ADAPTIVE MULTI-SOURCE DATA COLLECTION UNDER DISTRIBUTION SHIFTS

Anonymous authors
Paper under double-blind review

ABSTRACT

When labeled data from a target population are scarce, one can acquire additional training observations from related sites. These sites often differ in their covariate distributions P_X and conditional outcome distributions $P_{Y|X}$, making their value for target prediction unequal and initially unknown. We present an adaptive framework for planning batched data collection across source sites under a fixed budget. We formulate the problem as a Markov decision process, in which states are posterior beliefs about source response functions and actions are site selection for the next batch. Our objective is to reduce disagreement among predictors fitted to the available target labels, measured through their prediction variance over the target population. We focus on one-step lookahead, which selects the next batch to minimize the expected remaining predictor variance. In a tractable Gaussian model, we show that exact one-step lookahead is asymptotically optimal for this objective under regularity and sufficient-sampling conditions. Our framework supports a range of posterior modules (e.g., BNNs and TabPFN) for simulating source labels, as well as different prediction models. Experiments with linear and neural synthetic data and U.S. census income records show improved target prediction over several data collection baselines.

1 INTRODUCTION

When labeled data from a target population are scarce, related sites can provide additional training observations (Desautels et al., 2017; Jean et al., 2016; Mansour et al., 2008). A hospital, for example, may use patient records from partner hospitals, while survey analysts may combine data from different regions to predict socioeconomic indicators such as poverty or income. However, these sites may differ in the populations they represent (covariate distribution shifts in P_X) and in how outcomes depend on covariates (conditional outcome shifts in $P_{Y|X}$) (Zech et al., 2018). Data from different sites can therefore vary in their utility for target prediction, and data from a mismatched site can even increase error on the target (Compton et al., 2023; Shen et al., 2024). Under a fixed budget for collecting data across sites, the key decision is which sites should supply additional labeled data (Mahmood et al., 2022).

Because data typically arrive in batches, site selection is naturally sequential. Each newly acquired batch reveals how outcomes depend on covariates, and offers an opportunity to reassess each site's value for prediction on target (Frazier et al., 2008). Crucially, this value depends not only on a source's similarity to the target. A source may be relevant to target prediction but sufficiently noisy that additional observations offer little benefit. Another source may be poorly understood, but nonetheless contribute more to the target predictor. Useful acquisitions therefore resolve uncertainty in source responses that matters for prediction on the target population. As observations arrive, both our understanding of the sources and the value of further collection change.

We formulate this multi-source data collection as a Markov decision process (MDP), where the state tracks posterior beliefs about each site's response function and observation noise. Each action requests a batch of labeled observations from the sources, and the returned data update these beliefs (Figure 2). Using the limited labeled data available from the target site, we fit or combine source models for target prediction. Because the source response functions are unknown, the posterior over these functions induces a distribution over fitted target predictors. The MDP aims to reduce disagreement among these potential predictors, measured at the end of collection by their prediction variance averaged

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.

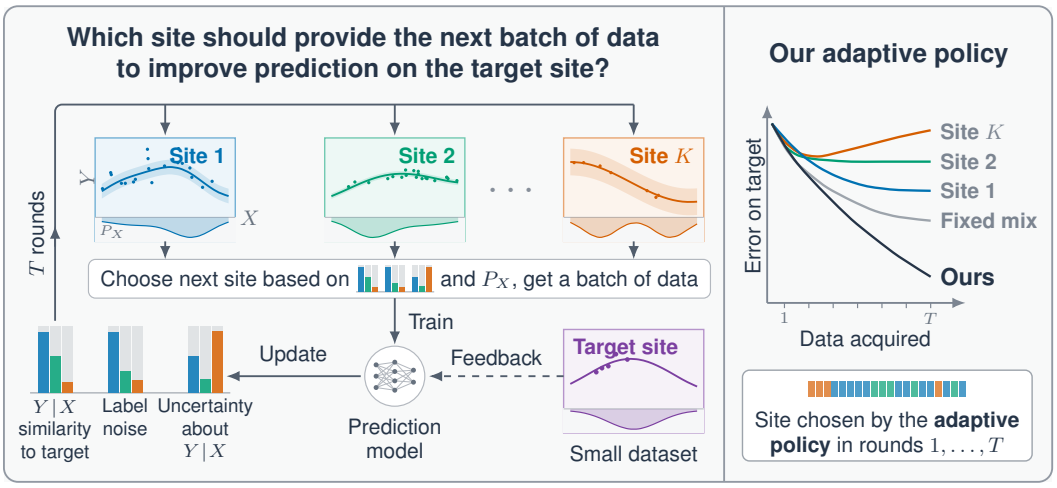

Figure 1: **Adaptive data collection under distribution shift. Left.** With few target labels, we collect batches from K source sites to improve target prediction. Source sites differ in their covariate distributions P_X (lower strips) and conditional outcome distributions $P_{Y|X}$. The bar chart summarizes, schematically, what we currently believe about each site: how similar its $P_{Y|X}$ is to the target's, how noisy its labels are, and how uncertain its response function remains (the width of its band). In each of T rounds, we choose a site based on this belief, collect a batch from it, and update the prediction model and our belief based on all data collected so far. **Right.** Schematic target-error curves. The errors for the single-site data collection often plateaus, and can even rise when the site's $P_{Y|X}$ differs substantially from the target's (Site K). A fixed mix does better but keeps spending labels on sites that are poorly matched or already well estimated. Our adaptive policy reassesses the sites after every batch. The strip shows its choices: It first explores the poorly understood Site K, then alternates between Sites 1 and 2, giving more rounds to Site 1 because its noisier labels require more data. It later revisits Site K to reassess it.

over the target covariate distribution (Section 3). This disagreement reflects epistemic uncertainty in the fitted target predictor arising from uncertainty about the source functions.

Planning over the full horizon is computationally demanding because the space of posterior beliefs is continuous and the number of possible acquisition sequences grows exponentially with the horizon. We therefore focus on one-step lookahead, which selects the next batch to minimize the expected predictor variance after that acquisition (Figure 3 and Section 3.1). To understand the behavior of one-step lookahead, we analyze a tractable Gaussian model and prove that one-step lookahead is asymptotically optimal and recovers the allocation of an oracle that knows the source parameters (Section 4).

Our approach is agnostic to the method used to quantify source uncertainty (posterior module) and can incorporate advances in Bayesian modeling, including tabular foundation models such as TabPFN (Hollmann et al., 2023; 2025). We do not require closed-form posterior updates. The framework can use approximate posterior inference or predictive distributions over unobserved labels without explicit parameter draws (Section 5). We demonstrate this flexibility using Bayesian linear regression, Bayesian neural networks, and TabPFN as posterior modules, paired with linear-mixture and random-forest predictors. Experiments on synthetic data and ACS survey records show that one-step lookahead can improve target prediction and label efficiency over uniform and frequentist-greedy collection (Section 6). We summarize **our contributions** below:

1. **An MDP framework for adaptive data collection.** We formulate multi-source data collection as a belief MDP for efficient adaptive acquisition (Section 3). Our framework accommodates different posterior modules for simulating source data and different target prediction models (Section 5).
2. **Asymptotic optimality of one-step lookahead.** In a tractable Gaussian setting, we show that exact one-step lookahead is asymptotically optimal for reducing predictor variance under regularity and sufficient-sampling conditions (Theorem 3).

3. **Empirical gains from adaptive source allocation.** Experiments with linear and neural synthetic data-generating processes and real survey data show improved target prediction over several collection baselines. These gains extend to different posterior and prediction models (Section 6).

2 Related Work

Multi-source learning. Domain adaptation studies transfer across distributions (Ben-David et al., 2010; Mansour et al., 2008), but covariate alignment alone can fail when conditional outcome mechanisms differ (Zhao et al., 2019). This motivates deciding which shifted datasets to include (Shen et al., 2024) and how to allocate collection effort (Mahmood et al., 2022). Closest to our setting, Zang et al. (2025) sequentially allocate a fixed budget across sources to optimize an aggregate objective. We instead evaluate each source through its contribution to uncertainty in a predictor constructed for a given target population, making both source uncertainty and target relevance central to allocation.

Target-aware experimental design. Bayesian experimental design evaluates acquisitions through posterior expected utility (Chaloner & Verdinelli, 1995); predictive-variance reduction evaluates their effect over a prediction region (Cohn et al., 1996). Expected predictive information gain (Kirsch et al., 2021) and goal-oriented design (Zhu & Hainy, 2026; Zhang et al., 2025) likewise focus on a downstream quantity. Our criterion shares this target-aware perspective, but tracks how uncertainty about source response mechanisms translates into uncertainty in the target predictor. It measures the variance that would remain in that predictor over target covariates after a candidate acquisition. Under conditional outcome shift, this distinction makes the acquisition value depend on how each source enters the target predictor: a source the target predictor does not use contributes no remaining variance, however uncertain its response mechanism.

Sequential design and planning. As labels update source beliefs, they also change which acquisition is most valuable. Knowledge-gradient policies use one-step improvement in a terminal value (Frazier et al., 2008), while deep adaptive design learns history-dependent policies over a full horizon (Foster et al., 2021); online planning can combine amortized policies with lookahead (Guo et al., 2026). We specialize belief-state planning to shifted-source collection: after each batch, posterior-predictive simulations identify the request with the smallest expected remaining predictor variance. This formulation supports oracle allocation and asymptotic guarantees in the Gaussian setting, while allowing different models for source simulation and target prediction.

3 Adaptive Data Collection as Belief MDP

We study how to allocate a fixed collection budget across K source sites to obtain labeled data for prediction on a target population. Here $X \in \mathcal{X}$ denotes covariates and $Y \in \mathbb{R}$ the outcome. Source $k \in \{1, \dots, K\}$ has an unknown distribution P^k, and the target has distribution P^{tgt}. For a prediction loss ℓ, the population target error of a predictor $\psi : \mathcal{X} \to \mathbb{R}$ is

$$\mathbb{E}_{(X,Y)\sim P^{\mathrm{tgt}}}[\ell(Y, \psi(X))].$$

Write P^k_X and $P^k_{Y|X}$ for the source covariate and conditional outcome laws, with target counterparts P^{tgt}_X and $P^{\mathrm{tgt}}_{Y|X}$. Sources may differ from the target in both who they represent and how outcomes depend on covariates. Superscripts index sites; subscripts index acquisition rounds.

Initially, we observe labeled source data $\mathcal{D}_0 = \{\mathcal{D}^1_0, \dots, \mathcal{D}^K_0\}$ and a small labeled target sample $\mathcal{D}^{\mathrm{tgt}} = \{(x_i, y_i)\}_{i=1}^m$. We also have access to unlabeled source and target covariates. The target distribution and target data remain fixed throughout collection; additional labeled data can be acquired only from the sources. Each acquisition is a request that returns one labeled observation or a batch before the next request is chosen. It may obtain labels for specified covariates or newly sampled covariate–outcome pairs. We consider up to T acquisition rounds. With equal acquisition costs, the collection budget counts additional labeled observations beyond initialization; T rounds with a common batch size b yield Tb additional observations.

We use regression with squared loss as a representative setting to develop the formulation and its intuition. Classification and other losses can be incorporated in the same belief MDP; Section C.3 gives the corresponding formulation for classification and other losses.



Figure 2: **Collection as a belief MDP.** Each feasible source request D_t returns observations that update the source posterior and target predictor. The target data stay fixed. The terminal reward $-C(\mu_T)$ measures prediction variance over target covariates.

Bayesian framework. At each source $k \in \{1, \ldots, K\}$ and at the target $k = \mathrm{tgt}$, covariates follow P_X^k. For regression, we write the conditional outcome law $P_{Y|X}^k$ in terms of a fixed but unknown parameter θ^k as

$$Y = f_{\theta^k}(X) + \varepsilon^k, \qquad \mathbb{E}[\varepsilon^k \mid X, \theta^k] = 0.$$

The function f_{θ^k} is the conditional mean, and the noise law describes the remaining outcome variation. The target response $f_{\theta^{\mathrm{tgt}}}$ remains fixed and unknown; we place no prior on it. We form the source likelihood conditional on observed covariates and include unknown source noise parameters in θ. A prior represents uncertainty about the source parameters only:

$$\theta = (\theta^1, \ldots, \theta^K) \sim \mu.$$

The prior may couple sources or factorize into independent source priors. For observed source datasets $\mathcal{D} = \{\mathcal{D}^1, \ldots, \mathcal{D}^K\}$, the posterior is $\mu_{\mathcal{D}} = \mu(\cdot \mid \mathcal{D})$.

Predictor model. Given fixed target data $\mathcal{D}^{\mathrm{tgt}}$ and a draw $\theta \sim \mu_{\mathcal{D}}$, we construct a predictor ψ_θ. The rule may depend on the target data, the posterior draw, or both; we suppress the fixed target data in the predictor notation. Examples include weighted combinations of sampled source functions, neural combiners that also take X as input, and pooled regression using simulated source observations and target data (Section C.1). A fixed construction rule induces a distribution $\nu_{\mathcal{D}}$ over predictors, with mean

$$\psi_\theta \sim \nu_{\mathcal{D}}, \qquad \bar{\psi}_{\mathcal{D}}(x) = \mathbb{E}_{\theta \sim \mu_{\mathcal{D}}}[\psi_\theta(x)].$$

We deploy $\bar{\psi}_{\mathcal{D}}$; variability across draws measures uncertainty in the constructed predictor.

3.1 The belief MDP and one step lookahead

We formulate this problem as a finite-horizon Markov decision process (Puterman, 1994; Mittal et al., 2025), with initial state represented by $\mu_0 = \mu(\cdot \mid \mathcal{D}_0)$. At round t, the state retains the posterior $\mu_t = \mu(\cdot \mid \mathcal{D}_t)$ and the source data collected so far. We use $\mu_{\mathcal{D}}$ and μ_t interchangeably when $\mathcal{D} = \mathcal{D}_t$. The state also contains the available unlabeled covariates and feasibility constraints, although we suppress these in the notation. The target data are fixed context shared by all states.

An action $D_t \in \mathcal{A}_t$ may specify a source from which to collect labeled data, or a source and covariates whose labels are requested. Requests may collect batches and span several sources. Feasibility can restrict batch size, acquisition costs, or available covariates. We add the returned observations to the corresponding source datasets and update $\mu_{t+1} = \mu(\cdot \mid \mathcal{D}_{t+1})$, along with the remaining pools and feasible requests (Figure 2).

Since the source parameters θ are unknown, the fitted predictor for target population ψ_θ is also uncertain. We collect source data to reduce this uncertainty over the space of fitted predictors. We therefore define our terminal objective as the variance of their predictions, averaged over the target covariate distribution:

$$C(\mu_T) = \mathbb{E}_{X \sim P_X^{\mathrm{tgt}}} \mathrm{Var}_{\psi \sim \nu_{\mathcal{D}_T}}[\psi(X)]. \tag{1}$$

This variance measures the expected squared deviation of predictor draws from their mean. Classification and other fitting losses ℓ admit corresponding measures of predictor disagreement (Section C.3). A *non-anticipating policy* π chooses each feasible request using only information available before its observations arrive. Our objective is

$$\min_{\pi} \ \mathbb{E}^{\pi}[C(\mu_T) \mid \mu_0]. \tag{2}$$

Exact planning is generally computationally intractable: beliefs are continuous, and batches drawn across sources can produce a large action space. We instead focus on simplest planning approximation, one-step lookaheads.

One-step lookahead. We focus on the simplest lookahead approximation: evaluate each acquisition as if it were the last, then repeat after observing the returned data. This follows the tradition of one-step expected-utility methods in Bayesian experimental design (Chaloner & Verdinelli, 1995) and the knowledge-gradient policy (Frazier et al., 2008), which selects measurements by their expected improvement in terminal utility. For our minimization objective (2), we choose the feasible request with the smallest expected remaining predictor variance after one additional acquisition.

Definition 1 (One-step lookahead)**.** *At round t, let μ_t be the current posterior and $\mathcal{A}_t$ the feasible request set. For each $D \in \mathcal{A}_t$, let $\mu_{t+1} = \mu(\cdot \mid \mathcal{D}_{t+1})$ be the posterior after adding the observations returned by request D to $\mathcal{D}_t$. With a fixed rule for ties, the one-step lookahead policy selects*

$$D_t^{\text{1-LAH}} \in \arg\min_{D \in \mathcal{A}_t} \mathbb{E}[C(\mu_{t+1}) \mid \mu_t, D_t = D]. \tag{3}$$

4 Understanding One-step lookahead through a Gaussian model

We study when one-step lookahead learns an oracle allocation across Gaussian sources.

Gaussian sources and priors. Source k has fixed true coefficient $\theta^{k,\star} \in \mathbb{R}^d$ and an independent iid stream

$$X^k \sim N(0, \Sigma^k), \qquad Y^k = (X^k)^\top \theta^{k,\star} + \varepsilon^k, \qquad \varepsilon^k \sim N(0, (\sigma^k)^2),$$

with known $\Sigma^k \succ 0$, $(\sigma^k)^2 > 0$, and noise independent of covariates. The source priors are independent $\theta^k \sim N(\bar{\theta}_0^k, V_0^k)$, with $V_0^k \succ 0$. An acquisition selects a source before observing its next covariate and returns the labeled data.

The target has $X^{\text{tgt}} \sim N(0, \Sigma^{\text{tgt}})$ and $Y^{\text{tgt}} = (X^{\text{tgt}})^\top \theta^{\text{tgt},\star} + \varepsilon^{\text{tgt}}$, where $\Sigma^{\text{tgt}} \succ 0$ and the independent noise is $\varepsilon^{\text{tgt}} \sim N(0, (\sigma^{\text{tgt}})^2)$. We evaluate the policy at fixed true parameters under the source-sampling law $P_{\theta^\star}$.

Predictor class. For a posterior draw $\theta \sim \mu_t$, let $\phi \in \mathbb{R}^K$ denote weights on the source response functions $f_{\theta^k}(x) = x^\top \theta^k$. We consider predictions of the form

$$\sum_{k=1}^{K} \phi^k f_{\theta^k}(x) = x^\top \theta \phi, \tag{4}$$

where $\theta = [\theta^1 \ \cdots \ \theta^K] \in \mathbb{R}^{d \times K}$. For each draw, we fit the weights to the fixed target sample $\mathcal{D}^{\text{tgt}}$, whose m observations form $(\mathbf{X}^{\text{tgt}}, \mathbf{y}^{\text{tgt}})$: $\widehat{\phi}(\theta) \in \arg\min_{\phi \in \mathbb{R}^K} \|\mathbf{y}^{\text{tgt}} - \mathbf{X}^{\text{tgt}} \theta \phi\|^2$. In this linear setting, we identify each predictor with its coefficient vector and write $\psi_\theta = \theta \widehat{\phi}(\theta) \in \mathbb{R}^d$, $\bar{\psi}_t = \mathbb{E}_{\theta \sim \mu_t}[\psi_\theta]$. Their predictions at x are $x^\top \psi_\theta$ and $x^\top \bar{\psi}_t$, respectively.

Remaining prediction uncertainty. For the linear predictor, (1) becomes

$$C(\mu_t) = \mathbb{E}_{X \sim P_X^{\text{tgt}}} \operatorname{Var}_{\theta \sim \mu_t}[X^\top \psi_\theta] = \operatorname{tr}\left(\Sigma^{\text{tgt}} \operatorname{Cov}_{\mu_t}[\psi_\theta]\right). \tag{5}$$

Define source k's predictor sensitivity and allocation coefficient by $J^k = \frac{\partial \psi_\theta}{\partial \theta^k}\Big|_{\theta = \theta^\star}$ and $c^k = \operatorname{tr}\left[\Sigma^{\text{tgt}} J^k (\sigma^k)^2 (\Sigma^k)^{-1} (J^k)^\top\right]$.

Assumption 1 (Regular fit and relevant sources)**.** The target design $\mathbf{X}^{\text{tgt}}$ has rank d, $\theta^\star = [\theta^{1,\star} \ \cdots \ \theta^{K,\star}]$ has rank $K \le d$, and $J^k \neq 0$ for every source k.

Theorem 1 (Expected uncertainty reduction)**.** *Under Assumption 1, every feasible source acquisition D satisfies*

$$C(\mu_t) - \mathbb{E}[C(\mu_{t+1}) \mid \mathcal{D}_t, D_t = D] = \operatorname{tr}\left(\Sigma^{\text{tgt}} \operatorname{Cov}[\bar{\psi}_{t+1} \mid \mathcal{D}_t, D_t = D]\right) \ge 0. \tag{6}$$

Let r_t^k count labeled observations from source k, including initialization, and let $n_t = \sum_k r_t^k$. An oracle knows the coefficients c^k and allocates observations to minimize the leading term of C.

Lemma 2 (Oracle allocation)**.** *Under Assumption 1, a deterministic allocation with $r_t^k / n_t \to \alpha^k > 0$ for every k satisfies*

$$n_t C(\mu_t) \longrightarrow \sum_k \frac{c^k}{\alpha^k}, \tag{7}$$

$P_{\theta^\star}$-almost surely conditional on $\mathcal{D}^{\text{tgt}}$. The limiting constant is uniquely minimized by $\alpha^{k,\star} = \sqrt{c^k} / \sum_j \sqrt{c^j}$, with minimum $c^\star = (\sum_k \sqrt{c^k})^2$.

Under the oracle allocation, $C(\mu_t) \sim c^\star/n_t$ vanishes. More generally, as every source count diverges, the mean predictor converges to $\psi_{\theta^\star}$, the coefficient that minimizes empirical target error within the span of the fully known source functions. Its population prediction error can remain positive:

$$\mathbb{E}_{(X,Y)\sim P^{\mathrm{tgt}}}[(Y - X^\top \psi_{\theta^\star})^2] = (\sigma^{\mathrm{tgt}})^2 + \left\|\theta^{\mathrm{tgt},\star} - \psi_{\theta^\star}\right\|^2_{\Sigma^{\mathrm{tgt}}}. \tag{8}$$

This floor depends on the predictor class and on the size and representativeness of the fixed target sample. Learning the sources removes uncertainty about this fitted predictor, rather than the remaining error of the fitting procedure itself.

The oracle depends on unknown source parameters through c^k. We next show that one-step lookahead learns its allocation, provided every source is sampled sufficiently often.

Assumption 2 (Sufficient samples from every source)**.** For some $a > 0$ and $0 < \gamma < 1$, $\min_k r_t^k \geq an_t^\gamma$ eventually, $P_{\theta^\star}$-almost surely.

Theorem 3 (Lower bound and one-step lookahead optimality)**.** *Under Assumptions 1 and 2, the following statements hold $P_{\theta^\star}$-almost surely. Every non-anticipating policy obeys*

$$\liminf_{t\to\infty} n_t C(\mu_t) \geq c^\star. \tag{9}$$

The one-step policy in Definition 1 attains this bound, with $\frac{r_t^k}{n_t} \to \alpha^{k,\star}$ and $n_t C(\mu_t) \to c^\star$. The same limits hold for fixed batches of $b \geq 1$ iid observations from the selected source, using exact joint predictive expectations and observation counts r_t^k.

One-step lookahead learns the oracle proportions and leading constant for remaining prediction uncertainty. Assumption 2 can be ensured by reserving a vanishing fraction of rounds for forced exploration, without changing the oracle limits (Corollary 10). The expressions for $\alpha^{k,\star}$, $c^\star$, and (8) show how allocation and the prediction floor depend on the model parameters.

5 Implementing one-step lookahead

Implementing one-step lookahead requires three components: a posterior module representing our uncertainty about the sources through $\mu_{\mathcal{D}}$, a prediction model ψ_θ for the target population, and a procedure for estimating the expected objective under each candidate action.

Posterior module. The posterior module maintains μ_t, which represents uncertainty about the source functions given the data $\mathcal{D}_t$. Our framework does not require a particular method for constructing this posterior. We can use classical models, such as Gaussian processes and Bayesian linear regression (BLR) (Rasmussen & Williams, 2006), as well as Bayesian neural networks with approximate posterior inference. We can also use more advanced models such as TabPFN, which implement posterior inference implicitly and directly by providing predictive distributions for unobserved labels given the available source data (Hollmann et al., 2025). We can sample labels from these distributions without drawing explicit model parameters. Section G explains how to formulate posterior updates and score acquisitions using these predictive distributions.

Prediction models. Our framework is agnostic to the prediction model family. Examples include predictors $h_\phi(x, \theta)$ that take a posterior draw $\theta \sim \mu_{\mathcal{D}}$ directly as input; weighted combinations of source functions, $\sum_{k=1}^K \phi^k f_{\theta^k}(x)$; nonlinear combinations, $g_\phi(x, f_{\theta^1}(x), \ldots, f_{\theta^K}(x))$, where g_ϕ may be a neural network; and regression models fitted to simulated source observations together with fixed target labels. We fix the simulation and fitting rules before collection, so source observations affect predictor draws through the posterior. Section C.1 describes these families and the corresponding predictor distributions in more detail.

Simulated rollout and objective estimation. To implement one-step lookahead (Definition 1), we use simulated rollouts to estimate the expected objective after each candidate action. A rollout generates an imagined batch (respecting the feasibility constraints) from the current posterior, updates a copy of the posterior, constructs predictor draws using the fixed fitting rule, and evaluates their remaining variance (Figure 3). A batch may contain observations from multiple sites. An action can specify how many observations to collect from each site, or additionally choose their covariates. Here we focus on selecting a single site for the entire batch: $D = k$ requests a batch from source k. For each candidate, we simulate M possible batches and update copies of μ_t to obtain $\widetilde{\mu}_D^{(1)}, \ldots, \widetilde{\mu}_D^{(M)}$. At each update, we use a Monte Carlo estimate $\widehat{C}$ of the objective in (1). We select

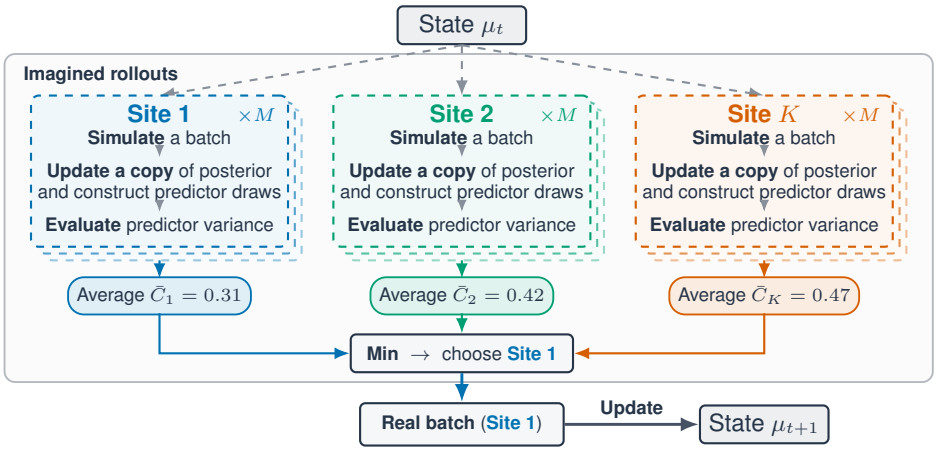

Figure 3: **One-step lookahead by simulation.** For each candidate site, simulate a batch from the current posterior, update a copy of the posterior, and evaluate predictor variance over target covariates. Stacked cards indicate M such simulated batches per site; averaging their scores estimates the objective C. The site with the smallest average is selected (Site 1 here; the averages shown are illustrative), a real batch is collected from it, and the state is updated to μ_{t+1}.

$D_t \in \arg\min_{D \in \mathcal{A}_t} \frac{1}{M} \sum_{r=1}^{M} \widehat{C}(\widetilde{\mu}_D^{(r)})$. We then collect the actual batch and update the posterior. Section F gives the overall algorithm, including a short extension to choosing covariates as well.

6 Experiments

We evaluate one-step lookahead on synthetic data and ACS survey records under a fixed collection budget. We begin with source and prediction models that match the data-generating structure, then change the posterior module, predictor, and both before considering a nonlinear target relationship and real data.

Setup. Each synthetic DGP has five sources with $Y = f_{\theta^k}(X) + \varepsilon^k$: source functions are linear or neural, with site-specific noise variances, and the target is a linear mixture or neural combination of these functions. Posterior modules are BLR, BNN (frozen neural features with a Bayesian last layer), or TabPFN; predictors are linear mixtures or random forests (RF) combining source predictions. ACS sources are state–age populations. Each round acquires a batch from one source; the labeled target sample stays fixed. Section H gives the specifications. TabPFN variants also change acquisition scoring and simulation budgets, so comparisons do not isolate the posterior module alone.

Baselines. We compare four collection policies: *Uniform* samples sources with equal probability; *one-step lookahead (OSL, ours)* simulates acquisitions and minimizes expected remaining predictor variance; *frequentist greedy (FG)* maximizes estimated target-error reduction by withholding and restoring collected observations; and *DAD* selects the source least distinguishable from the target by a classifier (Shen et al., 2024). FG and DAD also adapt after acquisitions, so these comparisons test the value of lookahead-based uncertainty reduction beyond adaptivity alone (Section H.3). OSL (TabPFN) uses TabPFN predictive draws for acquisition simulation; FG (TabPFN) uses TabPFN source predictions for retrospective scoring.

Results. Each synthetic setting has two noise configurations. We show the one with the larger FG–OSL gap in mean MSE averaged over collection rounds; the appendix reports the other (Section H.5).

The linear setting establishes collection efficiency when the source and prediction models match the data-generating structure (Figure 4). OSL approaches the oracle allocation for the experiment's variance proxy and reaches the same target error with substantially less collection: FG and uniform require 17 and 29 rounds, respectively, to match OSL's round-10 error. To test dependence on a source model that matches the data-generating structure, we replace BLR with TabPFN while retaining the linear-mixture predictor. Much of the early advantage persists, demonstrating effective collection with an alternative posterior module.

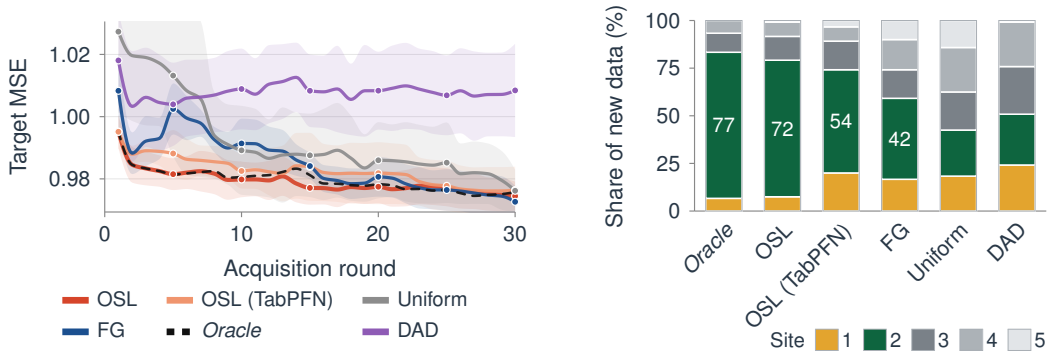

Figure 4: **Linear sources with a linear-mixture predictor.** Left: target MSE (mean and pointwise standard error over four seeds; bands are clipped at the axis limits). Right: mean final allocation of additional observations. The dashed curve shows the saved oracle allocation.

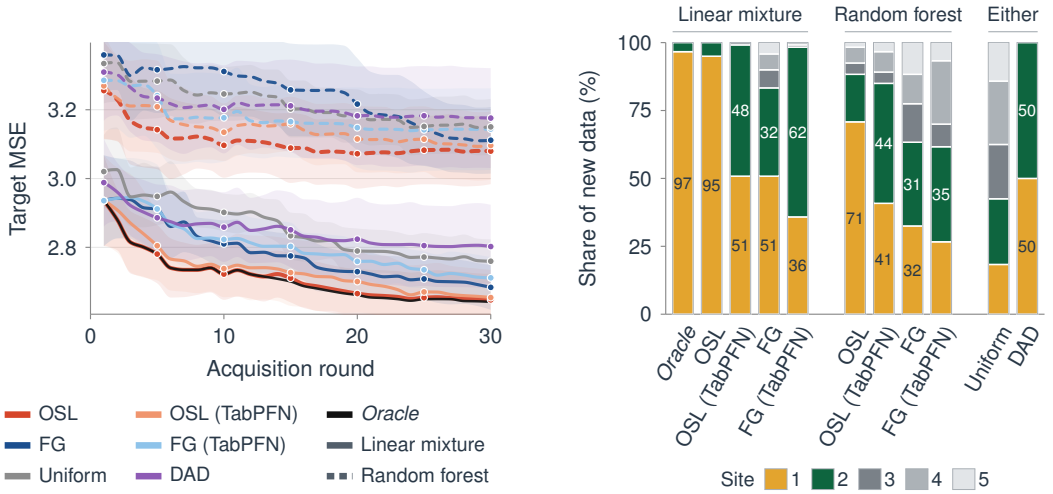

Figure 5: **Neural sources with a linear target mixture.** Left: target MSE (mean and pointwise standard error, four seeds), with solid lines for the linear-mixture predictor and dashed lines for RF. Right: mean final source allocation, grouped by predictor.

(a) Linear-mixture predictor

Method	Round 10	Round 20	Round 30
OSL	$\mathbf{2.72} \pm 0.1$	$\mathbf{2.67} \pm 0.1$	$\mathbf{2.65} \pm 0.1$
OSL (TabPFN)	2.74 ± 0.1	2.70 ± 0.1	$\mathbf{2.65} \pm 0.1$
FG	2.81 ± 0.1	2.73 ± 0.1	2.68 ± 0.1
FG (TabPFN)	2.82 ± 0.1	2.76 ± 0.1	2.71 ± 0.1
Uniform	2.90 ± 0.1	2.79 ± 0.1	2.76 ± 0.1
DAD	2.86 ± 0.1	2.82 ± 0.1	2.80 ± 0.1
Oracle	2.72 ± 0.1	2.66 ± 0.1	2.64 ± 0.1

(b) RF predictor

Method	Round 10	Round 20	Round 30
OSL	$\mathbf{3.10} \pm 0.1$	$\mathbf{3.07} \pm 0.1$	$\mathbf{3.08} \pm 0.1$
OSL (TabPFN)	3.13 ± 0.1	3.11 ± 0.1	3.10 ± 0.1
FG	3.31 ± 0.1	3.22 ± 0.1	3.11 ± 0.1
FG (TabPFN)	3.18 ± 0.1	3.15 ± 0.1	3.14 ± 0.1
Uniform	3.25 ± 0.1	3.18 ± 0.1	3.15 ± 0.1
DAD	3.20 ± 0.1	3.18 ± 0.1	3.18 ± 0.1

Table 1: Target MSE for Figure 5. Bold marks the lowest rounded mean, excluding Oracle.

With neural sources and a linear target mixture, we vary the posterior, predictor, and both (Figure 5 and Table 1). With BNN and a linear-mixture predictor, FG needs 22 rounds to match OSL's round-10 error; uniform and DAD never reach it within budget. Replacing the predictor with RF raises error, but OSL retains lower mean error than the baselines in every round. TabPFN lookahead remains competitive under both predictors, extending the collection benefit across joint model changes.

Allocations connect these gains to source relevance and uncertainty. Sources 1 and 2 contribute equally to the target, but source 1 is noisier. With the mixture predictor, OSL assigns 95% of new labels to source 1, close to the oracle's 97%, while DAD splits them equally. Thus, equally relevant

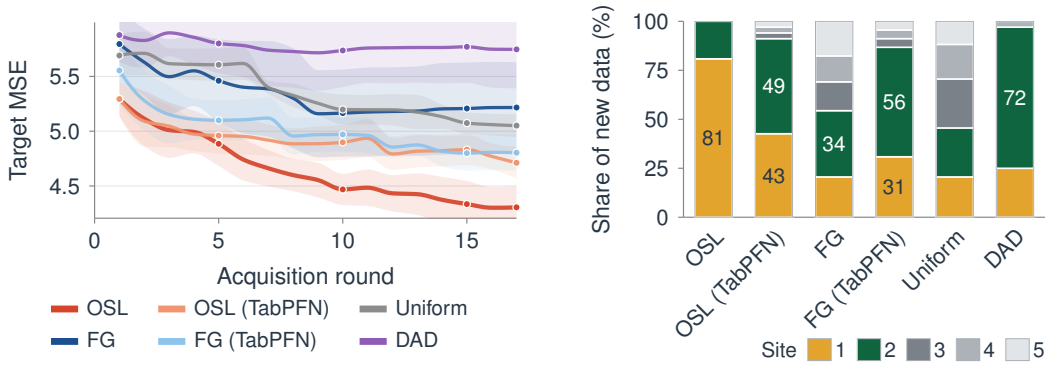

Figure 6: **Neural sources and a neural target combination with an RF predictor.** Left: target MSE (mean and pointwise standard error, four seeds). Right: mean final source allocation.

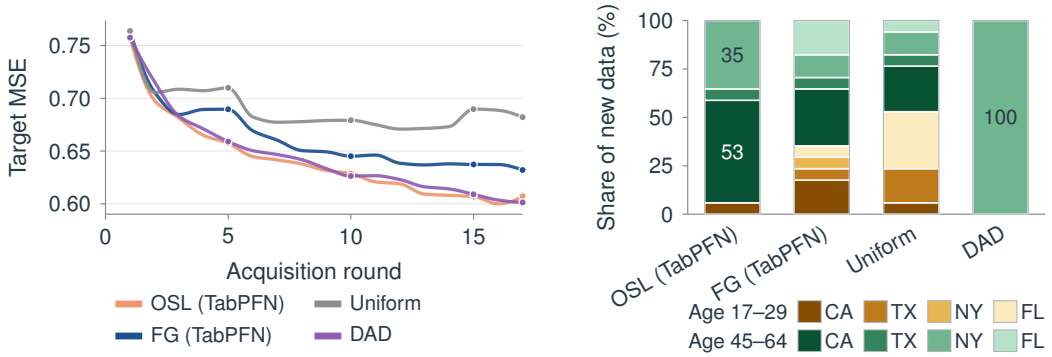

Figure 7: **ACS collection across state–age populations.** Left: target loss for the single available run. Right: final source allocation; gold and green denote younger and older age groups, respectively.

sources need not warrant equal collection. TabPFN lookahead also concentrates almost entirely on these sources, but divides the budget more evenly.

We next replace the linear target mixture with a neural combination, retaining RF to test collection beyond linear target relationships (Figure 6 and Table 7). OSL has lower mean error than FG, uniform, and DAD throughout collection, ending at 4.31 versus 5.22 for FG and 5.05 for uniform. TabPFN lookahead reaches 4.71, close to FG (TabPFN) at 4.80: the advantage with the BNN posterior module extends to nonlinear targets, but the TabPFN rules are less separated. TabPFN lookahead does not outperform the baselines in every noise configuration (Section H.6).

Finally, ACS uses observed outcomes without a constructed source–target relationship (Figure 7). The target mixes CA_45–64 and NY_45–64. TabPFN lookahead allocates 15 of 17 batches to these populations, finishing 11% below uniform and 4% below FG (TabPFN). DAD always selects NY_45–64 and finishes 1% lower, although TabPFN lookahead leads it in 15 of 17 rounds. The populations are nearly interchangeable: their income models, fitted on all available data, have correlation 0.99 on target covariates, and NY alone is within about 1% of the best source combination. This similarity makes total collection from the pair more consequential for remaining predictor variance than its split, explaining the small cost of DAD's single-source allocation. DAD selects by similarity without assessing this redundancy; in the synthetic worlds, where similarity misses consequential source differences, it has the highest loss. Lookahead remains competitive across both cases without requiring interchangeable sources. The ACS comparison is based on a single run.

7 Conclusion

We formulate multi-source data collection as a belief MDP and use one-step lookahead to reduce uncertainty in predictors fitted to target data. In the Gaussian setting, this policy asymptotically attains the oracle allocation for predictor variance under regularity and sufficient sampling. Experiments show that adaptive collection can improve target prediction across posterior and prediction models.

AI USE STATEMENT

We used generative AI to assist with manuscript organization, drafting and editing, literature searches, review of the formulation and proofs, experimental protocol design, and preparation of LaTeX figures. This assistance included revisions to theoretical statements and their proofs; it did not include running new experiments. The authors reviewed and verified the resulting text, citations, mathematics, figures, and protocols and take responsibility for the final manuscript.

REPRODUCIBILITY STATEMENT

The collection objective, available information, and acquisition rule are specified in Sections 3 and 3.1. Posterior-predictive simulation and Monte Carlo estimation are detailed in Sections 5 to G, and the assumptions and proofs are provided in Sections D and E. Section H documents the experimental settings, baselines, and evaluation procedures. The manuscript repository includes exported numerical results and plotting scripts for regenerating the reported figures and tables. We will release the experimental code and configurations upon publication. Sections H.2 and H.4 report the generator specifications, sample sizes, and ACS protocol used in the experiments.

ETHICS STATEMENT

This work studies data collection for predictive modeling through synthetic experiments and an ACS survey-data study. The healthcare example is motivational; we report no clinical deployment. The acquisition objective averages predictor uncertainty over the target covariate distribution and does not enforce subgroup coverage or equitable prediction accuracy. A policy may therefore allocate few observations to some populations. Applications involving people should assess subgroup performance and incorporate appropriate coverage constraints. For survey data, the choice of source populations, target population, and sampling weights determines which populations the findings describe. The ACS results are based on a single run and do not establish the suitability of the method for individual-level or policy decisions.

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*, 79:151–175, 2010. doi: 10.1007/s10994-009-5152-4. URL https://link.springer.com/article/10.1007/s10994-009-5152-4.

Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995. doi: 10.1214/ss/1177009939. URL https://www.stat.cmu.edu/tr/tr599/tr599.html.

Yuxin Chen, S. Hamed Hassani, Amin Karbasi, and Andreas Krause. Sequential information maximization: When is greedy near-optimal? In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 338–363, 2015. URL https://proceedings.mlr.press/v40/Chen15b.html.

David A. Cohn, Zoubin Ghahramani, and Michael I. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4:129–145, 1996. doi: 10.1613/jair.295. URL https://doi.org/10.1613/jair.295.

Rhys Compton, Lily Zhang, Aahlad Puli, and Rajesh Ranganath. When more is less: Incorporating additional datasets can hurt performance by introducing spurious correlations. In *Proceedings of the 8th Machine Learning for Healthcare Conference*, volume 219 of *Proceedings of Machine Learning Research*, pp. 110–127. PMLR, 2023. URL https://proceedings.mlr.press/v219/compton23a.html.

Thomas Desautels, Jacob Calvert, Jana Hoffman, Qingqing Mao, Melissa Jay, Grant Fletcher, Chris Barton, Uli Chettipally, Yaniv Kerem, and Ritankar Das. Using transfer learning for improved mortality prediction in a data-scarce hospital setting. *Biomedical Informatics Insights*, 9: 1178222617712994, 2017. doi: 10.1177/1178222617712994. URL https://journals.sagepub.com/doi/10.1177/1178222617712994.

Adam Foster, Desi R. Ivanova, Ilyas Malik, and Tom Rainforth. Deep adaptive design: Amortizing sequential bayesian experimental design. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3386–3396. PMLR, 2021. URL https://arxiv.org/abs/2103.02438.

Peter I. Frazier, Warren B. Powell, and Savas Dayanik. A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization*, 47(5):2410–2439, 2008. doi: 10.1137/070693424. URL https://optimallearning.princeton.edu/Papers/FrazierPowellDayanik_KnowledgeGradientSICON.pdf.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437. URL https://sites.stat.washington.edu/people/raftery/Research/PDF/Gneiting2007jasa.pdf.

Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011. doi: 10.1613/jair.3278. URL https://arxiv.org/abs/1003.3967.

Google Research. TabFM: A zero-shot foundation model for tabular data. Model release, 2026. URL https://research.google/blog/introducing-tabfm-a-zero-shot-foundation-model-for-tabular-data/.

Yujia Guo, Daolang Huang, Xinyu Zhang, Sammie Katt, Samuel Kaski, and Ayush Bharti. Constrained bayesian experimental design via online planning. *arXiv preprint arXiv:2605.26990*, 2026. doi: 10.48550/arXiv.2605.26990. URL https://arxiv.org/abs/2605.26990. Accepted at the 43rd International Conference on Machine Learning.

Noah Hollmann, Samuel Müller, Katharina Eggensperger, and Frank Hutter. TabPFN: A transformer that solves small tabular classification problems in a second. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2207.01848.

Noah Hollmann, Samuel Müller, Lennart Purucker, Arjun Krishnakumar, Max Körfer, Shi Bin Hoo, Robin Tibor Schirrmeister, and Frank Hutter. Accurate predictions on small data with a tabular foundation model. *Nature*, 637:319–326, 2025. doi: 10.1038/s41586-024-08328-6. URL https://www.nature.com/articles/s41586-024-08328-6.

Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. Bayesian active learning for classification and preference learning. *arXiv preprint arXiv:1112.5745*, 2011. URL https://arxiv.org/abs/1112.5745.

Neal Jean, Marshall Burke, Michael Xie, W. Matthew Alampay Davis, David B. Lobell, and Stefano Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301): 790–794, 2016. doi: 10.1126/science.aaf7894. URL https://www.science.org/doi/10.1126/science.aaf7894.

Andreas Kirsch, Tom Rainforth, and Yarin Gal. Test distribution-aware active learning: A principled approach against distribution shift and outliers. *arXiv preprint arXiv:2106.11719*, 2021. doi: 10.48550/arXiv.2106.11719. URL https://arxiv.org/abs/2106.11719.

Rafid Mahmood, James Lucas, Jose M. Alvarez, Sanja Fidler, and Marc T. Law. Optimizing data collection for machine learning. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://arxiv.org/abs/2210.01234.

Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation with multiple sources. In *Advances in Neural Information Processing Systems*, volume 21, 2008. URL https://proceedings.neurips.cc/paper/2008/hash/0e65972dce68dad4d52d063967f0a705-Abstract.html.

Daksh Mittal, Yuanzhe Ma, Shalmali Joshi, and Hongseok Namkoong. A planning framework for adaptive labeling. *arXiv preprint arXiv:2502.06076*, 2025. URL https://arxiv.org/abs/2502.06076.

Samuel Müller, Noah Hollmann, Sebastian Pineda Arango, Josif Grabocka, and Frank Hutter. Transformers can do Bayesian inference. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2112.10510.

Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html.

Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994. doi: 10.1002/9780470316887. URL https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316887.

Jingang Qu, David Holzmüller, Gaël Varoquaux, and Marine Le Morvan. TabICL: A tabular foundation model for in-context learning on large data. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 50817–50847, 2025. URL https://proceedings.mlr.press/v267/qu25d.html.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006. URL https://gaussianprocess.org/gpml/chapters/RW.pdf.

Judy Hanwen Shen, Inioluwa Deborah Raji, and Irene Y. Chen. The data addition dilemma. In *Proceedings of the 9th Machine Learning for Healthcare Conference*, volume 252 of *Proceedings of Machine Learning Research*. PMLR, 2024. URL https://proceedings.mlr.press/v252/shen24a.html.

Naimeng Ye and Hongseok Namkoong. Exchangeable sequence models quantify uncertainty over latent concepts. *arXiv preprint arXiv:2408.03307*, 2024. URL https://arxiv.org/abs/2408.03307.

Chengbo Zang, Mehmet Kerem Turkcan, Gil Zussman, Zoran Kostic, and Javad Ghaderi. Adaptive data collection for robust learning across multiple distributions. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 73974–73994. PMLR, 2025. URL https://proceedings.mlr.press/v267/zang25a.html.

John R. Zech, Marcus A. Badgeley, Manway Liu, Anthony B. Costa, Joseph J. Titano, and Eric Karl Oermann. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLOS Medicine*, 15(11):e1002683, 2018. doi: 10.1371/journal.pmed.1002683. URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002683.

Zheyu Zhang, Jiayuan Dong, Jie Liu, and Xun Huan. Goal-oriented sequential bayesian experimental design for causal learning. *arXiv preprint arXiv:2507.07359*, 2025. doi: 10.48550/arXiv.2507.07359. URL https://arxiv.org/abs/2507.07359.

Han Zhao, Remi Tachet des Combes, Kun Zhang, and Geoffrey Gordon. On learning invariant representations for domain adaptation. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 7523–7532. PMLR, 2019. URL https://proceedings.mlr.press/v97/zhao19a.html.

Hao Zhu and Markus Hainy. Sequential bayesian experimental design for prediction in physical experiments informed by computer models. *arXiv preprint arXiv:2603.16756*, 2026. doi: 10.48550/arXiv.2603.16756. URL https://arxiv.org/abs/2603.16756.

A Information and notation

The observed source data are $\mathcal{D}_t = \{\mathcal{D}_t^1, \ldots, \mathcal{D}_t^K\}$. The target data $\mathcal{D}^{\mathrm{tgt}}$ are fixed and kept separate. Source parameters have prior μ and posterior $\mu_{\mathcal{D}} = \mu(\cdot \mid \mathcal{D})$. The target response $f_{\theta^{\mathrm{tgt}}}$ is fixed and unknown, with no prior; its parameter is not part of the source vector θ. We write $\mu_t = \mu_{\mathcal{D}_t}$, as in Section 3. Sites appear in superscripts and acquisition rounds in subscripts; any additional sample or simulation index is defined locally. All statements condition on the fixed target data unless an expectation over their sampling distribution is explicitly stated.

The same fixed target sample $\mathcal{D}^{\mathrm{tgt}} = \{(x_i, y_i)\}_{i=1}^m$ is used to define the fixed prediction rule. Unlabeled target covariates define where its posterior variance is measured.

Symbol	Meaning
$\mathcal{D}_t^k, \mathcal{D}_t$	Source-k labeled data and the collection of source datasets after round t.
$\mathcal{D}^{\mathrm{tgt}}$	Fixed labeled target sample of size m, used for fitting the predictor.
$U^{\mathrm{tgt}}, \mathcal{D}^{\mathrm{tgt}}_{\mathrm{test}}$	Unlabeled target integration grid and separate evaluation data.
θ, μ_t	Source parameters and their posterior, with observed context retained in state notation.
$D_t, \mathcal{A}_t$	Source acquisition request and currently feasible requests.
$\nu_{\mathcal{D}}, \bar{\psi}_{\mathcal{D}}$	Constructed predictor distribution and its mean.
$C(\mu_t)$	Predictor variance averaged over P_X^{tgt}.
$\psi_\theta, \bar{\psi}_t$	Gaussian coefficient vector $\psi_\theta = \theta\widehat{\phi}(\theta)$ and its posterior mean; predictions at x are $x^\top\psi_\theta$ and $x^\top\bar{\psi}_t$.
r_t^k, n_t	Counts of labeled source observations and their total, including initialization.
$\widehat{\theta}_t^k, \widehat{V}_t^k$	Source-k posterior mean and covariance after round t; the prior uses $\bar{\theta}_0^k, V_0^k$.
$c^k, \alpha^{k,\star}$	Allocation coefficient and oracle proportion for the predictor-variance objective.
$R_{\infty,m}$	Limiting true target MSE with source responses known and target fit fixed.

Table 2: Notation for source-only collection with fixed target data.

A site-only request selects a source and batch size. A request with specified covariates is a set $D = \{(k, x_j^k) : 1 \le j \le b^k,\ 1 \le k \le K\}$. It specifies items to label, not their unknown outcomes. Feasibility restricts batch size, cost, and remaining source pools. Item identities are retained when covariates repeat. Acquired items leave their pools; their observations enter only the corresponding source datasets. In the experiments, each source offers one candidate batch with visible covariates, and the learner selects one of them. With equal acquisition costs, the collection budget counts additional labeled observations. Each full round returns b observations, so T rounds yield Tb additional observations. The theoretical counts include initialization: $n_T - n_0 = Tb$ for these fixed-size batches.

The shorthand state μ_t retains the posterior, observed source data, available pools, and feasibility constraints. The round determines the remaining horizon, and feasibility incorporates the remaining budget. Fixed target data define the prediction rule at every state. The source posterior determines its distribution; the observed pool and feasibility context are still needed when actions depend on depleted pools. Under independent source priors and site-specific likelihoods, an acquisition updates only the selected source; coupled source priors can update other source beliefs. Non-anticipating acquisition choices contribute no extra likelihood for the unknown parameters once the observed history is retained.

B Additional related work

Section 2 covers multi-source learning, active learning, and experimental design. Here we discuss guarantees for one-step policies and predictive models that can support simulated acquisitions.

B.1 Guarantees for one-step decisions

Adaptive submodularity extends diminishing returns to sequential decisions with partially observed outcomes. Under the required conditions, adaptive greedy policies admit approximation guarantees relative to an optimal adaptive policy (Golovin & Krause, 2011). For sequential information max-

imization with noisy observations, Chen et al. (2015) give near-optimality guarantees based on a separability condition on the observation channel. Their examples also show that greedy information gathering can fail when this condition is weak. These results tie the quality of a one-step decision to properties of the utility and observation model.

Our variance objective does not assume adaptive submodularity or the separability condition above. Under coherent updating of a fixed predictor map, its conditional gain is nonnegative but can be zero (Section E.3). The analysis in Section 4 derives the posterior-conditional gain for the fitted predictor and proves an allocation guarantee under positive allocation coefficients and sufficient sampling. The oracle comparison concerns the remaining predictor variance C. The limits hold at fixed true source parameters; they do not assert finite-horizon optimality for the general collection problem. Vanishing forced exploration supplies the sampling condition (Section E.1).

B.2 Tabular foundation models and predictive simulation

Prior-data fitted networks approximate Bayesian prediction by training on tasks sampled from a prior (Müller et al., 2022). The training examples teach a network to predict missing labels conditional on an observed dataset. At inference time, the dataset supplies context for prediction without requiring an explicit posterior over model parameters. TabPFN applies this approach to tabular classification (Hollmann et al., 2023), and its later version supports both classification and regression (Hollmann et al., 2025). TabICL extends in-context classification to larger tables (Qu et al., 2025). Google's TabFM also supports in-context classification and regression, combining row and column attention with compressed row representations (Google Research, 2026).

These models motivate a predictive interface to data collection. Ye & Namkoong (2024) connect exchangeable sequence prediction to Bayesian uncertainty over latent environments, with forward generation serving as posterior simulation. For our planner, the relevant operation is to generate the labels a candidate acquisition would return and then condition on those labels. Batch acquisitions require their joint predictive law: matching the marginal distribution of each label does not determine the dependence within the batch. Section G describes joint label simulation and predictive updates.

The label simulator and the prediction model can play different roles. Some implementations use TabPFN to generate hypothetical source labels and a parametric model for updates and scoring. Others score predictive trajectories directly, with a different model used for final prediction. The model substitutions examine how each component affects source selection. Predictive completion and refitting also permit learners such as random forests and CatBoost, provided their induced predictor distribution is specified. Section C.1 describes these constructions and distinguishes variation across refitted predictors from uncertainty in the latent source responses.

C Formulation, predictive states, and interpretation

C.1 Prediction model families and predictor distributions

Fix the labeled target sample and a measurable rule $\theta \mapsto \psi_\theta$. The rule, its hyperparameters, and any randomness convention are specified before collection. Source observations enter this construction only through $\mu_{\mathcal{D}}$; drawing $\theta \sim \mu_{\mathcal{D}}$ and returning ψ_θ induces $\nu_{\mathcal{D}}$. Assume the required integrated second moments are finite. Then

$$\mathrm{Var}_{\psi\sim\nu_{\mathcal{D}}}[\psi(x)] = \mathrm{Var}_{\theta\sim\mu_{\mathcal{D}}}[\psi_\theta(x)]. \tag{10}$$

This is uncertainty conditional on the observed data. Merely making a learner a function of the posterior is not enough for a fixed-map interpretation: the same θ must give the same predictor across acquisition rounds.

Using parameter draws directly. A family $h_\phi(x, \theta)$ can use sampled source parameters directly. Fit $\widehat{\phi}(\theta)$ by minimizing $\sum_{(x,y)\in\mathcal{D}^{\mathrm{tgt}}} \ell(y, h_\phi(x, \theta))$, with a fixed rule for ties, and set $\psi_\theta(x) = h_{\widehat{\phi}(\theta)}(x, \theta)$. For example, h_ϕ may be a neural network taking covariates and concatenated source parameters as inputs. This defines a fixed map even though the fitted parameter varies across source draws.

Combining source models. For a source draw θ, fit

$$\widehat{\phi}(\theta) \in \arg\min_{\phi} \sum_{(x,y)\in\mathcal{D}^{\mathrm{tgt}}} \ell\big(y, g_\phi(x, f_{\theta^1}(x), \ldots, f_{\theta^K}(x))\big),$$

with fixed constraints, regularization, and tie-breaking. Then

$$\psi_\theta(x) = g_{\widehat{\phi}(\theta)}(x, f_{\theta^1}(x), \ldots, f_{\theta^K}(x)). \tag{11}$$

A weighted sum, neural network, or forest on source predictions gives an example. Target fitting need not assert that the true target response belongs to this class. The Gaussian theory uses unrestricted linear weights; simplex constraints or other learners require their own regularity and allocation analysis.

Pooling simulated source observations. A source-agnostic learner can instead pool source observations simulated conditional on θ, together with $\mathcal{D}^{\mathrm{tgt}}$. Let $\widetilde{\mathcal{D}}(\theta, \xi)$ denote the simulated data, where ξ contains simulation and training random numbers with a fixed law ρ independent of the acquired data conditional on θ. Fix the simulated sample sizes, source proportions, covariate rules, and fitting procedure $\mathcal{T}$ before collection, and write

$$h(\theta, \xi) = \mathcal{T}(\widetilde{\mathcal{D}}(\theta, \xi), \mathcal{D}^{\mathrm{tgt}}).$$

One fixed-map construction holds $\xi = \xi_0$ fixed and sets $\psi_\theta = h(\theta, \xi_0)$. Another averages over this randomness:

$$\psi_\theta(x) = \mathbb{E}_{\xi\sim\rho}[h(\theta, \xi)(x)]. \tag{12}$$

The first depends on the chosen simulation; the second may require an inner Monte Carlo approximation. These constructions apply to BLR, random forests (Breiman, 2001), CatBoost (Prokhorenkova et al., 2018), and other learners. Actual source labels update $\mu_{\mathcal{D}}$ rather than being added separately to the predictor's training set. Predictive simulation offers an alternative to explicit parameter draws (Ye & Namkoong, 2024).

If ξ is redrawn for each predictor, the induced distribution has variance

$$\mathrm{Var}_{\theta,\xi}[h(\theta, \xi)(x)] = \mathrm{Var}_{\mu_{\mathcal{D}}}[\mathbb{E}_\xi h(\theta, \xi)(x)] + \mathbb{E}_{\mu_{\mathcal{D}}}[\mathrm{Var}_\xi\, h(\theta, \xi)(x)]\,.$$

The second term can remain after the source parameters are known. With coherent updating of the fixed joint latent quantity (θ, ξ), total variance still applies, but the objective then includes simulation and training variability. It is not solely uncertainty about the sources. A learner refitted directly on a growing observed source dataset can change the map itself; its variance score remains computable but does not inherit the fixed-map argument automatically.

C.2 The transition law and Bellman recursion

Let Z_D denote the observations returned by a source request. For known covariates this is its labels; for site-only requests it also includes the sampled covariates. Write $p_D(dz \mid \theta)$ for this sampling law. The predictive law and source posterior update are

$$p_{\mu_t}(dz \mid D) = \int p_D(dz \mid \theta)\mu_t(d\theta), \qquad \mu_{t+1}(d\theta) \propto p_D(z \mid \theta)\mu_t(d\theta).$$

The second expression uses a likelihood density with positive normalizer. Updating source data, pools, and feasibility gives the state $F(\mu_t, D, z)$, as in Section F.1. Target data do not change. A shared source parameter can induce dependence between labels within a batch, so the simulator must preserve their joint predictive law.

Let $V_t(\mu_t)$ be the smallest expected terminal criterion attainable from the retained state at round t. Dynamic programming gives (Puterman, 1994, Ch. 4)

$$\begin{aligned} V_T(\mu_T) &= C(\mu_T), \\ V_t(\mu_t) &= \min_{D\in\mathcal{A}_t} \mathbb{E}_{Z\sim p_{\mu_t}(\cdot|D)}[V_{t+1}(F(\mu_t, D, Z))], \qquad t < T. \end{aligned} \tag{13}$$

The feasible set is recomputed at each hypothetical state. We assume finite feasible families and finite expectations. Pool or budget exhaustion can be represented by empty requests that leave the state unchanged. The one-step rule replaces V_{t+1} by the variance terminal criterion C of (1). For a site-only request, simulation averages over covariates and labels; with specified covariates only labels are simulated. Approximate predictive models define an approximate transition on the retained context.

C.3 Classification and other losses

The regression model and squared loss in Section 3 provide a concrete setting for the belief MDP. The same construction applies to other outcomes by specifying their conditional laws and choosing a loss for fitting and a measure of predictor variability for acquisition.

Bayesian outcome model. Let $Y \in \mathcal{Y}$. At source k, covariates follow P_X^k, and a fixed unknown parameter θ^k specifies the conditional law $p_{\theta^k}(\cdot \mid x)$ of Y given $X = x$. Place a working prior μ on $\theta = (\theta^1, \ldots, \theta^K)$; it may couple sources or factorize into independent priors. Assuming conditionally independent labels given the parameters and covariates, Bayes' rule gives

$$\mu_{\mathcal{D}}(d\theta) \propto \left[\prod_{k=1}^{K} \prod_{(x,y)\in\mathcal{D}^k} p_{\theta^k}(y \mid x)\right] \mu(d\theta).$$

Here $p_{\theta^k}(y \mid x)$ denotes a probability mass or density, and the normalizing constant is assumed finite and positive. The likelihood conditions on the observed covariates. As in Section 3, the target conditional law remains fixed and unknown, with no prior. The fixed target sample $\mathcal{D}^{\mathrm{tgt}}$ is used to define the fixed prediction rule.

For binary classification, take $\mathcal{Y} = \{0, 1\}$ and

$$Y \mid X = x, \theta, k \sim \mathrm{Bernoulli}\big(f_{\theta^k}(x)\big), \qquad f_{\theta^k}(x) \in [0, 1].$$

For vector covariates, Bayesian logistic regression uses $f_{\theta^k}(x) = (1 + \exp(-x^\top \theta^k))^{-1}$ with a Gaussian prior on each source coefficient vector. For J classes, let $f_{\theta^k}(x)$ be a vector of J nonnegative probabilities summing to one and take $Y \mid X = x, \theta, k$ to be categorical with these probabilities. A softmax model with a prior on its weights is one choice. These likelihoods specify label randomness directly; no additive noise model is needed. The framework does not require closed-form posteriors.

Predictor draws. Given $\theta \sim \mu_{\mathcal{D}}$, apply the fixed rule ψ_θ defined using $\mathcal{D}^{\mathrm{tgt}}$. For classification, its output is a class-one probability or a vector of class probabilities. For example, the combiner in Section C.1 can use sampled source probabilities and covariates, with weights fitted on $\mathcal{D}^{\mathrm{tgt}}$. Nonnegative mixture weights summing to one preserve valid probabilities; a neural combiner can enforce them through a sigmoid or softmax output. The draws induce $\nu_{\mathcal{D}}$ as before. Their mean, computed coordinatewise for several classes, is also a valid probability prediction. If a class label is required, use the class with largest mean probability, with a fixed rule for ties.

Acquisition objective. The target fitting loss ℓ determines the rule ψ_θ. For acquisition, choose a nonnegative discrepancy $\delta(a, b)$ between predictions, with $\delta(a, a) = 0$, and specify it before collection. With $\bar{\psi}_{\mathcal{D}}(x) = \mathbb{E}_{\mu_{\mathcal{D}}}[\psi_\theta(x)]$, define

$$C_\delta(\mu_{\mathcal{D}}) = \mathbb{E}_{X\sim P_X^{\mathrm{tgt}}} \mathbb{E}_{\theta\sim\mu_{\mathcal{D}}}[\delta(\psi_\theta(X), \bar{\psi}_{\mathcal{D}}(X))].$$

The target integration uses unlabeled covariates. Taking $\delta(a, b) = (a-b)^2$ recovers (1). No empirical fitting loss is added to the acquisition score.

Absolute error. A fitting loss $\ell(y, a) = |y - a|$ can be combined with the same variance acquisition objective. Alternatively, if acquisition is to measure absolute estimation error, let $a_{\mathcal{D}}(x)$ be a posterior median of $\psi_\theta(x)$, chosen with a fixed rule for ties, and use

$$C_{\mathrm{abs}}(\mu_{\mathcal{D}}) = \mathbb{E}_X \mathbb{E}_{\mu_{\mathcal{D}}} |\psi_\theta(X) - a_{\mathcal{D}}(X)|.$$

The posterior median minimizes this posterior absolute loss. Expected Bayes risk cannot increase after an acquisition: at the updated state, the previous action $a_{\mathcal{D}}$ remains feasible, and averaging its risk uses the tower property. This argument relies on a fixed random predictor and differs from the quadratic covariance argument.

Brier loss. For $Y \in \{0, 1\}$, let $\psi(x) \in [0, 1]$ be a predicted probability of class one. The binary Brier loss $\ell(y, p) = (y - p)^2$ is a proper scoring rule (Gneiting & Raftery, 2007). With $\delta(p, q) = (p - q)^2$, the acquisition objective averages the posterior variance of the predicted probability. This

variance is zero when all predictor draws agree, even if their common probability is $1/2$. For J classes, let e_y be the unit vector for class y. Use $\ell(y,p) = \|e_y - p\|^2$ and $\delta(p,q) = \|p-q\|^2$. The acquisition objective is then the sum of posterior variances of the class probabilities, and the squared-error decomposition holds coordinatewise.

Log loss. For probabilities in $(0,1)$, use $\ell(y,p) = -y\log p - (1-y)\log(1-p)$, another proper scoring rule (Gneiting & Raftery, 2007). Use this loss for target fitting. One acquisition discrepancy is the Kullback–Leibler divergence between Bernoulli predictions,

$$\delta(p,q) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}.$$

Write $H(p) = -p\log p - (1-p)\log(1-p)$. Substituting $q = \bar{\psi}_{\mathcal{D}}(x) = \mathbb{E}_{\psi\sim\nu_{\mathcal{D}}}[\psi(x)]$ gives

$$\mathbb{E}_{\psi\sim\nu_{\mathcal{D}}}\big[\delta\big(\psi(x),\bar{\psi}_{\mathcal{D}}(x)\big)\big] = H\big(\bar{\psi}_{\mathcal{D}}(x)\big) - \mathbb{E}_{\psi\sim\nu_{\mathcal{D}}}[H(\psi(x))].$$

This has the entropy-difference form used in Bayesian active learning by disagreement (Houlsby et al., 2011). It is nonnegative by concavity of H and vanishes when all draws agree. Here it measures remaining predictor disagreement in the terminal objective; the acquisition rule still compares expected objectives after hypothetical updates. For J classes with positive probabilities, use $\ell(y,p) = -\log p_y$, $\delta(p,q) = \sum_{c=1}^{J} p_c\log(p_c/q_c)$, and $H(p) = -\sum_{c=1}^{J} p_c\log p_c$. The same entropy-difference identity holds.

One-step lookahead. To simulate a candidate batch, draw source parameters from μ_t and generate its labels from the corresponding conditional laws. Use one shared parameter draw across the batch to preserve dependence between labels after averaging over the posterior. For a site-only request, also draw its covariates from that source. Update the posterior with the simulated observations, construct the resulting predictor distribution, and evaluate C_δ. Thus Definition 1 becomes

$$D_t^{\text{1-LAH}} \in \operatorname*{arg\,min}_{D\in\mathcal{A}_t} \mathbb{E}\big[C_\delta(\mu_{t+1}) \mid \mu_t, D_t = D\big],$$

with the same feasibility constraints and fixed rule for ties. After acquiring the selected labels, update the posterior by the same Bayes rule and repeat. Only source observations enter these updates.

These choices extend the model and objective; they do not assert the squared-error decomposition for arbitrary ℓ. The loss alone need not determine a unique discrepancy. The allocation guarantees in Section 4 remain specific to the Gaussian regression model with the specified profiled predictor and variance objective.

D Proofs of the main Gaussian results

We prove the expected uncertainty reduction in Theorem 1, the oracle allocation in Lemma 2, and the one-step guarantee in Theorem 3. Supporting results are stated and proved where they enter each argument. Additional results follow in Section E.

All source asymptotics hold at fixed true source parameters, conditional on the realized target sample. Predictive expectations use the working Bayesian source model.

In the Gaussian model, the subscript 0 in $\bar{\theta}_0^k, V_0^k$ precedes all observations; μ_0 is the posterior after initialization, with source moments $\widehat{\theta}_0^k, \widehat{V}_0^k$. We retain t for acquisition rounds and r_t^k for counts of labeled source observations. The analysis selects a source before its next covariate arrives and computes exact posterior moments and predictive expectations. Selection among known covariate candidates and approximate inference require additional analysis. We fit the weights in (4) without constraints, using the minimum Euclidean norm to resolve nonuniqueness. Simplex constraints used in experiments can change the fitted map and its allocation coefficients.

We use the target design and labels $(\mathbf{X}^{\text{tgt}}, \mathbf{y}^{\text{tgt}})$ from Section 4 for the fixed fitting rule. We retain its coefficient-vector convention: $\psi_\theta = \theta\widehat{\phi}(\theta)$ and $\bar{\psi}_t = \mathbb{E}_{\mu_t}[\psi_\theta]$. Their predictions at x are $x^\top\psi_\theta$ and $x^\top\bar{\psi}_t$. For the proofs, let

$$W_t = \operatorname{Cov}_{\mu_t}[\psi_\theta], \qquad C(\mu_t) = \operatorname{tr}(\Sigma^{\text{tgt}} W_t).$$

The empirical target second moment appears in fitting the predictor, not as a subtraction from the acquisition objective. All predictor moments in this analysis are exact.

D.1 Proof of Theorem 1: expected uncertainty reduction

The fixed fitting rule makes ψ_θ the same random quantity before and after acquisition. Its posterior covariance therefore obeys the law of total covariance.

Proof of Theorem 1. For the linear predictor and an independent target covariate,

$$\mathbb{E}_{X\sim P_X^{\mathrm{tgt}}}\operatorname{Var}_{\mu_t}[X^\top\psi_\theta] = \mathbb{E}_X[X^\top W_t X] = \operatorname{tr}(\Sigma^{\mathrm{tgt}} W_t),$$

which proves (5). The target design has full column rank, and the fitted target values are a projection of $\mathbf{y}^{\mathrm{tgt}}$. Consequently $\|\psi_\theta\| \le \|\mathbf{y}^{\mathrm{tgt}}\|/s_{\min}(\mathbf{X}^{\mathrm{tgt}})$ for every source matrix, including singular matrices (Proposition 5). Thus all required moments are finite.

Fix the current data and a feasible action D. Under the joint law $\theta\sim\mu_t$ and $Z\mid\theta\sim p_D(\cdot\mid\theta)$, the posterior after the returned observations Z is μ_{t+1}. The fitting rule and target sample do not change. Total expectation and total covariance give

$$\mathbb{E}[\bar\psi_{t+1}\mid\mathcal{D}_t, D_t=D] = \bar\psi_t,$$
$$W_t = \mathbb{E}[W_{t+1}\mid\mathcal{D}_t, D_t=D] + \operatorname{Cov}[\bar\psi_{t+1}\mid\mathcal{D}_t, D_t=D].$$

Taking the trace against Σ^{tgt} proves (6). The covariance on the right is positive semidefinite, so its trace against $\Sigma^{\mathrm{tgt}}\succ 0$ is nonnegative. The same argument applies to a batch, provided its posterior is updated under the joint predictive law. For a non-anticipating randomized action, condition first on the chosen action and then average. Hence $C(\mu_t)$ is a nonnegative supermartingale under the working Bayesian acquisition law. □

This expectation is under the source posterior predictive law, whereas the asymptotic allocation results evaluate sample paths under fixed true source parameters. Nonnegative expected gains do not alone imply the coverage condition in Assumption 2.

D.2 Proof of Lemma 2: oracle allocation

We first establish posterior consistency and regularity of the fitted coefficient. These give the leading predictor variance for any allocation with sufficient sampling. Minimizing that leading term gives the oracle.

D.2.1 Source posterior and consistency

Proposition 4 (Source posterior and pathwise consistency)**.** *At round t, source k has r_t^k observations and posterior $N(\widehat\theta_t^k, \widehat V_t^k)$, where*

$$\widehat V_t^k = \left[(V_0^k)^{-1} + (\sigma^k)^{-2}\sum_{j=1}^{r_t^k} X_j^k (X_j^k)^\top\right]^{-1},$$
$$\widehat\theta_t^k = \widehat V_t^k\left[(V_0^k)^{-1}\bar\theta_0^k + (\sigma^k)^{-2}\sum_{j=1}^{r_t^k} X_j^k Y_j^k\right]. \tag{14}$$

Here j indexes observations in that source's stream. Along any diverging sequence of source counts, almost surely under the fixed true parameter,

$$\widehat\theta_t^k \to \theta^{k,\star}, \qquad r_t^k \widehat V_t^k \to (\sigma^k)^2(\Sigma^k)^{-1}. \tag{15}$$

These limits hold on one event along every diverging adaptive subsequence of counts.

Proof of Proposition 4. Complete the square in the Gaussian prior density times the conditional Gaussian likelihood to obtain (14). For each source stream, the strong law at deterministic sample sizes r gives

$$r^{-1}\sum_{j=1}^{r} X_j^k (X_j^k)^\top \to \Sigma^k, \qquad r^{-1}\sum_{j=1}^{r} X_j^k \varepsilon_j^k \to 0 \quad \text{almost surely.}$$

An adaptive policy observes a prefix of each stream. On the common probability-one event for the K streams, these limits therefore hold along every sequence of counts $r_t^k \to \infty$. Hence

$$r_t^k \widehat{V}_t^k = \left[\frac{(V_0^k)^{-1}}{r_t^k} + \frac{(\sigma^k)^{-2}}{r_t^k} \sum_{j=1}^{r_t^k} X_j^k (X_j^k)^\top \right]^{-1} \longrightarrow (\sigma^k)^2 (\Sigma^k)^{-1}.$$

The error of the posterior mean is

$$\widehat{\theta}_t^k - \theta^{k,\star} = \widehat{V}_t^k (V_0^k)^{-1} (\bar{\theta}_0^k - \theta^{k,\star}) + (\sigma^k)^{-2} (r_t^k \widehat{V}_t^k) \left(\frac{1}{r_t^k} \sum_{j=1}^{r_t^k} X_j^k \varepsilon_j^k \right),$$

which tends to zero on the same event. This argument does not substitute a stopping time into a moment or central limit theorem. □

D.2.2 The fitted coefficient and its derivative

Proposition 5 (Bounded fit and its derivative)**.** *Under the rank conditions in Assumption 1, the map $\theta \mapsto \psi_\theta$ is bounded and measurable on $\mathbb{R}^{d \times K}$, including rank-deficient source matrices. In particular,*

$$\|\psi_\theta\| \le \frac{\|\mathbf{y}^{\mathrm{tgt}}\|}{s_{\min}(\mathbf{X}^{\mathrm{tgt}})}. \tag{16}$$

It is smooth on a neighborhood of $\theta^\star$. At a full-rank θ, set $\mathbf{Z} = \mathbf{X}^{\mathrm{tgt}}\theta$, $A = \mathbf{Z}^\top \mathbf{Z}$, $\phi = \widehat{\phi}(\theta) = A^{-1}\mathbf{Z}^\top \mathbf{y}^{\mathrm{tgt}}$, and $e = \mathbf{y}^{\mathrm{tgt}} - \mathbf{Z}\phi$. Write $\mathrm{d}\psi_\theta[H]$ for the derivative of $\theta \mapsto \psi_\theta$ in direction H, and define $\mathrm{d}\widehat{\phi}(\theta)[H]$ similarly. Then

$$\begin{aligned} \mathrm{d}\widehat{\phi}(\theta)[H] &= A^{-1}\{(\mathbf{X}^{\mathrm{tgt}}H)^\top e - \mathbf{Z}^\top \mathbf{X}^{\mathrm{tgt}} H\phi\}, \\ \mathrm{d}\psi_\theta[H] &= H\phi + \theta\, \mathrm{d}\widehat{\phi}(\theta)[H]. \end{aligned} \tag{17}$$

Thus $J^k u = \mathrm{d}\psi_{\theta^\star}[u e_k^\top]$, matching the source Jacobian defined in Section 4.

Proof of Proposition 5. The fitted vector $\mathbf{X}^{\mathrm{tgt}}\psi_\theta$ is the orthogonal projection of $\mathbf{y}^{\mathrm{tgt}}$ onto the column space of $\mathbf{X}^{\mathrm{tgt}}\theta$, including at rank-deficient matrices. Thus $\|\mathbf{X}^{\mathrm{tgt}}\psi_\theta\| \le \|\mathbf{y}^{\mathrm{tgt}}\|$. The least singular value bound for the full-column-rank matrix $\mathbf{X}^{\mathrm{tgt}}$ proves (16). All minimizing weights give the same coefficient because $\mathbf{X}^{\mathrm{tgt}}$ is injective. The minimum-norm solution is $\widehat{\phi}(\theta) = \mathbf{Z}^\dagger \mathbf{y}^{\mathrm{tgt}}$, where $\dagger$ denotes the Moore–Penrose pseudoinverse. It is measurable: for instance it is the pointwise limit of the ridge solutions $(\mathbf{Z}^\top\mathbf{Z} + \epsilon I)^{-1}\mathbf{Z}^\top \mathbf{y}^{\mathrm{tgt}}$ as $\epsilon \downarrow 0$.

Near $\theta^\star$, $\mathbf{Z}^\top\mathbf{Z}$ is invertible, and its inverse and the profile maps are smooth. Differentiating $\mathbf{Z}^\top(\mathbf{y}^{\mathrm{tgt}} - \mathbf{Z}\phi) = 0$ yields

$$(\mathbf{Z}^\top\mathbf{Z})\,\mathrm{d}\phi = (\mathrm{d}\mathbf{Z})^\top(\mathbf{y}^{\mathrm{tgt}} - \mathbf{Z}\phi) - \mathbf{Z}^\top(\mathrm{d}\mathbf{Z})\phi.$$

Substitute $\mathrm{d}\mathbf{Z} = \mathbf{X}^{\mathrm{tgt}}H$ and differentiate $\theta\phi$ to get (17). □

For Gaussian target covariates, $m \ge d$ gives the required target-design rank almost surely. The bound above controls the fitted coefficient even near singular source matrices, while full rank of $\theta^\star$ gives local smoothness. Neither condition requires the true target response to lie in the source span.

The allocation coefficients also make the parameter dependence explicit. Holding the target sample, covariances, and true coefficients fixed, multiplying source k's noise standard deviation by a positive factor multiplies its unnormalized allocation weight $\sqrt{c^k}$ by that factor. Source noise and covariate distributions affect the approach to the prediction floor; the floor itself depends on the source span, true target response, target distribution, and finite target fit. Fixed proper priors disappear from the leading covariance in (15) and hence from c^k.

D.2.3 EXPANSION OF THE PREDICTOR VARIANCE

Proposition 6 (Predictor-variance expansion)**.** *Under the rank conditions in Assumption 1 and Assumption 2, every non-anticipating policy satisfies, $P_{\theta^\star}$-almost surely,*

$$C(\mu_t) = \sum_k \frac{c^k}{r_t^k} + o\left(\sum_k \frac{1}{r_t^k}\right). \tag{18}$$

Proof. **Localization at the true source matrix.** We identify matrices with vectors of their columns for Gaussian calculations. At round t, stack the source posterior means in $\widehat{\theta}_t = (\widehat{\theta}_t^1, \dots, \widehat{\theta}_t^K)$ and write $\widehat{V}_t = \operatorname{diag}(\widehat{V}_t^1, \dots, \widehat{V}_t^K)$ for the joint covariance. These joint moments use the source moments in (14). Work on a sample path from Proposition 4 along a policy with polynomial coverage. Then

$$\widehat{\theta}_t \to \theta^\star, \qquad \left\|\widehat{V}_t\right\| = O(n_t^{-\gamma}), \qquad \left\|\widehat{V}_t^k\right\| = O((r_t^k)^{-1}). \tag{19}$$

All constants in the following pathwise bounds can depend on the fixed target sample, model, and sample path, but not on the hypothetical observation or on the acquisition index once it is large enough.

Choose a smooth cutoff equal to one on a ball about $\theta^\star$ and supported on a slightly larger ball where the profile is smooth. The product of this cutoff and ψ_θ, extended by zero, defines a globally smooth bounded map $\theta \mapsto \widetilde{\psi}_\theta$ with bounded first and second derivatives. It equals ψ_θ on the smaller ball. Gaussian tail bounds and (19) imply, for some $c > 0$,

$$\mathbb{E}_{\mu_t}\left\|\psi_\theta - \widetilde{\psi}_\theta\right\|^2 = O(\exp(-cn_t^\gamma)). \tag{20}$$

Indeed, eventually $\widehat{\theta}_t$ is within half the smaller ball's radius of the truth. A deviation outside that ball requires some one of the fixed dK Gaussian coordinates to deviate by a fixed amount. Each coordinate has variance at most $\left\|\widehat{V}_t\right\|$; a union bound gives the displayed exponential bound. Boundedness of both maps then controls the second moment on the exceptional event.

For later use in the one-step gain, fix a hypothetical covariate x at source k and let μ_{t+1} denote the posterior after its label Y is observed. The posterior tower identity gives

$$\mathbb{E}_{Y|x,\mathcal{D}_t,k}\mathbb{E}_{\mu_{t+1}}\left\|\psi_\theta - \widetilde{\psi}_\theta\right\|^2 = \mathbb{E}_{\mu_t}\left\|\psi_\theta - \widetilde{\psi}_\theta\right\|^2.$$

The choice of x carries no information about the source coefficients in this model. Jensen's inequality therefore gives the same bound for the squared difference between the updated posterior means of the two maps, uniformly in x. If A, B are bounded random vectors, expanding their covariance matrices and using Cauchy–Schwarz bounds $\|\mathrm{Cov}(A) - \mathrm{Cov}(B)\|$ by a constant times $(\mathbb{E}\,\|A - B\|^2)^{1/2}$. It follows that replacing ψ_θ by $\widetilde{\psi}_\theta$ changes either its current posterior covariance or the covariance of its updated posterior mean by $O(\exp(-c'n_t^\gamma))$ for some $c' > 0$. This error is smaller than n_t^{-p} for every fixed p. It remains negligible even after multiplication by $(r_t^k)^2 \le n_t^2$.

Posterior covariance expansion. For the current posterior, write $\theta = \widehat{\theta}_t + \eta$ with $\eta \sim N(0, \widehat{V}_t)$. Taylor expand $\widetilde{\psi}_\theta$ at $\widehat{\theta}_t$. The remainder is bounded by a constant times $\|\eta\|^2$. Gaussian moment bounds and Cauchy–Schwarz therefore give

$$\mathrm{Cov}_{\mu_t}[\widetilde{\psi}_\theta] = \mathrm{d}\widetilde{\psi}_{\widehat{\theta}_t}\,\widehat{V}_t\,\mathrm{d}\widetilde{\psi}_{\widehat{\theta}_t}^\top + O\Big((\operatorname{tr}\widehat{V}_t)^{3/2} + (\operatorname{tr}\widehat{V}_t)^2\Big).$$

Since $\mathrm{d}\widetilde{\psi}_{\widehat{\theta}_t} \to \mathrm{d}\psi_{\theta^\star}$ and $\widehat{V}_t$ is block diagonal, this equals

$$\sum_{k=1}^K J^k \widehat{V}_t^k (J^k)^\top + o(\operatorname{tr}\widehat{V}_t).$$

Here $\operatorname{tr}\widehat{V}_t$ is comparable to $\sum_k 1/r_t^k$ by (15). It is also bounded below by a positive constant times $1/n_t$ eventually. The localization error is therefore $o(\operatorname{tr}\widehat{V}_t)$. Restoring ψ_θ, substituting $\widehat{V}_t^k = (\sigma^k)^2(\Sigma^k)^{-1}/r_t^k + o(1/r_t^k)$, and taking the trace against Σ^{tgt} proves (18). □

Polynomial coverage controls posterior mass near singular source matrices even when source counts are unequal. Positivity of the allocation coefficients is not needed for this expansion; it is needed for the oracle comparison.

D.2.4 Optimal source proportions

Proof of Lemma 2. For a deterministic allocation with every limiting proportion positive, Assumption 2 holds, for example with $\gamma = 1/2$. Hence Proposition 6 applies. Multiplying (18) by n_t gives (7) for $n_t C(\mu_t)$, since $n_t \sum_k 1/r_t^k$ stays bounded. For positive allocation coefficients, Cauchy–Schwarz gives

$$\left(\sum_k \sqrt{c^k}\right)^2 \le \left(\sum_k \frac{c^k}{\alpha^k}\right)\left(\sum_k \alpha^k\right) = \sum_k \frac{c^k}{\alpha^k}.$$

Equality holds exactly when α^k is proportional to $\sqrt{c^k}$. All allocation coefficients are positive, so the minimizer is unique and interior. Integer allocations converging to these proportions exist by rounding. Thus the minimum is attained asymptotically. □

D.3 Proof of Theorem 3: one-step recovery

We derive the exact conditional gain and its leading term, then show that maximizing these gains recovers the oracle proportions. The allocation argument also allows a vanishing fraction of departures from exact lookahead, which will give the forced-exploration result in Section E.1.

D.3.1 Exact conditional gain

For a source request k, define the conditional expected decrease

$$\Delta_t^k = C(\mu_t) - \mathbb{E}[C(\mu_{t+1}) \mid \mathcal{D}_t, k], \tag{21}$$

where the expectation averages over the posterior predictive covariate and label returned by that request. Conditioning on $\mathcal{D}_t, k$ means the retained current state and the site-only action $D_t = k$. By Theorem 1, the exact conditional gain is

$$\Delta_t^k = \operatorname{tr}\big(\Sigma^{\mathrm{tgt}} \operatorname{Cov}[\bar{\psi}_{t+1} \mid \mathcal{D}_t, k]\big) = \mathbb{E}_{X \sim P_X^{\mathrm{tgt}}} \operatorname{Var}[X^\top \bar{\psi}_{t+1} \mid \mathcal{D}_t, k, X] \ge 0. \tag{22}$$

The covariance averages over the next source covariate and its label; the evaluation covariate X is independent of that acquisition. The proof below controls this gain directly, rather than differencing an expansion of the current objective.

D.3.2 Expansion of the one-step gain

Theorem 7 (Expansion of the conditional gain)**.** *Under the rank conditions in Assumption 1 and Assumption 2, every non-anticipating policy satisfies, $P_{\theta^\star}$-almost surely,*

$$(r_t^k)^2 \Delta_t^k \longrightarrow c^k \qquad \textit{for each } k. \tag{23}$$

Proof. Fix source k and a covariate x. Abbreviate its current covariance by $V = \widehat{V}_t^k$ and noise variance by $\sigma^2 = (\sigma^k)^2$. Gaussian updating gives

$$V^+ = V - A_x, \qquad A_x = \frac{V x x^\top V}{\sigma^2 + x^\top V x},$$
$$\widehat{\theta}_{t+1}^k = \widehat{\theta}_t^k + \frac{V x}{\sigma^2 + x^\top V x}(Y - x^\top \widehat{\theta}_t^k). \tag{24}$$

Under the current predictive distribution, $Y - x^\top \widehat{\theta}_t^k \sim N(0, \sigma^2 + x^\top V x)$. Thus the mean increment δ_x has distribution $N(0, A_x)$. All other posterior means and covariance blocks stay unchanged.

Use the localized map $\theta \mapsto \widetilde{\psi}_\theta$ from Section D.2.3. Let $\widehat{V}_{t+1}(x)$ be $\widehat{V}_t$ with block k replaced by $V - A_x$, and set

$$F_x(v) = \mathbb{E}_{\eta \sim N(0, \widehat{V}_{t+1}(x))}[\widetilde{\psi}_{v+\eta}].$$

The updated posterior mean of $\widetilde{\psi}_\theta$ is $F_x(\widehat{\theta}_t + \delta_x^{[k]})$, where the superscript means insertion into block k. Write $\partial_k F_x$ for the Jacobian with respect to that block, so its limit is the source Jacobian J^k defined in Section 4. Differentiating under the Gaussian integral is justified by the bounded derivatives of $\widetilde{\psi}_\theta$. Taylor expansion gives

$$F_x(\widehat{\theta}_t + \delta_x^{[k]}) = F_x(\widehat{\theta}_t) + \partial_k F_x(\widehat{\theta}_t)\delta_x + R_x, \qquad \|R_x\| \le c_1 \|\delta_x\|^2,$$

with c_1 independent of x, t. Since $\widehat{V}_{t+1}(x) \preceq \widehat{V}_t$ and the first derivative is Lipschitz,

$$\sup_x \left\| \partial_k F_x(\widehat{\theta}_t) - J^k \right\| \le c_1 \left(\left\| \widehat{\theta}_t - \theta^\star \right\| + \sqrt{\operatorname{tr} \widehat{V}_t} \right) \longrightarrow 0. \tag{25}$$

The Gaussian fourth moment obeys $\mathbb{E}\|\delta_x\|^4 \le 3(\operatorname{tr} A_x)^2$. Expanding the covariance of the Taylor expression and bounding the linear–remainder terms by Cauchy–Schwarz yields

$$\begin{aligned} \operatorname{Cov}_Y[F_x(\widehat{\theta}_t + \delta_x^{[k]})] &= \partial_k F_x(\widehat{\theta}_t) A_x \partial_k F_x(\widehat{\theta}_t)^\top + E_x, \\ \|E_x\| &\le c_2\{(\operatorname{tr} A_x)^{3/2} + (\operatorname{tr} A_x)^2\}. \end{aligned} \tag{26}$$

The bound uses centered covariances; a nonzero mean of R_x does not change it, since its variance is bounded by its second moment.

Now average over the acquisition covariate $X \sim N(0, \Sigma^k)$. The denominator in A_X is at least σ^2, so

$$\operatorname{tr} A_X \le \sigma^{-2}\|V\|^2 \|X\|^2, \quad \mathbb{E}(\operatorname{tr} A_X)^{3/2} = O((r_t^k)^{-3}), \quad \mathbb{E}(\operatorname{tr} A_X)^2 = O((r_t^k)^{-4}).$$

Also,

$$\mathbb{E}_X A_X = \sigma^{-2} V \Sigma^k V + O(\|V\|^3). \tag{27}$$

To verify the remainder, use $|1/(\sigma^2 + x^\top V x) - 1/\sigma^2| \le \sigma^{-4}\|V\|\|x\|^2$ and $\|Vxx^\top V\| \le \|V\|^2\|x\|^2$, then integrate the finite fourth Gaussian moment. By (15),

$$(r_t^k)^2 \mathbb{E}_X A_X \longrightarrow (\sigma^k)^2 (\Sigma^k)^{-1}.$$

The posterior tower identity gives $\mathbb{E}_Y[F_x(\widehat{\theta}_t + \delta_x^{[k]}) \mid x] = \mathbb{E}_{\mu_t}\widetilde{\psi}_\theta$, independent of x. Consequently the covariance over the whole pair $Z = (X, Y)$ is the average over X of the conditional covariance in (26); there is no extra covariance-of-means term from X.

Combining (25)–(27) and restoring ψ_θ with the exponential error bound proves

$$\operatorname{Cov}[\bar{\psi}_{t+1} \mid \mathcal{D}_t, k] = \frac{J^k (\sigma^k)^2 (\Sigma^k)^{-1} (J^k)^\top}{(r_t^k)^2} + o((r_t^k)^{-2}). \tag{28}$$

Taking the trace against Σ^{tgt} and using Theorem 1 proves (23). There are only finitely many sources, so the convergence holds for all of them on the same sample path. □

This expansion controls the gain directly at the $(r_t^k)^{-2}$ scale. Differencing a risk expansion with an $o(n_t^{-1})$ remainder would not justify it. As in Proposition 6, the allocation coefficients need not be strictly positive for the expansion itself.

D.3.3 From marginal gains to the oracle allocation

Lemma 8 (Allocation from relative marginal gains)**.** *Consider K counts r_t^k with one count increasing by one per round, and fixed positive numbers c^k. Suppose that all counts diverge, that only $o(n_t)$ rounds are exceptional, and that at every nonexceptional round a source maximizing scores d_t^k is selected. If $(r_t^k)^2 d_t^k / c^k \to 1$ for every k, then*

$$r_t^k / n_t \to \sqrt{c^k} / \sum_j \sqrt{c^j}.$$

Proof. Fix $0 < \epsilon < 1$. After a finite time, all scores satisfy

$$(1-\epsilon)\frac{c^k}{(r_t^k)^2} \le d_t^k \le (1+\epsilon)\frac{c^k}{(r_t^k)^2}.$$

Set $z_t^k = r_t^k/\sqrt{c^k}$ and $\rho = \sqrt{(1+\epsilon)/(1-\epsilon)}$. At a nonexceptional selection of k, score maximality implies $z_t^k \le \rho \min_j z_t^j$ before the increment.

For each k, look back from a large time t to its last nonexceptional selection after the finite threshold. At that selection, the inequality just displayed holds. All subsequent increases of this source are exceptional, and their total number is $o(n_t)$. The minimum of the z_t^j is nondecreasing. The last nonexceptional increment contributes at most $1/\sqrt{c^k}$. If no such selection occurred, its count is a fixed initial value plus $o(n_t)$. Since K is finite, these observations give

$$\max_k z_t^k \le \rho \min_k z_t^k + o(n_t).$$

Writing $s = \sum_k \sqrt{c^k}$ and using $n_t = \sum_k \sqrt{c^k} z_t^k$, we obtain

$$\frac{1}{\rho s} - o(1) \le \frac{z_t^k}{n_t} \le \frac{\rho}{s} + o(1).$$

Let $t \to \infty$ and then $\epsilon \downarrow 0$. This proves the claimed proportions for every k. □

Proof of Theorem 3. Assumption 2 ensures that every source count diverges and that Proposition 6 and Theorem 7 apply. For any non-anticipating policy satisfying this condition, (18) holds on a probability-one event. Let $c_{\min} = \min_k c^k > 0$. Since $\sum_k c^k/r_t^k \ge c_{\min} \sum_k 1/r_t^k$, the expansion can be written as

$$C(\mu_t) = \left(\sum_k \frac{c^k}{r_t^k} \right) (1 + o(1)).$$

Cauchy–Schwarz yields, for every positive count vector,

$$\left(\sum_k r_t^k \right) \left(\sum_k \frac{c^k}{r_t^k} \right) \ge \left(\sum_k \sqrt{c^k} \right)^2.$$

The preceding multiplicative approximation proves the liminf bound, without assuming that the competing proportions converge. This is a pathwise statement; no limit is passed through an expectation.

For pure lookahead, positivity of all c^k and Theorem 7 give $(r_t^k)^2 \Delta_t^k / c^k \to 1$. The policy maximizes Δ_t^k at every round, so Lemma 8 applies with no exceptional rounds and proves the limiting proportions. They are all strictly positive. Substitution into (18) gives

$$n_t C(\mu_t) \longrightarrow \sum_k \frac{c^k}{\alpha^{k,\star}} = \left(\sum_k \sqrt{c^k} \right)^2.$$

For a common fixed batch size, the lower-bound argument is unchanged because it uses observation counts. Corollary 11 proves the corresponding gain expansion and applies the same allocation argument in batch rounds, giving the stated limits in observation units. □

The result is pathwise at fixed true parameters. It establishes the leading constant of posterior predictor variance itself, rather than only a component of another criterion. It does not pass an almost-sure limit through a Bayesian expectation, prove finite-horizon optimality, or establish first-order optimality for true target MSE. The latter also depends on the error of the target fitting rule.

E Additional Gaussian results

The following results give an explicit exploration schedule, extend the analysis to batch acquisitions, and examine the roles of target covariates and the predictor class.

E.1 Ensuring sufficient sampling by forced exploration

Theorem 3 assumes that pure lookahead samples every source sufficiently often. Assumption 2 allows source proportions to tend to zero and does not specify the oracle allocation. A full-rank initial design identifies directions within a source, but does not ensure that the policy continues sampling it. We ensure coverage by reserving a vanishing fraction of rounds. Number acquisitions after initialization by $t = 1, 2, \ldots$. Fix $0 < \gamma < 1$ and choose an integer j_0 so large that $K\lceil 2^{j\gamma}\rceil < 2^j$ for all $j \geq j_0$. At the start of each block $[2^j, 2^{j+1})$, reserve $K\lceil 2^{j\gamma}\rceil$ rounds and take $\lceil 2^{j\gamma}\rceil$ labeled observations from each source in a fixed order. Use exact one-step lookahead at every other round, including rounds before 2^{j_0}.

Lemma 9 (Coverage from forced exploration)**.** *This schedule has $O(n_t^\gamma) = o(n_t)$ reserved acquisitions through time t and, for a constant $a > 0$, guarantees $\min_k r_t^k \geq a n_t^\gamma$ for all sufficiently large t.*

Proof of Lemma 9. At time $2^j \leq t < 2^{j+1}$ with $j > j_0$, the preceding block is complete. Each source has therefore received at least $\lceil 2^{(j-1)\gamma}\rceil$ reserved observations. This is bounded below by a fixed positive multiple of t^γ. Finite initialization changes n_t from t by only a constant. The number of reserved acquisitions through time t is at most

$$K \sum_{l=j_0}^{j} \lceil 2^{l\gamma}\rceil = O(2^{j\gamma}) = O(t^\gamma),$$

where the geometric sum dominates the extra $O(j)$ terms from rounding. Blocks do not overlap by the choice of j_0. Since $\gamma < 1$, the reserved fraction tends to zero. □

The schedule verifies Assumption 2 without assuming that pure lookahead discovers every source. Adaptive sampling rules enforcing comparable lower counts and departing from exact lookahead on only $o(n_t)$ rounds give the same allocation proof.

Corollary 10 (Vanishing forced exploration)**.** *Under Assumption 1, reserve the deterministic $O(n_t^\gamma)$ rounds specified above to sample every source, and use exact one-step lookahead otherwise. This policy ensures Assumption 2 and satisfies the limits in Theorem 3.*

Proof of Corollary 10. Lemma 9 establishes Assumption 2 for this policy and shows that only $o(n_t)$ rounds are forced. Thus Theorem 7 gives $(r_t^k)^2\Delta_t^k/c^k \to 1$ almost surely. At all nonforced rounds, the policy maximizes these gains. Lemma 8, with the forced rounds treated as exceptional, proves the proportions in Theorem 3. Substitution into (18) gives the same limit for $n_t C(\mu_t)$. The oracle lower bound in Theorem 3 holds for every policy satisfying Assumption 2, so it also provides the comparison for this policy. □

The reserved fraction tends to zero, so the leading allocation is unchanged.

E.2 Common batch size

Corollary 11 (A common fixed batch size)**.** *Suppose every acquisition returns $b \geq 1$ iid observations from the selected source, with b fixed and a coherent joint predictive update. Keep r_t^k as the number of observations, rather than the number of batches. Under the hypotheses of Theorem 7,*

$$(r_t^k)^2\Delta_{t,b}^k \longrightarrow bc^k \qquad P_{\theta^\star}\text{-almost surely.}$$

Under positive allocation coefficients, pure one-step lookahead satisfying Assumption 2 has the limiting allocation proportions and $n_t C(\mu_t)$ in Theorem 3. The same limits hold when the forced-exploration schedule of Corollary 10 is applied in batch rounds. In round units, $tC(\mu_t) \to c^\star/b$.

Proof of Corollary 11. For a batch from source k, write $\mathbf{X} = \mathbf{X}_D^k$ for the $b \times d$ design matrix, $V = \widehat{V}_t^k$, and $\sigma^2 = (\sigma^k)^2$. Conditional on this design, the posterior covariance reduction and the predictive covariance of its mean increment both equal

$$A_{\mathbf{X}} = V\mathbf{X}^\top(\sigma^2 I_b + \mathbf{X}V\mathbf{X}^\top)^{-1}\mathbf{X}V.$$

The increment is centered Gaussian under the coherent joint predictive law. The posterior covariance after the batch is $V - A_{\mathbf{X}}$, independently of its labels. These are the matrix forms of (24). Since the inverse in this display is bounded above by $\sigma^{-2} I_b$,

$$\operatorname{tr} A_{\mathbf{X}} \le \sigma^{-2} \|V\|^2 \|\mathbf{X}\|_{\mathrm{F}}^2, \qquad \mathbb{E} A_{\mathbf{X}} = b\sigma^{-2} V \Sigma^k V + O(\|V\|^3).$$

Here $\|\mathbf{X}\|_{\mathrm{F}}$ denotes the Frobenius norm of this batch design. For the second bound, the resolvent identity bounds

$$\left\| (\sigma^2 I_b + \mathbf{X} V \mathbf{X}^\top)^{-1} - \sigma^{-2} I_b \right\| \le \sigma^{-4} \|V\| \|\mathbf{X}\|^2 .$$

Sandwiching by $V\mathbf{X}^\top$ and $\mathbf{X}V$ and integrating the finite fourth moment proves the remainder; $\mathbb{E}\mathbf{X}^\top\mathbf{X} = b\Sigma^k$ gives the leading term. The third and fourth moment bounds for the Taylor remainder in (26) remain $O((r_t^k)^{-3})$ and $O((r_t^k)^{-4})$, since b is fixed. Localization and derivative convergence are unchanged. Repeating the conditional covariance proof therefore multiplies the leading matrix in (28) by b.

If s_t^k is the number of batches after initialization, then $r_t^k = b s_t^k + O(1)$. Thus $(s_t^k)^2 \Delta_{t,b}^k \to c^k / b$. For pure lookahead, Assumption 2 in observation counts also ensures polynomial coverage in batch counts, since $n_t = bt + O(1)$. Apply Lemma 8 with no exceptional rounds. For the policy with forced exploration, Lemma 9 ensures polynomial coverage and $o(t)$ forced rounds in batch units, so the same lemma applies. The common factor $1/b$ cancels from normalized square roots, so the limiting allocation proportions are unchanged. The posterior variance expansion is still in observation counts. Since $n_t = bt + O(1)$, its limiting constant in round units is divided by b. □

E.3 Positive gains and uninformative predictors

Proposition 12 (Nonnegative gains and zero uncertainty)**.** *Under the rank conditions in Assumption 1, every exact acquisition gain is nonnegative. Since $\Sigma^{\mathrm{tgt}} \succ 0$, a gain is zero exactly when the updated posterior mean coefficient is almost surely constant under that acquisition's predictive law. If $K = d$ and $\theta^\star$ has full rank, the profiled predictor has zero posterior variance and zero gains at every finite Gaussian posterior state.*

Proof. The first two claims follow from (22), because the trace of $\Sigma^{\mathrm{tgt}} H$ for $H \succeq 0$ vanishes exactly when $H = 0$. For $K = d$, each finite source posterior is nondegenerate Gaussian, so θ is invertible almost surely. Its span is all of $\mathbb{R}^d$. The fitted coefficient is therefore the same target least-squares coefficient $((\mathbf{X}^{\mathrm{tgt}})^\top \mathbf{X}^{\mathrm{tgt}})^{-1} (\mathbf{X}^{\mathrm{tgt}})^\top \mathbf{y}^{\mathrm{tgt}}$ for almost every draw. Its covariance is zero before and after every acquisition. □

More generally, the definition of c^k gives $c^k > 0$ exactly when $J^k \neq 0$. The nondegeneracy assumption excludes sources with zero first-order influence on the fitted predictor. It does not assert that pure lookahead ensures sufficient coverage from every initial state.

E.4 A fixed target integration grid

Proposition 13 (Dependence on target covariates)**.** *Fix the target fitting sample and source posterior. For a positive semidefinite integration matrix S, define $C_S(\mu) = \operatorname{tr}(S \operatorname{Cov}_\mu[\psi_\theta])$. Then the exact acquisition gain is $\Delta_t^k(S) = \operatorname{tr}(S H_t^k)$, where $H_t^k = \operatorname{Cov}[\bar{\psi}_{t+1} \mid \mathcal{D}_t, k] \succeq 0$. For any two such matrices,*

$$|\Delta_t^k(S_1) - \Delta_t^k(S_2)| \le \|S_1 - S_2\| \, \frac{\|\mathbf{y}^{\mathrm{tgt}}\|^2}{s_{\min}(\mathbf{X}^{\mathrm{tgt}})^2}.$$

For an iid unlabeled target grid $\{u_j\}_{j=1}^{N_U}$ with finite second moment, $S_{N_U} = N_U^{-1} \sum_j u_j u_j^\top \to \Sigma^{\mathrm{tgt}}$ almost surely. Its exact gains consequently converge to the population-integrated gains as $N_U \to \infty$.

Proof. Total covariance proves the gain identity for every fixed S. Since $H_t^k \succeq 0$, $|\operatorname{tr}[(S_1 - S_2) H_t^k]| \le \|S_1 - S_2\| \operatorname{tr} H_t^k$. Also $H_t^k \preceq W_t$ and Proposition 5 gives $\operatorname{tr} W_t \le \mathbb{E}_{\mu_t} \|\psi_\theta\|^2 \le \|\mathbf{y}^{\mathrm{tgt}}\|^2 / s_{\min}(\mathbf{X}^{\mathrm{tgt}})^2$. This proves the bound. The strong law applied entrywise to the fixed finite-dimensional matrix $u_j u_j^\top$ proves the grid convergence. □

A grid with deficient rank may fail to detect uncertainty in some prediction directions. Using the labeled target covariates as the grid does not itself force zero gain: with a full-rank target design, $S = \widehat{\Sigma}^{\mathrm{tgt}} \succ 0$, so gains vanish only when the updated posterior mean does not change.

F Implementation details

F.1 Simulated rollouts for site selection

We focus on actions that select a source: $D = k$ requests b observations from source k, and $\mathcal{A}_t \subseteq \{1, \dots, K\}$ contains the currently available sources. For one simulated batch under iid source sampling, draw

$$\widetilde{\theta} \sim \mu_t, \qquad \widetilde{x}_j \overset{\mathrm{iid}}{\sim} P_X^k, \qquad \widetilde{y}_j = f_{\widetilde{\theta}^k}(\widetilde{x}_j) + \widetilde{\varepsilon}_j, \quad j = 1, \dots, b.$$

The noise follows the selected source's model. All labels use the same draw $\widetilde{\theta}$, preserving their dependence through the unknown source function. For finite pools, replace the iid covariate draws by the specified batch sampling rule. If each source offers a fixed batch of covariates, retain those covariates in every simulation of that source. In both cases the learner chooses the site; the batch sampling rule determines its covariates.

Let $p_{\mu_t}(\cdot \mid D)$ denote the resulting joint law of the returned covariate–label batch Z_D. For a simulated batch $\widetilde{Z}_D$, add its observations to the selected source dataset and compute

$$\widetilde{\mu}_D = \mu(\cdot \mid \mathcal{D}_t \cup \widetilde{Z}_D).$$

The union updates the source identified by D. Write $F(\mu_t, D, Z_D)$ for the full state update, retaining the source data, available pools, and feasibility constraints. The target labels and target integration covariates stay fixed. Repeat from a copy of μ_t for each of the M simulated batches; simulated observations never enter another rollout or the actual dataset. With a predictive model, sample labels sequentially conditional on earlier simulated labels, as detailed in Section G.1.

Algorithm 1 One-step site selection (1-LAH)

Require: Initial belief μ_0, horizon T, batch size b, available sources, rollouts M
Require: Fixed target data $\mathcal{D}^{\mathrm{tgt}}$, target integration covariates U^{tgt}, prediction draws L
for $t = 0, \dots, T-1$ **do**
  for each feasible site request $D \in \mathcal{A}_t$ **do**
    for $r = 1, \dots, M$ **do**
      Simulate a covariate–label batch $\widetilde{Z}_D^{(r)} \sim p_{\mu_t}(\cdot \mid D)$ jointly
      $\widetilde{\mu}_D^{(r)} \leftarrow F(\mathrm{copy}(\mu_t), D, \widetilde{Z}_D^{(r)})$
      Construct ν at this updated state and evaluate $\widehat{C}(\widetilde{\mu}_D^{(r)})$ by sample variance over predictor draws, averaged over U^{tgt}
    end for
  end for
  Select D_t with the smallest average simulated criterion, using the fixed rule for ties
  Acquire the actual batch Z_{D_t} from source D_t
  $\mu_{t+1} \leftarrow F(\mu_t, D_t, Z_{D_t})$
end for
return Monte Carlo approximation to the mean predictor $\bar{\psi}_{\mathcal{D}_T}$

Choosing covariates as well. If the learner also chooses the covariates, the action is $D = (k, x_{1:b})$ and $\mathcal{A}_t$ contains feasible site–batch pairs. Hold the requested covariates fixed and simulate only their labels:

$$\widetilde{\theta} \sim \mu_t, \qquad \widetilde{y}_j = f_{\widetilde{\theta}^k}(x_j) + \widetilde{\varepsilon}_j, \quad j = 1, \dots, b.$$

The same update and Monte Carlo comparison apply, now minimizing over these site–batch pairs.

G Posterior-predictive formulation

The planner can represent uncertainty through distributions over unobserved source labels without maintaining an explicit parameter posterior. This requires a law for the observations an acquisition would return and an update after observing them. The resulting predictor distribution is defined in Section C.1; the label simulator and final prediction model may differ.

G.1 Predictive states and batch observations

Index the current predictive context by the observed source data $\mathcal{D}$, with available covariates and feasibility constraints retained but suppressed in the notation. We focus on site selection, where a request $D = k$ acquires a batch from source k. Let $p_{\mathcal{D}}(\cdot \mid D)$ be the joint law of its returned observations Z, including covariates and labels. First draw the batch covariates using the source sampling rule, or retain them if a fixed batch is offered at that source. Then draw their labels conditionally. After observing $Z = z$, add these observations to the relevant source datasets and update the predictive context. We write this as $\mathcal{D} \cup z$, with site identities and the corresponding pool update understood. Target labels and target integration covariates remain fixed.

Conditional on the resulting batch covariates $x_{1:b}$ at source k, sequential label sampling gives

$$p_{\mathcal{D}}(y_{1:b} \mid x_{1:b}, k) = \prod_{j=1}^{b} p\big(y_j \mid x_{1:b}, k, \mathcal{D}, y_{1:j-1}\big). \tag{29}$$

For a coherent predictive model, this is the chain rule for one joint distribution; updating the context conditions that same distribution on the observed labels. In the source models considered here, future batch covariates carry no information about the response parameters, so each factor can be evaluated from x_j and the preceding labeled observations $(x_{1:j-1}, y_{1:j-1})$. A parameter posterior gives an alternative joint sampler: draw the source parameters once, then generate all batch labels conditional on that draw.

If covariates are also chosen by the learner, use actions $D = (k, x_{1:b})$ and hold $x_{1:b}$ fixed. The conditional sampling rule (29) is unchanged; the feasible set now contains site–batch pairs, as in Section F.1.

Independent draws from predictive marginals generally omit the dependence induced by shared uncertain parameters. Correct marginals therefore do not suffice for correct batch lookahead. A learned predictive model may also depend on the order in which simulated labels are added to its context. Ordering, block size, and whether updates occur within a batch are part of its implementation specification.

H Experimental details and additional results

H.1 Experiment matrix

Table 3 separates the source generator, target response, posterior module, and evaluated predictor, and lists where each setting is reported. We use "linear mixture" for a weighted sum of source functions. The source functions themselves may be linear or neural. BNN refers to the saved experiments' frozen tanh features with a Bayesian last layer, not posterior inference over all network weights. RF denotes a random forest applied to source predictions. Changing the posterior or predictor tests robustness to that modeling choice; it does not by itself establish that the replacement model cannot represent the true response.

Every setting records uniform, DAD, and FG. The figures report FG with OLS source models in the linear-mixture settings (W1, W2, P1), FG with random-forest source models in the RF settings (F1, F2), and FG (TabPFN) alongside TabPFN lookahead with neural sources (P2, J1, J2) and in ACS; the ACS table also reports FG with OLS source models. The neural-source runs also record a fixed split, and the F1 and J1 runs a variant whose forest also receives the covariates; neither is reported.

ID	Sources	Target	Posterior	Evaluated predictor	Figures
W1	Linear	Linear mix.	BLR	Linear mix. of BLRs	4, 8
W2	Neural	Linear mix.	BNN	Linear mix. of BNNs	5, 9
P1	Linear	Linear mix.	TabPFN	Linear mix. of BLRs	4, 8†
P2	Neural	Linear mix.	TabPFN	Linear mix. of BNNs	5, 9†
F1	Neural	Linear mix.	BNN	RF	5, 9
F2	Neural	Neural mix.	BNN	RF	6, 10
J1	Neural	Linear mix.	TabPFN	RF	5, 9†
J2	Neural	Neural mix.	TabPFN	RF	6, 10†
R1	ACS	Observed labels	TabPFN	Linear mix. of BLRs	7†

Table 3: Experiment matrix. W, P, F, J, and R correspond to matched models, posterior substitution, predictor substitution, joint substitution, and real data. The first figure listed is in the main text and the second, in the other noise configuration, in Section H.6. †The saved TabPFN variants also change the functions used for acquisition scoring and the simulation budget.

H.2 Synthetic sources, target construction, and collection budgets

Covariates follow $N(0, I_{10})$ at all sources and the target, so the synthetic shifts concern response mechanisms. Linear-source coefficients are drawn from $N(0, 5I)$. Neural sources share the frozen feature map $\varphi(x) = \tanh(Wx + b)/\sqrt{32}$, with output coefficients drawn from $N(0, 100I)$. Target noise has variance $\sum_k w_k^2 \sigma_k^2$. Source-noise inference uses a 48-point logarithmic variance grid on $[0.005, 200]$.

For linear sources, the response is $f_{\theta^k}(x) = x^\top \theta^k$. For neural sources, it is $f_{\theta^k}(x) = \varphi(x)^\top \theta^k$, where φ is the shared frozen tanh feature map defined above. The BLR and BNN posteriors are Gaussian conditional on the source noise variance; the saved models account for unknown noise using a grid over its variance. The linear target constructions take $f_{\theta^{\mathrm{tgt}}}(x) = \sum_k \phi^{k,\star} f_{\theta^{k,\star}}(x)$. The true weights describe the target response, while fitted weights define the prediction model and allocation shares describe collection. These three quantities serve different purposes.

The linear-source reference file records target weights $(0.4, 0.3, 0.15, 0.1, 0.05)$. The neural-source reference with a linear-mixture predictor records $(0.5, 0.5, 0, 0, 0)$. In both reference files, source noise variances are $(0.1, 10, 1, 1, 1)$ in configuration A and $(10, 0.1, 1, 1, 1)$ in configuration B, so source 2 is noisier in A and source 1 in B. The RF runs share these generators, seeds, and configurations; their uniform allocations match the linear-mixture runs seed by seed. The neural-target RF experiment adds a width-16 tanh MLP of standardized contributing-source outputs to the linear mixture, scaled to match its standard deviation. The additional two-contributor BLR setting uses weights $(0.5, 0.5, 0, 0, 0)$ and the same noise configurations and budgets as the five-contributor setting. Synthetic test sets contain 2,000 observations; labeled target samples contain 100 observations for linear sources and 500 for neural sources.

Experiment	Sources	Initial/source	Batch size	Rounds	New observations
W1/P1: linear mixture	5	100	100	30	3,000
W2/P2: neural mixture	5	100	50	30	1,500
F1/J1: linear target, RF	5	100	50	30	1,500
F2/J2: neural target, RF	5	100	50	17	850
R1: ACS, BLR mixture	8	150	50	17	850

Table 4: Collection budgets. Each synthetic setting has four seeds (0–3); ACS uses seed 0 with 150 initial observations per source. Initialization is excluded from the final allocation shares and from the additional-observation counts.

H.3 Baseline definitions

Uniform. Uniform samples one of the K sources with probability $1/K$ at each round, so its realized allocation varies across seeds.

Frequentist greedy. The recorded FG implementations use five retrospective splits. The repository implementation withholds a common number $h = \min\{b, \min_k \lfloor |\mathcal{D}_t^k|/2 \rfloor\}$ of acquired observations from each source. It fits source models to the reduced data, then restores source k's withheld observations and measures the improvement on target validation data. It averages the improvement across splits and chooses the source with the largest value. The target sample is divided for fitting and validation within this retrospective calculation; this is a baseline implementation detail.

The FG variants differ in their source models, all fitted to raw observations without the generator's features: unregularized OLS with an intercept (FG (OLS)), a random forest with 100 trees (FG (RF)), or the TabPFN mean prediction with eight estimators (FG (TabPFN)). Where only one of the OLS and RF variants appears, figures and tables label it FG. In the linear-mixture runs FG combines source predictions with fitted simplex weights; in the RF runs it fits a forest to the source predictions. The source model changes FG's acquisition choices even when the final evaluated predictor is shared. The current code uses random tie breaking and has an optional exploration probability. Experiment seeds are listed in Table 4; the exported tables do not specify the FG exploration probability or a separate tie-breaking seed.

Oracle allocation. The linear- and neural-source mixture runs (W1, W2) contain an oracle allocation: the fixed allocation that minimizes a variance proxy given the true source weights and noise variances. The linear-mixture figures show it as a black loss curve and a leading allocation bar, and the rounds tables list it as Oracle; the RF runs have no oracle. The oracle is optimal for the variance proxy, not necessarily for realized target MSE, so its loss is not a lower bound for the policies: in Table 6, FG ends below it. It has not been recomputed using the fitted-predictor Jacobians defined in Section 4.

DAD. DAD implements the source-selection heuristic of Shen et al. (2024). At every round it fits a logistic classifier that separates each source's standardized observations (x, y) from the target sample, scores the source by this separation, and acquires a batch from the closest source, re-scoring after each acquisition. It uses no posterior simulation and does not model source noise.

H.4 ACS POPULATIONS AND PROTOCOL

The recorded ACS sources cross four states with two age groups (Table 5). Each acquisition selects one population and obtains 50 additional observations. The evaluated predictor is a simplex mixture of per-source hidden-noise BLRs. TabPFN lookahead plans with a simplex mixture of per-source TabPFN predictors; FG (TabPFN) and FG (OLS) fit TabPFN or OLS source models to raw observations, and DAD scores sources with the classifier described above.

Site	State	Age	Site	State	Age
1	California	17–29	5	California	45–64
2	Texas	17–29	6	Texas	45–64
3	New York	17–29	7	New York	45–64
4	Florida	17–29	8	Florida	45–64

Table 5: ACS source names from the candidate-score records. Figure and text indices are one-based; saved CSV source indices are zero-based.

The outcome is inflation-adjusted log personal income from 2018 PUMS, using the ACSIncome adult filter and 62 features fixed on 2017 data. The target is an equal mixture of California and New York residents aged 45–64. Survey weights define the target population and test MSE. Households are assigned to disjoint source, calibration, and test blocks. Each source starts with 150 labels; the labeled target and test samples contain 500 and 5,000 observations, respectively. Evaluation uses 200 per-source BLR draws. BLR specifies the predictor, not a generative assumption about observed ACS outcomes. Only seed 0 is reported, so the ACS comparison has no across-run uncertainty estimate.

H.5 AGGREGATION, PLOTS, AND DATA PROVENANCE

All curves are redrawn from numerical tables, without averaging across rounds or reconstructing values from images. At each round, a synthetic loss curve shows the mean over four seeds and a band

of one standard error, $\mathrm{sd}(L_1, \ldots, L_4)/\sqrt{4}$. The bands are pointwise, not simultaneous confidence bands. Loss axes are fitted to the mean curves, so bands can extend beyond the plotted range. The rounds tables report the same mean and standard error, rounded to two and one decimals; the linear-source tables use three decimals for both, because their differences are smaller. The standard errors include seed-to-seed variation shared by all policies, so they overstate the uncertainty of differences between policies. Each setting is summarized separately; the four seeds are its replicate units. No significance test is claimed.

Final allocation bars stack the mean fraction of additional observations assigned to each source, from source 1 at the bottom; initial observations are excluded. Segments of at least 30% show their rounded value. Loss curves start after the first acquisition, since initial losses are not included in the exported tables. Each curve passes through the mean at every recorded round; monotone cubic (PCHIP) interpolation joins consecutive rounds without overshooting them. Markers show the first round, every fifth round, and the last round.

Each main-text figure uses one configuration, selected from the two available for the larger FG–OSL gap in mean target MSE averaged over acquisition rounds. This gives configuration A for the linear-source comparison and configuration B for the other synthetic figures. This selection favors OSL and is not an independent evaluation of its advantage. The appendix shows the other configuration for each of these comparisons, with the same rounds tables. The configuration mapping used to produce the figures is retained with the plotting code.

Source snapshots. The paper includes the original exported tables under `data/experiments/`, with source paths and SHA-256 hashes in `data/experiments/manifest.json`. The snapshot directories map to the following folders in `distshift-uq`:

- `blr`, `blr_tabpfn`: `dgp-blr-5-domains;post-blr;pred-blr` and `dgp-blr-5-domains;post-tabpfn;pred-blr`.
- `bnn`, `bnn_tabpfn`: `dgp-bnn;post-bnn;pred-bnn` and `dgp-bnn;post-tabpfn;pred-bnn`.
- `rf`, `rf_tabpfn`: `dgp-bnn;post-bnn;pred-rf` and `dgp-bnn;post-tabpfn;pred-rf`.
- `mlp_rf`, `mlp_rf_tabpfn`: `dgp-bnnmlp;post-bnn;pred-rf` and `dgp-bnnmlp;post-tabpfn;pred-rf`.
- `acs`: `dgp-acs-blr;post-tabpfn-blr;pred-blr`.

The plotting script `scripts/build_experiment_figures.py` validates unique records, complete acquisition paths, budget totals, agreement between final and round-by-round losses, and consistency between choices and count increments. The snapshots are unmodified copies of the exported tables. Uniform, DAD, oracle, and fixed-split records are identical across the matched and TabPFN runs of each generator. Figures are vector PDFs generated with PGFPlots. The snapshots support reproducing the reported figures, while the generator specifications, planning settings, and ACS protocol are described in Sections H.2 and H.4.

H.6 Additional experimental results

We give selected-round results for the main-text figures and report the other configuration for each synthetic comparison. Plots use the conventions in Section H.5: loss means and pointwise standard errors over four seeds, and stacked mean shares of additional observations by source. FG (TabPFN) is drawn in a lighter blue than FG. Tables report target MSE at selected rounds as mean $\pm$ standard error over four seeds. No missing policy is assigned a synthetic result. The linear-source comparison of Figure 4 and Table 6 appears for configuration B in Figure 8 and Table 8. The neural-source, linear-target settings with the linear-mixture or RF predictor appear in Figure 5 and Table 1 for configuration B and in Figure 9 and Table 9 for configuration A, and the neural-target comparison of Figure 6 and Table 7 appears for configuration A in Figure 10 and Table 10.

H.6.1 Main-text comparisons at selected rounds

Method	Round 10	Round 20	Round 30
OSL	$\mathbf{0.980} \pm 0.006$	$\mathbf{0.977} \pm 0.005$	0.974 ± 0.005
OSL (TabPFN)	0.983 ± 0.006	0.982 ± 0.010	0.976 ± 0.008
FG	0.991 ± 0.008	0.981 ± 0.006	$\mathbf{0.973} \pm 0.005$
Uniform	0.989 ± 0.014	0.986 ± 0.012	0.976 ± 0.007
DAD	1.009 ± 0.013	1.008 ± 0.013	1.008 ± 0.015
Oracle	0.981 ± 0.006	0.978 ± 0.005	0.976 ± 0.007

Table 6: Target MSE for Figure 4 (mean $\pm$ standard error, four seeds). Bold marks the lowest rounded mean, excluding Oracle.

Method	Round 5	Round 10	Round 15
OSL	$\mathbf{4.89} \pm 0.2$	$\mathbf{4.47} \pm 0.1$	$\mathbf{4.33} \pm 0.2$
OSL (TabPFN)	4.96 ± 0.2	4.90 ± 0.2	4.83 ± 0.1
FG	5.46 ± 0.4	5.17 ± 0.4	5.21 ± 0.4
FG (TabPFN)	5.10 ± 0.2	4.97 ± 0.2	4.80 ± 0.2
Uniform	5.61 ± 0.1	5.20 ± 0.1	5.07 ± 0.1
DAD	5.80 ± 0.3	5.74 ± 0.3	5.77 ± 0.3

Table 7: Target MSE for Figure 6 (mean $\pm$ standard error, four seeds). Bold marks the lowest rounded mean.

H.6.2 LINEAR SOURCES IN CONFIGURATION B

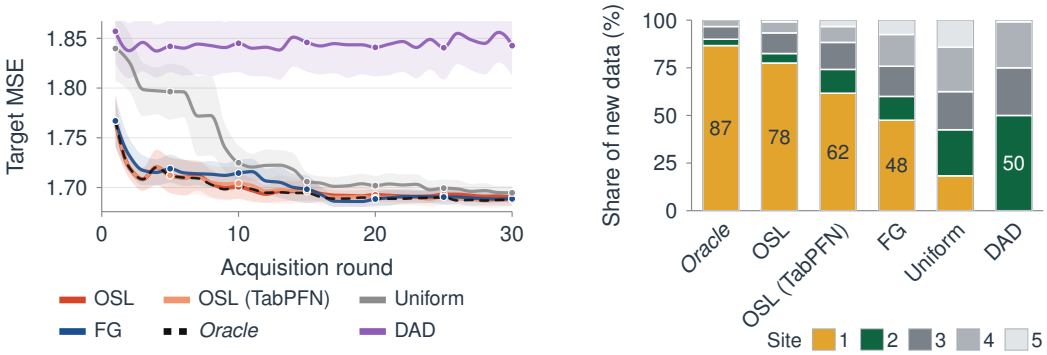

Figure 8: **Linear sources in configuration B.** The counterpart of Figure 4 for the configuration where site 1 is the noisier contributing source; panels follow that figure. Averaged over rounds, OSL and OSL (TabPFN) both reach 1.700, compared with 1.703 for FG and 1.732 for uniform. Final errors other than DAD's range from 1.688 to 1.695.

Method	Round 10	Round 20	Round 30
OSL	**1.701** $\pm$ 0.012	1.692 $\pm$ 0.007	1.690 $\pm$ 0.006
OSL (TabPFN)	1.704 $\pm$ 0.016	1.690 $\pm$ 0.008	**1.688** $\pm$ 0.007
FG	1.715 $\pm$ 0.013	**1.688** $\pm$ 0.007	1.689 $\pm$ 0.007
Uniform	1.725 $\pm$ 0.017	1.702 $\pm$ 0.010	1.695 $\pm$ 0.007
DAD	1.845 $\pm$ 0.029	1.841 $\pm$ 0.029	1.843 $\pm$ 0.031
Oracle	1.701 $\pm$ 0.012	1.689 $\pm$ 0.008	1.688 $\pm$ 0.009

Table 8: Target MSE at rounds 10, 20, and 30 for the runs in Figure 8, as mean $\pm$ standard error over four seeds, following Table 6.

H.6.3 Neural sources with a linear target in configuration A

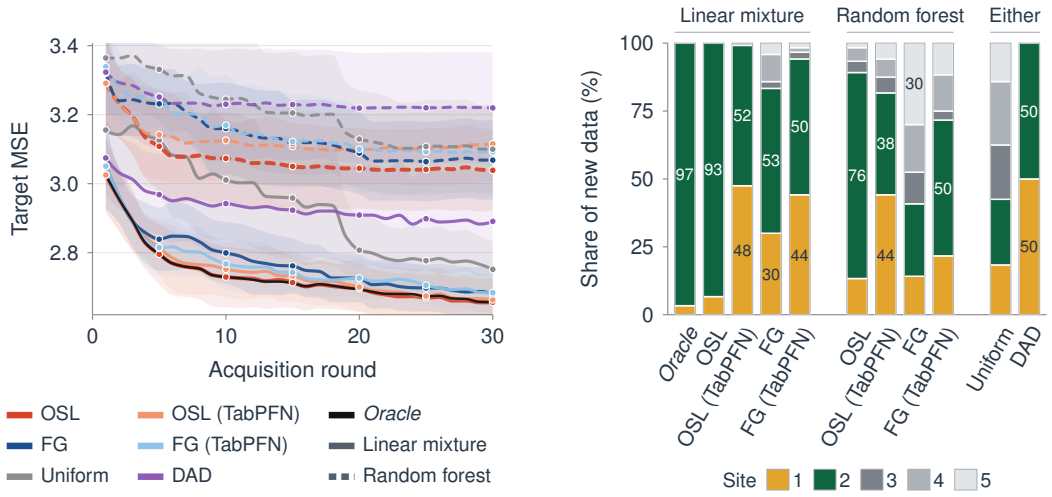

Figure 9: **Neural sources with a linear target in configuration A.** The counterpart of Figure 5 for configuration A, where site 2 is the noisier contributing source; panels follow that figure. Matched OSL has the lowest final MSE under both predictors. With the RF predictor, OSL (TabPFN) ends above both FG variants and uniform.

(a) Linear-mixture predictor

Method	Round 10	Round 20	Round 30
OSL	$\mathbf{2.73} \pm 0.1$	$\mathbf{2.70} \pm 0.1$	$\mathbf{2.66} \pm 0.1$
OSL (TabPFN)	2.75 ± 0.1	$\mathbf{2.70} \pm 0.1$	$\mathbf{2.66} \pm 0.1$
FG	2.80 ± 0.1	2.73 ± 0.1	2.68 ± 0.1
FG (TabPFN)	2.77 ± 0.1	2.73 ± 0.1	2.68 ± 0.1
Uniform	3.01 ± 0.2	2.81 ± 0.1	2.75 ± 0.1
DAD	2.94 ± 0.1	2.91 ± 0.1	2.89 ± 0.2
Oracle	2.73 ± 0.1	2.69 ± 0.1	2.66 ± 0.1

(b) RF predictor

Method	Round 10	Round 20	Round 30
OSL	$\mathbf{3.07} \pm 0.1$	$\mathbf{3.04} \pm 0.1$	$\mathbf{3.04} \pm 0.1$
OSL (TabPFN)	3.13 ± 0.1	3.10 ± 0.1	3.12 ± 0.1
FG	3.16 ± 0.1	3.09 ± 0.1	3.07 ± 0.1
FG (TabPFN)	3.17 ± 0.1	3.10 ± 0.1	3.09 ± 0.1
Uniform	3.24 ± 0.2	3.13 ± 0.1	3.10 ± 0.1
DAD	3.23 ± 0.2	3.22 ± 0.2	3.22 ± 0.2

Table 9: Target MSE at rounds 10, 20, and 30 for the runs in Figure 9, as mean $\pm$ standard error over four seeds, following Table 1.

H.6.4 Neural target combination and RF prediction

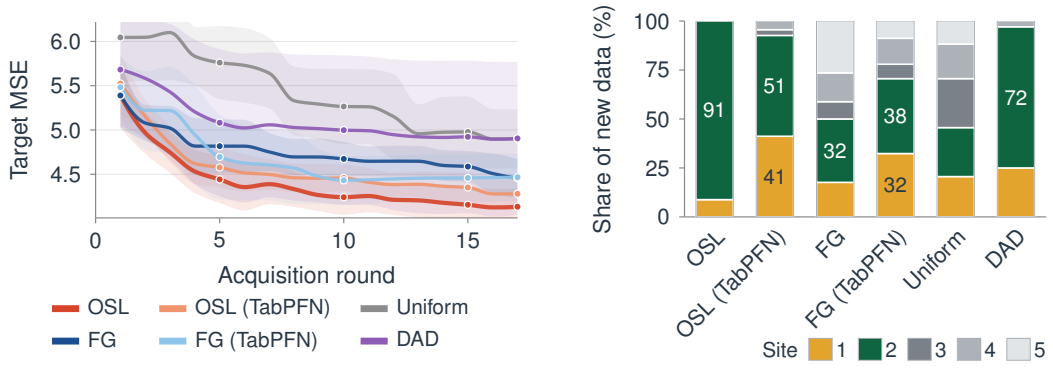

Figure 10: **Neural target with a random-forest predictor in configuration A.** The counterpart of Figure 6 for the configuration where site 2 is the noisier contributing source; panels follow that figure. Matched OSL has the lowest final MSE, followed by OSL (TabPFN).

Method	Round 5	Round 10	Round 15
OSL	$\mathbf{4.44} \pm 0.3$	$\mathbf{4.24} \pm 0.2$	$\mathbf{4.16} \pm 0.2$
OSL (TabPFN)	4.58 ± 0.2	4.46 ± 0.1	4.35 ± 0.2
FG	4.82 ± 0.3	4.67 ± 0.2	4.59 ± 0.2
FG (TabPFN)	4.70 ± 0.5	4.43 ± 0.3	4.46 ± 0.3
Uniform	5.76 ± 0.4	5.27 ± 0.4	4.98 ± 0.4
DAD	5.08 ± 0.8	5.00 ± 0.8	4.92 ± 0.9

Table 10: Target MSE at rounds 5, 10, and 15 for the runs in Figure 10, as mean $\pm$ standard error over four seeds, following Table 7.

H.6.5 ACS pilot

Method	Final loss	Difference from uniform
OSL (TabPFN)	0.6073	-0.0749
FG (TabPFN)	0.6320	-0.0501
Uniform	0.6822	0
DAD	0.6013	-0.0808
FG (OLS)	0.6533	-0.0289

Table 11: ACS final losses for the single available run, with 17 acquisitions per method. No standard error is reported because there are no independent replicate runs in the export.

The loss curves and all eight source allocations appear in Figure 7. OSL (TabPFN) allocates nine batches to site 5 and six to site 7. FG (TabPFN) spreads its batches, with five to site 5 and three each to sites 1 and 8; FG (OLS) allocates nine to site 8 and five to site 5. DAD allocates all 17 batches to site 7. The different choices are associated with different final losses in this run. They do not establish that one of these populations will be preferable across target samples.