_{J_t,t} = \eta_{J_t}$.
 Return $\max_{t\in\mathcal{C}_{k+1}} \widehat{TP}(\pi)$

Algorithm 2: Active FDR control in persistent and bounded noise settings.

Theorem 2 *Assume that for each $i \le n$ there is an associated $\eta_i \in [0, 1]$ and $\{Y_{i,j}\}_{j=1}^{\infty}$ is an i.i.d. sequence of random variables such that $Y_{i,j} \sim Ber(\eta_i)$. For any $\pi \in \Pi$ define $\Delta_{\pi,\alpha} = |FDR(\pi) - \alpha|$, and $\widetilde{\Delta}_\pi = |TP(\pi^*_\alpha) - TP(\pi)|/|\pi\Delta\pi^*| = |TP(\pi^*_\alpha \setminus \pi) - TP(\pi \setminus \pi^*_\alpha)|/|\pi\Delta\pi^*|$, and*

$$s^{FDR}_\pi = \frac{V_\pi}{|\pi|}\frac{1}{\Delta^2_{\pi,\alpha}}\log\left(n\log(\Delta^{-2}_{\pi,\alpha})/\delta\right), \quad s^{TP}_\pi = \frac{V_{\pi,\pi^*_\alpha}}{|\pi\Delta\pi^*_\alpha|}\frac{1}{\widetilde{\Delta}^2_\pi}\log\left(n\log(\widetilde{\Delta}^{-2}_\pi)/\delta\right)$$

*In addition define $T^{FDR}_\pi = \min\{s^{FDR}_\pi,\ \max\{s^{TP}_\pi,\ s^{FDR}_{\pi^*_\alpha}\},\ \min_{\substack{\pi'\in\Pi_\alpha\\ \pi\subset\pi'}} s^{FDR}_{\pi'}\}$ and $T^{TP}_\pi = \min\{\max\{s^{TP}_\pi,\ s^{FDR}_{\pi^*_\alpha}\},\ \min_{\substack{\pi'\in\Pi_\alpha\\ \pi\subset\pi'}} s^{FDR}_{\pi'}\}$. Using $C_1(\pi, t, \delta) := \sqrt{\frac{4|\pi|nV_\pi\log\left(\frac{n}{\delta}\right)}{t}} + \frac{4nV_\pi\log\left(\frac{n}{\delta}\right)}{3t}$ and $C_2 = C$ for C defined in Theorem 1, for a fixed constant c, with probability at least $1 - \delta$, in the **stochastic noise** setting Algorithm 2 returns π^*_α after a number of samples no more than*

$$\underbrace{c\sum_{i=1}^{n}\max_{\pi\in\Pi:i\in\pi} T^{FDR}_\pi}_{FDR-Control} + \underbrace{c\sum_{i=1}^{n}\max_{\pi\in\Pi_\alpha:i\in\pi\Delta\pi^*_\alpha} T^{TP}_\pi}_{TPR-Elimination}$$

*and in the **persistent noise** setting returns π^*_α after no more than $c\sum_{i=1}^{n}\min\left\{1, \left(\max_{\pi\in\Pi:i\in\pi} T^{FDR}_\pi + \max_{\pi\in\Pi_\alpha:i\in\pi\Delta\pi^*_\alpha} T^{TP}_\pi\right)\right\}$*

Though this result is complicated, each term is understood by considering each way a set can be removed and the time at which an arm i will stop being sampled. We now unpack underbraced terms.

Sample Complexity of FDR-Control In any round where there exists a set $\pi \in \mathcal{A}_k \setminus \mathcal{C}_k$ with arm $i \in \pi$, i.e. π is not yet FDR controlled, there is the potential for sampling $i \in \mathcal{S}_k$. A set π only leaves $\mathcal{A}_k$ if $i)$ it is shown to not be FDR controlled (condition 1 of the algorithm), $ii)$ because an FDR

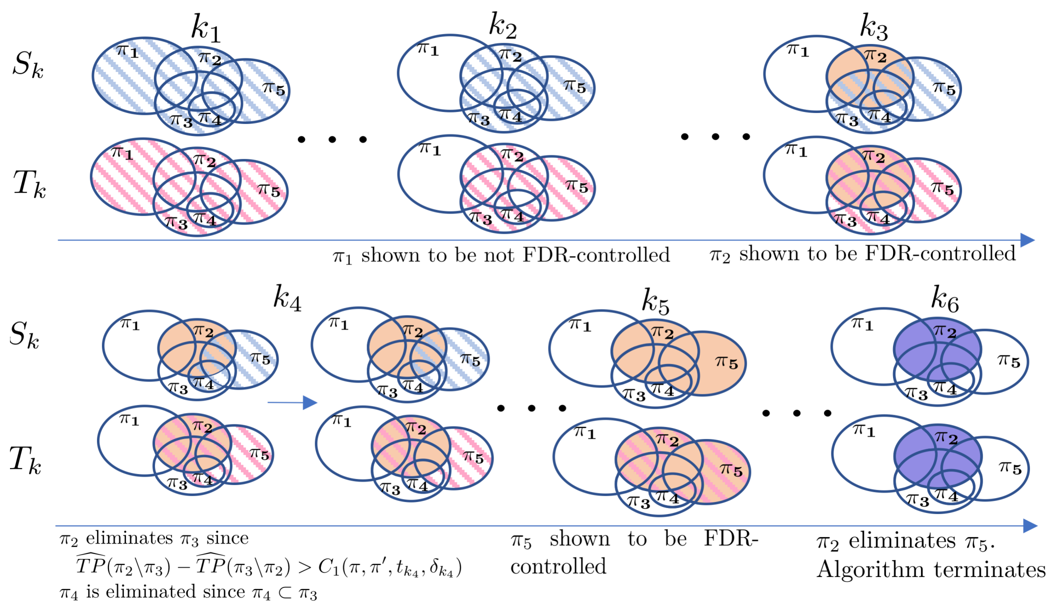

Figure 2: Example run of Algorithm 2, showing the evolution of sampling regions S_k (blue stripes), T_k (pink stripes) and FDR controlled sets C_k (orange fill) at each time k_t.

controlled set eliminates it on the basis of TP (condition 2), or iii) it is contained in an FDR controlled set (condition 3). These three cases reflect the three arguments of the $\min$ in the defined quantity T_π^{FDR}, respectively. Taking the maximum over all sets containing an arm i and summing over all i gives the total FDR-control term. This is a large savings relative to naive non-adaptive algorithms that sample until every set π in Π was FDR controlled which would take $O(n \max_{\pi \in \Pi} s_\pi^{FDR})$ samples.

Sample Complexity of TPR-Elimination An FDR-controlled set $\pi \in \Pi_\alpha$ is only removed from $\mathcal{C}_k$ when eliminated by an FDR-controlled set with higher TP or if it is removed because it is contained in an FDR-controlled set. In general we can upper bound the former time by the samples needed for π_α^* to eliminate π once we know π_α^* is FDR controlled - this gives rise to $\max_{\pi \in \Pi_\alpha : i \in \pi \Delta \pi_\alpha^*} T_\pi^{TP}$. Note that sets are removed in a procedure mimicking active classification and so the active gains there apply to this setting as well. A naive passive algorithm that continues to sample until both the FDR of every set is determined, and π_α^* has higher TP than every other FDR-controlled set gives a significantly worse sample complexity of $O(n \max\{\max_{\pi \in \Pi_\alpha} s_\pi^{FDR}, \max_{\pi \notin \Pi_\alpha} s_\pi^{TP}\})$.

Comparison with [4]. Similar to our proposed algorithm, [4] samples in the union of all active sets and maintains statistics on the empirical FDR of each set, along the way removing sets that are not FDR-controlled or have lower TPR than an FDR-controlled set. However, they fail to sample in the symmetric difference, missing an important link between FDR-control and active classification. They also only consider the case of persistent noise. Their proven sample complexity results are no better than those achieved by the passive algorithm that samples each item uniformly, which is precisely the sample complexity described at the end of the previous paragraph.

One Dimensional Thresholds Consider a stylized modeling of the topology $\beta\alpha\beta\beta$ from the introduction in the persistent noise setting where $\Pi = \{[t] : t \leq n\}$, $\eta_i \sim$ Ber$(\beta \mathbf{1}\{i \leq z\})$ with $\beta < .5$, and $z \in [n]$ is assumed to be small, i.e., we assume that there is only a small region in which positive labels can be found and the Bayes classifier is just to predict 0 for all points. Assuming $\alpha > 1 - \beta$, one can show the sample complexity of Algorithm 2 satisfies $O((1-\alpha)^{-2}(\log(n/(1-\alpha)) + (1+\beta)z/(1-\alpha)))$ while any naive non-adaptive sampling strategy will take at least $O(n)$ samples.

Implementation. For simple classes Π such as thresholds or axis aligned rectangles, our algorithm can be made computationally efficient. But for more complex classes there may be a wide gap between theory and practice, just as in classification [35, 19]. However, the algorithm motivates two key ideas - sample in the union of potentially good sets to learn which are FDR controlled, and sample in the symmetric difference to eliminate sets. The latter insight was originally made by A^2 in the case of classification and has justified heuristics such as uncertainty sampling [35]. Developing analogous heuristics for the former case of FDR-control is an exciting avenue of future work.

References

[1] Shivani Agarwal, Thore Graepel, Ralf Herbrich, Sariel Har-Peled, and Dan Roth. Generalization bounds for the area under the roc curve. *Journal of Machine Learning Research*, 6(Apr):393–425, 2005.

[2] Maria-Florina Balcan, Alina Beygelzimer, and John Langford. Agnostic active learning. *Journal of Computer and System Sciences*, 75(1):78–89, 2009.

[3] Rémi Bardenet, Odalric-Ambrym Maillard, et al. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.

[4] Paul N Bennett, David M Chickering, Christopher Meek, and Xiaojin Zhu. Algorithms for active classifier selection: Maximizing recall with precision constraints. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 711–719. ACM, 2017.

[5] Alina Beygelzimer, Sanjoy Dasgupta, and John Langford. Importance weighted active learning. *arXiv preprint arXiv:0812.4952*, 2008.

[6] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[7] Olivier Bousquet. A bennett concentration inequality and its application to suprema of empirical processes. *Comptes Rendus Mathematique*, 334(6):495–500, 2002.

[8] Kendrick Boyd, Kevin H Eng, and C David Page. Area under the precision-recall curve: Point estimates and confidence intervals. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 451–466. Springer, 2013.

[9] Diogo M Camacho, Katherine M Collins, Rani K Powers, James C Costello, and James J Collins. Next-generation machine learning for biological networks. *Cell*, 2018.

[10] Tongyi Cao and Akshay Krishnamurthy. Disagreement-based combinatorial pure exploration: Efficient algorithms and an analysis with localization. *arXiv preprint arXiv:1711.08018*, 2017.

[11] Rui M Castro et al. Adaptive sensing performance lower bounds for sparse signal detection and support estimation. *Bernoulli*, 20(4):2217–2246, 2014.

[12] Rui M Castro and Robert D Nowak. Minimax bounds for active learning. *IEEE Transactions on Information Theory*, 54(5):2339–2353, 2008.

[13] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[14] Lijie Chen, Anupam Gupta, and Jian Li. Pure exploration of multi-armed bandit under matroid constraints. In *Conference on Learning Theory*, pages 647–669, 2016.

[15] Lijie Chen, Anupam Gupta, Jian Li, Mingda Qiao, and Ruosong Wang. Nearly optimal sampling algorithms for combinatorial pure exploration. In *Conference on Learning Theory*, pages 482–534, 2017.

[16] Lijie Chen, Jian Li, and Mingda Qiao. Nearly instance optimal sample complexity bounds for top-k arm selection. In *Artificial Intelligence and Statistics*, pages 101–110, 2017.

[17] Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems*, pages 379–387, 2014.

[18] Sanjoy Dasgupta. Coarse sample complexity bounds for active learning. In *Advances in neural information processing systems*, pages 235–242, 2006.

[19] Sanjoy Dasgupta, Daniel J Hsu, and Claire Monteleoni. A general agnostic active learning algorithm. In *Advances in neural information processing systems*, pages 353–360, 2008.

[20] Devdatt P Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[21] Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of machine learning research*, 7(Jun):1079–1105, 2006.

[22] Victor Gabillon, Alessandro Lazaric, Mohammad Ghavamzadeh, Ronald Ortner, and Peter Bartlett. Improved learning complexity in combinatorial pure exploration bandits. In *Artificial Intelligence and Statistics*, pages 1004–1012, 2016.

[23] Steve Hanneke. Teaching dimension and the complexity of active learning. In *International Conference on Computational Learning Theory*, pages 66–81. Springer, 2007.

[24] Steve Hanneke et al. Theory of disagreement-based active learning. *Foundations and Trends® in Machine Learning*, 7(2-3):131–309, 2014.

[25] Tzu-Kuo Huang, Alekh Agarwal, Daniel J Hsu, John Langford, and Robert E Schapire. Efficient and parsimonious agnostic active learning. In *Advances in Neural Information Processing Systems*, pages 2755–2763, 2015.

[26] Kevin Jamieson and Lalit Jain. A bandit approach to multiple testing with false discovery control. In *Advances in Neural Information Processing Systems*, 2018.

[27] Kevin Jamieson, Matthew Malloy, Robert Nowak, and Sébastien Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory*, pages 423–439, 2014.

[28] Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. Pac subset selection in stochastic multi-armed bandits. In *ICML*, volume 12, pages 655–662, 2012.

[29] Zohar Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *International Conference on Machine Learning*, pages 1238–1246, 2013.

[30] Armaghan W Naik, Joshua D Kangas, Devin P Sullivan, and Robert F Murphy. Active machine learning-driven experimentation to determine compound effects on protein patterns. *Elife*, 5:e10047, 2016.

[31] Robert D Nowak. The geometry of generalized binary search. *IEEE Transactions on Information Theory*, 57(12):7893–7906, 2011.

[32] Gabriel J Rocklin, Tamuka M Chidyausiku, Inna Goreshnik, Alex Ford, Scott Houliston, Alexander Lemak, Lauren Carter, Rashmi Ravichandran, Vikram K Mulligan, Aaron Chevalier, et al. Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science*, 357(6347):168–175, 2017.

[33] Ashish Sabharwal and Yexiang Xue. Adaptive stratified sampling for precision-recall estimation. pages 825–834, 2018.

[34] Christoph Sawade, Niels Landwehr, and Tobias Scheffer. Active estimation of f-measures. In *Advances in Neural Information Processing Systems*, pages 2083–2091, 2010.

[35] Burr Settles. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1):1–114, 2012.

[36] Max Simchowitz, Kevin Jamieson, and Benjamin Recht. The simulator: Understanding adaptive sampling in the moderate-confidence regime. In *Conference on Learning Theory*, pages 1794–1834, 2017.

[37] Yuriy Sverchkov and Mark Craven. A review of active learning approaches to experimental design for uncovering biological networks. *PLoS computational biology*, 13(6):e1005466, 2017.

[38] Lorillee Tallorin, JiaLei Wang, Woojoo E Kim, Swagat Sahu, Nicolas M Kosa, Pu Yang, Matthew Thompson, Michael K Gilson, Peter I Frazier, Michael D Burkart, et al. Discovering de novo peptide substrates for enzymes using machine learning. *Nature communications*, 9(1):5253, 2018.

[39] Lu Zhang, Jianjun Tan, Dan Han, and Hao Zhu. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug discovery today*, 22(11):1680–1685, 2017.

[40] Martin J Zhang, James Zou, and David Tse. Adaptive monte carlo multiple testing via multi-armed bandits. *arXiv preprint arXiv:1902.00197*, 2019.

A Proofs

A.1 Connections to Combinatorial Bandits

A closely related problem to classification is the *pure-exploration combinatorial bandit* problem. As above we have access to a set of arms $[n]$, and associated to each arm is an unknown distribution ν_i. We let $\{R_{i,j}\}_{j=1}^{\infty}$ be a sequence of random variables where $R_{i,j} \sim \nu_i$ is the jth draw from ν_i satisfying $\mathbb{E}[R_{i,j}] = \mu_i \in [-1, 1]$. In the persistent noise setting we assume that ν_i is a point mass at $\mu_i \in [-1, 1]$. Given a collection of sets $\Pi \subseteq 2^{[n]}$, for each $\pi \in \Pi$ we define $\mu_\pi := \sum_{i\in\pi} \mu_i$ the sum of means in π.

> **Problem 3: (Combinatorial Bandits)** Given a hypothesis class $\Pi \subseteq 2^{[n]}$ identify $\pi^* = \underset{\pi\in\Pi}{\operatorname{argmax}} \mu_\pi$ by requesting as few labels as possible.

The combinatorial bandit extends many problems considered in the multi-armed bandit literature. For example if $\Pi = \{\{i\} : i \in [n]\}$ then this is equivalent to the best-arm identification problem.

As discussed in Section C, returning to the classification setting for a moment: for each i define $\mu_i := 2\eta_i - 1 \in [-1, 1]$ so $\eta_i = \frac{1+\mu_i}{2}$. By a simple manipulation of the definition of $R(\pi)$ above we have

$$R(\pi) = \frac{1}{n}\sum_{i=1}^{n} \eta_i + \frac{1}{n}\sum_{i\in\pi}(2\eta_i - 1) = \frac{1}{n}\sum_{i=1}^{n} \eta_i + \frac{1}{n}\sum_{i\in\pi} \mu_i$$

so that $\underset{\pi\in\Pi}{\operatorname{argmin}} \quad R(\pi) = \underset{\pi\in\Pi}{\operatorname{argmax}} \sum_{i\in\pi} \mu_i$. Hence, if for some $i \in [n]$ we map the jth draw of its label $Y_{i,j} \mapsto 2Y_{i,j} - 1$ then the $\mathbb{E}[2Y_{i,j} - 1] = \mu_i$ and returning an optimal classifier in the set is equivalent to returning a subset π with the largest μ_π.

The connection between FDR control and combinatorial bandits is more direct: we are seeking to find $\pi \in \Pi$ with maximum η_π subject to FDR-constraints. This already highlights a key difference between classification and FDR-control. In one we choose to sample to maximize η_π subject to FDR constraints where each $\eta_i \in [0, 1]$, whereas in classification we are trying to maximize μ_π where each $\mu_i \in [-1, 1]$. A major consequence of this difference is that $\eta_\pi \leq \eta_{\pi'}$ whenever $\pi \subseteq \pi'$, but such a condition does not hold for $\mu_\pi, \mu_{\pi'}$.

Motivating the sample complexity: As mentioned above, the general combinatorial bandit problem is considered in [10]. There they present an algorithm with sample complexity,

$$C\sum_{i=1}^{n} \max_{\pi: i\in\pi\Delta\pi^*} \frac{1}{|\pi\Delta\pi^*|}\frac{1}{\tilde{\Delta}_\pi^2} \log\left(\max(|B(|\pi\Delta\pi^*|, \pi)|, |B(|\pi\Delta\pi^*|, \pi^*)|)\frac{n}{\delta}\right)$$

This complexity parameter is difficult to interpret directly so we compare it to one more familiar in statistical learning - the VC dimension. To see how this sample complexity relates to ours in Theorem 1, note that $\log_2|B(k, \pi^*)| \leq \log_2\binom{n}{k} \lesssim k\log_2(n)$. Thus by the Sauer-Shelah lemma, $V(B(r, \pi^*)) \lesssim \log_2(|B(r, \pi^*)|) \lesssim \min\{V(B(r, \pi^*)), r\}\log_2(n)$ where $\lesssim$ hides a constant. The proof of Lemma 1 below effectively combines these two facts along with a union bound over all sets in $B(r, \pi^*)$.

It's natural to ask whether the $\log(n)$ on the right can be dropped. In specific examples,like nested classes, tools from empirical process theory (see Theorem 13.7 in [6]) imply that it can be improved to a $\log\log(n)$. We give such an example where the $\log(n)$ is not necessary in Appendix C for the case of one-dimensional thresholds.

A.2 Confidence Bounds for Combinatorial Bandits

In this section, we build confidence intervals useful in our general combinatorial bandit setup discussed in the previous section. The union bounds presented are motivated by those in [10]. The constants used in the case without replacement are motivated by Corollary 3.6 in [3].

Lemma 1 *Assume that for each arm $i \leq n$ there is an associated distribution ν_i with support $[-1, 1]$, mean μ_i and variance $\sigma_i^2 \leq 1$. Assume access to the observations $(I_1, y_{I_1}) \cdots, (I_t, y_{I_t})$ in two different but related settings, let $s \leq t$,*

1. ***Stochastic Noise*** *$I_s \sim$ Unif$([n])$ and $y_{I_s} \sim \nu_{I_s}$.*

2. ***Persistent Noise*** *$I_s \in [n]$ are drawn without replacement, $y_{I_s} = \mu_{I_s}$, $s \leq n$*

Let $\widehat{\mu}_\pi = \frac{n}{T}\sum_{k=1}^{T} y_s \mathbf{1}\{I_s \in \pi\}$. Then

1. *With probability greater than $1-\delta$ for all $\pi \in \Pi$*

$$|\widehat{\mu}_\pi - \mu_\pi| \leq C_1(\pi, t, \delta) := \sqrt{\frac{4\rho_t|\pi|nV_\pi \log\left(\frac{n}{\delta}\right)}{t}} + \frac{4n\kappa_t V_\pi \log\left(\frac{n}{\delta}\right)}{3t} \tag{1}$$

2. *Fix $\pi' \in \Pi$. With probability greater than $1-\delta$ for all $t > 0$ and $\pi \in \Pi$*

$$\left|\widehat{\mu}_{\pi'\setminus\pi} - \widehat{\mu}_{\pi\setminus\pi'} - (\mu_{\pi'\setminus\pi} - \mu_{\pi\setminus\pi'})\right| \leq C_2(\pi, \pi', t, \delta) := \sqrt{\frac{8\rho_t|\pi\Delta\pi'|nV_{\pi,\pi'} \log\left(\frac{n}{\delta}\right)}{t}} \tag{2}$$

$$+ \frac{4\kappa_t nV_{\pi,\pi'} \log\left(\frac{n}{\delta}\right)}{3t} \tag{3}$$

where $\rho_t, \kappa_t = 1$ in the stochastic case and in the persistent case

$$\rho_t = \begin{cases} 1 - \frac{t-1}{n} & t \leq n/2 \\ 1 - \frac{t}{n} & t \geq n/2 \end{cases} \qquad \kappa_t = \frac{4}{3} + \begin{cases} \sqrt{\frac{t(t-1)}{n(n-t+1)}} & t \leq n/2 \\ \sqrt{\frac{(n-t-1)(n-t)}{(t+1)n}} & t \geq n/2 \end{cases}$$

Note that by negative associativity the confidence bounds that hold in the case of sampling with replacement also hold when sampling without replacement.

Proof: Define the complexity measures

$$B_1(k) = \{\pi \in \mathcal{A} : |\pi| = k\}, B_2(k, \pi') = \{\pi \in \mathcal{A} : |\pi\Delta\pi'| = k\}.$$

Firstly note that for any $\pi \in \Pi$

$$\begin{aligned} \text{var}(\widehat{\mu}_\pi) &= \frac{n^2}{T}\text{var}\left(y_1 \mathbf{1}\{I_1 \in \pi\}\right) \\ &= \frac{n^2}{T}\left(\mathbb{E}[y_1^2 \mathbf{1}\{I_1 \in \pi \setminus \pi'\}] - \left(\frac{1}{n}\sum_{i \in \pi\setminus\pi'} \mu_i\right)^2\right) \\ &\leq \frac{n^2}{T}\left(\frac{1}{n}\sum_{i\in\pi}(\sigma_i^2 + \mu_i^2)\right) \leq \frac{2|\pi|n}{T} \end{aligned}$$

Thus by Bernstein's inequality and a union bound,

$$\begin{aligned} \mathbb{P}\left(\exists \pi \in \Pi : |\widehat{\mu}_\pi - \mu_\pi| > \sqrt{\frac{2|\pi|n\log(nB_1(|\pi|)/\delta)}{T}} + \frac{2n\log(nB_1(|\pi|)/\delta)}{3T}\right) &\leq \sum_{\pi\in\Pi} \frac{\delta}{nB_1(|\pi|)} \\ &\leq \sum_{k=1}^{n} B_1(k)\frac{\delta}{nB_1(k)} \leq \delta \end{aligned}$$

For the second assertion, firstly note that for any π, π', $\widehat{\mu}_\pi - \widehat{\mu}_{\pi'} = \widehat{\mu}_{\pi \setminus \pi'} - \widehat{\mu}_{\pi' \setminus \pi}$ and so

$$\begin{aligned}
&\mathrm{var}(\widehat{\mu}_\pi - \widehat{\mu}_{\pi'}) \\
&\quad = \mathrm{var}(\widehat{\mu}_{\pi\setminus\pi'} - \widehat{\mu}_{\pi'\setminus\pi}) \\
&\quad = \mathrm{var}(\widehat{\mu}_{\pi\setminus\pi'}) + \mathrm{var}(\widehat{\mu}_{\pi'\setminus\pi})) \\
&\quad = \frac{n^2}{T}\mathrm{var}\left(y_1 \mathbf{1}\{I_1 \in \pi \setminus \pi'\}\right) + \frac{n^2}{T}\mathrm{var}\left(y_1 \mathbf{1}\{I_1 \in \pi \setminus \pi'\}\right) \\
&\quad = \frac{n^2}{T}\left(\mathbb{E}[y_1^2 \mathbf{1}\{I_1 \in \pi \setminus \pi'\}] - \left(\frac{1}{n}\sum_{i \in \pi\setminus\pi'} \mu_i\right)^2 + \mathbb{E}[y_1^2 \mathbf{1}\{I_1 \in \pi' \setminus \pi\}] - \left(\frac{1}{n}\sum_{i \in \pi'\setminus\pi} \mu_i\right)^2\right) \\
&\quad \le \frac{n^2}{T}\left(\frac{1}{n}\sum_{i\in\pi\setminus\pi'}(\sigma_i^2 + \mu_i^2) + \frac{1}{n}\sum_{i\in\pi'\setminus\pi}(\sigma_i^2+\mu_i^2)\right) \\
&\quad \le \frac{4|\pi\Delta\pi'|n}{T}
\end{aligned}$$

Let $b_\pi = \max\{|B_2(|\pi\Delta\pi'|, \pi)|, |B_2(|\pi\Delta\pi'|, \pi')|)\}$

$$\begin{aligned}
\mathbb{P}&\left(\exists \pi \in \Pi : |\widehat{\mu}_{\pi'\setminus\pi} - \widehat{\mu}_{\pi\setminus\pi'} - \mu_{\pi'\setminus\pi} - \mu_{\pi\setminus\pi'}| > \sqrt{\frac{8|\pi\Delta\pi'|\log(nb_\pi/\delta)}{T}} + \frac{2n\log(b_\pi/\delta)}{3T}\right) \\
&\le \sum_{\pi\in\Pi}\frac{\delta}{nb_\pi} \\
&\le \sum_{k=1}^{n}\sum_{\pi\in\Pi}\mathbf{1}\{|\pi\Delta\pi'| = k\}\frac{\delta}{nb_\pi} \\
&= \sum_{k=1}^{n}\sum_{\pi\in\Pi}\mathbf{1}\{|\pi\Delta\pi'| = k\}\frac{\delta}{n\max\{|B_2(|\pi\Delta\pi'|,\pi)|, |B_2(|\pi\Delta\pi'|,\pi')|\}} \\
&\le \sum_{k=1}^{n}\sum_{\pi\in\Pi}\mathbf{1}\{|\pi\Delta\pi'| = k\}\frac{\delta}{n|B_2(|\pi\Delta\pi'|,\pi')|} \\
&\le \sum_{k=1}^{n}\frac{\delta}{n} \le \delta
\end{aligned}$$

Now by the Sauer-Shelah Lemma for any k

$$\log(B_1(k)) \le V(B_1(k))\log(en/V(B_1(k))).$$

where $V(\cdot)$ denotes the VC-dimension. At the same time, $|B_1(k)| \le |\{\pi \in \Pi : |\pi| = k\}| \le n^k$. Hence

$$\begin{aligned}
\log(n|B_1(k)|/\delta) &\le \min\{V(B_1(k))\log(en/V(B_1(k))) + \log(n/\delta), (k+1)\log(n/\delta)\} \\
&\le 4\min\{V(B_1(k)), k\}\log(en/\delta)
\end{aligned}$$

Similarly for any k,

$$\log(B_2(k,\pi')) \le V(B_2(k,\pi'))\log(en/V(B_2(k,\pi')))$$

and $|\{\pi \in \Pi : |\pi\Delta\pi^*| = k\}| = \binom{n}{k} \le n^k$. In particular,

$$\begin{aligned}
\log(n|B_2(k,\pi')|/\delta) &\le \min\{V(B_2(k,\pi'))\log(en/V(B_2(k,\pi'))) + \log(n/\delta), (k+1)\log(n/\delta)\} \\
&\le 4\min\{V(B_2(k,\pi')), k\}\log(en/\delta)
\end{aligned}$$

So using identical logic

$$\begin{aligned}
\log(nb_\pi/\delta) &\le \log(n\max\{|B_2(|\pi\Delta\pi'|,\pi)|, |B_2(|\pi\Delta\pi'|,\pi')|)\}/\delta) \\
&\le \max\{\log(n|B_2(|\pi\Delta\pi'|,\pi)|/\delta), \log(n|B_2(|\pi\Delta\pi'|,\pi')|)/\delta)\} \\
&\le 4\min\{\max\{V(B_2(|\pi\Delta\pi'|,\pi)), V(B_2(|\pi\Delta\pi'|,\pi'))\}, |\pi\Delta\pi'|\}\log(en/\delta)
\end{aligned}$$

Finally, in the case of without replacement, we can use the confidence intervals from Theorem 3.6 of [3] and the result follows. □

A.3 Comparison to the Disagreement Coefficient

One of the foundational works on active learning is the DHM algorithm of [19] and the A^2 algorithm that preceded it [2]. In their setting a set of points, $x_1, x_2, \cdots$ are streamed to a learner who chooses whether to label a point or not. Similar in spirit to our algorithm, DHM determines whether it is certain or not about how π^* would label the current point, and if not, would request the label. Thus, DHM only requests the labels of any point that it is uncertain about given all the information up to that time. A key quantity arising in the sample complexity of DHM (and many previous works on active classification) has been that of the disagreement coefficient of the set π^*: $\theta = \theta(\epsilon, \pi^*) := \sup_{r \geq n(\epsilon+\nu)} \left\{ \frac{|x : x \in \pi\Delta\pi^*, \pi \in \Pi \text{ and } |\pi\Delta\pi^*| \leq r|}{r} \right\}$ where $\nu = \mathbb{P}(\pi^*(x) \neq y)$ and ϵ is a bound on the excess error of the set $\widehat{\pi}$ returned by an active learning algorithm. After being streamed m points, DHM returns a classifier with error at most $O(\nu + V(\Pi)\log(m/\delta)/m + \sqrt{V(\Pi)\nu\log(m/\delta)/m})$ after labeling $O\left(\theta\left(\nu m + V(\Pi)\log^2(m) + \log\left(\frac{\log(m)}{\delta}\right)\right)\right)$ samples (provided $\epsilon \leq \nu$–the realistic setting in the non-realizable noisy case). Ignoring log factors, this roughly says that a classifier with error at most $\nu + \epsilon$ is returned after $\theta V(\Pi)\nu \max\{\epsilon^{-1}, \nu\epsilon^{-2}\}$ requested labels.

In general the analysis of the DHM algorithm can not characterize the contribution of each arm to the overall sample complexity leading to sub-optimal sample complexity for combinatorial classes. Consider the case when $\Pi = \{\pi_i\}_{i=1}^n$, with $\pi_i = \{i\}$, and $\pi^* = \{i^*\}$ where $i^* = \text{argmax}_{i \leq n} \mu_i$. If we take $\mu_i \in [-1/2, 1/2]$ for all i then $\frac{1}{4} - \frac{1}{2n} \leq \nu \leq \frac{3}{4} + \frac{1}{2n}$ and for best-arm we necessarily have $\epsilon = \min_{j \neq i^*} \frac{1}{n}(\mu_{i^*} - \mu_j)$. One can show for this problem $\theta = \frac{1}{\nu+\epsilon}$ and so the bound of Theorem 1 of [19] scales like $\theta d\nu \max\{\epsilon^{-1}, \nu\epsilon^{-2}\} = \frac{\nu}{\nu+\epsilon}\max\{\epsilon^{-1}, \nu\epsilon^{-2}\} \approx \epsilon^{-2} = n^2 \max_{i \neq i_*} \Delta_i^{-2}$ for $\Delta_i = \mu_{i^*} - \mu_i$, which is substantially worse than our bound for this problem which scales like $\sum_{i \neq i_*} \Delta_i^{-2}$, describing the contribution from each individual item. Similar arguments can be made for other combinatorial classes such as all subsets of size k. We emphasize that it is not that we are particularly interested in applying algorithms like DHM to this specific problem, but that it exposes such a gross inconsistency with the best known algorithms that its application in general should be questioned.

A.4 Proof of Theorem 1

Since Active Classification is a specific case of the more general combinatorial bandit problem as described in A.1, we focus on the more general case throughout the following. Algorithm 1 is repeated in this more general case below - all that changes are the reward distributions are more general than just Bernoulli distributions.

Proof: Throughout the following, let $\Delta_\pi := \mu_{\pi^* \setminus \pi} - \mu_{\pi \setminus \pi^*}$. Define

$$\mathcal{E} = \bigcap_{k \in \mathbb{N}} \bigcap_{\pi \in \Pi} \{|(\widehat{\mu}_{\pi_*,k} - \widehat{\mu}_{\pi,k}) - (\mu_{\pi_*} - \mu_\pi)| \leq C(\pi_*, \pi, t_k, \delta_k)\}$$

where we recall $C(\pi_*, \pi, t_k, \delta_k) = C(\pi, \pi_*, t_k, \delta_k)$. By Lemma 1 we have that $\mathbb{P}(\mathcal{E}) \geq 1 - \sum_{k=1}^{\infty} \delta_k \geq 1 - \delta$ so assume $\mathcal{E}$ holds in what follows.

First we show $\pi_* \in \mathcal{A}_k$ for all k. Assume $\pi_* \in \mathcal{A}_k$. Then for any $\widehat{\pi} \in \mathcal{A}_k$ we have

$$\begin{aligned}\widehat{\mu}_{\widehat{\pi} \setminus \pi_*, k} - \widehat{\mu}_{\pi_* \setminus \widehat{\pi}, k} &\overset{\mathcal{E}}{\leq} \mu_{\widehat{\pi} \setminus \pi_*} - \mu_{\pi_* \setminus \widehat{\pi}} + C(\widehat{\pi}, \pi_*, t_k, \delta_k) \\ &\leq C(\widehat{\pi}, \pi_*, t_k, \delta_k)\end{aligned}$$

which implies that $\pi_* \in \mathcal{A}_{k+1}$. The result follows by the fact that $\pi_* \in \mathcal{A}_0$.

Now we bound the number of samples taken with high probability. For an arm i to be sampled at time t, there must be at least two policies $\pi, \pi' \in \mathcal{A}_t$ such that $i \in \pi\Delta\pi'$. Since we just showed that $\pi_* \in \mathcal{A}_t$ for all t, it follows that $\min\left\{k : \widehat{\mu}_{\pi_* \setminus \pi, k} - \widehat{\mu}_{\pi \setminus \pi_*, k} > C(\pi_*, \pi, t_k, \delta_k)\right\}$ is an upper bound on the number of rounds before π is removed from Π_t. Since $\mu_{\pi_*} > \mu_\pi$ for all $\pi \in \Pi$, for each $\pi \in \Pi$ there exists a random first round K_π when

$$\widehat{\mu}_{\pi_* \setminus \pi, K_\pi} - \widehat{\mu}_{\pi \setminus \pi_*, K_\pi} \geq C(\pi_*, \pi, t_{K_\pi}, \delta_{K_\pi}).$$

Input: δ, Confidence bound $C(\pi', \pi, t, \delta)$.
Let $\mathcal{A}_1 = \Pi, T_1 = (\cup_{\pi\in\mathcal{A}_1}\pi) - (\cap_{\pi\in\mathcal{A}_1}\pi)$, $k = 1$, $\mathcal{A}_k$ will be the active sets in round k
for $t = 1, 2, \cdots$
  if $t == 2^k$**:**
    Set $\delta_k = .5\delta/k^2$. Let $t_k = 2^k$. For each π, π' let
    $\widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} = \frac{n}{t}(\sum_{s=1}^{t} R_{I_s,s}\mathbf{1}\{I_s \in \pi' \setminus \pi\} - \sum_{s=1}^{t} R_{I_s,s}\mathbf{1}\{I_s \in \pi \setminus \pi'\})$
    Set $\mathcal{A}_{k+1} = \mathcal{A}_k - \big\{\pi \in \mathcal{A}_k : \exists \pi' \in \mathcal{A}_k \text{with } \widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} > C(\pi', \pi, t_k, \delta_k)\big\}$.
    Set $T_{k+1} = \big(\cup_{\pi\in\mathcal{A}_{k+1}}\pi\big) - \big(\cap_{\pi\in\mathcal{A}_{k+1}}\pi\big)$.
    $k \leftarrow k + 1$
  endif
  Stochastic Noise:
    If $T_k = \emptyset$, **Break**. Otherwise, draw I_t uniformly at random from $[n]$ and if $I_t \in T_k$ receive an associated reward $R_{I_t,t} \overset{iid}{\sim} \mu_{I_t}$.
  Persistent Noise:
    If $T_k = \emptyset$ or $t > n$, **Break**. Otherwise, draw I_t uniformly at random from $[n] \setminus \{I_s : 1 \le s < t\}$ and if $I_t \in T_k$ receive associated reward $R_{I_t,t} = \mu_{I_t}$.
Output: $\pi' \in \mathcal{A}_k$ such that $\widehat{\mu}_{\pi',k} - \widehat{\mu}_{\pi,k} \ge 0$ for all $\pi \in \mathcal{A}_k \setminus \pi'$

Algorithm 3: Action Elimination for Combinatorial Bandits

But for every $\pi \in \Pi$ and $k \in \mathbb{N}$ we have

$$\widehat{\mu}_{\pi_*\setminus\pi,k} - \widehat{\mu}_{\pi\setminus\pi_*,k} \overset{\mathcal{E}}{\ge} \Delta_\pi - C(\pi_*, \pi, t_k, \delta_k)$$

so define

$$k_\pi := \min\{k : \Delta_\pi/2 \ge C(\pi_*, \pi, t_k, \delta_k)\}.$$

Also define $k_{\max} = \max_\pi k_\pi$ and note that $k_{\max}$ is finite and deterministic since $C(\pi_*, \pi, t_k, \delta_k)$ is decreasing in k. Now we have that

$$\begin{aligned} S_k &= \{i \in [n] : \exists \pi \in \Pi : i \in \pi_*\Delta\pi, K_\pi \ge k\} \\ &\overset{\mathcal{E}}{\subseteq} \{i \in [n] : \exists \pi \in \Pi : i \in \pi_*\Delta\pi, k_\pi \ge k\} \\ &=: s_k \end{aligned}$$

Thus, we trivially have $\mathbf{1}\{I_s \in S_k\} \le \mathbf{1}\{I_s \in s_k\}$ where the right hand side is a deterministic function. Furthermore, whether or not I_s are drawn uniformly at random from $[n]$ (with replacement) or uniformly at random from $[n] \setminus \{i : I_s = i, 1 \le s < t\}$ (without replacement for persistent noise), the I_s indices are negatively associated random variables [20]. Consequently, standard multiplicative Chernoff bounds apply:

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{k_{\max}} \sum_{s=t_{k-1}+1}^{t_k} \mathbf{1}\{I_s \in S_k\} \ge (1+r)\sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}\right) \\ &\le \mathbb{P}\left(\sum_{k=1}^{k_{\max}} \sum_{s=t_{k-1}+1}^{t_k} \mathbf{1}\{I_s \in s_k\} \ge (1+r)\sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}\right) \\ &\le \exp\left(-\tfrac{\min\{r,r^2\}}{3} \sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}\right) \end{aligned}$$

Taking $r = \max\left\{\frac{3\log(1/\delta)}{\sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}}, \sqrt{\frac{3\log(1/\delta)}{\sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}}}\right\}$ we have with probability at least $1-\delta$ that

$$\sum_{k=1}^{k_{\max}} \sum_{s=t_{k-1}+1}^{t_k} \mathbf{1}\{I_s \in S_k\} \leq \max\left\{3\log(1/\delta), \sqrt{3\log(1/\delta)\sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}}\right\} + \sum_{k=1}^{k_{\max}} t_k \frac{|s_k|}{n}$$
$$\leq \frac{9}{2}\log(1/\delta) + \frac{3}{2}\sum_{k=1}^{\infty} t_k \frac{|s_k|}{n}$$

where the last inequality follows by the arithmetic-geometric mean inequality. Now

$$\begin{aligned}
\sum_{k=1}^{\infty} t_k \frac{|s_k|}{n} &= \sum_{k=1}^{\infty} t_k \sum_{i=1}^{n} \frac{1}{n}\mathbf{1}\{\exists \pi \in \Pi : i \in \pi_*\Delta\pi, k_\pi \geq k\} \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{n} \frac{t_k}{n}\mathbf{1}\{\exists \pi \in \Pi : i \in \pi_*\Delta\pi, k_\pi \geq k\} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{\infty} \frac{2^k}{n}\mathbf{1}\{\exists \pi \in \Pi : i \in \pi_*\Delta\pi, 2^{k_\pi} \geq 2^k\} \\
&\leq \sum_{i=1}^{n} \max_{\pi\in\Pi: i\in\pi_*\Delta\pi} \frac{2^{k_\pi+1}}{n}
\end{aligned}$$

Now, using the specific confidence interval $C_2(\pi', \pi, t_k, \delta_k)$ from 1

$$\begin{aligned}
2^{k^\pi} &\leq 2\min\{t \in \mathbb{N} : \Delta_\pi/2 < C_2(\pi_*, \pi, t, \delta_{\lceil \log_2 t\rceil})\} \\
&\leq c_1 n V_{\pi,\pi'}\left(\frac{|\pi^*\Delta\pi|}{\Delta_\pi^2} + \frac{1}{\Delta_\pi}\right)\log\left(\frac{n\log\left(\Delta_\pi^{-2}\right)}{\delta}\right) \\
&\leq c_2 n V_{\pi,\pi'} \frac{|\pi^*\Delta\pi|}{\Delta_\pi^2}\log\left(\frac{n\log\left(\Delta_\pi^{-2}\right)}{\delta}\right) \\
&\leq c_2 \frac{nV_{\pi,\pi'}}{|\pi^*\Delta\pi|}\frac{1}{\widetilde{\Delta}_\pi^2}\log\left(\frac{n\log\left(\widetilde{\Delta}_\pi^{-2}\right)}{\delta}\right)
\end{aligned}$$

where the second to last line follows from

$$\frac{|\pi^*\Delta\pi|}{\Delta_\pi^2} + \frac{1}{\Delta_\pi} \leq \frac{1}{\Delta_\pi}\left(\frac{|\pi^*\Delta\pi|}{\Delta_\pi} + 1\right) \leq \frac{2|\pi^*\Delta\pi|}{\Delta_\pi^2}$$

since $\Delta_\pi \leq |\pi_*\Delta\pi|$. But for the persistent noise case we have $k_\pi \leq \log_2(n)$ which implies for any i, $\max_{\pi\in\Pi:i\in\pi_*\Delta\pi} \frac{2^{k_\pi+1}}{n} \leq 2$. The result now follows. □

B Proof of Theorem 2

Proof: Step 1: Correctness Let $t_k = 2^k$. Let $\mathcal{E}$ be the event that, for each k and for each $\pi \in \Pi$,

$$\left|\widehat{FDR}(\pi) - FDR(\pi)\right| < C_1(\pi_t, n, t_k, \delta_k)/|\pi|$$

and

$$|(\widehat{TP}(\pi^* \setminus \pi) - \widehat{TP}(\pi \setminus \pi^*)) - (TP(\pi^* \setminus \pi) - TP(\pi \setminus \pi^*))| \leq C_2(\pi^*, \pi, t_k, \delta_k).$$

By Lemma 1 and a union bound,

$$\mathbb{P}(\mathcal{E}^c) \leq \sum_{k\geq 1} 2\frac{2\delta}{8k^2} \leq \delta$$

First we argue that π^* is never eliminated on event $\mathcal{E}$. Note that since $FDR(\pi^*) < \alpha$

$$\begin{aligned}\widehat{FDR}(\pi^*) - \alpha &\overset{\mathcal{E}}{\leq} FDR(\pi^*) - \alpha + C_1(\pi, t_k, \delta_k)/|\pi| \\ &< C_1(\pi, t_k, \delta_k)/|\pi|.\end{aligned}$$

Also for any $\pi \in \Pi_\alpha$,

$$\begin{aligned}\widehat{TP}(\pi \setminus \pi^*) - \widehat{TP}(\pi^* \setminus \pi) &\overset{\mathcal{E}}{\leq} TP(\pi \setminus \pi^*) - TP(\pi^* \setminus \pi) + C_2(\pi, \pi^*, t_k, \delta_k) \\ &= TP(\pi) - TP(\pi_*) + C_2(\pi, \pi^*, t_k, \delta_k) \\ &\leq C_2(\pi, \pi^*, t_k, \delta_k),\end{aligned}$$

and by definition π^* is the maximal TP set in Π_α so π^* will never be removed by another π.

Finally note that on event $\mathcal{E}$, any π' (not just π_*) can knock out π using line 2 or 3 of the algorithm iff $TP(\pi') > TP(\pi)$ and $\pi' \in \Pi_\alpha$.

We define a few key random rounds

$$\begin{aligned}&K_\pi := \max\{k : \pi \in \mathcal{A}_k\} \\ &K_\pi^{FDR,1} := \max\{k : \pi \in \mathcal{A}_k \setminus \mathcal{C}_k\} \\ &K_\pi^{FDR,2} := \min\{k : |\widehat{FDR}(\pi) - \alpha| > C_1(\pi, t_k, \delta_k)\} \\ &K_\pi^{TP} := \min\{k : \exists \pi' \in \mathcal{C}_k \text{ such that } \widehat{TP}(\pi' \setminus \pi) - \widehat{TP}(\pi \setminus \pi') > C_2(\pi', \pi, t_k, \delta_k)\} \\ &K_\pi^{<} := \min\{k : \exists \pi' \in \mathcal{C}_k \text{ with } \pi \subset \pi'\}\end{aligned}$$

Our objective is to bound $\max_{\pi \in \Pi \setminus \pi_*} K_\pi$, which marks the termination of the algorithm.

Bound on $K_\pi^{FDR,1}$: We begin by establishing a deterministic bound on $K_\pi^{FDR,1}$ that holds when event $\mathcal{E}$ is true. Note that $K_\pi^{FDR,1}$ is immediately before the first k such that $\pi \notin \mathcal{A}_k \setminus \mathcal{C}_k$. There are three ways this can occur: i) if π becomes FDR-controlled or if π is determined to not be FDR-controlled, and ii) a $\pi' \in C_k$ knocks out π using statistics about TP (i.e., line 2 of the algorithm), or iii) a $\pi' \in C_k$ knocks out π deterministically by line 3 of the algorithm. These cases are reflected with the $\min$ respectively:

$$K_\pi^{FDR,1} = \min\{K_\pi^{FDR,2}, K_\pi^{TP}, K_\pi^{<}\}.$$

We provide a bound for each one of these terms under $\mathcal{E}$.

- Since $C_1(\pi, t_k, \delta_k)$ is a decreasing function of k, note that

  $$|FDR(\pi) - \alpha| > 2C_1(\pi, t_k, \delta_k)/|\pi| \quad \Longrightarrow \quad |\widehat{FDR}(\pi) - \alpha| > C_1(\pi, t_k, \delta_k)/|\pi|$$

  so on event $\mathcal{E}$, $K_\pi^{FDR,2} < k_\pi^{FDR,2}$ where

  $$k_\pi^{FDR,2} := \min\{k : \Delta_{\pi,\alpha}/2 > C_1(\pi, t_k, \delta_k)/|\pi|\}.$$

- On event $\mathcal{E}$, only sets from Π_α will enter $\mathcal{C}_k$, so only they can be used to knock out other sets in Line 2 of the algorithm. Since π^* is never eliminated on event $\mathcal{E}$, we have that:

  $$K_\pi^{TP} \overset{\mathcal{E}}{\leq} \min\{k : \pi^* \in \mathcal{C}_k \text{ and } \widehat{TP}(\pi^* \setminus \pi) - \widehat{TP}(\pi \setminus \pi^*) > C_2(\pi^*, \pi, t_k, \delta_k)\}.$$

  Thus denoting $\Delta_\pi = TP(\pi^* \setminus \pi) - TP(\pi \setminus \pi^*)$let

  $$k_\pi^{TP} := \min\{k : \Delta_\pi/2 > C_2(\pi^*, \pi, t_k, \delta_k) \text{ and } \Delta_{\pi^*,\alpha}/2 > C_1(\pi^*, t_k, \delta_k)/|\pi^*|\}$$

  and note that $K_\pi^{TP} \overset{\mathcal{E}}{\leq} k_\pi^{TP}$(note that this is potentially infinite if $TP(\pi) > TP(\pi_*)$).

- Using similar logic, on event $\mathcal{E}$ a set π' will knock out a set π using Line 3 of the algorithm only if π' is in $\mathcal{C}_k \cup R$ and $\pi \subset \pi'$. If $\pi' \in \mathcal{C}_k$ then $TP(\pi') \geq TP(\pi)$ so we can remove π. If $\pi' \in R$ but $\pi' \notin \mathcal{C}_k$ yet, there exists a $\pi'' \in \mathcal{C}_k$ (in particular, the π'' that eliminated π' into R) with $TP(\pi'') > TP(\pi') > TP(\pi)$ so we can safely remove π. Either way this implies that the $K_\pi^{<}$ is bounded by the time it takes to guarantee that π' is FDR-controlled, hence

  $$K_\pi^{<} \overset{\mathcal{E}}{\leq} \min_{\substack{\pi' \in \Pi_\alpha \\ \pi \subset \pi'}} K_{\pi'}^{FDR,2} \overset{\mathcal{E}}{\leq} \min_{\substack{\pi' \in \Pi_\alpha \\ \pi \subset \pi'}} k_{\pi'}^{FDR,2}.$$

Putting all of this together we set

$$k_\pi^{FDR,1} := \min\{k_\pi^{FDR,2}, k_\pi^{TP}, \min_{\substack{\pi\in\Pi_\alpha \\ \pi\subset\pi'}} k_{\pi'}^{FDR,2}\} \tag{4}$$

This is necessarily finite since $k_\pi^{FDR,2}$ is finite.

Summarizing:, on event $\mathcal{E}$, $k_\pi^{FDR,1}$ is an upper bound on $K_\pi^{FDR,1}$, the minimal round where $\pi \notin \mathcal{A}_{k+1} \setminus \mathcal{C}_{k+1}$.

Part 2 Bound on K_π**:** If $\pi \in \Pi_\alpha$, on event $\mathcal{E}$, π will be removed from $\mathcal{A}_k$ only when it demonstrably has lower TP than some other set $\pi' \in \Pi_\alpha$ regardless of whether it is in $\mathcal{C}_k$ or not. If $\pi \notin \Pi_\alpha$, on event $\mathcal{E}$, $K_\pi^{FDR,1} = K_\pi$, since the moment it's FDR is confirmed to be greater than α it is removed. Hence using the exact same logic as above, we have $K_\pi \overset{\mathcal{E}}{\leq} k_\pi$ where

$$k_\pi := \begin{cases} \min\{k_\pi^{TP}, \min_{\substack{\pi'\in\Pi_\alpha \\ \pi\subset\pi'}} k_{\pi'}^{FDR,2}\} & \pi\in\Pi_\alpha \\ k_\pi^{FDR,1} & \pi\notin\Pi_\alpha \end{cases} \tag{5}$$

Summarizing: On event $\mathcal{E}$, k_π is an upper bound on K_π and thus the algorithm terminates at some random round $K \leq k_{\max} := \max_{\pi\in\Pi\setminus\pi_*} k_\pi$ and outputs π_*.

Part 3: Bound the contribution of each arm. By the last step, we clearly have that the total sample complexity is bounded by

$$\sum_{k=1}^{k_{\max}} \sum_{t=t_{k-1}+1}^{t_k} \mathbf{1}\{I_t \in S_k\} + \mathbf{1}\{J_t \in T_k\}.$$

Since I_t, J_t are uniformly distributed over $[n]$, we have $\mathbb{E}[\mathbf{1}\{I_t \in S_k\}|S_k] = \frac{|S_k|}{n}$ and $\mathbb{E}[\mathbf{1}\{J_t \in T_k\}|T_k] = \frac{|T_k|}{n}$. However, because $|S_k|$ and $|T_k|$ are random variables, we will upper bound them by deterministic quantities, and then show that the sample complexity concentrates.

For each $i \in [n]$, in round k, note that arm $i \in S_k$ if there is a set $\pi \in \mathcal{A}_k \setminus \mathcal{C}_k$ with $i \in \pi$. Hence

$$S_k = \{i \in [n] : \exists \pi \in \Pi : K_\pi^{FDR,1} > k\} \overset{\mathcal{E}}{\subset} \{i \in [n] : \exists \pi \in \Pi : k_\pi^{FDR,1} > k\} =: \psi_k$$

Similarly, $i \in T_k$ if there is $\pi, \pi' \in \mathcal{A}_k$ with $i \in \pi\Delta\pi'$. On event $\mathcal{E}$, $\pi^* \in \mathcal{A}_k$ for all k, thus $i \in T_k$ iff $i \in \pi\Delta\pi^*$ for some $\pi \in \mathcal{A}_k$. Thus

$$T_k = \{\pi \in \Pi : i \in \pi\Delta\pi^*, K_\pi > k\} \overset{\mathcal{E}}{\subset} \{\exists \pi \in \Pi : i \in \pi\Delta\pi^*, k_\pi > k\} =: \tau_k$$

We now follow an argument similar to that in the proof of Theorem 1. Thus $\mathbf{1}\{I_t \in S_k\} \leq \mathbf{1}\{I_t \in \psi_k\}$ and $\mathbf{1}\{J_t \in T_k\} \leq \mathbf{1}\{J_t \in \tau_k\}$ regardless of whether I_t, J_t are drawn uniformly at random from $[n]$ or uniformly at random from $[n]\setminus\{i : I_s = i, 1 \leq s \leq t\}$ respectively $[n]\setminus\{i : J_s = i, 1 \leq s \leq t\}$. In particular, I_t, J_t are negatively associated so we can apply standard multiplicative Chernoff Bounds. In particular,

$$\begin{aligned}
\mathbb{P}\Big(\sum_{k=1}^{k_{\max}} \sum_{t=t_{k-1}+1}^{t_k} \mathbf{1}\{I_t \in S_k\} \geq (1+r) \sum_{k=1}^{k_{\max}} t_k \frac{|\psi_k|}{n}\Big) \\
&\leq \mathbb{P}\left(\sum_{k=1}^{k_{\max}} \sum_{t=t_{k-1}+1}^{t_k} \mathbf{1}\{I_t \in \psi_k\} \geq (1+r) \sum_{k=1}^{k_{\max}} t_k \frac{|\psi_k|}{n}\right) \\
&\leq \exp\left(-\frac{\min\{r, r^2\}}{3} \sum_{k=1}^{k_{\max}} t_k \frac{|\psi_k|}{n}\right)
\end{aligned}$$

with the appropriate choice of r, with probability greater than $1-\delta$,

$$\sum_{k=1}^{k_{\max}} \sum_{t=t_{k-1}+1}^{t_k} \mathbf{1}\{I_t \in S_k\} \leq \frac{9}{2}\log(2/\delta) + \frac{3}{2}\sum_{k=1}^{\infty} t_k \frac{|\psi_k|}{n}$$

An identical argument gives that with probability greater than $1-\delta$,

$$\sum_{k=1}^{k_{\max}} \sum_{t=t_{k-1}+1}^{t_k} \mathbf{1}\{J_t \in T_k\} \leq \frac{9}{2}\log(2/\delta) + \frac{3}{2}\sum_{k=1}^{\infty} t_k \frac{|\tau_k|}{n}.$$

While we have provided a bound on the sample complexity in terms of deterministic quantities ψ_k and τ_k, we now want to provide natural and interpretable upper bounds on these quantities for a final result.

Putting it all together we have that

$$\begin{aligned}
\sum_{k=1}^{\infty} t_k \frac{\psi_k + \tau_k}{n} &= \sum_{k=1}^{\infty} \frac{2^k}{n}(\psi_k + \tau_k) \\
&= \sum_{i=1}^{n}\sum_{k=1}^{\infty} \frac{2^k}{n}(\mathbf{1}\{\exists \pi \in \Pi : i \in \pi, k_\pi^{FDR,1} > k\} \\
&\qquad\qquad + \mathbf{1}\{\exists \pi \in \Pi : i \in \pi\Delta\pi^*, k_\pi > k\}) \\
&\leq \sum_{i=1}^{n}\sum_{k=1}^{\infty} \frac{2^k}{n}(\mathbf{1}\{\exists \pi \in \Pi : i \in \pi, k_\pi^{FDR,1} > k\} \\
&\qquad\qquad + \mathbf{1}\{\exists \pi \in \Pi, \pi \in \Pi_\alpha : i \in \pi\Delta\pi^*, k_\pi > k\} \\
&\qquad\qquad + \mathbf{1}\{\exists \pi \in \Pi, \pi \notin \Pi_\alpha : i \in \pi\Delta\pi^*, k_\pi > k\}) \\
&\leq \sum_{i=1}^{n} \max_{i\in\pi} \frac{2^{k_\pi^{FDR,1}+1}}{n} + \max_{\substack{\pi\notin\Pi_\alpha \\ i\in\pi\Delta\pi^*}} \frac{2^{k_\pi^{FDR,1}+1}}{n} + \max_{\substack{\pi\in\Pi_\alpha \\ i\in\pi\Delta\pi^*}} \frac{2^{k_\pi+1}}{n} \\
&\leq \sum_{i=1}^{n} 2\max_{i\in\pi} \frac{2^{k_\pi^{FDR,1}+1}}{n} + \max_{\substack{\pi\in\Pi_\alpha \\ i\in\pi\Delta\pi^*}} \frac{2^{k_\pi+1}}{n}
\end{aligned}$$

The fourth line follows from Equation (5) and the last line follows from upper bounding the second term in the fourth line by the first. Solving for k, shows that for some constant c_1

$$\begin{aligned}
2^{k_\pi^{FDR,2}} &\leq \min\left\{m : 2C(\pi, n, m, \delta_{\lfloor \log_2(m)\rfloor}) < |FDR(\pi) - \alpha|\right\} \\
&\leq c_1 n V_\pi \frac{\log(n\log(\Delta_{\pi,\alpha}^{-2})}{|\pi|\Delta_{\pi,\alpha}^2}
\end{aligned}$$

An identical argument shows that for arbitrary π, π', there is a constant c_2 such that

$$\begin{aligned}
2^{k_\pi^{TP}} &\leq \max\left\{c_2 n V_{\pi,\pi^*}\left(\frac{|\pi\Delta\pi^*|}{\Delta_\pi^2} + \frac{1}{\Delta_\pi}\right)\log\left(\frac{n\log(\Delta_\pi^{-2})}{\delta}\right), 2^{k_{\pi^*}^{FDR,2}}\right\} \\
&= \max\left\{c_2 \frac{nV_{\pi,\pi^*}}{|\pi\Delta\pi^*|}\frac{1}{\widetilde{\Delta}_\pi^2}\log\left(\frac{n\log(\widetilde{\Delta}_\pi^{-2})}{\delta}\right), 2^{k_{\pi^*}^{FDR,2}}\right\}
\end{aligned}$$

Finally, for the persistent noise case we have $k_\pi, k^{FDR,2} \leq \log_2(n)$ which implies for any i, $\max_{\pi\in\Pi: i\in\pi_*\Delta\pi} \frac{2^{k_\pi+1}}{n} \leq 2$. The theorem now follows.

$\square$

C One-dimensional thresholds

We can get tighter characterizations of Lemma and consequently, better sample complexity guarantees for particular VC classes. In particular, those classes that have sets with substantial overlap like thresholds. In the case of **Thresholds** we have the following improvement that manages to remove the extra $\log(n)$ terms in Lemma 1.

Lemma 2 *Assume that for each $i \in [n]$ there is an associated distribution ν_i with support $[-1, 1]$, mean μ_i and variance $\sigma_i^2 \leq 1$. Assume access to the observations $(y_1, I_1) \cdots, (y_T, I_T)$ where $I_k \sim$ Unif$([n])$ and $y_k \sim \nu_{I_k}$. Let $\widehat{\mu}_t = \frac{1}{T}\sum_{k=1}^T y_k \mathbf{1}\{I_k \leq t\}$. Fix $t' \leq n$. Then with probability greater than $1-\delta$ for any $s \leq n$,*

$$|\widehat{\mu}_s - \widehat{\mu}_{t'} - (\mu_s - \mu_{t'})| \leq \sqrt{\tfrac{2|s-t'|}{nT}\left(43 + 2\sqrt{2}\log(2\log_2^2(4|s-t'|)/3\delta)\right)} + \tfrac{12+\log\left(2\log_2^2(4|s-t'|)/3\delta\right)}{3T}$$

An analogous result can be proven in the persistent noise case of sampling without replacement.

Active Classification for One-dimensional thresholds with Tsybakov Noise - Let $h \in (0, 1]$, $\alpha \geq 0$, $z \in [0, 1]$ for some $i \in [n-1]$ and assume that $X_{i,j} \in \{-1, 1\}$ are Bernoulli with $\mathbb{P}(X_{i,j} = \text{SIGN}(z - i/n)) = \frac{1}{2} + \frac{1}{2}h|z - i/n|^\alpha$ so that $\mu_i = h|z - i/n|^\alpha\text{SIGN}(z - i/n)$. Let $\Pi = \{[k] : k \leq n\}$. In this case, inspecting the dominating term of 1 for $i \in \pi^*$ we have $\arg\max_{\pi\in\Pi:i\in\pi\delta\pi^*} \frac{V_{\pi,\pi^*}}{|\pi\Delta\pi^*|}\frac{1}{\tilde{\Delta}_\pi^2} = [i]$ and takes a value of $\left(\frac{1+\alpha}{h}\right)^2 n^{-1}(z - i/n)^{-2\alpha-1}$. Trivially upper bounding the other terms and summing, the sample complexities can be calculated to be within a constant of

$$\text{if } \alpha = 0,\ \log(n)\log(\log(n)/\delta)/h^2 \qquad \text{if } \alpha > 0 \quad n^{2\alpha}\log(\log(n)/\delta)/h^2$$

These rates match the minimax lower bound rates given in [12] up to $\log\log$ factors. Note that unlike the algorithms given there, our algorithm works in the *agnostic* setting, i.e. it is making no assumptions about whether the Bayes classifier is in the class. In the case of non-adaptive sampling, the sum is replaced with the max times n yielding

$$\text{if } \alpha \geq 0\ n^{2\alpha+1}\log(\log(n)/\delta)/h^2$$

which is substantially worse than adaptive sampling.

We are now ready to prove the theorem.

Proof: Let

$$f_t(I_k, y_k) = \begin{cases} y_k\mathbf{1}\{I_k \in [t', t]\} & t \geq t' \\ -y_k\mathbf{1}\{I_k \in [t, t']\} & t \leq t' \end{cases}$$

In particular, $\widehat{\mu}_t - \widehat{\mu}_{t'} = \frac{1}{T}\sum_{k=1}^T f_t(I_k, y_k)$. Note that the random variables (y_s, I_s), for $s = 1, \cdots, n$ are by definition i.i.d. drawn from a distribution on $[n] \times \{0, 1\}$. Note

$$\mathbb{E}\left[\frac{1}{T}\sum_{k=1}^n f_t(I_k, y_k)\right] = \begin{cases} \frac{1}{n}\sum_{k=t'}^t \eta_i & t \geq t' \\ \frac{1}{n}\sum_{k=t}^{t'} -\eta_i & t \leq t' \end{cases}$$

and (assuming that $t \leq t'$, an identical computation applies when $t \geq t'$)

$$\begin{aligned} \text{var}(f_t) &= \text{var}(y_s\mathbf{1}\{I_s \in [t, t']\}) \\ &\leq \mathbb{E}[y_s^2\mathbf{1}\{I_s \in [t, t']\}] \\ &= \frac{1}{n}\sum_{i=t}^{t'}(\sigma_i^2 + \eta_i^2) \leq \frac{2}{n}|t' - t|. \end{aligned}$$

By Theorem 2.3 in [7], given $\delta > 0$, for each $\{s : s \leq n, |s - t'| \leq \tau\}$ we have that

$$\mathbb{P}\Bigg(\left|\frac{1}{T}\sum_{k=1}^T f_s(I_k, y_k) - \mathbb{E}[f_s]\right| > 2\mathbb{E}\left[\sup_{|s-t'|\leq\tau}\left|\frac{1}{T}\sum_{k=1}^T f_s(I_k, y_k) - \mathbb{E}[f_s]\right|\right] + \sqrt{\frac{2\tau\log(1/\delta)}{nT}} + \frac{7\log(1/\delta)}{3T}\Bigg) \leq \delta$$

To obtain a bound over all time, we now face two major tasks. Firstly, we must apply a peeling argument to the set of t's. Secondly, and perhaps more immediate, we need bounds on the empirical process

$$\mathbb{E}\left[\sup_{|s-t'|\leq\tau}\left|\frac{1}{T}\sum_{k=1}^{T} f_s(I_k,y_k)-\mathbb{E}[f_s]\right|\right]$$

Let's start with the latter. Denote $Z_t = \frac{1}{T}\sum_{k=1}^{T} f_t(I_k,y_k) - \mathbb{E}[\frac{1}{T}\sum_{k=1}^{T} f_t(I_k,y_k)]$. Firstly note that,

$$|(f_s-f_t)(I_k,y_k)| = \begin{cases} y_k\mathbf{1}\{I_k\in[t,s]\} & s>t \\ -y_k\mathbf{1}\{I_k\in[s,t]\} & t>s \end{cases}$$

In particular the computation above shows,

$$\mathrm{var}\left((f_s-f_t)(I_k,y_k)\right)\leq 2\frac{|t-s|}{n}.$$

Hence,

$$\begin{aligned}\mathrm{var}\left(\frac{1}{T}\sum_{k=1}^{T} f_t(I_k,y_k)-\mathbb{E}[f_t]-\left(\frac{1}{T}\sum_{k=1}^{T} f_s(I_k,y_k)-\mathbb{E}[f_s]\right)\right) &= \frac{\mathrm{var}(f_t(I_k,y_k)-f_s(I_k,y_k))}{T}\\ &\leq \frac{2|t-s|}{nT}\end{aligned}$$

In particular, since $|\frac{1}{T}f_t(I_s,y_s)|\leq\frac{1}{T}$, Bernstein's inequality implies,

$$\log(\mathbb{E}[e^{\lambda(Z_t-Z_s)})]\leq\frac{\lambda^2\frac{2|t-s|}{nT}}{2(1-\lambda/3T)}.$$

Let $d^2(t,s)=|\frac{t}{n}-\frac{s}{n}|$. Then, Lemma 13.1 of [6] with $\nu=2/T$ and $c=1/3T$ we have that,

$$\begin{aligned}\mathbb{E}\left[\sup_{|s-t'|\leq\tau}|Z_s|\right] &\leq \frac{12\sqrt{2}}{\sqrt{T}}\int_0^{\sqrt{\tau/n}/2}\sqrt{\log(\tfrac{\sqrt{\tau/n}}{2u})}du+\frac{4}{T}\int_0^{\sqrt{\tau/n}/2}\log(\tfrac{\sqrt{\tau/n}}{2u})du\\ &\leq \frac{12\sqrt{2}}{\sqrt{T}}\int_0^{\infty}\sqrt{\frac{\tau}{n}}v^2e^{-v^2}dv+\frac{4}{T}\int_0^{\infty}\frac{1}{2}\sqrt{\frac{\tau}{n}}ve^{-v}dv\\ &\leq \frac{12\sqrt{\pi}}{\sqrt{T}}\sqrt{\frac{\tau}{n}}+\frac{2}{T}\sqrt{\frac{\tau}{n}}\\ &\leq 12\sqrt{\pi}\sqrt{\frac{\tau}{nT}}+\frac{2}{T}\end{aligned}$$

the third line follows from the second by doing the substitution, $v=\sqrt{\log(\sqrt{\tau/n}/u))}$ and similarly $u=\log(\sqrt{\tau/n}/u))$ on the second integral.

Hence for all $s:|s-t'|\leq\tau$, using the fact that $\sqrt{a}+\sqrt{b}\leq\sqrt{2(a+b)}$

$$\mathbb{P}\left(\left|\tfrac{1}{T}\sum_{k=1}^{T} f_s(I_k,y_k)-\mathbb{E}[f_s]\right|>\sqrt{\tfrac{\tau}{nT}\left(43+2\sqrt{2}\log(\tfrac{1}{\delta})\right)}+\tfrac{12+\log(1/\delta)}{3T}\right)\leq\delta$$

At this point we need to apply a peeling argument. Let $S_r=\{s\leq n: 2^{r-1}\leq|s-t'|\leq 2^r\}$. Note that $r\leq\log_2(2|s-t'|+2)\leq\log_2(4|s-t'|)$. For each $s\in S_r$ simultaneously, since $2^r\leq 2|s-t'|$, with probability greater than $1-\frac{2\delta}{3r^2}$,

$$\begin{aligned}\left|\tfrac{1}{T}\sum_{k=1}^{T} f_s(I_k,y_k)-\mathbb{E}[f_s]\right| &< \sqrt{\frac{2|s-t'|}{nT}\left(43+2\sqrt{2}\log(\tfrac{2r^2}{3\delta})\right)}+\frac{12+\log\left(\frac{2r^2}{3\delta}\right)}{3T}\\ &\leq \sqrt{\frac{2|s-t'|}{nT}\left(43+2\sqrt{2}\log(\tfrac{2\log_2^2(4|s-t'|)}{3\delta})\right)}+\frac{12+\log\left(\frac{2\log_2^2(4|s-t'|)}{3\delta}\right)}{3T}\end{aligned}$$

Now union-bounding over each $r = 1, \cdots, \log_2(n-t')$, we have that

$$\left|\frac{1}{T}\sum_{k=1}^{T} f_s(I_k, y_k) - \mathbb{E}[f_s]\right| \leq \sqrt{\frac{2|s-t'|}{nT}\left(43 + 2\sqrt{2}\log(\frac{2\log_2^2(4|s-t'|)}{3\delta})\right)} + \frac{12+\log\left(\frac{2\log_2^2(4|s-t'|)}{3\delta}\right)}{3T}$$

with probability greater than

$$\sum_{k=1}^{\log_2(n-t')} \frac{2\delta}{3k^2} \leq \sum_{k=1}^{\infty} \frac{2\delta}{3} k^2 \leq \delta$$

$\square$