

000 001 002 003 004 005 A REVISIT OF ACTIVE SEQUENTIAL PREDICTION- 006 POWERED MEAN ESTIMATION 007 008 009

010 **Anonymous authors**
011 Paper under double-blind review
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026

ABSTRACT

027 In this work, we revisit the problem of active sequential prediction-powered mean
028 estimation, where at each round one must decide the query probability of the
029 ground-truth label upon observing the covariates of a sample. Furthermore, if
030 the label is not queried, the prediction from a machine learning model is used
031 instead. Prior work proposed an elegant scheme that determines the query prob-
032 ability by combining an uncertainty-based suggestion with a constant probability
033 that encodes a soft constraint on the query probability. We explored different val-
034 ues of the mixing parameter and observed an intriguing empirical pattern: the
035 smallest confidence width tends to occur when the weight on the constant prob-
036 ability is close to one, thereby reducing the influence of the uncertainty-based
037 component. Motivated by this observation, we develop a non-asymptotic analysis
038 of the estimator and establish a data-dependent bound on its confidence interval.
039 Our analysis further suggests that when a no-regret learning approach is used to
040 determine the query probability and control this bound, the query probability con-
041 verges to the constraint of the max value of the query probability when it is chosen
042 obliviously to the current covariates. We also conduct simulations that corroborate
043 these theoretical findings.
044

1 INTRODUCTION

045 The mean estimation problem is a classical inference task that has seen revived interest in machine
046 learning and statistics over the last few years. While the conventional setting is well-understood,
047 numerous works have explored this problem under diverse settings and assumptions, aiming to en-
048 hance our understanding of the inherent challenges of learning from limited data. Recently, a line of
049 work has investigated the design of efficient mean estimators under the robust framework, including
050 the setting of a constant fraction of adversarial outliers (Cheng et al., 2020), heavy-tailed symmetric
051 distributions without moment assumptions (Novikov et al., 2023), mean-shift contamination in mul-
052 tivariate identity Gaussian distributions (Diakonikolas et al., 2025), sparse mean estimation in high
053 dimensions (Pensia, 2024), online high-dimensional mean estimation (Kane et al., 2024). A few
054 other recent works have explored other structural considerations, such as collaborative normal mean
055 estimation in the presence of strategic agents (Chen et al., 2023), communication-efficient mean
056 estimation in a distributed setting (Ben-Basat et al., 2024), vector mean estimation under the shuf-
057 fle model of privacy (Asi et al., 2024), dynamic multi-group mean estimation (Aznag et al., 2023),
058 leveraging favorable distribution structure to improve sub-Gaussian rate (Dang et al., 2023), among
059 others. This variety of scopes in the mean estimation setup highlights its relevance as a foundational
060 task for inference.

061 A direction that has attracted substantial attention with the increasing integration of machine learning
062 is mean estimation through an active inference perspective (Zrnic & Candes, 2024). Specifically,
063 the problem focuses on estimating the mean label from a set of unlabeled observations, by leveraging
064 a limited label collection budget and the abundant but potentially biased predictions of a machine
065 learning model. Under this setup, active statistical inference provides a data collection strategy that
066 utilizes the budget more effectively by taking into consideration which labels it would be more ben-
067 efitial to acquire. Specifically, it prioritizes the collection of labels where the model exhibits higher
068 uncertainty and uses the model predictions for instances where the model is more confident. The
069 same work considers querying the labels in both the batch and sequential setting, where the latter
070 additionally allows updating the model as the ground-truth labels are obtained, and provides *asym-*

054 *totically* valid confidence intervals for the estimator in question in both cases. The non-asymptotic
 055 analysis of the estimator was not provided in the prior work.¹
 056

057 In this work, we draw on the sequential active statistical inference perspective by providing *non-*
 058 *asymptotic* guarantees for the sequential active mean estimation problem. While prior work of Zrnic
 059 & Candes (2024) have established the asymptotic normality of their proposed estimator, our work
 060 investigates the sequential mean estimation problem further under the light of an online updating
 061 scheme and provides a non-asymptotic analysis with guarantees that hold *at any time* while going
 062 through the data. More specifically, we first formulate the scheme of active sequential mean estima-
 063 tion as an online update step, and establish a convergence guarantee that incorporates the conditional
 064 variance of the update direction and achieves a rate of $\tilde{O}(1/\sqrt{t})$ for sufficiently large t .
 065

066 Furthermore, motivated by a series of experimental findings, which reveal an intriguing pattern of
 067 the label sampling rule considered by previous work, we are led to examine more closely the role of
 068 the model uncertainty component through the current covariate. In particular, to adhere to the budget
 069 constraint and ensure a small variance of the estimator, Zrnic & Candes (2024) derive a sampling
 070 rule that is a mixture of the uniform rule and a model uncertainty estimator weighted by a mixing
 071 constant. Our experimental findings across a variety of tasks indicates that employing the mixture
 072 rule or relying solely on the uniform policy results in comparable confidence interval widths, with
 073 the uniform policy occasionally producing marginally narrower intervals. This observation suggests
 074 that, the contribution of model uncertainty with respect to the label of the current covariate might
 075 be brittle in practice. Based on this insight, we formulate the problem of tuning the query policy as
 076 an online learning task that does not rely on the current covariate, and whose validity is supported
 077 by the strong sublinear regret guarantees of the classical Follow-the-Regularized-Leader (FTRL)
 078 algorithm (Abernethy et al., 2012). Remarkably, we demonstrate that under this no-regret learning
 079 approach, the query policy converges to the maximum value permitted by the budget constraint. Our
 080 theoretical findings are further validated through experiments on three real-world and one synthetic
 081 dataset.
 082

083 2 PRELIMINARIES AND NOTATION

084 We begin by introducing the problem setup. We consider the setting where we have access to a
 085 sequence of data points $x_1, x_2, \dots, x_T \in \mathcal{X} \subset \mathbb{R}$ from an unknown, fixed distribution \mathbb{P}_X . Each data
 086 point x_t is associated with a ground-truth label $y_t \in \mathcal{Y} \subset \mathbb{R}$, drawn from an also unknown, fixed
 087 distribution $\mathbb{P}_{Y|X}$. We assume that the ground-truth labels are not known a priori, and the cost to
 088 obtain them could be high. We are interested in estimating the mean label $\mu_y = \mathbb{E}[y_t]$. Additionally,
 089 we assume that at each time $t \in [T]$ we have access to a black-box predictive model $f_t(\cdot) : \mathcal{X} \rightarrow \mathcal{Y}$,
 090 which can continually evolve by using the samples collected up to round $t-1$ to update. More
 091 specifically, we require $f_t \in \mathcal{F}_{t-1}$, where \mathcal{F}_t denotes the σ -algebra generated by the first t data
 092 points x_s , $1 \leq s \leq t$.
 093

Sequential active mean estimation seeks to construct an efficient estimator by observing data points
 one at a time and deciding whether to query each ground-truth label. Under a limited labeling
 budget, the objective is to sequentially acquire labels in a way that most effectively improves the
 accuracy of the mean estimator. Specifically, if we denote by T_{lab} the total number of collected
 labels, we require that the policy of querying the ground-truth label ensures $\mathbb{E}[T_{lab}] \leq T_b$, where we
 assume that $T_b \ll T$. Let $\pi_t(x_t)$ denote the probability of collecting the label of data point x_t at time
 t , where $\pi_t \in \mathcal{F}_{t-1}$, and let $\xi_t \sim \text{Bernoulli}(\pi_t(x_t))$ denote the labeling decision used to indicate
 whether the ground-truth label y_t was collected ($\xi_t = 1$) or not ($\xi_t = 0$). Zrnic & Candes (2024)
 propose the sequential active mean estimator $\hat{w} = \frac{1}{T} \sum_{t=1}^T \left(f_t(x_t) + (y_t - f_t(x_t)) \frac{\xi_t}{\pi_t(x_t)} \right)$. It is
 easy to verify that \hat{w} is unbiased, i.e., $\mathbb{E}[\hat{w}] = \mu_y$. Notably, Zrnic & Candes (2024) show that the
 optimal choice of querying policy π_t^{opt} is according to the uncertainty of the prediction model, and
 104
 105

106 ¹However, we acknowledge that Appendix C of Zrnic & Candes (2024) outlines a related scheme based
 107 on estimating bounded means via testing by betting (Waudby-Smith & Ramdas, 2024), which may come with
 108 certain non-asymptotic guarantees. We further provide a discussion in Appendix B.

108 it satisfies

$$110 \pi_t^{opt}(x_t) \propto \sqrt{\mathbb{E}[(y_t - f_t(x_t))^2 | \mathcal{F}_{t-1}]},$$

112 where the above expression hides a normalization constant to ensure that $\mathbb{E}[\pi_t^{opt}(x_t)] \leq T_b/T$.
113 However, since $\mathbb{P}_{Y|X}$ is unknown, the authors suggest fitting a model on past data (x, y) to approxi-
114 mate the uncertainty $u_t(x_t)$ by $|y_t - f_t(x_t)|$ for a given x_t , and then setting the querying policy to be
115 proportional to that estimate. To ensure that the budget constraint is met in practice, the remaining
116 budget at time t is set as the difference between the expected budget to be used up to time t and the
117 budget already used up to time $t-1$, i.e. $T_{\Delta,t} = tT_b/T - T_{lab,t-1}$. Then, the querying policy is set
118 to

$$119 \pi_t(x_t) = \min \{\eta_t u_t(x_t), T_{\Delta,t}\}_{[0,1]},$$

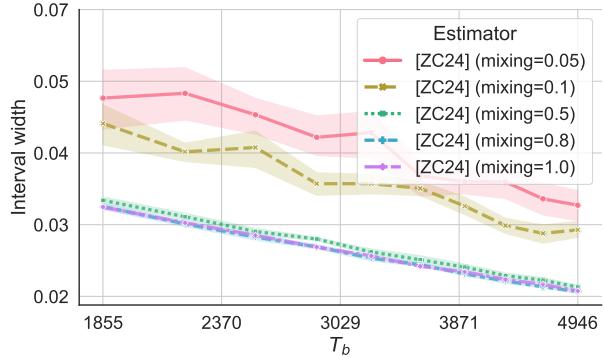
121 where the subscript $[0, 1]$ denotes clipping to $[0, 1]$, η_t is a normalizing constant set to $\eta_t =$
122 $T_b / (T \mathbb{E}[u_t(x_t)])$, and $\mathbb{E}[u_t(x_t)]$ is approximated empirically. This practical rule aims to balance
123 frequent sampling under high uncertainty against overusing the budget. However, since the esti-
124 mated uncertainty can be consistently low, and thus the budget could be underutilized, the policy
125 is occasionally set to $\pi_t(x_t) = (T_{\Delta,t})_{[0,1]}$. An additional concern is that an inaccurate uncertainty
126 estimate may be close to zero, while the quantity $|y_t - f_t(x_t)|$ is actually large, which would in turn
127 yield an amplified estimator variance. To address this issue, the authors suggest that the policy is set
128 to a mix of the described policy with the uniform rule as

$$129 \pi_t^{(\lambda)}(x_t) = (1 - \lambda)\pi_t(x_t) + \lambda\pi_t^{\text{unif}}(x_t),$$

131 where $\pi_t^{\text{unif}}(x_t) = T_b/T$, and $\lambda \in [0, 1]$. In the presence of sufficient historical data, λ can be tuned
132 by minimizing the empirical estimate of the estimator variance induced by the policy $\pi_t^{(\lambda)}$. However,
133 due to insufficient historical data in the sequential setting, λ is set to a fixed value.

135 While the related work simply sets the mixing parameter to 0.5, we explore different parameter values
136 for the mean estimation experiment in Zrnic & Candes (2024) by running their public implementation on
137 the same post-election survey dataset Center (2020) as in their experiments.
138 We keep all their other parameter choices unchanged. Figure 1 shows how the interval width varies with
139 the sampling budget T_b under different values of the mixing parameter in the query policy ². We find
140 that setting $\lambda = 1$, corresponding to using the uniform query policy that ignores model uncertainty,
141 produces confidence intervals that are slightly narrower than those obtained with $\lambda = 0.5$. We further eval-
142 uate the method on two additional real-world datasets and one synthetic dataset, and observe a similar
143 pattern, where the uniform query policy ($\lambda = 1$) yields confidence intervals that are comparable to,
144 and often tighter than those from $\lambda = 0.5$. Due to space constraints, the corresponding figures are
145 included in Appendix A.

146 These similar empirical patterns motivate us to further investigate this observation. A plausible
147 initial explanation is that the empirical performance improves when less weight is assigned to
148 the uncertainty-based component, possibly because the uncertainty predictor may not be reliable.
149 However, our theoretical analysis later suggests that this factor alone may not account for the phe-
150 nomenon, which might be surprising.



151 **Figure 1: Post-election survey dataset.** Interval width vs.
152 the sampling budget parameter T_b for different values of the
153 mixing parameter of the query probability scheme in Zrnic
154 & Candes (2024). Averaged over 10 repeated runs.

155 ²The code, provided in Jupyter notebook format, to reproduce Figure 1 as well as others in the paper is
156 available in the supplementary material.

162 **3 RELATED WORK**

164 **Active Statistical Inference.** The ideas in this work are motivated by the recent approach of Active
 165 Statistical Inference (Zrnic & Candes, 2024). Extending this framework, Angelopoulos et al. (2025)
 166 propose a method that optimizes the sampling rate between gold-standard and pseudo-labels rather
 167 than relying on a fixed label budget, and derive an improved active sampling policy. A recent work
 168 by Gligorić et al. (2025) utilizes LLM verbalized confidence scores to guide their sampling policy
 169 and subsequently performs active inference by combining the LLM and human annotations.

170 **Prediction Powered Inference.** The Active Statistical Inference framework is grounded in the idea
 171 of Prediction Powered Inference (PPI) Angelopoulos et al. (2023a), which differs from the former
 172 in that it assumes the availability of a small, pre-labeled dataset. The work of Angelopoulos et al.
 173 (2023b) introduces PPI++, which improves the computational efficiency of PPI by addressing the
 174 intractability of the original confidence interval construction and using a tuning parameter to con-
 175 trol the influence of the model predictions based on their quality. Subsequent works (Dorner et al.,
 176 2025; Mani et al., 2025) analyze the critical role of the correlation between the gold-standard and
 177 model-generated labels for the performance of PPI. Following the original PPI formulation, sev-
 178 eral works have proposed extensions and refinements in various directions, including addressing
 179 estimator bias in the few-label regime (Eyre & Madras, 2025), incorporating an inverse probability
 180 weighting (IPW) bias-correction term (Datta & Polson, 2025), combining predictions from multiple
 181 foundation models via a hybrid augmented IPW estimator (De Bartolomeis et al., 2025), applying
 182 a stratification approach (Fisch et al., 2024), exploring a bootstrap-based variant (Zrnic, 2024), em-
 183 ploying Bayes-assisted approaches (Cortinovis & Caron, 2025; Li & Ignatiadis, 2025) and extending
 184 the ideas of PPI to e-values (Csillag et al., 2025). Other applications of the PPI framework include
 185 LLM-assisted rank-set construction (Chatzi et al., 2024), average treatment effects from multiple
 186 datasets (Wang et al., 2025), clinic trial outcomes (Poulet et al., 2025), autoevaluation in machine
 187 learning systems (Boyeau et al., 2025; Park et al., 2025), machine learning generated surrogate
 188 rewards for multi-armed bandits (Ji et al., 2025b). A few other works have explored alternative
 189 machine-learning assisted estimators, e.g., Schmutz et al. (2023); Egami et al. (2023); Miao et al.
 190 (2023); Miao & Lu (2024); Gan et al. (2024).

191 We refer the reader to Appendix B for a more detailed discussion on related work.

192 **4 NON-ASYMPTOTIC ANALYSIS OF THE MEAN ESTIMATOR**

194 In this section, we provide the non-asymptotic analysis of the sequential active mean estimator.
 195 Since the sequential active estimation setting requires going over the data points sequentially, we
 196 can formulate the active mean estimator of Zrnic & Candes (2024) as an online update step at each
 197 time $t \in [T]$, as

$$199 \quad w_{t+1} = w_t + \frac{1}{T} \left(f_t(x_t) + (y_t - f_t(x_t)) \frac{\xi_t}{\pi_t(x_t)} \right), \quad (1)$$

201 where we set the initial point $w_1 = 0$, and let T be the horizon. We denote $g_t := f_t(x_t) + (y_t -$
 202 $f_t(x_t)) \frac{\xi_t}{\pi_t(x_t)}$. A simple calculation shows that $\mathbb{E}[g_t] = \mu_y$, the mean of the random variable $(y_t)_{t \geq 1}$.

204 Before proceeding with the asymptotic analysis of (1), we will need a technical lemma, known as
 205 Freedman's inequality, which is stated below for completeness.

206 **Lemma 1** (Freedman's inequality (Freedman, 1975), see also e.g., Lemma 3 in Rakhlin
 207 et al. (2012)). *Let ζ_1, \dots, ζ_T be a martingale difference sequence with a uniform upper
 208 bound $|\zeta_t| \leq b, \forall t$. Denote V_t the sum of conditional variances of ζ_s 's., i.e., $V_t =$
 209 $\sum_{s=1}^t \text{Var}(\zeta_s | \zeta_1, \dots, \zeta_{s-1})$. Also, denote $\sigma_t := \sqrt{V_t}$. Then, for any $0 < \delta < 1/e$ and $T \geq 4$,
 210 we have*

$$211 \quad \text{Prob} \left(\exists t \leq T : \sum_{s=1}^t \zeta_s \geq 2 \max \left\{ 2\sigma_t, b\sqrt{\log(1/\delta)} \right\} \sqrt{\log(1/\delta)} \right) \leq \log(T)\delta.$$

215 Lemma 1 provides a concentration inequality for martingales, yielding a high-probability bound
 on the deviation of a martingale sum from its mean that adapts to the accumulated conditional

216 variance. We are now ready to present the non-asymptotic analysis result for the sequential active
 217 mean estimator, as detailed in Theorem 1
 218

219 **Theorem 1.** Fix a time horizon $T \geq 4$. Assume each random variable g_t is bounded, i.e.,
 220 $|g_t| \leq G$ for a constant $G > 0$. Denote $\sigma_t^2 := \mathbb{E}[(g_t - \mu_y)^2 | \mathcal{F}_{t-1}]$ the conditional variance,
 221 where \mathcal{F}_{t-1} is the filtration up to $t-1$. Then, for any $\delta \in (0, 1/e)$, with probability at least
 222 $1 - \delta$, $\forall t \in [T]$:

$$223 \quad |w_{t+1} - \mu_y| \leq \frac{2 \max \left\{ 2\sqrt{S_t}, (G + |\mu_y|) \sqrt{\log \left(\frac{\log(T)}{\delta} \right)} \right\} \sqrt{\log \left(\frac{\log(T)}{\delta} \right)}}{T} + \left(1 - \frac{t}{T}\right) |\mu_y|, \quad (2)$$

$$228 \quad \text{where } S_t = \sum_{s=1}^t \sigma_s^2.$$

230 Theorem 1 demonstrates a data-dependent bound on the accuracy of the update w_t that holds with
 231 high-probability at any time $t \in [T]$. We make a couple remarks on this result. When $t \ll T$ and
 232 the $\left(1 - \frac{t}{T}\right) |\mu_y|$ term dominates the other term on (2), it is possible to observe a rate that is slower
 233 than $O(1/\sqrt{t})$ in the initial stage, i.e., $1 - \frac{t}{T} \approx 1$. However, after this burn-in stage (i.e., when t is
 234 sufficiently large), the first term will eventually become dominant. Furthermore, when $\sqrt{\sum_{s=1}^t \sigma_s^2}$
 235 dominates $(G + |\mu_y|) \sqrt{\log(\log(T)/\delta)}$, which happens easily when t is sufficiently large and δ is
 236 not too small, the rate becomes
 237

$$238 \quad |w_{t+1} - \mu_y| = O \left(\frac{\sqrt{\sum_{s=1}^t \sigma_s^2} \sqrt{\log(\log(T)/\delta)}}{T} \right). \quad (3)$$

242 Using the trivial bound $\sum_{s=1}^t \sigma_s^2 \leq 2t(G^2 + \mu_y^2)$, we can further express the rate as
 243 $O \left(\frac{\sqrt{t} \sqrt{(G^2 + \mu_y^2) \log(\log(T)/\delta)}}{T} \right) = O \left(\frac{1}{\sqrt{t}} \right).$
 244

247 5 POLICY OF QUERYING THE GROUND TRUTH

250 In the previous section, we discussed that the update w_t of (1) will have a rate of $O(\sqrt{t}/T) =$
 251 $O(1/\sqrt{t})$ in the worst case. An observation is that when $\sum_{s=1}^t \sigma_s^2 \ll 2t(G^2 + \mu_y^2)$, one might get
 252 an even faster rate than $O(1/\sqrt{t})$. This motivates us to control the sum of the conditional variances,
 253 i.e., $\sum_{s=1}^t \sigma_s^2 = \sum_{s=1}^t \mathbb{E}[(g_s - \mu_y)^2 | \mathcal{F}_{s-1}]$, by proposing an algorithm to determine the query
 254 probability of the ground-truth label online, which we detail next.

255 We begin by introducing the following observation on the decomposition of the conditional variance
 256 of the update step, which will subsequently guide the choice of the online query policy.
 257

258 **Lemma 2.** The conditional variance has the following decomposition:

$$259 \quad \mathbb{E}[(g_t - \mu_y)^2 | \mathcal{F}_{t-1}] = \mathbb{E} \left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1} \right] + \mathbb{E} \left[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1} \right] \\ 260 \quad + 2\mathbb{E} \left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1} \right] - \mu_y^2.$$

265 Furthermore, assume that the query policy at time t is \mathcal{F}_{t-1} -measurable, i.e., there exists a
 266 random variable $p_t \in [0, 1]$, measurable with respect to \mathcal{F}_{t-1} , such that $\pi_t(x_t) = p_t$. Then, we
 267 have

$$268 \quad \mathbb{E} \left[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1} \right] = \frac{1}{p_t} \mathbb{E} \left[(y_t - f_t(x_t))^2 \middle| \mathcal{F}_{t-1} \right].$$

We observe that the only term involving the query probability $\pi_t(x_t)$ that contributes to the conditional variance is $\mathbb{E}[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} | \mathcal{F}_{t-1}]$. We now consider the query policy at time t that is fully determined by the information up to $t-1$. With this, we can rewrite $\mathbb{E}[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} | \mathcal{F}_{t-1}] = \frac{1}{p_t} \mathbb{E}[(y_t - f_t(x_t))^2 | \mathcal{F}_{t-1}]$. However, we note that $\mathbb{E}[(y_t - f_t(x_t))^2 | \mathcal{F}_{t-1}]$ cannot be known since this depends on the unknown distributions of y_t and $f_t(x_t)$. Therefore, we assume that there is an oracle, denoted as $\Phi_t(x_t) \in \mathbb{R}_+$, which is available at time t and provides an approximation of the quantity of interest, i.e.,

$$\frac{1}{c_1} \Phi_t(x_t) \leq \mathbb{E} \left[(y_t - f_t(x_t))^2 \mid \mathcal{F}_{t-1} \right] \leq c_0 \Phi_t(x_t),$$

for some constants $c_0, c_1 > 0$ such that $\Phi_t(x_t) \approx \mathbb{E}[(y_t - f_t(x_t))^2 | \mathcal{F}_{t-1}]$. Equipped with such an oracle, we propose specifying the query probability p_t based on the following rule:

$$p_t \leftarrow \arg \min_{p \in [\beta, \tau]} \gamma \theta_{t-1} p + \frac{1}{2} p^2, \text{ where } \theta_{t-1} := - \sum_{s=1}^{t-1} \frac{\Phi_s(x_s)}{p_s^2}, \quad (4)$$

where $\gamma > 0$, $\tau \in (0, 1]$, $\beta \in (0, \tau]$ are user-specified parameters, and we let $\theta_0 := 0$. The following lemma shows that p_t has a closed-form expression.

Lemma 3. *The update (4) has a closed-form expression, which is*

$$p_t = \max\{\beta, \min\{\tau, -\gamma \theta_{t-1}\}\}.$$

We note that one can specify $\tau = \frac{T_b}{T}$, where T_b denotes the targeted maximum number of rounds in which the ground-truth is queried, which ensures that the query probability p_t at time t does not exceed the ratio $\frac{T_b}{T}$. This constraint is also akin to the sampling rule $\mathbb{E}[p_t] \leq \frac{T_b}{T}$ considered in Zrnic & Candes (2024). On the other hand, the parameter β encourages certain exploration at each round by preventing the query probability p_t from becoming too close to 0.

The update (4), in a nutshell, is one of the celebrated online learning algorithms called Follow-the-Regularized-Leader (FTRL), (see, e.g., Abernethy et al. (2012); Wang et al. (2024), and Chapter 7 of Orabona (2019)). FTRL is known to enjoy a sublinear regret bound when the sequence of loss functions is convex. We propose leveraging the strong guarantee of FTRL to determine the query probability p_t online. More specifically, in our scenario, one first determines the query probability p_t , after which it receives a loss function defined as $\tilde{\ell}_t(p) := \frac{\Phi_t(x_t)}{p}$, which is a convex loss function in $(0, 1]$. In online learning, a common goal is to minimize the regret. In our setting, the regret against a benchmark $p_* \in [\beta, \tau]$ over t rounds is defined as:

$$\text{Regret}_t(p_*) := \sum_{s=1}^t \tilde{\ell}_s(p_s) - \sum_{s=1}^t \tilde{\ell}_s(p_*) = \sum_{s=1}^t \frac{\Phi_s(x_s)}{p_s} - \sum_{s=1}^t \frac{\Phi_s(x_s)}{p_*}, \quad (5)$$

where the first sum is the cumulative loss of the updates $(p_s)_{s \geq 1}$ and the second one is that of the benchmark. A sublinear regret bound against any benchmark p_* in the same decision space $[\beta, \tau]$ of the learner implies that the sequence of query probabilities can *compete* with the best fixed query probability in hindsight. On the other hand, given that the oracle's output is non-negative, i.e., $\forall s : \Phi_s(\cdot) \geq 0$, it follows that $\arg \min_{p_* \in [\beta, \tau]} \tilde{\ell}_s(p) = \tau$. Combining these implies that an online learner may need to approach τ eventually to achieve a sublinear regret. In other words, to maintain sublinear regret, the query probabilities p_t will need to converge toward the constraint upper bound $\tau = \frac{T_b}{T}$. In particular, we have that the average regret is in fact vanishing (a.k.a. no-regret learning), i.e., $\frac{\text{Regret}_t(p_*)}{t} \rightarrow 0$ as $t \rightarrow \infty$, as the following lemma shows.

Lemma 4. (see e.g., Theorem 3 in Luo (2017)) FTRL satisfies

$$\text{Regret}_t(p_*) \leq \gamma \sum_{s=1}^t \left| \dot{\tilde{\ell}}_s \right|^2 + \frac{R(p_*) - \min_{p \in \mathcal{K}} R(p)}{\gamma},$$

324 for any comparator $p_* \in \mathcal{K} := [\beta, \tau]$, where $\gamma > 0$, $R(p) := \frac{1}{2}p^2$, and $\dot{\ell}_s := \frac{d\tilde{\ell}_s(p)}{dp} \Big|_{p=p_s}$.

326
327 We note that Lemma 4 is a classical result in online learning literature, see also Orabona (2019);
328 Shalev-Shwartz et al. (2012). The guarantee suggests that if one chooses $\gamma = \frac{1}{\sqrt{T}}$, then the regret
329 of FTRL is $O(\sqrt{T})$, which grows sublinearly with T , provided that the size of the derivative is
330 bounded. The following lemma shows that the size of the derivative in the regret bound is bounded
331 whenever the oracle’s output is bounded.

332 **Lemma 5.** Assume that the range of oracle’s output is bounded, i.e., $\forall t : \Phi(x_t) \leq B$, for a
333 constant $B > 0$. Then, $\forall t : \left| \dot{\ell}_t \right|^2 \leq \frac{B^2}{\beta^4}$.

335 In the following theorem, we denote $\sigma_{1:t}^{*2} := \sum_{s=1}^t \sigma_{s,(t)}^{*2}$, the cumulative conditional variance ob-
336 tained under a fixed query probability $p_{1:t}^* \in \mathcal{K} := [\beta, \tau]$, as if the method had committed the best
337 fixed probability *in hindsight* over t rounds rather than following the query policy from (4), i.e.,
338 $p_{1:t}^* = \arg \min_{p \in [\beta, \tau]} \sum_{s=1}^t \tilde{\ell}_s(p)$.

339 **Theorem 2.** Assume that there is an oracle that outputs $\Phi_t(x_t)$ at each t such that $\frac{1}{c_1} \Phi_t(x_t) \leq$
340 $\mathbb{E} \left[(y_t - f_t(x_t))^2 \mid \mathcal{F}_{t-1} \right] \leq c_0 \Phi_t(x_t)$ for some constant $c_0, c_1 > 0$ and that $\forall t : \Phi(x_t) \leq B$.
341 Set the parameter $\gamma = \frac{1}{\sqrt{T}} \frac{\beta^2}{B}$. Using the query policy (4), we have that, with probability at
342 least $1 - \delta$, $\forall t \in [T]$:

$$343 |w_{t+1} - \mu_y| \leq \frac{2 \max \left\{ 2\sqrt{\Psi_t}, (G + |\mu_y|) \sqrt{\log \left(\frac{\log(T)}{\delta} \right)} \right\} \sqrt{\log \left(\frac{\log(T)}{\delta} \right)}}{T} + \left(1 - \frac{t}{T} \right) |\mu_y|,$$

344 for any $\beta \in (0, \tau)$ and $\tau \in (0, 1)$, where $\Psi_t \leq c_0 c_1 \sigma_{1:t}^{*2} + 2c_0 \frac{t}{\sqrt{T}} \frac{B}{\beta^2} +$
345 $\frac{c_0 (R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p)) \sqrt{T} B}{\beta^2}$.

346 What Theorem 2 shows is a *data-dependent* bound. We note that $\sqrt{\Psi_t} \leq \sqrt{c_0 c_1 \sigma_{1:t}^{*2}} +$
347 $O(T^{1/4})$. From our earlier discussion, once a sufficient burn-in period has elapsed so that the
348 first term in the upper bound dominates, the non-asymptotic rate takes the form $O\left(\frac{\sqrt{\Psi_t}}{T}\right) =$
349 $O\left(\frac{\sqrt{c_0 c_1 \sigma_{1:t}^{*2}}}{T} + \frac{1}{T^{3/4}}\right)$, provided that δ is not too small.

350 6 EXPERIMENTS

351 In this section, we report experimental results by comparing the proposed method with two base-
352 lines. For clarity, Algorithm 1 presents the protocol for the task of active sequential mean estimation.
353 Compared to the procedure described in Algorithm 2 of Zrnic & Candes (2024), the key difference
354 is that we also update the uncertainty predictor whenever the ML model for label prediction is
355 updated. Furthermore, we split the dataset with ground-truth labels into two disjoint subsets, which
356 are accumulated as described on Line 13, and use these subsets to expand the data available for
357 updating the ML model f_{t+1} and its uncertainty predictor u_{t+1} , respectively. This treatment of the
358 disjoint training sets is intended to enable the uncertainty predictor to more accurately estimate the
359 uncertainty of the ML model when it is applied to *unseen* data at test time.

360 The first baseline was also considered in the prior work of Zrnic & Candes (2024).

$$361 w_T^{\text{Uniform}} := \frac{1}{T} \sum_{t=1}^T \left(f(x_t) + \frac{(y_t - f(x_t)) \xi_t}{T_b/T} \right), \quad \text{where } \xi_t \sim \text{Bernoulli} \left(\frac{T_b}{T} \right) \quad (6)$$

362 For this baseline, we note that the ML predictor is fixed. As discussed in Zrnic & Candes (2024),
363 this comparison can showcase the benefit of data collection. Following the terminology of Zrnic &
364 Candes (2024), we refer to this baseline as “uniform sampling.”

378 **Algorithm 1** Protocol of Active Sequential Mean Estimation

379 **Require:** Significance level parameter $\alpha \in (0, 1)$, target sampling budget $T_b > 0$, and batch size
380 B .

381 1: **Initialize** a machine learning (ML) model $f_1(\cdot) : \mathcal{X} \rightarrow \mathcal{Y}$ to predict the labels of data.

382 2: **Initialize** an uncertainty predictor $u_1(\cdot, \cdot) : \mathcal{X} \times f_1(\cdot) \rightarrow \mathbb{R}_+$ for the model's predictions.

383 3: **Initialize** the dataset for updating the model $\mathcal{D}_{\text{train}}$ and the dataset for the uncertainty predictor
384 $\mathcal{D}_{\text{uncertainty}}$.

385 4: **Set** $\mathcal{D}_{\text{tmp}} \leftarrow \emptyset$, $w_1 \leftarrow 0$

386 5: **for** $t = 1, \dots, T$ **do**

387 6: Observe features x_t of a sample and determine the query probability p_t for getting its label.

388 7: Sample the binary random variable $\xi_t \sim \text{Bernoulli}(p_t)$.

389 8: **if** $\xi_t = 1$ **then**

390 9: Obtain the ground-truth label y_t and set $\mathcal{D}_{\text{tmp}} \leftarrow \mathcal{D}_{\text{tmp}} \cup \{(x_t, y_t)\}$.

391 10: Increase b by 1.

392 11: **end if**

393 12: **if** $B = b$ **then**

394 13: Randomly split \mathcal{D}_{tmp} into two datasets with equal sizes, $\mathcal{D}_{\blacklozenge}$ and $\mathcal{D}_{\blacklozenge}$.

395 14: Set $\mathcal{D}_{\text{train}} \leftarrow \mathcal{D}_{\text{train}} \cup \mathcal{D}_{\blacklozenge}$ and set $\mathcal{D}_{\text{uncertainty}} \leftarrow \mathcal{D}_{\text{uncertainty}} \cup \mathcal{D}_{\blacklozenge}$.

396 15: Update the ML model to $f_{t+1}(\cdot)$ using the dataset $\mathcal{D}_{\text{train}}$; similarly, update the uncertainty
397 predictor $u_{t+1}(\cdot, \cdot) : \mathcal{X} \times f_{t+1}(\cdot) \rightarrow \mathbb{R}_+$ using the dataset $\mathcal{D}_{\text{uncertainty}}$.

398 16: Reset $b \leftarrow 0$.

399 17: **else**

400 18: $f_{t+1} \leftarrow f_t$ and $u_{t+1} \leftarrow u_t$.

401 19: **end if**

402 20: Update the estimate $w_{t+1} = w_t + \frac{1}{T} \left(f_t(x_t) + (y_t - f_t(x_t)) \frac{\xi_t}{p_t} \right)$.

403 21: **end for**

404 22: **Set** $\hat{\sigma}^2 \leftarrow \frac{1}{T} \sum_{t=1}^T \left(f_t(x_t) + (y_t - f_t(x_t)) \frac{\xi_t}{p_t} - w_{T+1} \right)^2$.

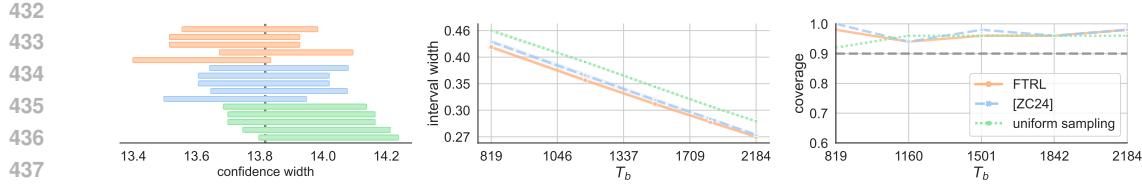
405 23: **Output:** $(1 - \alpha)$ -confidence interval $CI_\alpha = \left(w_{T+1} \pm z_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{T}} \right)$.

408 The second baseline is the scheme proposed in Zrnic & Candes (2024) for implementing Line 6
409 in Algorithm 1, which determines the query probability p_t , as described in the earlier preliminary
410 section.

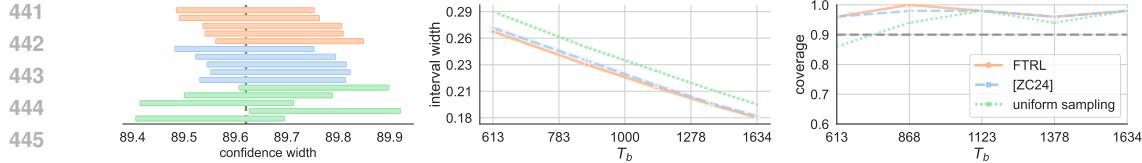
413 6.1 DATASETS AND EXPERIMENTAL SETUP

414 We compare the algorithms on three real-world datasets. The first dataset concerns the politeness
415 scores of texts based on human annotations, which is available from the works of Ji et al. (2025a)
416 and Gligorić et al. (2025). For each article, there is an associated 21-dimensional feature vector and
417 a score predicted by ChatGPT. We consider the task of regression for this dataset, where the ML
418 model is trained on the 22-dimensional vector (the 21 features plus the ChatGPT score) to predict
419 the average score of 5 human judgments. Following the suggestion in Zrnic & Candes (2024), an
420 uncertainty estimator $u_t(\cdot, \cdot) : \mathcal{X} \times f_t(\cdot) \rightarrow \mathbb{R}_+$ is used to predict the absolute error $|f_t(x_t) - y_t|$
421 from x_t without seeing the label y_t beforehand. This predicted uncertainty is then used as the
422 input to their proposed scheme for determining the query probability p_t at round t . The uncertainty
423 estimator is also updated based on the collected samples with queried ground-truth labels once every
424 batch of B labeled samples is collected, as depicted in Algorithm 1. For our proposed scheme, we
425 need to construct the approximation oracle $\Phi_t(x_t)$. In practice, we implement this by performing a
426 linear regression on the squared residual error $(f_t(x_t) - y_t)^2$ for the samples in $\mathcal{D}_{\text{uncertainty}}$, and this
427 estimator is updated regularly after every batch of size B .

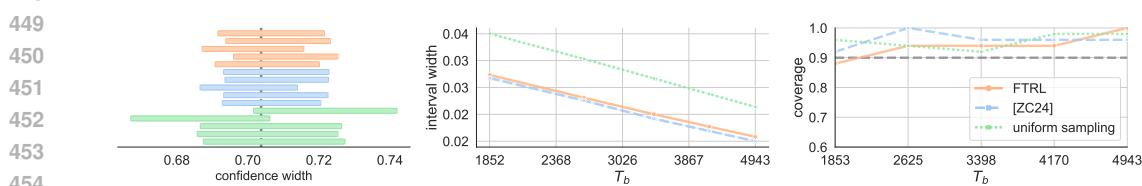
428 The second dataset concerns predicting the ratings of wine reviews, which is available in Ji et al.
429 (2025a). Each review is associated with the price of the wine and four additional binary attributes
430 representing the regions, along with the rating predicted by OpenAI's GPT-4o mini based on the
431 reviewers' comments. We also consider the task of regression for this dataset, where a linear regression
model is trained on the aforementioned covariates to predict the human ratings. The uncertainty



438 Figure 2: **Politeness score analysis.** Left: Intervals of randomly selected trials. Middle: Average confidence
439 width across repeated trials vs. sampling budget T_b . Right: Percentage of trials that cover the true mean.
440



446 Figure 3: **Wine review analysis.** Left: Intervals of randomly selected trials. Middle: Average confidence
447 width across repeated trials vs. sampling budget T_b . Right: Percentage of trials that cover the true mean.
448



455 Figure 4: **Post-election survey.** Left: Intervals of randomly selected trials. Middle: Average confidence
456 width across repeated trials vs. sampling budget T_b . Right: Percentage of trials that cover the true mean.
457

458 predictor and the approximation oracle $\Phi_t(x_t)$ have the same form and are updated in the same fash-
459 ion as for the first dataset.

460 The third dataset is a post-election survey dataset considered in Zrnic & Candes (2024), where the
461 original source is from Center (2020). This dataset includes the approval ratings of two politicians,
462 where approval is represented by $y_t \in \{0, 1\}$. Following the experimental setup in Zrnic & Candes
463 (2024), the ML model $f(\cdot)$ is implemented as an XGBoost model. Since the response y_t is binary,
464 the task can be treated as a classification problem. We hence follow the treatment in Zrnic & Candes
465 (2024) by using the uncertainty predictor as $u_t(x_t, f_t(\cdot)) = 2 \min\{f_t(x_t), 1 - f_t(x_t)\}$ for their
466 proposed scheme, where $f_t(x_t)$ is the predicted probability of $y_t = 1$ for x_t given by the XGBoost
467 model. On the other hand, the required approximation oracle $\Phi(x_t)$ is trained in the same fashion as
468 in the first two tasks.

469 We also conduct experiments on a synthetic dataset, the details of which can be found in Appendix E.
470

471 6.2 RESULTS

473 In this subsection, we report the results of the conducted experiments. Figures 2 - 4 and Figure 6
474 (provided in Appendix E due to space limitations) show the intervals of randomly selected trials,
475 average confidence width, and coverage for each of the datasets considered over 50 trials. Across all
476 four datasets examined, we find that the FTRL policy yields performance comparable to the mixture
477 policy proposed by Zrnic & Candes (2024), in the sense that both result in confidence intervals of
478 similar width, while both outperform the baseline policy. Notably, in two of the datasets, the FTRL
479 policy attains marginally narrower confidence intervals. With respect to coverage of the true mean,
480 all three policies yield a high proportion of confidence intervals that successfully include the true
481 value.

482 Our theoretical analysis and experimental findings consistently indicate that when the query prob-
483 ability p_t at time t is oblivious to the current covariates x_t , while still permitted to depend on past
484 covariates or past uncertainty estimates, the optimal strategy is simply to set $p_t = \frac{T_b}{T}$ in accordance
485 with the sampling budget. This result, implies that constructing an uncertainty predictor, or lever-
486 aging uncertainty estimates in any form, does not appear to provide a clear advantage for this type

486 of policy. Perhaps unexpectedly, this rules out any benefit from conditioning on past covariates or
 487 past uncertainty estimates. Furthermore, as our figures illustrate, even when the query probability
 488 ignores the current covariates, FTRL, which quickly converges to the constant $\frac{T_b}{T}$ and then maintains
 489 it, performs on par with the more sophisticated scheme of Zrnic & Candes (2024), which explicitly
 490 uses the current features x_t to determine the query probability.
 491

492 **REFERENCES**
 493

494 Jacob D Abernethy, Elad Hazan, and Alexander Rakhlin. Interior-point methods for full-information
 495 and bandit online learning. *IEEE Transactions on Information Theory*, 58(7):4164–4175, 2012.

496 Anastasios N Angelopoulos, Stephen Bates, Clara Fannjiang, Michael I Jordan, and Tijana Zrnic.
 497 Prediction-powered inference. *Science*, 382(6671):669–674, 2023a.

498 Anastasios N Angelopoulos, John C Duchi, and Tijana Zrnic. PPI++: Efficient prediction-powered
 499 inference. *arXiv preprint arXiv:2311.01453*, 2023b.

500 Anastasios N Angelopoulos, Jacob Eisenstein, Jonathan Berant, Alekh Agarwal, and Adam Fisch.
 501 Cost-optimal active ai model evaluation. *arXiv preprint arXiv:2506.07949*, 2025.

502 Hilal Asi, Vitaly Feldman, Jelani Nelson, Huy Nguyen, Kunal Talwar, and Samson Zhou. Private
 503 vector mean estimation in the shuffle model: Optimal rates require many messages. In *International
 504 Conference on Machine Learning (ICML)*, 2024.

505 Abdellah Aznag, Rachel Cummings, and Adam N Elmachtoub. An active learning framework for
 506 multi-group mean estimation. *Advances in Neural Information Processing Systems (NeurIPS)*,
 507 36:32602–32635, 2023.

508 William H. Beluch, Tim Genewein, Andreas Nurnberger, and Jan M. Kohler. The power of en-
 509 sembles for active learning in image classification. In *2018 IEEE/CVF Conference on Computer
 510 Vision and Pattern Recognition*, pp. 9368–9377, 2018. doi: 10.1109/CVPR.2018.00976.

511 Ran Ben-Basat, Amit Portnoy, Gil Einziger, Yaniv Ben-Itzhak, Michael Mitzenmacher, et al. Ac-
 512 celerating federated learning with quick distributed mean estimation. In *International Conference
 513 on Machine Learning (ICML)*, 2024.

514 Pierre Boreau, Anastasios N. Angelopoulos, Tianle Li, Nir Yosef, Jitendra Malik, and Michael I.
 515 Jordan. Autoeval done right: Using synthetic data for model evaluation. In *International Confer-
 516 ence on Machine Learning (ICML)*, 2025.

517 Pew Research Center. American trends panel (ATP) Wave 79, 2020. URL <https://www.pewresearch.org/science/dataset/american-trends-panel-wave-79/>.
 518 Accessed: 2025-09-17.

519 Ivi Chatzi, Eleni Straitouri, Suhas Thejaswi, and Manuel Gomez-Rodriguez. Prediction-powered
 520 ranking of large language models. In *Advances in Neural Information Processing Systems
 521 (NeurIPS)*, 2024.

522 Yiding Chen, Jerry Zhu, and Kirthevasan Kandasamy. Mechanism design for collaborative normal
 523 mean estimation. *Advances in Neural Information Processing Systems (NeurIPS)*, 36:49365–
 524 49402, 2023.

525 Yu Cheng, Ilias Diakonikolas, Rong Ge, and Mahdi Soltanolkotabi. High-dimensional robust mean
 526 estimation via gradient descent. In *International Conference on Machine Learning (ICML)*, pp.
 527 1768–1778, 2020.

528 Stefano Cortinovis and Francois Caron. Ab-ppi: Frequentist, assisted by bayes, prediction-powered
 529 inference. In *International Conference on Machine Learning (ICML)*, 2025.

530 Daniel Csillag, Claudio Jose Struchiner, and Guilherme Tegoni Goedert. Prediction-powered e-
 531 values. In *International Conference on Machine Learning (ICML)*, 2025.

540 Trung Dang, Jasper Lee, Maoyuan'Raymond' Song, and Paul Valiant. Optimality in mean estimation:
 541 beyond worst-case, beyond sub-gaussian, and beyond $1 + \alpha$ moments. *Advances in Neural*
 542 *Information Processing Systems (NeurIPS)*, 36:4150–4176, 2023.

543 Sanjoy Dasgupta. Two faces of active learning. *Theoretical Computer Science*, 412(19):1767–1781,
 544 2011. ISSN 0304-3975. doi: <https://doi.org/10.1016/j.tcs.2010.12.054>. URL <https://www.sciencedirect.com/science/article/pii/S0304397510007620>. Algorithmic
 545 Learning Theory (ALT 2009).

546 Jyotishka Datta and Nicholas G Polson. Prediction-powered inference with inverse probability
 547 weighting. *arXiv preprint arXiv:2508.10149*, 2025.

548 Piersilvio De Bartolomeis, Javier Abad, Guanbo Wang, Konstantin Donhauser, Raymond M Duch,
 549 Fanny Yang, and Issa J Dahabreh. Efficient randomized experiments using foundation models.
 550 *arXiv preprint arXiv:2502.04262*, 2025.

551 Ilias Diakonikolas, Giannis Iakovidis, Daniel Kane, and Thanasis Pittas. Efficient multivariate ro-
 552 bust mean estimation under mean-shift contamination. In *International Conference on Machine*
 553 *Learning (ICML)*, 2025.

554 Florian E Dorner, Vivian Yvonne Nastl, and Moritz Hardt. Limits to scalable evaluation at the
 555 frontier: Llm as judge won't beat twice the data. In *International Conference on Representation*
 556 *Learning (ICLR)*, 2025.

557 Melanie Ducoffe and Frederic Precioso. Adversarial active learning for deep networks: a margin
 558 based approach, 2018. URL <https://arxiv.org/abs/1802.09841>.

559 Naoki Egami, Musashi Hinck, Brandon Stewart, and Hanying Wei. Using imperfect surrogates for
 560 downstream inference: Design-based supervised learning for social science applications of large
 561 language models. *Advances in Neural Information Processing Systems (NeurIPS)*, 36:68589–
 562 68601, 2023.

563 Benjamin Eyre and David Madras. Regression for the mean: Auto-evaluation and inference with few
 564 labels through post-hoc regression. In *International Conference on Machine Learning (ICML)*,
 565 2025.

566 Adam Fisch, Joshua Maynez, R. Alex Hofer, Bhuvan Dhingra, Amir Globerson, and William W.
 567 Cohen. Stratified prediction-powered inference for effective hybrid evaluation of language mod-
 568 els. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, 2024.

569 David A Freedman. On tail probabilities for martingales. *the Annals of Probability*, pp. 100–118,
 570 1975.

571 Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep Bayesian active learning with image
 572 data. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Con-*
 573 *ference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp.
 574 1183–1192. PMLR, 06–11 Aug 2017. URL <https://proceedings.mlr.press/v70/gal17a.html>.

575 Feng Gan, Wanfeng Liang, and Changliang Zou. Prediction de-correlated inference: A safe ap-
 576 proach for post-prediction inference. *Australian & New Zealand Journal of Statistics*, 66(4):
 577 417–440, 2024.

578 Kristina Gligorić, Tijana Zrnic, Cinoo Lee, Emmanuel Candes, and Dan Jurafsky. Can unconfident
 579 llm annotations be used for confident conclusions? In *Proceedings of the 2025 Conference of*
 580 *the Nations of the Americas Chapter of the Association for Computational Linguistics: Human*
 581 *Language Technologies (Volume 1: Long Papers)*, pp. 3514–3533, 2025.

582 Steve Hanneke. Theory of disagreement-based active learning. *Found. Trends Mach. Learn.*, 7(2–3):
 583 131–309, June 2014. ISSN 1935-8237. doi: 10.1561/2200000037. URL <https://doi.org/10.1561/2200000037>.

584 Wenlong Ji, Lihua Lei, and Tijana Zrnic. Predictions as surrogates: Revisiting surrogate outcomes
 585 in the age of ai. *arXiv preprint arXiv:2501.09731*, 2025a.

594 Wenlong Ji, Yihan Pan, Ruihao Zhu, and Lihua Lei. Multi-armed bandits with machine learning-
 595 generated surrogate rewards. *arXiv preprint arXiv:2506.16658*, 2025b.
 596

597 Ajay J. Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. Multi-class active learning for image
 598 classification. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2372–
 599 2379, 2009. doi: 10.1109/CVPR.2009.5206627.

600 Daniel M Kane, Ilias Diakonikolas, Hanshen Xiao, and Sihan Liu. Online robust mean estimation.
 601 In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp.
 602 3197–3235. SIAM, 2024.

603 Sida Li and Nikolaos Ignatiadis. Prediction-powered adaptive shrinkage estimation. In *International
 604 Conference on Machine Learning (ICML)*, 2025.

605 Xiang Li, Yunai Li, Huiying Zhong, Lihua Lei, and Zhun Deng. Statistical inference under perfor-
 606 mativity. *arXiv preprint arXiv:2505.18493*, 2025.

607 Haipeng Luo. *Introduction to Online Learning CSCI 699*. Lecture Note, 2017. URL <https://haipeng-luo.net/courses/CSCI699/lecture2.pdf>.

608 Pranav Mani, Peng Xu, Zachary C Lipton, and Michael Oberst. No free lunch: Non-asymptotic
 609 analysis of prediction-powered inference. *arXiv preprint arXiv:2505.20178*, 2025.

610 Jiacheng Miao and Qiongshi Lu. Task-agnostic machine-learning-assisted inference. *Advances in
 611 Neural Information Processing Systems (NeurIPS)*, 37:106162–106189, 2024.

612 Jiacheng Miao, Xinran Miao, Yixuan Wu, Jiwei Zhao, and Qiongshi Lu. Assumption-lean and
 613 data-adaptive post-prediction inference. *arXiv preprint arXiv:2311.14220*, 2023.

614 Gleb Novikov, David Steurer, and Stefan Tiegel. Robust mean estimation without moments for
 615 symmetric distributions. *Advances in Neural Information Processing Systems (NeurIPS)*, 36:
 616 34371–34409, 2023.

617 Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*,
 618 2019.

619 Art B. Owen. *Monte Carlo theory, methods and examples*. <https://artowen.su.domains/mc/>, 2013.

620 Sangwoo Park, Matteo Zecchin, and Osvaldo Simeone. Adaptive prediction-powered autoeval with
 621 reliability and efficiency guarantees. *arXiv preprint arXiv:2505.18659*, 2025.

622 Ankit Pensia. A subquadratic time algorithm for robust sparse mean estimation. In *International
 623 Conference on Machine Learning (ICML)*, 2024.

624 Pierre-Emmanuel Poulet, Maylis Tran, Sophie Tezenas du Montcel, Bruno Dubois, Stanley Durrel-
 625 man, and Bruno Jedynak. Prediction-powered inference for clinical trials. *medRxiv*, 2025.

626 Alexander Rakhlin, Ohad Shamir, and Karthik Sridharan. Making gradient descent optimal for
 627 strongly convex stochastic optimization. In *International Conference on Machine Learning
 628 (ICML)*, 2012.

629 Pengzhen Ren, Yun Xiao, Xiaojun Chang, Po-Yao Huang, Zhihui Li, Brij B. Gupta, Xiaojiang Chen,
 630 and Xin Wang. A survey of deep active learning, 2021. URL <https://arxiv.org/abs/2009.00236>.

631 Hugo Schmutz, Olivier Humbert, and Pierre-Alexandre Mattei. Don't fear the unlabelled: Safe semi-
 632 supervised learning via debiasing. In *Proceedings of the International Conference on Learning
 633 Representations (ICLR)*, 2023.

634 Greg Schohn and David Cohn. Less is more: Active learning with support vector machines. In Pat
 635 Langley (ed.), *Proceedings of the Seventeenth International Conference on Machine Learning
 636 (ICML 2000)*, Stanford University, Stanford, CA, USA, June 29 - July 2, 2000, pp. 839–846.
 637 Morgan Kaufmann, 2000.

648 Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, Uni-
 649 versity of Wisconsin-Madison, 2009. URL [http://axon.cs.byu.edu/~martinez/](http://axon.cs.byu.edu/~martinez/classes/778/Papers/settles.activelearning.pdf)
 650 classes/778/Papers/settles.activelearning.pdf.

651

652 Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and*
 653 *Trends® in Machine Learning*, 4(2):107–194, 2012.

654

655 Simon Tong and Daphne Koller. Support vector machine active learning with application to text
 656 classification. In Pat Langley (ed.), *Proceedings of the Seventeenth International Conference on*
 657 *Machine Learning (ICML 2000)*, Stanford University, Stanford, CA, USA, June 29 - July 2, 2000,
 658 pp. 999–1006. Morgan Kaufmann, 2000.

659

660 Gokhan Tur, Dilek Hakkani-Tür, and Robert E. Schapire. Combining active and semi-supervised
 661 learning for spoken language understanding. *Speech Communication*, 45(2):171–186, 2005.
 662 ISSN 0167-6393. doi: <https://doi.org/10.1016/j.specom.2004.08.002>. URL <https://www.sciencedirect.com/science/article/pii/S0167639304000962>.

663

664 Jun-Kun Wang, Jacob Abernethy, and Kfir Y Levy. No-regret dynamics in the fenchel game: A
 665 unified framework for algorithmic convex optimization. *Mathematical Programming*, 205(1):
 666 203–268, 2024.

667

668 Yuxin Wang, Maresa Schröder, Dennis Frauen, Jonas Schweisthal, Konstantin Hess, and Stefan
 669 Feuerriegel. Constructing confidence intervals for average treatment effects from confounded and
 670 unconfounded data. In *International Conference on Representation Learning (ICLR)*, 2025.

671

672 Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting.
 673 *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(1):1–27, 2024.

674

675 Zichun Xu, Daniela Witten, and Ali Shojaie. A unified framework for semiparametrically efficient
 676 semi-supervised learning. *arXiv preprint arXiv:2502.17741*, 2025.

677

678 Tijana Zrnic. A note on the prediction-powered bootstrap. *arXiv e-prints*, arXiv:2405.18379, 2024.
 679 URL <https://arxiv.org/abs/2405.18379>.

680

681 Tijana Zrnic and Emmanuel Candès. Cross-prediction-powered inference. *Proceedings of the Na-*
 682 *tional Academy of Sciences*, 120(41):e2322083121, 2023.

683

684 Tijana Zrnic and Emmanuel Candes. Active statistical inference. In *International Conference on*
 685 *Machine Learning*, pp. 62993–63010. PMLR, 2024.

686

687

688

689

690

691

692

693

694

695

696

697

698

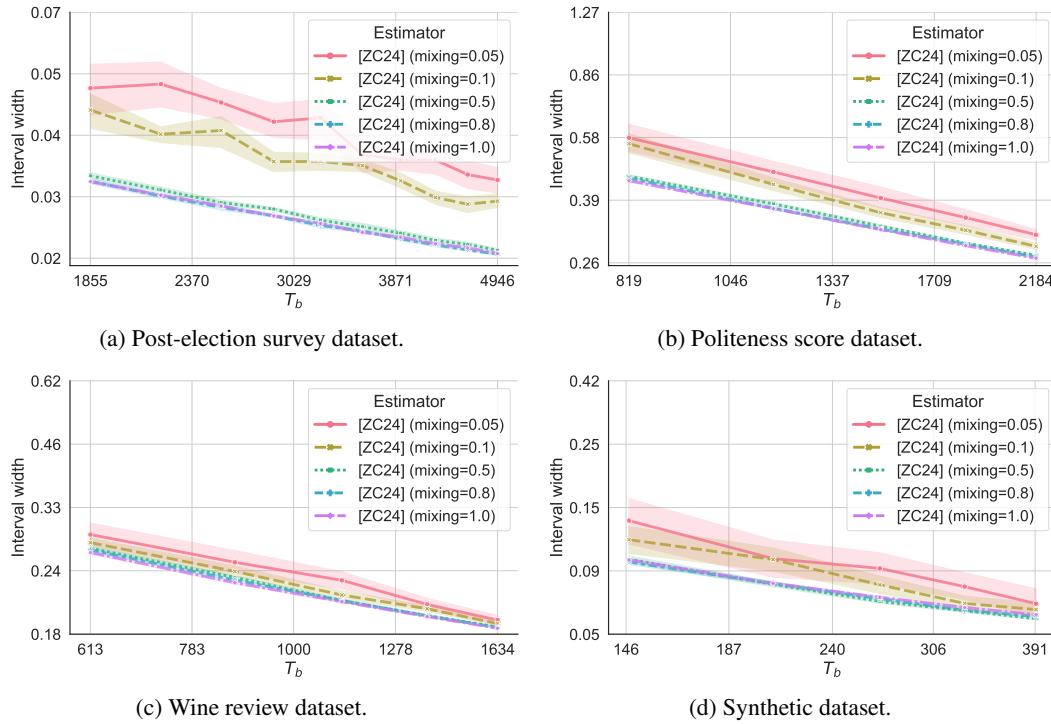
699

700

701

702
703 A EXPERIMENTS ON THE EFFECT OF THE MIXING PARAMETER
704

705 In this section, we report the results of experiments on tuning the mixing constant in the mixture
706 policy of Zrnic & Candes (2024), evaluated across four different datasets. A detailed description of
707 these datasets is provided in Section 6 and Appendix E.



732 Figure 5: Interval width vs. the sampling budget parameter T_b for different values of the mixing
733 parameter of the query probability scheme in Zrnic & Candes (2024). Averaged over 50 repeated
734 runs.

735
736 B EXTENDED RELATED WORK
737

738 **Active Statistical Inference.** The ideas in this work are motivated by the recent approach of Active
739 Statistical Inference (Zrnic & Candes, 2024). Extending this framework, Angelopoulos et al. (2025)
740 propose a method that optimizes the sampling rate between gold-standard and pseudo-labels rather
741 than relying on a fixed label budget, and derive an improved active sampling policy. A recent work
742 by Gligorić et al. (2025) utilizes LLM verbalized confidence scores to guide their sampling policy
743 and subsequently performs active inference by combining the LLM and human annotations.

744 **Prediction Powered Inference.** The Active Statistical Inference framework is grounded in the idea
745 of Prediction Powered Inference (PPI) Angelopoulos et al. (2023a), which differs from the former
746 in that it assumes the availability of a small, pre-labeled dataset. The work of Angelopoulos
747 et al. (2023b) introduces PPI++, which improves the computational efficiency of PPI by addressing
748 the intractability of the original confidence interval construction and using a tuning parameter to
749 control the influence of the model predictions based on their quality. Subsequent works (Dorner
750 et al., 2025; Mani et al., 2025) analyze the critical role of the correlation between the gold-standard
751 and model-generated labels for the performance of PPI. Focusing on the few-label regime, Eyre &
752 Madras (2025) argue that the PPI++ framework may lead to a significantly biased estimator that
753 is less efficient than classical inference by establishing its connection to univariate ordinary least
754 squares regression. Datta & Polson (2025) examine the use of an inverse probability weighted (IPW)
755 bias-correction term in the PPI mean estimator, inspired by classical Horvitz–Thompson and Hájek
estimators. The work of De Bartolomeis et al. (2025) establishes a connection of PPI++ with the

756 augmented inverse probability weighting (AIPW) estimator and propose an extension, which allows
 757 utilizing predictions from multiple foundation models. To address cases where the model accuracy
 758 varies across subdomains, Fisch et al. (2024) apply a stratification approach to PPI. Zrnic (2024)
 759 explore a bootstrap-based PPI variation to tackle arbitrary estimation problems. Xu et al. (2025)
 760 improve the PPI framework by proposing a safe PPI estimator that is always more efficient than
 761 the initial supervised estimator and can be used for arbitrary inferential problems. Li et al. (2025)
 762 generalize the ideas of PPI to a dynamic performative setting and show improved confidence regions
 763 in the task of performative prediction. In a semi-supervised context, Zrnic & Candès (2023) propose
 764 a method of using the labeled datapoints for cross-fitting and using the fitted models to compute
 765 the desired estimator. Cortinovis & Caron (2025) extend PPI by applying a Bayes-assisted frame-
 766 work that uses prior knowledge on the accuracy of the model predictions. In the case of compound
 767 estimation settings, Li & Ignatiadis (2025) adopt an approach that combines PPI with empirical
 768 Bayes shrinkage to correct noisy predictions within each problem and subsequently uses these as
 769 a shrinkage target. An interesting work by Csillag et al. (2025) presents a PPI framework based
 770 on e-values. Other applications of the PPI framework include LLM-assisted rank-set construction
 771 (Chatzi et al., 2024), average treatment effects from multiple datasets (Wang et al., 2025), clinic trial
 772 outcomes (Poulet et al., 2025), evaluating the accuracy of machine learning systems (Boyeau et al.,
 773 2025; Park et al., 2025), machine learning generated surrogate rewards for multi-armed bandits (Ji
 774 et al., 2025b). A few other works have explored alternative machine-learning assisted estimators,
 775 e.g. Schmutz et al. (2023); Egami et al. (2023); Miao et al. (2023); Miao & Lu (2024); Gan et al.
 776 (2024).

777 **Active Learning.** Similar to the Active Statistical Inference protocol is active learning (Settles,
 778 2009; Dasgupta, 2011; Hanneke, 2014), which frames learning as a process in which a machine
 779 learning model selectively queries unlabeled instances to be labeled by an oracle. In contrast to
 780 Active Statistical Inference, which focuses on enhancing statistical inference, the goal of active
 781 learning is to improve the predictive power of the machine learning model through the strategic use
 782 of labeled data. One of the main approaches in active learning, and the one most closely aligned with
 783 active estimation, is the uncertainty sampling strategy (Schohn & Cohn, 2000; Tong & Koller, 2000;
 784 Tur et al., 2005; Joshi et al., 2009; Gal et al., 2017; Ducoffe & Precioso, 2018; Beluch et al., 2018;
 785 Ren et al., 2021), where the model aims to query the labels of the most informative data points, i.e.,
 786 the ones which the model is most uncertain about.

787 **Importance Sampling.** The idea of prioritizing the most influential samples is also closely re-
 788 lated to adaptive importance sampling (Owen, 2013), which is a sequential scheme that updates
 789 the proposal distribution by learning from previously sampled values in order to better approximate
 790 some property of a target distribution. Active mean estimation is analogous to adaptive importance
 791 sampling, in that it adaptively selects samples that contribute most to reducing the variance of the
 792 estimator based on previously observed data to improve the accuracy of the mean estimator.

793 **Non-Asymptotic Results in Zrnic & Candès (2024).** We acknowledge that in Appendix C in
 794 Zrnic & Candès (2024), the authors consider incorporating the notion of actively querying the
 795 ground truth into the technique for estimating means of bounded random variables proposed by
 796 Waudby-Smith & Ramdas (2024). Waudby-Smith & Ramdas (2024) leverage the duality between
 797 sequential hypothesis testing and the construction of confidence intervals. More concretely, their
 798 method reduces the task of constructing a confidence interval for the mean to a potentially in-
 799 finite number of hypothesis testing problems. Each hypothesis testing problem corresponds to
 800 whether the observed samples have a population mean equal to a specific value. Hence, for a
 801 continuous random variable, this corresponds to an infinite number of hypotheses. In practice,
 802 a discretization is used. The confidence interval \mathcal{C}_t at time t then consists of those hypothe-
 803 sized mean values that have not been rejected based on the data observed up to time t , i.e.,
 804 $\mathcal{C}_t := \{\nu : H_0^{(\nu)} \text{ is not rejected based on observations up to time } t\}$, where $H_0^{(\nu)}$ denotes the hy-
 805 pothesis that the population mean is ν . While Zrnic & Candès (2024) provide the valuable idea
 806 of integrating these techniques and also provide some simulation results, the specific step-by-step
 807 algorithmic details and theoretical guarantees for active sequential mean estimation remain to be
 808 explicitly elaborated.

809

810 C PROOFS OF THE THEORETICAL RESULTS IN SECTION 4

812 **Theorem 1.** Fix a time horizon $T \geq 4$. Assume each random variable g_t is bounded, i.e., $|g_t| \leq G$
 813 for a constant $G > 0$. Denote $\sigma_t^2 := \mathbb{E}[(g_t - \mu_y)^2 | \mathcal{F}_{t-1}]$ the conditional variance, where \mathcal{F}_{t-1} is
 814 the filtration up to $t-1$. Then, for any $\delta \in (0, 1/e)$, with probability at least $1 - \delta$, $\forall t \in [T]$:

$$816 \quad |w_{t+1} - \mu_y| \leq \frac{2 \max \left\{ 2\sqrt{S_t}, (G + |\mu_y|) \sqrt{\log \left(\frac{\log(T)}{\delta} \right)} \right\} \sqrt{\log \left(\frac{\log(T)}{\delta} \right)}}{T} + \left(1 - \frac{t}{T}\right) |\mu_y|,$$

820 where $S_t = \sum_{s=1}^t \sigma_s^2$.

823 *Proof.* The difference of $w_t - \mu_y$ can be decomposed into two terms. Specifically, we have

$$825 \quad w_{t+1} - \mu_y = \left(\frac{1}{T} \sum_{s=1}^t g_s \right) - \mu_y = \frac{1}{T} \sum_{s=1}^t (g_s - \mu_y) - \left(1 - \frac{t}{T}\right) \mu_y. \quad (7)$$

828 Let us analyze the second-to-last-term $\sum_{s=1}^t (g_s - \mu_y)$ on (7). We note that $(g_s - \mu_y)_{s \geq 1}$ forms a
 829 martingale difference sequence, i.e., $\mathbb{E}[g_s - \mu_y | \mathcal{F}_{s-1}] = 0$, where \mathcal{F}_{s-1} encodes all information
 830 up to time $s-1$. Furthermore, $\forall s : |g_s - \mu_y| \leq G + |\mu_y|$. Also, the conditional variance is

$$832 \quad \text{Var}(g_s - \mu_y | \mathcal{F}_{s-1}) = \mathbb{E}[(g_s - \mu_y)^2 | \mathcal{F}_{s-1}] = \sigma_s^2$$

834 By Freedman's inequality (Lemma 1), we have, with probability $1 - \delta$,

$$835 \quad \forall t \in [T] : \sum_{s=1}^t (g_s - \mu_y) \leq 2 \max \left\{ 2 \sqrt{\sum_{s=1}^t \sigma_s^2}, (G + |\mu_y|) \sqrt{\log(\log(T)/\delta)} \right\} \sqrt{\log(\log(T)/\delta)}. \quad (8)$$

840 Combining (7) and (8), we obtain the following holds simultaneously at all $t \in [T]$, with probability
 841 at least $1 - \delta$:

$$843 \quad |w_{t+1} - \mu_y| \\ 844 \quad \leq \frac{2 \max \left\{ 2 \sqrt{\sum_{s=1}^t \sigma_s^2}, (G + |\mu_y|) \sqrt{\log(\log(T)/\delta)} \right\} \sqrt{\log(\log(T)/\delta)}}{T} + \left(1 - \frac{t}{T}\right) |\mu_y|.$$

848 \square

851 D PROOFS OF THE THEORETICAL RESULTS IN SECTION 5

853 **Lemma 2.** The conditional variance has the following decomposition:

$$855 \quad \mathbb{E}[(g_t - \mu_y)^2 | \mathcal{F}_{t-1}] \\ 856 \quad = \mathbb{E} \left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1} \right] + \mathbb{E} \left[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1} \right] + 2\mathbb{E} \left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1} \right] - \mu_y^2.$$

859 Furthermore, assume that the query policy at time t is \mathcal{F}_{t-1} -measurable, i.e., there exists a random
 860 variable $p_t \in [0, 1]$, measurable with respect to \mathcal{F}_{t-1} , such that $\pi_t(x_t) = p_t$. Then, we have

$$862 \quad \mathbb{E} \left[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1} \right] = \frac{1}{p_t} \mathbb{E} \left[(y_t - f_t(x_t))^2 \middle| \mathcal{F}_{t-1} \right].$$

864 *Proof.*

$$\begin{aligned}
 & \mathbb{E}[(g_t - \mu_y)^2 | \mathcal{F}_{t-1}] \\
 &= \mathbb{E}[g_t^2 | \mathcal{F}_{t-1}] - \mu_y^2 \\
 &= \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + \mathbb{E}\left[(y_t - f_t(x_t))^2 \frac{\xi_t^2}{\pi_t^2(x_t)} \middle| \mathcal{F}_{t-1}\right] + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \frac{\xi_t}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1}\right] \\
 &\quad - \mu_y^2 \\
 &= \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + \mathbb{E}\left[(y_t - f_t(x_t))^2 \frac{1}{\pi_t(x_t)} \middle| \mathcal{F}_{t-1}\right] + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1}\right] - \mu_y^2,
 \end{aligned}$$

875 where the first equality follows from that $\mathbb{E}[g_t | \mathcal{F}_{t-1}] = \mu_y$. \square

877 **Lemma 5.** Assume that the range of oracle's output is bounded, i.e., $\forall t : \Phi(x_t) \leq B$, for a constant
878 $B > 0$. Then, $\forall t : \left| \dot{\ell}_t \right|^2 \leq \frac{B^2}{\beta^4}$.

880 *Proof.*

$$\forall t : \left| \dot{\ell}_t \right|^2 = \frac{\Phi_t^2(x_t)}{p_t^4} \leq \frac{B^2}{p_t^4} \leq \frac{B^2}{\beta^4},$$

884 where the first inequality follows from that $\forall t : \Phi(x_t) \leq B$, and the last inequality uses that
885 $\forall t : p_t \geq \beta$. \square

886 **Theorem 2.** Assume that there is an oracle that outputs $\Phi_t(x_t)$ at each t such that $\frac{1}{c_1} \Phi_t(x_t) \leq$
887 $\mathbb{E}\left[(y_t - f_t(x_t))^2 \middle| \mathcal{F}_{t-1}\right] \leq c_0 \Phi_t(x_t)$ for some constant $c_0, c_1 > 0$ and that $\forall t : \Phi(x_t) \leq B$. Set
888 the parameter $\gamma = \frac{1}{\sqrt{T}} \frac{\beta^2}{B}$. Using the query policy (4), we have that, with probability at least $1 - \delta$,
889 $\forall t \in [T]$:

$$|w_{t+1} - \mu_y| \leq \frac{2 \max \left\{ 2\sqrt{\Psi_t}, (G + |\mu_y|) \sqrt{\log \left(\frac{\log(T)}{\delta} \right)} \right\} \sqrt{\log \left(\frac{\log(T)}{\delta} \right)}}{T} + \left(1 - \frac{t}{T} \right) |\mu_y|,$$

890 for any $\beta \in (0, \tau)$ and $\tau \in (0, 1)$, where $\Psi_t \leq c_0 c_1 \sigma_{1:t}^{*2} + 2c_0 \frac{t}{\sqrt{T}} \frac{B}{\beta^2} + \frac{c_0(R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p))\sqrt{T}B}{\beta^2}$.

891 *Proof.* By Lemma 2 and the constraint that the query probability p_t is fully determined in \mathcal{F}_{t-1} , we
892 have

$$\begin{aligned}
 \sigma_t^2 &= \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + \frac{1}{p_t} \mathbb{E}\left[(y_t - f_t(x_t))^2 \middle| \mathcal{F}_{t-1}\right] + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1}\right] - \mu_y^2 \\
 &\leq \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + c_0 \frac{\Phi_t(x_t)}{p_t} + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1}\right] - \mu_y^2 \\
 &= \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + c_0 \left(\frac{\Phi_t(x_t)}{p_{1:t}^*} + \frac{\Phi_t(x_t)}{p_t} - \frac{\Phi_t(x_t)}{p_{1:t}^*} \right) + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1}\right] \\
 &\quad - \mu_y^2 \\
 &\leq \mathbb{E}\left[f_t(x_t)^2 \middle| \mathcal{F}_{t-1}\right] + \frac{c_0 c_1}{p_{1:t}^*} \mathbb{E}\left[(y_t - f_t(x_t))^2 \middle| \mathcal{F}_{t-1}\right] + c_0 \left(\frac{\Phi_t(x_t)}{p_t} - \frac{\Phi_t(x_t)}{p_{1:t}^*} \right) \\
 &\quad + 2\mathbb{E}\left[f_t(x_t)(y_t - f_t(x_t)) \middle| \mathcal{F}_{t-1}\right] - \mu_y^2 \\
 &\leq c_0 c_1 \sigma_{t,(t)}^{*2} + c_0 \left(\frac{\Phi_t(x_t)}{p_t} - \frac{\Phi_t(x_t)}{p_{1:t}^*} \right),
 \end{aligned}$$

918 where the last inequality uses the fact that $c_0 c_1 \geq 1$. We note that the above inequality holds for all
 919 t . Hence, we have

$$920 \quad 921 \quad \sum_{s=1}^t \sigma_s^2 \leq c_0 c_1 \sigma_{1:t}^{*2} + c_0 \text{Regret}_t(p_{1:t}^*), \quad (9)$$

923 by summing up the above inequality for each round. To proceed, we use the regret bound that we
 924 have from Lemma 4:

$$925 \quad 926 \quad \text{Regret}_t(p_{1:t}^*) \leq 2\gamma \sum_{s=1}^t |\dot{\ell}_s|^2 + \frac{R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p)}{\gamma}$$

$$927 \quad 928 \quad \leq 2\gamma t \frac{B^2}{\beta^4} + \frac{R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p)}{\gamma}$$

$$929 \quad 930 \quad \stackrel{(i)}{=} 2\gamma t \frac{B}{\beta^2} + \frac{(R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p)) \sqrt{T} B}{\beta^2}, \quad (10)$$

$$931 \quad 932$$

933 where (i) is from Lemma 5 and (ii) is by the choice of $\gamma = \frac{1}{\sqrt{T}} \frac{\beta^2}{B}$. Combining (9) and (10), we have

$$934 \quad 935 \quad \sum_{s=1}^t \sigma_s^2 \leq c_0 c_1 \sigma_{1:t}^{*2} + 2c_0 \frac{t}{\sqrt{T}} \frac{B}{\beta^2} + \frac{c_0 (R(p_{1:t}^*) - \min_{p \in \mathcal{K}} R(p)) \sqrt{T} B}{\beta^2}.$$

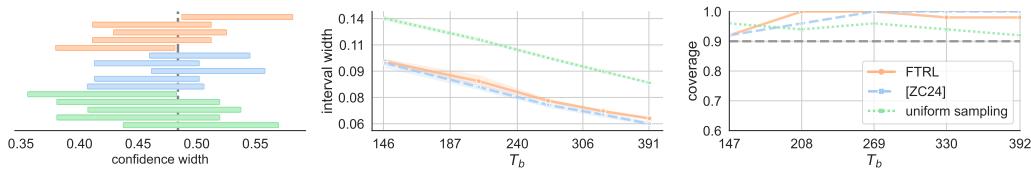
936 Using the above bound together with Theorem 1 leads to the result. This completes the proof. \square

937 E ADDITIONAL EXPERIMENTAL DETAILS

938 E.1 SYNTHETIC DATASET

939 The fourth dataset used in the experiments is a synthetic dataset that is generated for binary clas-
 940 sification according to a logistic model. More specifically, the covariates $x_t \in \mathbb{R}^d$ are inde-
 941 pendently drawn from a multivariate normal distribution with zero mean and identity covariance I_d ,
 942 where $d = 10$. The true parameter vector w^* is sampled independently from a normal distribution
 943 with zero mean and covariance $0.5 \cdot I_d$. Gaussian noise $\epsilon_t \sim \mathcal{N}(0, 10^{-5})$ is added to each $x_t^\top w^*$
 944 to produce the logits. The corresponding binary labels $y_t \in \{0, 1\}$ are then generated according to
 945 $y_t \sim \text{Bernoulli}(\sigma(x_t^\top w^* + \epsilon_t))$, where $\sigma(z) = 1/(1 + e^{-z})$ denotes the sigmoid function. The ML
 946 model $f_t(\cdot)$ is implemented as a logistic regression model whose uncertainty predictor is estimated
 947 in the same way as in the post-election survey dataset, i.e., $u_t(x_t, f_t(\cdot)) = 2 \min\{f_t(x_t), 1 - f_t(x_t)\}$
 948 using the predicted probabilities of $f_t(\cdot)$. A linear regression model is trained to predict the approx-
 949 imation oracle $\Phi(x_t)$ as in the previous tasks.

950 Figure 6 shows the experimental results on the synthetic dataset. \square



951 **Figure 6: Synthetic dataset.** Left: Intervals of randomly selected trials. Middle: Average confidence width
 952 across repeated trials vs. sampling budget T_b . Right: Percentage of trials that cover the true mean.

953 E.2 EXPERIMENTAL SETUP

954 All experiments were repeated over 50 trials, and reported results correspond to the averages across
 955 these trials. At the start of each trial, the data points were randomly permuted. For each experiment,

972 the budget T_b was varied over five uniformly spaced values between 15% and 40% of the total number
 973 of data points T . The interval width and coverage plots were obtained by linearly interpolating
 974 between the values at these grid points.

975 In the experimental setup, the ML model $f_t(\cdot)$, uncertainty estimator $u_t(\cdot, \cdot)$, and oracle $\Phi_t(\cdot)$ were
 976 updated periodically after observing a batch of B data points. For the first two datasets (politeness
 977 score and wine review analysis), the estimators were updated $N = 50$ times, while for the last two
 978 datasets (post-election survey and synthetic data), they were updated $N = 10$ times. Accordingly,
 979 the batch size was set to $B = \text{round}\left(\frac{T_b}{N}\right)$.
 980

981 For the FTRL policy, we set the upper bound hyperparameter to $\tau = \frac{T_b}{T}$ in accordance with our
 982 theoretical analysis to ensure that the query probability satisfies the sampling constraint. The lower
 983 bound hyperparameter was set to $\beta = \frac{\tau}{8} > 0$, to prevent the sampling probability from becoming
 984 too small, and thereby encouraging exploration, while still remaining sufficiently below τ so that the
 985 resulting sampling interval is non-trivial and allows the algorithm to adjust the sampling probability
 986 over time. Furthermore, the hyperparameter γ was chosen as $\gamma = \frac{1}{\sqrt{T}}$, in line with common practice
 987 in online learning, to guarantee sublinear regret growth with respect to T , which is necessary for
 988 achieving no-regret performance. For the policy of Zrnic & Candes (2024), we set the mixing
 989 hyperparameter to $\lambda = 0.5$, which is the recommended value used in their experimental setup, to
 990 enable a comparison with their proposed policy.
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000
 1001
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025