Active Online Domain Adaptation

Yining Chen [1] **Haipeng Luo** [2] **Tengyu Ma** [1] **Chicheng Zhang** [3]

Abstract

Online machine learning systems need to adapt to domain shifts. Meanwhile, acquiring label at every timestep is expensive. We propose a surprisingly simple algorithm that adaptively balances its regret and its number of label queries in settings where the data streams are from a mixture of hidden domains. For online linear regression with oblivious adversaries, we provide a *tight* tradeoff that depends on the durations and dimensionalities of the hidden domains. Our algorithm can adaptively deal with interleaving spans of inputs from different domains. We also generalize our results to non-linear regression for hypothesis classes with bounded eluder dimension and adaptive adversaries. Experiments on synthetic and realistic datasets demonstrate that our algorithm achieves lower regret than uniform queries and greedy queries with equal labeling budget.

1. Introduction

Domain shift, the difference between training and testing distributions, is a major bottleneck for many machine learning applications (Kouw & Loog, 2018). For example, bio-imaging model learned from one hospital may not transfer to machines in another hospital (Xu & Yang, 2011). Even imperceptible natural distributional shift can cause big drops in accuracy for many existing models (Recht et al., 2018; 2019; Engstrom et al., 2020). Principled algorithms for domain adaptation require strong assumptions on the types of domain shift (Ben-David et al., 2010). On the other hand, learning the shift requires strong generative models (Hoffman et al., 2017).

Online learning is a classical theoretical framework to deal with worst-case domain shift (Hazan, 2019). In online learning, even though the data is assumed to be given adversarially, strong regret bounds are attainable for many problems. So far, practical deployments of fully online learning systems has been somewhat limited, because labels are expensive to obtain; see Strickland (2018) for an example in fake news detection. Even if a system only acquires a small set of labels periodically, a label budget linear in time is still a luxury.

Cesa-Bianchi et al. (2004b) study label-efficient online learning for prediction with expert advice. Their algorithm queries the label of every example with a fixed probability, which, as they show, achieves minimax-optimal regret and query complexity for this problem. However, querying with uniform probability does not take into account the algorithm's uncertainty on each individual example, and thus can be suboptimal when the problem has certain favorable structures. For example, a sequence of online news may come from the mixture of a few topics or trends, and some news topics may require more samples to categorize well compared to others.

We aim to improve label-efficiency in online regression by exploiting hidden domain structures in the data. We assume that each input is from one of m potentially overlapping domains. For each input, the learner makes a prediction, incurs a loss, and decides whether to query its label. The regret of the learner is defined as the difference between its cumulative loss and that of the best predictor in hindsight. We assume realizability, i.e., there exists a predictor that is Bayes optimal across all the domains. This is a reasonable assumption in modern machine learning, since features can be high-dimensional (so that different domain may rely on different features), and models are often overparameterized (Zhang et al., 2016). Our goal is to trade off between regret and query complexity: given a fixed label budget, we hope to achieve a regret as low as possible.

We propose QuFUR (Query in the Face of Uncertainty for Regression), a surprisingly simple query scheme based on uncertainty quantification. We start with online linear regression from $\mathbb{R}^d$ to $\mathbb{R}$ with an oblivious adversary. In the realizable setting, with additional regularity conditions, we provide the following regret guarantee of QuFUR with label budget B: for *any* partition of $[T]$ into domains, $I_1, \ldots, I_m$, if for every u in $[m]$, the u-th domain $S_u = \{x_t : t \in I_u\}$ has T_u examples and lies in a d_u-dimensional subspace of

[1]Stanford University, USA [2]University of Southern California, USA [3]University of Arizona, USA. Correspondence to: Yining Chen <cynnjjs@stanford.edu>.

Accepted at Lifelong ML Workshop @ICML, Vienna, Austria, 2020.
Copyright 2020 by the author(s).

$\mathbb{R}^d$, the regret is $\tilde{O}((\sum_{u=1}^m \sqrt{d_u T_u})^2/B)$ (Theorem 2).[1]

When choosing $m = 1$ and $I_1 = [T]$, we see that the regret of QuFUR is at most $\tilde{O}(dT/B)$, matching minimax lower bounds (Theorem 7) in this setting. The advantage of QuFUR's adaptive regret guarantees becomes significant when the domains have heterogeneous time spans and dimensions. For example, $(\sum_{u=1}^m \sqrt{d_u T_u})^2$ can be substantially less than dT when the T_u/d_u's are heterogeneous across different u's. As an example, if $d = m$, $d_1 = d_2 = \ldots = d_m = 1$, $T_1 = T - m + 1$, $T_2 = \ldots = T_m = 1$, $(\sum_{u=1}^m \sqrt{d_u T_u})^2 = O\left(T + m^2\right)$, which is substantially less than $dT = O\left(mT\right)$ when $1 \ll m \ll T$. Using standard online-to-batch conversion (Cesa-Bianchi et al., 2004a), we also obtain novel results in batch active learning for regression (Theorem 12). Furthermore, we also define a stronger notion of minimax optimality, namely *hidden domain minimax optimality*, and show that QuFUR is optimal in this sense (Theorem 3), for a wide range of domain structure specifications.

We generalize our results to online regression with general hypothesis classes against an adaptive adversary. We obtain a similar regret-query complexity tradeoff, where the analogue of d_u is (roughly) the *eluder dimension* (Russo & Van Roy, 2013) of the hypotheses class with respect to the support of domain u (Theorem 4).

Experimentally, we show that our algorithm outperforms the baselines of uniform and greedy query strategies, in a synthetic experiment and several LIBSVM datasets (Chang & Lin, 2011).

2. Related works

Active learning. We refer the readers to Balcan et al. (2009); Hanneke (2014); Dasgupta et al. (2008); Beygelzimer et al. (2010) and the references therein for background on active learning. For classification, a line of work (Dasgupta & Hsu, 2008; Minsker, 2012; Kpotufe et al., 2015; Locatelli et al., 2017) performs hierarchical sampling for nonparametric active learning. The main idea is to maintain a hierarchical partitioning over the instance domain (either a pre-defined dyadic partition or a pre-clustering over the data), and "zooms" into uncertainty regions whose confidence interval of $P(y = 1|x)$ contains $1/2$. For regression, many works (Fedorov & Hackl, 2012; Chaudhuri et al., 2015) study the utility of active learning for maximum likelihood estimation in the realizable setting. Recent works also study active linear regression in nonrealizable (Drineas et al., 2006; Derezinski et al., 2018; Dereziński & Warmuth, 2018; Sabato & Munos, 2014) and heteroscedastic (Chaudhuri et al., 2017; Fontaine et al., 2019) settings. These works do not consider domain structures except for Sabato & Munos (2014), who propose a domain-aware active sampling scheme. Their algorithm needs to know the domain partition a priori, and its performance depends on the quality of the partition. The empirical works of Rai et al. (2010); Saha et al. (2011); Xiao & Guo (2013) study stream-based active learning when inputs comes from pre-specified source and target domains. Our algorithm handles multiple domains, and does not require knowledge of which domain the inputs come from.

Active online learning. Earlier works on selective sampling when iid data arrives in a stream and a label querying decision has to be made after seeing each example (Cohn et al., 1994; Dasgupta et al., 2008; Hanneke, 2011) implicitly provide online regret and label complexity guarantees. Works on worst-cast analysis of selective sampling for linear classification (Cesa-Bianchi et al., 2006) provide regret guarantees similar to that of popular online linear classification algorithms such as Perceptron and Winnow, but their label complexity guarantees are runtime-dependent and therefore cannot be easily converted to a guarantee that only involves problem parameters defined apriori. Subsequent works (Cesa-Bianchi et al., 2009; Dekel et al., 2010; Cavallanti et al., 2011; Agarwal, 2013) study the setting where there is a parametric model on $P(y|x, \theta)$ with unknown parameter θ, and the x's shown can be adversarial. Under those assumptions, they obtain regret and query complexity guarantees dependent on the fraction of examples with low margins. Yang (2011) gives a worst-case analysis of active online learning for classification with drifting distributions, under the assumption that the Bayes optimal classifier is in the hypothesis class given to the learner. In contrast, our work gives adaptive regret guarantees in terms of the hidden domain structure in the data, and focuses on regression instead of classification.

KWIK model. In the KWIK model (Li et al., 2011), at each time step, the algorithm is asked to either query the label (and output "Don't know") or predict an output with at most ϵ error. In contrast, in our setting, the learner's goal is to minimize its cumulative regret, as opposed to making pointwise-accurate predictions. Cesa-Bianchi et al. (2009) study linear regression in the KWIK-model, and propose an algorithm similar to ours; unlike our work, they do not consider the algorithm's adaptivity to domain structures. Szita & Szepesvári (2011) propose an algorithm that works in an agnostic setting, where the error guarantee at every round depends on the agnosticity of the problem. A relaxed KWIK model that allows a prespecified number of mistakes has been studied in (Sayedi et al., 2010; Zhang & Chaudhuri, 2016).

[1] Throughout this paper, denote by $[n] := \{1, \ldots, n\}$; notations $\tilde{O}$ and $\tilde{\Omega}$ hide logarithmic factors.

Adaptive/Switching Regret. Adaptive regret (Hazan & Seshadhri, 2007; Daniely et al., 2015) is the excessive loss of an online algorithm compared to the locally optimal solution over any continuous timespan. Our algorithm can be interpreted as being competitive with the locally optimal solution on every domain, even if the timespans of the domains are not continuous, which is closer to the concept of switching regret with long-term memory studied in e.g, (Bousquet & Warmuth, 2002; Zheng et al., 2019). However, typical bounds for this regret measure have a polynomial dependence on the number of domain switches, which does not appear in our bounds at all.

Online linear regression. Literature on fully-supervised online linear regression has a long history (Vovk, 2001; Azoury & Warmuth, 2001). As is implicit in Cesa-Bianchi et al. (2004b), we can reduce from fully-supervised online regression to active online regression by querying uniformly randomly with a fixed probability. Combining this reduction with existing online linear regression algorithms (Hazan et al., 2007), we get $\tilde{O}(dT/B)$ regret with $O(B)$ queries for any $B \leq T$. Our bound matches this in the realizable and oblivious setting when there is one domain, and is potentially better with more domain structures.

3. Setup and Preliminaries

3.1. Setup

Active online learning with squared loss Let $\mathcal{F} = \{f : \mathcal{X} \to [-1, 1]\}$ be a hypotheses class. We consider the realizable setting where $y_t = f^*(x_t) + \xi_t$ for some $f^* \in \mathcal{F}$ and random noise ξ_t. The adversary decides $f^* \in \mathcal{F}$ before interaction starts. ξ_t's are independent zero-mean, sub-Gaussian random variables with variance proxy η^2.

The learner is given a label budget B. For each $t = 1, \ldots, T$:

- x_t is revealed to the learner.
- The learned predicts $\hat{y}_t = \hat{f}_t(x_t)$ using predictor $\hat{f}_t \in \mathcal{F}$, incurring loss $(\hat{y}_t - y_t)^2$.
- The learner sets a query indicator $q_t \in \{0, 1\}$. If $q_t = 1$, y_t is revealed.

The performance of the learner is measured by its query complexity $Q = \sum_{t=1}^{T} q_t$ and regret $R = \sum_{t=1}^{T} (\hat{y}_t - f^*(x_t))^2$. By our realizability assumption, our notion of regret coincides with the one usually used in online learning when taking expectations; see Appendix D. Our goal is to design a learner that has low regret R subject to its budget constraint $Q \leq B$.

Oblivious vs. adaptive adversary In the *oblivious* setting, the adversary decides the sequence $\{x_t\}_{t=1}^{T}$ beforehand. In the *adaptive* setting, the adversary can choose x_t depending on $\{x_{1:t-1}, f_{1:t-1}, \xi_{1:t-1}\}$.

Miscellaneous notations For a vector $v \in \mathbb{R}^d$ and a positive semidefinite matrix $M \in \mathbb{R}^{d \times d}$, $\|v\|_M := \sqrt{v^\top M v}$. For vectors $\{z_t\}_{t=1}^{T} \subseteq \mathbb{R}^l$, and $S = \{i_1, \ldots, i_n\} \subseteq [T]$, denote by Z_S the $n \times l$ matrix whose rows are $z_{i_1}^\top, \ldots, z_{i_n}^\top$. Define $\mathrm{clip}(z) := \min(1, \max(-1, z))$ and $\tilde{\eta} := \max\{1, \eta\}$.

3.2. Baselines

We first study linear regression with oblivious adversary, and generalize to non-linear case with adaptive adversary in Section 5. For now, hypothesis class $\mathcal{F} = \{\langle x, \theta \rangle : \theta \in \mathbb{R}^d, \|\theta\|_2 \leq C\}$. Let the ground truth hypothesis be θ^*, and input space $\mathcal{X}$ be a subset of $\{x \in \mathbb{R}^d : \|x\|_2 \leq 1, \langle x, \theta^* \rangle \leq 1\}$. [2]

Uniform query is minimax-optimal with no domain structure As a baseline, consider the algorithm that always queries and returns the regularized least squared estimator $\hat{\theta}_t = \mathrm{argmin}_\theta \sum_{i=1}^{t-1} (\langle \theta, x_i \rangle - y_i)^2 + \lambda \|\theta\|^2$, where $\lambda = 1/C^2$. It is well-known (Vovk, 1990; Azoury & Warmuth, 2001) that this fully-supervised algorithm achieves $R = \tilde{O}(\tilde{\eta}^2 d)$ with $Q = T$. Consider an active learning extension of the above algorithm that queries uniformly randomly with probability μ, and always predicts with the regularized least squared estimator computed based on all queried examples $\hat{\theta}_t = \mathrm{argmin}_\theta \sum_{i \in [t-1], q_i = 1} (\langle \theta, x_i \rangle - y_i)^2 + \lambda \|\theta\|^2$. We show that the above uniform querying strategy achieves $R = \tilde{O}(\tilde{\eta}^2 dT/B)$ with $Q = \tilde{O}(B) = \tilde{O}(\mu T)$ in Appendix A.4. As shown in Theorem 7, this tradeoff is minimax optimal if $\tilde{\eta}$ is a constant. Although this guarantee is optimal in the worst case, one major weakness is that it is too pessimistic: as we will see next, when the data has some hidden domain structure, it is possible to achieve substantially better regret guarantees than the worst-case ones, if the learner has access to auxiliary domain information.

Oracle baseline when domain structure is known Suppose the learner is given the following knowledge from an oracle: there are m domains; for each u in $[m]$, there are a total of T_u examples from domain u from a subspace of $\mathbb{R}^d$ dimension d_u. In addition, for every t, the learner is given the index of the domain example x_t comes from. In this setting, the learner can use the following domain-aware querying scheme: for any example in domain u, it queries its label independently with probability $\mu_u \in (0, 1]$. Within domain u, the learner incurs $O(\mu_u T_u)$ queries and $\tilde{O}(\tilde{\eta}^2 d_u / \mu_u)$ regret.

[2] The constraint $\|x\|_2 \leq 1$ can be relaxed by only increasing the logarithmic terms in the regret and query complexity guarantees.

Summing over domains, its achieves a label complexity of $O(\sum_{u=1}^m \mu_u T_u)$, and a regret of $\tilde{O}(\tilde{\eta}^2 \sum_{u=1}^m d_u/\mu_u)$. This motivates the following optimization problem:

$$\min_{\mu} \sum_{u=1}^m d_u/\mu_u, \text{s.t.} \sum_{u=1}^m \mu_u T_u \leq B, \mu_u \in [0,1], \forall u \in [m].$$

i.e., we choose domain-dependent query probabilities that minimize the learner's regret guarantee, subject to its query complexity being at most B. When $B \leq \sum_{u=1}^m \sqrt{d_u T_u} \cdot \min_u \sqrt{T_u/d_u}$, the optimal μ_u is proportional to $\sqrt{d_u/T_u}$. This yields a regret guarantee of $O(\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2/B)$. Although this strategy can sometimes achieve much smaller regret than uniform query (as we have seen in Section 1, $(\sum_u \sqrt{d_u T_u})^2$ can be substantially smaller than dT), it still has two crucial drawbacks: first, it is not clear if this guarantee is always no worse than uniform querying, especially when $\sum_{u=1}^m d_u \gg d$; second, the domain membership of each example is rarely known in practice. We will develop algorithms that overcome these drawbacks.

4. Active online domain adaptation for linear regression: algorithms, analysis, and matching lower bounds

We start by presenting an algorithm parameterized by α in Section 4.1, which has a natural cost minimization interpretation. We then present a fixed-budget variant of it in Section 4.2. Section 4.3 shows that our algorithm is minimax-optimal under a wide range of domain structure specifications.

4.1. Main Algorithm: Query in the Face of Uncertainty for Regression (QuFUR)

We propose QuFUR (Query in the Face of Uncertainty for Regression), namely Algorithm 1. At each time step t, the algorithm first computes $\hat{\theta}_t$, a regularized empirical risk minimizer on the labeled data obtained so far, then predict using $\hat{f}_t(x) = \text{clip}(\langle \hat{\theta}_t, x \rangle)$. It makes label queries with probability proportional to a high-probability upper bound of the instantaneous regret $(\hat{y}_t - \langle \theta^*, x_t \rangle)^2$, which can also be interpreted as the uncertainty on x_t. Intuitively, when the algorithm is already confident about the current prediction, it will save budget for learning from less certain inputs in the future.

QuFUR measures the uncertainty of x_t using $\Delta_t := \tilde{\eta}^2 \min\left(1, \|x_t\|^2_{M_t^{-1}}\right)$, where $M_t = \lambda I + \sum_{i \in \mathcal{Q}_t} x_i x_i^\top$, and $\mathcal{Q}_t$ is the set of labeled examples seen up to time step $t-1$. We will show in Lemma 1 that with high probability, the instantaneous regret on x_t is at most $\tilde{O}(\Delta_t)$. QuFUR queries the label y_t with probability $\min\{1, \alpha \Delta_t\}$, where α is a parameter that tradeoffs query complexity and regret.

Algorithm 1 Query in the Face of Uncertainty for Regression (QuFUR(α))

Require: Total dimension d, time horizon T, θ^*'s norm bound C, noise level η, parameter α.

$M \leftarrow \frac{1}{C^2} I$, queried dataset $\mathcal{Q} \leftarrow \emptyset$.
for $t = 1$ to T **do**
  Compute regularized least squares solution $\hat{\theta}_t \leftarrow M^{-1} X_{\mathcal{Q}}^\top Y_{\mathcal{Q}}$.
  Let $\hat{f}_t(x) = \text{clip}(\langle \hat{\theta}_t, x \rangle)$ be the predictor at time t, and predict $\hat{y}_t \leftarrow \hat{f}_t(x_t)$.
  Uncertainty estimate $\Delta_t \leftarrow \tilde{\eta}^2 \min\{1, \|x_t\|^2_{M^{-1}}\}$.
  With probability $\min\{1, \alpha \Delta_t\}$, set $q_t \leftarrow 1$; otherwise $q_t \leftarrow 0$.
  if $q_t = 1$ **then**
    Query y_t. $M \leftarrow M + x_t x_t^\top$, $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

Perhaps surprisingly, the simple query strategy of QuFUR can leverage hidden domain structure, as shown by the following theorem.

Theorem 1. *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u\}_{u=1}^m$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ lie in a subspace of dimension d_u. Suppose $\alpha \in \left[1/\tilde{\eta}^2 \left(1/(\sum_u \sqrt{d_u T_u})\right)^2, 1/\tilde{\eta}^2 \min_{u \in [m]} T_u/d_u\right]$. If Algorithm 1 receives inputs dimension d, time horizon T, norm bound C, noise level η, parameter α, then, with probability $1-\delta$:*
1. Its query complexity $Q = \tilde{O}(\tilde{\eta} \cdot \sqrt{\alpha} \sum_u \sqrt{d_u T_u})$.
2. Its regret $R = \tilde{O}(\tilde{\eta} \cdot \sum_u \sqrt{d_u T_u}/\sqrt{\alpha})$.

The proof of the theorem is deferred to Section A.1. We make a series of remarks below:

Novel notion of adaptive regret. The above tradeoff is novel in that it holds for *any* meaningful domain partition. Our proof actually shows that for any (not necessarily contiguous) subsequence $I \subseteq [T]$, QuFUR has $Q = \tilde{O}(\tilde{\eta} \cdot \sqrt{d_I |I|} \cdot \sqrt{\alpha})$ and $R = \tilde{O}(\tilde{\eta}\sqrt{d_I |I|})/\sqrt{\alpha}$ within I, where d_I is the dimension of $\text{span}(\{x_t : t \in I\})$. This type of guarantee is stronger than the adaptive regret guarantees provided by e.g. (Hazan & Seshadhri, 2007), where the regret guarantee are only with respect to continuous intervals.

Matching uniform querying baseline and minimax optimality. Our tradeoff is never worse than the uniform querying baseline; this can be seen by applying the theorem with the trivial partition $\{[T]\}$ yields $Q = \tilde{O}(\tilde{\eta}\sqrt{\alpha d T})$ and $R = \tilde{O}(\tilde{\eta}\sqrt{dT/\alpha})$. Therefore, same as the uniform query baseline, this guarantee is also minimax optimal for constant η, in light of Theorem 7 in Appendix B.2.

Matching oracle baseline. QuFUR matches the domain-aware oracle baseline even without prior knowledge of domain structure. We show in Theorem 3 that in a wide range of problem specifications, this baseline, as well as QuFUR, is minimax-optimal in our problem formulation with domain structure.

Fixed-cost-ratio interpretation. The tradeoff in Theorem 1 can be interpreted in a fixed-cost-ratio formulation. Suppose a practitioner decides that the cost ratio between 1 unit of loss and 1 label query is $c : 1$. The performance of the algorithm is then measured by its total cost $cR + Q$. Theorem 1 shows that QuFUR(α) balances the cost incurred by prediction and the cost incurred by label queries, in that $Q \approx \alpha R$. We show in Appendix C that QuFUR with input $\alpha = c$ achieves near-optimal total cost, for a wide range of domain structure parameters.

Dependence on η. Our query complexity and regret bounds have a dependence on $\tilde{\eta} = \max(\eta, 1)$. Similar dependence also appears in online least-squares regression literature (Vovk, 1990; Azoury & Warmuth, 2001).

Requirements on α. The requirement that $\alpha \geq 1/\tilde{\eta}^2 \left(1/(\sum_u \sqrt{d_u T_u})\right)^2$ is immaterial: if α is smaller than this threshold, the regret guarantee is at least $\tilde{O}(\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2)$; as $\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2 \geq T$, the above guarantee can in fact be trivially achieved by an algorithm that performs no label queries and always predicts $\hat{y}_t = 0$. The requirement that $\alpha \leq 1/\tilde{\eta}^2 \min_{u\in[m]} T_u/d_u$ corresponds to a label usage of order $\tilde{O}\left(\sum_u \sqrt{d_u T_u} \cdot \min_{u\in[m]} \sqrt{T_u/d_u}\right)$, which also matches the label budget range of the oracle baseline discussed in Section 3.2.

Running time. The most computationally intensive operation in QuFUR is the $d \times d$ matrix inversion in line 8, which occurs T times. We can apply the same optimization as in Abbasi-Yadkori et al. (2011), i.e. recompute $\hat{\theta}_t$ only when $\det(M_t)$ increases by a constant factor. Using this, only $O(d \log T)$ matrix inversions are required; the regret of the modified algorithm is of the same order as the original algorithm, up to log factors.

4.2. QuFUR with a fixed label budget

The label complexity bound in Theorem 1 involve parameters $\{d_u, T_u\}_{u=1}^m$, which may be unknown in advance. In many practical settings, the learner is given a label budget B. For such settings, we propose a fixed-budget version of QuFUR, Algorithm 2, that takes B as input, and achieves near-optimal regret bound subject to the budget constraint, under a wide range of domain structure specifications.

Algorithm 2 Fixed-Budget QuFUR

Require: Total dimension d, time horizon T, label budget B, θ^*'s norm bound C, noise level η.
Number of copies $k \leftarrow \lceil 4 \log_2 T \rceil$.
for $i = 0$ to k **do**
  Parameter $\alpha_i \leftarrow 2^i/T$.
Initialize $M \leftarrow \frac{1}{C^2} I$, $\mathcal{Q} \leftarrow \emptyset$.
for $t = 1$ to T **do**
  Compute regularized least squares solution $\hat{\theta}_t \leftarrow M^{-1} X_{\mathcal{Q}}^\top Y_{\mathcal{Q}}$.
  Let $\hat{f}_t(x) = \text{clip}(\langle \hat{\theta}_t, x \rangle)$ be the predictor at time t, and predict $\hat{y}_t \leftarrow \hat{f}_t(x_t)$.
  Uncertainty estimate $\Delta_t \leftarrow \tilde{\eta}^2 \min\{1, \|x_t\|_{M^{-1}}^2\}$.
  for $i = 0$ to k **do**
    if $\sum_{j=1}^{t-1} q_j^i < \lfloor B/k \rfloor$ **then**
      With probability $\min\{1, \alpha_i \Delta_t\}$, set $q_t^i = 1$.
  if $\sum_i q_t^i > 0$ **then**
    Query y_t. $M \leftarrow M + x_t x_t^\top$, $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

Algorithm 2 is a master algorithm that aggregates over $k = O(\log T)$ copies of QuFUR (α). Each copy uses a different value of α lying in an exponentially increasing grid $\{2^i/T^3 : i = 0, \ldots, k\}$. The grid ensures that each copy still has label budget $\lfloor B/k \rfloor = \tilde{\Omega}(B)$, and there is always a copy that takes full advantage of its budget to achieve low regret. The algorithm queries whenever one of the copies issues a query, and predicts using a model learned on all historical labeled data. A copy can no longer query when its budget is exhausted. The regret of the master algorithm is no worse that of the copy running on a parameter α_i that make $\tilde{\Theta}(B)$ queries; this insight yields the following theorem.

Theorem 2. *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u\}_{u=1}^m$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t\in I_u}$ lie in a subspace of dimension d_u. Moreover, integer B satisfies*

$$B \leq \tilde{O}\left(\sum_u \sqrt{d_u T_u} \min_{u\in[m]} \sqrt{T_u/d_u}\right). \tag{1}$$

If Algorithm 2 receives inputs dimension d, time horizon T, label budget B, norm bound C, noise level η, then:
1. Its query complexity Q is at most B.
2. With probability $1 - \delta$, its regret $R = \tilde{O}(\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2/B)$.

The proof of the theorem is deferred to Appendix A.2. We now compare the guarantee of QuFUR with the that of of the oracle baseline in Section 3.2: for any budget $B \in [0, \tilde{O}(\sum_u \sqrt{d_u T_u} \min_u \sqrt{T_u/d_u})]$, Fixed-Budget QuFUR achieves a regret guarantee no worse than that of domain-aware uniform sampling, while being agnostic to $\{d_u, T_u\}_{u=1}^m$ and the domain memberships of the examples.

4.3. Lower bound

Our development so far establishes domain structure-aware regret upper bounds $R = \tilde{O}(\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2/B)$, achieved by Fixed-Budget QuFUR and domain-aware uniform sampling baseline (the latter requires extra knowledge about the domain structure and domain membership of each example, whereas the former does not). In this section, we study optimality properties of the above upper bounds. Specifically, we show via Theorem 3 that they are tight up to logarithmic factors, for a wide range of domain structure specifications. Its proof can be found in Appendix B.1.

Theorem 3. *For any noise level $\eta \geq 1$, set of positive integers $\{(d_u, T_u)\}_{u=1}^m$ and integer B that satisfy*

$$d_u \leq T_u, \forall u \in [m]; \sum_{u=1}^m d_u \leq d; \tag{2}$$

$$B \geq \sum_{u=1}^m \sqrt{d_u T_u} \cdot \sqrt{\max_{u \in [m]} d_u / T_u}, \tag{3}$$

there exists an oblivious adversary such that:
1. It uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \leq \sqrt{d}$, and for all t, $\left|\langle \theta^*, x_t \rangle\right| \leq 1$; in addition, the noises $\{\xi_t\}_{t=1}^T$ are sub-Gaussian with variance proxy η^2.
2. It shows example sequence $\{x_t\}_{t=1}^T$ such that $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u\}_{u=1}^m$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ lie in a subspace of dimension d_u.
3. Any online active learning algorithm $\mathcal{A}$ with label budget B has regret $\Omega((\sum_{u=1}^m \sqrt{d_u T_u})^2/B)$.

The above theorem is a domain structure-aware refinement of the $\Omega(dT/B)$ minimax lower bound (Theorem 7 in Appendix B.2), in that it further constrains the adversary to present sequences of examples with domain structure parametrized by $\{d_u, T_u\}_{u=1}^m$. In fact, $\Omega(dT/B)$ minimax lower bound is a special case of the lower bound of Theorem 3, by taking $m = 1, d_1 = d$, and $T_1 = T$.

To discuss the tightness of the upper and lower bounds we obtained so far in more detail, we first set up some useful notations. Denote by $\mathrm{E}(\{T_u, d_u\}_{u=1}^m)$ the set of oblivious adversaries that shows example sequences $\{x_t\}_{t=1}^T$, such that $[T]$ can be decomposed to m subsets $\{I_u\}_{u=1}^m$, where for every u in $[m]$, $|I_u| = T_u$, and examples in subset $\{x_t : t \in I_u\}$ lie in a subspace of dimension d_u. Additionally, denote by $\mathrm{A}(B)$ the set of online active learning algorithms that uses a label budget of B. Finally, for an algorithm $\mathcal{A}$ and an oblivious adversary $\mathcal{E}$, define $\mathrm{R}(\mathcal{A}, \mathcal{E})$ as the expected regret of $\mathcal{A}$ in the environment induced by $\mathcal{E}$. Consider the regime when the noise variance proxy $\eta \in [1, O(1)]$. In this case, Theorem 3 shows that for all $\{d_u, T_u\}_{u=1}^m$ and B such that 3 holds, we have

$$\min_{\mathcal{A} \in \mathrm{A}(B)} \max_{\mathcal{E} \in \mathrm{E}\left(\{d_u, T_u\}_{u=1}^m\right)} \mathrm{R}\left(\mathcal{A}, \mathcal{E}\right) \geq \Omega\left(\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)^2 / B\right).$$

On the other hand, Theorem 2 shows that for all $\{d_u, T_u\}_{u=1}^m$ and B such that (1) holds, we have

$$\max_{\mathcal{E} \in \mathrm{E}\left(\{d_u, T_u\}_{u=1}^m\right)} \mathrm{R}\left(\text{Fixed-Budget QuFUR}(B), \mathcal{E}\right) \leq \tilde{O}\left(\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)^2 / B\right).$$

This shows that, if $\eta \in [1, O(1)]$, for a wide range of domain structure specifications $\{d_u, T_u\}_{u=1}^m$ and budgets B (i.e., $B/(\sum_u \sqrt{d_u T_u}) \in [\sqrt{\max_u d_u/T_u}, \sqrt{\min_u T_u/d_u}]$), the regret guarantee of Fixed-Budget QuFUR is optimal; furthermore, the algorithm requires no knowlege on the domain structure. We call this property of Fixed-Budget QuFUR its *hidden-domain minimax optimality*.

5. Extension to realizable non-linear regression with adaptive adversary

QuFUR's design principle, querying with probability proportional to uncertainty estimates of unlabeled data, can be easily generalized to deal with other active online learning problems. In summary, QuFUR uses a suitable loss upper bound to yield query-regret tradeoff adaptive to the complexities and durations of individual domains. We now generalize QuFUR to non-linear regression with adaptive adversaries, extending Russo & Van Roy (2013).

In this section, we require the domain partition $\{I_u : u \in [m]\}$ to have a property we call *admissibility*.

Definition 1. *The partition $\{I_u : u \in [m]\}$ is admissible, if the domain membership of the t-th example, $u_t = \{u : t \in I_u\}$ depends on the interaction history up to $t-1$ and unlabeled example x_t; formally, u_t is $\sigma(H_{t-1}, x_t)$-measurable.*

Domain complexity measure. Analogous to the dimension of the support in linear regression, we use $d'_u = \dim_u^E(\mathcal{F}, 1/T_u^2)$, the *eluder dimension* of $\mathcal{F}$ with respect to domain $u \in [m]$ with support $\mathcal{X}_u \subseteq \mathcal{X}$. Formally,

Definition 2. *An input $x \in \mathcal{X}$ is ϵ-dependent of on another set of inputs $\{x_i\}_{i=1}^n \subseteq \mathcal{X}$ with respect to $\mathcal{F}$ if for all $f_1, f_2 \in \mathcal{F}$, $\sqrt{\sum_{i=1}^n (f_1(x_i) - f_2(x_i))^2} \leq \epsilon \implies f_1(x) - f_2(x) \leq \epsilon$.*

Algorithm 3 QuFUR(α) for Nonlinear Regression

Require: Hypothesis set $\mathcal{F}$, time horizon T, parameters α, δ, η.

Labeled dataset $\mathcal{Q} \leftarrow \emptyset$.
for $t = 1$ to T **do**
  Predict $\hat{f}_t \leftarrow \operatorname{argmin}_{f \in \mathcal{F}} \sum_{i \in \mathcal{Q}} (f(x_i) - y_i)^2$.
  Confidence set $\mathcal{F}_t \leftarrow \{f \in \mathcal{F} : \sum_{i \in \mathcal{Q}} (f(x_i) - \hat{f}_t(x_i))^2 \leq \beta_{|\mathcal{Q}|}(\mathcal{F}, \delta)\}$,
  where $\beta_k := 8\eta^2 \log(4\mathcal{N}(\mathcal{F}, 1/T^2, \|\cdot\|_\infty)/\delta) + 2k/T^2(16 + \sqrt{2\eta^2 \ln(16k^2/\delta)})$.
  Uncertainty $\Delta_t = \sup_{f_1, f_2 \in \mathcal{F}_t} \left| f_1(x_t) - f_2(x_t) \right|^2$.
  With probability $\min\{1, \alpha\Delta_t\}$, set $q_t = 1$.
  if $q_t = 1$ **then**
    Query y_t. $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

Definition 3. *The ϵ-eluder dimension of $\mathcal{F}$ with respect to support $\mathcal{X}_u$, $\dim_u^E(\mathcal{F}, \epsilon)$, is defined as the length of the longest sequence of elements in $\mathcal{X}_u$ such that for some $\epsilon' > \epsilon$, every element is ϵ'-independent of of its predecessors.*

The above domain-dependent eluder dimension notion captures how effective the potential value of acquiring a new label can be estimated from labeled examples in domain u.[3]

The Algorithm. The master algorithm, Algorithm 4 in Appendix A.3, runs $O(\log T)$ copies of Algorithm 3. At round t, Algorithm 3 predicts using the empirical risk minimizer $\hat{f}_t$ based on all previously queried examples. Same as Algorithm 1, Algorithm 3 queries with probability $\min\{1, \alpha\Delta_t\}$, where Δ_t is an uncertainty measure of the algorithm on example x_t. To compute the uncertainty measure, it constructs a confidence set $\mathcal{F}_t$, so that with high probability, the ground truth $f^* \in \mathcal{F}_t$ for all t. The uncertainty measure Δ_t is the squared maximum disagreement between two hypotheses in $\mathcal{F}_t$ on x_t. It can be seen that with high probability, its regret and query complexity are bounded by $O(\sum_{t=1}^T \Delta_t)$ and $O(\sum_{t=1}^T \min\{1, \alpha\Delta_t\})$, respectively.

We can bound the regret of the algorithm on example from domain u, in terms of domain complexity measure $R_u = \tilde{O}(\tilde{\eta}^2 d'_u \log \mathcal{N}(\mathcal{F}, T^{-2}, \|\cdot\|_\infty))$. Here $\mathcal{N}(\mathcal{F}, \epsilon, \|\cdot\|_\infty)$ is the *ϵ-covering number* of $\mathcal{F}$ with respect to $\|\cdot\|_\infty$. Specifically, we have the following theorem.

Theorem 4. *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ has an admissible partition $\{I_u : u \in [m]\}$, where for each u, $|I_u| = T_u$, and the eluder dimension of $\mathcal{F}$ w.r.t. $\{x_t\}_{t \in I_u}$ is d'_u. Then, given label budget $B \leq \tilde{O}(\sum_u \sqrt{R_u T_u} \min_u \sqrt{R_u/T_u})$, Algorithm 4 satisfies:*

1. It has query complexity $Q \leq B$;
2. With probability $1 - \delta$, its regret $R = \tilde{O}((\sum_u \sqrt{R_u T_u})^2/B)$.

The proof of the theorem can be found in Section A.3. Specializing the theorem to linear hypothesis class $\mathcal{F} = \{\langle x, \theta \rangle : \theta \in \mathbb{R}^d, \|\theta\|_2 \leq 1\}$, if $\mathcal{X}_u$ is a subset of a d_u-dimensional subspace of $\mathbb{R}^d$, we have $\dim_u^E(\mathcal{F}, 1/T_u^2) = \tilde{O}(d_u)$, $\log \mathcal{N}(\mathcal{F}, 1/T_u^2, \|\cdot\|_\infty) = \tilde{O}(d)$, implying $R_u = \tilde{O}(\tilde{\eta}^2 d_u d)$, which in turn implies that $R = \tilde{O}(\tilde{\eta}^2 d (\sum_u \sqrt{d_u T_u})^2/B)$. Compared to Theorem 2, the additional factor d is due to increased difficulty with adaptive adversary.

6. Experiments

We test the query-regret tradeoffs of QuFUR, the uniform query baseline (Section 3.2), and naive greedy query (i.e., always querying until labeling budget is exhausted) on several linear regression tasks.

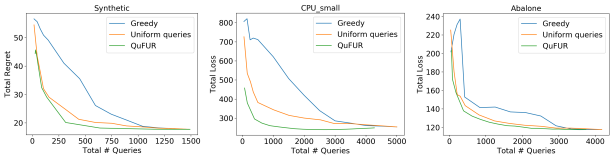

Figure 1. Total regret vs. total number of queries in synthetic dataset (**left**), cpu-small dataset (**middle**), and Abalone dataset (**right**), averaged across 5 runs. QuFUR is best and has more advantage on cpu-small potentially due to latent domain structure, whereas Abalone is more homogeneous.

Datasets We create a synthetic dataset with 20 domains. Each domain has either $T_u = 100$ and $d_u = 6$, or $T_u = 50$ and $d_u = 3$. Inputs from each domain spans a random subset of d_u out of $d = 40$ dimensions, with potential overlap between domains. θ^* is a random vector on the unit sphere in $\mathbb{R}^d$, as are x_i's from domain u in $\mathbb{R}^{d_u}$. Noise ξ_t's are iid zero-mean Gaussian with variance $\eta^2 = 0.1$.

We also experiment on two real-world LIBSVM datasets (Chang & Lin, 2011) **cpu-small** and **Abalone**. **cpu-small** uses 12 features, such as system reads/writes per second, to predict portion of time that cpu runs in user mode. **Abalone** uses 8 features (physical measurements) to predict animal ages.

Results We run QuFUR(α) for $\alpha \in [1/400, 400]$ and uniform queries with probability $\mu \in [0.01, 1]$. Figure 1 shows that QuFUR achieves the lowest total regret under the same labeling budget across all 3 datasets. Notably, QuFUR's advantage is more significant on **cpu-small**. We conjecture that this task has underlying domain structure, as different CPU usage modes may be predicted from a subset of

[3] Appendix D in Russo & Van Roy (2013) gives eluder dimension bounds for common function classes.

metrics. QuFUR potentially exploits this latent structure without knowledge of its existence.

7. Conclusion

We formulate a novel task of active online learning with latent domain structure. We propose a surprisingly simple algorithm that adapts to domain shifts, and give matching upper and lower bounds in a wide range of domain structure specifications for linear regression. The strategy is readily generalizable to other problems, as we did for non-linear regression, simply relying on a suitable uncertainty estimate for unlabeled data. We believe that our problem and solution can spur future work on making active online learning more practical.

References

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, pp. 2312–2320, 2011.

Agarwal, A. Selective sampling algorithms for cost-sensitive multiclass prediction. In *International Conference on Machine Learning*, pp. 1220–1228, 2013.

Azoury, K. S. and Warmuth, M. K. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43(3):211–246, 2001.

Balcan, M.-F., Beygelzimer, A., and Langford, J. Agnostic active learning. *Journal of Computer and System Sciences*, 75(1):78–89, 2009.

Bartlett, P. L., Dani, V., Hayes, T., Kakade, S., Rakhlin, A., and Tewari, A. High-probability regret bounds for bandit online linear optimization. 2008.

Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., and Vaughan, J. W. A theory of learning from different domains. *Machine learning*, 79(1-2):151–175, 2010.

Beygelzimer, A., Hsu, D. J., Langford, J., and Zhang, T. Agnostic active learning without constraints. In *Advances in neural information processing systems*, pp. 199–207, 2010.

Bousquet, O. and Warmuth, M. K. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3(Nov):363–396, 2002.

Cavallanti, G., Cesa-Bianchi, N., and Gentile, C. Learning noisy linear classifiers via adaptive and selective sampling. *Machine learning*, 83(1):71–102, 2011.

Cesa-Bianchi, N., Conconi, A., and Gentile, C. On the generalization ability of on-line learning algorithms. *IEEE Transactions on Information Theory*, 50(9):2050–2057, 2004a.

Cesa-Bianchi, N., Lugosi, G., and Stoltz, G. Minimizing regret with label efficient prediction. In *International Conference on Computational Learning Theory*, pp. 77–92. Springer, 2004b.

Cesa-Bianchi, N., Gentile, C., and Zaniboni, L. Worst-case analysis of selective sampling for linear classification. *Journal of Machine Learning Research*, 7(Jul):1205–1230, 2006.

Cesa-Bianchi, N., Gentile, C., and Orabona, F. Robust bounds for classification via selective sampling. In *Proceedings of the 26th annual international conference on machine learning*, pp. 121–128, 2009.

Chang, C.-C. and Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chaudhuri, K., Kakade, S. M., Netrapalli, P., and Sanghavi, S. Convergence rates of active learning for maximum likelihood estimation. In *Advances in Neural Information Processing Systems*, pp. 1090–1098, 2015.

Chaudhuri, K., Jain, P., and Natarajan, N. Active heteroscedastic regression. In *International Conference on Machine Learning*, pp. 694–702, 2017.

Cohn, D., Atlas, L., and Ladner, R. Improving generalization with active learning. *Machine learning*, 15(2): 201–221, 1994.

Daniely, A., Gonen, A., and Shalev-Shwartz, S. Strongly adaptive online learning. In *International Conference on Machine Learning*, pp. 1405–1411, 2015.

Dasgupta, S. and Hsu, D. Hierarchical sampling for active learning. In *Proceedings of the 25th international conference on Machine learning*, pp. 208–215, 2008.

Dasgupta, S., Hsu, D. J., and Monteleoni, C. A general agnostic active learning algorithm. In *Advances in neural information processing systems*, pp. 353–360, 2008.

Dekel, O., Gentile, C., and Sridharan, K. Robust selective sampling from single and multiple teachers. In *COLT*, pp. 346–358, 2010.

Dereziński, M. and Warmuth, M. K. Reverse iterative volume sampling for linear regression. *The Journal of Machine Learning Research*, 19(1):853–891, 2018.

Derezinski, M., Warmuth, M. K., and Hsu, D. J. Leveraged volume sampling for linear regression. In *Advances in Neural Information Processing Systems*, pp. 2505–2514, 2018.

Drineas, P., Mahoney, M. W., and Muthukrishnan, S. Sampling algorithms for l 2 regression and applications. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pp. 1127–1136. Society for Industrial and Applied Mathematics, 2006.

Engstrom, L., Ilyas, A., Santurkar, S., Tsipras, D., Steinhardt, J., and Madry, A. Identifying statistical bias in dataset replication. *arXiv preprint arXiv:2005.09619*, 2020.

Fedorov, V. V. and Hackl, P. *Model-oriented design of experiments*, volume 125. Springer Science & Business Media, 2012.

Fontaine, X., Perrault, P., and Perchet, V. Active linear regression. *arXiv preprint arXiv:1906.08509*, 2019.

Hanneke, S. Rates of convergence in active learning. *The Annals of Statistics*, 39(1):333–361, 2011.

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

Hazan, E. Introduction to online convex optimization. *arXiv preprint arXiv:1909.05207*, 2019.

Hazan, E. and Seshadhri, C. Adaptive algorithms for online decision problems. In *Electronic colloquium on computational complexity (ECCC)*, volume 14, 2007.

Hazan, E., Agarwal, A., and Kale, S. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A. A., and Darrell, T. Cycada: Cycle-consistent adversarial domain adaptation. *arXiv preprint arXiv:1711.03213*, 2017.

Kouw, W. M. and Loog, M. An introduction to domain adaptation and transfer learning. *arXiv preprint arXiv:1812.11806*, 2018.

Kpotufe, S., Urner, R., and Ben-David, S. Hierarchical label queries with data-dependent partitions. In *Conference on Learning Theory*, pp. 1176–1189, 2015.

Lattimore, T. and Szepesvári, C. Bandit algorithms. *preprint*, 2018.

Li, L., Littman, M. L., Walsh, T. J., and Strehl, A. L. Knows what it knows: a framework for self-aware learning. *Machine learning*, 82(3):399–443, 2011.

Locatelli, A., Carpentier, A., and Kpotufe, S. Adaptivity to noise parameters in nonparametric active learning. *Proceedings of Machine Learning Research vol*, 65:1–34, 2017.

Minsker, S. Plug-in approach to active learning. *The Journal of Machine Learning Research*, 13(1):67–90, 2012.

Rai, P., Saha, A., Daumé III, H., and Venkatasubramanian, S. Domain adaptation meets active learning. In *Proceedings of the NAACL HLT 2010 Workshop on Active Learning for Natural Language Processing*, pp. 27–32. Association for Computational Linguistics, 2010.

Recht, B., Roelofs, R., Schmidt, L., and Shankar, V. Do cifar-10 classifiers generalize to cifar-10? *arXiv preprint arXiv:1806.00451*, 2018.

Recht, B., Roelofs, R., Schmidt, L., and Shankar, V. Do imagenet classifiers generalize to imagenet? *arXiv preprint arXiv:1902.10811*, 2019.

Russo, D. and Van Roy, B. Eluder dimension and the sample complexity of optimistic exploration. In *Advances in Neural Information Processing Systems*, pp. 2256–2264, 2013.

Sabato, S. and Munos, R. Active regression by stratification. In *Advances in Neural Information Processing Systems*, pp. 469–477, 2014.

Saha, A., Rai, P., Daumé, H., Venkatasubramanian, S., and DuVall, S. L. Active supervised domain adaptation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 97–112. Springer, 2011.

Sayedi, A., Zadimoghaddam, M., and Blum, A. Trading off mistakes and don't-know predictions. In *Advances in Neural Information Processing Systems*, pp. 2092–2100, 2010.

Strickland, E. Ai-human partnerships tackle "fake news": Machine learning can get you only so far-then human judgment is required - [news]. *IEEE Spectrum*, 55:12–13, 2018.

Szita, I. and Szepesvári, C. Agnostic kwik learning and efficient approximate reinforcement learning. In *Proceedings of the 24th Annual Conference on Learning Theory*, pp. 739–772, 2011.

Vovk, V. Competitive on-line statistics. *International Statistical Review*, 69(2):213–248, 2001.

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

Xiao, M. and Guo, Y. Online active learning for cost sensitive domain adaptation. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pp. 1–9, 2013.

Xu, Q. and Yang, Q. A survey of transfer and multitask learning in bioinformatics. *Journal of Computing Science and Engineering*, 5(3):257–268, 2011.

Yang, L. Active learning with a drifting distribution. In *Advances in Neural Information Processing Systems*, pp. 2079–2087, 2011.

Zhang, C. and Chaudhuri, K. The extended littlestone's dimension for learning with mistakes and abstentions. In *Conference on Learning Theory*, pp. 1584–1616, 2016.

Zhang, C., Bengio, S., Hardt, M., Recht, B., and Vinyals, O. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2016.

Zheng, K., Luo, H., Diakonikolas, I., and Wang, L. Equipping experts/bandits with long-term memory. In *Advances in Neural Information Processing Systems*, pp. 5927–5937, 2019.

A. Proofs for upper bounds

A.1. Proof of Theorem 1

We provide the proof of Theorem 1 in this section. We focus on regret and query complexity bounds on one domain I_u, and sum over domain u to obtain Theorem 1. Recall that we define the interaction history between the learner and the environment up to time t be $H_t = \{x_{1:t}, f_{1:t}, \xi_{1:t}\}$; we abbreviate $\mathbb{E}[\cdot|x_t, H_{t-1}]$ as $\mathbb{E}_{t-1}[\cdot]$.

The following lemma upper bounds the regret with sum of uncertainty estimates, $\Delta_t = \tilde{\eta}^2 \min\left(1, \|x_t\|^2_{M_t^{-1}}\right)$. A similar lemma has appeared in Cesa-Bianchi et al. (2009, Lemma 1).

Lemma 1. *In the setting of Theorem 1, with probability* $1 - \frac{\delta}{2}$*, for all* $t \in [T]$*,* $(\hat{y}_t - \langle \theta^*, x_t\rangle)^2 = \tilde{O}(\Delta_t)$.

Proof of Lemma 1. Denote the value of $M, \mathcal{Q}$ at the beginning of round t as $M_t, \mathcal{Q}_t$. Let $\lambda = 1/C^2$, $V_t = M_t - \lambda I = \sum_{s\in\mathcal{Q}_t} x_s x_s^\top$. Therefore, $\hat{\theta}_t = M_t^{-1}(\sum_{s\in\mathcal{Q}_t} x_s y_s) = M_t^{-1}(V_t\theta^* + \sum_{s\in\mathcal{Q}_t}\xi_s x_s)$, and

$$\langle x_t, \hat{\theta}_t - \theta^*\rangle = \sum_{s\in\mathcal{Q}_t} \xi_s (x_t^\top M_t^{-1} x_s) - \lambda x_t^\top M_t^{-1}\theta^*. \tag{4}$$

The first term is a sum over a set of independent sub-Gaussian random variables, so it is $(\eta\sigma)^2$-sub-Gaussian with $\sigma^2 = \sum_{s\in\mathcal{Q}_t} x_t^\top M_t^{-1} x_s x_s^\top M_t^{-1} x_t \le x_t^\top M_t^{-1} x_t$. Define event

$$E_t = \left\{ \left| \sum_{s\in\mathcal{Q}_t} \xi_s (x_t^\top M_t^{-1} x_s) \right| \le \eta\sqrt{2\ln\frac{4T}{\delta}} \|x_t\|_{M_t^{-1}} \right\}.$$

By standard concentration of subgaussian random variables, we have $\mathbb{P}(E_t) \ge 1 - \frac{\delta}{2T}$. Define $E = \cap_{t=1}^T E_t$. By union bound, we have $\mathbb{P}(E) \ge 1 - \frac{\delta}{2}$. We henceforth condition on E happening, in which case the first term of Equation (4) is bounded by $\eta\sqrt{2\ln(4T/\delta)}\|x_t\|_{M_t^{-1}}$ at every time step t.

Meanwhile, the second term of Equation (4) can be bounded by Cauchy-Schwarz:

$$\left|\lambda x_t^\top M_t^{-1}\theta^*\right| = \lambda\left|\langle M_t^{-1/2}x_t, M_t^{-1/2}\theta^*\rangle\right| \le \lambda\|x_t\|_{M_t^{-1}}\|\theta^*\|_{M_t^{-1}} \le \sqrt{\lambda}\|\theta^*\|_2\|x_t\|_{M_t^{-1}},$$

which is at most $\|x_t\|_{M_t^{-1}}$, since $\|\theta^*\|_2 \le C$ and $\lambda = 1/C^2$. Using the basic fact that $(A+B)^2 \le 2A^2 + 2B^2$,

$$(\langle x_t, \hat{\theta}_t\rangle - \langle x_t, \theta^*\rangle)^2 \le (4\eta^2\ln(4T/\delta) + 2)\|x_t\|^2_{M_t^{-1}}.$$

Since $\hat{y}_t = \mathrm{clip}(\langle x_t, \hat{\theta}_t\rangle) \in [-1, 1]$ and $\left|\langle x_t, \theta^*\rangle\right| \le 1$, we also trivially have $(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 \le 4$. Therefore,

$$\begin{aligned}
(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 &\le \min\left(4, (4\eta^2\ln(4T/\delta') + 2)\|x_t\|^2_{M_t^{-1}}\right) \\
&\le (4\eta^2\ln(2T/\delta') + 4)\cdot\min\left(1, \|x_t\|^2_{M_t^{-1}}\right) \\
&\le \tilde{O}\left(\tilde{\eta}^2\min\left(1, \|x_t\|^2_{M_t^{-1}}\right)\right) = \tilde{O}(\Delta_t).
\end{aligned}$$

□

The following lemma bounds the sum of uncertainty estimates for k *queried* examples in a domain:

Lemma 2. *Let* $a_1, \ldots, a_k$ *be* k *vectors in* $\mathbb{R}^d$*. For* $i \in [k]$*, define* $N_i = \lambda I + \sum_{j=1}^{i-1} a_j a_j^\top$*. Then, for any* $S \subseteq [k]$*,* $\sum_{i\in S}\min\left(1, \|a_i\|^2_{N_i^{-1}}\right) \le \ln(\det(\lambda I + \sum_{i\in S} a_i a_i^\top)/\det(\lambda I))$.

Proof of Lemma 2. We denote by $N_{i,S} = \lambda I + \sum_{j\in S: j\le i-1} a_j a_j^\top$. As S is a subset of $[k]$, we have that $N_{i,S} \preceq N_i$. Consequently, $\|a_i\|_{N_i^{-1}} \le \|a_i\|_{N_{i,S}^{-1}}$. Therefore,

$$\sum_{i\in S}\min\left(1, \|a_i\|^2_{N_i^{-1}}\right) \le \sum_{i\in S}\min\left(1, \|a_i\|^2_{N_{i,S}^{-1}}\right) \le \ln\left(\frac{\det(\lambda I + \sum_{i\in S} a_i a_i^\top)}{\det(\lambda I)}\right),$$

where the second inequality is well-known (see e.g. Lattimore & Szepesvári, 2018, Lemma 19.4). □

Proof of Theorem 1. Let $p_t = \min(1, \alpha\Delta_t)$ be the learner's query probability at time t; it is easy to see that $\mathbb{E}_{t-1}[q_t] = p_t$. Let random variable $Z_t = q_t\Delta_t$. We have the following simple facts:

1. $Z_t \le \tilde{\eta}^2$,
2. $\mathbb{E}_{t-1} Z_t = p_t \Delta_t$,
3. $\mathbb{E}_{t-1} Z_t^2 \le \tilde{\eta}^2 \cdot \mathbb{E}_{t-1} Z_t \le \tilde{\eta}^2 p_t \Delta_t$.

For every $u \in [m]$, define event

$$F_u = \left\{ \left| \sum_{t \in I_u} p_t \Delta_t - \sum_{t \in I_u} q_t \Delta_t \right| \le O\left(\tilde{\eta} \sqrt{\sum_{t \in I_u} p_t \Delta_t \ln \frac{T}{\delta}} + \tilde{\eta}^2 \ln \frac{T}{\delta} \right) \right\}. \tag{5}$$

Applying Freedman's inequality to $\{Z_t\}_{t \in I_u}$ (see e.g. Bartlett et al., 2008, Lemma 2), we have that $\mathbb{P}(F_u) \ge 1 - \frac{\delta}{4m}$.

Similarly, define

$$G = \left\{ \left| \sum_{t=1}^{T} p_t - \sum_{t=1}^{T} q_t \right| \le O\left(\sqrt{\sum_{t=1}^{T} p_t \ln \frac{T}{\delta}} + \ln \frac{T}{\delta} \right) \right\}. \tag{6}$$

Applying Freedman's inequality to $\{q_t\}_{t \in I_u}$, we have that $\mathbb{P}(G) \ge 1 - \frac{\delta}{4}$.

Furthermore, define $H = E \cap (\cap_{u=1}^{m} F_u) \cap G$, where E is the event defined in the proof of Lemma 1. By union bound, $\mathbb{P}(H) \ge 1 - \delta$. We henceforth condition on H happening.

By the definition of F_u, Solving for $\sum_{t \in I_u} p_t \Delta_t$ in Equation (5), we get that

$$\sum_{t \in I_u} p_t \Delta_t = \tilde{O}\left(\sum_{t \in I_u} q_t \Delta_t + \tilde{\eta}^2 \right). \tag{7}$$

Using Lemma 2 with $\{a_i\}_{i=1}^{k} = \{x_t\}_{t \in \mathcal{Q}_T}$, and $S = I_u \cap \mathcal{Q}_T$, we get that

$$\begin{aligned} \sum_{t \in I_u} q_t \Delta_t &\le \tilde{\eta}^2 \cdot \ln\det\left(I + C^2 \sum_{t \in I_u \cap \mathcal{Q}_T} x_t x_t^\top \right) \\ &\le 2\tilde{\eta}^2 d_u \ln\left(1 + C^2 T_u / d_u\right) = \tilde{O}(\tilde{\eta}^2 d_u). \end{aligned}$$

In combination with Equation (7), we have $\sum_{t \in I_u} p_t \Delta_t = \tilde{O}(\tilde{\eta}^2 d_u)$.

We divide the examples in domain u into high and low risk subsets with index sets $I_{u,+}$ and $I_{u,-}$ (abbrev. I_+ and I_- hereafter). Formally,

$$I_+ = \{t \in I_u : \alpha\Delta_t > 1\}, \quad I_- = I - I_+.$$

We consider bounding the regrets and the query complexities in these two sets respectively:

1. For every t in I_+, as $p_t = 1$, label y_t is queried, so

$$\sum_{t \in I_+} \Delta_t = \sum_{t \in I_+} q_t \Delta_t \le \sum_{t \in I_u} q_t \Delta_t = \tilde{O}(\tilde{\eta}^2 d_u).$$

   Since for every t in I_-, $\Delta_t > 1/\alpha$, we have $\sum_{t \in I_+} \Delta_t > |I_+|/\alpha$. This implies that $\sum_{t \in I_+} p_t = |I_+| = \tilde{O}(\alpha\tilde{\eta}^2 d_u)$.

2. For every t in I_-, $p_t = \alpha\Delta_t$. Therefore, $\sum_{t\in I_-} \alpha\Delta_t^2 = \sum_{t\in I_-} p_t\Delta_t \le \sum_{t\in I_u} p_t\Delta_t = \tilde{O}(\tilde{\eta}^2 d_u)$. By Cauchy-Schwarz, and the fact that $|I_-| \le T_u$, we get that $\sum_{t\in I_-} \Delta_t \le \sqrt{|I_-| \cdot (\sum_{t\in I_-} \Delta_t^2)} = \tilde{O}(\tilde{\eta}\sqrt{d_u T_u/\alpha})$.

   Consequently, $\sum_{t\in I_-} p_t = \sum_{t\in I_-} \alpha\Delta_t \le \tilde{O}(\tilde{\eta}\sqrt{\alpha d_u T_u})$.

Summing over the two cases, we have

$$\sum_{t\in I_u} p_t \le \tilde{O}\left(\alpha\tilde{\eta}^2 d_u + \tilde{\eta}\sqrt{\alpha d_u T_u}\right), \quad \sum_{t\in I_u} \Delta_t \le \tilde{O}\left(\tilde{\eta}^2 d_u + \tilde{\eta}\sqrt{d_u T_u/\alpha}\right),$$

By the assumption that $\alpha \le \frac{1}{\tilde{\eta}^2}\min_u \frac{T_u}{d_u}$, for every u, we have $\alpha\tilde{\eta}^2 d_u \le \tilde{\eta}\sqrt{\alpha d_u T_u}$. Therefore, the above bounds can be simplified to

$$\sum_{t\in I_u} p_t \le \tilde{O}\left(\tilde{\eta}\sqrt{\alpha d_u T_u}\right), \quad \sum_{t\in I_u} \Delta_t \le \tilde{O}\left(\tilde{\eta}\sqrt{d_u T_u/\alpha}\right). \tag{8}$$

For the query complexity, from the definition of event G, applying AM-GM inequality on Equation (6), we also have

$$Q = \sum_{t=1}^{T} q_t = \tilde{O}\left(\sum_{t=1}^{T} p_t + 1\right) = \tilde{O}\left(\tilde{\eta}\sum_{u=1}^{m}\sqrt{\alpha d_u T_u} + 1\right) = \tilde{O}\left(\tilde{\eta}\sum_{u=1}^{m}\sqrt{\alpha d_u T_u}\right).$$

where in the last equality we use the assumption that $\alpha \ge \frac{1}{\tilde{\eta}^2} \cdot \frac{1}{(\sum_u \sqrt{d_u T_u})^2}$.

For the regret guarantee, we have by the definition of event E and Lemma 1 that

$$\sum_{t=1}^{T}(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 \le \tilde{O}\left(\sum_{t=1}^{T}\Delta_t^2\right) = \tilde{O}\left(\sum_{u=1}^{m}\left(\sum_{t\in I_u}\Delta_t^2\right)\right).$$

Using the second inequality of Equation (8), we get

$$\sum_{t=1}^{T}(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 \le \tilde{O}\left(\tilde{\eta}\sum_{u=1}^{m}\sqrt{d_u T_u/\alpha}\right).$$

The theorem follows. □

A.2. Proof of Theorem 2

Before going into the proof, we set up some useful notations. Define $I = \{0, 1, \ldots, k\}$ as the index set of the α_i's of interest. Recall the number of copies $k = 1 + \lceil 4\log T\rceil \le 2 + 4\log T$. Recall also that $B' = \lfloor B/k\rfloor$ is the label budget for each copy.

Let $p_t^i = \min(1, \alpha_i\Delta_t)$ be the intended query probability of copy i at time step t; let $r_t^i \sim \text{Bernoulli}(p_t^i)$ be the attempted query decision of copy i at time step t; let $A_t^i = \mathbb{1}\left[\sum_{j=1}^{t-1} r_j^i < B'\right]$, i.e. the indicator that copy i has not reached its budget limit at time step t. Using this notation, the actual query decision of copy i, q_t^i, can be written as $r_t^i A_t^i$.

We have the following useful observation that gives a sufficient condition for copy i to be within its label budget:

Lemma 3. *Given $i \in [k]$, if $\sum_{t=1}^{T} A_t^i r_t^i < B'$, the following items hold:*

1. $\sum_{t=1}^{T} r_t^i < B'$.
2. *For all $t \in [T]$, $A_t^i = 1$, i.e. copy i does not run of label budget throughout.*

Proof. Suppose for the sake of contradiction that $\sum_{t=1}^{T} r_t^i \geq B'$. Consider the first B' occurrences of $r_j^i = 1$; call them $J = \{j_1, \ldots, j_{B'}\}$. It can be seen that for all $j \in J$, $A_j^i = 1$. Therefore,

$$\sum_{t=1}^{T} A_t^i r_t^i \geq \sum_{j \in J} A_j^i r_j^i \geq |J| = B',$$

which contradicts with the premise that $\sum_{t=1}^{T} A_t^i r_t^i < B'$.

The second item immediately follows from the first item, as $\sum_{j=1}^{T} r_j^i < B'$ implies that $\sum_{j=1}^{t-1} r_j^i < B'$ for every $t \in [T]$. □

Complementary to the above lemma, we can also see that for every $i \in [k]$, $\sum_{t=1}^{T} A_t^i r_t^i = \sum_{t=1}^{T} q_t^i \leq B'$ is trivially true. We next give a key lemma that generalizes Theorem 1, and upper bounds $\sum_{t=1}^{T} A_t^i r_t^i$ for all i's beyond the above trivial B' bound.

Lemma 4. *There exists* $C = \mathrm{polylog}(T, \frac{1}{\delta}) \geq 1$*, such that with probability* $1 - \delta/2$*,*

$$\sum_{t=1}^{T} A_t^i \Delta_t \leq C \cdot \tilde{\eta} \sum_u \sqrt{d_u T_u}/\sqrt{\alpha_i}, \text{ and } \sum_{t=1}^{T} A_t^i r_t^i \leq C \cdot \tilde{\eta}\sqrt{\alpha_i} \sum_u \sqrt{d_u T_u},$$

for every $i \in I$ *such that* $\alpha_i \in \left[\frac{1}{\tilde{\eta}^2}\left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2, \frac{1}{\tilde{\eta}^2} \min_{u \in [m]} \frac{T_u}{d_u}\right]$.

Proof. Applying Freedman's inequality to the martingale difference sequence $\{A_t^i(r_t^i - p_t^i)\}_{t=1}^{T}$, we get that with probability $1 - \delta/4$,

$$\sum_{t=1}^{T} A_t^i r_t^i = \tilde{O}\left(\sum_{t=1}^{T} A_t^i p_t^i + 1\right). \tag{9}$$

Applying Freedman's inequality to $\{A_t^i(r_t^i - p_t^i)\Delta_t \mathbb{1}[t \in I_u]\}_{t=1}^{T}$, and take a union bound over all $u \in [m]$, we get that with probability $1 - \delta/4$,

$$\sum_{t \in I_u} A_t^i p_t^i \Delta_t = \tilde{O}\left(\sum_{t \in I_u} A_t^i r_t^i \Delta_t + \tilde{\eta}^2\right).$$

Using Lemma 2 with $\{a_i\}_{i=1}^{k} = \{x_t\}_{t \in \mathcal{Q}_T}$, and $S = I_u \cap \mathcal{Q}_T$ we get that, deterministically, $\sum_{t \in I_u} A_t^i r_t^i \Delta_t \leq \sum_{t \in I_u} q_t \Delta_t = \tilde{O}(\tilde{\eta}^2 d_u)$. So with probability $1 - \delta/4$,

$$\sum_{t \in I_u} A_t^i p_t^i \Delta_t = \tilde{O}(\tilde{\eta}^2 d_I). \tag{10}$$

We henceforth condition on Equations (9) and (10) occuring, which happens with probability $1 - \delta/2$ by union bound. Let $I_+ = \{t \in I_u : \alpha_i \Delta_j > 1\}$, and $I_- = I_u - I_+$.

1. For I_+, by Equation (10), $\sum_{t \in I_+} A_j^i \Delta_j = \tilde{O}(\tilde{\eta}^2 d_u) \implies \sum_{j \in I_+} A_j^i p_j^i = \tilde{O}(\alpha_i \tilde{\eta}^2 d_u)$.
2. For I_-, by Equation (10), $\sum_{j \in I_-} A_j^i \alpha_i \Delta_j^2 = \sum_{j \in I_-} A_j^i p_j \Delta_j = \tilde{O}(\tilde{\eta}^2 d_u)$; this implies that $\sum_{j \in I_-} A_j^i \Delta_j = \tilde{O}(\tilde{\eta}\sqrt{d_u T_u/\alpha_i})$. In this event, we also have $\sum_{j \in I_-} A_j^i p_j^i = \sum_{j \in I_-} A_j^i \alpha_i \Delta_j = \tilde{O}(\tilde{\eta}\sqrt{d_u T_u \alpha_i})$.

Summing over the two cases, we have

$$\sum_{t \in I_u} A_t^i p_t^i \leq \tilde{O}(\alpha \tilde{\eta}^2 d_u + \tilde{\eta}\sqrt{\alpha_i d_u T_u}), \quad \sum_{t \in I_u} A_t^i \Delta_t \leq \tilde{O}(\tilde{\eta}^2 d_u + \tilde{\eta}\sqrt{d_u T_u/\alpha_i}),$$

By the assumption that $\alpha_i \leq \frac{1}{\tilde{\eta}^2} \min_u \frac{T_u}{d_u}$, for every u, we have, $\alpha_i \tilde{\eta}^2 d_u \leq \tilde{\eta}\sqrt{\alpha_i d_u T_u}$. This implies that

$$\sum_{t \in I_u} A_t^i p_t^i \leq \tilde{O}(\tilde{\eta}\sqrt{\alpha_i d_u T_u}), \quad \sum_{t \in I_u} A_t^i \Delta_t \leq \tilde{O}(\tilde{\eta}\sqrt{d_u T_u / \alpha_i}). \tag{11}$$

Summing over $u \in [m]$, we get

$$\sum_{t=1}^{T} A_t^i p_t^i \leq \tilde{O}(\tilde{\eta} \sum_{u=1}^{m} \sqrt{\alpha_i d_u T_u}), \quad \sum_{t=1}^{T} A_t^i \Delta_t \leq \tilde{O}(\tilde{\eta} \sum_{u=1}^{m} \sqrt{d_u T_u / \alpha_i}).$$

Therefore, using Equation (9), we have

$$\sum_{t=1}^{T} A_t^i r_t^i \leq \tilde{O}\left(\sum_{t=1}^{T} A_t^i p_t^i + 1\right) \leq \tilde{O}\left(\tilde{\eta} \sum_{u=1}^{m} \sqrt{\alpha_i d_u T_u} + 1\right) \leq \tilde{O}\left(\tilde{\eta} \sum_{u=1}^{m} \sqrt{\alpha_i d_u T_u}\right),$$

where the last inequality uses the assumption that $\alpha_i \geq \frac{1}{\tilde{\eta}^2}\left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2$. The lemma follows. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. First, the query complexity of Fixed-Budget QuFUR is B by construction, as the algorithm maintains k copies of QuFUR, and each copy consumes at most $B' = \lfloor B/k \rfloor$ labels.

We now bound the regret of Fixed-Budget QuFUR. We consider $\overline{B} = Ck(\sum_u \sqrt{d_u T_u}) \cdot \min_{u \in [m]} \sqrt{T_u / d_u} = \tilde{O}\left(\sum_u \sqrt{d_u T_u}) \cdot \min_{u \in [m]} \sqrt{T_u / d_u}\right)$, where $C = \text{polylog}(T, \frac{1}{\delta}) \geq 1$ is defined in Lemma 4. We will show that if $B \in (0, \overline{B}]$, with probability $1 - \delta$, the regret of Fixed-Budget QuFUR is at most $\tilde{O}\left(\frac{\tilde{\eta}^2 (\sum_u \sqrt{d_u T_u})^2}{B}\right)$.

If $B < 2C\tilde{\eta}^2$, the regret of the algorithm is trivially upper bounded by $4T$, which is clearly $\tilde{O}\left(\frac{\tilde{\eta}^2 (\sum_u \sqrt{d_u T_u})^2}{B}\right)$. Therefore, throughout the rest of the proof, we consider $B \in [2C\tilde{\eta}^2, \overline{B}]$.

Recall that $I = \left\{\frac{2^i}{T^3} : i \in \{0, 1, \ldots, k\}\right\}$. We denote by $\alpha_{\min} = \frac{1}{T^3}$ the minimum element of I, and $\alpha_{\max} = \frac{2^k}{T^3} \geq T$ the maximum element of I.

Denote by

$$i_B = \max\left\{i \in I : C\tilde{\eta}\sqrt{\alpha_i} \sum_{u=1}^{m} \sqrt{d_u T_u} < B'\right\} = \max\left\{i \in I : \alpha_i < \left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2\right\}.$$

As $B \in [2C\tilde{\eta}^2, \overline{B}]$, we have $\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \in (\alpha_{\min}, \alpha_{\max}]$. Indeed, $\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \leq \left(\frac{\overline{B}}{Ck\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \leq T \leq \alpha_{\max}$, $\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \geq \left(\frac{\tilde{\eta}}{\sum_u \sqrt{d_u T_u}}\right)^2 > \alpha_{\min}$, as $\sum_u \sqrt{d_u T_u} \leq \sum_u T_u = T$.

Therefore, by the definition of i_B, we have

$$\alpha_{i_B} \in \left[\frac{1}{2}\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2, \left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2\right) \tag{12}$$

Again by our assumption on B, $\frac{1}{2}\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \geq \tilde{\eta}^2 \left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2 \geq \frac{1}{\tilde{\eta}^2}\left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2$, $\left(\frac{B'}{C\tilde{\eta} \sum_u \sqrt{d_u T_u}}\right)^2 \leq$

$\left(\frac{\overline{B}}{Ck\tilde{\eta}\sum_u \sqrt{d_u T_u}}\right)^2 \leq \frac{1}{\tilde{\eta}^2}\min_{u\in[m]}\frac{T_u}{d_u}$. Therefore,

$$\alpha_{i_B} \in \left[\frac{1}{\tilde{\eta}^2}\left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2, \frac{1}{\tilde{\eta}^2}\min_{u\in[m]}\frac{T_u}{d_u}\right].$$

Hence, the premises of Lemma 4 is satisfied for $i = i_B$; this gives that with probability $1-\delta/2$,

$$\sum_{t=1}^{T} A_t^{i_B}\Delta_t \leq C\cdot\tilde{\eta}\sum_u \sqrt{d_u T_u}/\sqrt{\alpha_{i_B}}, \tag{13}$$

and

$$\sum_{t=1}^{T} A_t^{i_B} r_t^{i_B} \leq C\cdot\tilde{\eta}\sqrt{\alpha_{i_B}}\sum_u \sqrt{d_u T_u}. \tag{14}$$

Now from Equation (14) and the definition of i_B, we have

$$\sum_{t=1}^{T} A_t^{i_B} r_t^{i_B} \leq C\cdot\tilde{\eta}\sqrt{\alpha_{i_B}}\sum_u \sqrt{d_u T_u} < B'.$$

Applying Lemma 3, we deduce that for all t in $[T]$, $A_t^{i_B} = 1$. Plugging this back to Equation (13), we have

$$\begin{aligned}\sum_{t=1}^{T}\Delta_t =& \sum_{t=1}^{T} A_t^{i_B}\Delta_t \\ \leq& C\cdot\tilde{\eta}\sum_u \sqrt{d_u T_u}/\sqrt{\alpha_{i_B}} \\ \leq& \tilde{O}\left(\frac{\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2}{B}\right).\end{aligned}$$

where the second inequality is from the lower bound of α_{i_B} in Equation (12).

Combining the above observation with Lemma 1, along with the union bound, we get that with probability $1-\delta$,

$$R = \sum_{t=1}^{T}(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 = \tilde{O}\left(\sum_{t=1}^{T}\Delta_t\right) = \tilde{O}\left(\frac{\tilde{\eta}^2(\sum_u \sqrt{d_u T_u})^2}{B}\right). \qquad \square$$

A.3. Proof of Theorem 4

Recall that we define

$$\beta_k := 8\eta^2\log\left(4\mathcal{N}(\mathcal{F}, 1/T^2, \|\cdot\|_\infty)/\delta\right) + 2k/T^2(16 + \sqrt{2\eta^2\ln(16k^2/\delta)}),$$

and

$$R_u := \frac{T_u}{T^2} + 4\min(d'_u, T_u) + 4d'_u\beta_T\ln T_u = \tilde{O}\left(\eta^2 d'_u\log\mathcal{N}(\mathcal{F}, T^{-2}, \|\cdot\|_\infty)\right).$$

Analogous to Theorem 1, the following theorem provides the query and regret guarantees of of Algorithm 3.

Theorem 5. *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ has an admissible partition $\{I_u : u\in[m]\}$, where for each u, $|I_u| = T_u$, and the eluder dimension of $\mathcal{F}$ w.r.t. $\{x_t\}_{t\in I_u}$ is d'_u. Suppose $\alpha \geq \tilde{\eta}^2\max_{u\in[m]} R_u/T_u$. With probability $1-\delta$, Algorithm 3 satisfies:*
1. Its query complexity $Q = \tilde{O}(\tilde{\eta}\cdot\sqrt{\alpha}\sum_u\sqrt{R_u T_u})$.
2. Its regret $R = \tilde{O}(\tilde{\eta}\cdot\sum_u\sqrt{R_u T_u})/\sqrt{\alpha}$.

We shall prove Theorem 4 directly below; the proof of Theorem 5 follows as a corollary, using the same argument in the proof of Theorem 2; we note that the admissibility condition on domain partition $\{I_u\}_{u=1}^m$ ensures that $\{A_t^i(r_t^i - p_t^i)\mathbb{1}[t \in I]\}_{t=1}^T$ and $\{A_t^i(r_t^i - p_t^i)\Delta_t \mathbb{1}[t \in I]\}_{t=1}^T$ are still martingale difference sequences in our proof.

Proof of Theorem 4. We focus on proving the analogues of Lemma 1 and Lemma 2; the rest of proof follows the same argument as the proof of Theorem 2 and is therefore omitted.

Lemma 5 (Analogue of Lemma 1)**.** *With probability* $1 - \delta/2$*,* $R \le \sum_{t=1}^T \Delta_t$.

Proof. Recall that the confidence set at time t is $\mathcal{F}_t = \{f \in \mathcal{F} : \sum_{i \in \mathcal{Q}_t} (f(x_i) - \hat{f}_t(x_i))^2 \le \beta_{|\mathcal{Q}_t|}(\mathcal{F}, \delta)\}$. By Russo & Van Roy (2013, Proposition 2), we have that with probability $1 - \delta/2$, $f^* \in \mathcal{F}_t$, for all $t \in [T]$.

Meanwhile, if $f^* \in \mathcal{F}_t$, for all $t \in [T]$, $(\hat{f}_t(x_t) - f^*(x_t)) \le \sup_{f_1, f_2 \in \mathcal{F}_t} (f_1(x_t) - f_2(x_t))^2 = \Delta_t$. This implies that the regret is bounded by $R \le \sum_{t=1}^T \Delta_t$. □

Lemma 6 (Analogue of Lemma 2)**.** $\sum_{t \in I_u} q_t \Delta_t \le R_u$.

Proof. Let $k = |I_u \cap \mathcal{Q}_T|$ and write $d = d'_u$ as a shorthand. Let $(D_1, \ldots, D_k)$ be $\{\Delta_t : t \in I_u \cap \mathcal{Q}_T\}$ sorted in non-increasing order. We have

$$\sum_{t \in I_u \cap \mathcal{Q}_T} \Delta_t = \sum_{j=1}^k D_j = \sum_{j=1}^k D_j \mathbb{1}[D_j \le 1/T^4] + \sum_{j=1}^k D_j \mathbb{1}[D_j > 1/T^4].$$

Clearly, $\sum_{j=1}^k D_j \mathbb{1}[D_j \le 1/T^4] \le \frac{T_u}{T^2}$.

We know for all $j \in [k]$, $D_j \le 4$. In addition, $D_j > \epsilon^2 \iff \sum_{t \in I_u \cap \mathcal{Q}_T} \mathbb{1}[\Delta_t > \epsilon^2] \ge j$. By Lemma 7 below, this can only occur if $j < (4\beta_T/\epsilon^2 + 1)d$. Thus, when $D_j > \epsilon^2$, $j < (4\beta_T/\epsilon^2 + 1)d$, which implies $\epsilon^2 < \frac{4\beta_T d}{j-d}$. This shows that if $D_j > 1/T^4$, $D_j \le \min\left\{4, \frac{4\beta_T d}{j-d}\right\}$. Therefore $\sum_j D_j \mathbb{1}[D_j > 1/T^4] \le 4d + \sum_{j=d+1}^k \frac{4\beta_T d}{j-d} \le 4d + 4d\beta_T \log T_u$.

Consequently,

$$\sum_{t \in I_u} q_t \Delta_t = \sum_{t \in I_u \cap \mathcal{Q}_T} \Delta_t \le \min\left\{4T_u, \frac{T_u}{T^2} + 4d'_u + 4d'_u \beta_T \log T_u\right\} \le R_u. \quad \square$$

□

The following lemma generalizes Russo & Van Roy (2013, Proposition 3), in that it considers a subsequence of examples coming from a subdomain of $\mathcal{X}$. We define $\dim_I^E$ as the eluder dimension of $\mathcal{F}$ with respect to support $\{x_t : t \in I\}$. It can be easily seen that $\dim_{I_u}^E \le \dim_u^E$.

Lemma 7. *Fix* $I \subseteq [T]$*. If* $\{\beta_t \ge 0\}_{t=1}^T$ *is a nondecreasing sequence and* $\mathcal{F}_t := \{f \in \mathcal{F} : \sum_{i \in \mathcal{Q}_t} (f(x_i) - \hat{f}_t(x_i))^2 \le \beta_{|\mathcal{Q}_t|}(\mathcal{F}, \delta)\}$*, then*

$$\forall \epsilon > 0, \sum_{t \in I \cap \mathcal{Q}_T} \mathbb{1}[\Delta_t > \epsilon^2] < \left(\frac{4\beta_T}{\epsilon^2} + 1\right) dim_I^E(\mathcal{F}, \epsilon).$$

Proof. Let $k = |I \cap \mathcal{Q}_T|$, $(a_1, \ldots, a_k) = (x_t : t \in I \cap \mathcal{Q}_T)$, and $(b_1, \ldots, b_k) = (\Delta_t : t \in I \cap \mathcal{Q}_T)$. First, we show that if $b_j > \epsilon^2$ then a_j is ϵ-dependent on fewer than $4\beta_T/\epsilon^2$ disjoint subsequences of $(a_1, \ldots, a_{j-1})$, for $j \le k$. If $b_j > \epsilon^2$ and $a_j = x_t$, there are $f_1, f_2 \in \mathcal{F}_t$ such that $f_1(a_j) - f_2(a_j) > \epsilon$. By definition, if a_j is ϵ-dependent on a subsequence $(a_{i_1}, \ldots, a_{i_p})$ of $(a_1, \ldots, a_{j-1})$, then $\sum_{l=1}^p (f_1(a_{i_l}) - f_2(a_{i_l}))^2 > \epsilon^2$. Thus, if $a_j = x_t$ is ϵ-dependent on K subsequences of $(a_1, \ldots, a_{j-1})$, then $\sum_{i \in \mathcal{Q}_t} (f_1(x_i) - f_2(x_i))^2 > K\epsilon^2$. By the triangle inequality,

$$\sqrt{\sum_{i \in \mathcal{Q}_t} (f_1(x_i) - f_2(x_i))^2} \le \sqrt{\sum_{i \in \mathcal{Q}_t} (f_1(x_i) - f^*(x_i))^2} + \sqrt{\sum_{i \in \mathcal{Q}_t} (f_2(x_i) - f^*(x_i))^2} \le 2\sqrt{\beta_T}.$$

Algorithm 4 Fixed-budget QuFUR for general function class

Require: Hypotheses set $\mathcal{F}$, time horizon T, label budget B, parameter δ, noise level η.
Labeled dataset $\mathcal{Q} \leftarrow \emptyset$.
Number of copies $k \leftarrow 4\lceil \log_2 T \rceil$.
for $i = 0$ to k **do**
  Parameter $\alpha_i \leftarrow 2^i/T^2$.
for $t = 1$ to T **do**
  Predict $\hat{f}_t \leftarrow \operatorname{argmin}_{f\in\mathcal{F}} \sum_{i\in\mathcal{Q}} (f(x_i) - y_i)^2$.
  Confidence set $\mathcal{F}_t \leftarrow \{f \in \mathcal{F} : \sum_{i\in\mathcal{Q}} (f(x_i) - \hat{f}(x_i))^2 \le \beta_{|\mathcal{Q}|}(\mathcal{F}, \delta)\}$,
where $\beta_k := 8\eta^2 \log(4\mathcal{N}(\mathcal{F}, 1/T^2, \|\cdot\|_\infty)/\delta) + 2k/T^2(16 + \sqrt{2\eta^2 \ln(16k^2/\delta)})$.
  Uncertainty estimate $\Delta_t = \sup_{f_1,f_2\in\mathcal{F}_t} \left|f_1(x_t) - f_2(x_t)\right|^2$.
  for $i = 0$ to k **do**
    if $\sum_{j=1}^{t-1} q_j^i < \lfloor B/k \rfloor$ **then**
      With probability $\min\{1, \alpha_i \Delta_t\}$, set $q_t^i = 1$.
  if $\sum_i q_t^i > 0$ **then**
    Query y_t. $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

Thus, $K < 4\beta_T/\epsilon^2$.

Next, we show that in any sequence of elements in I, $(c_1, \ldots, c_\tau)$, there is some c_j that is ϵ-dependent on at least $\tau/d - 1$ disjoint subsequences of $(c_1, \ldots, c_{j-1})$, where $d := \dim_I^E(\mathcal{F}, \epsilon)$. For any integer K satisfying $Kd + 1 \le \tau \le Kd + d$, we will construct K disjoint subsequences $C_1, \ldots, C_K$. First let $C_i = (c_i)$ for $i \in [K]$. If c_{K+1} is ϵ-dependent on $C_1, \ldots, C_K$, our claim is established. Otherwise, select a C_i such that c_{K+1} is ϵ-independent and append c_{K+1} to C_i. Repeat for all $j > K + 1$ until c_j is ϵ-dependent on each subsequence or $j = \tau$. In the latter case $\sum |C_i| \ge Kd$, and $|C_i| = d$. In this case, c_τ must be ϵ-dependent on each subsequence, by the definition of $\dim_I^E$.

Now take $(c_1, \ldots, c_\tau)$ to be the subsequence $(a_{t_1}, \ldots, a_{t_\tau})$ of $(a_1, \ldots, a_k)$ consisting of elements a_j for which $b_j > \epsilon^2$. We proved that each a_{t_j} is ϵ-dependent on fewer than $4\beta_T/\epsilon^2$ disjoint subsequences of $(a_1, \ldots, a_{t_j-1})$. Thus, each c_j is ϵ-dependent on fewer than $4\beta_T/\epsilon^2$ disjoint subsequences of $(c_1, \ldots, c_{j-1})$. Combining this with the fact that there is some c_j that is ϵ-dependent on at least $\tau/d - 1$ disjoint subsequences of $(c_1, \ldots, c_{j-1})$, we have $\tau/d - 1 < 4\beta_T/\epsilon^2$. Thus, $\tau < (4\beta_T/\epsilon^2 + 1)d$. □

A.4. Analysis of uniform query strategy for online active linear regression with oblivious adversary

Theorem 6. *With probability $1 - \delta$, the uniformly querying strategy with probability μ achieves $\mathbb{E}[R] = \tilde{O}\left(\frac{\tilde{\eta}^2 d}{\mu}\right)$ and $\mathbb{E}[Q] = \mu T$.*

Proof sketch. As $Q = \sum_{t=1}^T q_t$ is a sum of T iid Bernoulli random variables with means μ, $\mathbb{E}[Q] = \mu T$.

We now bound the regret of the algorithm. We still define $\Delta_t = \tilde{\eta}^2 \min\{1, \|x_t\|^2_{M_t^{-1}}\}$.

Using Lemma 2 with $\{a_i\}_{i=1}^k = \{x_t\}_{t=1}^T$, and $S = \mathcal{Q}_T$, $\sum_t q_t\Delta_t = \tilde{O}(\tilde{\eta}^2 d)$. Let $Z_t = q_t\Delta_t$. We have $Z_t \le \Delta_t \le \tilde{\eta}^2$, $\mathbb{E}_{t-1}Z_t = \mu\Delta_t$, and $\mathbb{E}_{t-1}Z_t^2 \le \tilde{\eta}^2\mu\Delta_t$. Applying Freedman's inequality, with probability $1 - \delta/2$,

$$\sum_{t=1}^T \mu\Delta_t - \sum_{t=1}^T q_t\Delta_t = O\left(\tilde{\eta}\sqrt{\sum_{t=1}^T \mu\Delta_t \ln(\ln T/\delta)} + \tilde{\eta}^2 \ln(\ln T/\delta))\right).$$

The above inequality implies that $\sum_{t=1}^T \Delta_t = \tilde{O}\left(\frac{\tilde{\eta}^2 d}{\mu}\right)$. Now, applying Lemma 1 and take the union bound, we have that

with probability $1-\delta$,

$$R = \tilde{O}\left(\sum_{t=1}^{T} \Delta_t\right) = \tilde{O}\left(\frac{\tilde{\eta}^2 d}{\mu}\right).$$

Use the basic relationship between the expectation and tail probability $\mathbb{E}[R] = \int_0^\infty \mathbb{P}(R \geq a)da$, we conclude that $\mathbb{E}[R] = \tilde{O}\left(\frac{\tilde{\eta}^2 d}{\mu}\right)$. □

B. Proofs for lower bounds

B.1. Proof of Theorem 3

We restate a slightly strengthened version of Theorem 3 here: the assumption $B \geq \sum_{u=1}^m \sqrt{d_u T_u} \cdot \sqrt{\max_{u\in[m]} d_u/T_u}$ is weakened to $B \geq \sum_{u=1}^m d_u$.

Theorem 3. *For any $\eta \geq 1$, any set of positive integers $\left\{(d_u, T_u)\right\}_{u=1}^m$ and integer B that satisfy*

$$d_u \leq T_u, \forall u \in [m], \quad \sum_{u=1}^m d_u \leq d, \quad B \geq \sum_{u=1}^m d_u,$$

there exists an oblivious adversary such that:
1. It uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \leq \sqrt{d}$, and $\left|\langle \theta^*, x_t\rangle\right| \leq 1$; in addition, the noises $\{\xi_t\}_{t=1}^T$ are sub-Gaussian with variance proxy η^2.
2. It shows example sequence $\{x_t\}_{t=1}^T$ such that $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u\}_{u=1}^m$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t\in I_u}$ lie in a subspace of dimension d_u.
3. Any online active learning algorithm $\mathcal{A}$ with label budget B has regret $\Omega((\sum_{u=1}^m \sqrt{d_u T_u})^2/B)$.

Proof. Our proof is inspired by Vovk (2001, Theorem 2). For $u \in [m]$ and $i \in [d_u]$, define $c_{u,i} = e_{\sum_{v=1}^{u-1} d_v + i}$, where e_j denotes the j-th standard basis of $\mathbb{R}^d$. It can be easily seen that all $c_{u,i}$'s are orthonormal. In addition, for a vector $\theta \in \mathbb{R}^d$, denote by $\theta_{u,i} = \theta_{\sum_{v=1}^{u-1} d_v + i}$.

For task u, we construct domain $\mathcal{X}_u = \mathrm{span}(c_{u,i} : i \in [d_u])$. The sequence of examples shown by the adversary is the following: it is divided to m blocks, where the u-th block occupies a time interval $I_u = [\sum_{v=1}^{u-1} T_v + 1, \sum_{v=1}^{u} T_v]$; Each block is further divided to d_u subblocks, where for $i \in [d_u - 1]$, subblock (u,i) spans time interval $I_{u,i} = [\sum_{v=1}^{u-1} T_v + (i-1)\lfloor T_u/d_u\rfloor + 1, \sum_{v=1}^{u-1} T_v + i\lfloor T_u/d_u\rfloor]$, and subblock (u, d_u) spans time interval $I_{u,d_u} = [\sum_{v=1}^{u-1} T_v + (d_u - 1)\lfloor T_u/d_u\rfloor + 1, \sum_{v=1}^{u-1} T_v + T_u]$. At block u, examples from domain $\mathcal{X}_u$ are shown; furthermore, for every t in $I_{u,i}$, i.e. in the (u,i)-th subblock, example $c_{u,i}$ is repeatedly shown to the learner. Observe that (u,i)-th subblock contains at least $\lfloor \frac{T_u}{d_u}\rfloor \geq \frac{T_u}{2d_u}$ examples, as $T_u \geq d_u$.

We first choose θ^* from distribution D_θ, such that for every coordinate $j \in [d]$, $\theta_i^* \sim \mathrm{Beta}(1,1)$, which is also the uniform distribution over $[0,1]$. Given θ^*, the adversary reveals labels using the following mechanism: given x_t, it draws $y_t \sim \mathrm{Bernoulli}(\langle \theta^*, x_t\rangle)$ independently and optionally reveals it to the learner upon learner's query. Specifically, given θ^*, if $t \in I_{u,i}$, $y_t \sim \mathrm{Bernoulli}(\theta^*_{u,i})$. By Hoeffding's Lemma, $\xi_t = y_t - \theta^*_{u,i}$ is zero mean subgaussian with variance proxy $\frac{1}{4} \leq \eta^2$.

Denote by $N_{u,i}(t) = \sum_{s\in I_{u,i}: s\leq t} q_s$ the number of label queries of the learner in domain (u,i) up to time t. Because the learner satisfies a budget constraint of B under all environments, we have

$$\mathbb{E}\left[\sum_{u=1}^m \sum_{i=1}^{d_u} N_{u,i}(T) \mid \theta^*\right] \leq B.$$

Adding $2\sum_{u=1}^m d_u$ on both sides and by linearity of expectation, we get

$$\sum_{u=1}^m \sum_{i=1}^{d_u} \mathbb{E}\left[(N_{u,i}(T) + 2) \mid \theta^*\right] \leq B + 2\sum_{u=1}^m d_u \leq 3B. \tag{15}$$

On the other hand, we observe that the expected regret of the algorithm can be written as follows:

$$\mathbb{E}\left[R\right] = \mathbb{E}\left[\sum_{u=1}^{m}\sum_{i=1}^{d_u}\sum_{t\in I_{u,i}}(\hat{y}_t - \theta^*_{u,i})^2\right],$$

where the expectation is with respect to both the choice of θ^* and the random choices of $\mathcal{A}$.

We define a filtration $\{\mathcal{F}_t\}_{t=1}^T$, where $\mathcal{F}_t$ is the σ-algebra generated by $\left\{(x_s, q_s, y_s q_s)\right\}_{s=1}^t$, which encodes the informative available to the *learner* up to time step t.[4] We note that $\hat{y}_t$ is $\mathcal{F}_{t-1}$-measurable. Denote by $N^+_{u,i}(t) = \sum_{s\in I_{u,i}:s\leq t} q_s \cdot \mathbb{1}\left(y_s = 1\right)$, which is the number of 1 labels seen on example $c_{u,i}$ by the learner up to round $t-1$. Observe that both $N^+_{u,i}(t-1)$ and $N_{u,i}(t-1)$ are $\mathcal{F}_{t-1}$-measurable.

Observe that conditioned on the interaction logs $(x_s, q_s, y_s q_s)_{s=1}^{t-1}$, the posterior distribution of $\theta^*_{u,i}$ is $\mathrm{Beta}(1 + N^+_{u,i}(t-1), 1 + N_{u,i}(t-1) - N^+_{u,i}(t-1))$. Therefore, define random variable $\hat{y}^*_t = \mathbb{E}\left[\theta^*_{u,i} \mid \mathcal{F}_{t-1}\right] = \frac{1+N^+_{u,i}}{2+N_{u,i}}$, we have by bias-variance decomposition,

$$\begin{aligned}\mathbb{E}\left[(\hat{y}_t - \theta_{u,i})^2 \mid \mathcal{F}_{t-1}\right] &= \mathbb{E}\left[(\hat{y}^*_t - \theta^*_{u,i})^2 \mid \mathcal{F}_{t-1}\right] + (\hat{y}_t - \hat{y}^*_t)^2 \\ &\geq \mathbb{E}\left[(\hat{y}^*_t - \theta^*_{u,i})^2 \mid \mathcal{F}_{t-1}\right]\end{aligned}$$

Summing over all time steps, we have

$$\mathbb{E}\left[R\right] \geq \mathbb{E}\left[\sum_{u=1}^{m}\sum_{i=1}^{d_u}\sum_{t\in I_{u,i}}(\hat{y}^*_t - \theta^*_{u,i})^2\right].$$

On the other hand, from Lemma 8, we have for all $t \in I_{u,i}$,

$$\mathbb{E}\left[(\hat{y}_t - \theta^*_{u,i})^2 \mid N_{u,i}(T), \theta^*\right] \geq \frac{f(\theta^*_{u,i})}{2(N_{u,i}(T)+2)},$$

where $f(\gamma) = \min(\gamma\cdot(1-\gamma), (2\gamma-1)^2)$.

By the tower property of conditional expectation and conditional Jensen's inequality, we have

$$\mathbb{E}\left[(\hat{y}_t - \theta_{u,i})^2 \mid \theta^*\right] \geq \mathbb{E}\left[\frac{f(\theta^*_{u,i})}{N_{u,i}(T)+2} \mid \theta^*\right] \geq \frac{f(\theta^*_{u,i})}{2(\mathbb{E}\left[N_{u,i}(T) \mid \theta^*\right]+2)}.$$

Summing over all t in $I_{u,i}$, and then summing over all subblocks $(u,i) : u \in [m], i \in [d_u]$, and using the aforementioned fact that the (u,i) subblock has at least $\frac{T_u}{2d_u}$ examples, we have

$$\begin{aligned}\mathbb{E}\left[R \mid \theta^*\right] &= \sum_{u=1}^{m}\sum_{i=1}^{d_u}\sum_{t\in I_{u,i}}\mathbb{E}\left[(\hat{y}_t - \theta_{u,i})^2 \mid \theta^*\right] \\ &\geq \sum_{u=1}^{m}\sum_{i=1}^{d_u}\frac{T_u/d_u \cdot f(\theta^*_{u,i})}{4(\mathbb{E}\left[N_{u,i}(T) \mid \theta^*\right]+2)}.\end{aligned} \tag{16}$$

[4]This notion should be distinguished from the history notion H_t defined before, in that it does not include the labels not queried by the learner up to time step t. For s in $[t]$, we use $y_s q_s$ to indicate the labeled data information acquired at time step s; if $q_s = 1$, $y_s q_s = y_s$, encoding the fact that the learner has access to label y_s; otherwise $q_s = 0$, $y_s q_s$ is always 0, meaning that the learner does not have label y_s available.

Combining the above inequality with Equation (15), we have:

$$3B \cdot \mathbb{E}\left[R \mid \theta^*\right] \geq \left(\sum_{u=1}^{m}\sum_{i=1}^{d_u} \frac{T_u/d_u \cdot f(\theta^*_{u,i})}{4(\mathbb{E}\left[N_{u,i}(T) \mid \theta^*\right]+2)}\right) \cdot \left(\sum_{u=1}^{m}\sum_{i=1}^{d_u} \mathbb{E}\left[(N_{u,i}(T) \mid \theta^*\right]+2)\right)$$
$$\geq \frac{1}{4}\left(\sum_{u=1}^{m}\sum_{i=1}^{d_u}\left(\sqrt{T_u/d_u}\cdot\sqrt{f(\theta^*_{u,i})}\right)\right)^2.$$

where the second inequality is from Cauchy-Schwarz. Now taking expectation over θ, using Jensen's inequality and Lemma 9 that $\mathbb{E}\sqrt{f(\theta^*_{u,i})} \geq \frac{1}{25}$, and some algebra yields

$$3B \cdot \mathbb{E}\left[R\right] \geq \frac{1}{2}\left(\sum_{u=1}^{m}\sum_{i=1}^{d_u}\left(\sqrt{T_u/d_u}\cdot\mathbb{E}\left[\sqrt{f(\theta^*_{u,i})}\right]\right)\right)^2 \geq \frac{1}{2500}\left(\sum_{u=1}^{m}\sqrt{d_u T_u}\right)^2.$$

In conclusion, we have

$$\mathbb{E}\left[R\right] \geq \frac{\left(\sum_{u=1}^{m}\sum_{i=1}^{d_u}\sqrt{T_u/d_u}\right)^2}{7500\cdot B}.$$

As the above expectation is over θ^* chosen randomly from D_θ, there must exists an θ^* from $\mathrm{supp}(D_\theta) = [0,1]^d$ such that

$$\mathbb{E}\left[R \mid \theta^*\right] \geq \frac{\left(\sum_{u=1}^{m}\sum_{i=1}^{d_u}\sqrt{T_u/d_u}\right)^2}{7500\cdot B}$$

holds. This θ^* has ℓ_2 norm at most $\sqrt{\sum_{j=1}^{d}(\theta^*_j)^2} \leq \sqrt{d}$. □

Lemma 8. *If t is in $I_{u,i}$, then*

$$\mathbb{E}\left[(\hat{y}^*_t - \theta^*_{u,i})^2 \mid N_{u,i}(T), \theta^*\right] \geq \frac{f(\theta^*_{u,i})}{2(N_{u,i}(T)+2)},$$

where $f(\gamma) = \min\left(\gamma(1-\gamma), (2\gamma-1)^2\right)$.

Proof. We condition on $N_{u,i}(T) = m$, and a value of θ^*. Recall that $\hat{y}^*_t = \frac{1+N^+_{u,i}}{2+N_{u,i}} = \frac{1+N^+_{u,i}}{2+m}$, where $N^+_{u,i}$ can be seen as drawn from the binomial distribution $\mathrm{Bin}(m, \theta^*_{u,i})$.

$$\begin{aligned}
&\mathbb{E}\left[(\hat{y}^*_t - \theta^*_{u,i})^2 \mid N_{u,i}(T) = m, \theta^*\right] \\
=&\mathbb{E}\left[\left(\frac{1+N^+_{u,i}}{2+m} - \theta^*_{u,i}\right)^2 \mid N_{u,i}(T) = m, \theta^*\right] \\
=&\frac{m\theta^*_{u,i}(1-\theta^*_{u,i})}{(m+2)^2} + \frac{(2\theta^*_{u,i}-1)^2}{(m+2)^2} \\
\geq&\frac{m+1}{(m+2)^2} f(\theta^*_{u,i}) \geq \frac{f(\theta^*_{u,i})}{2(m+2)}.
\end{aligned}$$

□

Lemma 9. *Suppose $Z \sim \mathrm{Beta}(1,1)$. Then $\mathbb{E}\left[\sqrt{f(Z)}\right] \geq \frac{1}{25}$.*

Proof. We observe that

$$\mathbb{E}\left[\sqrt{f(Z)}\right] = \int_{[0,1]}\sqrt{f(z)}dz \geq \int_{[\frac{1}{5},\frac{2}{5}]}\sqrt{f(z)}dz,$$

Now, for all $z \in [\frac{1}{5}, \frac{2}{5}]$, $\sqrt{f(z)} \geq \sqrt{\frac{1}{25}} = \frac{1}{5}$, which implies that the above integral is at least $\frac{1}{25}$. □

B.2. Lower bound for unstructured domains

We have the following lower bound in the case when there is no domain structure.

Theorem 7. *For any set of positive integers d, T, B such that $d \leq T$ and $d \leq B$, there exists an oblivious adversary such that:*

1. *it uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \leq \sqrt{d}$, and $\left|\langle \theta^*, x_t \rangle\right| \leq 1$.*
2. *any online active learning algorithm $\mathcal{A}$ with label budget B has regret at least $\Omega\left(\frac{dT}{B}\right)$.*

Proof. This is an immediate consequence of Theorem 3, by setting $m = 1$, $d_1 = d$, $T_1 = T$, and the label budget equal to B. □

C. The c-cost model for online active learning

We consider the following variant of our learning model, which models settings where the cost ratio between a unit of square loss regret and a label query is c to 1. In this setting, the interaction protocol between the learner and the environment remains the same, with the goal of the learner modified to minimizing the total cost, formally $C = cR + Q$. We call the above model the *c-cost model*. We will show that Algorithm 1 achieves optimal cost up to constant factors, for a wide range of values of η and c.

Theorem 8. *For any $\eta \geq 1$, set of positive integers $\left\{(d_u, T_u)\right\}_{u=1}^m$ such that $d_u \leq T_u, \forall u \in [m]$, $\sum_{u=1}^m d_u \leq d$, cost ratio $c \geq \max_u \frac{d_u}{T_u}$, there exists an oblivious adversary such that:*

1. *it uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \leq \sqrt{d}$, and $\left|\langle \theta^*, x_t \rangle\right| \leq 1$; in addition, the subgaussian variance proxy of noise is η^2.*
2. *it shows example sequence $\{x_t\}_{t=1}^T$ such that $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u\}_{u=1}^m$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ lie in a subspace of dimension d_u.*
3. *any online active learning algorithm $\mathcal{A}$ has total cost $\Omega\left(\sqrt{c} \cdot (\sum_{u=1}^m \sqrt{d_u T_u})\right)$.*

Proof. Consider any algorithm $\mathcal{A}$. Same as in the proof of Theorem 3, we will choose θ^* randomly where each of its coordinates is drawn independently from the $\mathrm{Beta}(1,1)$ distribution, and show the exact same sequence of instances $\{x_t\}_{t=1}^T$ and reveals the labels the same say as in that proof. It can be seen that the η_t's are subgaussian with variance proxy 1, which is also subgaussian with variance proxy η^2.

As $\mathcal{A}$ can behave differently under different environments, we define $\mathbb{E}\left[Q \mid \theta^*\right]$ as $\mathcal{A}$'s query complexity conditioned on the adversary choosing ground truth linear predictor θ^*.

We conduct a case analysis on the random variable $\mathbb{E}\left[Q \mid \theta^*\right]$:

1. If there exists some $\theta^* \in [0,1]^d$, $\mathbb{E}\left[Q \mid \theta^*\right] \geq \sqrt{c}\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)$, then we are done: under the environment where the ground truth linear predictor is θ^*, the total cost of $\mathcal{A}$, $\mathbb{E}\left[C \mid \theta^*\right]$, is clearly at least $\mathbb{E}\left[Q \mid \theta^*\right] \geq \Omega\left(\sqrt{c}\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)\right)$.
2. If for every $\theta^* \in [0,1]^d$, $\mathbb{E}\left[Q \mid \theta^*\right] \leq \sqrt{c}\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)$, $\mathcal{A}$ can be viewed as an algorithm with label budget $B = \sqrt{c}\left(\sum_{u=1}^m \sqrt{d_u T_u}\right)$. By the premise that $c \geq \max_u \frac{d_u}{T_u}$, we get that $B \geq \sum_{u=1}^m \sqrt{d_u T_u} \cdot \sqrt{\frac{d_u}{T_u}} = \sum_{u=1}^m d_u$. Therefore, from the proof of Theorem 3, we get that there exists a θ^* in $[0,1]^d$, such that

$$\mathbb{E}\left[R \mid \theta^*\right] \geq \frac{(\sum_u \sqrt{d_u T_u})^2}{B} \geq \Omega\left(\frac{1}{\sqrt{c}}\left(\sum_u \sqrt{d_u T_u}\right)\right),$$

which implies that the total cost of $\mathcal{A}$, under the environment where the ground truth linear predictor is θ^*, $\mathbb{E}\left[C \mid \theta^*\right]$, is at least $c \cdot \mathbb{E}\left[R \mid \theta^*\right] \geq \Omega\left(\sqrt{c}\left(\sum_u \sqrt{d_u T_u}\right)\right)$.

In summary, in both cases, there is an oblivious adversary that uses θ^* in $[0,1]^d$, under which $\mathcal{A}$ has a expected cost of $\Omega\left(\sqrt{c}\left(\sum_u \sqrt{d_u T_u}\right)\right)$. □

In the theorem below, we discuss the optimality of Algorithm 1 in the c-cost for a range of problem parameters.

Theorem 9. *Suppose $\eta \in [1, O(1)]$; in addition, consider a set of $\left\{(T_u, d_u)\right\}_{u=1}^m$, such that $\min_u T_u/d_u \geq \eta$. Fix $c \in [\max_u \frac{d_u}{T_u}, \frac{1}{\eta^2}\min_u \frac{T_u}{d_u}]$. We have*

1. *Under all environments with domain dimension and duration $\left\{(T_u, d_u)\right\}_{u=1}^m$, such that $\|\theta^*\| \leq C$ and $\max_{t\in[T]}\left|\langle\theta^*, x_t\rangle\right| \leq 1$, QuFUR$(c)$ (with the knowledge of norm bound C) has the guarantee that*

$$C \leq \tilde{O}\left(\sqrt{c} \cdot \sum_u \sqrt{T_u d_u}\right),$$

2. *For any algorithm, there exists an environment with domain dimension and duration $\left\{(T_u, d_u)\right\}_{u=1}^m$ such that $\|\theta^*\| \leq \sqrt{d}$ and $\max_{t\in[T]}\left|\langle\theta^*, x_t\rangle\right| \leq 1$, under which the algorithm must have the following cost lower bound:*

$$C \geq \Omega\left(\sqrt{c} \cdot \sum_u \sqrt{T_u d_u}\right),$$

Proof. We show the two items respectively:

1. As $c \leq \tilde{\eta}^2 \min_u \frac{T_u}{d_u}$, and $c \geq \max_u \frac{d_u}{T_u} \geq \frac{1}{\tilde{\eta}^2}(\frac{1}{(\sum_u \sqrt{d_u T_u})^2})$, applying Theorem 1, we have that QuFUR(c) achieves the following regret and query complexity guarantees:

$$Q \leq O\left(\tilde{\eta}\sqrt{c}\sum_u \sqrt{T_u d_u}\right), \quad R \leq O\left(\tilde{\eta}\sum_u \sqrt{T_u d_u}/\sqrt{c}\right).$$

This implies that

$$C = cQ + R \leq O\left(\tilde{\eta}\sum_u \sqrt{T_u d_u} \cdot \sqrt{c}\right) = O\left(\sqrt{c} \cdot \sum_u \sqrt{T_u d_u}\right).$$

2. By the condition that $c \geq \max_u \frac{d_u}{T_u}$, applying Theorem 8, we get the item. □

D. The regret definition

Recall that in the main text, we define the regret of an algorithm as $R = \sum_{t=1}^T (\hat{y}_t - f^*(x_t))^2$. This is different from the usual definition of regret in online learning, which measures the difference between the loss of the learner and that of the predictor θ^*: $\text{Reg} = \sum_{t=1}^T (\hat{y}_t - y_t)^2 - \sum_{t=1}^T (f^*(x_t) - y_t)^2$.

We show a standard a result in this section that the expectation of these two notions coincide.

Theorem 10. $\mathbb{E}[R] = \mathbb{E}[\text{Reg}]$.

Proof. Denote by $\mathcal{F}_{t-1}$ be the σ-algebra generated by all observations up to time $t-1$, and x_t. As a shorthand, denote by $\mathbb{E}_{t-1}[\cdot] = \mathbb{E}[\cdot \mid \mathcal{F}_{t-1}]$.

Let $Z_t = (\hat{y}_t - y_t)^2 - (f^*(x_t) - y_t)^2$; we have

$$\begin{aligned}\mathbb{E}_{t-1} Z_t &= \mathbb{E}_{t-1}\left[(\hat{y}_t - f^*(x_t) + f^*(x_t) - y_t)^2 - (f^*(x_t) - y_t)^2\right] \\ &= \mathbb{E}_{t-1}\left[(f^*(x_t) - \hat{y}_t)^2 + 2(\hat{y}_t - f^*(x_t))(f^*(x_t) - y_t)\right] \\ &= (f^*(x_t) - \hat{y}_t)^2\end{aligned}$$

where the last inequality uses the fact that $\mathbb{E}_{t-1}(f^*(x_t) - y_t) = 0$ and $\hat{y}_t - f^*(x_t)$ is $\mathcal{F}_{t-1}$-measurable. Consequently, $\mathbb{E} Z_t = \mathbb{E}(f^*(x_t) - \hat{y}_t)^2$. The theorem is concluded by summing over all time steps t from 1 to T. □

E. Online to batch conversion

In this section we show that by an standard application of online to batch conversion (Cesa-Bianchi et al., 2004a) on QuFUR, we obtain new results on active linear regression under the batch learning setting.

First we recall a standard result on online to batch conversion; for completeness we provide its proof here.

Theorem 11. *Suppose online active learning algorithm $\mathcal{A}$ sequentially receives a set of iid examples $(x_t, y_t)_{t=1}^T$ drawn from D, and at every time step t, it outputs predictor $\hat{f}_t : \mathcal{X} \to \mathcal{Y}$. In addition, suppose $\ell : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ is a loss function. Define regret $\mathrm{Reg} = \sum_{t=1}^T \ell(\hat{f}_t(x_t), y_t) - \sum_{t=1}^T \ell(f^*(x_t), y_t)$, and define $\ell_D(f) = \mathbb{E}_{(x,y)\sim D}\ell(f(x), y)$. If $\mathbb{E}[\mathrm{Reg}] \le R_0$, then,*

$$\mathbb{E}\left[\mathbb{E}_{f\sim \mathrm{uniform}(\hat{f}_1,\dots,\hat{f}_T)}\ell_D(f)\right] - \ell_D(f^*) \le \frac{R_0}{T}.$$

Proof. As $\mathrm{Reg} = \sum_{t=1}^T \ell(\hat{f}_t(x_t), y_t) - \sum_{t=1}^T \ell(f^*(x_t), y_t)$, We have

$$\begin{aligned}R_0 \ge \mathbb{E}[\mathrm{Reg}] &= \sum_{t=1}^T \mathbb{E}\left[\ell_D(\hat{f}_t)\right] - \mathbb{E}\left[\sum_{t=1}^T \ell(f^*(x_t), y_t)\right] \\ &= T \cdot \left(\frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[\ell_D(\hat{f}_t)\right] - \mathbb{E}_{(x,y\sim D}\ell(f^*(x), y).\right)\end{aligned}$$

The theorem is proved by dividing both sides by T and recognizing that

$$\frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[\ell_D(\hat{f}_t)\right] = \mathbb{E}_{f\sim \mathrm{uniform}(\hat{f}_1,\dots,\hat{f}_T)}\ell_D(f).$$ □

Combining Theorem 11 with Theorem 2, we have the following adaptive excess loss guarantee of Fixed-Budget QuFUR (Algorithm 2) when run on iid data with hidden domain structure.

Theorem 12. *Suppose the unlabeled data distribution D_X is a mixture distribution: $D_X = \sum_{u=1}^m p_u D_u$, where D_u is a distribution supported on a subspace of $\mathbb{R}^d$ of dimension d_u and is a subset of $\left\{x : \|x\|_2 \le 1, \left|\langle \theta^*, x\rangle\right| \le 1\right\}$. The conditional distribution of y given x is $y = \langle \theta^*, x\rangle + \xi$ where ξ is a subgaussian with variance proxy η^2. In addition, suppose we are given integer B, T_0 such that $T_0 \ge \Omega\left(\max\left(\frac{B}{\sum_u \sqrt{d_u p_u}\cdot \min_u \sqrt{\frac{p_u}{d_u}}}, \frac{\ln m}{\min_u p_u}\right)\right)$. If Algorithm 2 is given dimension d, time horizon $T \ge T_0$, label budget B, norm bound C, noise level η as input, then:*
1. It uses T unlabeled examples.
2. Its query complexity Q is at most B.
3. Denote by $\ell(\hat{y}, y) = (\hat{y} - y)^2$ the square loss. We have,

$$\mathbb{E}\left[\mathbb{E}_{f\sim \mathrm{uniform}(\hat{f}_1,\dots,\hat{f}_T)}\ell_D(f)\right] - \ell_D(f^*) \le O\left(\frac{\tilde{\eta}^2(\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

Proof sketch. From Theorem 11 it suffices to show that

$$\mathbb{E}\left[\text{Reg}\right] \leq O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

By Theorem 10, $\mathbb{E}\left[\text{Reg}\right] = \mathbb{E}\left[R\right]$, it therefore suffices to show that

$$\mathbb{E}\left[R\right] \leq O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

We first show a high probability upper bound of R. Given a sequence of unlabeled examples $\{x_t\}_{t=1}^T$, we denote by S_u the subset of examples drawn from component D_u, and denote by T_u the size of S_u. From the assumption of D_u, we know that S_u all lies in a subspace of dimension d_u.

Define event E as follows:

$$E = \left\{\forall u \in [m] \,.\, T_u \in \left[\frac{Tp_u}{2}, 2Tp_u\right]\right\}.$$

From the assumption that $T \geq T_0 \geq \Omega(\frac{\ln m}{\min_u p_u})$, we have that by Chernoff bound and union bound, $\mathbb{P}(E) \geq 1 - \frac{1}{T^2}$.

Conditioned on event E happening, we have that by the assumption that $T \geq T_0 \geq \frac{B}{\sum_u \sqrt{d_u p_u} \cdot \min_u \sqrt{\frac{p_u}{d_u}}}$,

$$B \leq \tilde{O}\left(T \cdot \sum_u \sqrt{d_u p_u} \min_u \sqrt{\frac{p_u}{d_u}}\right) \leq \tilde{O}\left(\sum_u \sqrt{d_u T_u} \min_u \sqrt{\frac{T_u}{d_u}}\right).$$

Therefore, applying Theorem 2, we have that conditioned on event E happening, with probability $1 - \frac{1}{T^2}$ over the draw of $\{y_t\}_{t=1}^T$,

$$R \leq O\left(\frac{\tilde{\eta}^2 \cdot (\sum_u \sqrt{d_u T_u})^2}{B}\right) \leq O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

Combining the above two equations and using union bound, we conclude that with probability $1 - \frac{2}{T^2}$,

$$R \leq O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

Observe that with probability 1, $\hat{y}_t \in [-1, 1]$ and $\langle \theta^*, x_t \rangle \in [-1, 1]$. Therefore, $R = \sum_{t=1}^T (\hat{y}_t - \langle \theta^*, x_t \rangle)^2 \in [0, 4T]$. Hence,

$$\mathbb{E}[R] \leq \left(1 - \frac{2}{T^2}\right) \cdot O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right) + \frac{2}{T^2} \cdot 4T = O\left(\frac{\tilde{\eta}^2 T \cdot (\sum_u \sqrt{d_u p_u})^2}{B}\right).$$

The theorem follows. □