ONLINE MULTI-SOURCE TRANSFER LEARNING UNDER SHIFTING SOURCE RELIABILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Transfer learning leverages labeled source data to improve online prediction when labeled target data are limited. However, both target and source distributions may change over time. A source that helps early in the stream may then become unhelpful, making fixed transfer weights unreliable. We propose Adaptive Multi-Source Transfer Learning (AMTL), which combines a target-only predictor with multiple source-specific transfer predictors. After each target batch, AMTL updates predictor weights based on their losses and uses a structured transition to adjust how much to transfer and which sources to trust. The theoretical analysis establishes a pathwise cumulative-loss bound relative to target-only prediction and examines the sources of error for sparse logistic models. Experiments on synthetic streams, CIFAR-10, and hospital records show that AMTL performs well in online prediction relative to existing methods.

1 INTRODUCTION

Transfer learning aims to improve prediction in a target domain with limited labeled data by using information from related source domains (Pan & Yang, 2010). Its effectiveness depends on how closely the sources match the target. A well-matched source leads to positive transfer, whereas a mismatched source can degrade prediction, a phenomenon known as negative transfer (Zhang et al., 2023). Therefore, the key to effective transfer is to exploit helpful sources while guarding against harmful ones. This challenge becomes more pronounced when multiple sources are available, as their relevance to the target may vary. Multi-source methods address this variation by weighting source contributions (Mansour et al., 2008; Hoffman et al., 2018; Liu & Lin, 2024) or identifying informative sources (Mansour et al., 2021; Li et al., 2022; Tian & Feng, 2023; Li et al., 2024).

In many applications, target observations arrive sequentially, motivating online transfer learning. Zhao et al. (2014) introduce a setting in which predictions on incoming target examples use a source model and previously observed target data. Wu et al. (2017) extend this setting to multiple sources. Subsequent work analyzed online performance and the conditions under which negative transfer occurs (Wu et al., 2021; 2023). However, source domains may change gradually over time (Gama et al., 2014). For example, a target hospital updating a cardiovascular risk model may draw on records from another hospital. When local labeled data are scarce, records from a well-matched source hospital can improve prediction. As that hospital's patient population or measurement practices change, its records may later harm prediction at the target hospital (Finlayson et al., 2021). This raises a natural question: how can we perform online transfer learning as source domains evolve?

Existing online transfer methods rarely address changes in source reliability, and many assume that source distributions remain fixed over time (Zhao et al., 2014; Wu et al., 2017; 2021). Recent work considers time-varying transfer (Wu et al., 2023) or drifting multi-source streams through adaptive ensembles (Du et al., 2019) and representation alignment (Xie & Pratama, 2022). Despite these advances, determining whether transfer helps and which sources remain useful as their distributions change remains challenging.

To address this challenge, we propose Adaptive Multi-Source Transfer Learning (AMTL), which combines a target-only predictor with a separate transfer predictor for each source. As target labels arrive, AMTL adjusts both the overall reliance on transfer and the weights assigned to individual sources.

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.

Our contributions are as follows.

- An adaptive weighting method based on a structured state graph that separately adjusts the overall reliance on transfer and the weights assigned to individual sources as their usefulness changes.
- A pathwise log-loss bound that compares the predictions made by AMTL with changing sequences of target-only and source predictors, including an explicit comparison with target-only prediction.
- A finite-sample analysis of the transfer predictors that accounts for estimation error, transfer bias, and numerical error, and connects these errors to cumulative online loss.
- An empirical evaluation across eight synthetic settings, heart disease records from four hospitals, and frozen ResNet-18 features from CIFAR-10. AMTL achieves the best average online rank in the synthetic settings and lower cumulative online loss than all seven non-ablation baselines on the hospital data.

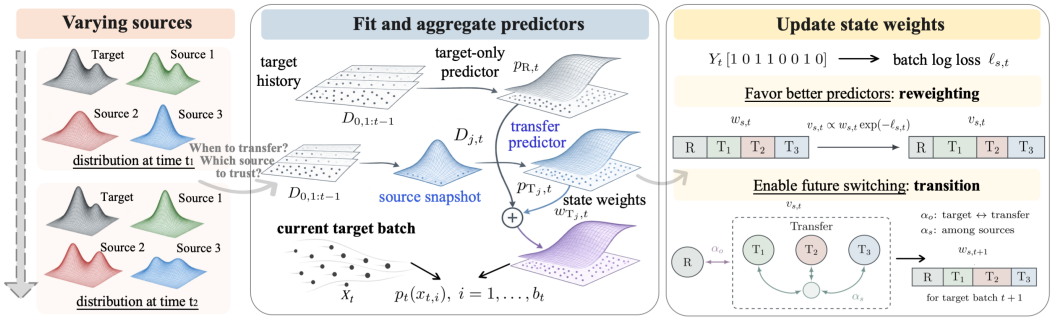


Figure 1: AMTL overview under changing source distributions. AMTL fits and aggregates predictors to predict each target batch before its labels are revealed, then updates state weights through reweighting and a transition step after observing the labels.

2 Problem Formulation

We consider a multi-source transfer learning problem for binary prediction in which the source distributions may change over time. Let $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ denote the sample space, where $\mathcal{X} \subseteq \mathbb{R}^d$ and $\mathcal{Y} = \{0, 1\}$. The problem comprises a single target domain, indexed by 0, and m source domains, indexed by $j = 1, \ldots, m$. The target distribution is P_0, and the distribution of source j at time t is $P_{j,t}$. The distributions P_0 and $P_{j,t}$ may differ both in the marginal distribution of X and in the conditional distribution of Y given X.

We use logistic regression as the prediction class:

$$\Pr_\theta(Y = 1 \mid X = x) = \sigma(x^\top \theta),$$

where $\theta \in \mathbb{R}^d$ and $\sigma(a) := (1 + e^{-a})^{-1}$ is the logistic function. For $z = (x, y) \in \mathcal{Z}$, the logistic loss is

$$\ell(\theta; z) = -yx^\top \theta + \log(1 + e^{x^\top \theta}).$$

Consider an online setting over times $t = 1, \ldots, T$. At time t, $D_{j,t}$ denotes the current labeled snapshot of source j, which may persist until it is replaced. The target batch $D_{0,t}$ becomes complete after its labels are revealed. Write

$$D_{j,t} = \{(x_{j,t,k}, y_{j,t,k})\}_{k=1}^{n_{j,t}}, \quad j = 1, \ldots, m, \qquad D_{0,t} = \{(x_{t,i}, y_{t,i})\}_{i=1}^{b_t},$$

where the observations in $D_{0,t}$ are drawn from P_0 and those in $D_{j,t}$ are drawn from $P_{j,t}$. At each time t, estimation and prediction proceed in the following three steps.

1. *Estimation.* The available data are the current source snapshots $D_{1:m,t} = (D_{1,t}, \ldots, D_{m,t})$ and the target history $D_{0,1:t-1} = (D_{0,1}, \ldots, D_{0,t-1})$.

2. *Prediction.* The target covariates $X_t = (x_{t,1}, \ldots, x_{t,b_t})$ are observed. The learner outputs $p_t(x_{t,i}) \in [0,1]$ for each $i = 1, \ldots, b_t$.
3. *Feedback.* The true labels $Y_t = (y_{t,1}, \ldots, y_{t,b_t})$ are revealed, and $D_{0,t}$ becomes available from time $t+1$ onward.

Our goal is to predict the target labels accurately. The difficulty is that relying on the target data alone is inefficient when the target history is short, whereas relying on the source data may be harmful when the source distributions differ from P_0. Moreover, since the source distributions change over time, the usefulness of each source is unknown and may vary with t. The central question is therefore how to determine, at each time t and using only the information available at that time, how much to rely on each source.

3 The AMTL Method

This section presents AMTL. Section 3.1 constructs a target-only predictor based on the target data and a Transfer predictor for each source. Section 3.2 aggregates the two predictors in the single-source case, and Section 3.3 extends the aggregation to multiple sources.

3.1 Target-only and Transfer Predictors

At each time t, before Y_t is revealed, two types of predictors are constructed: a target-only predictor and a Transfer predictor for each source. For $t \geq 2$, let $N_t := \sum_{r=1}^{t-1} b_r$ denote the number of target observations available before time t. The empirical target risk and the empirical risk of source j at time t are defined as

$$\widehat{L}_{0,t-1}(\theta) = \frac{1}{N_t} \sum_{r=1}^{t-1} \sum_{i=1}^{b_r} \ell\big(\theta; (x_{r,i}, y_{r,i})\big), \widehat{L}_{j,t}(\theta) = \frac{1}{n_{j,t}} \sum_{k=1}^{n_{j,t}} \ell\big(\theta; (x_{j,t,k}, y_{j,t,k})\big), \quad j = 1, \ldots, m.$$

Following weighted empirical risk formulations (Ben-David et al., 2010) and high-dimensional transfer methods for generalized linear models (Li et al., 2022; Tian & Feng, 2023), we fit the following penalized objectives:

$$F_{\mathrm{R},t}(\theta) := \widehat{L}_{0,t-1}(\theta) + \lambda_{\mathrm{R},t}\|\theta\|_1, \tag{1}$$

$$F_{j,t}(\theta) := \gamma\, \widehat{L}_{0,t-1}(\theta) + (1-\gamma)\, \widehat{L}_{j,t}(\theta) + \lambda_{j,t}\|\theta\|_1, \quad j = 1, \ldots, m, \tag{2}$$

where $\lambda_{\mathrm{R},t}, \lambda_{j,t} > 0$ are regularization parameters and $\gamma \in [0,1)$ balances the target and source risks. Let $\bar{\theta}_{\mathrm{R},t}$ and $\bar{\theta}_{j,t}$ be exact minimizers of these objectives. The fitted estimators $\widehat{\theta}_{\mathrm{R},t}$ and $\widehat{\theta}_{j,t}$ are their numerical approximations. The target-only fit uses the target history; Transfer fit j also uses its current source snapshot.

To bound the aggregation log loss, define the clipping map $\mathrm{clip}_\epsilon(a) := \min\{1-\epsilon, \max\{\epsilon, a\}\}$ for $0 < \epsilon < 1/2$. The target-only predictor and the Transfer predictors are then given by

$$p_{\mathrm{R},t}(x) := \mathrm{clip}_\epsilon\big\{\sigma(x^\top \widehat{\theta}_{\mathrm{R},t})\big\}, \quad p_{\mathrm{T}_j,t}(x) := \mathrm{clip}_\epsilon\big\{\sigma(x^\top \widehat{\theta}_{j,t})\big\}, \quad j = 1, \ldots, m. \tag{3}$$

At $t = 1$, only the source snapshots are available. Thus, the target-only predictor is set to $p_{\mathrm{R},1} \equiv 1/2$, and each Transfer estimator is obtained by approximately minimizing $\widehat{L}_{j,1}(\theta) + \lambda_{j,1}\|\theta\|_1$, which uses the source snapshot alone.

3.2 Single Source Aggregation

Consider first the case $m = 1$, in which the target-only predictor $p_{\mathrm{R},t}$ and a single Transfer predictor $p_{\mathrm{T},t}$ are available at each time t. The two predictors are combined into the final predictor:

$$p_t(x) := (1-q_t)\, p_{\mathrm{R},t}(x) + q_t\, p_{\mathrm{T},t}(x),$$

where the weight $q_t \in [0,1]$ controls the reliance on the source. The resulting predictions are $p_t(x_{t,i})$, $i = 1, \ldots, b_t$. A larger q_t places more weight on the Transfer predictor and less on the target-only predictor. The initial weight $q_1 \in (0,1)$ is a user-specified parameter that reflects the initial reliance on the source. For instance, $q_1 = 1/2$ treats the two predictors equally.

After Y_t is revealed, the weight q_t is updated to q_{t+1} in two steps. The first step is a reweighting step. Each predictor is evaluated by its average log loss on the target batch at time t,

$$\ell_t(p) := \frac{1}{b_t} \sum_{i=1}^{b_t} \Big[- y_{t,i} \log p(x_{t,i}) - (1 - y_{t,i}) \log\{1 - p(x_{t,i})\} \Big], \tag{4}$$

For any predictor p taking values in $[\epsilon, 1-\epsilon]$, $0 \le \ell_t(p) \le B := -\log \epsilon$. Exponential weighting (Cesa-Bianchi & Lugosi, 2006) gives the intermediate Transfer weight

$$v_t = \frac{q_t \exp\{-\ell_t(p_{\mathrm{T},t})\}}{(1-q_t)\exp\{-\ell_t(p_{\mathrm{R},t})\} + q_t \exp\{-\ell_t(p_{\mathrm{T},t})\}}.$$

Second, since the usefulness of the source may change over time, the predictor that performs better at time t may not remain so at time $t+1$. To account for such changes, a transition step is applied, in which a fraction $\alpha_o \in [0, 1/2]$ of the weight of each predictor is moved to the other,

$$q_{t+1} = (1 - \alpha_o)\, v_t + \alpha_o (1 - v_t). \tag{5}$$

This is the fixed-share method of Herbster & Warmuth (1998). It guarantees that $\alpha_o \le q_{t+1} \le 1 - \alpha_o$, so that neither predictor is discarded and the aggregation can shift back to either predictor when the usefulness of the source changes.

3.3 Multi-Source Aggregation

With m sources, the target-only predictor and the m Transfer predictors are aggregated in the same manner. Let $\mathcal{S} := \{\mathrm{R}, \mathrm{T}_1, \ldots, \mathrm{T}_m\}$ denote the state set. Here R is the target-only reference state and T_j corresponds to source j. Each state $s \in \mathcal{S}$ has predictor $p_{s,t}$ from (3) and carries a weight $w_{s,t} \ge 0$ with $\sum_{s \in \mathcal{S}} w_{s,t} = 1$. The final predictor is

$$p_t(x) := \sum_{s \in \mathcal{S}} w_{s,t} p_{s,t}(x) = w_{\mathrm{R},t} p_{\mathrm{R},t}(x) + \sum_{j=1}^{m} w_{\mathrm{T}_j,t} p_{\mathrm{T}_j,t}(x). \tag{6}$$

The total weight on the Transfer predictors, $q_t := \sum_{j=1}^{m} w_{\mathrm{T}_j,t}$, plays the role of the transfer mass in Section 3.2 and controls the overall reliance on the sources. For $q_t > 0$, the ratios $w_{\mathrm{T}_j,t}/q_t$ describe how this reliance is divided among the sources. The weights are initialized at $w_{\mathrm{R},1} = 1 - q_1$ and $w_{\mathrm{T}_j,1} = q_1/m$, so that the initial transfer mass is spread evenly across the sources.

After Y_t is revealed, the weights are updated in the same two steps as in Section 3.2. The first step is a reweighting step. Let $\ell_{s,t} := \ell_t(p_{s,t})$ denote the average log loss (4) of the predictor $p_{s,t}$. The intermediate weights are

$$v_{s,t} := \frac{w_{s,t} \exp(-\ell_{s,t})}{\sum_{a \in \mathcal{S}} w_{a,t} \exp(-\ell_{a,t})}, \quad s \in \mathcal{S}. \tag{7}$$

The second step is a transition step. With multiple sources, two kinds of movement are needed: between the target-only predictor and the Transfer predictors, controlled by $\alpha_o \in [0, 1/2]$, and among the Transfer predictors of different sources, controlled by $\alpha_s \in [0, 1]$. Let $Q(s, s')$ denote the fraction of the weight of state s moved to state s':

$$\begin{aligned} Q(\mathrm{R}, \mathrm{R}) &= 1 - \alpha_o, \qquad Q(\mathrm{R}, \mathrm{T}_k) = \alpha_o / m, \\ Q(\mathrm{T}_j, \mathrm{R}) &= \alpha_o, \qquad Q(\mathrm{T}_j, \mathrm{T}_k) = (1 - \alpha_o)\big[(1 - \alpha_s)\, \mathbf{1}\{j = k\} + \alpha_s / m\big], \end{aligned} \tag{8}$$

for $j, k = 1, \ldots, m$. The updated weights are

$$w_{s',t+1} := \sum_{s \in \mathcal{S}} v_{s,t}\, Q(s, s'), \quad s' \in \mathcal{S}. \tag{9}$$

When $m = 1$, this update reduces to (5).

Example 1. Consider $m = 2$ sources, and take $q_1 = 1/2$, $\alpha_o = 0.1$ and $\alpha_s = 0.2$. At $t = 1$, the initial weights are $w_{\mathrm{R},1} = 0.5$, $w_{\mathrm{T}_1,1} = 0.25$, and $w_{\mathrm{T}_2,1} = 0.25$. Since $p_{\mathrm{R},1} \equiv 1/2$, the target-only predictor has loss $\ell_{\mathrm{R},1} = \log 2$. Suppose that the Transfer predictors have losses $\ell_{\mathrm{T}_1,1} = -\log 0.7$

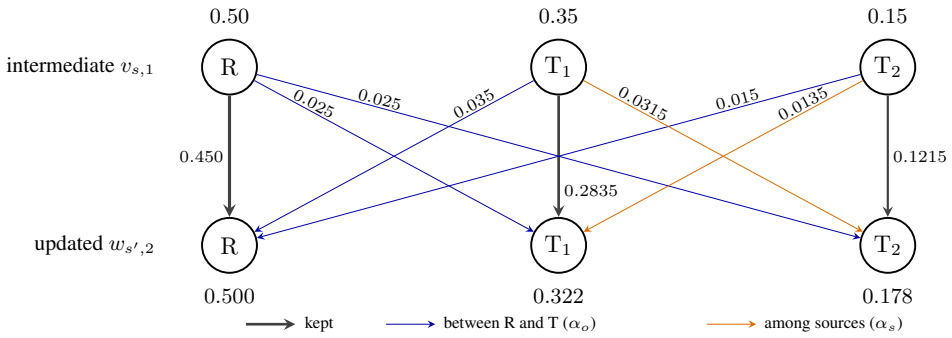


Figure 2: Transition step in Example 1. Each arrow is labeled by the amount $v_{s,1}\, Q(s, s')$ moved from state s to state s', and each updated weight is the sum of its incoming amounts.

Algorithm 1 Adaptive Multi-Source Transfer Learning (AMTL)

Require: initial transfer mass $q_1 \in (0, 1)$, transition rates $\alpha_o \in [0, 1/2]$ and $\alpha_s \in [0, 1]$, weighting constant $\gamma \in [0, 1)$, regularization parameters, clipping level $\epsilon \in (0, 1/2)$

Set $w_{\mathrm{R},1} = 1 - q_1$ and $w_{\mathrm{T}_j,1} = q_1/m$ for $j = 1, \ldots, m$
for $t = 1, \ldots, T$ **do**
  Apply any available source-snapshot replacements; retain other snapshots
  if $t = 1$ **then**
    Set $p_{\mathrm{R},1} \equiv 1/2$ and fit each Transfer model j from $D_{j,1}$
  else
    Fit the target-only model from $D_{0,1:t-1}$ and Transfer model j from $D_{0,1:t-1}$ and $D_{j,t}$ using (1)–(2)
  end if
  Form each fitted model's clipped predictor by (3)
  Observe X_t and output $p_t(x_{t,i})$ for $i = 1, \ldots, b_t$ by (6)
  Observe Y_t and compute the losses $\ell_{s,t}$, $s \in \mathcal{S}$, by (4)
  Compute the intermediate weights $v_{s,t}$ by (7)
  Update the weights $w_{s,t+1}$ by (9)
  Add $D_{0,t}$ to the target history
end for

and $\ell_{\mathrm{T}_2,1} = -\log 0.3$, so that source 1 is useful while source 2 performs worse than the constant prediction $1/2$.

In the reweighting step, $w_{s,1} \exp(-\ell_{s,1})$ are $0.5 \times 0.5 = 0.25$, $0.25 \times 0.7 = 0.175$ and $0.25 \times 0.3 = 0.075$, which sum to 0.5. By (7), the intermediate weights are

$$v_{\mathrm{R},1} = 0.5, \; v_{\mathrm{T}_1,1} = 0.35, \; v_{\mathrm{T}_2,1} = 0.15.$$

In the transition step, each state distributes its intermediate weight according to (8), as shown in Figure 2. After the update, source 1 gains weight and source 2 loses weight. The transfer mass stays at $q_2 = 0.322 + 0.178 = 0.5$, so the weight gained by source 1 comes entirely from source 2. Source 2 still retains a weight of 0.178 and can regain influence if it becomes useful later.

Algorithm 1 gives the complete online procedure. We instantiate the state predictors with sparse logistic regression, while the aggregation step depends only on predicted probabilities and losses and therefore applies to other probabilistic classifiers as well. Figure 1 summarizes the prediction, feedback, and transition steps.

4 THEORY

4.1 STRUCTURED PATH BOUND

For a state path $s_{1:T} = (s_1, \dots, s_T)$ with $s_t \in \mathcal{S}$, define its prior mass and loss $\Pi(s_{1:T}) = w_{s_1,1} \prod_{t=1}^{T-1} Q(s_t, s_{t+1})$, $\mathcal{L}(s_{1:T}) = \sum_{t=1}^{T} \ell_{s_t,t}$.

Theorem 1 (Mixable structured path bound)**.** *For any state predictors formed before the current target labels and any path with $\Pi(s_{1:T}) > 0$, the updates (7)–(9) satisfy, pathwise,*

$$\sum_{t=1}^{T} \ell_t(p_t) \le \mathcal{L}(s_{1:T}) - \log \Pi(s_{1:T}).$$

The pathwise bound applies to adaptive target streams and changing source snapshots. For the constant target-only path, $\Pi(\mathrm{R}, \dots, \mathrm{R}) = (1-q_1)(1-\alpha_o)^{T-1}$, hence

$$\sum_{t=1}^{T} \{\ell_t(p_t) - \ell_{\mathrm{R},t}\} \le -\log(1-q_1) + (T-1)\log\frac{1}{1-\alpha_o} \le \log 2 + 1,$$

where the last inequality uses $q_1 = 1/2$, $\alpha_o = 1/T$, and $T \ge 2$. Let K_o count $\mathrm{R} \leftrightarrow \mathrm{T}_j$ transitions and K_s count $\mathrm{T}_j \to \mathrm{T}_k$ transitions with $j \ne k$. Proposition 1 in Appendix A decomposes $-\log \Pi(s_{1:T})$ by K_o and K_s, including a $\log m$ cost for each $\mathrm{R} \to \mathrm{T}_j$ transition.

4.2 TRANSFER-MODEL ESTIMATION AND TARGET RISK

For statistical bounds, target observations are i.i.d. from P_0, source laws are exogenous, and new snapshots are conditionally independent. The full conditions are in Appendix A.1. Define

$$L_0(\theta) = \mathbb{E}_{P_0}\ell(\theta; Z),\ L_{j,t}(\theta) = \mathbb{E}_{P_{j,t}}\ell(\theta; Z),\ L_{\gamma,j,t}(\theta) = \gamma L_0(\theta) + (1-\gamma)L_{j,t}(\theta).$$

Let θ_0^* and $\theta_{\gamma,j,t}^*$ be finite minimizers of L_0 and $L_{\gamma,j,t}$. The population transfer bias and effective sample size are

$$\tau_{\gamma,j,t} = \|\theta_{\gamma,j,t}^* - \theta_0^*\|_2, \qquad n_{\mathrm{eff},j,t} = \left(\frac{\gamma^2}{N_t} + \frac{(1-\gamma)^2}{n_{j,t}}\right)^{-1} \quad (N_t \ge 1).$$

If $\theta_{j,t}^* \in \arg\min_\theta L_{j,t}(\theta)$, the integrated curvature conditions of Lemma 1 in Appendix A give

$$\tau_{\gamma,j,t} \le (1-\gamma)\frac{M}{\mu}\|\theta_{j,t}^* - \theta_0^*\|_2.$$

Assumption 1 in Appendix A.1 defines $s_{\gamma,j,t}$ (support size), ν (score scale), κ (restricted curvature), and M_0 (target smoothness).

Theorem 2 (Transfer-model estimation and target risk)**.** *Fix $\delta \in (0,1)$ and consider rounds with $N_t \ge 1$. For the exact minimizer $\bar{\theta}_{j,t}$ in Section 3.1, suppose $\|\widehat{\theta}_{j,t} - \bar{\theta}_{j,t}\|_2 \le \eta_{\mathrm{opt},j,t}$. Under Assumption 1, take $\lambda_{j,t} = c_\lambda\sqrt{\log(4dmT/\delta)/n_{\mathrm{eff},j,t}}$ with $c_\lambda \ge 2\sqrt{2}\,\nu$, and assume $3\lambda_{j,t}\sqrt{s_{\gamma,j,t}}/\kappa < r$ for every such j, t. Write*

$$e_{j,t} := \frac{3c_\lambda}{\kappa}\sqrt{\frac{s_{\gamma,j,t}\log(4dmT/\delta)}{n_{\mathrm{eff},j,t}}} + \tau_{\gamma,j,t} + \eta_{\mathrm{opt},j,t}.$$

With probability at least $1 - \delta/2 - \delta_{\mathrm{RSC}}$, simultaneously over all such j, t,

$$\|\widehat{\theta}_{j,t} - \theta_0^*\|_2 \le e_{j,t}, \tag{10}$$

$$L_0(\widehat{\theta}_{j,t}) - L_0(\theta_0^*) \le \frac{M_0}{2}e_{j,t}^2. \tag{11}$$

In $e_{j,t}$, the three terms are estimation error, population transfer bias, and numerical error. At fixed γ, a larger $n_{j,t}$ reduces the estimation term. Set $E_{j,t} = M_0 e_{j,t}^2/2$ and $\xi_\epsilon = -\log(1-\epsilon)$. Lemma 2 in Appendix A defines the target-only bound $E_{R,t}$.

Corollary 1 (End-to-end prequential statistical bound)**.** *Suppose the conditions of Theorem 2 and Lemma 2 hold for every $t \geq 2$ and every j. Set $E_{R,1} = E_{j,1} = B$. With probability at least $1 - 2\delta - \delta_{\mathrm{RSC}}$,*

$$\sum_{t=1}^{T} \ell_t(p_t) - TL_0(\theta_0^*) \leq \sum_{t=1}^{T} w_{\mathrm{R},t} E_{R,t} + \sum_{t=1}^{T} \sum_{j=1}^{m} w_{\mathrm{T}_j,t} E_{j,t} + (T-1)\xi_\epsilon + B\sqrt{\frac{T}{2} \log \frac{2}{\delta}}.$$

Since the state weights sum to one, the weighted state-risk term equals

$$\sum_{t=1}^{T} E_{R,t} + \sum_{t=1}^{T} \sum_{j=1}^{m} w_{\mathrm{T}_j,t}(E_{j,t} - E_{R,t}).$$

A Transfer state with $E_{j,t} < E_{R,t}$ contributes a negative correction, while one with $E_{j,t} > E_{R,t}$ contributes a positive correction, each scaled by its online weight.

The predictable-path bound is Corollary 2 in Appendix A.

5 Experiments

5.1 Sources with Varying Reliability

The synthetic experiments examine whether AMTL improves online prediction with four sources. We generate a synthetic binary classification task with one target stream and four labeled source datasets. In every domain, covariates follow the same correlated Gaussian distribution, while binary labels are sampled from Bernoulli distributions with probabilities given by domain-specific logistic models. The target contains 400 examples, and each source contains 1,000 labeled examples.

We consider five settings with fixed sources and three in which the source data change over time. The fixed settings are **Near** (four similar sources), **Mixed** (two similar and two distant), **Far** (four distant), **Opposite** (four opposing sources), and **Scarce** (the Near setting with only 100 target examples). In the changing settings, all four sources turn from similar to opposing before round 11 (**Flip**); the source closest to the target changes from source 1 to source 2 and then source 3 (**Rotate**); or both the useful source and whether any source is useful change over time (**Switch**). Detailed configurations are given in Appendix B.

To examine whether the source data help, we first compare AMTL with **Target**, which uses only target labels revealed in earlier batches. Table 1 reports AMTL minus Target for cumulative online log loss, the largest cumulative loss difference over consecutive batches, and test log loss across 50 independent simulations. Cumulative online loss measures predictive performance throughout the stream, whereas test log loss evaluates the final predictor on an independent held-out target set. Negative values mean AMTL has lower loss. Across the fixed-source settings, AMTL's online advantage over Target decreases as the sources become less relevant and turns into a loss in Opposite. Notably, in Flip, AMTL's cumulative online log loss is 2.232 lower than Target's over rounds 1-10, when the sources are close to the target. After the reversal, AMTL incurs 1.559 ± 0.038 more cumulative loss than Target over rounds 11-20.

Table 1: Log loss differences between AMTL and Target. Values are means, with 95% confidence intervals shown as $\pm$ half-widths.

Setting	Cumulative log loss ↓	Max. cumulative log loss ↓	Test log loss ↓
Near	-2.817 ± 0.091	0.030 ± 0.009	-0.0318 ± 0.0019
Mixed	-2.133 ± 0.085	0.039 ± 0.007	-0.0125 ± 0.0018
Far	-0.664 ± 0.069	0.338 ± 0.031	0.0151 ± 0.0012
Opposite	1.476 ± 0.028	1.477 ± 0.028	0.0154 ± 0.0003
Flip	-0.673 ± 0.070	1.561 ± 0.038	0.0157 ± 0.0003
Rotate	-1.295 ± 0.114	0.586 ± 0.043	-0.0781 ± 0.0075
Scarce	-3.346 ± 0.116	0.015 ± 0.010	-0.1666 ± 0.0137
Switch	-0.454 ± 0.135	1.346 ± 0.103	0.0070 ± 0.0010

5.2 Comparison with Baselines

We compare AMTL with seven baselines. **Target** trains only on previously observed target data, whereas **Pool** trains one logistic model using both target and source data. **Equal-WL** and **Dist-WL** (Liu & Lin, 2024) also train one model, but differ in how they weight the domains: Equal-WL assigns equal weight to each domain, while Dist-WL gives more weight to sources closer to the target. The other three baselines retain multiple predictors. **Credit-TL** uses one predictor for each source, **SafePool** uses Target and Pool, and **Flat-FS** (Herbster & Warmuth, 1998) uses Target and one predictor for each source. Each of these methods forms its final prediction by weighting the probabilities from its predictors, with weights updated after target labels are observed. Further details on the baselines and their implementation are given in Appendix B.

Table 2 summarizes online performance across the eight settings. Each method is ranked by cumulative online loss in each setting. The table reports each method's average rank, mean batch log loss, and worst excess (its largest cumulative loss difference from Target across settings). AMTL has the best average rank and the lowest mean batch loss.

Table 2: Aggregate online performance across eight settings.

Method	Average rank ↓	Mean batch log loss ↓	Worst excess ↓
Target	5.25	0.523	0.000
Pool	6.75	0.750	23.272
Equal-WL (Liu & Lin, 2024)	6.00	0.636	13.674
Dist-WL (Liu & Lin, 2024)	5.13	0.567	9.206
Credit-TL	2.88	0.523	10.621
SafePool	5.00	0.479	1.924
Flat-FS (Herbster & Warmuth, 1998)	3.25	0.449	**1.340**
AMTL (ours)	**1.75**	**0.436**	1.476

5.3 Effect of Separate Transition Rates

To assess the role of separate transition rates, we compare AMTL with AMTL-Flat, a simplified version that shifts weight among all predictors with one common rate. AMTL uses one rate to move weight between Target and the group of source models, and another to move weight from one source model to another. The two methods use the same fitted predictors, initial weights, and loss updates. Table 3 compares the cumulative online log loss of AMTL and AMTL-Flat. AMTL has lower loss in five of the eight settings, with the largest gain in Switch. These results suggest that separate rates help when sources are unreliable or change in usefulness, although AMTL-Flat performs slightly better in Near, Mixed, and Scarce.

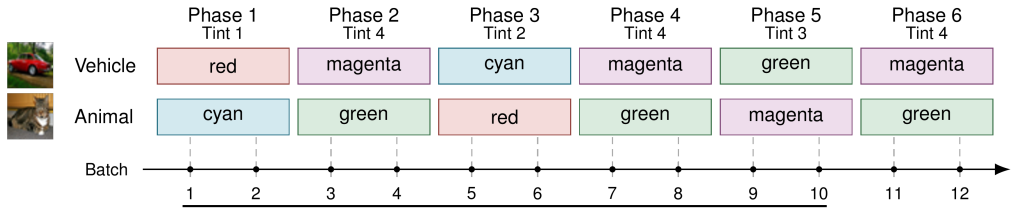
Table 3: Cumulative online log loss of AMTL minus AMTL-Flat across eight synthetic settings. Values are means, with 95% confidence intervals shown as ± half-widths.

Setting	**Near**	**Mixed**	**Far**	**Opposite**
Cumulative log loss (AMTL–AMTL-Flat) ↓	0.062 ± 0.004	0.122 ± 0.008	-0.213 ± 0.012	-0.482 ± 0.009
Setting	**Flip**	**Rotate**	**Scarce**	**Switch**
Cumulative log loss (AMTL–AMTL-Flat) ↓	-0.656 ± 0.007	-0.164 ± 0.007	0.044 ± 0.002	-1.222 ± 0.050

5.4 Image Classification with Changing Colors

We next evaluate AMTL on an online CIFAR-10 image-classification task, where the ten classes are grouped into animals and vehicles and class-dependent color tints induce controlled distribution shifts across phases, while the true animal-versus-vehicle labels remain unchanged. Each target batch contains 10 images, providing limited feedback after each rule change. Fixed source datasets are available for three of the four tint rules. For each represented rule, one source has 300 correctly labeled images and four have 50 images each with reversed labels. Rule 4 has no matching source. The target

contains 120 images in 12 batches, two per phase, with rule sequence $1 \to 4 \to 2 \to 4 \to 3 \to 4$. Target and AMTL predict each batch before its labels are revealed and are evaluated by batch-mean log loss. Figure 3 shows the target stream and phase-wise online loss. Across the full stream, AMTL has lower cumulative online log loss than Target (6.195 versus 7.705), although Target performs better in the first two phases. Construction details are in Appendix C.



Method	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Target	**0.809**	**0.247**	3.491	0.277	2.367	0.514
AMTL	0.852	0.271	**2.159**	**0.269**	**2.165**	**0.479**

Figure 3: CIFAR-10 classification under changing color rules. Each phase reports the sum of its two batch-mean online log losses, averaged over 30 repetitions.

5.5 Real data: Multi-hospital heart disease records

In this section, we evaluate AMTL on the UCI Heart Disease collection, which contains records from Cleveland, the Hungarian Institute of Cardiology, Switzerland, and VA Long Beach (Janosi et al., 1989; Detrano et al., 1989). The label indicates whether heart disease is present. Cleveland provides the target stream, while the other three hospitals provide fixed source datasets. Across 30 paired stratified splits, 180 Cleveland records arrive in 12 batches and 123 records are reserved for testing. We use ten variables shared across hospitals. Figure 4 shows online batch log loss and AMTL's source weights. AMTL and AMTL-Flat have lower cumulative online loss than all seven non-ablation baselines. More details are given in Appendix D.

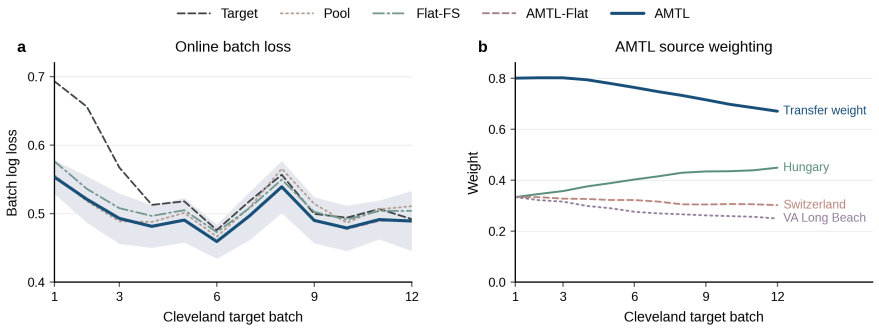


Figure 4: Multi-hospital heart disease experiment across 30 stratified splits. Left: mean online batch log loss; shading shows 95% confidence intervals for AMTL and AMTL-Flat. Right: AMTL's total transfer weight and each source's share of that weight.

6 Conclusion

AMTL adaptively combines target-only and source-specific transfer information under shifting source reliability. Its theory gives pathwise and finite-sample guarantees, while experiments demonstrate robust online performance across diverse settings. The current analysis assumes a fixed i.i.d. target law and exogenous source changes; extending AMTL to nonstationary targets and richer learned representations is an important direction for future work.

Reproducibility Statement

The main text describes the method and evaluation protocol. The appendix provides proofs, experimental settings, data construction, and implementation details.

Use of AI Tools

Generative AI was used for auxiliary checking and language editing. All research ideas, analyses, experiments, and conclusions were developed and verified by the authors.

References

Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine Learning*, 79(1–2):151–175, 2010. doi: 10.1007/s10994-009-5152-4.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Robert Detrano, Andras Janosi, Walter Steinbrunn, Matthias Pfisterer, Johann-Jakob Schmid, Sarbjit Sandhu, Kern H. Guppy, Stella Lee, and Victor Froelicher. International application of a new probability algorithm for the diagnosis of coronary artery disease. *The American Journal of Cardiology*, 64(5):304–310, 1989. doi: 10.1016/0002-9149(89)90524-9.

Honghui Du, Leandro L. Minku, and Huiyu Zhou. Multi-source transfer learning for non-stationary environments. In *International Joint Conference on Neural Networks (IJCNN)*, 2019.

Samuel G. Finlayson, Adarsh Subbaswamy, Karandeep Singh, et al. The clinician and dataset shift in artificial intelligence. *New England Journal of Medicine*, 385(3):283–286, 2021.

João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4):1–37, 2014.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016. doi: 10.1109/CVPR.2016.90.

Mark Herbster and Manfred K. Warmuth. Tracking the best expert. *Machine Learning*, 32:151–178, 1998. doi: 10.1023/A:1007424614876.

Judy Hoffman, Mehryar Mohri, and Ningshan Zhang. Algorithms and theory for multiple-source adaptation. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Andras Janosi, William Steinbrunn, Matthias Pfisterer, and Robert Detrano. Heart disease. UCI Machine Learning Repository, 1989.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.

Sai Li, T. Tony Cai, and Hongzhe Li. Transfer learning for high-dimensional linear regression: Prediction, estimation and minimax optimality. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84(1):149–173, 2022. doi: 10.1111/rssb.12479.

Sai Li, Linjun Zhang, T. Tony Cai, and Hongzhe Li. Estimation and inference for high-dimensional generalized linear models with knowledge transfer. *Journal of the American Statistical Association*, 119(546), 2024. doi: 10.1080/01621459.2023.2184373.

Zhaolei Liu and Lu Lin. Weighted likelihood transfer learning for high-dimensional generalized linear models. *Statistics*, 2024. doi: 10.1080/02331888.2024.2361861.

Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation with multiple sources. In *Advances in Neural Information Processing Systems*, volume 21, 2008.

Yishay Mansour, Mehryar Mohri, Jae Ro, Ananda Theertha Suresh, and Ke Wu. A theory of multiple-source adaptation with limited target labeled data. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2332–2340, 2021.

Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010. doi: 10.1109/TKDE.2009.191.

Ye Tian and Yang Feng. Transfer learning under high-dimensional generalized linear models. *Journal of the American Statistical Association*, 118(544):2684–2697, 2023. doi: 10.1080/01621459.2022.2071278.

Qingyao Wu, Hanrui Wu, Xiaoming Zhou, Mingkui Tan, Yonghui Xu, Yuguang Yan, and Tianyong Hao. Online transfer learning with multiple homogeneous or heterogeneous sources. *IEEE Transactions on Knowledge and Data Engineering*, 29(7):1494–1507, 2017.

Xuetong Wu, Jonathan H. Manton, Uwe Aickelin, and Jingge Zhu. Online transfer learning: Negative transfer and effect of prior knowledge. In *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021. doi: 10.1109/ISIT45174.2021.9518065.

Xuetong Wu, Jonathan H. Manton, Uwe Aickelin, and Jingge Zhu. A bayesian approach to (online) transfer learning: Theory and algorithms. *Artificial Intelligence*, 324:103991, 2023. doi: 10.1016/j.artint.2023.103991.

Renchunzi Xie and Mahardhika Pratama. Automatic online multi-source domain adaptation. *Information Sciences*, 582:480–494, 2022. doi: 10.1016/j.ins.2021.09.031.

Wen Zhang, Lingfei Deng, Lei Zhang, and Dongrui Wu. A survey on negative transfer. *IEEE/CAA Journal of Automatica Sinica*, 10(2):305–329, 2023. doi: 10.1109/JAS.2022.106004.

Peilin Zhao, Steven C. H. Hoi, Jialei Wang, and Bin Li. Online transfer learning. *Artificial Intelligence*, 216:76–102, 2014.

A Technical Conditions and Proofs

A.1 Statistical sampling model and regularity conditions

Target batch sizes and source snapshot sizes are fixed in advance. The source-law sequence $\{P_{j,t} : j \le m, t \le T\}$ is independent of the target data. Conditional on these laws, the collection of source snapshots is independent of the target stream, with

$$Z_{0,1}, \dots, Z_{0,\sum_{t=1}^{T} b_t} \overset{\text{i.i.d.}}{\sim} P_0, \qquad Z_{j,t,1}, \dots, Z_{j,t,n_{j,t}} \overset{\text{i.i.d.}}{\sim} P_{j,t}.$$

A snapshot and its law persist until replacement. All population minimizers below are finite; choose one when nonunique.

For a differentiable function f, define

$$\mathcal{D}_f(\theta, h) = f(\theta + h) - f(\theta) - \langle \nabla f(\theta), h \rangle, \qquad \widehat{L}_{\gamma,j,t} = \gamma \widehat{L}_{0,t-1} + (1-\gamma)\widehat{L}_{j,t}.$$

For a matrix A, $\|A\|_{\text{op}}$ is its spectral norm; for symmetric A, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ are its extreme eigenvalues.

Assumption 1 (Sparse logistic regularity)**.** There are constants $\nu, \kappa, \kappa_0, r, r_0, M_0 > 0$ and $\delta_{\text{RSC}} \in [0, 1)$ such that the following conditions hold for every j and every round with $N_t \ge 1$.

(A1) **Score tails.** Let $Z_0 \sim P_0$, $Z_{j,t} \sim P_{j,t}$, and $\psi(\theta; Z) = \nabla_\theta \ell(\theta; Z)$. Conditional on the source-law sequence, for every coordinate k, the following centered scores are ν^2-sub-Gaussian:

$$[\psi(\theta; Z_0) - \mathbb{E}_{P_0}\psi(\theta; Z_0)]_k, \quad \theta \in \{\theta_0^*, \theta_{\gamma,j,t}^*\},$$
$$[\psi(\theta_{\gamma,j,t}^*; Z_{j,t}) - \mathbb{E}_{P_{j,t}}\psi(\theta_{\gamma,j,t}^*; Z_{j,t})]_k.$$

(A2) **Support sets.** Write $\text{supp}(\theta) = \{k : \theta_k \ne 0\}$. For $S \subseteq \{1, \dots, d\}$, let h_S be the restriction of h to S and S^c its complement. Define $\mathcal{C}(S) = \{h : \|h_{S^c}\|_1 \le 3\|h_S\|_1\}$. Set $S_{j,t} = \text{supp}(\theta_{\gamma,j,t}^*)$, $s_{\gamma,j,t} = |S_{j,t}|$, $S_0 = \text{supp}(\theta_0^*)$, and $s_0 = |S_0|$. Assume $\max\{s_0, \max_{j,t} s_{\gamma,j,t}\} < d$.

(A3) **Local restricted curvature.** There is an event $\mathcal{E}_{\text{RSC}}$ with $\Pr(\mathcal{E}_{\text{RSC}}) \ge 1 - \delta_{\text{RSC}}$ on which, simultaneously over j, t,

$$\mathcal{D}_{\widehat{L}_{\gamma,j,t}}(\theta_{\gamma,j,t}^*, h) \ge \frac{\kappa}{2}\|h\|_2^2 \quad (h \in \mathcal{C}(S_{j,t}),\ \|h\|_2 \le r),$$
$$\mathcal{D}_{\widehat{L}_{0,t-1}}(\theta_0^*, h) \ge \frac{\kappa_0}{2}\|h\|_2^2 \quad (h \in \mathcal{C}(S_0),\ \|h\|_2 \le r_0).$$

(A4) **Target smoothness.** For every $\theta \in \mathbb{R}^d$, $\|\nabla^2 L_0(\theta)\|_{\text{op}} \le M_0$. For logistic loss, this holds whenever $\lambda_{\max}\{\mathbb{E}_{P_0}(XX^\top)\} \le 4M_0$.

A.2 Proof of Theorem 1

For item i in batch t, define

$$z_{s,t,i} = p_{s,t}(x_{t,i})^{y_{t,i}}[1 - p_{s,t}(x_{t,i})]^{1-y_{t,i}}, \qquad e^{-\ell_{s,t}} = \left(\prod_{i=1}^{b_t} z_{s,t,i}\right)^{1/b_t}.$$

The geometric mean is concave on $\mathbb{R}_+^{b_t}$. Therefore

$$\begin{aligned} e^{-\ell_t(p_t)} &= \left\{\prod_{i=1}^{b_t}\left(\sum_s w_{s,t} z_{s,t,i}\right)\right\}^{1/b_t} \\ &\ge \sum_s w_{s,t}\left(\prod_{i=1}^{b_t} z_{s,t,i}\right)^{1/b_t} = \sum_s w_{s,t} e^{-\ell_{s,t}} =: Z_t. \end{aligned} \tag{12}$$

Set $\mathcal{W}_0 = 1$ and, for $t \geq 1$, define the truncated path partition

$$\mathcal{W}_t = \sum_{s_{1:t}} w_{s_1,1} \left\{ \prod_{u=1}^{t-1} Q(s_u, s_{u+1}) \right\} \exp\left(- \sum_{u=1}^{t} \ell_{s_u,u} \right).$$

The unnormalized mass entering state s at round t is

$$A_t(s) = \sum_{\substack{s_{1:t} \in \mathcal{S}^t \\ s_t = s}} w_{s_1,1} \left\{ \prod_{u=1}^{t-1} Q(s_u, s_{u+1}) \right\} \exp\left(- \sum_{u=1}^{t-1} \ell_{s_u,u} \right).$$

For $t = 1$, $A_1(s) = w_{s,1} = \mathcal{W}_0 w_{s,1}$. If $A_t(s) = \mathcal{W}_{t-1} w_{s,t}$, then

$$\mathcal{W}_t = \sum_s A_t(s) e^{-\ell_{s,t}} = \mathcal{W}_{t-1} Z_t, \tag{13}$$

$$\frac{A_{t+1}(s')}{\mathcal{W}_t} = \sum_s \frac{A_t(s) e^{-\ell_{s,t}}}{\mathcal{W}_t} Q(s, s') = \sum_s \frac{w_{s,t} e^{-\ell_{s,t}}}{Z_t} Q(s, s') = w_{s',t+1}.$$

Induction and (13) give

$$\mathcal{W}_T = \prod_{t=1}^{T} Z_t = \sum_{s_{1:T}} \Pi(s_{1:T}) e^{-\mathcal{L}(s_{1:T})}. \tag{14}$$

Combining (12) and (14), for every path with $\Pi(s_{1:T}) > 0$,

$$\begin{aligned} \sum_{t=1}^{T} \ell_t(p_t) &\leq - \sum_{t=1}^{T} \log Z_t = - \log \mathcal{W}_T \\ &\leq - \log\{\Pi(s_{1:T}) e^{-\mathcal{L}(s_{1:T})}\} = \mathcal{L}(s_{1:T}) - \log \Pi(s_{1:T}). \end{aligned}$$

□

A.3 Path-penalty decomposition

Proposition 1 (Mode and source switch complexity)**.** *Suppose $0 < \alpha_o < 1$ and $0 < \alpha_s \leq 1$. For a path with $\Pi(s_{1:T}) > 0$, let K_o count transitions between R and a Transfer state, and let K_s count transitions $\mathrm{T}_j \to \mathrm{T}_k$ with $j \neq k$. Let N_{RT} and N_{TR} count transitions from the target-only state to Transfer and from Transfer to the target-only state, respectively, and let N_{same} count $\mathrm{T}_j \to \mathrm{T}_j$ transitions. Then*

$$\begin{aligned} - \log \Pi(s_{1:T}) = & - \log w_{s_1,1} + (T - 1 - K_o) \log \frac{1}{1 - \alpha_o} + K_o \log \frac{1}{\alpha_o} \\ & + N_{RT} \log m + K_s \log \frac{m}{\alpha_s} + N_{\mathrm{same}} \log \frac{1}{1 - \alpha_s + \alpha_s/m}. \end{aligned} \tag{15}$$

Proof. Let N_{RR} count $\mathrm{R} \to \mathrm{R}$ transitions and set $K_o = N_{RT} + N_{TR}$. Since $N_{RR} + N_{\mathrm{same}} + K_s = T - 1 - K_o$, the transition probabilities in (8) give

$$\begin{aligned} - \log \Pi(s_{1:T}) = & - \log w_{s_1,1} + (N_{RR} + N_{\mathrm{same}} + K_s) \log \frac{1}{1 - \alpha_o} \\ & + (N_{RT} + N_{TR}) \log \frac{1}{\alpha_o} + N_{RT} \log m + K_s \log \frac{m}{\alpha_s} \\ & + N_{\mathrm{same}} \log \frac{1}{1 - \alpha_s + \alpha_s/m}. \end{aligned}$$

Using the two count identities gives (15). □

A.4 CURVATURE-AWARE TRANSFER BIAS

Let $\theta^*_{j,t} \in \arg\min_\theta L_{j,t}(\theta)$ and $\Delta_{j,t} = \theta^*_{j,t} - \theta^*_0$. Define

$$H^{\Delta}_{j,t} = \int_0^1 \nabla^2 L_{j,t}\big(\theta^*_{j,t} + u(\theta^*_0 - \theta^*_{j,t})\big)\, du,$$
$$G_{\gamma,j,t} = \int_0^1 \nabla^2 L_{\gamma,j,t}\big(\theta^*_0 + u(\theta^*_{\gamma,j,t} - \theta^*_0)\big)\, du.$$

Lemma 1 (Exact curvature-aware bias)**.** *If $G_{\gamma,j,t}$ is invertible, then*

$$\theta^*_{\gamma,j,t} - \theta^*_0 = (1-\gamma) G^{-1}_{\gamma,j,t} H^{\Delta}_{j,t} \Delta_{j,t},$$
$$\tau_{\gamma,j,t} = (1-\gamma)\left\| G^{-1}_{\gamma,j,t} H^{\Delta}_{j,t} \Delta_{j,t} \right\|_2.$$

If $\lambda_{\min}(G_{\gamma,j,t}) \geq \mu$ and $\|H^{\Delta}_{j,t}\|_{\mathrm{op}} \leq M$, then $\tau_{\gamma,j,t} \leq (1-\gamma)(M/\mu)\|\Delta_{j,t}\|_2$.

Proof. Since $\nabla L_{j,t}(\theta^*_{j,t}) = 0$,

$$\begin{aligned}\nabla L_{j,t}(\theta^*_0) &= \int_0^1 \nabla^2 L_{j,t}\big(\theta^*_{j,t} + u(\theta^*_0 - \theta^*_{j,t})\big)(\theta^*_0 - \theta^*_{j,t})\, du \\ &= -H^{\Delta}_{j,t}\Delta_{j,t}.\end{aligned}$$

Since $\nabla L_0(\theta^*_0) = 0$,

$$\nabla L_{\gamma,j,t}(\theta^*_0) = -(1-\gamma) H^{\Delta}_{j,t}\Delta_{j,t}.$$

The second integral identity is

$$0 = \nabla L_{\gamma,j,t}(\theta^*_{\gamma,j,t}) = \nabla L_{\gamma,j,t}(\theta^*_0) + G_{\gamma,j,t}(\theta^*_{\gamma,j,t} - \theta^*_0).$$

so

$$\theta^*_{\gamma,j,t} - \theta^*_0 = (1-\gamma) G^{-1}_{\gamma,j,t} H^{\Delta}_{j,t}\Delta_{j,t}.$$

Under the stated spectral conditions,

$$\tau_{\gamma,j,t} \leq (1-\gamma)\|G^{-1}_{\gamma,j,t}\|_{\mathrm{op}}\|H^{\Delta}_{j,t}\|_{\mathrm{op}}\|\Delta_{j,t}\|_2 \leq (1-\gamma)\frac{M}{\mu}\|\Delta_{j,t}\|_2.$$

□

A.5 PROOF OF THEOREM 2

Let

$$\mathcal{G}_{\mathrm{src}} = \sigma(P_{j,t} : j \leq m, t \leq T), \qquad \theta^* = \theta^*_{\gamma,j,t}, \qquad \widehat{L} = \gamma \widehat{L}_{0,t-1} + (1-\gamma)\widehat{L}_{j,t}.$$

Condition on $\mathcal{G}_{\mathrm{src}}$. Write the target-prefix observations as $Z_{0,1}, \ldots, Z_{0,N_t}$ and the current source observations as $Z_{j,t,1}, \ldots, Z_{j,t,n_{j,t}}$. Since $\nabla L_{\gamma,j,t}(\theta^*) = 0$,

$$\begin{aligned}g_{j,t} := \nabla \widehat{L}(\theta^*) = \frac{\gamma}{N_t}\sum_{i=1}^{N_t}\{\psi(\theta^*; Z_{0,i}) - \mathbb{E}_{P_0}\psi(\theta^*; Z)\} \\ + \frac{1-\gamma}{n_{j,t}}\sum_{i=1}^{n_{j,t}}\{\psi(\theta^*; Z_{j,t,i}) - \mathbb{E}_{P_{j,t}}\psi(\theta^*; Z)\}.\end{aligned}$$

For every coordinate k and $u \in \mathbb{R}$, Assumption 1(A1) and conditional independence imply

$$\log \mathbb{E}[\exp\{u[g_{j,t}]_k\} \mid \mathcal{G}_{\mathrm{src}}] \leq \frac{\nu^2 u^2}{2}\left\{\frac{\gamma^2}{N_t} + \frac{(1-\gamma)^2}{n_{j,t}}\right\} = \frac{\nu^2 u^2}{2 n_{\mathrm{eff},j,t}}.$$

Consequently,

$$\Pr\left(|[g_{j,t}]_k| > \nu\sqrt{\frac{2\log(4dmT/\delta)}{n_{\mathrm{eff},j,t}}} \,\middle|\, \mathcal{G}_{\mathrm{src}}\right) \leq \frac{\delta}{2dmT}.$$

Define

$$\mathcal{E}_{\mathrm{src}} = \left\{ \max_{j,t:N_t \geq 1} \frac{\|g_{j,t}\|_\infty}{\sqrt{\log(4dmT/\delta)/n_{\mathrm{eff},j,t}}} \leq \sqrt{2}\,\nu \right\}.$$

A union bound, followed by removing the conditioning, gives

$$\Pr(\mathcal{E}_{\mathrm{src}} \mid \mathcal{G}_{\mathrm{src}}) \geq 1 - \frac{\delta}{2}, \qquad \Pr(\mathcal{E}_{\mathrm{src}}) = \mathbb{E}\{\Pr(\mathcal{E}_{\mathrm{src}} \mid \mathcal{G}_{\mathrm{src}})\} \geq 1 - \frac{\delta}{2}.$$

The choice $c_\lambda \geq 2\sqrt{2}\,\nu$ ensures, on $\mathcal{E}_{\mathrm{src}}$,

$$\|g_{j,t}\|_\infty \leq \frac{\lambda_{j,t}}{2} \qquad \text{for every } j,t \text{ with } N_t \geq 1. \tag{16}$$

Fix j,t with $N_t \geq 1$ and set $h = \bar{\theta}_{j,t} - \theta^*$ and $S = S_{j,t}$. Optimality of $\bar{\theta}_{j,t}$ gives

$$\mathcal{D}_{\widehat{L}}(\theta^*, h) + \langle g_{j,t}, h\rangle + \lambda_{j,t}\{\|\theta^* + h\|_1 - \|\theta^*\|_1\} \leq 0. \tag{17}$$

On (16),

$$-\langle g_{j,t}, h\rangle \leq \frac{\lambda_{j,t}}{2}(\|h_S\|_1 + \|h_{S^c}\|_1), \qquad \|\theta^*\|_1 - \|\theta^* + h\|_1 \leq \|h_S\|_1 - \|h_{S^c}\|_1.$$

Substitution into (17), together with convexity of $\widehat{L}$, gives

$$0 \leq \mathcal{D}_{\widehat{L}}(\theta^*, h) \leq \frac{3\lambda_{j,t}}{2}\|h_S\|_1 - \frac{\lambda_{j,t}}{2}\|h_{S^c}\|_1, \tag{18}$$

hence

$$\|h_{S^c}\|_1 \leq 3\|h_S\|_1, \qquad h \in \mathcal{C}(S_{j,t}). \tag{19}$$

The local RSC condition can be invoked directly when $\|h\|_2 \leq r$. If $\|h\|_2 > r$, let

$$a = \frac{r}{\|h\|_2} \in (0,1), \qquad \widetilde{h} = ah, \qquad F(\theta) = \widehat{L}(\theta) + \lambda_{j,t}\|\theta\|_1.$$

Convexity and $F(\theta^* + h) \leq F(\theta^*)$ imply

$$\begin{aligned} F(\theta^* + \widetilde{h}) &= F\{(1-a)\theta^* + a(\theta^* + h)\} \\ &\leq (1-a)F(\theta^*) + aF(\theta^* + h) \leq F(\theta^*). \end{aligned}$$

Thus $\widetilde{h}$ satisfies (18)–(19). For either $v = h$ when $\|h\|_2 \leq r$, or $v = \widetilde{h}$ otherwise,

$$\frac{\kappa}{2}\|v\|_2^2 \leq \mathcal{D}_{\widehat{L}}(\theta^*, v) \leq \frac{3\lambda_{j,t}}{2}\|v_S\|_1 \leq \frac{3\lambda_{j,t}}{2}\sqrt{s_{\gamma,j,t}}\|v\|_2,$$

and therefore

$$\|v\|_2 \leq \frac{3\lambda_{j,t}}{\kappa}\sqrt{s_{\gamma,j,t}}. \tag{20}$$

The estimation-radius condition in Theorem 2 makes the right side of (20) smaller than r. Hence the choice $v = \widetilde{h}$ would give $r = \|\widetilde{h}\|_2 < r$, a contradiction. Thus

$$\|\bar{\theta}_{j,t} - \theta^*_{\gamma,j,t}\|_2 \leq \frac{3c_\lambda}{\kappa}\sqrt{\frac{s_{\gamma,j,t}\log(4dmT/\delta)}{n_{\mathrm{eff},j,t}}}.$$

By the definition of $\tau_{\gamma,j,t}$ and the numerical-error condition,

$$\begin{aligned} \|\widehat{\theta}_{j,t} - \theta^*_0\|_2 &\leq \|\widehat{\theta}_{j,t} - \bar{\theta}_{j,t}\|_2 + \|\bar{\theta}_{j,t} - \theta^*_{\gamma,j,t}\|_2 + \|\theta^*_{\gamma,j,t} - \theta^*_0\|_2 \\ &\leq \eta_{\mathrm{opt},j,t} + \frac{3c_\lambda}{\kappa}\sqrt{\frac{s_{\gamma,j,t}\log(4dmT/\delta)}{n_{\mathrm{eff},j,t}}} + \tau_{\gamma,j,t}. \end{aligned} \tag{21}$$

This is (10). Let $d_{j,t} = \widehat{\theta}_{j,t} - \theta_0^*$. Global target smoothness applies along $\theta_0^* + u d_{j,t}$ for every $u \in [0,1]$. Since $\nabla L_0(\theta_0^*) = 0$,

$$L_0(\widehat{\theta}_{j,t}) - L_0(\theta_0^*) = \int_0^1 (1-u) d_{j,t}^\top \nabla^2 L_0(\theta_0^* + u d_{j,t}) d_{j,t}\, du$$
$$\leq \int_0^1 (1-u) M_0 \|d_{j,t}\|_2^2\, du = \frac{M_0}{2}\|d_{j,t}\|_2^2. \tag{22}$$

Equations (21) and (22) give (11). Finally,

$$\Pr(\mathcal{E}_{\mathrm{src}} \cap \mathcal{E}_{\mathrm{RSC}}) \geq 1 - \frac{\delta}{2} - \delta_{\mathrm{RSC}}.$$

□

A.6 Target-only bound and proof of Lemma 2

Lemma 2 (Explicit target-only error). *Fix $\delta \in (0,1)$. For $N_t \geq 1$, let $\bar{\theta}_{\mathrm{R},t}$ be an exact minimizer of (1), with*

$$\lambda_{\mathrm{R},t} = c_R \sqrt{\frac{\log(4dT/\delta)}{N_t}}, \qquad c_R \geq 2\sqrt{2}\,\nu,$$

and suppose $\|\widehat{\theta}_{\mathrm{R},t} - \bar{\theta}_{\mathrm{R},t}\|_2 \leq \eta_{\mathrm{opt},R,t}$. Assume $3\lambda_{\mathrm{R},t}\sqrt{s_0}/\kappa_0 < r_0$ for every such t. Under Assumption 1, with probability at least $1 - \delta/2 - \delta_{\mathrm{RSC}}$, simultaneously over these rounds,

$$\|\widehat{\theta}_{\mathrm{R},t} - \theta_0^*\|_2 \leq \frac{3c_R}{\kappa_0}\sqrt{\frac{s_0 \log(4dT/\delta)}{N_t}} + \eta_{\mathrm{opt},R,t}, \tag{23}$$

$$L_0(\widehat{\theta}_{\mathrm{R},t}) - L_0(\theta_0^*) \leq E_{R,t} := \frac{M_0}{2}\left(\frac{3c_R}{\kappa_0}\sqrt{\frac{s_0 \log(4dT/\delta)}{N_t}} + \eta_{\mathrm{opt},R,t}\right)^2. \tag{24}$$

Proof. Let

$$g_{R,t} = \nabla \widehat{L}_{0,t-1}(\theta_0^*) = \frac{1}{N_t}\sum_{i=1}^{N_t}\{\psi(\theta_0^*; Z_{0,i}) - \mathbb{E}\psi(\theta_0^*; Z_0)\}, \qquad \mathbb{E}\psi(\theta_0^*; Z_0) = \nabla L_0(\theta_0^*) = 0.$$

Assumption 1(A1) gives, for every k, t with $N_t \geq 1$,

$$\Pr\left(|[g_{R,t}]_k| > \nu\sqrt{\frac{2\log(4dT/\delta)}{N_t}}\right) \leq \frac{\delta}{2dT}.$$

Therefore the event

$$\mathcal{E}_R = \left\{\max_{t:N_t \geq 1} \frac{\|g_{R,t}\|_\infty}{\sqrt{\log(4dT/\delta)/N_t}} \leq \sqrt{2}\,\nu\right\}$$

satisfies

$$\Pr(\mathcal{E}_R) \geq 1 - \frac{\delta}{2}, \qquad \|g_{R,t}\|_\infty \leq \frac{\lambda_{\mathrm{R},t}}{2} \quad \text{on } \mathcal{E}_R \text{ for } N_t \geq 1. \tag{25}$$

Fix t, set $h_R = \bar{\theta}_{\mathrm{R},t} - \theta_0^*$ and $S_0 = \mathrm{supp}(\theta_0^*)$. Exact optimality gives

$$0 \geq F_R(\theta_0^* + h_R) - F_R(\theta_0^*)$$
$$= \mathcal{D}_{\widehat{L}_{0,t-1}}(\theta_0^*, h_R) + \langle g_{R,t}, h_R\rangle + \lambda_{\mathrm{R},t}\{\|\theta_0^* + h_R\|_1 - \|\theta_0^*\|_1\}, \tag{26}$$

where $F_R(\theta) = \widehat{L}_{0,t-1}(\theta) + \lambda_{\mathrm{R},t}\|\theta\|_1$. Using (25) and $\|\theta_0^*\|_1 - \|\theta_0^* + h_R\|_1 \leq \|(h_R)_{S_0}\|_1 - \|(h_R)_{S_0^c}\|_1$ in (26), and using convexity of $\widehat{L}_{0,t-1}$, we obtain

$$0 \leq \mathcal{D}_{\widehat{L}_{0,t-1}}(\theta_0^*, h_R) \leq \frac{3\lambda_{\mathrm{R},t}}{2}\|(h_R)_{S_0}\|_1 - \frac{\lambda_{\mathrm{R},t}}{2}\|(h_R)_{S_0^c}\|_1,$$
$$\|(h_R)_{S_0^c}\|_1 \leq 3\|(h_R)_{S_0}\|_1.$$

For $\|h_R\|_2 \le r_0$, local RSC gives

$$\frac{\kappa_0}{2}\|h_R\|_2^2 \le \frac{3\lambda_{\mathrm{R},t}}{2}\|(h_R)_{S_0}\|_1 \le \frac{3\lambda_{\mathrm{R},t}}{2}\sqrt{s_0}\|h_R\|_2,$$

and hence

$$\|h_R\|_2 \le \frac{3c_R}{\kappa_0}\sqrt{\frac{s_0\log(4dT/\delta)}{N_t}}. \tag{27}$$

If $\|h_R\|_2 > r_0$, set $a_0 = r_0/\|h_R\|_2$ and $\widetilde{h}_R = a_0 h_R$. Convexity gives

$$\begin{aligned} F_R(\theta_0^* + \widetilde{h}_R) &= F_R\{(1-a_0)\theta_0^* + a_0(\theta_0^* + h_R)\} \\ &\le (1-a_0)F_R(\theta_0^*) + a_0 F_R(\theta_0^* + h_R) \le F_R(\theta_0^*). \end{aligned}$$

The basic inequality for $\widetilde{h}_R$ has the same form as (26); hence $\widetilde{h}_R \in \mathcal{C}(S_0)$ and

$$\begin{aligned} \frac{\kappa_0}{2}r_0^2 \le \mathcal{D}_{\widehat{L}_{0,t-1}}(\theta_0^*, \widetilde{h}_R) &\le \frac{3\lambda_{\mathrm{R},t}}{2}\|(\widetilde{h}_R)_{S_0}\|_1 \\ &\le \frac{3\lambda_{\mathrm{R},t}}{2}\sqrt{s_0}\, r_0, \end{aligned}$$

which implies $r_0 \le 3\lambda_{\mathrm{R},t}\sqrt{s_0}/\kappa_0 < r_0$. Thus (27) holds. The numerical error gives

$$\begin{aligned} \|\widehat{\theta}_{\mathrm{R},t} - \theta_0^*\|_2 &\le \|\widehat{\theta}_{\mathrm{R},t} - \bar{\theta}_{\mathrm{R},t}\|_2 + \|\bar{\theta}_{\mathrm{R},t} - \theta_0^*\|_2 \\ &\le \eta_{\mathrm{opt},R,t} + \frac{3c_R}{\kappa_0}\sqrt{\frac{s_0\log(4dT/\delta)}{N_t}}, \end{aligned}$$

which is (23). With $d_R = \widehat{\theta}_{\mathrm{R},t} - \theta_0^*$, global target smoothness gives

$$\begin{aligned} L_0(\widehat{\theta}_{\mathrm{R},t}) - L_0(\theta_0^*) &= \int_0^1 (1-u)d_R^\top \nabla^2 L_0(\theta_0^* + u d_R) d_R \, du \\ &\le \frac{M_0}{2}\|d_R\|_2^2 \le E_{R,t}. \end{aligned}$$

The event probability is

$$\Pr(\mathcal{E}_R \cap \mathcal{E}_{\mathrm{RSC}}) \ge 1 - \frac{\delta}{2} - \delta_{\mathrm{RSC}}.$$

□

A.7 Proof of Corollary 1

Let

$$\mathcal{H}_t = \sigma\big(\{D_{j,r} : 1 \le j \le m,\ 1 \le r \le t\}, D_{0,1:t-1}\big), \qquad \mathcal{R}_t(p) = \mathbb{E}[\ell_t(p) \mid \mathcal{H}_t].$$

With a fixed fitting rule, state weights and predictors are $\mathcal{H}_t$-measurable. Transfer fit j uses $D_{j,t}$ and $D_{0,1:t-1}$. Clipping a binary probability increases either label's log loss by at most $\xi_\epsilon = -\log(1-\epsilon)$. The current target batch is independent of $\mathcal{H}_t$. For each θ measurable with respect to $\mathcal{H}_t$ and $q = p_\theta(x)$,

$$\begin{aligned} -y\log \mathrm{clip}_\epsilon(q) - (1-y)\log\{1 - \mathrm{clip}_\epsilon(q)\} &\le \ell(\theta;(x,y)) + \xi_\epsilon, \\ \mathcal{R}_t(\mathrm{clip}_\epsilon \circ p_\theta) &\le L_0(\theta) + \xi_\epsilon. \end{aligned} \tag{28}$$

Equations (24), (11), and (28) give, for $t \ge 2$ on $\mathcal{E}_{\mathrm{src}} \cap \mathcal{E}_R \cap \mathcal{E}_{\mathrm{RSC}}$,

$$\mathcal{R}_t(p_{\mathrm{R},t}) - L_0(\theta_0^*) \le E_{R,t} + \xi_\epsilon, \tag{29}$$

$$\mathcal{R}_t(p_{\mathrm{T}_j,t}) - L_0(\theta_0^*) \le E_{j,t} + \xi_\epsilon. \tag{30}$$

By (6) and convexity of binary log loss,

$$\mathcal{R}_t(p_t) \le w_{\mathrm{R},t}\mathcal{R}_t(p_{\mathrm{R},t}) + \sum_{j=1}^m w_{\mathrm{T}_j,t}\mathcal{R}_t(p_{\mathrm{T}_j,t}). \tag{31}$$

Combining (29)–(31),

$$\mathcal{R}_t(p_t) - L_0(\theta_0^*) \le w_{\mathrm{R},t} E_{R,t} + \sum_{j=1}^{m} w_{\mathrm{T}_j,t} E_{j,t} + \xi_\epsilon.$$

At $t = 1$, $0 \le \mathcal{R}_1(p_1) \le B$ and $L_0(\theta_0^*) \ge 0$, so

$$\mathcal{R}_1(p_1) - L_0(\theta_0^*) \le B = w_{\mathrm{R},1} E_{R,1} + \sum_{j=1}^{m} w_{\mathrm{T}_j,1} E_{j,1}.$$

Let

$$d_t = \ell_t(p_t) - \mathcal{R}_t(p_t).$$

Then $\mathbb{E}[d_t \mid \mathcal{H}_t] = 0$, and conditionally on $\mathcal{H}_t$,

$$d_t \in [-\mathcal{R}_t(p_t), B - \mathcal{R}_t(p_t)], \qquad \{B - \mathcal{R}_t(p_t)\} - \{-\mathcal{R}_t(p_t)\} = B.$$

Hoeffding–Azuma therefore gives

$$\Pr\left(\left|\sum_{t=1}^{T} d_t\right| \ge u\right) \le 2\exp\left(-\frac{2u^2}{TB^2}\right),$$
$$\Pr\left(\sum_{t=1}^{T} d_t \le B\sqrt{\frac{T}{2}\log\frac{2}{\delta}}\right) \ge 1 - \delta. \tag{32}$$

Denote the event in the second line of (32) by $\mathcal{E}_M$. On $\mathcal{E}_{\mathrm{src}} \cap \mathcal{E}_R \cap \mathcal{E}_{\mathrm{RSC}} \cap \mathcal{E}_M$,

$$\begin{aligned}\sum_{t=1}^{T} \ell_t(p_t) - TL_0(\theta_0^*) &= \sum_{t=1}^{T} \{\mathcal{R}_t(p_t) - L_0(\theta_0^*)\} + \sum_{t=1}^{T} d_t \\ &\le \sum_{t=1}^{T} w_{\mathrm{R},t} E_{R,t} + \sum_{t=1}^{T}\sum_{j=1}^{m} w_{\mathrm{T}_j,t} E_{j,t} + (T-1)\xi_\epsilon + B\sqrt{\frac{T}{2}\log\frac{2}{\delta}}.\end{aligned}$$

Finally,

$$\Pr(\mathcal{E}_{\mathrm{src}} \cap \mathcal{E}_R \cap \mathcal{E}_{\mathrm{RSC}} \cap \mathcal{E}_M) \ge 1 - \frac{\delta}{2} - \frac{\delta}{2} - \delta_{\mathrm{RSC}} - \delta = 1 - 2\delta - \delta_{\mathrm{RSC}}.$$

□

A.8 Predictable paths and proof of Corollary 2

Corollary 2 (Statistical guarantee for a predictable structured path)**.** *Fix a state-selection rule $s_t \in \mathcal{S}$ that is $\mathcal{H}_t$-measurable at every round and satisfies $\Pi(s_{1:T}) > 0$ almost surely. Define*

$$E_{s,t} = \begin{cases} E_{R,t}, & s_t = \mathrm{R}, \\ E_{j,t}, & s_t = \mathrm{T}_j. \end{cases}$$

Under Assumption 1, suppose the estimation-radius conditions in Theorem 2 and Lemma 2 hold for every $t \ge 2$ and every j. With probability at least $1 - 2\delta - \delta_{\mathrm{RSC}}$,

$$\sum_{t=1}^{T} \ell_t(p_t) - TL_0(\theta_0^*) \le \sum_{t=1}^{T} E_{s,t} - \log\Pi(s_{1:T}) + (T-1)\xi_\epsilon + B\sqrt{\frac{T}{2}\log\frac{1}{\delta}}.$$

Proof. For the specified rule, define

$$r_{s,t} = \mathcal{R}_t(p_{s_t,t}), \qquad d_{s,t} = \ell_{s_t,t} - r_{s,t}.$$

Measurability of s_t implies

$$\mathbb{E}[d_{s,t} \mid \mathcal{H}_t] = 0, \qquad d_{s,t} \in [-r_{s,t}, B - r_{s,t}].$$

The one-sided Hoeffding–Azuma inequality gives an event $\mathcal{E}_{M,s}$ with

$$\Pr(\mathcal{E}_{M,s}) \geq 1-\delta, \qquad \sum_{t=1}^{T} d_{s,t} \leq B\sqrt{\frac{T}{2}\log\frac{1}{\delta}}. \tag{33}$$

For $t \geq 2$, (29)–(30) imply

$$\begin{aligned} r_{s,t} - L_0(\theta_0^*) &\leq E_{R,t} + \xi_\epsilon, \quad s_t = \mathrm{R}, \\ r_{s,t} - L_0(\theta_0^*) &\leq E_{j,t} + \xi_\epsilon = E_{s,t} + \xi_\epsilon, \quad s_t = \mathrm{T}_j. \end{aligned} \tag{34}$$

At $t=1$, $r_{s,1} \leq B = E_{s,1}$. Theorem 1, (33), and (34) give

$$\begin{aligned} \sum_{t=1}^{T} \ell_t(p_t) - TL_0(\theta_0^*) &\leq \sum_{t=1}^{T} \ell_{s_t,t} - TL_0(\theta_0^*) - \log\Pi(s_{1:T}) \\ &= \sum_{t=1}^{T} d_{s,t} + \sum_{t=1}^{T}\{r_{s,t} - L_0(\theta_0^*)\} - \log\Pi(s_{1:T}) \\ &\leq \sum_{t=1}^{T} E_{s,t} - \log\Pi(s_{1:T}) + (T-1)\xi_\epsilon + B\sqrt{\frac{T}{2}\log\frac{1}{\delta}}. \end{aligned}$$

The event probability is

$$\Pr(\mathcal{E}_{\mathrm{src}} \cap \mathcal{E}_R \cap \mathcal{E}_{\mathrm{RSC}} \cap \mathcal{E}_{M,s}) \geq 1 - 2\delta - \delta_{\mathrm{RSC}}.$$

□

B Experiment Settings in Section 5.1

Synthetic design. The eight settings vary source relevance, source replacements, and sample size. Table 4 summarizes their configurations.

Table 4: Source configurations in the eight synthetic settings.

Setting	Source composition and temporal structure
Near	Four relevant sources with coefficient gap 0.35 throughout.
Mixed	Two relevant sources with gap 0.35 and two distant sources with gap at least 2.75, fixed throughout.
Far	Four distant sources with gap at least 2.75 throughout.
Opposite	Four antagonistic sources with coefficients near $-1.25\theta_0^*$ from round 1.
Flip	All four sources start relevant and are replaced by antagonistic snapshots before round 11 of 20.
Rotate	In each third, one source has gap 0.35, one has gap 0.90, and two are antagonistic. The design path is $\mathrm{T}_1 \to \mathrm{T}_2 \to \mathrm{T}_3$, with $K_o = 0$ and $K_s = 2$.
Scarce	The all-relevant source structure with 100 target observations, compared with 400 in the standard stream.
Switch	Eight sources across eight phases. Transfer phases have one exact, one gap-0.35, and six antagonistic sources; target-only phases have eight antagonistic sources. The design path is $\mathrm{T}_1 \to \mathrm{T}_2 \to \mathrm{T}_3 \to \mathrm{R} \to \mathrm{T}_4 \to \mathrm{T}_5 \to \mathrm{R} \to \mathrm{T}_6$.

Experimental settings. The standard synthetic setting uses clipping 0.02, $d = 160$, target sparsity 8, four sources with 1,000 samples each, 400 target samples in batches of 20, and 3,000 held-out target samples. Scarce uses 100 targets and Rotate uses 120, both in batches of 10. Covariates are Gaussian with Toeplitz correlation 0.2. For every domain with coefficient θ, labels are generated independently as $Y \mid X \sim \mathrm{Bernoulli}\{\sigma(X^\top\theta)\}$. Relevant coefficients have distance 0.35 from the target. Irrelevant coefficients have distance at least 2.75. Their overlapping active coordinates, when present, have opposite signs. Harmful coefficients are near $-1.25\theta_0^*$. Rotate uses source gaps 0.35 and 0.90 plus two adversarial sources in each third. Switch data use eight sources, 160 targets in batches of 5, 1,500 observations per source, and eight phases; transfer phases contain one exact, one gap-0.35, and six adversarial sources, while all sources are adversarial in target-only phases.

Method configurations. The final settings for all methods are given in Table 5. AMTL uses $(\gamma, \alpha_o, \alpha_s, q_1) = (0.05, 0.025, 0.10, 0.8)$ and Transfer-model penalty scale $c_\lambda = 0.10$. AMTL-Flat uses the same fitted models and initial weights, with flat share rate $\alpha = 0.08$.

Table 5: Method configurations for the synthetic experiments.

Method	Definition
Target	Sparse logistic model trained only on past target batches.
Pool	One model on the target prefix and all current snapshots; domain weights are proportional to their sample counts.
Equal-WL	One model with weight $1/(m+1)$ on the target and each source domain.
Dist-WL	Target weight is 0.35. For source-only fits β_j and the past-target fit β_R, let $d_j = \|\beta_R - \beta_j\|_1$ and $a_j \propto \exp[-d_j/(\mathrm{median}_k\, d_k + 10^{-8})]$; source weight is $0.65a_j$.
Credit-TL	Source-indexed Transfer predictors with $\gamma = 0.35$ and source share $1/T$, without a target-only state. Source weights restart uniformly after replacement.
SafePool	Two-state fixed share over the target-only and Pool predictors, with $(\alpha_o, q_1) = (0.05, 0.7)$.
Flat-FS	Flat fixed share over the target-only and all Transfer states, with $(\gamma, \alpha, q_1) = (0, 1/T, 0.7)$.
AMTL-Flat	Same fitted models, $q_1 = 0.8$, clipping, and exponential-weights update as AMTL; only the transition is replaced by flat share $\alpha = 0.08$.
AMTL	Structured transition with $(\gamma, \alpha_o, \alpha_s, q_1) = (0.05, 0.025, 0.10, 0.8)$ and Transfer-model penalty scale $c_\lambda = 0.10$.

Synthetic results. Figure 5 shows how AMTL's transfer mass and source weights change in Rotate.

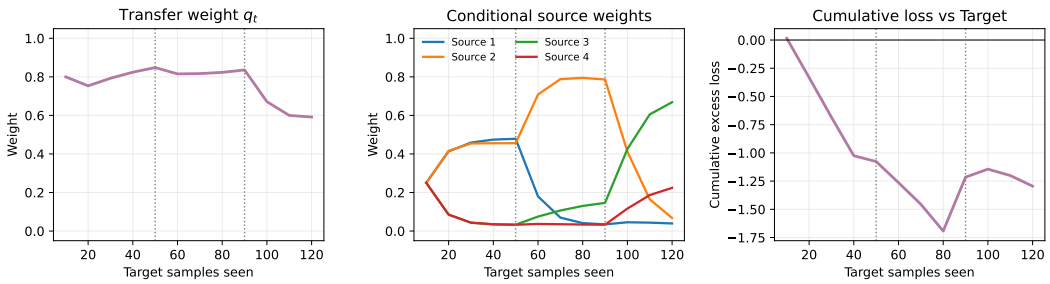


Figure 5: Rotating useful source. Vertical lines mark snapshot replacements. Left: transfer mass. Middle: conditional source weights, whose useful identity changes $1 \to 2 \to 3$. Right: full-stream cumulative AMTL loss minus the matching target-only state, including round one.

In the rotating regime, the useful source is the top-weight source on $71.5\% \pm 2.9\%$ of rounds. Its design-oracle path loss is 3.88 ± 0.13, compared with 7.56 for the target-only baseline. AMTL has cumulative loss 6.27, compared with 6.54 for Flat-FS and 6.43 for AMTL-Flat. In Switch, the paired differences are $L_{\mathrm{AMTL}} - L_{\mathrm{FlatFS}} = -0.636 \pm 0.080$, $L_{\mathrm{AMTL}} - L_{\text{AMTL-Flat}} = -1.222 \pm 0.050$, and the paired AMTL minus target-only difference is -0.454 ± 0.135. Moreover, Figure 6 shows online batch log loss in Opposite, Flip, Rotate, and Switch.

AMTL has the best average online rank across the eight settings, while its cumulative loss in Opposite remains close to Target's. Credit-TL performs best in Rotate partly because it resets its source weights uniformly when snapshots change, preventing weights learned for the previous sources from carrying over. Table 6 gives the per-setting cumulative online losses behind the aggregate comparison in Table 2.

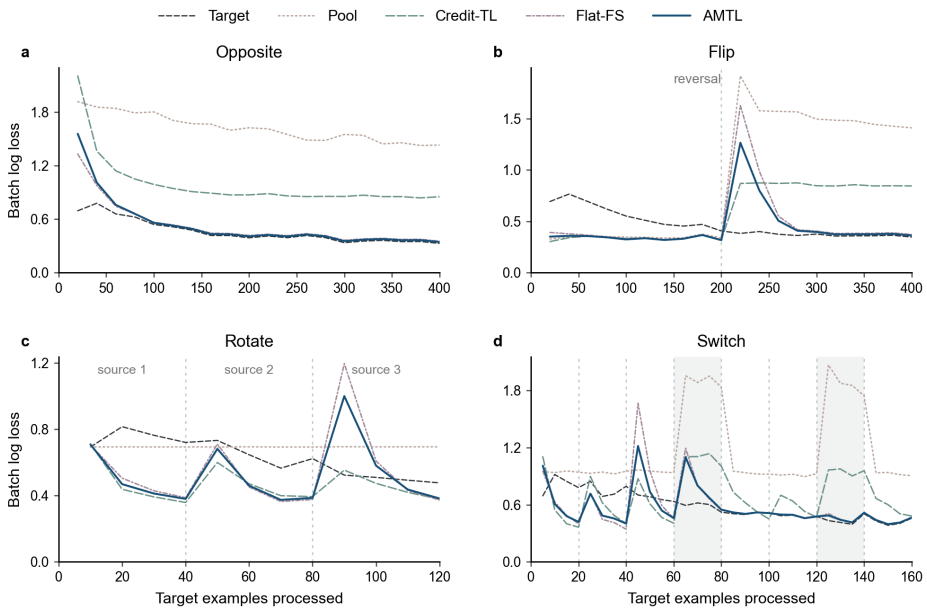


Figure 6: Mean prequential batch log loss over 50 paired repetitions in four mechanism tests: harmful sources, a source flip, a rotating useful source, and joint mode/source switching. Each batch is scored before its labels update any model or state weight. Shading gives 95% normal intervals for the target-only baseline, Flat-FS, and AMTL.

Table 6: Cumulative online log loss across eight synthetic settings, averaged over 50 repetitions. Boldface marks the lowest unrounded mean in each column.

Method	Near	Mixed	Far	Opposite	Flip	Rotate	Scarce	Switch
Target	9.22	9.18	9.34	**9.19**	9.31	7.56	6.61	18.63
Pool	6.70	8.17	9.75	32.46	18.85	8.33	3.33	37.59
Equal-WL	6.71	7.96	9.32	22.34	14.90	8.15	3.34	32.31
Dist-WL	6.76	7.59	8.98	18.06	12.66	7.44	3.40	27.84
Credit-TL	6.41	6.98	8.82	19.81	11.92	**5.59**	3.23	22.24
SafePool	6.96	8.14	9.19	10.53	8.94	7.72	3.61	20.56
Flat-FS	6.52	7.07	8.97	10.53	9.38	6.54	3.37	18.82
AMTL-Flat (ours)	**6.34**	**6.93**	8.89	11.14	9.29	6.43	**3.22**	19.40
AMTL (ours)	6.40	7.05	**8.67**	10.66	**8.63**	6.27	3.26	**18.18**

C Experiment Settings in Section 5.4

Experimental design. We turn CIFAR-10 (Krizhevsky, 2009) into a class-balanced animal-versus-vehicle task. Images are encoded with a frozen ImageNet ResNet-18 (He et al., 2016); the online methods fit only logistic heads and aggregation weights. Four class-dependent color tints define the domains. For each of the first three tints, one source has 300 correctly labeled images and four sources have 50 images each with reversed labels. The fourth tint has no matching source. The target stream contains 120 distinct images in 12 batches of 10, with tint order $1 \to 4 \to 2 \to 4 \to 3 \to 4$ across six phases. Source, validation, and target images are disjoint. We use 30 paired evaluation repetitions. Tinting precedes standard ResNet-18 preprocessing. We replace the final fully connected layer with the identity, normalize each 512-dimensional feature vector to unit length, and fit standardization and a common 64-dimensional PCA map using only unlabeled source features.

Online dynamics and source-pool scaling. Figure 7 shows online loss across the color phases and the effect of increasing the source pool. In the scaling experiment, Flat-FS has lower cumulative loss

at 6 and 9 sources, while AMTL has lower loss at 27 sources when the number of images per source is fixed.

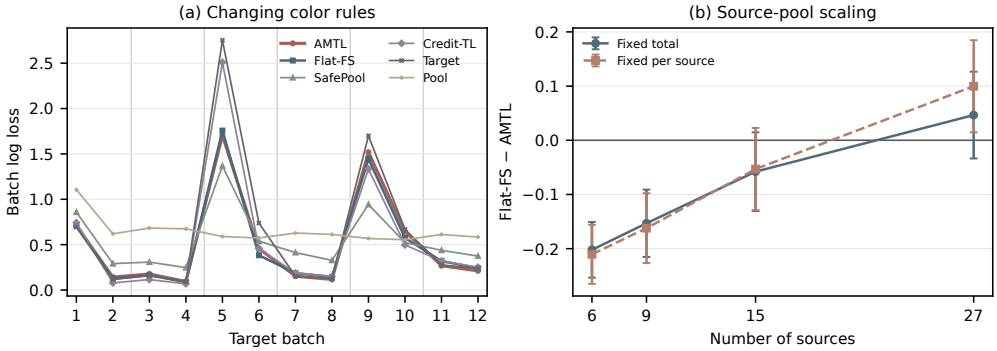


Figure 7: Online behavior under changing color rules. Left: mean batch log loss over 30 repetitions; vertical lines separate the color phases. Right: paired cumulative loss of Flat-FS minus AMTL, with 95% intervals; positive values favor AMTL. Fixed total uses 1,500 source images, while fixed per source holds the sample count for each source constant.

D Experiment Settings in Section 5.5

Table 7 gives the cohort sizes and label prevalence for the four processed UCI Heart Disease cohorts (https://doi.org/10.24432/C52P4X). We omit `slope`, `ca`, and `thal`, which are missing in 190/294, 291/294, and 266/294 Hungarian records; 17/123, 118/123, and 52/123 Swiss records; and 102/200, 198/200, and 166/200 VA records. For the retained continuous variables, all imputation medians, means, and scales are computed from the three source cohorts only. Each continuous and categorical variable has an explicit missingness indicator. Known categorical levels are fixed before splitting, the first level is the reference, and an intercept is appended.

Table 7: Multi-hospital cohorts. A positive label means recorded heart disease.

Cohort	Role	Records	Positive rate
Cleveland	Target	303	0.459
Hungary	Source	294	0.361
Switzerland	Source	123	0.935
VA Long Beach	Source	200	0.745

For repetition $r = 0, \ldots, 29$, the seed is $7301 + 37r$. The Cleveland split is stratified by disease label, and the target order is randomized. Each batch contains 15 records. Every method receives the same sources and target order within a repetition. Stratification is used only to construct the stream; the learner receives each batch's labels after prediction. Table 8 reports cumulative online loss for each method.

Table 8: Cumulative online log loss over 30 paired Cleveland splits. Values are means $\pm$ 95% t-based interval half-widths; lower is better. Boldface marks the lowest mean.

Method	Target	Pool	Equal-WL	Dist-WL	Credit-TL
Online loss	6.490 ± 0.090	6.110 ± 0.072	6.130 ± 0.071	6.126 ± 0.075	6.064 ± 0.064

Method	SafePool	Flat-FS	AMTL-Flat	AMTL
Online loss	6.048 ± 0.070	6.154 ± 0.061	$\mathbf{5.976} \pm 0.064$	5.983 ± 0.064