ACQUISITION POLICIES CHANGE THE BAYES RULE: POLICY-CONDITIONED POSTERIOR CORRECTION FOR MULTIMODAL LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Multimodal models usually predict from available modalities as if an observed subset had fixed predictive meaning. We show that this can fail after an acquisition-policy shift. The same observed content and mask can imply a different class posterior. We formulate policy-conditioned multimodal learning and define the observed acquisition mass, the class-conditional probability that a policy produces an observed record after marginalizing unobserved content. Its ratio transforms a source posterior into a policy-specific posterior, while a posterior-weighted Jensen-Shannon divergence quantifies irreducible log-loss without policy information. Our Policy-Conditioned Observed-Mass Model (PCOM) combines a frozen source predictor, source-trained mass estimation, and policy inference from unlabeled context-mask logs. Record-level mixtures are corrected in mass space before normalization. Across CMU-MOSEI, MELD, MM-IMDb, and UPMC Food-101, PCOM improves endpoint log loss and retains positive mean gains at held-out mixture proportions. Ablations verify the roles of candidate-class conditioning and policy state, while stress tests show that state fidelity alone does not guarantee utility. Thus, policy variation should not be removed through invariance when it changes an observation's predictive meaning.

1 INTRODUCTION

Multimodal predictors are commonly designed to operate on whichever modalities are observed, implicitly treating the observed subset as a stable conditioning variable. In practice, however, availability is generated by a selection process. Sensors decide what to record, platforms what to retain, and institutions what to measure. Missingness therefore reflects not only absent content but also the acquisition mechanism that produced the observation (Rubin, 1976; Seaman et al., 2013). When this mechanism changes after training, the same observed content and mask can have different class-conditional likelihoods even under an unchanged complete-data distribution. We study prediction under such shifts as *Generalization to Unseen Acquisition Policies* (GUAP). Closely related non-ignorable missingness shifts can alter the Bayes predictor itself (Rockenschaub et al., 2024), so the challenge is not merely to process incomplete inputs, but to interpret them under the mechanism that made them observable.

This perspective exposes a limitation of treating missing-modality robustness purely as invariance to the observed subset. Existing methods improve incomplete-input prediction through modality dropping, modular fusion, recovery, or prompting (Neverova et al., 2016; Swamy et al., 2023; Lee et al., 2023), while robust objectives reduce dependence on unstable cues or environments (Sagawa et al., 2020; Geirhos et al., 2020). These approaches remain well motivated when the meaning of an observed subset is stable across environments. When the acquisition mechanism instead changes the class-conditional meaning of the same observation, however, one policy-agnostic prediction rule must compromise between policy-specific posteriors. In this regime, stronger optimization alone cannot recover the correct prediction without information about the observation process, which motivates a policy-conditioned rather than purely policy-invariant view of multimodal prediction.

Motivated by this distinction, we formulate GUAP through *policy-conditioned multimodal learning* and instantiate it with the Policy-Conditioned Observed-Mass Model (PCOM). PCOM separates a

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.

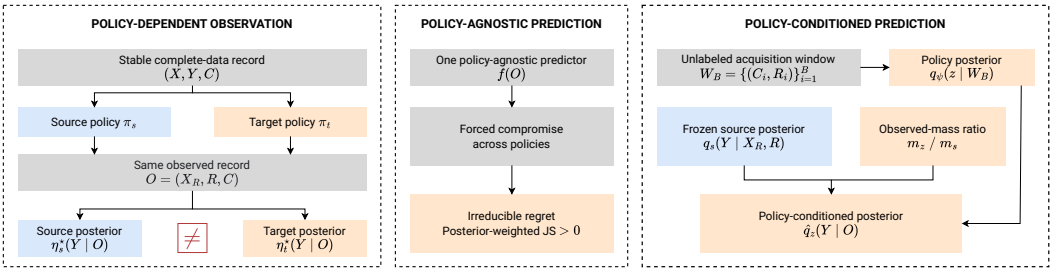

Figure 1: **Availability can have policy-dependent meaning.** The same observed record can require different posteriors under different acquisition policies. A policy-agnostic predictor must compromise.

frozen source prediction from the class-dependent acquisition effect that changes its interpretation, estimating observed acquisition mass from source data and inferring policy information from unlabeled context-mask logs. For record-level mixtures, it combines masses before posterior normalization. This formulation leads to four linked research questions (RQs). **RQ1** asks whether acquisition policy can change the correct prediction for the same observed record. **RQ2** asks whether the learned correction improves predictive log loss at learned endpoints and held-out mixture proportions. **RQ3** asks whether matching a learned state-known predictor is sufficient to improve on the source anchor. **RQ4** examines which components and acquisition-log relationships explain the observed gains and failures. We answer these questions through controlled posterior-conflict experiments, four non-toy multimodal benchmarks, mechanism ablations, interpolation studies, and stress tests. In summary, our contributions are threefold:

- We study GUAP as a policy-dependent prediction problem and formalize the mechanism by which acquisition shift changes the predictive posterior. We introduce class-conditional observed acquisition mass, derive the source-to-policy posterior correction, and characterize the irreducible log-loss incurred when one policy-agnostic predictor must serve conflicting policy-specific posteriors.
- We propose PCOM, a structured realization of policy-conditioned multimodal learning. PCOM combines a frozen source predictor with source-trained observed-mass estimation and policy inference from unlabeled acquisition logs, while distinguishing uncertainty about an active policy from record-level mixtures of acquisition mechanisms.
- We provide a comprehensive empirical study across four multimodal benchmarks. The results establish a predictive role for acquisition-policy information, demonstrate useful correction at learned endpoints and held-out mixture proportions, and show through ablations and stress tests when policy-state fidelity does and does not translate into downstream predictive utility.

2 When Acquisition Policies Change the Bayes Rule

We first examine how acquisition changes the meaning of an observed record. Separating the underlying data distribution from the policy that reveals its modalities yields a policy-specific posterior correction and explains why prediction without policy information can be intrinsically suboptimal. It also separates the value of policy information from learned-correction gains.

Separating content from acquisition. Let $X = (X^{(1)}, \ldots, X^{(M)})$, Y, and C denote complete multimodal content, the task label, and retained acquisition context. We hold the pre-acquisition law $P_0(X, Y, C)$ fixed across policy states to isolate acquisition-induced changes from changes in the underlying population. Policy state z generates a feasible mask through $R \sim \pi_z(\cdot \mid X, Y, C)$, leaving the observation $O = (X_r, r, C)$ when $R = r$. Dependence on Y represents class-dependent selection. In label-aware simulations, labels generate masks offline rather than serving as prediction-time inputs.

From acquisition to observed posterior. The complete-data policy may depend on content that the predictor cannot observe. We therefore average over that content to obtain the policy effect relevant to the observed record. For a fixed mask r, define the class-conditional *observed acquisition mass*

$$m_z(r \mid x_r, y, c) = \mathbb{E}_{P_0}[\pi_z(r \mid X, y, c) \mid X_r = x_r, Y = y, C = c]\,. \tag{1}$$

The expectation is taken before acquisition, rather than under the distribution already selected by $R = r$. Integrating out unobserved content then gives $p_z(o, y) = p_0(x_r, y, c) m_z(r \mid x_r, y, c)$,

separating the shared content distribution from the policy-specific selection effect. This factorization yields the following correction.

Proposition 1 (Policy-conditioned posterior)**.** *Suppose P_0 is invariant across policy states and $m_{z_s}(r \mid x_r, y, c) > 0$ for every class with positive target joint probability. Writing $\eta^\star_z(y \mid o) = p_z(y \mid o)$, we have*

$$\eta^\star_z(y \mid o) = \frac{\eta^\star_{z_s}(y \mid o)\, m_z(r \mid x_r, y, c)/m_{z_s}(r \mid x_r, y, c)}{\sum_{y'} \eta^\star_{z_s}(y' \mid o)\, m_z(r \mid x_r, y', c)/m_{z_s}(r \mid x_r, y', c)}. \tag{2}$$

The mass ratio adjusts each source class probability according to how the target policy changes the likelihood of the same observation. The shared content factor cancels in the joint source-to-target ratio, and normalization over classes gives Eq. (2). The full derivation and label-wise multilabel specialization appear in section B. This identity is the population target approximated by PCOM.

Why policy information matters. Having identified the policy-specific posterior, we next ask what is lost when a predictor cannot distinguish the policies that produced an observation. Under a specified policy prior and the resulting observation law, let $\rho_z(o) = P(Z = z \mid O = o)$ and $\eta^\star_0(\cdot \mid o) = \sum_z \rho_z(o)\eta^\star_z(\cdot \mid o)$. For any common prediction $q(\cdot \mid o)$, with D_{KL} denoting Kullback-Leibler divergence,

$$\sum_z \rho_z D_{\mathrm{KL}}(\eta^\star_z \| q) = \sum_z \rho_z D_{\mathrm{KL}}(\eta^\star_z \| \eta^\star_0) + D_{\mathrm{KL}}(\eta^\star_0 \| q). \tag{3}$$

The first term is a generalized Jensen-Shannon divergence (Lin, 1991), which is positive whenever positive-weight policy posteriors disagree. The second measures how far q is from the best predictor using the observation alone. Better optimization can eliminate the second term, but not the first. The weights are conditional on the observation, so equal global policy priors need not produce equal $\rho_z(o)$. Averaging the minimum excess loss over observations gives $I(Y; Z \mid O)$, the predictive information supplied by policy state beyond the observed record.

From population identities to measured gains. We distinguish these population quantities from empirical gains to evaluate what the learned correction actually achieves. Let $\mathcal{R}_z(q) = \mathbb{E}_{P_z}[-\log q(Y \mid O)]$, with $\mathcal{R}$ additionally averaging over the policy prior. The information value is $\mathcal{R}(\eta^\star_0) - \mathcal{R}(\eta^\star_Z) = I(Y; Z \mid O)$, whereas source-transfer headroom is $\mathcal{R}_z(\eta^\star_{z_s}) - \mathcal{R}_z(\eta^\star_z) = \mathbb{E}_{O \sim P_z} D_{\mathrm{KL}}(\eta^\star_z \| \eta^\star_{z_s})$. Both are nonnegative ideal-posterior contrasts. Our experiments compare learned predictors.

$$G^{\mathrm{known}}_z = \widehat{\mathcal{R}}_z(q_{\mathrm{anchor}}) - \widehat{\mathcal{R}}_z(\widehat{q}_z), \quad \text{and} \quad G^{\mathrm{inf}}_z = \widehat{\mathcal{R}}_z(q_{\mathrm{anchor}}) - \widehat{\mathcal{R}}_z(\widehat{p}). \tag{4}$$

Here $\widehat{q}_z$ receives the known state, $\widehat{p}$ infers it, and q_{anchor} is the frozen source predictor used as the empirical reference. Either gain can be negative because the learned correction may be inaccurate even when policy information is useful. Comparing the two gains therefore separates the cost of state inference from the utility of the state-conditioned predictor itself. We reserve *policy-agnostic shared predictor* for the separately trained comparator in RQ1. We next position this correction principle relative to existing missing-modality approaches before describing its implementation.

3 Related Work

Missing-data theory and graphical analyses establish when selection can be ignored or quantities identified from incomplete observations (Rubin, 1976; Mohan & Pearl, 2021). Ignorable shifts preserve the Bayes predictor under stated assumptions, whereas non-ignorable shifts may change it (Rockenschaub et al., 2024). Related adaptation methods use structured source-target models (Zhou et al., 2023) or reduce missing-not-at-random (MNAR) adaptation to imputation under additional assumptions (Stokes et al., 2025). These results motivate the policy-specific selection effect in Eq. (2).

For multimodal inputs, modality dropping, severe-missingness training, and fusion design improve incomplete-record prediction (Neverova et al., 2016; Ma et al., 2021; 2022). Alignment, recovery, prompting, shared/specific representations, and modular experts provide complementary strategies (Lin & Hu, 2023; Wang et al., 2023b; Lee et al., 2023; Guo et al., 2024; Hu et al., 2024; Dai et al., 2025; Wang et al., 2023a; Swamy et al., 2023; Yun et al., 2024; Han et al., 2024). We additionally ask whether changing the observation mechanism changes the appropriate posterior.

Informative-missingness models use observation patterns directly (Che et al., 2018; Liang et al., 2025). In Case You Multimodal Missed It (ICYM^{2}I) studies their effect on apparent multimodal

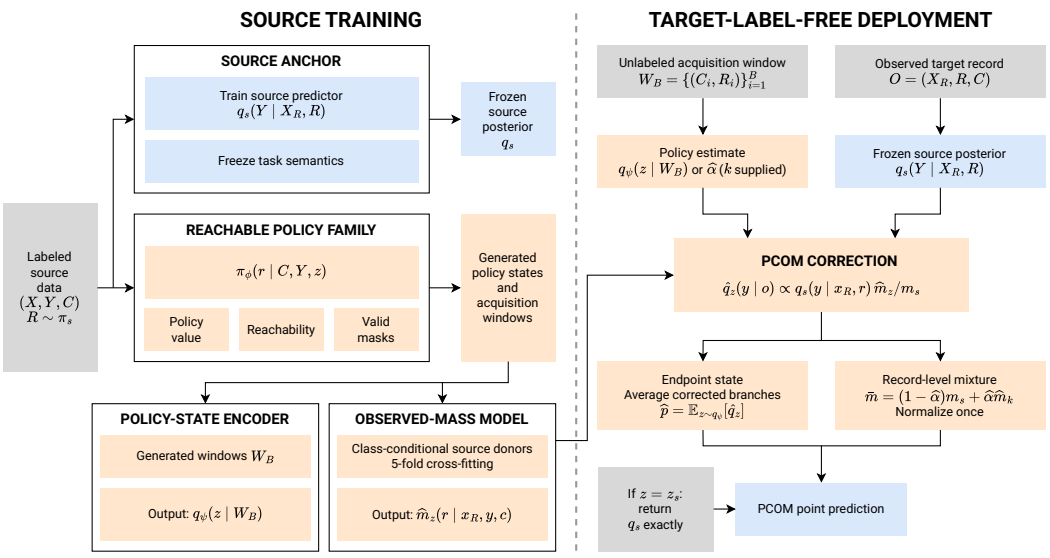

Figure 2: **Policy-conditioned correction with PCOM.** Source-generated environments train state and observed-mass estimators around a frozen predictor. Endpoint predictions average corrected branches. Mixtures infer their proportions from logs and combine masses before normalization.

informativeness (Choi et al., 2026), and MIRRAMS regularizes against unseen tabular missingness shifts without target data (Lee et al., 2025). PCOM instead infers acquisition state from unlabeled context-mask logs and estimates class-conditional observed masses to adjust a frozen source predictor. Broader comparisons and information-access details appear in sections F and F.1.

4 Policy-Conditioned Observed-Mass Correction

The posterior identity in Eq. (2) suggests separating source prediction from the acquisition effect that changes its interpretation. We implement this separation with PCOM, and Fig. 2 shows how these components form the correction.

Source anchor and learned policy family. We select the source anchor $q_s(y \mid x_r, r)$ by source-policy validation and freeze it to keep task learning separate from policy correction. The anchor pools embeddings of present modalities and includes the availability mask, but does not receive C. We use a uniform source policy over the main catalog of $A = 2^M - 1$ nonempty masks to give every acquisition action positive source probability. This choice also provides an analytic reference mass, $m_{z_s} = 1/A$, avoiding an additional learned denominator in the correction. To train the policy and mass estimators without target labels, we construct challenging acquisition environments from labeled source records. A generator learns eight endpoint codes and a zero source code:

$$\pi_\phi(a \mid C, Y, z) = \mathrm{softmax}_a[h_\phi(C, Y, z) - h_\phi(C, Y, 0)]\,. \tag{5}$$

The subtraction recovers the uniform source policy at $z = 0$. We increase expected frozen-anchor loss to expose policies under which the source prediction becomes less reliable. Action-marginal penalties limit changes in overall mask frequencies, path-overlap penalties encourage locally supported transitions, and a diversity penalty discourages redundant endpoints. These regularized environments provide source-only training data for the correction modules. Section C details their training.

Inferring state from retained logs. An individual availability pattern may be compatible with several policies, so we infer acquisition state from a collection of context-mask pairs. A permutation-invariant encoder (Zaheer et al., 2017) transforms each (C_i, R_i) and pools the feature means and standard deviations to represent the window's context-availability relationships. It then returns a nine-state distribution. To reduce sensitivity to the composition of a single sampled window, we average these distributions over 64 bootstrap windows, each containing $B = 128$ records drawn with replacement from the unlabeled target log pool. This fixed resampling budget produces the pool-level state weights $q_\psi(z \mid W)$ used throughout the correction.

Estimating observed mass. We approximate the expectation in Eq. (1) using complete source records whose observed features and context resemble the query. For a fixed query mask r, candidate class y, and policy fingerprint $\varphi(z)$, a donor estimator averages acquisition probabilities over $K = 16$ neighbors matching the candidate class or label bit. These averages provide soft targets for a mass model trained with categorical cross-entropy. The model predicts the full action-probability vector, while the correction reads only the realized-action coordinate. Thus, the query mask determines which content is used for matching, rather than restricting the model to a single output action.

To limit reuse of the same records for fitting and out-of-fold prediction, we split complete-record IDs into five folds and exclude each held-out fold from both fitting queries and donor pools. At target prediction, we average the five models rather than rely on one split. Section C specifies the shared preprocessing, record-level cross-fitting, and dataset-specific outer grouping rules.

Posterior correction. For the realized mask, let $\widehat{m}_z^y(o) = \widehat{m}_z(r \mid x_r, y, c)$. The estimated mass ratio adjusts the source probabilities according to the class-specific acquisition effect:

$$\begin{aligned} \widehat{q}_z(y \mid o) &= \mathrm{Normalize}_y \left[q_{\mathrm{s}}(y \mid x_r, r) \frac{\widehat{m}_z^y(o)}{m_{z_{\mathrm{s}}}^y(o)} \right], \\ \bar{p}(y \mid o, W) &= \sum_z q_\psi(z \mid W) \widehat{q}_z(y \mid o). \end{aligned} \tag{6}$$

We use stabilized masses and approximate the state average with $L = 16$ sampled branches, retaining uncertainty rather than committing to one state. Candidate-class dependence is essential as a multiplier shared by all classes cancels during normalization and exactly recovers the source prediction.

At $z = z_{\mathrm{s}}$, we return q_{s} directly so that mass-estimation error cannot alter the source-state branch. For the exact state average, total variation distance TV obeys $\mathrm{TV}(\bar{p}, q_{\mathrm{s}}) \leq 1 - q_\psi(z_{\mathrm{s}} \mid W)$. The sampled analogue uses its source-state fraction, so source-environment behavior also depends on inferred state. Because the learned anchor omits C and the masses are estimated, the update is a Bayes-derived *plug-in* correction. Section B.6 connects these approximations to prediction error. If the combined anchor and log-mass-ratio error is bounded by ϵ after removing a class-independent offset, the target-posterior KL is at most 2ϵ for each evaluated record.

Record-level mixtures require mass-space correction. Uncertainty about one active policy differs from a mechanism that mixes policies across records. For a known endpoint k, let $\pi_\alpha = (1-\alpha)\pi_{z_{\mathrm{s}}} + \alpha\pi_k$. Linearity of Eq. (1) determines the corresponding observed mass:

$$\begin{aligned} m_\alpha^y(o) &= (1-\alpha) m_{z_{\mathrm{s}}}^y(o) + \alpha m_k^y(o), \\ \widehat{q}_\alpha(y \mid o) &= \mathrm{Normalize}_y \left[q_{\mathrm{s}}(y \mid x_r, r) \frac{(1-\widehat{\alpha}) m_{z_{\mathrm{s}}}^y(o) + \widehat{\alpha}\, \widehat{m}_k^y(o)}{m_{z_{\mathrm{s}}}^y(o)} \right]. \end{aligned} \tag{7}$$

A source-trained scalar encoder for endpoint k estimates $\widehat{\alpha}$ from logs. Combining masses before normalization preserves the acquisition-mixture semantics. Averaging normalized endpoint posteriors generally does not, because their normalizers differ. Continuous latent paths, $z_\lambda = \lambda z_k$, are evaluated separately because a nonlinear generator need not map code interpolation to a mixture of acquisition mechanisms. The experiments therefore measure both predictive improvement and the accuracy with which inferred policies reproduce state-known predictions.

5 Experimental Setup

To answer the research questions introduced in section 1, we combine a controlled posterior-conflict study with non-toy evaluations of learned-policy correction. Endpoint and mixture experiments measure predictive gains, while continuous-path, hidden-content, and component studies separate state-inference fidelity from correction quality.

Data and information access. We evaluate binary, multiclass, and multilabel prediction using fixed processed features from CMU-MOSEI (Bagher Zadeh et al., 2018), MELD (Poria et al., 2019), MM-IMDb (Arevalo et al., 2017), and UPMC Food-101 (Wang et al., 2015). The tasks are binary sentiment after neutral exclusion, seven-class emotion, 23-label genre prediction, and 101-class recognition, respectively. This tests correction across distinct prediction tasks without coupling it to end-to-end feature learning. Section D details every dataset-specific protocol.

Table 1: **Controlled posterior conflict.** Mean ± sample SD over six seeds. This experiment is separate from the five-seed feature study reported below.

Method	Source NLL	Target NLL	Source AUROC	Target AUROC
Source-specific	0.4408 ± 0.0024	0.6814 ± 0.0061	0.8764 ± 0.0012	0.7090 ± 0.0061
Target-specific	0.6748 ± 0.0106	0.4411 ± 0.0016	0.7092 ± 0.0094	0.8760 ± 0.0007
Policy-agnostic shared	0.4957 ± 0.0039	0.4913 ± 0.0049	0.8381 ± 0.0028	0.8409 ± 0.0036
State-known	0.4407 ± 0.0019	0.4417 ± 0.0018	0.8764 ± 0.0011	0.8760 ± 0.0010
Inferred state	0.4516 ± 0.0038	0.4485 ± 0.0047	0.8696 ± 0.0025	0.8717 ± 0.0033

The policy log C contains retained acquisition metadata, modality-level summary statistics, source-fitted triage features, while the task predictor receives only the observed modalities and their availability mask. Label-aware generators use evaluation labels only to sample benchmark masks offline. Target labels are absent from adaptation, policy inference, checkpoint selection, and target-loss optimization. A separate hidden-content study uses public text context and label-blind policies.

Policy families and held-out settings. To distinguish correction at learned states from interpolation beyond fitted coordinates, each of five confirmation seeds evaluates eight known-family endpoints per dataset. For each endpoint, scalar encoders train on $\{0, .2, .4, .6, .8, 1\}$ and test on $\{.1, .25, .5, .75, .9\}$, without refitting the mass ensemble at test coordinates. Mixtures therefore test held-out proportions, while latent paths test held-out coordinates along a supplied endpoint direction. Neither study holds out endpoint identities or policy families. We retain every evaluated cell in the unconditional performance summaries so that interpolation failures remain visible.

Comparators and diagnostic references. We compare against empirical risk minimization with a mask (ERM+Mask), ModDrop-M (Neverova et al., 2016), Random-LPG, AdvDrop-M (Park et al., 2018), MultiMoDN-F (Swamy et al., 2023), GroupDRO-M (Sagawa et al., 2020), and domain adaptation under missingness shift with mask importance weighting (DAMS-Mask-IW) (Zhou et al., 2023). These methods cover conventional prediction, incomplete-modality training, and robustness or adaptation to availability patterns. Suffixes distinguish modality-level adaptations from feature-level reproductions. Endpoint effects use DAMS as a fixed named comparator. DAMS receives target masks R, whereas PCOM also receives paired context-mask logs and a source-trained policy family.

For mixtures, paths, and hidden-content policies, we measure gains against the frozen source anchor because no corresponding DAMS runs are available. We also supply the true state or coefficient to the same learned correction to form the state-known reference. Comparing this reference with inferred PCOM measures the effect of withholding state information, while comparing either predictor with the anchor measures task utility. This state-known reference is learned, not a Bayes oracle.

Metrics and uncertainty. We use negative log-likelihood (NLL) as the primary metric because the correction estimates class probabilities and its population interpretation is defined under log loss. Area under the receiver operating characteristic curve (AUROC), Brier score, and expected calibration error (ECE) (Guo et al., 2017) provide complementary ranking and calibration measurements. MM-IMDb uses mean Bernoulli NLL and macro label AUROC.

Policy cells within a seed share fitted models, so we average them equally within each seed before reporting sample standard deviation (SD) or paired two-sided 95% Student-t confidence intervals (CIs) across the five confirmation seeds. These intervals describe seed variation under the fixed processed datasets. We report recovery $G^{\text{inf}}/G^{\text{known}}$ only when $G^{\text{known}} > 0.005$ to avoid unstable ratios with near-zero denominators. This offline diagnostic supplements all-cell NLL gains and neither selects predictions nor tunes models during deployment or evaluation.

6 Results

Effect of Policy State on the Correct Prediction (RQ1). To test whether policy state changes the prediction target, we use a controlled construction with matched global mask rates and common observation support. We compare source-specific, target-specific, policy-agnostic shared, and state-conditioned predictors in a separate six-seed study, isolating posterior conflict from changes in overall modality availability.

Table 2: **Current learned-family endpoints.** NLL mean ± sample SD over five seed means. First six rows are source-only baselines. DAMS uses target masks, and PCOM additionally uses retained context-mask logs. State-known PCOM is a diagnostic. Full metrics and feature-contract qualifications are in sections D and H, respectively.

Method	CMU-MOSEI	MELD	MM-IMDb	UPMC
ERM+Mask	0.6620 ± 0.0091	1.5938 ± 0.0310	0.2716 ± 0.0037	2.4748 ± 0.0603
ModDrop-M	0.6108 ± 0.0370	1.5487 ± 0.0341	0.2705 ± 0.0035	2.4425 ± 0.0608
Random-LPG	0.6417 ± 0.0461	1.5402 ± 0.0274	0.2728 ± 0.0031	2.4525 ± 0.0586
AdvDrop-M	0.5864 ± 0.0156	1.4732 ± 0.0397	0.2722 ± 0.0036	2.4327 ± 0.0611
MultiMoDN-F	0.6163 ± 0.0210	1.5395 ± 0.0265	0.2815 ± 0.0061	2.4352 ± 0.0674
GroupDRO-M	0.6360 ± 0.0197	1.5719 ± 0.0475	0.2721 ± 0.0037	2.6908 ± 0.0626
DAMS-Mask-IW	0.6410 ± 0.0199	1.5569 ± 0.0297	0.2721 ± 0.0037	2.4491 ± 0.0579
PCOM	$\mathbf{0.2280 \pm 0.0542}$	$\mathbf{1.0336 \pm 0.0289}$	$\mathbf{0.2627 \pm 0.0035}$	$\mathbf{2.2661 \pm 0.0612}$
State-known PCOM	0.2277 ± 0.0545	1.0336 ± 0.0290	0.2627 ± 0.0036	2.2661 ± 0.0612

Table 1 shows that the specialized predictors exchange performance across policies, whereas the shared predictor compromises. Supplying the state reduces source and target NLL from 0.4957 and 0.4913 to 0.4407 and 0.4417, nearly matching each specialized predictor in its own environment. The inferred-state predictor closes much of the same gap, showing that estimated policy information can improve prediction under both policies in this controlled setting.

A related controlled calculation estimates $I(Y; Z \mid O)$ as 0.052825 nats and the learned shared-to-state-known NLL reduction as 0.052780 nats. Their independently computed agreement is consistent with the log-loss identity in Eq. (3), rather than a numerical proof that the fitted predictors are Bayes-optimal. The full results are in section G. These observations establish a predictive role for policy conditioning in the controlled setting. We next test whether the learned correction retains that benefit on non-toy features, where its policy state and acquisition masses must be estimated.

Effect of Learned Correction on Predictive NLL (RQ2). The population identity specifies the required correction, but it does not guarantee that estimated masses and inferred states reduce prediction loss. We therefore evaluate PCOM first at learned policy endpoints and then at held-out mixture proportions, keeping the fitted mass models fixed during evaluation.

At endpoints, PCOM reduces NLL relative to DAMS by 0.4129, 0.5233, 0.0094, and 0.1831 on CMU-MOSEI, MELD, MM-IMDb, and UPMC Food-101 (Tables 2 and 3). All paired CIs are above zero. The improvement varies across tasks but is positive on average in each dataset under the declared information interfaces. Mean endpoint gaps to state-known PCOM are below 0.0004, indicating little additional average loss from inferring the state. As DAMS and PCOM receive different target information, their endpoint gap compares the full systems, and it does not isolate the mass correction.

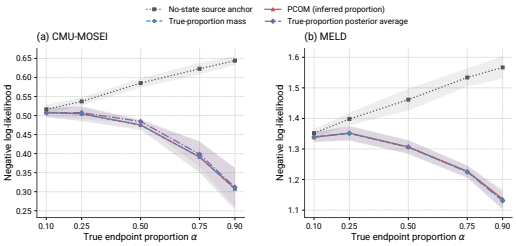

Figure 3: **Held-out mixtures.** PCOM infers α from logs. Bands show variability across five seeds.

The secondary metrics show that these endpoint gains are not confined to log loss. Relative to DAMS, AUROC rises from 0.7035 to 0.9633 on CMU-MOSEI, from 0.5927 to 0.8450 on MELD, from 0.7603 to 0.7871 on MM-IMDb, and from 0.9111 to 0.9232 on UPMC Food-101. PCOM also has the lowest reported ECE on CMU-MOSEI, MELD, and UPMC Food-101, while AdvDrop-M remains better on MM-IMDb (0.0109 versus 0.0130). Thus, the NLL result is accompanied by better ranking across all four datasets, but not by uniform dominance on every calibration summary. Section H reports the complete metric tables.

For record-level mixtures, we hold out proportions while keeping the endpoint identity known. Mean gains over the source anchor are 0.1437, 0.1902, 0.0027, and 0.0705 across the same datasets, with

Table 3: **Paired effects with distinct references.** Mean [two-sided 95% CI] over five seeds. Endpoints use fixed DAMS. Mixture and continuous gains use the same frozen source anchor. No DAMS mixture/path comparison is available, and every policy cell contributes before aggregation.

Dataset	Endpoint DAMS − PCOM	Mixture $G^{\inf}$	Continuous $G^{\inf}$
CMU-MOSEI	0.4129 [0.3638, 0.4620]	0.1437 [0.1206, 0.1669]	−0.0391 [−0.1383, 0.0601]
MELD	0.5233 [0.4581, 0.5886]	0.1902 [0.1665, 0.2139]	−0.0469 [−0.0617, −0.0320]
MM-IMDb	0.0094 [0.0064, 0.0124]	0.0027 [0.0012, 0.0043]	0.0023 [−0.0003, 0.0050]
UPMC Food-101	0.1831 [0.1740, 0.1921]	0.0705 [0.0690, 0.0720]	0.0652 [0.0470, 0.0834]

positive paired CIs when every tested cell is retained (Table 3). The corresponding mean gaps to known-proportion PCOM are 0.0003, 0.0010, 0.0003, and 0.0005. Thus, the correction retains utility between fitted proportions without refitting the mass ensemble.

This experiment isolates interpolation in acquisition-mechanism space rather than generalization to an unidentified policy family. The endpoint direction is supplied, the proportion is inferred only from held-out context-mask logs, and the mass models remain frozen. Consequently, the comparison tests whether the mixture construction in Eq. (7) remains useful between fitted proportions. It does not establish recovery of an arbitrary policy whose endpoint identity or functional form is absent from source training.

Fig. 3 also compares mass-space correction with averaging normalized endpoint posteriors. The latter has lower NLL on MELD and MM-IMDb despite representing a different probabilistic operation. This separates the correct population construction from the accuracy of its estimated implementation. PCOM estimates both masses and the mixture proportion, whereas this control receives the true proportion. The positive gains therefore support useful interpolation on known source-endpoint pairs, but not uniform dominance over that more informed control. We next test whether agreement with a learned state-known predictor also implies improvement over the source anchor.

Sufficiency of State-Known Fidelity for Utility (RQ3). We test whether matching a learned state-known predictor is sufficient for useful correction. Here, fidelity means a small NLL gap between inferred PCOM and the same model given the policy state. Utility means lower NLL than the frozen source anchor. We examine both quantities on continuous latent paths and a separate label-blind hidden-content family.

In dataset order, continuous-path mean gains are −0.0391, −0.0469, 0.0023, and 0.0652 (Table 3). MELD's CI lies below zero, the CMU-MOSEI and MM-IMDb CIs include zero, and the UPMC Food-101 CI remains positive. Hidden-content gains are −0.0033 (CMU-MOSEI) and −0.0091 (MELD), excluding the development seed. Thus, endpoint and mixture gains do not extend uniformly across the evaluated policy constructions, including continuous latent paths and the separate hidden-content family.

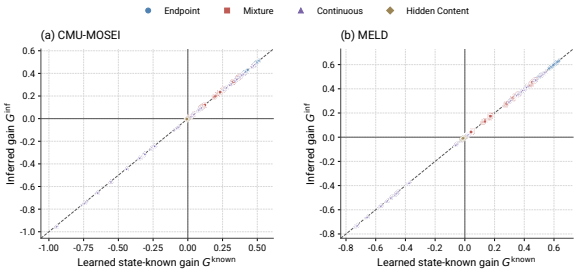

Figure 4: **State-known fidelity is not utility.**

The all-cell audit makes this distinction explicit. We evaluate 160 endpoint, 800 continuous, 800 mixture, and 80 hidden-content cells, retaining every confirmation cell in the unconditional summaries. Of these 1,840 cells, 438 have negative state-known or inferred gain. The failures are concentrated outside the endpoint setting and remain visible even when inferred predictions closely reproduce their state-known counterparts. This rules out an interpretation based only on average favorable cells and motivates reporting utility and state fidelity as separate outcomes.

Fig. 4 shows that inferred and state-known gains remain close even where both are negative. The direct NLL gaps in section I confirm this behavior. Accurate reproduction of a learned state-known predictor is therefore insufficient for positive utility when that predictor itself underperforms the

anchor. This is a limitation of the learned correction, not evidence that Bayes policy information is negative or absent. We therefore examine the update and its information channels to explain both the gains and the failures.

Analysis of Observed Behavior and Scope of Validity (RQ4). To explain the difference between state fidelity and predictive utility, we examine class-dependent reweighting, policy-state selection, and information in acquisition logs. Support diagnostics and failure cases then delimit the empirical scope of the method. In a three-seed endpoint-only ablation, removing candidate-class dependence or the state term yields zero correction. The first result follows from class normalization: a factor shared by all classes cancels. These interventions verify the structure of the fitted update, rather than the performance of fully retrained equal-capacity alternatives.

Hard state selection and marginalization give nearly identical CMU endpoint gains (0.3506 and 0.3503), while marginalization improves MM gain from 0.0082 to 0.0102. Its benefit is therefore dataset dependent. The analytic-mass comparator is stronger on CMU and weaker on MM, further separating the validity of the correction identity from the quality of a particular mass estimator (section J). Where analytic mass is available, the learned estimator is not automatically preferable.

To determine which log relationships support state inference, the hidden-content audit compares paired context-mask inputs with context-only, mask-only, and shuffled-pair controls. The controls remain near nine-state chance performance, while paired retrieval reaches 0.2453 on CMU and 0.3840 on MELD. Because this construction uses public text context and policies without label inputs, the contrast isolates the record-level context-mask relationship used for state retrieval.

Finally, continuous-policy audits record no unsupported or cap-exceeding estimated ratios on the sampled target records (section K). These measurements check the numerical behavior of evaluated corrections, while the reported target-sample weight dispersion remains distinct from source-law population ESS. Taken together, the component and input studies identify how policy-specific class reweighting produces the update, while the analytic-mass and hidden-content results show why this structure alone does not ensure an accurate or useful learned correction in every setting.

These results also determine the scope of the empirical claim. Endpoints test the complete learned system on source-generated policy states, and mixtures test interpolation between an identified source-endpoint pair. Continuous paths and hidden-content policies deliberately move beyond those interfaces and expose negative cells. We therefore interpret PCOM as a policy-conditioned correction within a learned acquisition family, not as a guarantee for arbitrary acquisition mechanisms. When the target mechanism cannot be located within that family, state fidelity alone is not evidence that point correction should be applied.

7 Discussion and Conclusion

Acquisition-policy changes can alter the class posterior associated with the same observed record. We express that change through observed acquisition mass and implement the resulting correction with PCOM. The controlled study establishes a predictive role for policy information, while the non-toy experiments show useful correction at learned endpoints and held-out proportions on known source-endpoint pairs.

The stress tests distinguish state-known fidelity from predictive utility. Inferred PCOM can closely reproduce a learned state-known predictor even when both underperform the source anchor. Component and log studies explain how class-dependent reweighting changes prediction and which context-mask relationships support state inference. These findings support policy-conditioned correction in the evaluated settings while showing why state fidelity and task utility must be assessed separately.

We freeze the source anchor so adaptation remains focused on interpreting observation changes rather than relearning the underlying feature representation. At deployment, we use retained context-mask logs to select or mix source-trained branches, preserving record-level associations between context and modality availability lost under mask-frequency aggregation. Our decomposition also makes two error sources explicit. The first arises from finite-neighborhood estimation of acquisition mass, while the second arises from finite-window inference of policy state. We therefore do not treat accurate endpoint retrieval as sufficient evidence of a specified mass surface, nor a small inferred-to-state-known gap as evidence of positive utility outside the learned acquisition law.

AI USE STATEMENT

AI tools were used only to assist with language editing and improve the clarity of the manuscript. All scientific content, analyses, results, and conclusions were produced and verified by the authors.

ETHICS STATEMENT

This work does not involve new human-subject data collection or deployment decisions. We use existing research datasets and follow their established research-use protocols.

REPRODUCIBILITY STATEMENT

We provide the information needed to reproduce the method and reported evaluation. The model architecture, training procedure, algorithms, and hyperparameters are described in section C. Dataset preprocessing and partition protocols are provided in section D. Acquisition-policy construction and evaluation settings are described in section E, while complete theoretical derivations are given in section B. Extended results and ablations are reported in sections H to J.

REFERENCES

Anastasios N. Angelopoulos and Stephen Bates. Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4):494–591, 2023. doi: 10.1561/2200000101.

John Arevalo, Thamar Solorio, Manuel Montes-y Gómez, and Fabio A. González. Gated multimodal units for information fusion. In *International Conference on Learning Representations Workshop*, 2017.

AmirAli Bagher Zadeh, Paul Pu Liang, Soujanya Poria, Erik Cambria, and Louis-Philippe Morency. Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pp. 2236–2246, 2018. doi: 10.18653/v1/P18-1208. URL https://aclanthology.org/P18-1208/.

Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2): 423–443, 2019. doi: 10.1109/TPAMI.2018.2798607.

David Blackwell. Conditional expectation and unbiased sequential estimation. *The Annals of Mathematical Statistics*, 18(1):105–110, 1947. doi: 10.1214/aoms/1177730497.

Remi Cadene, Corentin Dancette, Hedi Ben-younes, Matthieu Cord, and Devi Parikh. RUBi: Reducing unimodal biases for visual question answering. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8:6085, 2018. doi: 10.1038/s41598-018-24271-9.

Young Sang Choi, Vincent Jeanselme, Pierre Elias, and Shalmali Joshi. ICYM^{2}I: The illusion of multimodal informativeness under missingness. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/883cf7782b794759b98dd604c4adb5a1-Abstract-Conference.html.

Ruiting Dai, Chenxi Li, Yandong Yan, Lisi Mo, Ke Qin, and Tao He. Unbiased missing-modality multimodal learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 24507–24517, 2025. doi: 10.1109/ICCV51701.2025.02272.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016.

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A. Wichmann. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2:665–673, 2020. doi: 10.1038/s42256-020-00257-z.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1321–1330. PMLR, 2017.

Zirun Guo, Tao Jin, and Zhou Zhao. Multimodal prompt learning with missing modalities for sentiment analysis and emotion recognition. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1726–1736. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.94. URL https://aclanthology.org/2024.acl-long.94/.

Xing Han, Huy Nguyen, Carl Harris, Nhat Ho, and Suchi Saria. FuseMoE: Mixture-of-experts transformers for fleximodal fusion. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2167.

Lianyu Hu, Tongkai Shi, Wei Feng, Fanhua Shang, and Liang Wan. Deep correlated prompting for visual recognition with missing modalities. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2153.

Byungju Kim, Hyunwoo Kim, Kyungsu Kim, Sungjin Kim, and Junmo Kim. Learning not to learn: Training deep neural networks with biased data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9012–9020, 2019. doi: 10.1109/CVPR.2019.00922.

Jihye Lee, Minseo Kang, and Dongha Kim. MIRRAMS: Learning robust tabular models under unseen missingness shifts. *arXiv preprint arXiv:2507.08280*, 2025. URL https://arxiv.org/abs/2507.08280.

Yi-Lun Lee, Yi-Hsuan Tsai, Wei-Chen Chiu, and Chen-Yu Lee. Multimodal prompting with missing modalities for visual recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14943–14952, 2023. doi: 10.1109/CVPR52729.2023.01435.

Sijie Li, Chen Chen, and Jungong Han. SimMLM: A simple framework for multi-modal learning with missing modality. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 24068–24077, 2025. doi: 10.1109/ICCV51701.2025.02231.

Zihan Liang, Ziwen Pan, and Ruoxuan Xiong. Causal representation learning from multimodal clinical records under non-random modality missingness. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 28791–28808. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.emnlp-main.1465. URL https://aclanthology.org/2025.emnlp-main.1465/.

Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115.

Ronghao Lin and Haifeng Hu. MissModal: Increasing robustness to missing modality in multimodal sentiment analysis. *Transactions of the Association for Computational Linguistics*, 11:1686–1702, 2023. doi: 10.1162/tacl_a_00628. URL https://aclanthology.org/2023.tacl-1.94/.

Roderick J. A. Little and Donald B. Rubin. *Statistical Analysis with Missing Data*. Wiley, 3 edition, 2019. doi: 10.1002/9781119482260.

Mengmeng Ma, Jian Ren, Long Zhao, Sergey Tulyakov, Cathy Wu, and Xi Peng. SMIL: Multimodal learning with severely missing modality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 2302–2310, 2021. doi: 10.1609/aaai.v35i3.16330.

Mengmeng Ma, Jian Ren, Long Zhao, Davide Testuggine, and Xi Peng. Are multimodal transformers robust to missing modality? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18177–18186, 2022. doi: 10.1109/CVPR52688.2022.01764.

Karthika Mohan and Judea Pearl. Graphical models for processing missing data. *Journal of the American Statistical Association*, 116(534):1023–1037, 2021. doi: 10.1080/01621459.2021.1874961.

Natalia Neverova, Christian Wolf, Graham W. Taylor, and Florian Nebout. ModDrop: Adaptive multi-modal gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(8):1692–1706, 2016. doi: 10.1109/TPAMI.2015.2461544.

Sungrae Park, JunKeon Park, Su-Jin Shin, and Il-Chul Moon. Adversarial dropout for supervised and semi-supervised learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, pp. 3917–3924, 2018. doi: 10.1609/aaai.v32i1.11634.

Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, Gautam Naik, Erik Cambria, and Rada Mihalcea. MELD: A multimodal multi-party dataset for emotion recognition in conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 527–536. Association for Computational Linguistics, 2019. doi: 10.18653/v1/P19-1050. URL https://aclanthology.org/P19-1050/.

C. Radhakrishna Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37(3):81–91, 1945.

Alfréd Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pp. 547–561. University of California Press, 1961.

Patrick Rockenschaub, Zhicong Xian, Alireza Zamanian, Marta Piperno, Octavia-Andreea Ciora, Elisabeth Pachl, and Narges Ahmidi. Robust prediction under missingness shifts. *arXiv preprint arXiv:2406.16484*, 2024. URL https://arxiv.org/abs/2406.16484.

Yaniv Romano, Matteo Sesia, and Emmanuel J. Candès. Classification with valid and adaptive coverage. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Donald B. Rubin. Inference and missing data. *Biometrika*, 63(3):581–592, 1976. doi: 10.1093/biomet/63.3.581.

Shiori Sagawa, Pang Wei Koh, Tatsunori B. Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In *International Conference on Learning Representations*, 2020.

Shaun Seaman, John Galati, Dan Jackson, and John Carlin. What is meant by "missing at random"? *Statistical Science*, 28(2):257–268, 2013. doi: 10.1214/13-STS415.

Tyrel Stokes, Hyungrok Do, Saul Blecker, Rumi Chunara, and Samrachana Adhikari. Domain adaptation under MNAR missingness. *arXiv preprint arXiv:2504.00322*, 2025. URL https://arxiv.org/abs/2504.00322.

Vinitra Swamy, Malika Satayeva, Jibril Frej, Thierry Bossy, Thijs Vogels, Martin Jaggi, Tanja Käser, and Mary-Anne Hartley. MultiModN: Multimodal, multi-task, interpretable modular networks. In *Advances in Neural Information Processing Systems*, volume 36, pp. 28115–28138, 2023.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World*. Springer, 2005. doi: 10.1007/b106715.

Hu Wang, Yuanhong Chen, Congbo Ma, Jodie Avery, Louise Hull, and Gustavo Carneiro. Multi-modal learning with missing modality via shared-specific feature modelling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15878–15887, 2023a. doi: 10.1109/CVPR52729.2023.01524.

Xin Wang, Devinder Kumar, Nicolas Thome, Matthieu Cord, and Frédéric Precioso. Recipe recognition with large multimodal food dataset. In *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 1–6, 2015. doi: 10.1109/ICMEW.2015.7169757.

Yuanzhi Wang, Yong Li, and Zhen Cui. Incomplete multimodality-diffused emotion recognition. In *Advances in Neural Information Processing Systems*, volume 36, 2023b. doi: 10.52202/075280-0748. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/372cb7805eaccb2b7eed641271a30eec-Abstract.html.

Wenxin Xu, Hexin Jiang, and Xuefeng Liang. Leveraging knowledge of modality experts for incomplete multimodal learning. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 438–446, 2024. doi: 10.1145/3664647.3681683.

Sukwon Yun, Inyoung Choi, Jie Peng, Yangfan Wu, Jingxuan Bao, Qiyiwen Zhang, Jiayi Xin, Qi Long, and Tianlong Chen. Flex-MoE: Modeling arbitrary modality combination via the flexible mixture-of-experts. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3135.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabás Póczos, Ruslan Salakhutdinov, and Alexander J. Smola. Deep sets. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Helen Zhou, Sivaraman Balakrishnan, and Zachary C. Lipton. Domain adaptation under missingness shift. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 9577–9606. PMLR, 2023.

A NOTATION AND SETUP

We restate the notation for the Policy-Conditioned Observed-Mass Model (PCOM) to connect the acquisition mechanism, the posterior correction, and the quantities measured in the experiments. The derivations first identify the population correction. The implementation and additional results then explain how we estimate it and evaluate its predictive effect.

Symbol	Meaning
P_0	Pre-acquisition joint law of complete content, labels, and context
$X = (X^{(1)}, \dots, X^{(M)})$	Complete multimodal feature record
Y, C, R	Label, retained context, and observed-modality mask
$O = (X_r, r, C)$	Observation when the realized mask is r
$\pi_z(a \mid X, Y, C)$	Acquisition probability for action a under policy z
$m_z(r \mid x_r, y, c)$	Observed mass for the realized mask, marginalized under P_0
$\eta_z^\star(y \mid o)$	Ideal policy-specific class posterior
$q_{\mathrm{s}}(y \mid x_r, r)$	Frozen learned source predictor, excluding direct context input
$\widehat{q}_z$, $\bar{p}$, $\widehat{p}_L$	Learned known-state, exact state-averaged, and sampled predictions
$q_\psi(z \mid W)$	Distribution over policy states inferred from acquisition logs
G^{known}, G^{inf}	Learned state-known and inferred gains relative to the source anchor
α, λ	Record-mixture proportion and continuous latent-path coordinate
A, K, F, B, L	Action count, donor count, fold count, window size, and sampled states

Table 4: Notation for the acquisition model, corrected predictions, and evaluation metrics.

To isolate acquisition-induced changes, we keep the pre-acquisition distribution $P_0(X, Y, C)$ fixed while varying π_z. This leaves a shared content factor in the observed joint distributions, allowing the posterior change to be expressed through acquisition masses. The weaker condition $p_z(x, y \mid c) = p_0(x, y \mid c)$ also preserves the class-normalized correction, because the marginal factor $p_z(c)$ cancels across candidate labels. The main statement uses the fixed complete-data distribution of the controlled experiments.

Context is included in the observation because it supplies the retained information used to interpret availability. A class-dependent acquisition law describes how labels and selection are statistically related. In the label-aware experiments, labels generate masks offline, while prediction evaluates the correction for each candidate label.

B DERIVATIONS AND PROOFS

We derive the correction by separating content from selection, then examine the consequences for policy-agnostic prediction and learned approximation. The remaining results explain why mixtures must be formed in mass space, how the source bypass behaves under state uncertainty, and how the construction extends to multilabel tasks.

B.1 OBSERVED MASS UNDER THE PRE-ACQUISITION LAW

To express the acquisition effect using only observed content, fix a query mask r and integrate over the unavailable coordinates:

$$p_z(o, y) = \int p_0(x_r, x_{\bar{r}}, y, c)\pi_z(r \mid x_r, x_{\bar{r}}, y, c)\,\mathrm{d}x_{\bar{r}} \tag{8}$$

$$= p_0(x_r, y, c) \int \pi_z(r \mid x_r, x_{\bar{r}}, y, c)p_0(x_{\bar{r}} \mid x_r, y, c)\,\mathrm{d}x_{\bar{r}} \tag{9}$$

$$= p_0(x_r, y, c)m_z(r \mid x_r, y, c). \tag{10}$$

For discrete content, sums replace the integrals. The equalities hold almost everywhere on the relevant support. Conditioning under P_0 averages the acquisition probability over complete records compatible with the query. Conditioning instead on the already selected event $R = r$ would change that averaging distribution and incorporate selection a second time.

The mass regressor predicts the full action vector so that one model can share information across acquisition actions. With the query mask r held fixed, the population target for output action a is

$$m_z^{[r]}(a \mid x_r, y, c) = \mathbb{E}_{P_0}[\pi_z(a \mid X, Y, C) \mid X_r = x_r, Y = y, C = c]. \tag{11}$$

Because the conditioning information is the same for every output action, the coordinates sum to one. Prediction reads $a = r$, thereby distinguishing the action being scored from the mask that determines which query features are available.

B.2 Posterior-correction identity

The observed-joint factorization gives source and target distributions the same content factor. Taking their ratio therefore isolates the change due to acquisition:

$$\frac{p_z(o, y)}{p_{z_s}(o, y)} = \frac{m_z(r \mid x_r, y, c)}{m_{z_s}(r \mid x_r, y, c)}. \tag{12}$$

Substituting $p_{z_s}(o, y) = \eta^\star_{z_s}(y \mid o) p_{z_s}(o)$ and normalizing across candidate labels yields Eq. (2). Positivity is required wherever the target joint probability is positive so that this ratio is defined. If the ratio is identical across classes, it also cancels during normalization. A policy can change the acquisition distribution without changing the class posterior at that observation.

B.3 Posterior-weighted policy-agnostic regret

Having identified the policy-specific posterior, we ask what a predictor can achieve when it receives only the observation. For a fixed o, let $\eta^\star_0 = \sum_z \rho_z(o)\eta^\star_z$. Using Kullback-Leibler (KL) divergence, we expand the weighted objective as

$$\sum_z \rho_z D_{\mathrm{KL}}(\eta^\star_z \| q) = \sum_z \rho_z \sum_y \eta^\star_z(y) \log \frac{\eta^\star_z(y)}{\eta^\star_0(y)} + \sum_y \eta^\star_0(y) \log \frac{\eta^\star_0(y)}{q(y)} \tag{13}$$

$$= \mathrm{JS}_{\rho(o)}(\{\eta^\star_z\}_z) + D_{\mathrm{KL}}(\eta^\star_0 \| q). \tag{14}$$

The second term vanishes when $q = \eta^\star_0$, identifying the best observation-only prediction. The first is the generalized Jensen-Shannon divergence (Lin, 1991). It remains positive when positive-weight policy posteriors disagree. Better fitting can therefore remove the distance to the optimal common prediction, but the remaining conflict requires additional policy information.

B.4 Population risk and learned gains

We separate population risk from measured gains to distinguish useful policy information from the accuracy of the model that uses it. For any fixed policy z,

$$\mathcal{R}_z(q) = \mathbb{E}_{O \sim P_z}\left[H(\eta^\star_z(\cdot \mid O)) + D_{\mathrm{KL}}(\eta^\star_z(\cdot \mid O) \| q(\cdot \mid O))\right]. \tag{15}$$

Using the source Bayes posterior for q and subtracting the target Bayes risk leaves the expected KL from the target posterior to the source posterior. Under a joint distribution that also samples Z, observing the state reduces Bayes risk by

$$H(Y \mid O) - H(Y \mid O, Z) = I(Y; Z \mid O) \geq 0. \tag{16}$$

The empirical gains in Eq. (4) instead compare learned predictors with the frozen source anchor. They can be negative when the estimated correction is less accurate than the anchor. Their difference measures the additional loss from inferring the state, while each gain separately measures improvement over the anchor. This is why state-known fidelity and task utility are evaluated separately.

B.5 Conditional-expectation projection

The observed mass is also the squared-error-optimal summary of the complete-data acquisition probability under the query's information. For fixed r, z, define $A_{z,a} = \pi_z(a \mid X, Y, C)$, $\mathcal{G}_r = \sigma(X_r, Y, C)$, and $m = \mathbb{E}_{P_0}[A_{z,a} \mid \mathcal{G}_r]$. For any square-integrable $\mathcal{G}_r$-measurable g,

$$\mathbb{E}[(A_{z,a} - g)^2] = \mathbb{E}[(A_{z,a} - m)^2] + \mathbb{E}[(m - g)^2] + 2\mathbb{E}[(A_{z,a} - m)(m - g)] \tag{17}$$

$$= \mathbb{E}[(A_{z,a} - m)^2] + \mathbb{E}[(m - g)^2], \tag{18}$$

The cross term vanishes because $m - g$ is measurable with respect to $\mathcal{G}_r$ and $\mathbb{E}[A_{z,a} - m \mid \mathcal{G}_r] = 0$. This conditional-expectation projection (Rao, 1945; Blackwell, 1947) identifies the quantity that the donor estimator and mass network seek to approximate. Their finite-sample accuracy is assessed separately in section J.

B.6 Plug-in error and class-independent cancellation

The learned correction uses an approximate anchor and estimated masses. To relate these errors to the final posterior, define the combined log error on positive target support:

$$\delta_y(o) = \log \frac{q_{\mathrm{s}}(y \mid x_r, r)}{\eta^\star_{z_{\mathrm{s}}}(y \mid o)} + \log \frac{\widehat{m}^y_z(o)}{m^y_z(o)} - \log \frac{\widehat{m}^y_{z_{\mathrm{s}}}(o)}{m^y_{z_{\mathrm{s}}}(o)}. \tag{19}$$

The estimated masses denote the effective values used after numerical stabilization. The analytic uniform source mass contributes no estimation error away from numerical changes.

Proposition 2 (Class-centered plug-in error)**.** *Suppose $|\delta_y(o) - b(o)| \le \epsilon$ for every class y, where $b(o)$ is class independent. Then*

$$D_{\mathrm{KL}}(\eta^\star_z(\cdot \mid o) \| \widehat{q}_z(\cdot \mid o)) \le 2\epsilon. \tag{20}$$

Proof. Let a_y be exact corrected logits and write $e_y = \delta_y - b(o)$. Subtracting the same offset from every class leaves the softmax unchanged, so

$$\log \frac{\widehat{q}_z(y \mid o)}{\eta^\star_z(y \mid o)} = e_y - \log \sum_{y'} \eta^\star_z(y' \mid o) e^{e_{y'}}. \tag{21}$$

Both e_y and the log-sum term lie in $[-\epsilon, \epsilon]$. The absolute log-posterior error is therefore at most 2ϵ, and averaging $\log(\eta^\star_z / \widehat{q}_z)$ under $\eta^\star_z$ gives the KL bound. □

The bound explains why class-relative error matters. A common multiplicative error disappears after class normalization. We therefore examine centered log-mass error alongside posterior error rather than relying only on raw acquisition-distribution error.

B.7 Record-level mixture correction

For mixture weights independent of hidden record values, linearity of conditional expectation gives

$$\pi_{\boldsymbol{\alpha}} = \sum_k \alpha_k \pi_k \quad \Longrightarrow \quad m^y_{\boldsymbol{\alpha}} = \sum_k \alpha_k m^y_k. \tag{22}$$

To see why mixing normalized posteriors is different, let $a_k(y) = q_{\mathrm{s}}(y) m^y_k / m^y_{z_{\mathrm{s}}}$ and $Z_k = \sum_y a_k(y)$. The posterior obtained from the mixed masses is

$$\frac{\sum_k \alpha_k a_k(y)}{\sum_k \alpha_k Z_k} = \sum_k \frac{\alpha_k Z_k}{\sum_j \alpha_j Z_j} \widehat{q}_k(y). \tag{23}$$

It is thus a posterior average with coefficients reweighted by the observation-dependent normalizers, rather than the original α_k. For example, take $q_s = (1/2, 1/2)$ and positive weight vectors $(2, 1)$ and $(1, 4)$. Averaging their normalized posteriors gives $(0.4333, 0.5667)$, whereas normalizing their averaged unnormalized weights gives $(0.375, 0.625)$.

This distinction determines the two prediction procedures. Uncertainty about one fixed state is represented by averaging state-specific predictions. A record-level acquisition mixture first combines masses. A latent path $z_\lambda = \lambda z_k$ instead queries a nonlinear generator at a new coordinate. Its acquisition law need not be the corresponding convex mixture. The experiments keep these constructions separate.

B.8 SOURCE PRESERVATION UNDER STATE UNCERTAINTY

The source bypass fixes $\widehat{q}_{z_s} = q_s$. For the exact state average, the remaining deviation comes only from non-source branches. We measure it by total variation distance TV.

$$\bar{p} - q_s = \sum_{z \neq z_s} q_\psi(z \mid W)(\widehat{q}_z - q_s), \qquad \mathrm{TV}(\bar{p}, q_s) \leq 1 - q_\psi(z_s \mid W). \tag{24}$$

With L sampled states, let $\widehat{\rho}_L$ denote the observed fraction of source-state samples. The same argument gives the realization-wise bound

$$\mathrm{TV}(\widehat{p}_L, q_s) \leq 1 - \widehat{\rho}_L. \tag{25}$$

Averaging this bound over state sampling gives $1 - q_\psi(z_s \mid W)$. A one-sided Hoeffding bound further yields, with conditional probability at least $1 - \beta$,

$$\mathrm{TV}(\widehat{p}_L, q_s) \leq \min\left\{1,\, 1 - q_\psi(z_s \mid W) + \sqrt{\frac{\log(1/\beta)}{2L}}\right\}. \tag{26}$$

The bypass therefore removes mass-estimation error from the source branch, while the complete prediction also reflects how much inferred or sampled weight is assigned to that branch. The numerical identity check evaluates the source branch with the source state supplied.

B.9 SUPPORT UNDER THE UNIFORM SOURCE POLICY

A correction ratio must be supported by the source distribution. The population effective sample size (ESS) is defined under the source law. If $P_z \ll P_{z_s}$ and $w = \mathrm{d}P_z/\mathrm{d}P_{z_s}$, then

$$\rho_{\mathrm{ESS}}^{\mathrm{source}} = \frac{(\mathbb{E}_{P_{z_s}} w)^2}{\mathbb{E}_{P_{z_s}} w^2} = \exp\{-D_2(P_z \| P_{z_s})\}, \tag{27}$$

where $\mathbb{E}_{P_{z_s}} w = 1$ and D_2 is order-two Rényi divergence (Rényi, 1961). This population ESS uses a source-law expectation.

For the main uniform source catalog, $m_{z_s} = 1/A$ and $0 \leq w = Am_z \leq A$. Consequently, $\mathbb{E}_{P_{z_s}} w^2 \leq A\mathbb{E}_{P_{z_s}} w = A$, giving $\rho_{\mathrm{ESS}}^{\mathrm{source}} \geq 1/A$ on the common feasible catalog. This provides a positive source reference for every action. The sample-based weight measurements in section K describe the estimated ratios on evaluated target records and complement this population calculation.

B.10 LABEL-WISE MULTILABEL CORRECTION

To apply the same construction to a multilabel task, consider each Bernoulli coordinate Y_j separately and define

$$m_{z,j}(r \mid x_r, y_j, c) = \mathbb{E}_{P_0}[\pi_z(r \mid X, Y, C) \mid X_r = x_r, Y_j = y_j, C = c]. \tag{28}$$

Here the conditional expectation averages over both unavailable content and the remaining labels Y_{-j}. Applying Proposition 1 to the two outcomes of Y_j gives the label-wise correction used for MM-IMDb. The outputs are marginal Bernoulli probabilities. The procedure does not require an independence assumption among the labels.

C IMPLEMENTATION DETAILS

We keep task learning, policy estimation, and posterior correction separate so that each model supplies a defined part of the update. The following details specify how the source predictor is fixed, how source-generated policies train the correction modules, and how those modules are combined at evaluation.

C.1 SOURCE PREDICTOR

To distinguish an absent modality from an observed zero vector, the anchor applies the mask to each input, encodes the present features, and also includes the availability mask in its prediction head. Each modality uses two Linear/ReLU layers of width 128. The model averages present embeddings, concatenates the mask, and applies a width-128 head with dropout 0.1. CMU-MOSEI uses two softmax logits, MELD and UPMC Food-101 use seven and 101, and MM-IMDb uses 23 independent sigmoid outputs. Its final affine layer therefore has a 23×128 weight matrix and 23 bias terms.

Binary and multiclass tasks minimize ordinary cross-entropy. MM-IMDb averages binary cross-entropy first over labels and then over records. The losses use no class weighting, and this prediction path has no post-hoc temperature calibration. We train with AdamW at learning rate 10^{-3}, weight decay 10^{-5}, batch size 256, gradient clipping at five, and 30 epochs without a learning-rate schedule. Source-validation negative log-likelihood (NLL) under fixed source masks selects the checkpoint. After freezing, the same anchor evaluated under target masks supplies the source-anchor reference.

C.2 CONTEXT AND POLICY GENERATION

Context summarizes policy-side information retained alongside the availability mask. The main construction concatenates modality-vector summary statistics, available quality metadata, and three source-fitted triage axes from the first modality. Triage uses scaling, a 16-dimensional principal-component representation, and a source-supervised logistic model, with an additional fitted reduction for multiclass features. Scaling by the median and median absolute deviation (MAD) is fitted on model-training records, and values are clipped to $[-6, 6]$. The resulting descriptors are stored in the policy log separately from the masked task input.

The generator uses eight 16-dimensional endpoint codes, two width-128 SiLU layers, and the source-subtracted logits in Eq. (5). Its minimized objective is

$$-\mathbb{E}[\ell_{\text{anchor}}] + 50\Omega_{\text{marg}} + 25\Omega_{\text{path}} - .05\Omega_{\text{sep}}. \tag{29}$$

The anchor-loss term searches for policies that make the fixed predictor less reliable. The marginal penalty encourages the uniform source action distribution, the path penalty encourages overlap at coordinates $0, .25, .5, .75, 1$, and the separation term discourages redundant endpoints. Path overlap is measured using minibatch second moments with a soft target of 0.8. We optimize the weighted objective with AdamW at 10^{-3}, weight decay 10^{-6}, batch size 512, gradient clipping five, and 3,000 steps. The soft penalty allows the measured path quantity to differ from its target.

C.3 POLICY ESTIMATION FROM LOG WINDOWS

The discrete estimator represents the relationship between context and the observed action by embedding (C, R) through two width-128 SiLU layers. It pools feature means and population standard deviations within each window, then applies a width-128 nine-state softmax head. Source-generated training samples states uniformly and uses 48 windows per step, $B = 128$, 2,500 steps, and AdamW at 10^{-3} with weight decay 10^{-6}. We retain the final source-trained estimator.

For continuous coordinates and mixture proportions, a separate scalar estimator is fitted for each dataset, seed, endpoint, and family. It uses pooled log features, a sigmoid output, and smooth-L_1 regression on the six fitting coordinates in section E. The supplied endpoint identity selects the scalar estimator, which then infers the unknown coordinate.

To reduce dependence on one sampled window, evaluation averages predictions over 64 bootstrap windows drawn with replacement from the fixed unlabeled target log pool. Scalar estimation returns the mean and standard deviation of the coordinate estimates. Endpoint prediction averages discrete state distributions and draws $L = 16$ states. This batch-level protocol uses the same deployment log pool to estimate the policy state applied to its records.

C.4 DONOR TARGETS AND MASS ESTIMATION

The donor representation is designed to match complete source records using the content available in the query. Source-fitted scaling and principal-component analysis produce 16 dimensions per

modality. Missing query embeddings are zero-filled according to the fixed mask, and retained context is concatenated for Euclidean neighbor search. Eligible donors match the candidate class or label bit. The exact query identifier (ID) is excluded. Donor weights are $\exp(-d/s)$, where s is based on the median Kth-neighbor distance with a numerical floor.

The weighted average of donor policy-action vectors provides the soft target. We use $K = 16$ and five deterministic sample-ID folds. Each fold excludes held-out records from both fitting queries and donors, and target prediction averages the five fitted models. Source preprocessing is fitted once on model-training records. Fold exclusion applies to the subsequent query/donor and mass-model fitting stages. Dialogue-level grouping is described in section D.

A policy fingerprint flattens the action probabilities on 32 deterministic reference source-training records. The mass model uses label embeddings of size 24, candidate-value embeddings of size eight, and two width-384 record and policy encoders. Concatenation, elementwise product, and absolute difference combine the encoded query and policy representations. Each fold uses 8,000 training steps, batch size 1,024, AdamW at 10^{-3}, and weight decay 10^{-6}. The source state is omitted from neural mass fitting because its mass is analytically $1/A$.

For numerical stability, normalized probability denominators are floored at 10^{-12}, task-NLL probabilities at 10^{-7}, and selective-score probabilities at 10^{-12}. The weight threshold of 100 is used only in the diagnostic in section K. Model sizes are given in Table 5, and algorithms 1 and 2 summarize the fitting and prediction order.

Module	CMU-MOSEI	MELD	MM-IMDb	UPMC
Task anchor	206,210	574,471	156,951	167,013
Policy generator	21,511	22,535	23,299	33,539
Discrete state encoder	53,129	53,513	52,489	52,489
One mass fold	1,011,887	1,013,183	954,811	956,683
Number of mass folds	5	5	5	5

Table 5: Trainable parameter counts by dataset. The five-fold mass ensemble contains five copies of the listed mass model. Endpoint-specific scalar estimators are additional models.

Algorithm 1 Source fitting of PCOM

Require: Complete source feature partitions, source action catalog, fixed configuration
1: Fit source context and donor preprocessing on the declared training partition
2: Train the anchor under uniform source masks, select by source-validation NLL, and freeze
3: Train endpoint generator using anchor stress loss and soft policy penalties
4: Fit the discrete state encoder on source-generated context-mask windows
5: Assign record-ID folds before expanding any masked views
6: **for** each mass fold **do**
7: Exclude held-out-fold queries and donors, then fit the vector mass regressor
8: Freeze the generator, anchor, state encoder, and mass ensemble
9: **for** each endpoint and each scalar family (continuous or mixture) **do**
10: Fit a separate coordinate estimator on the six source-generated training coordinates

Algorithm 2 Endpoint, mixture, and continuous-coordinate prediction

Require: Unlabeled target log pool, observed query, frozen models, regime
Construct 64 bootstrap log windows of size 128
if regime is an endpoint **then**
    Average discrete state probabilities across windows and sample $L = 16$ states
    for each sampled state z **do**
        if $z = z_{\mathrm{s}}$ **then**
            Return the anchor for this branch
        else
            Read ensemble masses and normalize $q_{\mathrm{s}}\widehat{m}_z/m_{z_{\mathrm{s}}}$
    Average branch predictions
else if regime is a record-level mixture with supplied endpoint k **then**
    Infer $\widehat{\alpha}$ using the scalar estimator selected by k
    Form $(1-\widehat{\alpha})m_{z_{\mathrm{s}}} + \widehat{\alpha}\widehat{m}_k$
    Apply one class normalization after the mass mixture
else
    Infer $\widehat{\lambda}$ for the supplied endpoint direction
    Query the mass model at its fingerprint and apply the correction
Store predictions before evaluating task labels

D Datasets and Preprocessing

We use fixed processed features to study acquisition-policy correction separately from end-to-end representation learning. Table 6 summarizes the input dimensions and output spaces, while the following preparation details specify the data on which the source and correction models are fitted.

Dataset	Retained records	Outputs	Feature dimensions and order
CMU-MOSEI	17,866	2	Text 300, audio 74, visual 713
MELD	13,706	7	Text 300, audio 1,611, visual 2,048
MM-IMDb	15,228	23	Text 300, image 512
UPMC Food-101	90,704	101	Text 300, image 512

Table 6: Processed inputs and prediction tasks. CMU-MOSEI, MELD, and UPMC Food-101 have mutually exclusive classes. MM-IMDb has 23 Bernoulli label outputs.

D.1 Data partitions and grouping

We preserve the cache train/validation/test divisions, reserve about 20% of cache training records for source calibration, and partition cache test into source-test and target-test records. The resulting partitions are listed in Table 7. MELD uses split-qualified dialogue groups for these outer divisions. The other loaders use record identifiers. All masked views of one record retain the same outer partition and mass fold, so changing the mask does not move a record between fitting and evaluation sets.

Dataset	Model train	Source val.	Source cal.	Source test	Target test
CMU-MOSEI	10,233	1,438	2,558	1,819	1,818
MELD	7,832-8,010	1,108	1,978-2,156	1,241-1,329	1,281-1,369
MM-IMDb	8,801	1,942	2,200	1,143	1,142
UPMC Food-101	46,231	10,199	11,558	11,358	11,358

Table 7: Complete-record partitions used for fitting and evaluation. MELD ranges reflect indivisible dialogue groups across five seeds.

Inner mass folds use record IDs for all datasets, ensuring that each held-out record and all of its masked views are excluded from the corresponding fitting queries and donor pools. MELD additionally keeps target dialogues separate through the outer split, while its inner mass folds remain record-based. CMU-MOSEI, MM-IMDb, and UPMC use the same record-level inner-fold contract within their fixed outer partitions.

D.2 CMU-MOSEI

CMU-MOSEI provides language, audio, and visual features for sentiment prediction (Bagher Zadeh et al., 2018). We retain nonzero sentiment examples and use $Y = \mathbb{I}[\text{sentiment} > 0]$, excluding neutral examples. The evaluated cache contains 17,866 records. The preparation script specifies aligned timestamped word, acoustic, and facial streams, mean pooling, and standardization fitted on cache training records.

D.3 MELD

MELD extends the evaluation to seven-class conversational emotion recognition (Poria et al., 2019). The preparation starts from 13,708 utterances in comma-separated-value (CSV) files, comprising 9,989 training, 1,109 development, and 2,610 test records. Two records with missing features are excluded, leaving 13,706. One feature-label disagreement is resolved using the official CSV label. The class order is anger, disgust, fear, joy, neutral, sadness, and surprise.

The inputs combine 300-dimensional GloVe-average text, 1,611-dimensional selected audio features, and mean-pooled 2,048-dimensional MM-Align ResNet101 visual features. The original train/development/test dialogue divisions are retained before forming the additional source and target partitions.

D.4 MM-IMDB

MM-IMDb evaluates label-wise correction on plot/poster genre prediction (Arevalo et al., 2017). The processed task uses 23 genres, namely action, adventure, animation, biography, comedy, crime, documentary, drama, family, fantasy, film-noir, history, horror, music, musical, mystery, romance, sci-fi, short, sport, thriller, war, and western. Label restoration preserves the existing feature arrays and record IDs.

The supplied feature builder specifies term frequency-inverse document frequency (TF-IDF) features with singular value decomposition (SVD) for text and a deterministic random projection of resized image pixels. The evaluated cache uses project-specific train/validation/test partitions and 23 Bernoulli outputs. Training averages binary loss over labels, retaining the marginal prediction task used throughout the comparisons.

D.5 UPMC FOOD-101

UPMC Food-101 provides a 101-class image-text recognition task (Wang et al., 2015). The evaluated cache has 90,704 records, with 300-dimensional text and 512-dimensional image vectors. Text features are built from the dataset's recipe titles using TF-IDF/SVD, without filename or class-name inputs. Image features use a deterministic projection of resized pixels. Source-fitted normalization is applied before model training.

D.6 AVAILABILITY AND FITTED PREPROCESSING

The supplied experimental description uses complete base masks for the four caches. Main source masks are uniform over all nonempty actions, giving seven actions for three modalities and three for two modalities. The hidden-content study retains text and uses the four text-containing masks, allowing optional-channel quality to control audio/visual acquisition while text remains observed.

The context and donor transforms used by PCOM are fitted on the declared model-training partition and then frozen for calibration and evaluation.

E ACQUISITION POLICIES AND EVALUATION SETTINGS

We use distinct policy constructions to separate correction at learned endpoints, interpolation of acquisition probabilities, and interpolation of latent generator codes. Table 8 specifies what is fitted and what changes at evaluation, so each result can be tied to a particular form of generalization.

Regime	Fitting and supplied information	Evaluation variable
Endpoints	Nine learned-family states fit the state/mass interfaces, with state identity inferred	New target records and masks under learned states
Latent path	Endpoint k supplied, with scalar encoder fit at $0, .2, .4, .6, .8, 1$	Coordinate $\lambda \in \{.1, .25, .5, .75, .9\}$ on that known direction
Mass mixture	Pair (z_s, k) supplied, with scalar encoder fit on the same six proportions	Proportion $\alpha \in \{.1, .25, .5, .75, .9\}$ for the supplied pair
Hidden content	Source-fitted independent/correlated quality templates	Target records under the declared templates and severities
Dynamics	Synthetic schedules over known hidden-content states	Temporal arrangement and window composition

Table 8: Policy constructions and their fitting/evaluation variables. Scalar inference receives the endpoint direction and estimates the held-out coordinate or proportion.

The scalar encoders use source-generated fitting and validation grids. Held-out target task losses do not select their checkpoints. Candidate-class correction enumerates possible labels, while source fingerprints use labeled source reference records. In the label-aware simulations, the target record's true label is used offline to sample its mask and is excluded from prediction-time state estimation. Calibration masks are generated by the corresponding frozen acquisition simulator, as detailed in section L.

E.1 Label-blind hidden-content policies

To examine selection that depends on unavailable content without taking the label as an input, we use source-fitted independent and correlated quality templates. Public context contains the norm, mean absolute value, and standard deviation of always-observed text. Quality values derived from optional features and metadata affect acquisition but are withheld from the state encoder.

For optional-mask signs $s_a = 2r_{a,\text{optional}} - 1$, directional quality q, and public-context score v, the independent template scores actions by $q^\top s_a + .25v \sum_j s_{a,j}$. The correlated template adds a pair term that is positive when every optional channel is present, negative when none is present, and zero otherwise. Severities are $.5, 1, 1.5, 2$. Exponentiated scores are rebalanced on source fitting records to match the source marginal mask distribution. This construction varies the content-dependent selection relationship while controlling overall mask frequencies.

F Baseline Design and Related Approaches

The baselines test whether incomplete-input training, availability-based robustness, or target-mask reweighting can account for the observed prediction gains. Empirical risk minimization with mask input (ERM+Mask) and domain adaptation under missingness shift with mask importance weighting (DAMS-Mask-IW) are named baselines. Table 9 records target information alongside each fitted operation, making the comparison explicit before interpreting the task scores.

DAMS-Mask-IW specializes observed-mask adaptation using the target/source mask-frequency ratio (Zhou et al., 2023). With matched global mask marginals, this ratio can be near one even when the context-mask relationship changes. PCOM models that paired relationship through its retained context and source-trained policy family. The endpoint comparison therefore contrasts these complete systems and their declared inputs.

DAMS, MultiMoDN-F, and GroupDRO-M use matching endpoint seed and sampler inputs. Paired differences are formed from outcomes on the same policy cells. Mixture, continuous, and hidden-content studies instead use the frozen source anchor and state-known PCOM as references, matching the comparisons defined in Eq. (4).

F.1 Relation to missingness and multimodal learning

Missing-data theory identifies when selection can be ignored and which quantities can be recovered from observed information (Rubin, 1976; Little & Rubin, 2019; Seaman et al., 2013; Mohan &

Method	Target input	Implementation
ERM+Mask	None	Common feature anchor with mask input
ModDrop-M	None	Random complete-modality dropping composed with available masks
Random-LPG	None	Source-side random-policy exposure control
AdvDrop-M	None	Modality-level hard-mask adaptation of adversarial dropout
MultiMoDN-F	None	Feature-level modular fusion and missing-modality skipping
GroupDRO-M	None	Availability-pattern groups under a robust objective
DAMS-Mask-IW	R	Clipped target/source mask-frequency weighting
PCOM	(C, R)	Learned policy family, paired-log inference, and mass correction
State-known PCOM	State or coordinate	Learned reference with the state or coordinate supplied
Posterior-average control	True α	Source/endpoint posterior average, not inferred-α mass mixing

Table 9: Baseline implementations and target information. DAMS uses mask frequencies. PCOM uses paired context-mask logs and its source-trained policy family.

Pearl, 2021). Missingness-shift analysis further explains how a non-ignorable change can alter the Bayes predictor (Rockenschaub et al., 2024). These results motivate estimating the policy-dependent selection factor rather than assigning a fixed meaning to every availability pattern.

Multimodal learning addresses representation, alignment, fusion, and co-learning (Baltrušaitis et al., 2019). Methods for incomplete inputs use modality dropping (Neverova et al., 2016), severe-missingness training (Ma et al., 2021), shared/specific features (Wang et al., 2023a), and modular fusion (Swamy et al., 2023). Expert-routing models provide additional ways to combine available modalities (Xu et al., 2024; Yun et al., 2024; Han et al., 2024; Li et al., 2025), while prompting and recovery adapt or reconstruct incomplete representations (Lee et al., 2023; Hu et al., 2024; Dai et al., 2025). Our fixed-feature comparisons isolate posterior correction from these representation-learning choices.

Robust objectives and bias-mitigation methods reduce sensitivity to unstable signals (Sagawa et al., 2020; Ganin et al., 2016; Cadene et al., 2019; Kim et al., 2019; Park et al., 2018). Conversely, GRU-D and CRL-MMNAR exploit informative observation patterns (Che et al., 2018; Liang et al., 2025). This contrast motivates evaluating when availability should influence prediction and how that influence should change with the policy.

DAMS and missing-not-at-random (MNAR) domain-adaptation methods model changes in observation processes under specified assumptions (Zhou et al., 2023; Stokes et al., 2025), while MIRRAMS regularizes prediction against unseen tabular missingness shifts (Lee et al., 2025). PCOM uses source-trained observed masses and deployment logs to form an explicit policy-conditioned posterior update. Its defining choice is this correction target and information interface, rather than the use of missingness information alone.

G Controlled Posterior-Conflict Experiments

We use a controlled construction to test whether policy information changes the prediction target before attributing gains to a particular correction model. The source-specific, target-specific, shared, and state-aware predictors are evaluated with six seeds, numbered 41, 43, 47, 53, 59, and 61. Matching global mask rates and retaining common support helps separate posterior conflict from a change in overall modality availability.

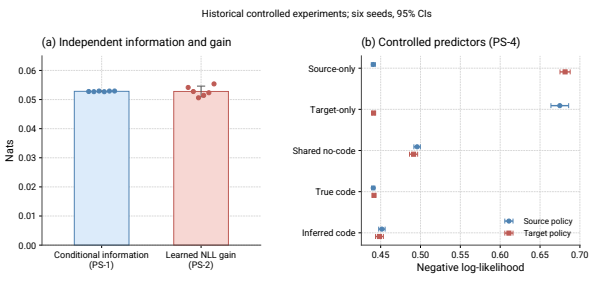

Figure 5: Controlled policy information and predictor NLL. Error bars are six-seed 95% confidence intervals (CIs).

The information comparison measures conditional information under the specified joint law and NLL reduction between

fitted predictors in separate calculations.
Pairing their values by seed tests whether
the learned improvement is consistent with the population log-loss identity.

Quantity	Mean $\pm$ sample standard deviation (SD)
$I(Y; Z \mid O)$	0.052825 ± 0.000112
Learned shared-to-state-known NLL reduction	0.052780 ± 0.001749
Paired difference (learned reduction $-$ information)	-0.000044 $[-0.001870,\ 0.001781]$

Table 10: Controlled information and learned-risk comparison over six seeds. The first two entries are mean $\pm$ sample SD. The difference uses a paired two-sided 95% CI.

The estimated information is 0.052825 nats and the learned reduction is 0.052780 nats. Their paired difference is close to zero and its CI includes zero. Together with the specialization/shared-prediction comparison, this pattern supports a role for policy information beyond the observed record, while retaining the distinction between a population information quantity and a fitted predictor's risk reduction.

H Extended Endpoint Results

We report area under the receiver operating characteristic curve (AUROC) and expected calibration error (ECE) alongside NLL to examine whether the correction changes more than the primary proper loss. Each table first averages the eight endpoints within a seed and then summarizes five seeds. ERM+Mask supplies the source-anchor reference, and state-known PCOM isolates the additional effect of providing the true state. Dashes indicate unavailable diagnostic metrics.

Method	NLL↓	AUROC↑	Brier↓	ECE↓
ERM+Mask	0.6620 ± 0.0091	0.6838 ± 0.0066	0.2273 ± 0.0032	0.1097 ± 0.0199
ModDrop-M	0.6108 ± 0.0370	0.7376 ± 0.0330	0.2076 ± 0.0118	0.0880 ± 0.0133
Random-LPG	0.6417 ± 0.0461	0.6970 ± 0.0469	0.2215 ± 0.0156	0.1006 ± 0.0109
AdvDrop-M	0.5864 ± 0.0156	0.7419 ± 0.0195	0.2008 ± 0.0054	0.0525 ± 0.0131
MultiMoDN-F	0.6163 ± 0.0210	0.7159 ± 0.0305	0.2162 ± 0.0057	0.1067 ± 0.0262
GroupDRO-M	0.6360 ± 0.0197	0.7084 ± 0.0272	0.2174 ± 0.0097	0.0969 ± 0.0187
DAMS-Mask-IW	0.6410 ± 0.0199	0.7035 ± 0.0240	0.2199 ± 0.0051	0.0976 ± 0.0112
PCOM	0.2280 ± 0.0542	0.9633 ± 0.0165	0.0701 ± 0.0163	0.0147 ± 0.0028
State-known PCOM	0.2277 ± 0.0545	0.9634 ± 0.0166	-	-

Table 11: CMU-MOSEI endpoint results. Mean $\pm$ sample SD over five seed means, each averaging eight endpoints.

Method	NLL↓	AUROC↑	Brier↓	ECE↓
ERM+Mask	1.5938 ± 0.0310	0.5552 ± 0.0270	0.1038 ± 0.0021	0.0789 ± 0.0174
ModDrop-M	1.5487 ± 0.0341	0.5841 ± 0.0328	0.1016 ± 0.0017	0.0796 ± 0.0161
Random-LPG	1.5402 ± 0.0274	0.5970 ± 0.0218	0.1002 ± 0.0016	0.0743 ± 0.0212
AdvDrop-M	1.4732 ± 0.0397	0.6167 ± 0.0440	0.0981 ± 0.0030	0.0651 ± 0.0059
MultiMoDN-F	1.5395 ± 0.0265	0.5868 ± 0.0260	0.1012 ± 0.0021	0.0798 ± 0.0136
GroupDRO-M	1.5719 ± 0.0475	0.5820 ± 0.0253	0.1032 ± 0.0033	0.0669 ± 0.0184
DAMS-Mask-IW	1.5569 ± 0.0297	0.5927 ± 0.0208	0.1019 ± 0.0022	0.0743 ± 0.0160
PCOM	1.0336 ± 0.0289	0.8450 ± 0.0175	0.0735 ± 0.0015	0.0430 ± 0.0110
State-known PCOM	1.0336 ± 0.0290	0.8451 ± 0.0175	-	-

Table 12: MELD endpoint results. Mean $\pm$ sample SD over five seed means, each averaging eight endpoints.

PCOM improves NLL and AUROC over the listed baselines on all four endpoint datasets. For example, CMU-MOSEI NLL decreases from 0.6410 for DAMS to 0.2280 for PCOM, accompanied

Method	NLL↓	AUROC↑	Brier↓	ECE↓
ERM+Mask	0.2716 ± 0.0037	0.7649 ± 0.0096	0.0799 ± 0.0013	0.0161 ± 0.0029
ModDrop-M	0.2705 ± 0.0035	0.7627 ± 0.0070	0.0795 ± 0.0012	0.0137 ± 0.0029
Random-LPG	0.2728 ± 0.0031	0.7575 ± 0.0098	0.0800 ± 0.0012	0.0159 ± 0.0017
AdvDrop-M	0.2722 ± 0.0036	0.7529 ± 0.0074	0.0798 ± 0.0012	0.0109 ± 0.0025
MultiMoDN-F	0.2815 ± 0.0061	0.7342 ± 0.0068	0.0826 ± 0.0020	0.0198 ± 0.0043
GroupDRO-M	0.2721 ± 0.0037	0.7599 ± 0.0055	0.0798 ± 0.0014	0.0155 ± 0.0034
DAMS-Mask-IW	0.2721 ± 0.0037	0.7603 ± 0.0067	0.0798 ± 0.0013	0.0158 ± 0.0020
PCOM	0.2627 ± 0.0035	0.7871 ± 0.0056	0.0777 ± 0.0012	0.0130 ± 0.0031
State-known PCOM	0.2627 ± 0.0036	0.7876 ± 0.0055	-	-

Table 13: MM-IMDb endpoint results. NLL averages Bernoulli losses across labels. Entries are mean $\pm$ sample SD over five endpoint-aggregated seed means.

Method	NLL↓	AUROC↑	Brier↓	ECE↓
ERM+Mask	2.4748 ± 0.0603	0.9094 ± 0.0054	0.0056 ± 0.0001	0.0254 ± 0.0027
ModDrop-M	2.4425 ± 0.0608	0.9115 ± 0.0046	0.0056 ± 0.0001	0.0250 ± 0.0020
Random-LPG	2.4525 ± 0.0586	0.9111 ± 0.0043	0.0056 ± 0.0001	0.0264 ± 0.0032
AdvDrop-M	2.4327 ± 0.0611	0.9120 ± 0.0046	0.0056 ± 0.0001	0.0246 ± 0.0016
MultiMoDN-F	2.4352 ± 0.0674	0.9122 ± 0.0048	0.0056 ± 0.0001	0.0247 ± 0.0032
GroupDRO-M	2.6908 ± 0.0626	0.8790 ± 0.0086	0.0061 ± 0.0001	0.0699 ± 0.0059
DAMS-Mask-IW	2.4491 ± 0.0579	0.9111 ± 0.0042	0.0056 ± 0.0001	0.0243 ± 0.0009
PCOM	2.2661 ± 0.0612	0.9232 ± 0.0046	0.0055 ± 0.0001	0.0140 ± 0.0016
State-known PCOM	2.2661 ± 0.0612	0.9232 ± 0.0046	-	-

Table 14: UPMC Food-101 endpoint results on the processed features described in section D. Mean $\pm$ sample SD over five seed means.

by an AUROC increase from 0.7035 to 0.9633. The smaller MM-IMDb NLL improvement is also accompanied by higher AUROC, from 0.7603 to 0.7871.

Calibration metrics provide a complementary view. PCOM has lower ECE than the listed baselines on CMU-MOSEI, MELD, and UPMC, whereas AdvDrop-M has lower MM-IMDb ECE (0.0109 versus 0.0130). Thus the NLL improvement need not imply the best binned calibration score in every task. All method comparisons use the information interfaces in section F.

I Predictive Gains and Held-Out Mixtures

We extend the main results by comparing all-cell utility, fidelity to state-known PCOM, and recovery where the reference gain is large enough to normalize reliably. Keeping these views separate allows a small state-inference gap to be interpreted without assuming that the underlying correction is useful.

I.1 Paired estimation across seeds

Policy cells have equal weight within each seed. A paired effect is computed on matching cells before the within-seed average, so both methods are compared on the same conditions. For seed differences d_s, we report

$$\bar{d} \pm t_{.975,S-1} \frac{s_d}{\sqrt{S}}. \tag{30}$$

where $S = 5$ for the feature experiments. These intervals describe variation across independently fitted seeds on the fixed processed data. Secondary intervals are unadjusted. Positive-cell counts describe the distribution of outcomes rather than independent training replications.

The evaluation contains 160 endpoint, 800 continuous, 800 mixture, and 80 hidden-content cells. All five confirmation seeds contribute to unconditional metrics. The development smoke seed is excluded. Among the 1,840 evaluated cells, 438 have negative known-state or inferred-state gain. The direct gaps show that inferred and known-state predictions can remain close even when both underperform the source anchor. This is why recovery is reported as a separate, conditional measurement.

Dataset	Regime	G^{known} [95% CI]	G^{inf} [95% CI]	Gap to known [95% CI]	Positive cells
CMU-MOSEI	endpoint	0.4343 [0.3639, 0.5047]	0.4339 [0.3639, 0.5039]	0.00036 [−0.00054, 0.00127]	40/40
CMU-MOSEI	mixture	0.1440 [0.1206, 0.1674]	0.1437 [0.1206, 0.1669]	0.00027 [−0.00033, 0.00088]	200/200
CMU-MOSEI	continuous	−0.0388 [−0.1372, 0.0595]	−0.0391 [−0.1383, 0.0601]	0.00027 [−0.00062, 0.00116]	112/200
CMU-MOSEI	hidden content	−0.0038 [−0.0060, −0.0016]	−0.0033 [−0.0050, −0.0017]	−0.00049 [−0.00105, 0.00007]	1/40
MELD	endpoint	0.5602 [0.4889, 0.6315]	0.5602 [0.4889, 0.6314]	0.00004 [−0.00004, 0.00012]	40/40
MELD	mixture	0.1912 [0.1678, 0.2146]	0.1902 [0.1665, 0.2139]	0.00099 [−0.00008, 0.00206]	200/200
MELD	continuous	−0.0469 [−0.0617, −0.0321]	−0.0469 [−0.0617, −0.0320]	−0.00002 [−0.00031, 0.00027]	101/200
MELD	hidden content	−0.0093 [−0.0122, −0.0064]	−0.0091 [−0.0114, −0.0068]	−0.00018 [−0.00108, 0.00071]	0/40
MM-IMDb	endpoint	0.0089 [0.0043, 0.0135]	0.0088 [0.0044, 0.0132]	0.00006 [−0.00011, 0.00024]	40/40
MM-IMDb	mixture	0.0030 [0.0013, 0.0048]	0.0027 [0.0012, 0.0043]	0.00030 [0.00004, 0.00055]	172/200
MM-IMDb	continuous	0.0025 [−0.0003, 0.0053]	0.0023 [−0.0003, 0.0050]	0.00019 [−0.00017, 0.00054]	123/200
UPMC Food-101	endpoint	0.2087 [0.1999, 0.2176]	0.2087 [0.1999, 0.2176]	0.00000 [−0.00000, 0.00000]	40/40
UPMC Food-101	mixture	0.0710 [0.0691, 0.0729]	0.0705 [0.0690, 0.0720]	0.00051 [−0.00018, 0.00120]	200/200
UPMC Food-101	continuous	0.0653 [0.0469, 0.0837]	0.0652 [0.0470, 0.0834]	0.00015 [−0.00011, 0.00041]	136/200

Table 15: All-cell gains relative to the frozen source anchor and NLL gaps to state-known PCOM. Each 95% CI uses five paired seed aggregates. Positive cells count settings with positive inferred gain.

Dataset	Regime	Recovery	Eligible seeds	Eligible cells
CMU-MOSEI	endpoint	0.9992 ± 0.0015	5	40/40
CMU-MOSEI	mixture	1.0086 ± 0.0155	5	192/200
CMU-MOSEI	continuous	1.0041 ± 0.0101	5	110/200
CMU-MOSEI	hidden content	-	0	0/40
MELD	endpoint	0.9999 ± 0.0001	5	40/40
MELD	mixture	1.0188 ± 0.0095	5	200/200
MELD	continuous	1.0018 ± 0.0046	5	95/200
MELD	hidden content	-	0	0/40
MM-IMDb	endpoint	0.9905 ± 0.0097	4	32/40
MM-IMDb	mixture	0.9086 ± 0.0409	4	56/200
MM-IMDb	continuous	1.0013 ± 0.0047	4	65/200
UPMC Food-101	endpoint	1.0000 ± 0.0000	5	40/40
UPMC Food-101	mixture	1.0008 ± 0.0199	5	168/200
UPMC Food-101	continuous	0.9968 ± 0.0065	5	136/200

Table 16: Recovery on cells with $G^{\text{known}} > 0.005$. Entries are mean $\pm$ sample SD over eligible seed means. Counts show the evaluated subset. Seeds without eligible cells remain in unconditional NLL summaries.

I.2 Recovery on material-gain cells

We normalize inferred gain by state-known gain only when $G^{\text{known}} > 0.005$, avoiding ratios dominated by a near-zero denominator. Ratios are computed per eligible cell and averaged within each seed before reporting the mean and SD across eligible seeds. A seed with no eligible cells has undefined recovery but remains part of all-cell performance.

MM-IMDb has fewer eligible cells than the other datasets, reflecting smaller learned state-conditioned gains. Its endpoint recovery excludes seed 827, which has no material endpoint cells, while its unconditional effects retain all five seeds. Neither hidden-content family has an eligible cell at this threshold. Values slightly above one are possible because the denominator uses a learned reference and the inferred prediction averages state branches.

Both axes in the gain map subtract the same anchor loss. We therefore use the paired NLL gaps in Table 15 to measure state-known fidelity, and use the sign of each unconditional gain to assess task improvement. This separates the two conclusions that can otherwise appear similar in a diagonal scatter plot.

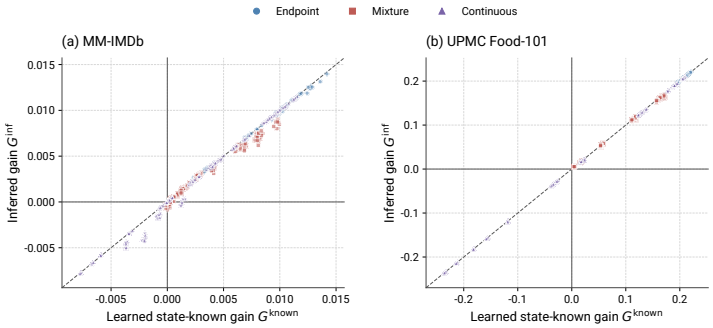

Figure 6: Inferred versus state-known gains for MM-IMDb and UPMC Food-101. All confirmation cells are shown. Direct gaps appear in Table 15.

I.3 Held-out mixture proportions

We test whether mass-space correction remains useful between fitted proportions on the eight known source-endpoint pairs. The six fitting and five evaluation proportions are given in section E. The endpoint identity is supplied and the proportion is inferred, while the task and mass models remain fixed. All four datasets have positive unconditional mean mixture gain over the source anchor, with positive five-seed CIs.

To compare the two normalization choices, we also average endpoint posteriors using the true proportion. The paired differences below contrast this control with PCOM's inferred-proportion mass correction.

Dataset	True-α posterior average − inferred-α mass mix
CMU-MOSEI	$0.0044\ [-0.0010,\ 0.0099]$
MELD	$-0.0016\ [-0.0030,\ -0.0002]$
MM-IMDb	$-0.0004\ [-0.0008,\ -0.0000]$
UPMC Food-101	$0.0006\ [-0.0008,\ 0.0019]$

Table 17: Paired NLL of the true-proportion posterior average minus inferred-proportion mass correction. Positive values favor PCOM. Brackets show two-sided 95% CIs over five seed means. Both methods receive the endpoint identity.

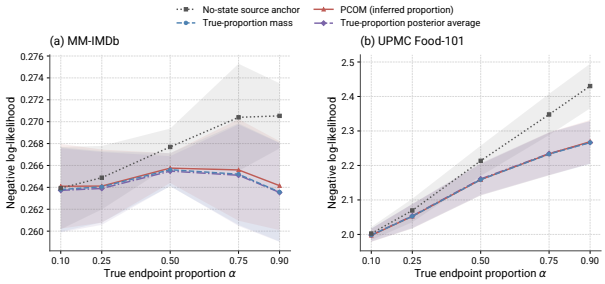

Figure 7: MM-IMDb and UPMC Food-101 mixtures. PCOM infers α. The control receives true α.

The posterior-average control has lower NLL on MELD and MM-IMDb, while the CMU-MOSEI and UPMC Food-101 intervals include zero. This separates the correct population construction from the quality of its estimated implementation. PCOM estimates both the mass and proportion, whereas the control receives the true proportion. Positive gains over the anchor and superiority over this control are therefore different experimental questions.

J Mechanism Studies and Mass Estimation

We examine the source of the correction by varying its class dependence, policy selection, source architecture, and mass estimator. The comparisons distinguish changes required by the update's algebra from choices whose practical benefit depends on estimation accuracy and the dataset.

J.1 Class dependence and policy selection

The endpoint ablation uses seeds 601, 607, and 613 and measures each variant's gain against the same study-specific anchor. Removing candidate-class dependence tests whether the update can change relative class probabilities. Removing the state term tests whether a policy-specific factor remains. Hard selection and the state-known reference then separate state uncertainty from mass estimation.

Variant	CMU-MOSEI gain	MM-IMDb gain	CMU-MOSEI AUROC	MM-IMDb AUROC
Full PCOM	0.3503 ± 0.0587	0.0102 ± 0.0019	0.9308 ± 0.0253	0.7800 ± 0.0052
No candidate-class input	0.0000 ± 0.0000	0.0000 ± 0.0000	0.6818 ± 0.0191	0.7518 ± 0.0056
No policy-state inference	0.0000 ± 0.0000	0.0000 ± 0.0000	0.6818 ± 0.0191	0.7518 ± 0.0056
Hard state selection	0.3506 ± 0.0589	0.0082 ± 0.0027	0.9309 ± 0.0253	0.7745 ± 0.0063
State-known PCOM	0.3502 ± 0.0586	0.0102 ± 0.0018	0.9308 ± 0.0252	0.7804 ± 0.0052
Analytic policy mass	0.5711 ± 0.0356	0.0074 ± 0.0003	0.9935 ± 0.0045	0.7789 ± 0.0049
No overlap regularizer	0.4294 ± 0.0960	0.0102 ± 0.0019	0.9597 ± 0.0334	0.7800 ± 0.0052
No source bypass	0.3512 ± 0.0590	0.0102 ± 0.0019	0.9311 ± 0.0253	0.7800 ± 0.0052

Table 18: Endpoint mechanism ablations over three seeds. Gains use the same study-specific source anchor. Analytic mass is directly available for the context/class-conditioned generator used here.

Both structural removals yield zero gain, as expected when the correction factor is absent or cancels during class normalization. Hard selection and marginalization are nearly tied on CMU-MOSEI (0.3506 and 0.3503 gain), while marginalization raises MM-IMDb gain from 0.0082 to 0.0102. Averaging state uncertainty is therefore useful in the MM-IMDb comparison, whereas a hard decision is already sufficient to match the endpoint average on CMU-MOSEI.

For the context/class-conditioned generator, analytic mass is directly available. Its CMU-MOSEI gain of 0.5711, compared with 0.3503 for learned PCOM, identifies room to improve the learned mass estimator. Its weaker MM-IMDb gain shows that replacing the fitted estimator is not uniformly beneficial across tasks. The additional overlap and source-bypass rows measure target-task effects. Branch preservation and weight behavior are examined by their corresponding derivations and diagnostics.

J.2 Changing the source architecture

To test whether correction also helps a different source predictor, we compare Deep Sets fusion and a MultiMoDN-style anchor on CMU-MOSEI and MM-IMDb. Endpoint results use seeds 601, 607, and 613. Each configuration is evaluated relative to its own frozen anchor, allowing us to ask whether it benefits from correction without assuming that the two anchors start at the same task risk.

Dataset	Anchor	Gain vs. own anchor	Recovery	AUROC
CMU-MOSEI	Deep Sets fusion	0.3472 ± 0.0643	0.9999 ± 0.0002	0.9289 ± 0.0283
CMU-MOSEI	MultiMoDN state fusion	0.4312 ± 0.0728	0.9998 ± 0.0000	0.9592 ± 0.0214
MM-IMDb	Deep Sets fusion	0.0104 ± 0.0022	0.9973 ± 0.0039	0.7796 ± 0.0064
MM-IMDb	MultiMoDN state fusion	0.0156 ± 0.0015	0.9979 ± 0.0039	0.7708 ± 0.0119

Table 19: Endpoint correction with two source architectures. Each gain uses the corresponding frozen anchor. Entries are mean $\pm$ sample SD over three seeds.

Both architectures benefit on both tasks. MultiMoDN-style correction gains are 0.4312 on CMU-MOSEI and 0.0156 on MM-IMDb, while Deep Sets gains are 0.3472 and 0.0104. Recovery is close

Dataset	Variant	NLL	Mass KL	Centered error	Posterior KL to reference
CMU-MOSEI	class_prior_donors	0.4292 ± 0.0107	0.1757 ± 0.0125	0.1621 ± 0.0337	0.0113 ± 0.0045
CMU-MOSEI	full_content_plugin	0.4241 ± 0.0092	1.1373 ± 0.0542	0.0897 ± 0.0041	0.0039 ± 0.0004
CMU-MOSEI	naive_observed_only	0.4241 ± 0.0092	0.1728 ± 0.0114	0.0897 ± 0.0041	0.0039 ± 0.0004
CMU-MOSEI	neighbors_32	0.4266 ± 0.0092	0.0135 ± 0.0017	0.0639 ± 0.0020	0.0021 ± 0.0002
CMU-MOSEI	neighbors_8	0.4301 ± 0.0095	0.0425 ± 0.0029	0.1046 ± 0.0049	0.0051 ± 0.0005
CMU-MOSEI	no_class_conditioning	0.4241 ± 0.0092	0.0255 ± 0.0024	0.0897 ± 0.0041	0.0039 ± 0.0004
CMU-MOSEI	no_cross_fitting	0.4320 ± 0.0090	0.0411 ± 0.0031	0.1108 ± 0.0043	0.0059 ± 0.0005
CMU-MOSEI	random_donor_representation	0.4262 ± 0.0096	0.1452 ± 0.0129	0.1035 ± 0.0047	0.0050 ± 0.0007
CMU-MOSEI	standard	0.4280 ± 0.0090	0.0241 ± 0.0025	0.0819 ± 0.0029	0.0033 ± 0.0003
MELD	class_prior_donors	1.2236 ± 0.0156	0.0914 ± 0.0043	0.2029 ± 0.0048	0.0236 ± 0.0017
MELD	full_content_plugin	1.1989 ± 0.0176	1.3602 ± 0.0633	0.1057 ± 0.0032	0.0065 ± 0.0007
MELD	naive_observed_only	1.1989 ± 0.0176	0.1156 ± 0.0037	0.1057 ± 0.0032	0.0065 ± 0.0007
MELD	neighbors_32	1.2050 ± 0.0182	0.0297 ± 0.0017	0.1175 ± 0.0022	0.0083 ± 0.0003
MELD	neighbors_8	1.2114 ± 0.0185	0.0509 ± 0.0033	0.1747 ± 0.0066	0.0155 ± 0.0007
MELD	no_class_conditioning	1.1989 ± 0.0176	0.0305 ± 0.0020	0.1057 ± 0.0032	0.0065 ± 0.0007
MELD	no_cross_fitting	1.2123 ± 0.0181	0.0458 ± 0.0031	0.1652 ± 0.0059	0.0147 ± 0.0013
MELD	random_donor_representation	1.2023 ± 0.0168	0.0362 ± 0.0015	0.1086 ± 0.0048	0.0069 ± 0.0007
MELD	standard	1.2082 ± 0.0178	0.0378 ± 0.0025	0.1427 ± 0.0035	0.0110 ± 0.0007

Table 20: Hidden-content mass estimators. Metrics are averaged over policies within each seed and summarized as mean ± SD over five seeds. Posterior KL uses a larger donor reference. The plug-in and observed-only variants retain the anchor in their task-prediction paths.

to one in all four cases, indicating that state inference preserves most of the improvement achieved by each configuration's state-known reference.

J.3 Donor representation and hidden-content estimation

The hidden-content study varies donor matching, neighborhood size, class dependence, and cross-fitting to examine how these choices affect estimated acquisition probabilities and task predictions. We compare mass error with a larger $K = 128$ donor approximation. That diagnostic uses a squared-distance kernel rather than the fitting kernel, may share source donors with training, and evaluates a deterministic target subset. Its posterior KL therefore measures agreement with the specified donor reference.

Increasing the neighborhood from the standard $K = 16$ to $K = 32$ lowers centered error and posterior KL to the reference on both datasets. On CMU-MOSEI, centered error decreases from 0.0819 to 0.0639 and posterior KL from 0.0033 to 0.0021. MELD changes from 0.1427 to 0.1175 and from 0.0110 to 0.0083. The associated NLL improvements show that donor size affects both the estimated correction and task prediction in these settings.

The `full_content_plugin` and `naive_observed_only` task paths return anchor probabilities, while their diagnostic mass vectors are repeated across candidate classes. The no-class variant also removes candidate dependence. Their equal downstream scores thus reflect a shared uncorrected task prediction despite differing acquisition-distribution diagnostics. The displayed labels are retained to match the estimator settings in the table.

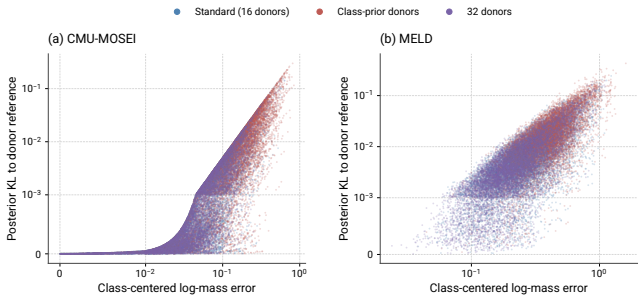

Figure 8: Class-centered log-mass error versus posterior KL to the $K = 128$ donor reference. All stored observations are shown.

The centered-error analysis follows the cancellation property in section B.6. Errors shared by all classes leave the normalized posterior unchanged, while class-relative errors can change it. The plotted relationship is measured against the larger donor reference rather than the population posterior. Across the tested hidden-content policies, mean task gains remain negative, so small reference discrepancies alone are insufficient to establish an improvement over the anchor. The measurements do not isolate a single cause of that behavior.

K Support and Policy Identifiability

We examine support and identifiability separately because a numerically stable ratio and an informative log window address different parts of prediction. The first concerns estimating the acquisition effect. The second concerns selecting which effect to apply.

K.1 Weight behavior on evaluated policies

For each continuous-policy target record, the audit divides the estimated target mass by the analytic source mass, using the realized class or label bit. It summarizes the resulting weights by

$$\widehat{\rho}_{\mathrm{target}} = \frac{(\sum_i w_i)^2}{n \sum_i w_i^2}, \tag{31}$$

This is a target-sample weight-dispersion statistic. The population ESS in section B instead takes its expectation under the source law. Minimum measured values are 0.4885, 0.4365, 0.8662, and 0.6218 on CMU-MOSEI, MELD, MM-IMDb, and UPMC Food-101. Unsupported and cap-exceeding fractions are zero on these continuous samples, with maximum weight 6.9628. The analytic uniform denominator provides positive support on the common catalog. The measurements describe the numerical weights used by the fitted correction.

K.2 Information in context-mask relationships

To determine where the state signal resides, the hidden-content study compares context-only, mask-only, shuffled-pair, and correctly paired log inputs. The shuffled control preserves the input types while disrupting their record-level relationship.

Input	CMU-MOSEI accuracy	MELD accuracy	CMU-MOSEI cross-entropy	MELD cross-entropy
Context only	0.1109 ± 0.0084	0.1055 ± 0.0101	2.1976 ± 0.0007	2.1981 ± 0.0007
Mask only	0.1082 ± 0.0169	0.1168 ± 0.0101	2.1978 ± 0.0008	2.1976 ± 0.0010
Shuffled pairs	0.1043 ± 0.0137	0.1047 ± 0.0134	2.1973 ± 0.0007	2.1973 ± 0.0007
Paired context-mask	0.2453 ± 0.0081	0.3840 ± 0.0218	1.9418 ± 0.0461	1.5251 ± 0.0477

Table 21: Log-input controls in the hidden-content study. Entries are mean $\pm$ sample SD over five seeds. Chance accuracy for nine equally sampled states is $1/9$.

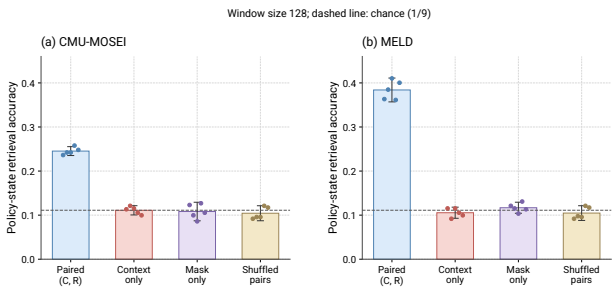

Figure 9: Hidden-content state retrieval from context and availability. Error bars are five-seed 95% CIs.

The individual and shuffled inputs remain near chance, whereas paired inputs reach accuracy 0.2453 on CMU-MOSEI and 0.3840 on MELD. Their cross-entropies also fall below the near-uniform control values. This pattern supports the role of the context-mask relationship in state estimation within the tested hidden-content interface.

The corresponding information boundary follows from equality of window laws. If two states induce the same one-record (C, R) distribution and records are independent within a fixed-state window, their product window distributions are also identical. Under equal binary priors, a window-only test then has chance Bayes error. For temporally dependent records the full window distributions must be considered, since one-record marginals alone do not characterize the joint window. The implemented encoder pools its inputs without using their order.

L Selective Prediction

Prediction sets supplement point estimates by expressing uncertainty through the set of labels retained. We first describe the source-resampling procedure used in the reported measurements, then derive how uncertainty over policy state enters a union of state-specific prediction sets.

L.1 Source-resampled calibration

For each continuous or mixture setting, source calibration records are disjoint from task-model training queries. The frozen simulator for that environment resamples their acquisition masks, after which PCOM infers the policy coordinate and computes calibration scores using the same prediction pipeline applied at test time. Bootstrap windows are formed from the corresponding fixed log pools.

The nonconformity score is $S_i = -\log \widehat{q}(Y_i \mid O_i)$. For n calibration records, the implementation uses zero-based sorted index $\min(n-1, \lceil (n+1)(1-\alpha) \rceil - 1)$ with $\alpha = .075$. If a prediction set is empty, its highest-probability label is added. The implementation clips to the largest available score for very small n. The formal split-conformal construction instead uses an infinite threshold when the required rank is $n+1$. Calibration is pooled for binary and multiclass tasks and performed separately for each MM-IMDb label.

Dataset	Regime	Min. coverage	Mean seed-min	Max. wrong singleton	Mean size	Below .90
CMU-MOSEI	continuous	0.8971	0.9110	0.1029	1.3433	3/200
CMU-MOSEI	mixture	0.8949	0.9085	0.1051	1.3270	4/200
MELD	continuous	0.9041	0.9199	0.0179	4.2789	0/200
MELD	mixture	0.9111	0.9212	0.0175	3.8424	0/200
MM-IMDb	continuous	0.9401	0.9411	0.0599	1.1268	0/200
MM-IMDb	mixture	0.9400	0.9411	0.0600	1.1265	0/200
UPMC Food-101	continuous	0.9186	0.9220	0.0035	25.6134	0/200
UPMC Food-101	mixture	0.9203	0.9228	0.0041	26.9776	0/200

Table 22: Simulator-matched source-resampling calibration. Nominal coverage is 0.925. The last column counts cells below the 0.90 diagnostic level. Wrong singletons are measured over all records, or label-wise record fractions for MM-IMDb.

The table reports both the minimum over evaluated environments and the mean of per-seed minima. CMU-MOSEI has three continuous and four mixture cells below the diagnostic level 0.90, with minima 0.8971 and 0.8949. This diagnostic level is distinct from the nominal 0.925. MELD and UPMC Food-101 attain larger coverage values with larger multiclass sets. MM-IMDb sizes refer to individual Bernoulli-label sets. Reading size together with coverage therefore requires the task's output space rather than a shared binary-size criterion.

L.2 COMBINING STATE-SPECIFIC SETS

To carry uncertainty about the acquisition state into a prediction set, consider state-specific split-conformal sets Γ_z and a candidate state collection $\mathcal{Z}(W)$ (Vovk et al., 2005; Romano et al., 2020; Angelopoulos & Bates, 2023). If $P(Y \notin \Gamma_Z) \leq \alpha$ and $P(Z \notin \mathcal{Z}(W)) \leq \delta$, then

$$P\left(Y \notin \bigcup_{z \in \mathcal{Z}(W)} \Gamma_z\right) \leq P(Y \notin \Gamma_Z) + P(Z \notin \mathcal{Z}(W)) \leq \alpha + \delta. \tag{32}$$

The first error event concerns label coverage under the true state. The second concerns excluding that state. Their union bound separates these two sources of uncertainty without requiring them to be independent. Obtaining the first bound through split conformal requires suitable calibration/test exchangeability for a fixed score. Obtaining the second requires a calibrated state-inclusion property.

The reported experiment evaluates simulator-matched label calibration for each acquisition environment. Applying the state-union construction to an unknown policy additionally requires a candidate state collection with the calibrated inclusion probability specified above.

M PREDICTION UNDER DYNAMIC POLICIES

We vary policy schedules and window sizes to examine how state inference responds when a window contains changing acquisition behavior. The hidden-content experiment uses a fixed state, an abrupt state-1-to-state-4 switch, a stepwise progression through rounded state IDs, recurring state-1/state-4 blocks, and windows mixing those states. The table retains the schedule name *gradual* for the stepwise progression. Window sizes are 16, 32, 64, 128, and 256.

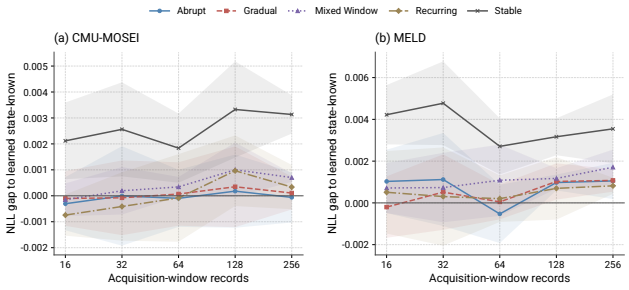

Figure 10: Window-size sensitivity of NLL gaps to state-known PCOM with five-seed 95% CIs.

Dataset	Schedule	$B = 16$	$B = 32$	$B = 64$	$B = 128$	$B = 256$
CMU-MOSEI	stable	0.00212 ± 0.00119	0.00256 ± 0.00146	0.00184 ± 0.00107	0.00333 ± 0.00149	0.00313 ± 0.00059
CMU-MOSEI	abrupt	-0.00031 ± 0.00083	-0.00001 ± 0.00154	-0.00010 ± 0.00089	0.00018 ± 0.00114	-0.00006 ± 0.00078
CMU-MOSEI	gradual	-0.00010 ± 0.00085	-0.00008 ± 0.00116	0.00007 ± 0.00098	0.00035 ± 0.00125	0.00011 ± 0.00051
CMU-MOSEI	recurring	-0.00074 ± 0.00064	-0.00041 ± 0.00106	-0.00008 ± 0.00136	0.00097 ± 0.00110	0.00034 ± 0.00069
CMU-MOSEI	mixed window	-0.00015 ± 0.00051	0.00019 ± 0.00074	0.00034 ± 0.00031	0.00100 ± 0.00085	0.00070 ± 0.00013
MELD	stable	0.00422 ± 0.00114	0.00477 ± 0.00161	0.00271 ± 0.00108	0.00317 ± 0.00071	0.00355 ± 0.00133
MELD	abrupt	0.00103 ± 0.00122	0.00112 ± 0.00180	-0.00053 ± 0.00112	0.00097 ± 0.00031	0.00105 ± 0.00070
MELD	gradual	-0.00020 ± 0.00119	0.00052 ± 0.00146	0.00005 ± 0.00067	0.00103 ± 0.00070	0.00108 ± 0.00044
MELD	recurring	0.00051 ± 0.00159	0.00030 ± 0.00190	0.00020 ± 0.00090	0.00070 ± 0.00121	0.00082 ± 0.00031
MELD	mixed window	0.00071 ± 0.00097	0.00073 ± 0.00134	0.00108 ± 0.00138	0.00117 ± 0.00040	0.00171 ± 0.00068

Table 23: Empirical NLL gaps to state-known PCOM under hidden-content policy schedules. Entries are mean $\pm$ sample SD across five seeds. The schedule labelled gradual is a stepwise progression through state IDs.

The mean NLL gaps remain small across schedules, but do not decrease monotonically with window size. For example, stable CMU-MOSEI has a gap of 0.00184 at $B = 64$ and 0.00333 at $B = 128$. MELD has 0.00271 and 0.00317 at the same sizes. These comparisons measure sensitivity to the window used for policy inference rather than a universal benefit from longer windows.

The reference is state-known PCOM, so the gaps measure tracking fidelity. Their occasional negative values indicate that the inferred prediction has lower empirical NLL than this learned reference. Because the hidden-content corrections can themselves underperform the source anchor, the task gains in section I remain necessary to judge whether close tracking is useful.

N Inference Cost

We time the anchor, state estimation, and correction separately to identify where the inference overhead arises. Measurements use seed 811, warm models and caches, three synchronized repetitions, and batches of at most 512 target records. Two preflight checks verify idle, low-utilization selected graphics processing units (GPUs). Unrelated host activity may remain. State-inference time includes 64 bootstrap windows and is amortized over the timed record batch.

Dataset	Anchor ms/record	State work ms/record	Correction ms/record
CMU-MOSEI	0.0140	0.0615	0.7035
MELD	0.0244	0.0554	1.8702
MM-IMDb	0.0606	0.0558	12.0145
UPMC Food-101	0.1218	0.0551	24.1150

Table 24: Amortized warm inference time in milliseconds per record. Stages are timed separately on selected preflight-idle devices. Feature extraction, model loading, and full calibration are excluded.

Correction is the largest timed stage at 0.70-24.12 ms per record. These warm-inference figures exclude feature extraction, loading, calibration, and training.