Active Online Domain Adaptation

Yining Chen CYNNJJS@STANFORD.EDU
Department of Computer Science
Stanford University
Stanford, CA 94305-9025, USA

Haipeng Luo HAIPENGL@USC.EDU
Computer Science Department
University of Southern California
Los Angeles, CA 90089, USA

Tengyu Ma TENGYUMA@STANFORD.EDU
Department of Computer Science
Stanford University
Stanford, CA 94305-9025, USA

Chicheng Zhang CHICHENGZ@CS.ARIZONA.EDU
Department of Computer Science
University of Arizona
Tucson, AZ 85721, USA

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.

Keywords: Active Learning, Online Learning, Bandit Algorithms

1. Introduction

Domain shift, the difference between training and testing distributions, is a major bottleneck for many machine learning applications (Kouw and Loog, 2018). Online learning is a classical framework to deal with worst-case domain shift. In online learning, even though the data is assumed to be given adversarially, strong regret bounds are attainable for many problems. So far, the practical deployments of fully online learning systems has been somewhat limited, because labels are expensive to obtain.

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 achieves minimax-optimal regret and query complexity for this problem. However, querying with uniform probability does not take into account the algorithm's uncertainty in each individual example, and thus can be suboptimal when the problem has certain favorable structures.

We aim to improve label-efficiency in online learning 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 balance regret and query complexity: given a fixed labeling 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 for 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 length T_u and lies in a d_u-dimensional subspace of $\mathbb{R}^d$, 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 $\tilde{O}(dT/B)$, which also matches minimax lower bounds (Theorem 19) in this setting. The advantage of QuFUR's adaptive regret guarantees becomes significant when the domains have heterogeneous time spans and dimensions. Using standard online-to-batch conversion (Cesa-Bianchi et al., 2004a), we also obtain novel results in batch active learning for regression (Theorem 24). 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 and Van Roy, 2013) of the hypotheses class with respect to the support of domain u (Theorem 6).

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

2. Setup and Preliminaries

2.1 Setup

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 $t = 1, \ldots, T$:
1. x_t is revealed to the learner.
2. The learned predicts $\hat{y}_t = \hat{f}_t(x_t)$ using predictor $\hat{f}_t \in \mathcal{F}$, incurring loss $(\hat{y}_t - y_t)^2$.
3. 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

1. $[n] := \{1, \ldots, n\}$; $\tilde{O}$ and $\tilde{\Omega}$ hide logarithmic factors.

one usually used in online learning when taking expectations; see Appendix C. Our goal is to design a learner that has low R subject to budget constraint $Q \leq B$.

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 $\{x_{1:t-1}, f_{1:t-1}, \xi_{1:t-1}\}$.

Some additional notations: For $v \in \mathbb{R}^d$ and PSD matrix $M \in \mathbb{R}^{d \times d}$, $\|v\|_M := \sqrt{v^\top M v}$. For $\{z_t\}_{t=1}^T \subseteq \mathbb{R}^l$, 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 $\tilde{\eta} := \max\{1, \eta\}$.

2.2 Baselines

We first study linear regression with oblivious adversary, where $\mathcal{F} = \{\langle x, \theta \rangle : \theta \in \mathbb{R}^d, \|\theta\|_2 \leq C\}$. Let the ground truth hypothesis be θ^*, and input space $\mathcal{X} \subseteq \{x \in \mathbb{R}^d : \|x\|_2 \leq 1, \langle x, \theta^* \rangle \leq 1\}$.

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. It is known Vovk (1990); Azoury and 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 on all queried examples. We show that the uniform querying strategy achieves $R = \tilde{O}(\tilde{\eta}^2 dT/B)$ with $Q = \tilde{O}(B) = \tilde{O}(\mu T)$. As shown in Theorem 19, this tradeoff is minimax optimal if $\tilde{\eta}$ is a constant. Although this guarantee is optimal in the worst case, when the data has some hidden domain structure, it is possible to achieve better regret guarantees, if the learner has access to domain information.

Oracle baseline when domain structure is known Suppose the learner knows: 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. Then, for any example in domain u, the learner can query 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. 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: $\min_\mu \sum_{u=1}^m d_u/\mu_u$, 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 may achieve much smaller regret than uniform query (as $(\sum_u \sqrt{d_u T_u})^2$ 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 develop algorithms that overcome these drawbacks.

3. Active online linear regression: algorithms, analysis, and matching lower bounds

3.1 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 $\hat{f}_t(x) = \mathrm{clip}(\langle \hat{\theta}_t, x \rangle)$. It makes label queries with probability proportional to a high-confidence upper bound of the instantaneous regret $(\hat{y}_t - \langle \theta^*, x_t \rangle)^2$, which

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.

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

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 less certain inputs in the future.

QuFUR measures the uncertainty of x_t using $\Delta_t := \tilde{\eta}^2 \min\{1, \|x_t\|_{M_t^{-1}}^2\}$, 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 7 that with high probability, the squared loss on x_t is at most $\tilde{O}(\Delta_t)$. QuFUR queries y_t with probability $\min\{1, \alpha \Delta_t\}$ where α is a parameter that tradeoffs query complexity and regret.

Perhaps surprisingly, the simple query strategy of QuFUR can leverage hidden domain structure:

Theorem 1 *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u : u \in [m]\}$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ lie in a subspace of dimension d_u. Suppose $\alpha \in \left[\frac{1}{\tilde{\eta}^2}\left(\frac{1}{\sum_u \sqrt{d_u T_u}}\right)^2, \frac{1}{\tilde{\eta}^2} \min_{u \in [m]} \frac{T_u}{d_u}\right]$. 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 is deferred to Section A.1. A few remarks:

Novel notion of adaptive regret The above tradeoff is novel; 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 $\mathrm{span}(\{x_t : t \in I\})$. This type of guarantee is stronger than the adaptive regret guarantees provided by e.g. Hazan and Seshadhri (2007), where the regret guarantee only holds for continguous intervals.

Matching uniform query 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 19 in Appendix A.6.

Matching oracle baseline. QuFUR matches the domain-aware baseline even without prior knowledge of domain structure. Theorem 3 says that both are optimal in our problem formulation (within a range of problem specifications).

Fixed-cost-ratio interpretation. 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 $Q \approx \alpha R$. We show in Appendix B that QuFUR with input $\alpha = c$ achieves near-optimal total cost, for a wide range of domain structure parameters.

3.2 QuFUR with a fixed label budget

The bounds 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 in A.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 $\{2^i/T^2 : i = 0, \ldots, k\}$. The grid ensures that each copy still has label budget $B/k = \tilde{\Omega}(B)$, 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 $\{I_u : u \in [m]\}$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ lie in a subspace of dimension d_u. Moreover, B satisfies $B \le \tilde{O}(\sum_u \sqrt{d_u T_u} \min_{u \in [m]} \sqrt{T_u/d_u})$.*

If Algorithm 2 receives inputs dimension d, time horizon T, label budget B, norm bound C, noise level η, then, with probability $1 - \delta$:
1. Its query complexity Q is at most B.
2. 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 compare this theorem with the guarantees of the oracle baseline in Section 2.2: 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.

3.3 Lower bound

We showed domain structure-aware regret upper bounds of the form $R = \tilde{O}(\tilde{\eta}^2 (\sum_u \sqrt{d_u T_u})^2 / B)$, 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). In this section, we show via Theorem 3 that they are tight up to constants, for a wide range of domain structure specifications. Its proof can be found in Appendix A.3.

Theorem 3 *For a set of positive integers $\{(d_u, T_u)\}_{u=1}^m$ and B such that $d_u \le T_u, \forall u \in [m]$, $\sum_{u=1}^m d_u \le d, B \ge \sum_{u=1}^m \sqrt{d_u T_u} / \sqrt{\min_{u \in [m]} T_u / d_u}$, there is an oblivious adversary such that:*
1. It uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \le \sqrt{d}$, and $|\langle \theta^*, x_t \rangle| \le 1$; the noises $\{\xi_t\}_{t=1}^T$ are sub-Gaussian with variance proxy η^2 for any $\eta \ge 1$.
2. It shows examples in m domains, where domain u has dimension d_u and time span T_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 refines the $\Omega(dT/B)$ minimax lower bound (Theorem 19 in Appendix A.6); it further constrains the adversary to present sequences of examples with domain structure specified by $\{d_u, T_u\}_{u=1}^m$. Theorem 3 subsumes the $\Omega(dT/B)$ lower bound by taking $m = 1, d_1 = d, T_1 = T$.

Combining Theorems 2 and 3, if $1 \le \eta \le O(1)$, for a wide range of $\{d_u, T_u\}_{u=1}^m$ 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*.

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

We generalize our algorithm to non-linear regression with adaptive adversaries, extending Russo and Van Roy (2013).

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 4 *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 $f_1, f_2 \in \mathcal{F}$, $\sqrt{\sum_{i=1}^n (f_1(x_i) - f_2(x_i))^2} \le \epsilon \implies f_1(x) - f_2(x) \le \epsilon$.*

Definition 5 *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.

The Algorithm The master algorithm, Algorithm 4 in Appendix A.4, runs $O(\log T)$ copies of Algorithm 3. Algorithm 3 queries with probability $\min\{1, \alpha\Delta_t\}$. At round t, Algorithm 3 predicts using the empirical risk minimizer $\hat{f}_t$ on all queried examples. 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 loss upper bound Δ_t is the squared maximum disagreement between two hypotheses in $\mathcal{F}_t$ on x_t. Therefore, 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 bound the regret on domain u with $R_u = \tilde{O}(\tilde{\eta}^2 d'_u \log \mathcal{N}(\mathcal{F}, T^{-2}, \|\cdot\|_\infty))$, where $\mathcal{N}(\mathcal{F}, \epsilon, \|\cdot\|_\infty)$ is the ϵ-*covering number* of $\mathcal{F}$ with respect to $\|\cdot\|_\infty$. We have the following theorem (proof is in Section A.4).

Theorem 6 *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ can be partitioned into m disjoint nonempty subsets $\{I_u : u \in [m]\}$, where for each u, $|I_u| = T_u$, and $\{x_t\}_{t \in I_u}$ all lie in $\mathcal{X}_u$. Given budget $B \le \tilde{O}(\sum_u \sqrt{R_u T_u} \min_u \sqrt{R_u/T_u})$, Algorithm 4 satisfies: (1) $Q \le B$; (2) with probability $1 - \delta$, $R = \tilde{O}((\sum_u \sqrt{R_u T_u})^2/B)$.*

5. Experiments

We test the query-regret tradeoffs of QuFUR, uniform queries (Section 2.2), and greedy queries (i.e., always querying until budget is exhausted) on a synthetic dataset, and LIBSVM datasets cpu-small and Abalone (Chang and Lin, 2011). QuFUR achieves the lowest total regret under the same labeling budget. See Appendix E for more details.

Appendix A. Missing proofs

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. First some notations. Let the interaction history between the learner and the environment up to time t be H_t and $\mathbb{E}_{t-1}[\cdot] := \mathbb{E}[\cdot|x_t, H_{t-1}]$.

The following lemma upper bounds the regret in a subdomain with sum of uncertainty estimates $\Delta_t = \tilde{\eta}^2 \min\left\{1, \|x_t\|_{M_t^{-1}}\right\}$.

Lemma 7 *In the setting of Theorem 1, Fix $I \subseteq [T]$. Then with probability $1 - \frac{|I|\delta}{2T}$, the regret in $I \subseteq [T]$ is upper bounded by $\sum_{t\in I}(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 = \tilde{O}(\sum_{t\in I}\Delta_t)$.*

Proof [Proof of Lemma 7]

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

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

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

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

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

Meanwhile, the second term of Equation (1) can be bounded by Cauchy-Schwarz:

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

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

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

Since $\hat{y}_t = \text{clip}(\langle x_t, \hat{\theta}_t\rangle)$ and $\left|\langle x_t, \theta^*\rangle\right| \leq 1$,

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

Therefore, on event E, we have $\sum_{t\in I}(\hat{y}_t - \langle\theta^*, x_t\rangle)^2 = \tilde{O}(\sum_{t\in I}\Delta_t)$. ■

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

Lemma 8 *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 [Proof of Lemma 8] 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,

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

where the second inequality follows from e.g. Lemma 19.4 of Lattimore and Szepesvári (2018). ■

Proof [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 $u \in [m]$, define event

$$F_u = \left\{\left|\sum_{t\in I_u} p_t\Delta_t - \sum_{t\in I_u} q_t\Delta_t\right| \le O\left(\tilde{\eta}\sqrt{\sum_{t\in I_u} p_t\Delta_t \ln\frac{T}{\delta'}} + \tilde{\eta}^2\ln\frac{T}{\delta'}\right)\right\}. \tag{2}$$

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

Similarly, define

$$G_u = \left\{\left|\sum_{t=1}^T p_t - \sum_{t=1}^T q_t\right| \le O\left(\sqrt{\sum_{t=1}^T p_t \ln\frac{T}{\delta'}} + \ln\frac{T}{\delta'}\right)\right\}. \tag{3}$$

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

Furthermore, define $H = \cap_{u=1}^m (E_{I_u} \cap F_u \cap G_u))$, where E_{I_u} is the event defined in the proof of Lemma 7 on subset I_u. 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 (2), we get that

$$\sum_{t\in I_u} p_t\Delta_t = \tilde{O}\left(\sum_{t\in I_u} q_t\Delta_t + \tilde{\eta}^2\right). \tag{4}$$

Using Lemma 8 with $\{a_i\}_{i=1}^k = \mathcal{Q}_T$, and $S = I_u \cap \mathcal{Q}_T$, we get that

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

In combination with Equation 4, 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 vs. low risk subsets $I_{u,+}$ and $I_{u,-}$ (abbrev. I_+ and I_-). Formally,

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

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

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

Summing over the two cases, we have

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

If $\alpha \le \frac{1}{\tilde{\eta}^2} \min_u \frac{T_u}{d_u}$, for every u, we have, $\alpha \tilde{\eta}^2 d_u \le \tilde{\eta} \sqrt{\alpha d_u T_u}$. This implies that

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

For the query complexity, from the definition of event G_t, applying AM-GM inequality on Equation (3), we also have

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

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

For the regret guarantee, for every $u \in [m]$, Lemma 7 with $I = I_u$ implies that

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

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 $k \leftarrow 4\lceil \log_2 T \rceil$.
for $i = 0$ to k **do**
  Parameter $\alpha_i \leftarrow 2^i / T^2$.
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 $\hat{f}_t(x) = \mathrm{clip}(\langle \hat{\theta}_t, x \rangle)$ 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. $M \leftarrow M + x_t x_t^\top$, $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

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

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

The theorem follows. ■

A.2 Proof of Theorem 2

Before going into the proof, we set up some useful notations. Define $I = \{0, 1, \ldots, \lceil \log T \rceil\}$ as the index set of the sub-algorithms. Recall the number of copies $k = 1 + \lceil 4 \log T \rceil \leq 2 + 4 \log T$. Recall also that $B' = B/k$ 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; $r_t^i \sim \mathrm{Bernoulli}(p_t^i)$ be the attempted query decision of copy i at time step t; $A_t^i = \mathbb{1}\left[\sum_{j=1}^{t-1} r_j^i < B'\right]$, 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 9 *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 $j \in J$, $A_j^i = 1$.

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

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 one, 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 upper bounds $\sum_{t=1}^T A_t^i r_t^i$ for all i's beyond the above trivial B' bound.

Lemma 10 *There exists a constant $C = \mathrm{polylog}(T, \frac{1}{\delta})$, such that with probability $1 - \delta$,*

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

for every $i \in \{\}$ 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 $1 - \delta/4$,

$$\sum_{t=1}^T A_t^i r_t^i = \tilde{O}\left(\sum_{t=1}^T A_t^i p_t^i + 1 \right). \tag{5}$$

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 $1 - \delta'$,

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

Using Lemma 8 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 $1 - \delta'$,

$$\sum_{t \in I_u} A_t^i p_t^i \Delta_t = \tilde{O}(\tilde{\eta}^2 d_I). \tag{6}$$

Let $I_+ = \{j \in I_u, \alpha_i \Delta_j > 1\}$, and $I_- = I_u - I_+$.

1. In I_+, by Equation (6), $\sum_{j \in I_+} A_j^i \Delta_j = \tilde{O}(\tilde{\eta}^2 d_u) \implies \sum_{j \in I_+} A_j^i p_j^i = \tilde{O}(\alpha_i \tilde{\eta}^2 d_u)$.

2. In I_-, by Equation (6), with probability $1-\delta'$, $\sum_{j\in I_-} A_j^i \alpha_i \Delta_j^2 = \sum_{j \in I_-} A_j^i p_j \Delta_j = \tilde{O}(\tilde{\eta}^2 d_u)$; this implies $\sum_{j\in I_-} A_j^i \Delta_j = \tilde{O}(\tilde{\eta}\sqrt{d_u T_u/\alpha_i})$. If this event happens, 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

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

If $\alpha_i \le \frac{1}{\tilde{\eta}^2}\min_u \frac{T_u}{d_u}$, for every u, we have, $\alpha_i \tilde{\eta}^2 d_u \le \tilde{\eta}\sqrt{\alpha d_u T_u}$. This implies that

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

Therefore, using Equation (5), we have

$$\sum_{t\in I_u} A_t^i q_t^i \le \tilde{O}\left(\sum_{t=1}^T A_t^i p_t^i + 1\right) \le \tilde{O}(\tilde{\eta}\sqrt{\alpha_i d_u T_u} + 1) \le \tilde{O}(\tilde{\eta}\sqrt{\alpha_i d_u T_u}),$$

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

We are now ready to prove Theorem 2.

Proof [Proof of Theorem 2] First, the query complexity of Fixed-Budget QuFUR is B by construction, as there are k copies running, and each copy consumes at most B' labels.

We now bound the regret of Fixed-Budget QuFUR. If $B \le 1$, the regret of the algorithm is trivially upper bounded by $4T$, which is $O((\sum_{u=1}^m \sqrt{d_u T_u})^2/B)$.

Thus throughout the rest of the proof, we consider $B \in [1, \tilde{O}(\sum_u \sqrt{d_u T_u}\min_{u\in[m]}\sqrt{T_u/d_u})]$.

Denote by

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

By the construction of i_B, and the assumption that $B \in [1, \sum_u \sqrt{d_u T_u}\min_{u\in[m]}\sqrt{T_u/d_u}]$, we have

$$\frac{B'}{2} \le C\tilde{\eta}\sqrt{\alpha_{i_B}}\sum_{u=1}^m \sqrt{d_u T_u} \le B'.$$

This implies that,

$$\alpha_{i_B} \in \left[\left(\frac{B'}{2C\tilde{\eta}\sum_u \sqrt{d_u T_u}}\right)^2, \left(\frac{B'}{C\tilde{\eta}\sum_u \sqrt{d_u T_u}}\right)^2\right]. \tag{7}$$

Again by our assumption on B, we deduce that

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

Therefore, the premises of Lemma 10 is satisfied for $i = i_B$; this gives that with probability $1 - \delta$:

$$\sum_{t=1}^{T} A_t^{i_B} \Delta_t \le C \cdot \tilde{\eta} \sum_u \sqrt{d_u T_u} / \sqrt{\alpha_{i_B}}, \tag{8}$$

and

$$\sum_{t=1}^{T} A_t^i r_t^{i_B} \le C \cdot \tilde{\eta} \sqrt{\alpha_{i_B}} \sum_u \sqrt{d_u T_u}. \tag{9}$$

Now from Equation (9) and the definition of i_B, we have

$$\sum_{t=1}^{T} A_t^i r_t^{i_B} \le C \cdot \tilde{\eta} \sqrt{\alpha_{i_B}} \sum_u \sqrt{d_u T_u} < B.$$

From Lemma 9, we know that for all t in $[T]$, $A_t^{i_B} = 1$. Plugging this back to Equation (8), we have

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

where the second inequality is from the lower bound of α_{i_B} in Equation (7).

Combining the above observation with Lemma 7 on $I = [T]$ gives that

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

■

A.3 Proof of Theorem 3

Proof 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 task u are shown; specifically, 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.

We first choose θ^* from distribution D_θ, such that for every coordinate $j \in [d]$, $\theta_i^* \sim \mathrm{Beta}(1, 1)$. 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 $t \in I_{u,i}$, $y_t \sim \mathrm{Bernoulli}(\theta_{u,i}^*)$. Denote by $N_{u,i}(t) = \sum_{s \in I_{u,i}: s \le t} q_s$ 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

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

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

$$\sum_{u=1}^{m} \sum_{i=1}^{d_u} \mathbb{E}\left[(N_{u,i}(T) + 2) \mid \theta^*\right] \le B + 2\sum_{u=1}^{m} d_u \le 3B. \tag{10}$$

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

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

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.[1] We note that $\hat{y}_t$ is $\mathcal{F}_{t-1}$-measurable. Denote by $N_{u,i}^{+}(t) = \sum_{s \in I_{u,i}: s \le t} q_s \cdot \mathbb{1}(y_s = 1)$, 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,

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

Summing over all time steps, we have

$$\mathrm{Reg}(T) \ge \mathbb{E}\left[\sum_{u=1}^{m} \sum_{i=1}^{d_u} \sum_{t \in I_{u,i}} (\hat{y}_t^* - \theta_{u,i}^*)^2\right].$$

1. We use $y_s q_s$ to indicate the labeled data information acquired at time step s; if $q_s = 1$, $y_s q_s = y_s$, in which case the learner has access to label y_s; otherwise $q_s = 0$, $y_s q_s = 0$, in which case the learner does not have label y_s available.

On the other hand, from Lemma 11, 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 $(u,i) : u \in [m], i \in [d_u]$, we have

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Reg}(T) \mid \theta^*\right] &= \sum_{u=1}^{m}\sum_{i=1}^{d_u}\sum_{t \in I_{u,i}} \mathbb{E}\left[(\hat{y}_t - \theta_{u,i})^2 \mid \theta^*\right] \\
&\geq \sum_{u=1}^{m}\sum_{i=1}^{d_u} \frac{\lfloor T_u/d_u \rfloor \cdot f(\theta^*_{u,i})}{2(\mathbb{E}\left[N_{u,i}(T) \mid \theta^*\right]+2)} \\
&\geq \sum_{u=1}^{m}\sum_{i=1}^{d_u} \frac{T_u/d_u \cdot f(\theta^*_{u,i})}{4(\mathbb{E}\left[N_{u,i}(T) \mid \theta^*\right]+2)} \qquad (11)
\end{aligned}$$

Combining the above inequality with Equation (10), we have:

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

where the second inequality is from Cauchy-Schwarz. Now taking expectation over θ, using Jensen's inequality and Lemma 12 that $\mathbb{E}\sqrt{f(\theta^*_{u,i})} \geq \frac{1}{25}$, and some algebra yields

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

In conclusion, we have

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

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

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

holds. This θ^* has ℓ_2 norm at most $\sqrt{\sum_{j=1}^d (\theta_j^*)^2} \le \sqrt{d}$. ■

Lemma 11 *If t is in $I_{u,i}$, then*

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

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

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 $\mathrm{Bin}(m, \theta_{u,i})$. Therefore,

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

■

Lemma 12 *Suppose $Z \sim \mathrm{Beta}(1,1)$. Then* $\mathbb{E}\sqrt{f(Z)} \ge \frac{1}{25}$.

Proof We observe that

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

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

A.4 Proof of Theorem 6

First, we clarify that in the setting of Theorem 6, we require that the partition $\{I_u : u \in [m]\}$ to have a properties that we call *admissibility*.

Definition 13 *The partition $\{I_u : u \in [m]\}$ is called* admissible*, if $u_t = \{u : t \in I_u\}$ is only dependent on the interaction history up to $t-1$ and unlabeled example x_t; formally, u_t is $\sigma(H_{t-1}, x_t)$-measurable.*

Recall that we define

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

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 14 *Suppose the example sequence $\{x_t\}_{t=1}^T$ has the following structure: $[T]$ has an admissible partition $\left\{I_u : u \in [m]\right\}$, 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 6 directly below; the proof of Theorem 14 follows as a corollary, using the same argument in the proof of Theorem 2; we note that the admissibility assumption on $\left\{I_u : u \in [m]\right\}$ 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 [Proof of Theorem 6] We focus on proving the analogues of Lemma 7 and Lemma 8; and the rest of proof follows the same argument as the proof of Theorem 2.

Lemma 15 (Analogue of Lemma 7) *With probability $1-\delta/2$, $R \leq \sum_{t=1}^T \Delta_t$.*

Proof Recall that the confidence set at time t is $\mathcal{F}_t = \{f \in \mathcal{F} : \sum_{i\in\mathcal{Q}_t}(f(x_i) - \hat{f}_t(x_i))^2 \leq \beta_{|\mathcal{Q}_t|}(\mathcal{F}, \delta)\}$. By Russo and 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 $t \in [T]$, $(\hat{f}_t(x_t) - f^*(x_t)) \leq \sup_{f_1, f_2\in\mathcal{F}_t}(f_1(x_t) - f_2(x_t))^2 = \Delta_t$. This implies that the regret is bounded by $R \leq \sum_{t=1}^T \Delta_t$. ■

Lemma 16 (Analogue of Lemma 8) $\sum_{t\in I_u} q_t\Delta_t \leq R_u$.

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

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

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

We know for all $j \in [k]$, $D_j \leq 4$. In addition, $D_j > \epsilon^2 \iff \sum_{t\in I_u\cap\mathcal{Q}_T} \mathbb{1}[\Delta_t > \epsilon^2] \geq j$. By Lemma 17 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 $D_j > 1/T^4$, $D_j \leq \min\left\{4, \frac{4\beta_T d}{j-d}\right\}$. Therefore $\sum_j D_j\mathbb{1}[D_j > 1/T^4] \leq 4d + \sum_{j=d+1}^k \frac{4\beta_T d}{j-d} \leq 4d + 4d\beta_T \log T_u$.

Consequently,

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

■

■

The following lemma generalizes Russo and Van Roy (2013, Proposition 3), in that it considers a subsequence of examples coming from a subdomains 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 17 *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 $k = |I \cap \mathcal{Q}_T|$, $(a_1, \dots, a_k) = (x_t : t \in I \cap \mathcal{Q}_T)$, and $(b_1, \dots, b_k) = (\Delta_t : t \in I \cap \mathcal{Q}_T)$. 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 $(a_1, \dots, a_{j-1})$, 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 $(a_{i_1}, \dots, a_{i_p})$ of $(a_1, \dots, a_{j-1})$, 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 $(a_1, \dots, a_{j-1})$, then $\sum_{i\in\mathcal{Q}_t} (f_1(x_i) - f_2(x_i))^2 > K\epsilon^2$. By the triangle inequality,

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

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

Next, we show that in any sequence of elements in I, $(c_1, \dots, c_\tau)$, there is some c_j that is ϵ-dependent on at least $\tau/d - 1$ disjoint subsequences of $(c_1, \dots, c_{j-1})$, 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 $C_1, \dots, C_K$. First let $C_i = (c_i)$ for $i \in [K]$. If c_{K+1} is ϵ-dependent on $C_1, \dots, C_K$, 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 $(c_1, \dots, c_\tau)$ to be the subsequence $(a_{t_1}, \dots, a_{t_\tau})$ of $(a_1, \dots, a_k)$ 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 $(a_1, \dots, a_{t_j-1})$. Thus, each c_j is ϵ-dependent on fewer than $4\beta_T/\epsilon^2$ disjoint subsequences of $(c_1, \dots, c_{j-1})$. Combining this with the fact that there is some c_j that is ϵ-dependent on at least $\tau/d - 1$ disjoint subsequences of $(c_1, \dots, c_{j-1})$, we have $\tau/d - 1 < 4\beta_T/\epsilon^2$. Thus, $\tau < (4\beta_T/\epsilon^2 + 1)d$. ■

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 $\hat{f}_t \leftarrow \operatorname{argmin}_{f\in\mathcal{F}} \sum_{i\in\mathcal{Q}} (f(x_i) - y_i)^2$.
  $\mathcal{F}_t \leftarrow \{f \in \mathcal{F} : \sum_{i\in\mathcal{Q}} (f(x_i) - \hat{f}_t(x_i))^2 \le \beta_{|\mathcal{Q}|}(\mathcal{F}, \delta)\}$,
where $\beta_k := 8\eta^2 \log (4\mathcal{N}(\mathcal{F}, 1/T^2, \|\cdot\|_\infty)/\delta) + 2k/T^2(16 + \sqrt{2\eta^2 \ln (16k^2/\delta)})$.
  $\Delta_t = \sup_{f_1,f_2\in\mathcal{F}_t} \left|f_1(x_t) - f_2(x_t)\right|^2$.
  With probability $\min\{1, \alpha\Delta_t\}$, set $q_t = 1$; otherwise $q_t = 0$.
  if $q_t = 1$ **then**
    Query y_t. $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

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. $\mathcal{Q} \leftarrow \mathcal{Q} \bigcup \{t\}$.

A.5 Analysis of uniform query strategy for online active linear regression with oblivious adversary

Theorem 18 *With probability $1 - \delta$, the uniformly querying strategy with probability μ achieves $R = \tilde{O}(\tilde{\eta}^2 d/\mu)$ and $Q = O(\mu T + 1)$.*

Proof [Proof sketch] Let $\delta' = \delta/3$.

As $Q = \sum_{t=1}^{T} q_t$ is a sum of T independent Bernoulli random variable with mean μ, by Chernoff bound, $Q = O(\mu T + \ln \frac{1}{\delta})$ with probability $1 - \delta'$.

We still define $\Delta_t = \tilde{\eta}^2 \min\{1, \|x_t\|^2_{M_t^{-1}}\}$.

Using Lemma 8 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 $1 - \delta'$,

$$\sum_{t=1}^{T} \mu\Delta_t - \sum_{t=1}^{T} q_t\Delta_t = O\left(\tilde{\eta}\sqrt{\sum_{t=1}^{T} \mu\Delta_t \ln(\ln T/\delta') + \tilde{\eta}^2 \ln(\ln T/\delta'))}\right)$$
$$\implies \sum_t \Delta_t = \tilde{O}(\tilde{\eta}^2 d/\mu).$$

In addition, applying Lemma 7, with probability $1 - \delta'$,

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

Union bound over the three events above completes the proof. ■

A.6 Lower bound for unstructured domains

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

Theorem 19 *For any set of positive integers d, T, B such that $d \le T$ and $d \le B$, there exists an oblivious adversary such that:*

1. *it uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \le \sqrt{d}$, and $\big|\langle\theta^*, x_t\rangle\big| \le 1$.*
2. *any online active learning algorithm $\mathcal{A}$ with label budget B has regret at least $\frac{dT}{7500B}$.*

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. ■

Appendix B. 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 20 *For any set of positive integers $\big\{(d_u, T_u)\big\}_{u=1}^m$ such that $d_u \le T_u, \forall u \in [m]$; $\sum_{u=1}^m d_u \le d$; cost ratio $c \ge \max_u \frac{d_u}{T_u}$; real number $\eta \ge 1$; there exists an oblivious adversary such that:*

1. *it uses a ground truth linear predictor $\theta^\star \in \mathbb{R}^d$ such that $\|\theta^*\|_2 \le \sqrt{d}$, and $\big|\langle\theta^*, x_t\rangle\big| \le 1$; in addition, the subgaussian variance proxy of noise is η^2.*
2. *it shows examples in m tasks, where the examples from task u has dimension d_u, and the duration of task u is T_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 21 *Suppose $1 \leq \eta \leq O(1)$; 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 $\max_{t \in [T]} \left|\langle \theta^*, x_t \rangle\right| \leq 1$, $\mathrm{QuFUR}(c)$ (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 parameters* $\left\{(T_u, d_u)\right\}_{u=1}^{m}$ *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 $c \leq \tilde{\eta}^2 \min_u \frac{T_u}{d_u}$, applying Theorem 1, we have that $\mathrm{QuFUR}(c)$ achieves the following regret and query complexity guarantees:

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

   This implies that

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

2. By the premise that $c \geq \max_u \frac{d_u}{T_u}$, applying Theorem 20, we get the item.

■

Appendix C. Clarification of 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^*: $\mathrm{Reg} = \sum_{t=1}^{T}(\hat{y}_t - y_t)^2 - \sum_{t=1}^{T}(f^*(x_t) - y_t)^2$.

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

Theorem 22 $\mathbb{E}[R] = \mathbb{E}[\mathrm{Reg}]$.

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. ■

Appendix D. Online to batch conversion

In this section we show a straightforward result on online to batch conversion in active learning setting.

Theorem 23 *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,*

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

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

Combining Theorem 23 with Theorem 2, we have the following theorem on fixed-budget QuFUR (Algorithm 2) when run on iid data with domain structure:

Theorem 24 *Suppose the unlabeled data distribution D_X - a mixture $D_X = \sum_{u=1}^m p_u D_u$, where D_u is supported on a subspace of $\mathbb{R}^d$ of dimension d_u - intersects with the set $\left\{x : \|x\|_2 \le 1, \left|\langle\theta^*, x\rangle\right| \le 1\right\}$. 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 integers B, T_0 satisfy $T_0 \ge \Omega\left(\max\left(\frac{B}{\sum_u \sqrt{d_u p_u}\cdot \min_u \sqrt{\frac{p_u}{d_u}}}, \frac{\ln m}{\min_u p_u}\right)\right)$.*
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, with probability $1-\delta$:
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,

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

Proof [Proof sketch] From Theorem 23 it suffices to show that

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

By Theorem 22, $\mathbb{E}[\mathrm{Reg}] = \mathbb{E}[R]$, it therefore reduces to showing that

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

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.

First, from the assumption that $T \ge T_0 \ge \Omega(\frac{1}{\min_u p_u})$, we have that by Chernoff bound, with probability $1 - \frac{1}{T^2}$, for all u, $T_u \in [Tp_u/2, 2Tp_u]$. We call the set of unlabeled examples *good*.

Conditioned on a set of good unlabeled examples, we have that

$$B \le \tilde{O}(T \cdot \sum_u \sqrt{d_u p_u} \min_u \sqrt{p_u/d_u}) \le \tilde{O}(\sum_u \sqrt{d_u T_u} \min_u \sqrt{T_u/d_u}).$$

Therefore, applying Theorem 2, we have that conditioned on a good sample, with probability $1 - \frac{1}{T^2}$ over the draw of $(y_t)_{t=1}^T$,

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

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

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

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,

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

The theorem follows. ■

Appendix E. Additional experimental details

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 and 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.

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

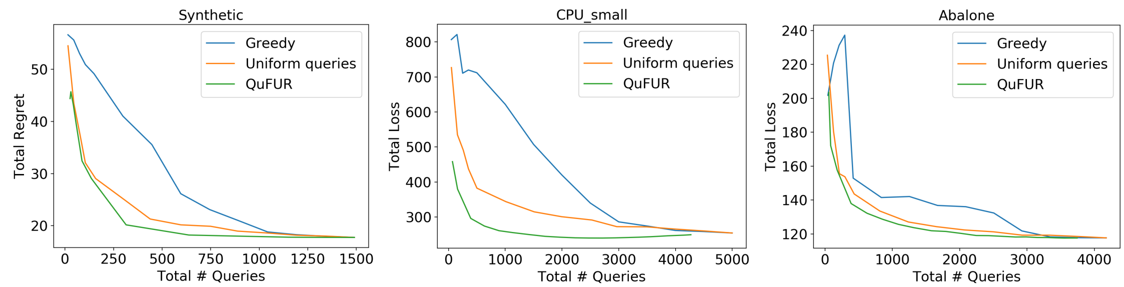

Figure 1: Total regret vs. total number of queries in synthetic dataset (**left**), cpu-small dataset (**middle**), and Abalone dataset (**right**), averaged across 5 runs. QuFUR is best and has more advantage on cpu-small potentially due to latent domain structure, whereas Abalone is more homogeneous.

References

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

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

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

Chih-Chung Chang and Chih-Jen Lin. 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`.

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

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

Tor Lattimore and Csaba Szepesvári. Bandit algorithms. *preprint*, page 28, 2018.

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

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

Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2016.