Online Adaptation to Class-Prior Shift from Positive and Unlabeled Data

Anonymous authors
Paper under double-blind review

ABSTRACT

Class-prior shift is a common distribution shift in which the class-conditional distributions remain stable while the class proportions change. Existing online label-shift methods assume labeled source data from every class. We ask: *what is the statistical price of online adaptation when the source supervision itself is incomplete?* We study this question when the source data contain only positive and unlabeled (PU) examples, while unlabeled target observations arrive sequentially and the class prior changes over time. Such data arise, for example, in protein screening, where only sequences that pass the screen are labeled. We show that labeled negatives are not needed to preserve the online tracking rate. Under a source-sample stability condition, the temporal component of the cumulative squared prior-tracking error retains the $\widetilde{\mathcal{O}}(T^{1/3}V_T^{2/3})$ dependence on the horizon T and total variation V_T achieved with positive and negative source data. The effect of PU supervision appears in the finite-source term: reconstructing the unobserved negative class amplifies source-estimation errors with weights that depend on where the target-prior path lies, beyond V_T alone. We further derive a dynamic classification-regret bound for online threshold adaptation against the best threshold of the same fixed score at each target prior. Experiments on image and text benchmarks show that threshold adaptation improves classification under substantial prior shift with low online cost without retraining the initial model.

1 INTRODUCTION

Distribution shift is a fundamental challenge in machine learning. The distribution encountered after training need not coincide with the one from which the training data were collected (Quiñonero-Candela et al., 2008; Koh et al., 2021). Among its basic forms is *label shift*, or *class-prior shift*, in which the class-conditional distributions remain unchanged while the class proportions change. Existing methods use this structure to adapt a classifier trained on the source distribution to a single fixed target distribution (Saerens et al., 2002; Lipton et al., 2018; Garg et al., 2020). However, in real-world scenarios on deployed systems, the target population may continue to evolve, which has motivated online adaptation to continuously changing covariate distributions (Zhang et al., 2023) and to time-varying label distributions (Wu et al., 2021; Bai et al., 2022; Baby et al., 2023).

These online label shift methods track a time-varying label marginal from unlabeled target observations. They need neither target labels nor prior knowledge of how much the label marginal varies. However, they assume labeled offline data from every class (Wu et al., 2021; Bai et al., 2022; Baby et al., 2023), from which source quantities such as the confusion matrix are estimated.

Source supervision may be incomplete. In *Positive–Unlabeled* (PU) learning, labeled observations are available from the positive class together with an unlabeled mixture, but no labeled negative examples are observed (Elkan & Noto, 2008; du Plessis et al., 2014; Bekker & Davis, 2020). For example, in a high-throughput protein screen, sequences that pass the screen are labeled positive, while the library as a whole is an unknown mixture of active and inactive sequences (Song et al., 2021). Negatives are hard to obtain, because it is easier to confirm activity than to rule it out (Li et al., 2022). The proportion of active sequences can also change across batches. PU class-prior estimation and classification have been studied in static settings (du Plessis et al., 2017; Charoenphakdee & Sugiyama, 2019; Nakajima & Sugiyama, 2023; Mielniczuk et al., 2026). However, substantially less is understood when such data must support adaptation to a target distribution that keeps changing.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

This leads to the question studied in this paper: *what is the statistical price of online adaptation when the source supervision is incomplete?* In particular, we ask whether weak source supervision makes a changing class prior fundamentally harder to track or only adds a separate statistical cost.

Our answer is that the two difficulties separate. Online adaptation has to track a class prior that changes over time. Since the class prior is not observed, it is estimated from the unlabeled online observations, and this estimate relies on statistics of the positive and negative classes that are estimated once from the finite PU data. We measure the accuracy of tracking by the cumulative squared prior-tracking error and split it into two terms: a *temporal term*, due to the change of the class prior over time, and a *finite-source term*, due to estimating the class statistics from finite PU data. Our analysis shows that the lack of labeled negatives does not change how the temporal term depends on T and V_T: up to logarithmic and conditioning factors, it grows as $T^{1/3}V_T^{2/3}$ in the horizon T and the total variation V_T of the class prior, as in online label shift with labeled offline data from every class (Bai et al., 2022; Baby et al., 2023).

Our second result concerns the finite-source term. Unlike the temporal term, it depends not only on T and V_T but also on where the class prior moves: up to logarithmic and conditioning factors, it scales as $(1-\pi_0)^{-2}\{\sum_t(1-\pi_t)^2/n_u + \sum_t(\pi_t-\pi_0)^2/n_+\}$, where π_0 is the class prior of the source distribution and n_+ and n_u are the numbers of positive and unlabeled offline examples. Two paths with the same T and V_T can therefore have different finite-source terms. We also prove a lower bound showing that, in general, no estimator can avoid this dependence on the path. A finite-source term is also present when labeled negatives are available. Under PU supervision, however, the negative class has to be reconstructed from the positive and unlabeled data, and this reconstruction amplifies the estimation error. This is the origin of the factor $(1-\pi_0)^{-2}$: the fewer negatives the unlabeled data contain, the larger the amplification.

The tracking guarantees also yield guarantees for classification. Existing online label shift methods (Bai et al., 2022; Baby et al., 2023) assume labeled offline data from every class, and they adapt the classifier in two different ways. Bai et al. (2022) update the parameters of the classifier by online gradient steps and use the labeled offline data at every round. Concurrent work on online PU learning (Anonymous authors, 2026) also updates classifier parameters, in a protocol where labeled positives keep arriving online. Baby et al. (2023) keep the initial model fixed and adjust only its output to the estimated label marginal. We take the same approach of Baby et al. (2023) with PU offline data and unlabeled online observations. In the binary case, adjusting the output of a fixed model amounts to moving its decision threshold. The threshold rule is computed once from the PU offline data, and the raw offline examples are not needed during online adaptation. The tracking bound then gives a bound on the dynamic regret against the best threshold at each round. Further related work and comparisons are given in Appendix A.

Our contributions are summarized as follows.

- We derive an upper bound on the cumulative squared prior-tracking error that separates into a temporal term and a finite-source term (Proposition 2 and Theorem 1). Using an online regression oracle that does not require V_T, the temporal term grows as $T^{1/3}V_T^{2/3}$ up to logarithmic and conditioning factors, which is the same dependence as with labeled offline data from every class. The finite-source term depends on the path of the class prior through the two sums in the expression above (denoted by $W_{\mathrm{U}}(\pi_{1:T})$ and $W_+(\pi_{1:T})$ in Section 3).
- We derive a lower bound for the finite-source term with the same dependence on the path and the sample sizes when the sample sizes are not too small (Theorem 2), and a separate lower bound for the temporal term (Theorem 3).
- We propose a threshold rule for a fixed classification score, which is computed once from the PU offline data, and derive a bound on its dynamic regret against the best threshold at each round (Theorem 4 and Corollary 2).
- Across five image and text datasets and four target-prior paths, our method reduces the classification error of the PU initial model from 15.2% to 5.5%, while requiring substantially less online computation than the optimization-based baselines and neither online retraining nor access to raw offline data. Our sensitivity analysis shows that prior-tracking error increases as π_0 approaches one (Section 6).

2 Problem Formulation and Performance Metric

Online class prior shift with positive-unlabeled (PU) data. We consider binary classification on a measurable feature space $\mathcal{X}$ with labels $Y \in \{0, 1\}$. Let P_+ and P_- denote the distributions of X conditional on $Y = 1$ and $Y = 0$, respectively. For a positive-class prior $\pi \in [0, 1]$, define the corresponding unlabeled distribution by

$$P_{\mathrm{U},\pi} := \pi P_+ + (1 - \pi)P_-.$$

Class prior shift means that only the proportion of the two classes changes over time, while P_+ and P_- stay fixed. We study its online version, in which the learner receives an unlabeled target stream and adapts to a time-varying prior (Wu et al., 2021; Bai et al., 2022; Baby et al., 2023).

Before online adaptation, the learner receives PU offline data: positive examples from P_+ and unlabeled examples from P_{U,π_0}, but no examples drawn directly from P_-. We split these data into two independent parts. The training data are used to learn the initial model. The holdout data $S_0 := (S_+, S_{\mathrm{U},0})$ are used to estimate the quantities needed for online adaptation, where

$$S_+ := \{X_i^+\}_{i=1}^{n_+}, \quad X_i^+ \overset{\text{iid}}{\sim} P_+, \qquad S_{\mathrm{U},0} := \{X_i^{\mathrm{U},0}\}_{i=1}^{n_u}, \quad X_i^{\mathrm{U},0} \overset{\text{iid}}{\sim} P_{\mathrm{U},\pi_0}.$$

We assume the source prior $\pi_0 \in (0, 1)$ is known. Estimating it from PU data is a separate mixture-proportion-estimation problem (Blanchard et al., 2010; Scott, 2015; Ramaswamy et al., 2016; du Plessis et al., 2017).

In the online adaptation stage, P_+ and P_- remain fixed while the positive-class prior changes along a sequence $\pi_{1:T} = (\pi_1, \ldots, \pi_T)$. Online adaptation is *unsupervised*. At round t, the learner receives a batch of $N_t \geq 1$ unlabeled observations

$$X_{t,1}, \ldots, X_{t,N_t} \overset{\text{iid}}{\sim} P_{\mathrm{U},t}, \qquad P_{\mathrm{U},t} = \pi_t P_+ + (1 - \pi_t)P_-,$$

where the class prior π_t is latent and no target labels are revealed. Its drift is measured by the total variation $V_T := \sum_{t=2}^{T} |\pi_t - \pi_{t-1}|$, which is unknown to the learner. We compare throughout with the *positive–negative* (PN) baseline, which receives the same positive sample and a labeled negative source sample in place of $S_{\mathrm{U},0}$. For simplicity, the theoretical analysis below is stated for $N_t = 1$.

Online fixed-score classification. We use a measurable classification score $g : \mathcal{X} \to [0, 1]$ and a bounded statistic $\phi : \mathcal{X} \to [0, 1]$. The statistic ϕ gives the scalar observation from which the current prior is estimated (Sections 3 and 4), and g is thresholded to classify (Section 5). Taking $\phi = g$ is admissible and recovers the single-score formulation. Both are fixed before S_0 is observed, for example learned from the training data, and neither is updated during online adaptation. The raw samples in S_0 are not accessed during online adaptation.

For a threshold $\tau \in \mathbb{R}$, define $h_\tau(x) := \mathbf{1}\{g(x) \geq \tau\}$, where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Before online adaptation, the learner computes a few quantities from S_0: the sample means of ϕ on S_+ and $S_{\mathrm{U},0}$ (specified later in Section 3) and a threshold map (Section 5). At the beginning of round t, the learner uses only these quantities and the past unlabeled observations $X_1, \ldots, X_{t-1}$ to choose a prior estimate $\widehat{\pi}_t \in [0, 1]$ and a threshold $\widehat{\tau}_t \in \mathbb{R}$. It then observes $X_t \sim P_{\mathrm{U},t}$ and predicts its label by $\widehat{h}_t(X_t) := \mathbf{1}\{g(X_t) \geq \widehat{\tau}_t\}$. Unless stated otherwise, the latent prior path is deterministic, and all source and target observations are mutually independent.

Prior-tracking error. We measure how well $\widehat{\pi}_t$ follows π_t by the cumulative squared prior-tracking error $L_T^{\mathrm{trk}} := \sum_{t=1}^{T} (\widehat{\pi}_t - \pi_t)^2$. This is the main quantity of our analysis: the temporal error and the finite-source error separate in it, and Section 4 gives upper and lower bounds on it. Section 5 turns these bounds into a bound on the classification regret defined next.

Dynamic classification regret. For a classifier $h : \mathcal{X} \to \{0, 1\}$, its population zero–one risk under prior π is $R_\pi(h) := \pi\, P_+(h(X) = 0) + (1 - \pi)\, P_-(h(X) = 1)$. Hence thresholding g has risk $R_\pi(h_\tau) = R_\pi^g(\tau)$ with

$$R_\pi^g(\tau) = \pi P_+(g(X) < \tau) + (1 - \pi)P_-(g(X) \geq \tau).$$

We evaluate the downstream classification performance by the *dynamic threshold-classification regret*, the cumulative excess risk against the best threshold at each round's true prior:

$$\mathrm{Reg}_T^{\mathrm{thr}} := \sum_{t=1}^{T} \Big(R_{\pi_t}^g(\widehat{\tau}_t) - \inf_{\tau \in \mathbb{R}} R_{\pi_t}^g(\tau) \Big).$$

The comparator is the best threshold of g for each π_t and changes with the target prior. It is restricted to thresholds of the fixed score g. For a general g, this regret differs from the regret against the Bayes classifier. The two coincide when g is the exact source posterior (Appendix D.4). The threshold rule, computed from the PU holdout data, is specified in Section 5.

3 Estimating the Statistic–Prior Map from PU Source Data

To estimate the class prior from unlabeled target data, we first need to determine how the mean of $\phi(X)$ varies with the prior. This section estimates this relation from the PU holdout data and quantifies the error caused by the finite source samples.

3.1 Identifying the Statistic–Prior Map from PU Data

Let $\mu_+ := \mathbb{E}_{P_+}[\phi(X)]$, $\mu_- := \mathbb{E}_{P_-}[\phi(X)]$, and $\mu_0 := \mathbb{E}_{P_{\mathrm{U},\pi_0}}[\phi(X)]$. Since $P_{\mathrm{U},\pi_0} = \pi_0 P_+ + (1-\pi_0)P_-$, the PU mixture identity (du Plessis et al., 2014; Bekker & Davis, 2020) gives

$$\mu_0 = \pi_0\mu_+ + (1-\pi_0)\mu_-, \qquad \mu_- = \frac{\mu_0 - \pi_0\mu_+}{1-\pi_0}.$$

Thus μ_+ and μ_0 are estimable from the positive and source-unlabeled holdout samples, while μ_- is reconstructed from them. For $\pi \in [0,1]$, the population statistic–prior map is

$$M(\pi) := \mathbb{E}_{P_{U,\pi}}[\phi(X)] = (1-\pi)\mu_- + \pi\mu_+ \tag{1}$$

$$= \frac{1-\pi}{1-\pi_0}\,\mu_0 + \frac{\pi-\pi_0}{1-\pi_0}\,\mu_+. \tag{2}$$

The map is affine in π, with slope $d := \mu_+ - \mu_- = (\mu_+ - \mu_0)/(1-\pi_0)$. We assume that d has a known lower bound $d_{\min} > 0$, meaning that ϕ is larger on average for positive examples. The separation d controls how errors in the source means propagate to the prior estimate. If μ_- and d were known, $(\phi(X_t) - \mu_-)/d$ would have conditional mean π_t, reducing prior estimation to scalar mean tracking—the binary analogue of black-box label-shift estimation (Lipton et al., 2018; Garg et al., 2020).

Let $\widehat{\mu}_+ := n_+^{-1}\sum_{i=1}^{n_+}\phi(X_i^+)$ and $\widehat{\mu}_0 := n_u^{-1}\sum_{i=1}^{n_u}\phi(X_i^{\mathrm{U},0})$ be the sample means on the holdout data. Replacing μ_0 and μ_+ by these in Eq. (2) gives the estimated map

$$\widehat{M}(\pi) := \frac{1-\pi}{1-\pi_0}\,\widehat{\mu}_0 + \frac{\pi-\pi_0}{1-\pi_0}\,\widehat{\mu}_+. \tag{3}$$

Its value at $\pi = 0$ and its slope provide estimates of μ_- and d:

$$\widehat{\mu}_- := \widehat{M}(0) = \frac{\widehat{\mu}_0 - \pi_0\widehat{\mu}_+}{1-\pi_0}, \qquad \widehat{d} := \widehat{M}(1) - \widehat{M}(0) = \widehat{\mu}_+ - \widehat{\mu}_-, \tag{4}$$

so that $\widehat{M}(\pi) = \widehat{\mu}_- + \widehat{d}\,\pi$. These quantities are computed once from the holdout data and used throughout online adaptation.

3.2 Path-dependent finite-source error

Since $\widehat{\mu}_-$ is reconstructed from $\widehat{\mu}_+$ and $\widehat{\mu}_0$, the map error $\widehat{M}(\pi) - M(\pi)$ depends only on the two source-estimation errors $\Delta_+ := \widehat{\mu}_+ - \mu_+$ and $\Delta_0 := \widehat{\mu}_0 - \mu_0$. Along a prior path $\pi_{1:T}$ their cumulative squared weights are

$$W_{\mathrm{U}}(\pi_{1:T}) := \sum_{t=1}^{T}(1-\pi_t)^2, \qquad W_+(\pi_{1:T}) := \sum_{t=1}^{T}(\pi_t - \pi_0)^2. \tag{5}$$

For $\delta \in (0,1)$, set $\mathrm{rad}_+(\delta) := \sqrt{\log(4/\delta)/(2n_+)}$ and $\mathrm{rad}_0(\delta) := \sqrt{\log(4/\delta)/(2n_u)}$, and define $\mathcal{E}_{\mathrm{src}}(\delta)$ as the event on which $|\Delta_+| \le \mathrm{rad}_+(\delta)$ and $|\Delta_0| \le \mathrm{rad}_0(\delta)$. Hoeffding's inequality gives $\Pr(\mathcal{E}_{\mathrm{src}}(\delta)) \ge 1-\delta$. The following proposition exposes the path dependence: the two source errors enter with coefficients $1-\pi_t$ and $\pi_t - \pi_0$. Their cumulative effect therefore depends on the location of the prior path, not only on its total variation.

Proposition 1. *For every $\pi \in [0,1]$,*

$$\widehat{M}(\pi) - M(\pi) = \frac{1-\pi}{1-\pi_0}\,\Delta_0 + \frac{\pi-\pi_0}{1-\pi_0}\,\Delta_+. \tag{6}$$

Moreover, on $\mathcal{E}_{\mathrm{src}}(\delta)$,

$$\sum_{t=1}^{T} |\widehat{M}(\pi_t) - M(\pi_t)|^2 \le \frac{\log(4/\delta)}{(1-\pi_0)^2}\left\{\frac{W_{\mathrm{U}}(\pi_{1:T})}{n_u} + \frac{W_+(\pi_{1:T})}{n_+}\right\}. \tag{7}$$

The online method in Section 4 divides $\phi(X_t) - \widehat{\mu}_-$ by $\widehat{d}$. Thus $\widehat{d}$ must remain bounded away from zero. The same source event provides the required uniform control. Define

$$\mathrm{rad}_{\mathrm{PU}}(\delta) := \max\left\{\mathrm{rad}_+(\delta),\ \frac{\mathrm{rad}_0(\delta) + \pi_0\,\mathrm{rad}_+(\delta)}{1-\pi_0}\right\}. \tag{8}$$

Let $\varepsilon_{\mathrm{map}} := \|\widehat{M} - M\|_{\infty,[0,1]}$ denote the uniform error of the estimated map. On $\mathcal{E}_{\mathrm{src}}(\delta)$,

$$\varepsilon_{\mathrm{map}} \le \mathrm{rad}_{\mathrm{PU}}(\delta), \qquad |\widehat{d} - d| \le 2\,\mathrm{rad}_{\mathrm{PU}}(\delta). \tag{9}$$

Thus, $\mathrm{rad}_{\mathrm{PU}}(\delta) \le d_{\min}/4$ ensures $\widehat{d} \ge d_{\min}/2 > 0$. Proofs and the comparison with PN source supervision are given in Appendices B.1 and B.2.

4 Online Prior Tracking with Finite PU Source Data

The map $\widehat{M}$ is estimated once from the holdout data S_0 and kept fixed throughout online adaptation. Its estimation error persists over all T rounds. We reduce prior tracking to online regression on the resulting distorted path and control its gap from the true path using Proposition 1.

4.1 Reduction to online regression and tracking a prior path

Fix S_0 and suppose that $\widehat{d} > 0$. Inverting the estimated map $\widehat{M}(\pi) = \widehat{\mu}_- + \widehat{d}\,\pi$ at the observed statistic gives the transformed observation

$$Z_t := \frac{\phi(X_t) - \widehat{\mu}_-}{\widehat{d}}. \tag{10}$$

Although Z_t need not lie in $[0,1]$, its conditional mean is

$$\pi_t^\dagger := \mathbb{E}[Z_t \mid S_0] = \frac{M(\pi_t) - \widehat{\mu}_-}{\widehat{d}}. \tag{11}$$

If $\widehat{M} = M$, then $\pi_t^\dagger = \pi_t$; their difference is caused by estimating M from the finite source data. Given S_0, the noise $Z_t - \pi_t^\dagger$ is independent across rounds, mean zero, and bounded.

At the beginning of round t, the online regression algorithm uses $Z_1, \dots, Z_{t-1}$ to produce a real-valued prediction $x_t \in \mathbb{R}$ of $\pi_t^\dagger$. Before observing X_t and hence Z_t, the learner outputs the valid prior estimate $\widehat{\pi}_t := \mathrm{proj}_{[0,1]}(x_t)$. The following proposition decomposes the tracking error L_T^{trk} into an online-regression error for $\pi_{1:T}^\dagger$ and a fixed source-map error.

Proposition 2. *Condition on S_0 and suppose that $\widehat{d} > 0$. For every t,*

$$\pi_t^\dagger - \pi_t = \frac{M(\pi_t) - \widehat{M}(\pi_t)}{\widehat{d}}. \tag{12}$$

For any predictions $x_{1:T} \in \mathbb{R}^T$, the projected estimates $\widehat{\pi}_t := \mathrm{proj}_{[0,1]}(x_t)$ satisfy, pointwise,

$$L_T^{\mathrm{trk}} = \sum_{t=1}^{T}(\widehat{\pi}_t - \pi_t)^2 \le 2\sum_{t=1}^{T}(x_t - \pi_t^\dagger)^2 + \frac{2}{\widehat{d}^2}\sum_{t=1}^{T}\left|M(\pi_t) - \widehat{M}(\pi_t)\right|^2. \tag{13}$$

Moreover, if $\varepsilon_{\mathrm{map}} \le r \le d_{\min}/4$, then $\widehat{d} \ge d_{\min}/2$ and $V_T(\pi^\dagger) \le 2V_T(\pi)$, where $V_T(\cdot)$ denotes the total variation of a path.

The first term in Eq. (13) measures how well the online predictor tracks the distorted path $\pi^{\dagger}_{1:T}$. The second one is the persistent error caused by estimating M from S_0, which Eq. (7) controls.

The main tracking guarantee below shows that, up to logarithmic and conditioning factors, *the temporal component retains the $T^{1/3}V_T^{2/3}$ dependence achieved with labeled source data from both classes, while finite PU source uncertainty appears as a separate path-dependent term involving W_U and W_+*. We control the temporal component using ALIGATOR (Baby et al., 2021), in the form of Baby et al. (2023, Proposition 20). Its guarantee depends on the variation of the path being tracked, and Proposition 2 gives $V_T(\pi^{\dagger}) \le 2V_T(\pi)$. The method adapts to V_T without taking it as an input. We use the learning-rate choice specified in Proposition 4, run ALIGATOR on $\widetilde{Z}_t := (d_{\min}/2)Z_t$, and rescale its prediction $\widetilde{x}_t$ to $x_t := 2\widetilde{x}_t/d_{\min}$ before projection.

Theorem 1. *Fix $\delta_{\mathrm{off}} \in (0,1)$ and suppose that $\mathrm{rad}_{\mathrm{PU}}(\delta_{\mathrm{off}}) \le d_{\min}/4$. For the ALIGATOR-based estimates $\widehat{\pi}_t$ above, there are universal constants $C, c > 0$ such that, with probability at least $1-\delta_{\mathrm{off}}$ over the holdout data S_0,*

$$\mathbb{E}\big[L_T^{\mathrm{trk}} \mid S_0\big] \le C\log^c(eT)\left\{\frac{1}{d_{\min}^2} + \frac{T^{1/3}V_T^{2/3}}{d_{\min}^{4/3}}\right\} + \frac{8\log(4/\delta_{\mathrm{off}})}{d_{\min}^2(1-\pi_0)^2}\left\{\frac{W_U(\pi_{1:T})}{n_u} + \frac{W_+(\pi_{1:T})}{n_+}\right\}. \quad (14)$$

The finite-source term pairs W_U/n_u with the error in $\widehat{\mu}_0$ and W_+/n_+ with the error in $\widehat{\mu}_+$. The factor $(1-\pi_0)^{-2}$ reflects the cost of reconstructing the unobserved negative mean. Since $W_U, W_+ \le T$, replacing both weights by T gives the worst-case form in Appendix C.5.

4.2 LOWER BOUNDS FOR SOURCE ESTIMATION AND TIME VARIATION

We show that each term of the upper bound is necessary through two lower bounds under different statistical experiments. In both, d is the separation of the constructed models and plays the role of $\mu_+ - \mu_-$. The first bound strengthens the learner by revealing the exact target distribution $P_{\mathrm{U},t}$ at every round. Even then, finitely many source samples cannot reliably distinguish the constructed models. The squared gaps between the induced prior paths add up to multiples of W_U and W_+. Therefore, additional target observations cannot remove the source term.

Theorem 2. *Fix $T, n_+, n_u \ge 1$, $\pi_0 \in (0,1)$, $d \in (0, 1/8]$, and any deterministic prior path $\pi_{1:T} \in [0,1]^T$. There exist three Bernoulli PU models Q_0, Q_U, Q_+. For each $j \in \{0, U, +\}$, its positive- and negative-class statistic means $\mu_+^{(j)}$ and $\mu_-^{(j)}$ lie in $[1/8, 7/8]$, and its separation $d_j := \mu_+^{(j)} - \mu_-^{(j)}$ lies in $[d, 3d/2]$. Let $\pi_{1:T}^{(j)}$ denote the target prior path under Q_j. The three models induce the same target distribution at every round, while $\pi_{1:T}^{(0)} = \pi_{1:T}$. For every estimator that observes the source holdout samples S_+ and $S_{\mathrm{U},0}$ and is also given the exact target distributions at all rounds,*

$$\max_{j\in\{0,U,+\}} \mathbb{E}_{Q_j} \sum_{t=1}^{T} (\widehat{\pi}_t - \pi_t^{(j)})^2 \ge c\left[W_U(\pi_{1:T}) \min\left\{1, \frac{1}{d^2(1-\pi_0)^2 n_u}\right\} + W_+(\pi_{1:T}) \min\left\{1, \frac{1}{d^2(1-\pi_0)^2 n_+}\right\}\right] \quad (15)$$

for a universal constant $c > 0$. Moreover, $V_T(\pi^{(U)}) \le V_T(\pi)$ and $V_T(\pi^{(+)}) \le V_T(\pi)$.

Proof sketch. The lower bound uses two pairs of models that are difficult to distinguish from finite source data. In the first pair, changing the negative-class mean affects the source-unlabeled distribution through the factor $1 - \pi_0$. Keeping the target distribution unchanged then changes the target prior in proportion to $1-\pi_t$, which yields the W_U term. In the second pair, changing both class means leaves the source-unlabeled distribution unchanged and changes the target prior in proportion to $\pi_t - \pi_0$, which yields the W_+ term. The full construction is given in Appendix C.7.

The second lower bound shows that the temporal term remains even when the class-conditional distributions are known. Define

$$\Pi_T(V) := \left\{\pi_{1:T} \in [3/8, 5/8]^T : \sum_{t=2}^{T} |\pi_t - \pi_{t-1}| \le V\right\}.$$

Theorem 3. *Let $T \geq 2$, $V \geq 0$, and $d \in (0, 1/8]$. Consider the Bernoulli label-shift submodel $P_+ = \mathrm{Bernoulli}(1/2 + d/2)$ and $P_- = \mathrm{Bernoulli}(1/2 - d/2)$, and reveal P_+ and P_- to the learner. There is a universal constant $c > 0$ such that even a batch estimator that observes the full target sequence before producing $\widehat{\pi}_{1:T}$ satisfies*

$$\inf_{\widehat{\pi}_{1:T}} \sup_{\pi_{1:T} \in \Pi_T(V)} \mathbb{E}_{\pi_{1:T}} \sum_{t=1}^{T} (\widehat{\pi}_t - \pi_t)^2 \geq c \min \left\{ T, \frac{1}{d^2} + \frac{T^{1/3} V^{2/3}}{d^{4/3}} \right\}. \tag{16}$$

Consequently, the same lower bound holds for online estimators.

Proof sketch. The constant-path term follows by applying Le Cam's two-point method to two nearby constant prior paths. For the variation term, we divide the horizon into blocks and let the prior take one of two nearby values on each block. The number of blocks and the gap between the two values are chosen so that every path satisfies the variation budget V and neighboring paths generate similar data distributions. Assouad's lemma then gives the $T^{1/3}V^{2/3}/d^{4/3}$ term (Baby & Wang, 2019; Besbes et al., 2015; Tsybakov, 2009). The complete proof is given in Appendix C.8.

When the source sample sizes are large enough for the lower bound to scale as $1/n_u$ and $1/n_+$, its dependence on W_U and W_+ matches that of the upper bound. Separately, the temporal lower bound shows that the $T^{1/3}V^{2/3}$ dependence is unavoidable even when the class-conditional distributions are known. The two lower bounds concern different statistical settings and establish the necessity of the source and temporal terms separately.

5 Dynamic Classification Regret

We next turn each prior estimate $\widehat{\pi}_t$ into a classification threshold for the fixed score g. Before online adaptation, the PU holdout data determine a map from candidate priors to estimated risk-minimizing thresholds. At round t, we evaluate this map at $\widehat{\pi}_t$, keeping g unchanged. The general prior-indexed decision reduction is given in Appendix D. Here we state its fixed-score threshold instantiation.

For a candidate prior $q \in [0, 1]$, recall that $R_q^g(\tau) = qP_+(g(X) < \tau) + (1-q)P_-(g(X) \geq \tau)$. The negative-class tail cannot be estimated directly from the PU holdout data. Let $\widehat{P}_+$ and $\widehat{P}_{\mathrm{U},0}$ denote the empirical distributions on S_+ and $S_{\mathrm{U},0}$. The PU mixture identity $P_{\mathrm{U},\pi_0} = \pi_0 P_+ + (1-\pi_0)P_-$ (du Plessis et al., 2014) gives the following empirical negative-class tail and threshold risk:

$$\widehat{P}_-^{\mathrm{PU}}(g(X) \geq \tau) := \frac{\widehat{P}_{U,0}(g(X) \geq \tau) - \pi_0 \, \widehat{P}_+(g(X) \geq \tau)}{1 - \pi_0}, \tag{17}$$

$$\widehat{R}_q^{g,\mathrm{PU}}(\tau) := q \, \widehat{P}_+(g(X) < \tau) + (1-q) \, \widehat{P}_-^{\mathrm{PU}}(g(X) \geq \tau). \tag{18}$$

Because $g \in [0, 1]$, the observed score values suffice to represent every threshold labeling of the holdout data; we add -1 and 2 for the all-positive and all-negative classifiers. Define

$$\mathcal{T}(S_0) := \{-1, 2\} \cup \{g(X) : X \in S_+ \cup S_{\mathrm{U},0}\}, \tag{19}$$

and, for each $q \in [0, 1]$, set

$$\widehat{\tau}^{\mathrm{PU}}(q) := \min \operatorname*{arg\,min}_{\tau \in \mathcal{T}(S_0)} \widehat{R}_q^{g,\mathrm{PU}}(\tau). \tag{20}$$

Choosing the smallest minimizer resolves ties and makes the resulting map $q \mapsto \widehat{\tau}^{\mathrm{PU}}(q)$ measurable.

For $\delta \in (0, 1)$, let $r_+(\delta) := \sqrt{\log(4/\delta)/(2n_+)}$ and $r_U(\delta) := \sqrt{\log(4/\delta)/(2n_u)}$ be the deviation radii for the positive and source-unlabeled score distributions, respectively. Define the PU risk radius at a candidate prior q by

$$\epsilon_R(q, \delta) := q r_+(\delta) + \frac{1-q}{1-\pi_0} \left\{ r_U(\delta) + \pi_0 r_+(\delta) \right\}. \tag{21}$$

Applying the Dvoretzky–Kiefer–Wolfowitz inequality to the two holdout score distributions gives, with probability at least $1 - \delta$ over the holdout data, $|\widehat{R}_q^{g,\mathrm{PU}}(\tau) - R_q^g(\tau)| \leq \epsilon_R(q, \delta)$ simultaneously for every $q \in [0, 1]$ and $\tau \in \mathbb{R}$. On the same event, minimizing the empirical risk gives excess threshold risk at most $2\epsilon_R(q, \delta)$ at every q (See Proposition 6).

Table 1: Classification error in percent on three representative datasets under a gradual Linear path and an abrupt Bernoulli path (lower is better). The smallest mean among the non-oracle methods in each row is bold. The PN columns use labeled negative offline data, whereas all PU columns use no labeled negatives. Complete results are in Table 2.

Dataset	STATIC	PU–ALIGATOR	PU–FLH–FTL	PN–ALIGATOR	PN–FLH–FTL	PU–UOGD	PU–ATLAS	PU PRIOR ORACLE	PN PRIOR ORACLE
Linear shift									
MNIST	2.92 ±0.39	2.47 ±0.39	**2.00 ±0.35**	2.10 ±0.41	2.06 ±0.40	2.18 ±0.32	2.21 ±0.29	2.05 ±0.37	2.18 ±0.40
CIFAR-10	13.53 ±0.93	7.97 ±0.17	7.03 ±0.07	6.75 ±0.10	**6.39 ±0.24**	7.26 ±0.36	7.12 ±0.23	6.31 ±0.13	6.18 ±0.29
AG News	25.05 ±1.83	10.46 ±0.13	9.52 ±0.27	9.41 ±0.70	**8.82 ±0.47**	9.30 ±0.19	8.98 ±0.10	8.71 ±0.25	8.70 ±0.34
Bernoulli shift									
MNIST	3.28 ±0.58	2.35 ±0.21	2.18 ±0.24	**1.72 ±0.47**	2.02 ±0.42	2.42 ±0.26	1.95 ±0.11	1.09 ±0.17	1.07 ±0.18
CIFAR-10	15.47 ±1.87	6.43 ±0.43	7.00 ±0.38	**5.48 ±0.35**	5.88 ±0.50	7.22 ±0.53	5.78 ±0.30	2.58 ±0.19	2.34 ±0.09
AG News	28.94 ±3.70	8.74 ±0.16	9.33 ±0.29	**7.57 ±0.31**	8.06 ±0.44	9.23 ±1.07	7.62 ±0.53	3.65 ±0.31	3.61 ±0.11

At the beginning of round t, before observing X_t, we set $\widehat{\tau}_t := \widehat{\tau}^{\mathrm{PU}}(\widehat{\pi}_t)$. Combining the preceding threshold-selection guarantee with the prior-indexed decision reduction in Proposition 5 gives the following dynamic regret bound.

Theorem 4. *Let* $\overline{\epsilon}_R(\delta) := \max\left\{r_+(\delta), \frac{r_U(\delta)+\pi_0 r_+(\delta)}{1-\pi_0}\right\}$. *With probability at least* $1-\delta$ *over* S_0,

$$\mathrm{Reg}_T^{\mathrm{thr}} \le 2\sum_{t=1}^{T}|\widehat{\pi}_t - \pi_t| + 2\sum_{t=1}^{T}\epsilon_R(\widehat{\pi}_t,\delta) \le 2\sqrt{TL_T^{\mathrm{trk}}} + 2T\overline{\epsilon}_R(\delta). \tag{22}$$

The comparator is the best threshold of the fixed g at each prior. With the tracking bound in Theorem 1, this theorem yields a bound on $\mathbb{E}[\mathrm{Reg}_T^{\mathrm{thr}} \mid S_0]$ (See Corollary 2): the temporal error and the path-dependent finite-source error enter through prior tracking, while PU threshold-risk estimation contributes additional terms. If g is the exact source posterior, the closed-form threshold rule is Bayes-optimal at each q. Evaluating this rule at $\widehat{\pi}_t$ requires no PU threshold-risk estimation, and the resulting regret against the Bayes classifier satisfies $\mathrm{Reg}_T^{\mathrm{Bayes}} \le 2\sqrt{TL_T^{\mathrm{trk}}}$ (See Appendix D.4). Finite-source error can still enter through L_T^{trk}. The empirical PU rule in Eq. (20) requires no exact-posterior assumption and applies to any fixed score, including the learned score in Section 6.

6 EXPERIMENTS

We evaluate whether the PU reduction provides accurate and efficient online adaptation across image and text benchmarks, and compare the two online regression oracles within the same framework.

Setup. We construct five binary tasks: digit 0 versus the remaining digits on MNIST, T-shirt/top versus the remaining classes on Fashion-MNIST, airplane versus the remaining classes on CIFAR-10, AnnualCrop versus the remaining classes on EuroSAT, and Sci/Tech versus the remaining classes on AG News. MNIST and Fashion-MNIST use scaled pixels, CIFAR-10 and EuroSAT use fixed ImageNet ResNet-18 features, and AG News uses fixed DistilBERT features. In every task the source prior is $\pi_0 = 0.1$. Each online run has $T = 10^3$ rounds and $N_t = 10$ per round. We use the Linear, Sine, Square, and Bernoulli prior paths in $[0.05, 0.95]$ and report results over three seeds. Target labels are used only for evaluation. Table 1 shows three representative datasets and the Linear and Bernoulli paths; the complete five-dataset, four-path results are reported in Table 2.

PU–ALIGATOR and PU–FLH–FTL instantiate our method with the corresponding online regression oracle. We compare against the static PU initial classifier; PN versions of Baby et al. (2023), which use labeled negative offline data; and natural PU extensions of UOGD and ATLAS from Bai

et al. (2022). The latter revisit raw PU offline training examples and update the final layer during online adaptation. In contrast, our methods precompute the required source quantities, do not access raw offline examples during online adaptation, and keep the classification score g fixed while adapting only its threshold. The full data construction, hyperparameters, method definitions, prior paths, prior-tracking results, computation times, and prior-oracle definitions are deferred to Appendix E.

Results. Across the complete five-dataset, four-path results in Table 2, online adaptation improves substantially over the static initial classifier: every adaptive method reduces classification error on every dataset–path pair. In aggregate, PU–ALIGATOR and PU–FLH–FTL reduce error from 15.24% to 5.66% and 5.50%, respectively. More importantly, despite using no labeled negative offline data, their errors are only 0.82 and 0.75 percentage points above the corresponding PN variants (4.84% and 4.75%). They are also within 1.83 and 1.67 percentage points of the PU prior oracle (3.83%), which supplies the true current class prior while retaining the same PU initial model and threshold map. These comparisons show that the proposed threshold-only adaptation remains competitive even under weaker offline supervision and restricted access to offline data during online adaptation.

The optimization-based PU baselines provide a complementary tradeoff. PU–UOGD and PU–ATLAS obtain aggregate errors of 5.09% and 4.62%, respectively, with PU–ATLAS giving the lowest aggregate error among the deployable methods. These methods, however, repeatedly draw mini-batches from the stored PU offline training data and update the final layer during online adaptation, whereas the proposed methods use only precomputed source quantities and leave the offline-trained classification score fixed. The prior-tracking results should therefore be read as a mechanism diagnostic rather than as a ranking of classification performance: the two optimization baselines share the same one-round-lagged local estimate and hence the same prior MSE in every setting, yet their aggregate classification errors differ because they use different online model-update strategies. For the threshold-based methods, the estimated prior is mapped to a decision threshold through a discrete empirical threshold map, so a lower prior MSE does not necessarily lead to a lower classification error. As a path-specific tracking result, PU–FLH–FTL achieves the lowest average prior MSE among the PU methods on the Linear path, 1.79×10^{-2}, compared with 2.70×10^{-2} for the local estimator shared by PU–UOGD and PU–ATLAS. A controlled sensitivity experiment in Appendix E.2 further isolates the dependence on the offline class prior: under an identical target-prior path and fixed V_T, PU tracking becomes less stable as π_0 approaches one, whereas the PN controls remain unchanged, consistent with the path-dependent finite-source term in our bound.

The proposed methods are considerably cheaper. PU–ALIGATOR is the fastest online method at 0.110 seconds per trajectory, followed by PU–FLH–FTL at 0.132 seconds, compared with 0.279 seconds for PU–UOGD and 0.889 seconds for PU–ATLAS.

Thus, both proposed variants consistently outperform the static no-adaptation classifier, while the optimization-based PU baselines require approximately 2.1–8.1 times as much online runtime; they also avoid retaining and repeatedly accessing the raw PU offline training examples. PU–ATLAS offers the lowest aggregate error at the highest computational and data-access cost, while the proposed methods offer competitive accuracy through prior tracking and threshold adaptation alone.

7 CONCLUSION AND LIMITATION

We asked whether online adaptation to class-prior shift fundamentally requires labeled examples from both classes. Our results show that positive and unlabeled source data suffice to preserve the variation-dependent prior-tracking rate, while the cost of reconstructing the missing class depends on the entire target-prior path. This distinction also yields a dynamic classification-regret guarantee for threshold adaptation of a fixed score.

Limitations. Our guarantees assume class-prior shift and a known source prior π_0. Changes in the class-conditional distributions or estimation of π_0 require further identification and error analysis. We also assume a fixed statistic ϕ with a known lower bound $d_{\min} > 0$ on its class-mean separation. Selecting such a statistic from PU data is beyond our analysis. Our classification guarantee compares thresholds of a fixed score and does not cover online model updates or Bayes regret for a general learned score. Establishing lower bounds for classification regret remains open.

AI USE STATEMENT

In this work, we used Claude Opus 5.5 (primarily High, occasionally Extra) mainly to draft and edit manuscript text and improve its English, and GPT-6 Sol (Extra High) for research brainstorming. We used generative AI tools to provide feedback on research methodology and experiments, assist with specific proof tasks, help implement methods, suggest manuscript structure, and assist with translation. The authors have developed the theoretical framework and principal mathematical claims. The problem setting and research questions were all well-discussed by the authors. We asked for assistance with proofs, where we wrote the relevant statements, lemmas, and intended proof steps in the manuscript clearly, then asked narrowly scoped questions about conditions for supporting lemmas or possible errors in author-written proofs. The authors independently checked proposed conditions and corrections. We likewise completed the experimental pseudocode before seeking AI feedback and reviewed implementation suggestions against it. The authors decided which suggestions to adopt and revised and approved the final wording. We take responsibility for all claims, proofs, code, results, and text in this paper.

ETHICS STATEMENT

This paper is a theoretical foundation of statistical machine learning. The authors believe that this work does not cause any ethical issue.

REPRODUCIBILITY STATEMENT

The problem setting and evaluation criteria are clearly specified in Section 2, and the whole procedure of proposed method is detailed in Algorithm 1. Complete proofs of all theoretical results are provided in the appendix, including the online tracking analysis in Appendix C and the tracking-to-classification reduction in Appendix D. Appendix E introduces the dataset construction and preprocessing, model architectures, training and tracking hyperparameters, class-prior paths, evaluation metrics, and reports the complete experimental results. The code implementation of all experiments is provided as supplementary material.

REFERENCES

Anonymous authors. Online positive-unlabeled learning under continuous class-prior shift, 2026.

Kamyar Azizzadenesheli, Anqi Liu, Fanny Yang, and Animashree Anandkumar. Regularized learning for domain adaptation under label shifts. In *International Conference on Learning Representations (ICLR)*, 2019.

Dheeraj Baby and Yu-Xiang Wang. Online forecasting of total-variation-bounded sequences. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32, pp. 11071–11081, 2019.

Dheeraj Baby, Xuandong Zhao, and Yu-Xiang Wang. An optimal reduction of TV-denoising to adaptive online learning. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2899–2907, 2021.

Dheeraj Baby, Saurabh Garg, Tzu-Ching Yen, Sivaraman Balakrishnan, Zachary C. Lipton, and Yu-Xiang Wang. Online label shift: Optimal dynamic regret meets practical algorithms. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Yong Bai, Yu-Jie Zhang, Peng Zhao, Masashi Sugiyama, and Zhi-Hua Zhou. Adapting to online label shift with provable guarantees. In *Advances in Neural Information Processing Systems*, 2022.

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: a survey. *Machine Learning*, 109(4):719–760, 2020.

Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations Research*, 63(5):1227–1244, 2015.

Gilles Blanchard, Gyemin Lee, and Clayton Scott. Semi-supervised novelty detection. *Journal of Machine Learning Research*, 11:2973–3009, 2010.

Nontawat Charoenphakdee and Masashi Sugiyama. Positive-unlabeled classification under class prior shift and asymmetric error. In *Proceedings of the SIAM International Conference on Data Mining (SDM)*, pp. 271–279, 2019.

Marthinus C. du Plessis, Gang Niu, and Masashi Sugiyama. Analysis of learning from positive and unlabeled data. In *Advances in Neural Information Processing Systems*, pp. 703–711. Curran Associates, Inc., 2014.

Marthinus Christoffel du Plessis, Gang Niu, and Masashi Sugiyama. Class-prior estimation for learning from positive and unlabeled data. *Machine Learning*, 106(4):463–492, 2017.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 213–220, 2008.

Saurabh Garg, Yifan Wu, Sivaraman Balakrishnan, and Zachary Lipton. A unified view of label shift estimation. In *Advances in Neural Information Processing Systems*, volume 33, pp. 3290–3300. Curran Associates, Inc., 2020.

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

Ryuichi Kiryo, Gang Niu, Marthinus C. du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 1674–1684, 2017.

Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Lanas Phillips, Irena Gao, Tony Lee, Etienne David, Ian Stavness, Wei Guo, Berton Earnshaw, Imran Haque, Sara M Beery, Jure Leskovec, Anshul Kundaje, Emma Pierson, Sergey Levine, Chelsea Finn, and Percy Liang. Wilds: A benchmark of in-the-wild distribution shifts. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 5637–5664, 2021.

Atsutoshi Kumagai, Tomoharu Iwata, Hiroshi Takahashi, Taishi Nishiyama, and Yasuhiro Fujiwara. Importance-weighted positive-unlabeled learning for distribution shift adaptation. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, pp. 1576–1584, 2025.

Fuyi Li, Shuangyu Dong, André Leier, Meiya Han, Xudong Guo, Jing Xu, Xiaoyu Wang, Shirui Pan, Cangzhi Jia, Yang Zhang, et al. Positive-unlabeled learning in bioinformatics and computational biology: a brief review. *Briefings in bioinformatics*, 23(1):bbab461, 2022.

Zachary C. Lipton, Yu-Xiang Wang, and Alexander J. Smola. Detecting and correcting for label shift with black box predictors. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, volume 80 of *PMLR*, pp. 3122–3130, 2018.

Pascal Massart. The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *The Annals of Probability*, 18(3):1269–1283, 1990. doi: 10.1214/aop/1176990746.

Jan Mielniczuk, Wojciech Rejchel, and Paweł Teisseyre. Prior shift estimation for positive unlabeled data through the lens of kernel embedding. In *The 29th International Conference on Artificial Intelligence and Statistics*, 2026.

Shota Nakajima and Masashi Sugiyama. Positive-unlabeled classification under class-prior shift: a prior-invariant approach based on density ratio estimation. *Machine Learning*, 112(3):889–919, 2023.

Joaquin Quiñonero-Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D Lawrence. *Dataset shift in machine learning*. Mit Press, 2008.

Harish G. Ramaswamy, Clayton Scott, and Ambuj Tewari. Mixture proportion estimation via kernel embeddings of distributions. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, volume 48 of *PMLR*, pp. 2052–2060, 2016.

Marco Saerens, Patrice Latinne, and Christine Decaestecker. Adjusting the outputs of a classifier to new a priori probabilities: A simple procedure. *Neural Computation*, 14(1):21–41, 2002.

Clayton Scott. A rate of convergence for mixture proportion estimation, with application to learning from noisy labels. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 38 of *PMLR*, pp. 838–846, 2015.

Hyebin Song, Bennett J Bremer, Emily C Hinds, Garvesh Raskutti, and Philip A Romero. Inferring protein sequence-function relationships with large-scale positive-unlabeled learning. *Cell systems*, 12(1):92–101, 2021.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, 2009.

Ruihan Wu, Chuan Guo, Yi Su, and Kilian Weinberger. Online adaptation to label distribution shift. In *Advances in Neural Information Processing Systems*, volume 34, pp. 11340–11351, 2021.

Yu-Jie Zhang, Zhen-Yu Zhang, Peng Zhao, and Masashi Sugiyama. Adapting to continuous covariate shift via online density ratio estimation. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

A Additional related work

Online label shift studies changing label marginals under fixed class-conditional distributions. Wu et al. (2021) initiated online adaptation to drifting label distributions; Bai et al. (2022); Baby et al. (2023) obtained optimal dynamic regret via reductions to online oracles under a known or population operator (Lipton et al., 2018; Saerens et al., 2002). Finite-sample confusion-operator error was analyzed offline (Azizzadenesheli et al., 2019; Garg et al., 2020). Bai et al. (2022) bound the resulting source-estimation error but treat it as negligible when the offline sample is sufficiently large, while the main regression analysis of Baby et al. (2023) uses population source quantities and discusses empirical replacement in a regime relating the horizon to the offline sample size. Our work focuses on the theoretical properties to explain how the weak supervision in offline data affects online label shift. Our novel finding is that when the same PU source estimates are reused at every round, their error accumulates with weights determined by the target-prior path; we characterize these weights and prove a corresponding lower bound. These theoretical characterizations were overlooked in the literature.

Class-prior/mixture-proportion estimation from PU data is classical (Elkan & Noto, 2008; du Plessis et al., 2017; Scott, 2015; Ramaswamy et al., 2016; Bekker & Davis, 2020), with identifiability under irreducibility (Blanchard et al., 2010). Since the source prior π_0 is known in our setting, our problem is different from source-prior identification. PU classification under class-prior shift has been studied in the batch setting (Charoenphakdee & Sugiyama, 2019; Nakajima & Sugiyama, 2023), and prior-shift estimation from PU data has recently been approached through kernel mean embeddings (Mielniczuk et al., 2026). Kumagai et al. (2025) address more general distribution shifts using PU samples from both the training and test distributions and iteratively retrain the classifier. That line of work estimates a changed prior in a single target environment, while we consider multiple online rounds where the prior changes at every round and the source quantities themselves are estimated once from finite PU data. Concurrent online PU work receives labeled positive mini-batches during deployment and updates the classifier head (Anonymous authors, 2026). Our online stage receives only unlabeled target observations, summarizes the PU source sample once, and adapts a fixed score through its prior estimate and decision threshold.

B Proofs for the statistic–prior map

This appendix collects the material deferred from Section 3: the geometry of the map, the reconstructed negative-mean error, the concentration and stability of the estimated map, the proof of Proposition 1, and the complete comparison with PN source supervision.

Geometry of the map. The map satisfies $M(0) = \mu_-$, $M(\pi_0) = \mu_0$, and $M(1) = \mu_+$. In Eq. (2) the two coefficients sum to one and are both nonnegative exactly when $\pi \in [\pi_0, 1]$, so for a target prior between the source prior and one, $M(\pi)$ interpolates between the two observed source means. For $\pi < \pi_0$ the coefficient of μ_+ is negative and the same affine formula extrapolates beyond the observed prior range, toward the reconstructed endpoint μ_-. The weights carried by the two source-estimation errors therefore depend on where the target prior lies relative to π_0.

Cancellation at $\pi = \pi_0$. Evaluating Eq. (6) at $\pi = \pi_0$ gives $\widehat{M}(\pi_0) - M(\pi_0) = \Delta_0$, so the direct Δ_+ term in the map error vanishes there. This is a statement about the map error alone. The prior-tracking error of Section 4 is formed with the estimated slope $\widehat{d}$, which is random and depends on $\widehat{\mu}_+$, so the dependence of the tracking error on n_+ does not disappear at $\pi = \pi_0$.

B.1 Concentration and stability of the PU map

Only the two direct deviations Δ_+ and Δ_0 appear in Eq. (6), but the reconstructed negative mean has a deviation of its own, which the concentration argument below uses. Evaluating that identity at $\pi = 0$ gives

$$\Delta_-^{\mathrm{PU}} := \widehat{\mu}_- - \mu_- = \frac{\Delta_0 - \pi_0 \Delta_+}{1 - \pi_0}, \tag{23}$$

so Δ_-^{PU} is not a separate sampling error: it is determined by the two direct estimation errors, and it carries the factor $(1-\pi_0)^{-1}$ of the reconstruction. Under PN supervision the negative mean is estimated directly from an independent negative sample, with deviation $\Delta_-^{\mathrm{PN}} := \widehat{\mu}_-^{\mathrm{PN}} - \mu_-$; this is used in Remark 1 below.

The next proposition proves the two bounds stated in Eq. (9) and adds a worst-case bound on the radius Eq. (8).

Proposition 3 (Concentration and stability of $\widehat{M}$)**.** *Let* $\mathrm{rad}_-^{\mathrm{PU}}(\delta) := \{\mathrm{rad}_0(\delta) + \pi_0\,\mathrm{rad}_+(\delta)\}/(1-\pi_0)$*, so that* $\mathrm{rad}_{\mathrm{PU}}(\delta) = \max\{\mathrm{rad}_+(\delta), \mathrm{rad}_-^{\mathrm{PU}}(\delta)\}$*. The event* $\mathcal{E}_{\mathrm{src}}(\delta)$ *on which* $|\Delta_+| \le \mathrm{rad}_+(\delta)$ *and* $|\Delta_0| \le \mathrm{rad}_0(\delta)$ *has probability at least* $1-\delta$ *over the holdout data* S_0*, and on it* $|\Delta_-^{\mathrm{PU}}| \le \mathrm{rad}_-^{\mathrm{PU}}(\delta)$*,* $\varepsilon_{\mathrm{map}} \le \mathrm{rad}_{\mathrm{PU}}(\delta)$*, and* $|\widehat{d} - d| \le 2\,\mathrm{rad}_{\mathrm{PU}}(\delta)$*. Moreover,*

$$\mathrm{rad}_{\mathrm{PU}}(\delta)^2 \le \frac{\log(4/\delta)}{(1-\pi_0)^2}\left(\frac{1}{n_+} + \frac{1}{n_u}\right). \tag{24}$$

Proof of Proposition 3. Recall that $\Delta_+ = \widehat{\mu}_+ - \mu_+$ and $\Delta_0 = \widehat{\mu}_0 - \mu_0$. Since ϕ is fixed before S_0 is observed, the values $\phi(X)$ for $X \in S_+$ and $X \in S_{\mathrm{U},0}$ are iid given ϕ and lie in $[0,1]$, so Hoeffding's inequality and a union bound give the two stated deviation bounds. On that event, Eq. (23) and the triangle inequality give $|\Delta_-^{\mathrm{PU}}| \le (\mathrm{rad}_0(\delta) + \pi_0\,\mathrm{rad}_+(\delta))/(1-\pi_0) = \mathrm{rad}_-^{\mathrm{PU}}(\delta)$. Since $\widehat{M} - M$ is affine on $[0,1]$, its largest absolute value is attained at $\pi = 0$ or $\pi = 1$, so $\varepsilon_{\mathrm{map}} = \max\{|\Delta_-^{\mathrm{PU}}|, |\Delta_+|\} \le \mathrm{rad}_{\mathrm{PU}}(\delta)$, and the bound on $\widehat{d} - d$ follows by subtracting the errors at $\pi = 0$ and $\pi = 1$. For Eq. (24), apply $\max\{a,b\}^2 \le a^2 + b^2$ and $(a+b)^2 \le 2a^2 + 2b^2$ to obtain

$$\mathrm{rad}_{\mathrm{PU}}(\delta)^2 \le \log(4/\delta)\left[\frac{1}{n_u(1-\pi_0)^2} + \frac{1}{n_+}\left(\frac{1}{2} + \frac{\pi_0^2}{(1-\pi_0)^2}\right)\right],$$

and note that $\frac{1}{2}(1-\pi_0)^2 + \pi_0^2 \le 1$ for every $\pi_0 \in [0,1]$, so the bracketed coefficient of $1/n_+$ is at most $(1-\pi_0)^{-2}$. □

Proof of Proposition 1. Subtracting Eq. (2) from Eq. (3) gives Eq. (6). On $\mathcal{E}_{\mathrm{src}}(\delta)$, that identity and the triangle inequality give

$$|\widehat{M}(\pi_t) - M(\pi_t)| \le \frac{(1-\pi_t)\,\mathrm{rad}_0(\delta) + |\pi_t - \pi_0|\,\mathrm{rad}_+(\delta)}{1-\pi_0}.$$

Applying $(a+b)^2 \le 2a^2 + 2b^2$, then we sum over t. Substituting $\mathrm{rad}_0(\delta)^2 = \log(4/\delta)/(2n_u)$ and $\mathrm{rad}_+(\delta)^2 = \log(4/\delta)/(2n_+)$ concludes the proof. □

B.2 Comparison with PN source supervision

Remark 1. The same path dependence also appears under PN supervision. The difference is that PU supervision changes the effective sample size for estimating the negative mean and shifts the cancellation point of the positive-sample error from $\pi = 0$ to $\pi = \pi_0$. Suppose instead that a labeled negative sample of size n_- is available. Then μ_- has its own sample mean $\widehat{\mu}_-^{\mathrm{PN}}$, computed directly from that sample rather than reconstructed, whose error $\Delta_-^{\mathrm{PN}} := \widehat{\mu}_-^{\mathrm{PN}} - \mu_-$ satisfies $|\Delta_-^{\mathrm{PN}}| \le \sqrt{\log(4/\delta)/(2n_-)}$ with probability at least $1-\delta/2$. Here the PN map is built from the two endpoint means, so the endpoint form Eq. (1), rather than the representation in Eq. (2), is the natural plug-in target, and

$$\widehat{M}^{\mathrm{PN}}(\pi) - M(\pi) = (1-\pi)\Delta_-^{\mathrm{PN}} + \pi\,\Delta_+.$$

Repeating the argument of Proposition 1 gives, with probability at least $1-\delta$ over the PN source sample,

$$\sum_{t=1}^{T}\left|\widehat{M}^{\mathrm{PN}}(\pi_t) - M(\pi_t)\right|^2 \le \log(4/\delta)\left\{\frac{W_U(\pi_{1:T})}{n_-} + \frac{\sum_{t=1}^{T}\pi_t^2}{n_+}\right\}.$$

Thus, in both PU and PN supervision, the negative-mean error is weighted by $W_U(\pi_{1:T})$. The negative-mean term scales as $1/n_-$ under PN supervision and as $1/\{(1-\pi_0)^2 n_u\}$ under PU supervision. The positive-sample error, however, is weighted by $\sum_t \pi_t^2$ under PN and by $W_+(\pi_{1:T})/(1-$

$\pi_0)^2$ under PU. Hence, it vanishes at $\pi = 0$ under PN but at $\pi = \pi_0$ under PU. In particular, $\pi_0 = 0$ reduces the PU setting to PN supervision. The exchange rate $n_u \asymp n_-/(1-\pi_0)^2$ therefore equates only the negative-side terms of the two bounds. It does not make the PU and PN bounds agree, because the positive-sample terms are weighted by $\sum_t \pi_t^2$ under PN and by $W_+(\pi_{1:T})/(1-\pi_0)^2$ under PU, and these differ for any path that does not stay at π_0. The comparison is also at a fixed score and fixed positive sample; on its own it does not equate a given n_u with the same n_- in annotation or sampling cost.

C Proofs for online prior tracking

C.1 The online regression result used in the upper bound

We use the following one-dimensional consequence of ALIGATOR. It is the specialization of the online-regression guarantee recorded in Baby et al. (2023, Proposition 20), which restates the guarantee of Baby et al. (2021). We state it here to make the dependence on the estimated separation explicit.

Proposition 4. *Fix holdout data S_0 such that*

$$\|\widehat{M} - M\|_{\infty,[0,1]} \le d_{\min}/4.$$

Set $a \coloneqq d_{\min}/2$ and

$$\widetilde{Z}_t \coloneqq aZ_t, \qquad \widetilde{\pi}_t^\dagger \coloneqq a\pi_t^\dagger.$$

Let ALIGATOR *use $\widetilde{Z}_1, \dots, \widetilde{Z}_{t-1}$ to produce $\widetilde{x}_t$ at round t, and set $x_t \coloneqq \widetilde{x}_t/a$.* ALIGATOR*'s learning rate is the one prescribed by Baby et al. (2021, Theorem 5),*

$$\eta = \frac{1}{8\big(B + \sigma\sqrt{\log(2T/\delta)}\big)^2},$$

which we instantiate with $B = 1$ and $\sigma = 1/2$, the bounds verified in the proof below, and $\delta = (eT)^{-2}$. There are universal constants $C, c > 0$ such that

$$\mathbb{E}\left[\sum_{t=1}^{T}(x_t - \pi_t^\dagger)^2 \,\middle|\, S_0\right] \le C\log^c(eT)\left\{\frac{1}{d_{\min}^2} + \frac{T^{1/3}V_T^{2/3}}{d_{\min}^{4/3}}\right\}. \tag{25}$$

The algorithm does not require V_T as an input.

Proof. The assumed event and Proposition 2 give $\widehat{d} \ge d_{\min}/2$ and $V_T(\pi^\dagger) \le 2V_T$. Moreover,

$$|\pi_t^\dagger - \pi_t| = \frac{|M(\pi_t) - \widehat{M}(\pi_t)|}{\widehat{d}} \le \frac{1}{2},$$

so $|\pi_t^\dagger| \le 3/2$. Conditional on S_0,

$$\widetilde{Z}_t = \widetilde{\pi}_t^\dagger + \widetilde{\xi}_t, \qquad \widetilde{\xi}_t \coloneqq a\frac{\phi(X_t) - M(\pi_t)}{\widehat{d}}.$$

The variables $\widetilde{\xi}_t$ are independent and mean zero. Since $\phi \in [0,1]$,

$$|\widetilde{\pi}_t^\dagger| \le \frac{3}{4}, \qquad |\widetilde{\xi}_t| \le \frac{a}{\widehat{d}} \le 1, \qquad \mathrm{Var}(\widetilde{\xi}_t \mid S_0) \le \frac{a^2}{4\widehat{d}^2} \le \frac{1}{4}.$$

Moreover,

$$V_T(\widetilde{\pi}^\dagger) = aV_T(\pi^\dagger) \le d_{\min}V_T.$$

These are exactly the bounded-signal, bounded-noise, zero-mean, and variance conditions required by Baby et al. (2023, Proposition 20), with $B = 1$ and $\sigma^2 = 1/4$. This proposition restates Baby et al. (2021, Theorem 5), where the learning rate above is prescribed. Conditionally on S_0, Baby et al. (2023, Proposition 20) gives the stated rate with probability at least $1 - \delta$.

To pass to an expectation, note that ALIGATOR predicts a convex combination of online averages of the past observations, started from 0, so $|\widetilde{x}_t| \le 7/4$. Together with $|\widetilde{\pi}_t^\dagger| \le 3/4$, we have

$$\sum_{t=1}^{T} (\widetilde{x}_t - \widetilde{\pi}_t^\dagger)^2 \le \frac{25T}{4}$$

on every sample path. With $\delta = (eT)^{-2}$ the complementary event therefore contributes at most $(25T/4)\delta = O(1/T)$ to the conditional expectation, and $\log(2T/\delta) = O(\log(eT))$ keeps the learning rate inside the same poly-logarithmic factor. Hence, conditionally on S_0,

$$\mathbb{E}\left[\sum_{t=1}^{T} (\widetilde{x}_t - \widetilde{\pi}_t^\dagger)^2 \,\middle|\, S_0\right] \le C \log^c(eT) \left\{1 + T^{1/3} (d_{\min} V_T)^{2/3}\right\}.$$

Since $x_t = \widetilde{x}_t / a$ and $a^{-2} = 4 d_{\min}^{-2}$, rescaling this inequality gives Eq. (25) after changing the universal constant. The same proposition states that the variation is not required by the algorithm. □

C.2 BATCHED ROUNDS

Algorithm 1 is run with a batch of $N_t \ge 1$ observations per round, whereas the analysis is stated for $N_t = 1$. Fix the batch sizes $N_1, \dots, N_T \ge 1$ in advance, so that they do not depend on the data, and keep the standing assumptions that the prior path is deterministic and that all source and target observations are mutually independent. With $\bar{\phi}_t := N_t^{-1} \sum_{i=1}^{N_t} \phi(X_{t,i})$ the rescaled noise is

$$\widetilde{\xi}_t = \frac{a}{\widehat{d}} \big(\bar{\phi}_t - M(\pi_t)\big).$$

Conditionally on S_0, the variables $\widetilde{\xi}_1, \dots, \widetilde{\xi}_T$ are independent across rounds, because the batches are drawn independently, and each has mean zero because $\mathbb{E}[\phi(X_{t,i})] = M(\pi_t)$ for every i. Since $\phi \in [0,1]$ and $a/\widehat{d} \le 1$ on the event of Proposition 4,

$$|\widetilde{\xi}_t| \le 1, \qquad \mathrm{Var}(\widetilde{\xi}_t \mid S_0) \le \frac{1}{4N_t} \le \frac{1}{4},$$

while the signal $\widetilde{\pi}_t^\dagger$ does not involve N_t at all. These are the conditions used in the proof of Proposition 4, so that proposition, and with it Theorem 1, applies verbatim for every fixed $N_t \ge 1$. The bound does not improve with N_t, since the variance reduction $1/(4N_t)$ is discarded as soon as it is bounded by $1/4$. We claim nothing about batch sizes chosen adaptively from the data, and nothing about the temporal lower bound of Theorem 3 under general batches.

C.3 PROOF OF THE REDUCTION TO THE CONDITIONAL-MEAN SEQUENCE

Proof of Proposition 2. Eq. (12) follows by substituting $M(\pi_t) = \mu_- + d\pi_t$ and $\widehat{M}(\pi_t) = \widehat{\mu}_- + \widehat{d}\pi_t$. Projection cannot increase the distance to $\pi_t \in [0,1]$. Applying $(a+b)^2 \le 2a^2 + 2b^2$ to $x_t - \pi_t = (x_t - \pi_t^\dagger) + (\pi_t^\dagger - \pi_t)$ and summing proves Eq. (13). Finally, $\pi_t^\dagger$ is affine in π_t with slope $d/\widehat{d}$. The assumed uniform map bound gives $|\widehat{d} - d| \le 2r \le d/2$, and hence $\widehat{d} \ge d_{\min}/2$ and $d/\widehat{d} \le 2$. □

C.4 PROOF OF THE PATH-DEPENDENT PU TRACKING BOUND

Proof of Theorem 1. Condition on holdout data S_0 in the event $\mathcal{E}_{\mathrm{src}}(\delta_{\mathrm{off}})$ of Proposition 3. Then $\widehat{d} \ge d_{\min}/2$ and Proposition 2 gives $V_T(\pi^\dagger) \le 2V_T$. Conditional on S_0, $Z_t = \pi_t^\dagger + \xi_t$, where the ξ_t are independent, mean zero, and bounded in absolute value by $1/\widehat{d}$. The normalized Aligator construction in Proposition 4 therefore gives

$$\mathbb{E}\left[\sum_{t=1}^{T} (x_t - \pi_t^\dagger)^2 \,\middle|\, S_0\right] \le C \log^c(eT) \left\{\frac{1}{d_{\min}^2} + \frac{T^{1/3} V_T^{2/3}}{d_{\min}^{4/3}}\right\}.$$

Take the conditional expectation in Eq. (13). Since $2/\widehat{d}^2 \le 8/d_{\min}^2$, the remaining term is bounded by Eq. (7). This proves Eq. (14) on a holdout-data event of probability at least $1 - \delta_{\mathrm{off}}$. □

C.5 WORST-CASE FORM OF THE TRACKING BOUND

We record the worst-case consequence of Theorem 1, obtained by bounding both path weights by T.

Since $W_U(\pi_{1:T}) \le T$ and $W_+(\pi_{1:T}) \le T$, the same source event gives the worst-case consequence

$$\mathbb{E}\big[L_T^{\mathrm{trk}} \mid S_0\big] \le C \log^c(eT) \left(\frac{1}{d_{\min}^2} + \frac{T^{1/3} V_T^{2/3}}{d_{\min}^{4/3}} \right) + \frac{8T \log(4/\delta_{\mathrm{off}})}{d_{\min}^2 (1-\pi_0)^2} \left(\frac{1}{n_u} + \frac{1}{n_+} \right). \tag{26}$$

Our main upper bound is the more precise path-dependent result in Eq. (14).

C.6 TWO ELEMENTARY LOWER-BOUND TOOLS

For $p, q \in (0,1)$, let $\mathrm{kl}(p \,\|\, q) := p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$.

Lemma 1. *If $p, q \in [1/8, 7/8]$, then*

$$\mathrm{kl}(p \,\|\, q) \le 10(p-q)^2.$$

Proof. The standard chi-square upper bound gives

$$\mathrm{kl}(p \,\|\, q) \le \frac{(p-q)^2}{q(1-q)} \le \frac{64}{7}(p-q)^2 \le 10(p-q)^2.$$

□

Next, we state the vector-valued Le Cam inequality.

Lemma 2. *Let Q_0, Q_1 be probability laws and let $\theta_0, \theta_1 \in \mathbb{R}^T$. Every possibly randomized estimator $\widehat{\theta} \in \mathbb{R}^T$ satisfies*

$$\max_{j \in \{0,1\}} \mathbb{E}_j \|\widehat{\theta} - \theta_j\|_2^2 \ge \frac{\|\theta_1 - \theta_0\|_2^2}{8} \left(1 - \|Q_0 - Q_1\|_{\mathrm{TV}}\right).$$

In particular, if $D_{\mathrm{KL}}(Q_0 \,\|\, Q_1) \le 1/8$, the right-hand side is at least $3\|\theta_1 - \theta_0\|_2^2/32$.

Proof. Let $\Delta = \|\theta_1 - \theta_0\|_2$ and choose model 1 when $\widehat{\theta}$ is at least as close to θ_1 as to θ_0. Whenever this test is wrong, the estimation error under the true model is at least $\Delta/2$. The sum of its two error probabilities is at least $1 - \|Q_0 - Q_1\|_{\mathrm{TV}}$. Summing the two risks and then using that their maximum is at least half their sum gives the first claim. Pinsker's inequality gives $\|Q_0 - Q_1\|_{\mathrm{TV}} \le \sqrt{D_{\mathrm{KL}}(Q_0 \,\|\, Q_1)/2} \le 1/4$ for the second claim; see Tsybakov (2009, Lemma 2.5(i)). □

C.7 PROOF OF THE PATH-SPECIFIC PU SOURCE LOWER BOUND

Proof of Theorem 2. Set

$$q_+^{(0)} := \frac{1}{2} + \frac{d}{2}, \qquad q_-^{(0)} := \frac{1}{2} - \frac{d}{2}.$$

Under Q_0, let $P_+^{(0)} = \mathrm{Bernoulli}(q_+^{(0)})$, $P_-^{(0)} = \mathrm{Bernoulli}(q_-^{(0)})$, and use the prescribed path $\pi_{1:T}$. We take $c_0 := 1/16$ and construct one alternative for each source sample.

Let

$$h_u := \min\left\{ \frac{d}{2}, \frac{c_0}{(1-\pi_0)\sqrt{n_u}} \right\}.$$

The model Q_U keeps $q_+^{(U)} = q_+^{(0)}$ and sets

$$q_-^{(U)} := q_-^{(0)} - h_u, \qquad \pi_t^{(U)} := \frac{d\pi_t + h_u}{d + h_u}.$$

Its separation is $d + h_u \in [d, 3d/2]$. Since $d \le 1/8$ and $h_u \le d/2$, its class-conditional Bernoulli parameters belong to $[3/8, 5/8] \subset [1/8, 7/8]$. For every t,

$$q_-^{(U)} + (d + h_u)\pi_t^{(U)} = q_-^{(0)} + d\pi_t.$$

Thus all target-unlabeled laws are identical under Q_0 and Q_U. The positive-source law is also unchanged. Only the source-unlabeled Bernoulli mean changes, by $(1 - \pi_0)h_u$. Since both source-unlabeled means lie in $[1/8, 7/8]$, Lemma 1 gives

$$D_{\mathrm{KL}}(Q_0 \,\|\, Q_U) \le 10 n_u (1 - \pi_0)^2 h_u^2 \le 10 c_0^2 < \frac{1}{8}.$$

Moreover,

$$\pi_t^{(U)} - \pi_t = \frac{h_u(1 - \pi_t)}{d + h_u},$$

and hence

$$\|\pi^{(U)} - \pi\|_2^2 = \frac{h_u^2}{(d + h_u)^2} W_U(\pi_{1:T}) \ge \frac{4h_u^2}{9d^2} W_U(\pi_{1:T}).$$

By Lemma 2, the maximum risk over Q_0, Q_U is at least

$$c_U W_U(\pi_{1:T}) \min\left\{1, \frac{1}{d^2(1 - \pi_0)^2 n_u}\right\}$$

for a universal constant $c_U > 0$.

Let

$$h_+ \coloneqq \min\left\{\frac{d}{2}, \frac{c_0}{(1 - \pi_0)\sqrt{n_+}}\right\}.$$

The model Q_+ uses

$$q_+^{(+)} \coloneqq q_+^{(0)} + (1 - \pi_0)h_+, \qquad q_-^{(+)} \coloneqq q_-^{(0)} - \pi_0 h_+, \qquad \pi_t^{(+)} \coloneqq \frac{d\pi_t + \pi_0 h_+}{d + h_+}.$$

Again, the separation is in $[d, 3d/2]$, and both class-conditional parameters lie in $[3/8, 5/8]$. The source-unlabeled law is unchanged because

$$\pi_0 q_+^{(+)} + (1 - \pi_0) q_-^{(+)} = \pi_0 q_+^{(0)} + (1 - \pi_0) q_-^{(0)}.$$

The target-unlabeled law at every round is also unchanged:

$$q_-^{(+)} + (d + h_+)\pi_t^{(+)} = q_-^{(0)} + d\pi_t.$$

Only the positive-source mean changes, by $(1 - \pi_0)h_+$. Therefore

$$D_{\mathrm{KL}}(Q_0 \,\|\, Q_+) \le 10 n_+ (1 - \pi_0)^2 h_+^2 \le 10 c_0^2 < \frac{1}{8}.$$

The path separation is

$$\pi_t^{(+)} - \pi_t = \frac{h_+(\pi_0 - \pi_t)}{d + h_+},$$

and another application of Lemma 2 gives a lower bound

$$c_+ W_+(\pi_{1:T}) \min\left\{1, \frac{1}{d^2(1 - \pi_0)^2 n_+}\right\}$$

on the maximum risk over Q_0, Q_+.

For every estimator, its maximum risk over the three models is at least the maximum of the two pairwise lower bounds, and hence at least half their sum. This proves Eq. (15) after decreasing the universal constant.

Finally,

$$\pi_t^{(U)} = \frac{d}{d + h_u}\pi_t + \frac{h_u}{d + h_u}, \qquad \pi_t^{(+)} = \frac{d}{d + h_+}\pi_t + \frac{h_+}{d + h_+}\pi_0.$$

Both paths are affine contractions of a path in $[0, 1]^T$ toward a point in $[0, 1]$. They therefore belong to $[0, 1]^T$, and

$$V_T(\pi^{(U)}) = \frac{d}{d + h_u} V_T(\pi), \qquad V_T(\pi^{(+)}) = \frac{d}{d + h_+} V_T(\pi).$$

Thus neither perturbation increases total variation. □

C.8 Proof of the oracle variation lower bound

Proof of Theorem 3. For a path $\pi_{1:T}$, the target observations are independent with

$$X_t \sim \text{Bernoulli}\left(\frac{1}{2} + d\left(\pi_t - \frac{1}{2}\right)\right).$$

It is enough to lower-bound batch estimators, as in the statement.

Let

$$a_0 \coloneqq \min\left\{\frac{1}{8}, \frac{c_0}{d\sqrt{T}}\right\},$$

where $c_0 > 0$ is a sufficiently small universal constant, and consider $\pi_t^\varepsilon = 1/2 + \varepsilon a_0$ for $\varepsilon \in \{-1, 1\}$. Both paths belong to $\Pi_T(V)$ for every $V \geq 0$. Their observation laws, denoted by Q_- and Q_+, are products of Bernoulli laws with means $1/2 - da_0$ and $1/2 + da_0$. By Lemma 1,

$$D_{\mathrm{KL}}(Q_- \,\|\, Q_+) \leq 40Td^2a_0^2 \leq 40c_0^2.$$

Choose c_0 so that this is at most $1/8$. The two parameter vectors have squared Euclidean separation $4Ta_0^2$. Hence Lemma 2 gives

$$R_{\mathrm{bat}}(T, d, V) \geq cTa_0^2 \geq c\min\{T, d^{-2}\}, \tag{27}$$

where R_{bat} denotes the batch minimax risk in Theorem 3.

Fix $K \in \{1, \ldots, T\}$ and partition $\{1, \ldots, T\}$ into K consecutive nonempty blocks $B_1, \ldots, B_K$ whose sizes differ by at most one. Put $m = \lfloor T/K \rfloor$. Then

$$m \leq |B_j| \leq \frac{2T}{K}, \qquad mK \geq \frac{T}{2}.$$

For $\theta \in \{-1, 1\}^K$, define

$$\pi_t^\theta = \frac{1}{2} + a\theta_j \quad (t \in B_j), \qquad a \coloneqq \min\left\{\frac{1}{8}, \frac{V}{2K}, c_1\sqrt{\frac{K}{Td^2}}\right\},$$

where $c_1 > 0$ is another sufficiently small universal constant. Each path belongs to $[3/8, 5/8]^T$, and its total variation is at most $2a(K-1) \leq V$. Thus all these paths belong to $\Pi_T(V)$.

Let Q_θ be the corresponding observation law, and let $\theta^{(j)}$ be obtained by flipping coordinate j. These two laws differ only on B_j. Their Bernoulli means on that block differ by $2da$. Thus Lemma 1 gives

$$D_{\mathrm{KL}}(Q_\theta \,\|\, Q_{\theta^{(j)}}) \leq 40|B_j|d^2a^2 \leq 80c_1^2.$$

Choose c_1 so that the last quantity is at most $1/8$. Pinsker's inequality then gives

$$\max_{\theta, j} \|Q_\theta - Q_{\theta^{(j)}}\|_{\mathrm{TV}} \leq \frac{1}{4}. \tag{28}$$

Given an estimator $\widetilde{\pi}_{1:T}$, estimate θ_j by the sign of the average of $\widetilde{\pi}_t - 1/2$ over B_j. If this sign is wrong, Jensen's inequality gives

$$\sum_{t \in B_j} (\widetilde{\pi}_t - \pi_t^\theta)^2 \geq ma^2.$$

The total-variation form of Assouad's lemma (Tsybakov, 2009, Theorem 2.12(ii)), together with Eq. (28), gives

$$\inf_{\widehat{\theta}} \sup_{\theta \in \{-1,1\}^K} \mathbb{E}_\theta \sum_{j=1}^K \mathbf{1}\{\widehat{\theta}_j \neq \theta_j\} \geq \frac{K}{2}\left(1 - \frac{1}{4}\right) = \frac{3K}{8}.$$

Using $mK \geq T/2$, we therefore obtain

$$R_{\mathrm{bat}}(T, d, V) \geq cTa^2 \geq cT\min\left\{1, \frac{V^2}{K^2}, \frac{K}{Td^2}\right\} \quad \text{for every } K \in \{1, \ldots, T\}. \tag{29}$$

Set

$$A \coloneqq d^{-2}, \qquad B \coloneqq T^{1/3}V^{2/3}d^{-4/3}, \qquad z \coloneqq (Td^2V^2)^{1/3} = d^2B.$$

If $B \le A$, Eq. (27) already implies $R_{\mathrm{bat}}(T,d,V) \ge c\min\{T, A+B\}$, after changing the universal constant. Suppose now that $B > A$, so $z > 1$.

If $B \le T$, take $K = \lceil z \rceil$. As $d \le 1/8$, we have $z = d^2B \le T/64$ and $z \le K \le 2z$. Hence we have that $\frac{K}{Td^2} \ge \frac{B}{T}$, $\frac{V^2}{K^2} \ge \frac{B}{4T}$, and $1 \ge \frac{B}{T}$.

Eq. (29) gives $R_{\mathrm{bat}}(T,d,V) \ge cB$. Combining this with Eq. (27) and using $\max\{x,y\} \ge (x+y)/2$ yields $R_{\mathrm{bat}}(T,d,V) \ge c\min\{T, A+B\}$.

It remains to consider $B > T$. If $A \ge T$, the constant-path bound gives $R_{\mathrm{bat}}(T,d,V) \ge cT$. Otherwise $Td^2 > 1$. Take $K = \lceil Td^2 \rceil$. Then $1 \le K \le 2Td^2 \le T$, while $B > T$ is equivalent to $V > Td^2$. Consequently, $\frac{K}{Td^2} \ge 1$, and $\frac{V^2}{K^2} \ge \frac{1}{4}$. The block bound again gives $R_{\mathrm{bat}}(T,d,V) \ge cT$. These cases prove Eq. (16) for every $V \ge 0$. We obtained the desired result. □

D Proofs for dynamic classification regret

This appendix collects the general prior-indexed decision reduction that Section 5 instantiates, the proofs of the results stated there, the uniform PU threshold-risk estimate, the end-to-end PU classification bound, and the exact-score special case referred to at the end of that section.

D.1 From prior tracking to prior-indexed decisions

The reduction does not depend on the downstream decision family. Any prior-indexed rule whose excess risk is controlled at the estimated prior inherits the prior-tracking bound, with an additive decision-approximation term. Fixed-score thresholding is the main concrete and computationally simple instance.

Let $\mathcal{H}$ be a class of decision rules $h : \mathcal{X} \to \mathcal{A}$ and let $\ell(\cdot,\cdot) \in [0,1]$ be a bounded loss that does not depend on the prior. Under the class-prior-shift model $P_{U,\pi} = \pi P_+ + (1-\pi)P_-$, write the class-conditional risks

$$R_+(h) := \mathbb{E}_{P_+}[\ell(h(X),1)], \qquad R_-(h) := \mathbb{E}_{P_-}[\ell(h(X),0)],$$

so that the risk under prior q decomposes as

$$R_q(h) = qR_+(h) + (1-q)R_-(h), \qquad R_+(h), R_-(h) \in [0,1]. \tag{30}$$

Because $R_+(h)$ and $R_-(h)$ do not depend on the prior, for every $h \in \mathcal{H}$ and $\pi, q \in [0,1]$,

$$|R_\pi(h) - R_q(h)| = |\pi - q|\,|R_+(h) - R_-(h)| \le |\pi - q|. \tag{31}$$

The zero–one threshold risk $R^q_\pi(\tau) = R_\pi(h_\tau)$ used below is the instance $\mathcal{H} = \mathcal{H}_g$ with ℓ the zero–one loss.

Proposition 5. *Let $q \mapsto \widehat{h}(q) \in \mathcal{H}$ be any, possibly data-dependent, prior-indexed rule that is approximately optimal at its own prior. There is $\alpha(q) \ge 0$ with*

$$R_q(\widehat{h}(q)) \le \inf_{h\in\mathcal{H}} R_q(h) + \alpha(q). \tag{32}$$

Then for every $\pi, q \in [0,1]$,

$$R_\pi(\widehat{h}(q)) - \inf_{h\in\mathcal{H}} R_\pi(h) \le 2|\pi - q| + \alpha(q). \tag{33}$$

Applying Eq. (33) at each round with $q = \widehat{\pi}_t$ and the deployed rule $\widehat{h}_t = \widehat{h}(\widehat{\pi}_t)$ yields dynamic decision regret against the per-round best-in-class comparator.

Corollary 1. *With $\widehat{h}_t = \widehat{h}(\widehat{\pi}_t)$ and $\mathrm{Reg}_T^{\mathcal{H}} := \sum_{t=1}^T \big(R_{\pi_t}(\widehat{h}_t) - \inf_{h\in\mathcal{H}} R_{\pi_t}(h)\big)$,*

$$\mathrm{Reg}_T^{\mathcal{H}} \le 2\sum_{t=1}^T |\widehat{\pi}_t - \pi_t| + \sum_{t=1}^T \alpha(\widehat{\pi}_t) \le 2\sqrt{TL_T^{\mathrm{trk}}} + \sum_{t=1}^T \alpha(\widehat{\pi}_t), \tag{34}$$

where the last step is Cauchy–Schwarz with $L_T^{\mathrm{trk}} = \sum_{t=1}^T (\widehat{\pi}_t - \pi_t)^2$.

The path-dependent PU source uncertainty of Section 4 thus propagates to *any* prior-indexed decision rule through L_T^{trk}; the downstream family enters only through the approximation term $\alpha(\widehat{\pi}_t)$. The main text, Section 5, instantiates α for the fixed-score threshold family.

Proof of Proposition 5. Fix π, q and let $h^\star \in \mathcal{H}$ satisfy $R_\pi(h^\star) \le \inf_{h\in\mathcal{H}} R_\pi(h) + \epsilon$ for an arbitrary $\epsilon > 0$. Decompose

$$R_\pi(\widehat{h}(q)) - R_\pi(h^\star) = \big[R_\pi(\widehat{h}(q)) - R_q(\widehat{h}(q))\big] + \big[R_q(\widehat{h}(q)) - R_q(h^\star)\big] + \big[R_q(h^\star) - R_\pi(h^\star)\big].$$

The first and third terms are at most $|\pi - q|$ by Eq. (31). For the middle term, $\inf_{h\in\mathcal{H}} R_q(h) \le R_q(h^\star)$ and Eq. (32) give $R_q(\widehat{h}(q)) - R_q(h^\star) \le \alpha(q)$. Hence $R_\pi(\widehat{h}(q)) - \inf_{h\in\mathcal{H}} R_\pi(h) \le 2|\pi - q| + \alpha(q) + \epsilon$, and letting $\epsilon \to 0$ proves Eq. (33). □

Remark 2. If a uniform PU risk estimator satisfies $\sup_{h\in\mathcal{H}} |\widehat{R}_q^{\mathrm{PU}}(h) - R_q(h)| \le \epsilon_{\mathcal{H}}(q)$ on a holdout-data event, then its empirical PU risk minimizer $\widehat{h}(q) \in \arg\min_{h\in\mathcal{H}} \widehat{R}_q^{\mathrm{PU}}(h)$ satisfies Eq. (32) with $\alpha(q) \le 2\epsilon_{\mathcal{H}}(q)$, by the usual two-deviation argument. The threshold family $\mathcal{H}_g$ used below has VC dimension one, so the Dvoretzky–Kiefer–Wolfowitz inequality yields a sharp bound without retraining the initial model. Here we adopt it as the paper's main implementable instance and do not pursue the general-classifier training, nonconvex PU optimization, or deep-network retraining that a richer $\mathcal{H}$ would require.

D.2 Uniform PU threshold-risk estimation

Recall the empirical PU threshold risk in Eq. (18), the candidate set in Eq. (19), the threshold map in Eq. (20), and the radius in Eq. (21).

Because $g \in [0,1]$, the two extra values give the all-positive and all-negative rules, and every distinct labeling of the holdout data produced by a threshold appears in this finite set.

Signed empirical PU risk. The estimator in Eq. (18) is used in its raw signed form and can take values outside $[0,1]$. This affects neither the uniform deviation bound below nor the finite-threshold empirical-risk-minimization argument. Nonnegative corrections matter when a flexible model is trained by minimizing a signed PU risk (Kiryo et al., 2017). Here the score g is fixed and only the threshold is selected.

Proposition 6. *With probability at least $1-\delta$ over the holdout data, simultaneously for every $q \in [0,1]$,*

$$\sup_{\tau\in\mathbb{R}} \left|\widehat{R}_q^{g,\mathrm{PU}}(\tau) - R_q^g(\tau)\right| \le \epsilon_R(q,\delta).$$

Consequently,

$$R_q^g\big(\widehat{\tau}^{\mathrm{PU}}(q)\big) - \inf_{\tau\in\mathbb{R}} R_q^g(\tau) \le 2\epsilon_R(q,\delta). \tag{35}$$

The proof applies the Dvoretzky–Kiefer–Wolfowitz inequality, with the sharp constant of Massart (1990), to the positive and source-unlabeled score distributions. A union bound gives the stated event for the two empirical tails. The bound is simultaneous in q because its right-hand side comes pointwise from those same two events. The excess-risk display then follows from the usual two-deviation empirical-risk-minimization argument.

Proof of Theorem 4. Work on the event of Proposition 6, which has probability at least $1-\delta$ over the holdout data. Instantiate Proposition 5 with $\mathcal{H} = \mathcal{H}_g$, the zero–one loss, $q = \widehat{\pi}_t$, and $\alpha(q) = 2\epsilon_R(q,\delta)$ from Eq. (35). Eq. (33) becomes the per-round bound

$$R_{\pi_t}^g(\widehat{\tau}_t) - \inf_{\tau} R_{\pi_t}^g(\tau) \le 2|\pi_t - \widehat{\pi}_t| + 2\epsilon_R(\widehat{\pi}_t,\delta).$$

Summing over t gives the first line. Then, the Cauchy–Schwarz with $L_T^{\mathrm{trk}} = \sum_t (\widehat{\pi}_t - \pi_t)^2$ and $\epsilon_R(\widehat{\pi}_t,\delta) \le \overline{\epsilon}_R(\delta)$ gives the second, which completes the proof. □

D.3 END-TO-END PU CLASSIFICATION BOUND

Corollary 2. *Let $\mathcal{B}_T^{\mathrm{PU}}(\delta_{\mathrm{off}})$ denote the right-hand side of the tracking bound in Theorem 1. With probability at least $1-\delta_{\mathrm{off}}-\delta_{\mathrm{cls}}$ over the holdout data,*

$$\mathbb{E}\Big[\mathrm{Reg}_T^{\mathrm{thr}} \,\Big|\, S_0\Big] \le 2\big(1+\overline{\epsilon}_R(\delta_{\mathrm{cls}})\big)\sqrt{T\mathcal{B}_T^{\mathrm{PU}}(\delta_{\mathrm{off}})} + 2T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}}). \tag{36}$$

Here $\bar{\pi} := \frac{1}{T}\sum_{t=1}^{T}\pi_t$ is the time-averaged prior. Consequently, the explicit tracking bound in Eq. (14) *yields*

$$\begin{aligned}\mathbb{E}\Big[\mathrm{Reg}_T^{\mathrm{thr}} \,\Big|\, S_0\Big] \le{}& 2\sqrt{C}\,\big(1+\overline{\epsilon}_R(\delta_{\mathrm{cls}})\big)\log^{c/2}(eT)\left(\frac{\sqrt{T}}{d_{\min}}+\frac{T^{2/3}V_T^{1/3}}{d_{\min}^{2/3}}\right)\\ &+4\sqrt{2}\,\big(1+\overline{\epsilon}_R(\delta_{\mathrm{cls}})\big)\frac{\sqrt{T\log(4/\delta_{\mathrm{off}})}}{d_{\min}(1-\pi_0)}\sqrt{\frac{W_U(\pi_{1:T})}{n_u}+\frac{W_+(\pi_{1:T})}{n_+}}\\ &+2T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}}),\end{aligned} \tag{37}$$

where $c, C > 0$ is the universal constants.

Proof of Corollary 2. Let $\mathcal{E}_{\mathrm{trk}}$ and $\mathcal{E}_{\mathrm{cls}}$ denote the holdout-data events in Theorem 1 and Proposition 6, respectively. Both events are defined on the same source samples and need not be independent. The union bound nevertheless gives

$$\Pr_{S_0}(\mathcal{E}_{\mathrm{trk}}\cap\mathcal{E}_{\mathrm{cls}}) \ge 1-\delta_{\mathrm{off}}-\delta_{\mathrm{cls}}.$$

Fix holdout data S_0 in $\mathcal{E}_{\mathrm{trk}}\cap\mathcal{E}_{\mathrm{cls}}$. The uniform threshold-risk bound and the first line of Eq. (22) give

$$\mathrm{Reg}_T^{\mathrm{thr}} \le 2\sum_{t=1}^{T}|\widehat{\pi}_t-\pi_t| + 2\sum_{t=1}^{T}\epsilon_R(\widehat{\pi}_t,\delta_{\mathrm{cls}}).$$

The map $q\mapsto\epsilon_R(q,\delta)$ is affine, so $\sum_t\epsilon_R(\pi_t,\delta_{\mathrm{cls}}) = T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}})$, and its slope has absolute value at most $\overline{\epsilon}_R(\delta_{\mathrm{cls}})$; hence

$$\sum_t\epsilon_R(\widehat{\pi}_t,\delta_{\mathrm{cls}}) \le T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}}) + \overline{\epsilon}_R(\delta_{\mathrm{cls}})\sum_t|\widehat{\pi}_t-\pi_t|.$$

Combining and applying Cauchy–Schwarz to $\sum_t|\widehat{\pi}_t-\pi_t| \le \sqrt{TL_T^{\mathrm{trk}}}$,

$$\mathrm{Reg}_T^{\mathrm{thr}} \le 2\big(1+\overline{\epsilon}_R(\delta_{\mathrm{cls}})\big)\sqrt{TL_T^{\mathrm{trk}}} + 2T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}}).$$

The remaining randomness, conditional on S_0, is that of the online target observations and any randomization used by the tracking algorithm. Taking conditional expectations and using conditional Jensen's inequality,

$$\mathbb{E}\Big[\mathrm{Reg}_T^{\mathrm{thr}} \,\Big|\, S_0\Big] \le 2\big(1+\overline{\epsilon}_R(\delta_{\mathrm{cls}})\big)\sqrt{T\,\mathbb{E}[L_T^{\mathrm{trk}} \mid S_0]} + 2T\,\epsilon_R(\bar{\pi},\delta_{\mathrm{cls}}).$$

On the event $\mathcal{E}_{\mathrm{trk}}$, Theorem 1 ensures that $\mathbb{E}[L_T^{\mathrm{trk}} \mid S_0] \le \mathcal{B}_T^{\mathrm{PU}}(\delta_{\mathrm{off}})$. Substitution proves Eq. (36).

Finally, substituting the explicit expression from Eq. (14) and applying $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$ yields Eq. (37). That concludes the proof. □

Discussion. The first line of Eq. (37) is the cost of tracking the changing prior. Up to logarithmic factors and the conditioning parameter $d_{\min}$, it has the same dependence on T and V_T as the PN result of Baby et al. (2023). Thus, the absence of labeled negatives does not change the temporal-tracking exponents. PU attains this PN-type online rate whenever the two source-estimation terms in the remaining lines are of no larger order.

PU supervision instead changes the offline identification cost. Let $s^2 := T^{-1}\sum_{t=1}^{T}(\pi_t - \bar{\pi})^2$ be the path variance. Then

$$W_U(\pi_{1:T}) = T\big[(1-\bar{\pi})^2 + s^2\big], \qquad W_+(\pi_{1:T}) = T\big[(\bar{\pi}-\pi_0)^2 + s^2\big].$$

Hence the tracking contribution of the finite PU source sample depends on both the average location and the fluctuation of the target-prior path. For $\pi_t \in [\pi_0, 1]$, $M(\pi_t)$ interpolates between the two observed source means μ_0 and μ_+. For $\pi_t < \pi_0$, the same representation extrapolates toward the unobserved negative endpoint; this is where the $(1-\pi_0)^{-1}$ factor can be costly. The threshold-risk term has the analogous factor $(1-\bar{\pi})/(1-\pi_0)$, but depends only on the average prior rather than on s^2.

For fixed $d_{\min}$, $\pi_0 < 1$, and confidence levels, sufficient conditions for the displayed bound to be $o(T)$ are

$$V_T = o(T), \qquad \frac{W_U(\pi_{1:T})}{n_u} + \frac{W_+(\pi_{1:T})}{n_+} = o(T), \qquad \epsilon_R(\bar{\pi}, \delta_{\mathrm{cls}}) = o(1).$$

With fixed source sizes these conditions generally fail. Indeed, $W_U + W_+ \geq T(1-\pi_0)^2/2$, and the threshold-risk radius also stays constant, so the corollary does not guarantee sublinear classification regret. Such accumulation of a fixed source-estimation error also occurs under PN supervision. The PU-specific limitation is the path-dependent finite-source term: the matching lower bound in Theorem 2 shows that the W_U, W_+, and $(1-\pi_0)^{-2}$ dependence is unavoidable for squared prior-tracking loss in its unsaturated regime.

D.4 Exact density-ratio special case

Suppose that the fixed score is the exact source posterior,

$$g(x) = \pi_0 \frac{dP_+}{dP_{U,\pi_0}}(x).$$

Adjusting a posterior or density-ratio threshold when the class prior changes is a classical label-shift correction (Saerens et al., 2002; Charoenphakdee & Sugiyama, 2019; Nakajima & Sugiyama, 2023). In this notation, the Bayes classifier under target prior q thresholds g at

$$\tau^\star(q) = \frac{\pi_0(1-q)}{\pi_0(1-q) + (1-\pi_0)q}. \tag{38}$$

Writing

$$\mathrm{Reg}_T^{\mathrm{Bayes}} := \sum_{t=1}^{T} \left\{ R_{\pi_t}(h_{\widehat{\tau}_t}) - R^\star_{\pi_t} \right\},$$

the choice $\widehat{\tau}_t = \tau^\star(\widehat{\pi}_t)$ gives

$$\mathrm{Reg}_T^{\mathrm{Bayes}} \leq 2\sqrt{TL_T^{\mathrm{trk}}}.$$

In this special case, threshold selection adds no finite-source risk-estimation term, so classification regret inherits the prior-tracking bound directly. The empirical PU rule in Eq. (20) applies more generally to an arbitrary fixed score and does not require a posterior or density-ratio scale.

Remark 3. Under the standing class-prior-shift assumption, the exact-score rule above is prior reweighting written as a moving threshold. For a binary probabilistic source classifier $f_0(\cdot \mid x)$ with $g(x) = f_0(1 \mid x)$ and source prior $q_0(1) = \pi_0$, reweighting to a target prior $q(1) = q$,

$$f_q^{\mathrm{rw}}(y \mid x) \propto \frac{q(y)}{q_0(y)}\, f_0(y \mid x),$$

has deterministic $\arg\max$ decision "predict 1" exactly when $g(x) \geq \tau^\star(q)$, with $\tau^\star(q)$ as in Eq. (38). Thus, for binary deterministic decisions, reweighting a fixed source posterior as the prior drifts *is* threshold adaptation of g. When g is the exact posterior this threshold is the target Bayes rule, so the Bayes classifier lies in $\mathcal{H}_g$ for every prior and $\inf_{h\in\mathcal{H}_g} R_\pi(h) = R^\star_\pi$. Instantiating Proposition 5 with $\alpha(q) \equiv 0$ then recovers $\mathrm{Reg}_T^{\mathrm{Bayes}} \leq 2\sqrt{TL_T^{\mathrm{trk}}}$ as the zero-approximation corollary of the general reduction. The oracle here is the exact score rather than the target prior, and the threshold still uses the tracked $\widehat{\pi}_t$.

We note that this equivalence is limited to binary, deterministic $\arg\max$ decisions under the zero–one loss. It does not subsume the multiclass confusion-matrix reweighting, randomized prediction, or general proper-loss decisions of Baby et al. (2023). Reweighted predictive probabilities and proper-loss acts are not a single threshold and lie outside our formulation. In the shared binary deterministic setting, the relevant new difficulty lies in the source supervision rather than in the decision interface: we reconstruct the prior-observation map from the two PU source means rather than from labeled positive–negative data or a known confusion matrix, and this reconstruction carries the path-dependent finite-source cost of Section 4. This comparison does not identify the two tracking algorithms or their statistics outside that common decision setting.

Algorithm 1 Online prior tracking and threshold adaptation from PU source data, with a black-box online-regression oracle $\mathcal{A}$.

Require: holdout data $S_0 = (S_+, S_{\mathrm{U},0})$; known source prior π_0; statistic ϕ and score g, both fixed before S_0 is observed; separation floor $d_{\min}$; batch sizes $N_t \geq 1$; rescaled online-regression oracle $\mathcal{A}$ (Remark 4)

Offline stage (once, before round 1)

$\widehat{\mu}_+ \leftarrow \frac{1}{n_+}\sum_{X\in S_+}\phi(X)$, $\widehat{\mu}_0 \leftarrow \frac{1}{n_u}\sum_{X\in S_{\mathrm{U},0}}\phi(X)$ ▷ the two source-mean estimates used for tracking

$\widehat{\mu}_- \leftarrow \dfrac{\widehat{\mu}_0 - \pi_0\widehat{\mu}_+}{1-\pi_0}$, $\widehat{d} \leftarrow \widehat{\mu}_+ - \widehat{\mu}_-$ ▷ endpoints of $\widehat{M}$, Eq. (4)

if $\widehat{d} < d_{\min}/2$ **then abort**: the separation condition behind Eq. (9) fails on the realised holdout data

$\Gamma_0^{\mathrm{trk}} \leftarrow (\widehat{\mu}_-, \widehat{d})$

precompute $\widehat{\tau}^{\mathrm{PU}} : [0,1] \to \mathbb{R}$, $\widehat{\tau}^{\mathrm{PU}}(q) = \min\arg\min_{\tau\in\mathcal{T}(S_0)} \widehat{R}_q^{g,\mathrm{PU}}(\tau)$; $\Gamma_0^{\mathrm{cls}} \leftarrow \widehat{\tau}^{\mathrm{PU}}$ ▷ Eqs. (19) and (20)

discard S_0; initialise $\mathcal{A}$

Online stage (unsupervised; π_t latent, V_T unknown)

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

  $x_t \leftarrow \mathcal{A}.\textsc{Predict}()$ ▷ a function of $Z_1, \ldots, Z_{t-1}$ only

  $\widehat{\pi}_t \leftarrow \mathrm{proj}_{[0,1]}(x_t)$, $\widehat{\tau}_t \leftarrow \widehat{\tau}^{\mathrm{PU}}(\widehat{\pi}_t)$ ▷ $\mathcal{F}_{t-1}$-measurable

  receive $X_{t,1}, \ldots, X_{t,N_t} \overset{\mathrm{iid}}{\sim} P_{\mathrm{U},t}$ and predict $\widehat{h}_t(X_{t,i}) = \mathbf{1}\{g(X_{t,i}) \geq \widehat{\tau}_t\}$ ▷ labels never revealed

  $\bar{\phi}_t \leftarrow \frac{1}{N_t}\sum_{i=1}^{N_t}\phi(X_{t,i})$, $Z_t \leftarrow (\bar{\phi}_t - \widehat{\mu}_-)/\widehat{d}$ ▷ $\mathbb{E}[Z_t \mid S_0] = \pi_t^{\dagger}$, Eq. (11)

  $\mathcal{A}.\textsc{Update}(Z_t)$

end for

E Experimental details

The implemented procedure. Algorithm 1 states the method analysed in Sections 4 and 5 in the form in which it is run. Everything estimated from the source data is computed once, before round 1, and the raw holdout samples are then discarded; each online round costs one oracle update, one table lookup and one scoring pass over the batch, and never refits the initial model. The online-regression step is deliberately left as a black box $\mathcal{A}$: the reduction of Proposition 2 holds for arbitrary predictions $x_{1:T}$, so the finite-source part of the guarantee is the same whichever oracle fills that slot (Remark 4).

Remark 4 (Instantiating the online-regression oracle)**.** Throughout, $\mathcal{A}$ denotes the *rescaled* oracle: with $a := d_{\min}/2$, the call $\mathcal{A}.\textsc{Update}(Z_t)$ passes $\widetilde{Z}_t = aZ_t$ to the underlying online-regression routine, and $\mathcal{A}.\textsc{Predict}()$ returns $\widetilde{x}_t/a$, where $\widetilde{x}_t$ is that routine's prediction. Only the returned x_t is projected onto $[0,1]$, never $\widetilde{x}_t$. The rescaling is what places the inputs in the range both guarantees assume: on $\mathcal{E}_{\mathrm{src}}(\delta_{\mathrm{off}})$ the rescaled signal and noise satisfy the bounded, zero-mean and variance conditions with $B = 1$ and $\sigma^2 = 1/4$, as established in Proposition 4.

Two routines fit this interface. ALIGATOR (Baby et al., 2021, Figure 4) is the one our analysis uses; Proposition 4 is its one-dimensional specialization, via the restatement in Baby et al. (2023, Proposition 20); its learning rate is fixed in Proposition 4. FLH–FTL, that is, Follow the Leading History (Hazan & Seshadhri, 2007) with running averages as base learners, is the practical alternative, in the form specialized to squared losses by Baby et al. (2023, Appendix F, Algorithm 7). There the mixing learning rate is set to the exp-concavity factor $1/(8B^2)$ of the squared loss; since $|\widetilde{Z}_t| \leq 7/4$ and every prediction is a convex combination of past inputs, $B = 2$ is admissible and the rate is $1/32$. Neither routine takes the total variation V_T as an input, which is what lets Theorem 1 hold without tuning.

The analysis is stated for $N_t = 1$, but Algorithm 1 runs with any batch sizes $N_t \geq 1$ fixed in advance: averaging within a round leaves the conditional mean and the independence across rounds intact and only reduces the noise variance, so the constants above remain valid and the guarantee holds verbatim, without becoming sharper as N_t grows (Appendix C.2). The remaining parameters

are ours rather than the oracle's: the separation floor $d_{\min}$, fixed before the holdout data are seen and never read off $\widehat{d}$, and the confidence levels δ_{off} and δ_{cls}.

E.1 Datasets and full implementation details

We evaluate the complete procedure in Algorithm 1, including estimation from finite PU holdout data, class-prior tracking, and online threshold adaptation. The experiments cover image and text representations and compare the two online regression oracles discussed above. Target labels are used only to evaluate the methods and are never revealed during online adaptation.

Data and offline training. We form five binary tasks: digit 0 versus the remaining digits on MNIST, T-shirt/top versus the remaining classes on Fashion-MNIST, airplane versus the remaining classes on CIFAR-10, AnnualCrop versus the remaining classes on EuroSAT, and Sci/Tech versus the remaining classes on AG News. MNIST and Fashion-MNIST pixels are scaled to $[0, 1]$ and flattened into 784-dimensional vectors. CIFAR-10 and EuroSAT use standardized 512-dimensional features from a fixed ImageNet-pretrained ResNet-18, while AG News uses standardized 768-dimensional `[CLS]` features from a fixed DistilBERT encoder. We use the official train–test splits except for EuroSAT, which is divided into a stratified $80/20$ split with seed 42. The source prior is $\pi_0 = 0.1$ in every task.

For MNIST, Fashion-MNIST, and CIFAR-10, the PU initial model is trained from $2{,}000$ positive and $5{,}000$ source-unlabeled examples; the PU holdout contains $1{,}000$ positive and $2{,}000$ source-unlabeled examples. The PN model uses $5{,}000$ labeled source examples and a holdout with $1{,}000$ positive and $2{,}000$ negative examples. On EuroSAT, the PU training counts are $800/3{,}000$, the PN model uses $3{,}000$ labeled source examples, and the positive/source-unlabeled/negative holdout counts are $400/1{,}000/1{,}000$. The corresponding counts on AG News are $500/2{,}000$, $2{,}000$, and $300/800/800$, respectively. Offline training and holdout sets are mutually disjoint.

Both initial models are ReLU multilayer perceptrons with hidden widths $(100, 100, 20)$ and a scalar logistic output. We train for three epochs with mini-batches of size 256, learning rate 0.03, momentum 0.9, and weight decay 10^{-4}. The PU model uses the nonnegative PU logistic objective (Kiryo et al., 2017), with correction parameters $\beta = 0$ and $\gamma = 1$; the PN model uses ordinary binary logistic loss. For all PU methods we take $\phi = g$, where g is the logistic score of the PU initial model. The orientation of ϕ is fixed using only the offline training data, before the holdout data are used. We set $d_{\min} = 0.02$. Across all runs, the estimated separation satisfies $\widehat{d} \in [0.392, 0.777]$, which passes the separation check in Algorithm 1.

Online prior paths. Each run has horizon $T = 1{,}000$ and batch size $N_t = 10$. At round t, we first draw each binary label with positive-class probability π_t and then sample with replacement from the corresponding class in the test split. All methods receive exactly the same unlabeled online batches for a given seed and prior path. We use three seeds and four paths, the priors are all with values in $[0.05, 0.95]$:

- **Linear:** $\pi_t = 0.05 + 0.9(t-1)/T$.
- **Sine:** two repeated low–high–low half-sine cycles, with period $p = T/2$ and $\pi_t = 0.05 + 0.9\sin\big(\pi((t-1) \bmod p)/p\big)$.
- **Square:** $\pi_t = 0.05$ for the first 500 rounds and $\pi_t = 0.95$ for the remaining 500 rounds.
- **Bernoulli:** a two-state path on $\{0.05, 0.95\}$, initialized at 0.05 and switched after each round with probability $1/\sqrt{T}$.

Proposed methods. PU–Aligator and PU–FLH–FTL both implement Algorithm 1. They reconstruct $\widehat{\mu}_-$ and $\widehat{d}$ from the PU holdout data, precompute the empirical threshold map $\widehat{\tau}^{\mathrm{PU}}$, and form the raw observation

$$Z_t = \frac{\bar{\phi}_t - \widehat{\mu}_-}{\widehat{d}}.$$

At round t, the online regression oracle predicts from $Z_1, \dots, Z_{t-1}$, its output is projected onto $[0, 1]$, and the resulting $\widehat{\pi}_t$ selects the threshold $\widehat{\tau}^{\mathrm{PU}}(\widehat{\pi}_t)$. The raw Z_t itself is not projected. Both variants use the $d_{\min}/2$ rescaling in Remark 4 and initialize the underlying online averages at zero.

PU–ALIGATOR uses the learning rate in Proposition 4; PU–FLH–FTL uses mixing rate $1/32$. In both cases the classification score g remains fixed throughout online adaptation, and only the class-prior estimate and threshold change.

Comparison methods. We use the following comparisons.

- STATIC applies the PU-trained initial model with threshold $1/2$ in every round and performs no online adaptation.

- PN–FLH–FTL is the binary soft-confusion implementation of the UOLS construction of Baby et al. (2023). We temperature-scale the PN score on its positive and negative holdout data, estimate the two conditional score means, invert their affine relation to obtain a local class-prior estimate, and project this estimate onto $[0, 1]$ as in binary BBSE. FLH–FTL starts from the source prior and uses mixing rate $1/2 = 1/K$ for $K = 2$ classes. Classification uses prior-ratio reweighting of the fixed PN score. PN–ALIGATOR is a tracker-swapped diagnostic that uses the same PN initial model, local estimate, and reweighted classifier, but replaces FLH–FTL by Aligator initialized at the source prior with learning rate $1/8$. It is not a separate method proposed by Baby et al. (2023).

- PU–UOGD and PU–ATLAS are natural PU extensions of the online optimization methods of Bai et al. (2022). PU score is used to form the clipped one-round-lagged local class-prior estimate. For an augmented last-layer feature vector and head w, the negative conditional logistic risk and gradient are reconstructed from positive and source-unlabeled offline training mini-batches through

$$\widehat{R}_{-}(w) = \frac{\widehat{R}^{-}_{\mathrm{U}}(w) - \pi_0 \widehat{R}^{-}_{+}(w)}{1 - \pi_0}.$$

  The estimated target risk is then $\widehat{\pi}_t \widehat{R}^{+}_{+}(w) + (1 - \widehat{\pi}_t)\widehat{R}_{-}(w)$. Consequently, both methods retain and revisit raw PU offline training examples during online adaptation. This differs from the proposed methods, which use only precomputed source quantities and keep the classification score g fixed while adapting its threshold. Both methods keep the learned representation fixed and update only the final affine layer after predicting the current batch. UOGD uses step size 0.05, source mini-batches of size 256, and projection onto an ℓ_2 ball of radius 20. ATLAS combines six UOGD learners whose step sizes form a factor-of-two geometric grid centered at 0.05, with Hedge learning rate 0.5.

- PU PRIOR ORACLE supplies the generating π_t to the PU threshold map while retaining the same PU initial model and holdout estimates as proposed methods. PN PRIOR ORACLE supplies π_t to the empirical PN threshold map built from labeled positive and negative holdout data. These oracle methods use the true current class prior, which is unavailable to the learner at prediction time. They are included as evaluation references and are not deployable.

Metrics. Classification error is averaged over all TN_t online examples. Prior MSE is $T^{-1}\sum_{t=1}^{T}(\widehat{\pi}_t - \pi_t)^2$. We report the mean and standard deviation over three seeds. Runtime is the mean time of the online routine for one complete trajectory; it excludes initial-model training and holdout estimation, and is included only as an indicative comparison on the same server.

Table 2: Complete classification error in percent on all five datasets under all four class-prior paths (lower is better). The smallest mean among the non-oracle methods in each row is bold. The PN columns use labeled negative offline data, whereas all PU columns use no labeled negatives.

Dataset	STATIC	PU–ALIGATOR	PU–FLH–FTL	PN–ALIGATOR	PN–FLH–FTL	PU–UOGD	PU–ATLAS	PU PRIOR ORACLE	PN PRIOR ORACLE
Linear shift									
MNIST	2.92 ±0.39	2.47 ±0.39	**2.00 ±0.35**	2.10 ±0.41	2.06 ±0.40	2.18 ±0.32	2.21 ±0.29	2.05 ±0.37	2.18 ±0.40
Fashion-MNIST	17.29 ±3.53	8.11 ±0.54	7.42 ±0.42	7.33 ±0.32	7.11 ±0.40	6.82 ±0.13	**6.70 ±0.13**	6.80 ±0.20	6.90 ±0.35
CIFAR-10	13.53 ±0.93	7.97 ±0.17	7.03 ±0.07	6.75 ±0.10	**6.39 ±0.24**	7.26 ±0.36	7.12 ±0.23	6.31 ±0.13	6.18 ±0.29
EuroSAT	10.98 ±4.96	4.35 ±0.22	3.60 ±0.27	3.96 ±0.49	3.84 ±0.43	3.47 ±0.34	**3.35 ±0.15**	3.65 ±0.28	3.80 ±0.19
AG News	25.05 ±1.83	10.46 ±0.13	9.52 ±0.27	9.41 ±0.70	**8.82 ±0.47**	9.30 ±0.19	8.98 ±0.10	8.71 ±0.25	8.70 ±0.34
Sine shift									
MNIST	3.67 ±0.46	2.43 ±0.46	2.26 ±0.26	**2.13 ±0.32**	2.15 ±0.27	2.44 ±0.26	2.29 ±0.38	2.02 ±0.37	2.04 ±0.23
Fashion-MNIST	20.97 ±4.71	7.36 ±0.67	7.21 ±0.71	6.59 ±0.20	6.75 ±0.39	5.90 ±0.22	**5.70 ±0.31**	5.87 ±0.55	6.03 ±0.28
CIFAR-10	16.61 ±1.02	7.12 ±0.22	6.74 ±0.33	6.42 ±0.31	**6.33 ±0.36**	6.89 ±0.13	6.36 ±0.28	5.60 ±0.31	5.60 ±0.19
EuroSAT	13.05 ±6.39	3.76 ±0.25	3.86 ±0.17	3.37 ±0.27	3.35 ±0.24	3.28 ±0.07	**3.03 ±0.04**	3.24 ±0.16	3.18 ±0.18
AG News	30.70 ±2.19	9.71 ±0.31	9.60 ±0.37	8.81 ±0.67	8.63 ±0.59	9.26 ±0.62	**8.50 ±0.54**	7.83 ±0.34	7.81 ±0.54
Square shift									
MNIST	2.97 ±0.42	1.93 ±0.40	1.60 ±0.21	1.22 ±0.17	**1.08 ±0.02**	1.22 ±0.21	1.19 ±0.25	1.13 ±0.28	0.96 ±0.02
Fashion-MNIST	17.66 ±3.46	5.20 ±0.28	5.25 ±0.38	4.03 ±0.20	3.43 ±0.34	3.33 ±0.14	**3.16 ±0.09**	3.01 ±0.25	2.97 ±0.41
CIFAR-10	13.57 ±1.03	4.39 ±0.10	4.65 ±0.21	3.64 ±0.24	**2.90 ±0.07**	4.24 ±0.04	4.00 ±0.12	2.64 ±0.08	2.45 ±0.09
EuroSAT	10.47 ±4.95	2.98 ±0.24	2.48 ±0.14	1.98 ±0.13	**1.70 ±0.07**	2.04 ±0.11	1.94 ±0.24	1.76 ±0.09	1.64 ±0.10
AG News	24.96 ±1.43	6.22 ±0.19	6.28 ±0.31	5.12 ±0.39	**4.24 ±0.59**	5.67 ±0.36	5.09 ±1.09	3.61 ±0.20	3.60 ±0.19
Bernoulli shift									
MNIST	3.28 ±0.58	2.35 ±0.21	2.18 ±0.24	**1.72 ±0.47**	2.02 ±0.42	2.42 ±0.26	1.95 ±0.11	1.09 ±0.17	1.07 ±0.18
Fashion-MNIST	21.13 ±5.91	7.59 ±0.36	7.91 ±0.48	6.27 ±0.12	6.81 ±0.25	5.97 ±0.20	**4.50 ±0.11**	3.30 ±0.18	3.13 ±0.22
CIFAR-10	15.47 ±1.87	6.43 ±0.43	7.00 ±0.38	**5.48 ±0.35**	5.88 ±0.50	7.22 ±0.53	5.78 ±0.30	2.58 ±0.19	2.34 ±0.09
EuroSAT	11.58 ±4.94	3.61 ±0.26	4.10 ±0.43	3.00 ±0.53	3.37 ±0.43	3.70 ±0.45	**2.97 ±0.26**	1.79 ±0.11	1.76 ±0.23
AG News	28.94 ±3.70	8.74 ±0.16	9.33 ±0.29	**7.57 ±0.31**	8.06 ±0.44	9.23 ±1.07	7.62 ±0.53	3.65 ±0.31	3.61 ±0.11

Table 3: Class-prior MSE (multiplied by 10^2) and online computation time. The smallest mean in each row is bold. The final block reports mean online runtime in seconds per trajectory, averaged over seeds and paths.

Dataset	PU–ALIGATOR	PU–FLH–FTL	PN–ALIGATOR	PN–FLH–FTL	PU–UOGD	PU–ATLAS
Linear shift						
MNIST	3.44 ±0.08	1.61 ±0.04	2.19 ±0.04	**0.70 ±0.04**	2.13 ±0.15	2.13 ±0.15
Fashion-MNIST	3.81 ±0.24	1.88 ±0.26	2.56 ±0.19	**0.95 ±0.16**	2.97 ±0.23	2.97 ±0.23
CIFAR-10	3.75 ±0.07	1.76 ±0.10	2.43 ±0.12	**0.85 ±0.11**	2.78 ±0.19	2.78 ±0.19
EuroSAT	3.56 ±0.18	1.74 ±0.21	2.25 ±0.11	**0.77 ±0.12**	2.46 ±0.22	2.46 ±0.22
AG News	3.84 ±0.28	1.95 ±0.26	2.42 ±0.24	**0.84 ±0.25**	3.18 ±0.28	3.18 ±0.28
Sine shift						
MNIST	4.25 ±0.11	4.60 ±0.11	2.65 ±0.16	2.29 ±0.15	**1.77 ±0.07**	**1.77 ±0.07**
Fashion-MNIST	4.59 ±0.09	4.68 ±0.09	3.03 ±0.04	**2.45 ±0.15**	2.64 ±0.09	2.64 ±0.09
CIFAR-10	4.62 ±0.03	4.62 ±0.04	3.02 ±0.03	**2.42 ±0.09**	2.44 ±0.11	2.44 ±0.11
EuroSAT	4.32 ±0.14	4.71 ±0.08	2.71 ±0.01	2.44 ±0.13	**2.21 ±0.25**	**2.21 ±0.25**
AG News	4.69 ±0.32	4.64 ±0.15	3.11 ±0.22	**2.53 ±0.20**	2.87 ±0.43	2.87 ±0.43
Square shift						
MNIST	4.50 ±0.14	6.92 ±0.18	2.65 ±0.10	0.87 ±0.03	**0.62 ±0.02**	**0.62 ±0.02**
Fashion-MNIST	4.95 ±0.19	7.43 ±0.18	3.20 ±0.12	**1.06 ±0.02**	1.35 ±0.23	1.35 ±0.23
CIFAR-10	4.74 ±0.11	7.16 ±0.23	3.03 ±0.12	**1.01 ±0.06**	1.14 ±0.10	1.14 ±0.10
EuroSAT	4.52 ±0.25	7.02 ±0.25	2.74 ±0.11	0.89 ±0.03	**0.81 ±0.06**	**0.81 ±0.06**
AG News	4.96 ±0.42	7.29 ±0.32	3.12 ±0.26	**1.08 ±0.14**	1.46 ±0.31	1.46 ±0.31
Bernoulli shift						
MNIST	10.66 ±0.75	16.38 ±0.42	9.40 ±0.81	12.26 ±0.63	**3.10 ±0.41**	**3.10 ±0.41**
Fashion-MNIST	11.34 ±1.03	16.37 ±0.41	10.01 ±0.76	12.54 ±0.67	**3.84 ±0.22**	**3.84 ±0.22**
CIFAR-10	11.08 ±0.98	16.36 ±0.43	9.83 ±0.89	12.49 ±0.68	**3.46 ±0.38**	**3.46 ±0.38**
EuroSAT	10.73 ±0.83	16.47 ±0.40	9.53 ±0.76	12.36 ±0.66	**3.19 ±0.43**	**3.19 ±0.43**
AG News	11.05 ±0.57	16.36 ±0.42	9.84 ±0.58	12.39 ±0.60	**3.61 ±0.29**	**3.61 ±0.29**
Online runtime (seconds per trajectory)						
MNIST	**0.112**	0.134	0.123	0.146	0.280	0.871
Fashion-MNIST	**0.113**	0.136	0.125	0.148	0.290	0.896
CIFAR-10	**0.107**	0.129	0.118	0.141	0.274	0.913
EuroSAT	**0.107**	0.126	0.119	0.141	0.267	0.880
AG News	**0.110**	0.133	0.126	0.149	0.282	0.884

Results. Table 1 in the main paper reports representative classification results for three datasets and two class-prior paths. Tables 2 and 3 report the complete classification results and the corresponding prior-tracking error and online runtime. The overall conclusions are given in the main experimental discussion.

E.2 SENSITIVITY TO THE OFFLINE CLASS PRIOR

Setup. We isolate the dependence of prior tracking on the offline class prior π_0. On AG News, we train the PU initial model once using the same offline training configuration as above, whose source-unlabeled training sample has positive-class proportion 0.1, and then keep the statistic $\phi = g$ fixed. For each of 20 replicates, we construct PU holdout data with $\pi_0 \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, using $n_+ = 300$ positive and $n_u = 800$ source-unlabeled observations. The PN controls use the same fixed statistic, the same online regression oracles, and $n_- = 800$ labeled negative holdout observations; they differ only in estimating μ_- directly rather than reconstructing it from PU holdout

data. Within each replicate, every value of π_0 uses exactly the same Linear target-prior path and the same unlabeled online batches. The path ranges from 0.05 to 0.95, with $T = 1{,}000$, $N_t = 10$, and $V_T = 0.8991$. Thus, only the holdout distribution and the PU reconstruction through π_0 change. This experiment evaluates class-prior MSE only; it does not compare classification error.

For reference, the last column of Table 4 reports the π_0-dependent part of the path-dependent finite-source term, up to constants common to all rows,

$$\frac{1}{T(1-\pi_0)^2}\left\{\frac{W_U(\pi_{1:T})}{n_u}+\frac{W_+(\pi_{1:T})}{n_+}\right\}.$$

Table 4: Sensitivity of class-prior tracking to the offline class prior on AG News. Class-prior MSE and the finite-source bound proxy are multiplied by 10^2. MSE entries are mean $\pm$ standard deviation over 20 replicates. The target-prior path, online batches, fixed statistic, and V_T are shared across π_0 within each replicate.

π_0	PU–ALIGATOR	PU–FLH–FTL	PN–ALIGATOR	PN–FLH–FTL	Finite-source bound proxy
0.1	3.643 $\pm$0.189	1.761 $\pm$0.241	3.635 $\pm$0.215	1.745 $\pm$0.278	0.1425
0.3	3.657 $\pm$0.144	1.785 $\pm$0.195	3.635 $\pm$0.215	1.745 $\pm$0.278	0.1541
0.5	3.676 $\pm$0.176	1.831 $\pm$0.355	3.635 $\pm$0.215	1.745 $\pm$0.278	0.2490
0.7	3.782 $\pm$0.438	2.058 $\pm$0.852	3.635 $\pm$0.215	1.745 $\pm$0.278	0.8404
0.9	5.271 $\pm$3.173	5.085 $\pm$6.033	3.635 $\pm$0.215	1.745 $\pm$0.278	11.5697

Results. The PN baselines are numerically unchanged across π_0, confirming that the target-prior path and online observations are held fixed. The PU methods remain close to their PN counterparts for $\pi_0 \leq 0.5$, begin to deteriorate at $\pi_0 = 0.7$, and show substantially larger error and variability at $\pi_0 = 0.9$. Relative to $\pi_0 = 0.1$, the class-prior MSE at $\pi_0 = 0.9$ increases by 44.7% for PU–ALIGATOR and 188.8% for PU–FLH–FTL, while the finite-source bound proxy increases by a factor of 81.2. This behavior is qualitatively consistent with the $(1-\pi_0)^{-2}$ dependence in Theorem 1: as π_0 approaches one, reconstructing μ_- from PU holdout data becomes increasingly ill-conditioned. The increase is modest at intermediate values of π_0; the main empirical effect is the predicted instability when the offline positive-class prior is close to one.