ACQUISITION-ORIENTED INCOMPLETE MULTI-VIEW CLUSTERING UNDER LIMITED BUDGETS

Anonymous authors
Paper under double-blind review

ABSTRACT

Incomplete multi-view clustering (IMVC) typically recovers missing view information from the observed data, overlooking practical scenarios where missing sample-view entries can instead be directly acquired at a cost. This raises a fundamental question: how should a limited acquisition budget be allocated to improve multi-view clustering? To address this problem, we propose Evidence-Adaptive Acquisition (EAA), a fully unsupervised framework for acquisition-aware incomplete multi-view clustering. Specifically, EAA first assesses whether the available observations in each view are sufficient to support reliable entry-level probabilistic modeling. With sufficient observations, EAA quantifies the acquisition value of each missing entry through its conditional information gain about the latent cluster assignment and selectively acquires the most informative entries. Under insufficient observations, EAA avoids unreliable fine-grained utility estimation and instead selects views by residuals and samples eligible missing entries uniformly within those views. The acquisition policy iteratively incorporates newly acquired entries to update the model and reassess the remaining candidates, enabling adaptive budget allocation under a strict acquisition budget. Extensive experiments on multiple datasets demonstrate that EAA effectively allocates limited acquisition budgets and improves clustering performance over existing acquisition and incomplete multi-view clustering methods.

1 INTRODUCTION

Multi-view clustering (MVC) exploits agreement and complementarity across multiple views to discover a common partition without class labels, and has been extensively studied (Eisenberg et al., 2025; Geng et al., 2025; Liu et al., 2026). However, conventional MVC generally assumes that every sample is observed in every view. This assumption is violated when measurements are unavailable because of sensor failure, omitted examinations, or missing descriptors. Incomplete multi-view clustering (IMVC) therefore addresses this gap by recovering missing representations or integrating the structures revealed by observed views (Hu & Chen, 2018; Liu et al., 2019; Wu et al., 2024).

Existing IMVC methods primarily differ in how they exploit observed entries to compensate for missing ones (Wen et al., 2023; Yu et al., 2025). For example, COMPLETER (Lin et al., 2021) learns consistent cross-view representations via contrastive learning and reconstructs missing views through dual prediction. SFCR (Jiang et al., 2025) instead aggregates similarities among observed samples into a unified graph to guide missing-sample recovery, while ToRES (Yu et al., 2024) uses prototype-based learning to avoid costly sample-level relationship modeling and directly optimize cluster assignments. These approaches progressively exploit cross-view representations, sample relationships, or cluster prototypes to compensate for missing information. In many applications, however, a missing sample–view entry is not necessarily permanently unavailable. It may instead be obtained through an additional measurement, computation, or data collection process, albeit at an extra cost. This setting suggests a different question:

Key Question. *Rather than inferring all missing entries from the existing observations, can we selectively acquire the entries that are most valuable for clustering under a limited budget?*

This shift from passively recovering missing information to actively acquiring informative observations gives rise to the problem of *active incomplete multi-view clustering* (AIMVC).

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.

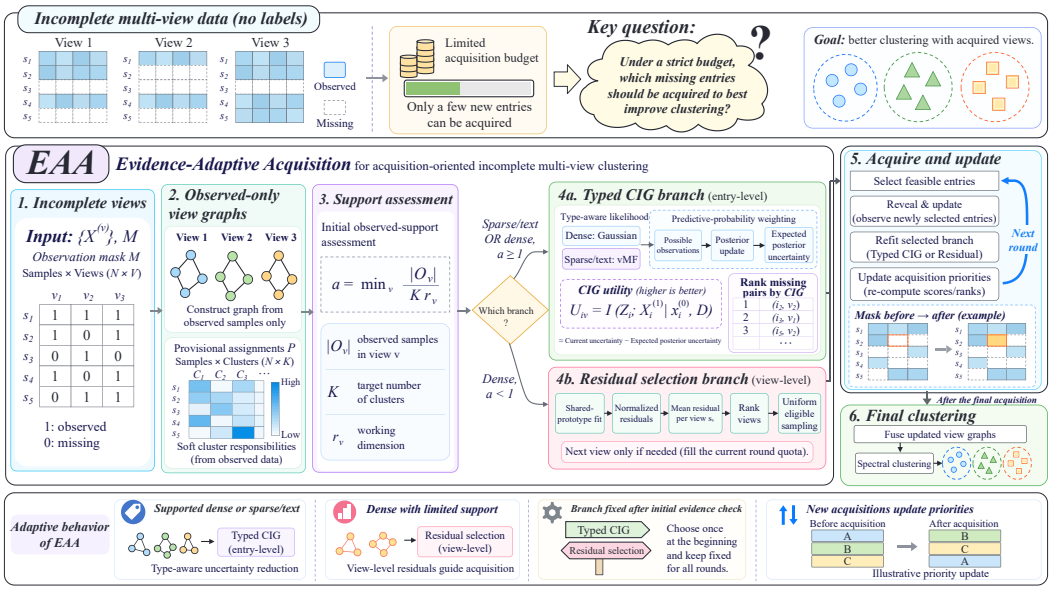

Figure 1: Overview of Evidence-Adaptive Acquisition (EAA).

As a relatively new research problem, AIMVC raises a simple but challenging question: given a limited acquisition budget, which missing entries should we acquire? The difficulty is that different missing entries can contribute very differently to clustering. Some entries may provide useful information for identifying the underlying clusters, whereas others may add little beyond what is already observed. However, before acquiring an entry, its actual benefit is unknown and may be difficult to estimate when the available observations are limited or uneven across views. Moreover, the decision cannot be made once and for all. After some missing entries are acquired, the newly available information changes what we already know about the data, and therefore can also change which of the remaining entries are most useful. Thus, an effective AIMVC method should identify promising entries based on the current observations and update its acquisition decisions as new information is obtained.

To address these challenges, we propose Evidence-Adaptive Acquisition (EAA), a budgeted acquisition framework for AIMVC that selects missing entries according to how reliably their potential benefits can be assessed. Figure 1 provides an overview of the proposed framework. EAA first examines whether the observed samples in each view provide sufficient evidence to support reliable entry-level evaluation. When sufficient evidence is available, EAA estimates the benefit of acquiring each missing entry through typed conditional information gain, with likelihood models tailored to the characteristics of the corresponding view. This enables the acquisition budget to focus on entries that are expected to provide the most useful information about the latent cluster structure. When the available evidence is insufficient for reliable entry-level evaluation, EAA avoids making potentially unreliable fine-grained decisions. Instead, it first identifies views with larger unexplained structural residuals, indicating that additional observations in these views may provide greater complementary information, and then acquires eligible missing entries from the selected views. After each acquisition round, the newly acquired entries are incorporated into the observed data, the corresponding models are updated, and the remaining candidates are reevaluated. In this way, EAA progressively updates its acquisition decisions, allowing the limited budget to be allocated to missing entries that are both informative and reliably assessable. Our contributions are:

- We introduce active incomplete multi-view clustering, a new setting that enables missing sample–view entries to be actively acquired under a limited budget. This shifts the focus from passively recovering missing information to actively acquiring the most informative missing entries.
- We propose Evidence-Adaptive Acquisition (EAA), a fully unsupervised framework that adapts acquisition granularity to the available evidence. With sufficient evidence, EAA

estimates candidate values via typed conditional information gain; otherwise, it performs view-level selection based on unexplained structural residuals to avoid unreliable fine-grained decisions.

- We develop an iterative acquisition strategy that incorporates newly acquired entries into the observed data and updates the acquisition model and remaining candidates after each round, enabling adaptive budget allocation as new information becomes available.

2 Related Work

Incomplete multi-view clustering. Incomplete multi-view clustering (IMVC) learns a common partition from partially observed multi-view data under a fixed observation pattern. Existing methods include traditional and deep approaches (Jin et al., 2023; Xu et al., 2024). Traditional methods use sample relationships to recover missing information for clustering. Graph- and subspace-based approaches learn shared structures from view-specific representations (Zhao et al., 2023). Specifically, Liu et al. (2025) remove biased errors from view-specific affinity representations and model cross-view interactions through consensus-guided tensor learning, while Wang et al. (2026) couple view recovery with tensorial subspace learning through a shared latent representation. Anchor- and prototype-based approaches reduce computational cost using compact representatives (Wang et al., 2022). Yu et al. (2024) replace sample-level self-expression with prototype–sample affinities and directly optimize discrete cluster assignments. Xu et al. (2025) jointly learn and complete anchor–sample similarities through latent embedding reconstruction, while Wang et al. (2025) leverage anchors for both missing-view completion and direct cluster-label generation. Deep IMVC methods employ neural networks to learn latent representations that integrate missing-information recovery with clustering (Jin et al., 2025). Lin et al. (2021) combine cross-view contrastive learning with dual prediction to recover missing-view representations, while Tang & Liu (2022) dynamically impute missing views from semantic neighbors and use bi-level optimization to account for the reliability of imputed samples. Du et al. (2026) couple missing-view recovery with clustering by estimating missing cluster assignments and reconstructing latent features from within-view and within-cluster statistics.

Active clustering and information acquisition. Active information acquisition selects additional observations or supervision to improve learning under a budget, and the acquisition object distinguishes the relevant settings. In active clustering, queries typically provide semantic supervision or reduce the information required to recover a partition. For example, Basu et al. (2004) select informative must-link and cannot-link constraints, Ashtiani et al. (2016) study the query complexity of partition recovery through same-cluster queries under separation assumptions, Eriksson et al. (2011) reduce the information required for hierarchical clustering through adaptively selected pairwise similarities, and Zhao et al. (2024) acquire human annotations for multi-view clustering. In feature-value acquisition, Saar-Tsechansky et al. (2009) rank candidate measurements by their expected utility for classifier induction, and Ma et al. (2019) model the conditional distribution of unobserved outcomes to guide acquisition through expected information gain. Li & Oliva (2021) use generative surrogate models to plan sequential feature acquisition, with an unsupervised extension whose goal is to infer the remaining unobserved values. At the modality level, Rheude et al. (2025) allocate missing measurements across a test cohort to improve supervised prediction. Within this broader acquisition setting, AIMVC considers fully unsupervised clustering of incomplete multi-view data, where each query reveals a missing sample–view representation and its value must be estimated solely from the incomplete observations, without access to ground-truth cluster assignments.

3 Method

3.1 Problem formulation

We consider an incomplete multi-view dataset with N aligned samples and V views, denoted by $\mathcal{X} = \{\mathbf{X}^{(v)}\}_{v=1}^{V}$, where $\mathbf{X}^{(v)} \in \mathbb{R}^{N \times d_v}$. The observation status of each sample-view pair is represented by a binary mask $\mathbf{M} \in \{0,1\}^{N \times V}$, where $M_{iv} = 1$ indicates that view v of sample i is observed and $M_{iv} = 0$ otherwise. We assume that every sample has at least one observed view. The set of initially missing entries is therefore $\Omega(\mathbf{M}) = \{(i,v) : M_{iv} = 0\}$.

Unlike conventional IMVC, which treats the missingness pattern as fixed, AIMVC allows selected missing entries to be actively acquired before clustering. Specifically, acquiring $(i, v) \in \Omega(\mathbf{M})$ reveals the corresponding view $\mathbf{x}_i^{(v)}$ of sample i and incurs an acquisition cost. Given a budget B, the objective is to select $\mathcal{S} \subseteq \Omega(\mathbf{M})$ with $|\mathcal{S}| = B$ such that the acquired entries are most beneficial for subsequent clustering. In our experiments, all entries have the same acquisition cost; a heterogeneous-cost extension would replace the cardinality constraint with $\sum_{(i,v)\in\mathcal{S}} c_{iv} \leq B$.

The acquisition process is fully unsupervised: ground-truth labels are unavailable during mask generation, candidate evaluation, acquisition, and clustering, and are used only to evaluate the final clustering. We perform acquisition iteratively over $T = 3$ rounds. After each round, newly acquired entries are incorporated into the observed data and the remaining candidates are reevaluated. Within each round, each sample can acquire at most one new view.

3.2 Evidence-Adaptive Acquisition

EAA is designed around the principle that the granularity of an acquisition decision should match the amount of evidence available for evaluating that decision. When the observed data provide sufficient support, EAA evaluates individual missing sample-view entries; when such evidence is insufficient, it avoids unreliable fine-grained comparisons and makes decisions at the view level. Figure 1 summarizes the initial branch assignment and the subsequent acquisition rounds.

3.2.1 Provisional structure and support assessment

At each acquisition stage, EAA constructs a k-nearest-neighbor graph for each view v using only observed samples, $\mathcal{O}_v = \{i : M_{iv} = 1\}$. Dense continuous data use Gaussian affinities; sparse count/text data use cosine affinities. No edges are introduced for unobserved samples. Averaging the available view graphs gives a sparse consensus graph $\mathbf{A}$, whose spectral embedding provides provisional cluster assignments $\mathbf{P} \in [0, 1]^{N\times K}$ for acquisition-value estimation.

Provisional assignments do not ensure that each view has enough observations to fit a reliable cluster-conditional model. Our observed-support check compares observed counts with model capacity. With a globally fixed projection cap $r_{\max} = 32$, we define

$$\begin{aligned} r_v &= \min\{r_{\max}, d_v, |\mathcal{O}_v| - 1\}, \\ a &= \min_v \frac{|\mathcal{O}_v|}{K r_v}. \end{aligned} \tag{1}$$

The ratio a is the minimum observed-row count per cluster and working dimension across views. For dense data, the dataset-level condition $a \geq 1$ selects typed CIG, whereas $a < 1$ selects the residual branch. Declared sparse count/text data use directional typed CIG irrespective of a. The branch is determined once from the initial observations, the specified K, and declared feature type, without labels or downstream outcomes. This branch remains fixed during acquisition; only its fitted model and candidate priorities are updated as additional entries become observed.

3.2.2 Typed conditional information gain

For entry-level evaluation, EAA measures expected reduction in latent-cluster uncertainty through conditional mutual information under a cluster-conditional generative model.

Let $Z_i \in \{1, \ldots, K\}$ be the latent cluster. Conditioned on Z_i, observed views are modeled independently:

$$p(Z_i = k) = \pi_k, \qquad p(\mathbf{x}_i^{\mathcal{O}_i} \mid Z_i = k) = \prod_{u\in\mathcal{O}_i} p_u(\mathbf{x}_i^{(u)} \mid k). \tag{2}$$

For dense continuous views, observed features are standardized, projected to a common PCA working dimension, and modeled with a diagonal Gaussian

$$p_v(\mathbf{x} \mid k) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{vk}, \mathrm{diag}(\boldsymbol{\sigma}_{vk}^2)). \tag{3}$$

For sparse count/text views, features are ℓ_2-normalized and modeled with a directional von Mises–Fisher (vMF) likelihood (Banerjee et al., 2005),

$$p_v(\mathbf{x} \mid k) = C_{d_\star}(\kappa_{vk}) \exp\big(\kappa_{vk} \boldsymbol{\mu}_{vk}^\top \mathbf{x}\big), \qquad \|\boldsymbol{\mu}_{vk}\|_2 = \|\mathbf{x}\|_2 = 1, \tag{4}$$

Algorithm 1 EAA acquisition and refitting.

Require: Views $\mathcal{X}$, mask $\mathbf{M}$, clusters K, budget B, rounds T
Build observed-only view graphs and provisional assignments
Compute support score a from the initial mask by Equation 1
Set typed CIG branch if sparse/count or $a \geq 1$; otherwise set residual branch
for $t = 1, \ldots, T$ **do**
  if typed CIG branch **then**
    Fit the typed mixture on the current observed entries
    Score feasible pairs using Equation 6
    Select the top b_t pairs, at most one per sample
  else
    Refit shared prototypes and residual view scores
    Commit view quota and sample feasible pairs within each view
  end if
  Acquire entries and update $\mathbf{M}$
end for
Build the final consensus graph and return spectral clusters

where κ_{vk} is the concentration, $\boldsymbol{\mu}_{vk}$ is the unit mean direction, and $C_{d_\star}$ is the vMF normalizer in the common working dimension $d_\star$. This choice aligns the likelihood with cosine graph geometry. We refer to this feature-dependent choice between Gaussian and vMF models as the type-aware likelihood. The common projection dimension $d_\star$ is capped by $r_{\max}$, view dimensions, and current observed counts, consistently with the capacity used in Equation 1. Parameters and posterior responsibilities are fitted by EM from graph-aligned initialization. Only observed rows update each view's parameters.

For a missing pair (i, v), let $q_i(k) = p(Z_i = k \mid \mathbf{x}_i^{\mathcal{O}_i}, \mathcal{D})$ be the current posterior. Its utility is the conditional mutual information

$$U_{iv} = I\Big(Z_i; X_i^{(v)} \mid \mathbf{x}_i^{\mathcal{O}_i}, \mathcal{D}\Big) \tag{5}$$

$$= \mathcal{H}(q_i) - \mathbb{E}_{x \sim p(X_i^{(v)} \mid \mathbf{x}_i^{\mathcal{O}_i}, \mathcal{D})} \left[\mathcal{H}\Big(p(Z_i \mid \mathbf{x}_i^{\mathcal{O}_i}, x, \mathcal{D})\Big)\right]. \tag{6}$$

Larger U_{iv} indicates greater expected reduction in cluster uncertainty from revealing view v for sample i. We evaluate Equation 6 by normalizing cluster-conditional likelihoods over observed outcomes. On large datasets, deterministic grouping in likelihood-signature space gives at most 256 outcomes, each assigned the sum of its members' cluster-conditional masses. All observations contribute, but scoring uses the coarsened outcomes; Appendix F gives the construction.

CIG-based entry ranking orders missing entries by U_{iv}. We greedily accept the highest-scoring feasible pair until the round quota is met, skipping additional views for a sample already selected in that round. Ties are deterministic. The next round refits the model after revealing the selected entries.

3.2.3 Residual-based view selection

For dense data with $a < 1$, EAA fits an observed-only shared-prototype representation and prioritizes views with greater unexplained structure, avoiding under-supported entry-level posterior comparisons. Let e_{iv} be the normalized reconstruction residual of observed pair (i, v). The view residual is

$$s_v = \frac{1}{|\mathcal{O}_v|} \sum_{i \in \mathcal{O}_v} e_{iv} \tag{7}$$

A larger s_v indicates more observed structure unexplained by the shared representation. Residual-based view ranking therefore allocates budget in descending order of s_v.

Within each committed view, we draw uniformly at random from eligible missing entries without replacement, excluding samples already selected in the same round. If a view cannot fill the round quota, selection continues in the next-highest-residual view. This residual branch is used when dense data fail the support check in Equation 1.

3.2.4 ITERATIVE ACQUISITION AND EXACT BUDGET

Algorithm 1 summarizes the full procedure. Let $q = \lfloor B/T \rfloor$, $r = B - Tq$, and $b_t = q + \mathbf{1}[t \leq r]$ for $t = 1, \ldots, T$. The first r rounds receive one additional acquisition, so $\sum_{t=1}^{T} b_t = B$. Each branch selects b_t distinct feasible missing pairs. After the final acquisition, the updated observed-only view graphs are fused into a consensus affinity, from which a normalized spectral embedding and the final cluster assignments are computed (Ng et al., 2001).

With $G \leq 256$ predictive outcome groups on large data, typed scoring costs $O(NVGK)$ per round after fitting and grouping; sparse graph construction stores $O(VNk)$ edges. The residual branch uses prototype-sample quantities rather than an $N \times N$ self-representation. These choices allow the same implementation to process the two 30,000-sample benchmarks. By contrast, the official CAL and ARSL-IMVC formulations contain $N \times N$ optimization variables and are therefore not reported at that scale.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Datasets and acquisition protocol. We evaluate on ten multi-view datasets spanning different sample sizes, view counts, feature types, and cluster numbers, as summarized in Table 1.

For every dataset and seed, we generate a canonical mask with exactly 50% missing sample-view entries while preserving at least one observed view per sample. The acquisition budget is 30% of the initially missing entries. We use five seeds and three acquisition rounds. All methods receive the same initial mask, exact budget, revealed feature values, and final clustering pipeline. No class label is used before evaluation.

Table 1: Dataset statistics.

Dataset	Samples	Views	Clusters
Yale	165	3	15
ORL	400	3	40
BBCSport	544	2	5
Reuters-1200	1,200	5	6
COIL20	1,440	3	20
Mfeat	2,000	6	10
Caltech101-20	2,386	6	20
scene15	4,485	3	15
FashionMNIST-30k	30,000	3	10
EMNIST-30k	30,000	3	10

Baselines. We compare heuristic policies (Random, View-rate, and Uncertainty), protocol-adapted criteria (CAMA (Rheude et al., 2025), CAL (Liu et al., 2025), ASCR (Xu et al., 2025), and JVILG (Wang et al., 2025)), and mechanism-inspired proxies (DIMVC-HIA (Du et al., 2026) and ARSL-IMVC (Wang et al., 2026)). CAL, ASCR, and JVILG use outputs of released implementations; the two proxies do not invoke the original optimizers. Appendix G specifies each adaptation and score. All policies update priorities after each round and share its quota and one-new-view-per-sample limit.

Downstream clustering. All policies use observed-only k-nearest-neighbor graphs, affinity averaging over jointly observed views, and normalized spectral embedding (Ng et al., 2001). Normalized association selects among candidate discretizations on the eight standard datasets; K-means discretizes the embedding on the two 30,000-sample datasets. Appendix F gives the fixed decoder settings.

Implementation and evaluation. Shared settings are 15 graph neighbors, projection cap $r_{\max} = 32$, dense-view support threshold $a = 1$, and at most 256 predictive outcome groups on large data. We report five-seed clustering accuracy (ACC; optimal label permutation), normalized mutual information (NMI), and adjusted Rand index (ARI). Full means, sample standard deviations, and additional analyses appear in the appendix.

4.2 MAIN RESULTS

How much does the acquisition policy matter when the budget and the downstream decoder are held constant? Table 2 reports five-seed mean ACC, NMI, and ARI on ten datasets; best and second-best results are **bold** and underlined. EAA achieves the highest mean NMI on nine datasets, ARI on eight, and ACC on six. On ORL, NMI rises from Random's 77.31% to 83.79%, a gain of 6.48

Table 2: Clustering performance (five-seed means). Best results are **bold**; second-best results are underlined.

Metric	Method	Yale	ORL	BBCSport	Reuters	COIL20	Mfeat	Caltech20	scene15	FMNIST	EMNIST
ACC	EAA	**0.6255**	**0.7260**	**0.9691**	**0.3932**	0.7435	0.8272	**0.4545**	0.4164	0.5945	**0.6781**
	Random	0.5588	0.6355	0.9640	0.3700	0.7339	0.8366	0.4394	0.4088	0.6231	0.6595
	View-rate	0.5297	0.6290	0.9474	0.3438	0.7371	0.8327	0.4353	0.4077	0.6156	0.6730
	Uncertainty	0.5612	0.6400	0.9574	0.2770	**0.7464**	0.8254	0.4303	0.4154	**0.6407**	0.6731
	CAMA†	0.5515	0.6580	0.8860	0.3898	0.6879	0.8325	0.4334	0.3644	0.5989	0.5860
	CAL†	0.5261	0.7105	0.7982	0.2747	0.6889	0.8387	0.4189	0.4083	–	–
	ASCR†	0.4945	0.5950	0.9452	0.2625	0.7360	0.8284	0.4347	**0.4173**	0.5830	0.6110
	JVILG†	0.5248	0.6225	0.9353	0.3880	0.7264	0.8253	0.4340	0.4094	0.6000	0.6327
	DIMVC-HIA‡	0.5564	0.6400	0.9537	0.2730	0.7174	0.8552	0.4483	0.3667	0.5763	0.5645
	ARSL-IMVC‡	0.5164	0.6240	0.6618	0.2717	0.7186	**0.8655**	0.4267	0.4136	–	–
NMI	EAA	**0.6398**	**0.8379**	**0.9037**	**0.2722**	0.8343	**0.8729**	**0.5912**	**0.4185**	**0.6700**	**0.6453**
	Random	0.5746	0.7731	0.8835	0.2485	0.8129	0.8424	0.5670	0.3855	0.6430	0.6195
	View-rate	0.5619	0.7687	0.8627	0.2318	0.8154	0.8546	0.5642	0.3858	0.6492	0.6235
	Uncertainty	0.5785	0.7698	0.8677	0.1951	0.8238	0.8492	0.5590	0.3790	0.6298	0.6143
	CAMA†	0.5763	0.7910	0.7861	0.2595	0.8112	0.8306	0.5606	0.3470	0.5889	0.5269
	CAL†	0.5532	0.8209	0.7366	0.1822	0.7942	0.8695	0.5772	0.3902	–	–
	ASCR†	0.5153	0.7556	0.8454	0.1624	0.8275	0.8620	0.5772	0.3941	0.6084	0.5454
	JVILG†	0.5541	0.7685	0.8461	0.2538	0.8356	0.8441	0.5580	0.3755	0.6310	0.5934
	DIMVC-HIA‡	0.5783	0.7717	0.8604	0.1868	0.8324	0.8467	0.5628	0.3387	0.6097	0.5188
	ARSL-IMVC‡	0.5524	0.7669	0.5907	0.1860	**0.8359**	0.8544	0.5682	0.3972	–	–
ARI	EAA	**0.4383**	**0.6076**	**0.9176**	**0.1598**	**0.6989**	0.8076	**0.3230**	**0.2383**	0.4790	**0.5313**
	Random	0.3404	0.4849	0.9058	0.1383	0.6747	0.7859	0.3178	0.2342	0.4830	0.5085
	View-rate	0.3190	0.4788	0.8829	0.1092	0.6811	0.7976	0.3047	0.2318	0.4845	0.5213
	Uncertainty	0.3478	0.4893	0.8873	0.0653	0.6963	0.7905	0.2985	0.2340	**0.4894**	0.5159
	CAMA†	0.3458	0.5290	0.8173	0.1577	0.6308	0.7717	0.3050	0.1937	0.4448	0.3994
	CAL†	0.3146	0.5792	0.6712	0.0411	0.6180	**0.8104**	0.2973	0.2331	–	–
	ASCR†	0.2716	0.4467	0.8577	0.0310	0.6880	0.8003	0.3051	0.2345	0.4303	0.4391
	JVILG†	0.3108	0.4746	0.8743	0.1377	0.6813	0.7860	0.3034	0.2244	0.4645	0.4795
	DIMVC-HIA‡	0.3498	0.4914	0.8808	0.0460	0.6740	0.7925	0.3227	0.1883	0.4348	0.3876
	ARSL-IMVC‡	0.3026	0.4784	0.4563	0.0406	0.6639	0.8078	0.3034	0.2356	–	–

Full mean ± sample SD: Tables A3 and A4. Caltech20 = Caltech101-20; FMNIST/EMNIST = FashionMNIST-30k/EMNIST-30k; Reuters = Reuters-1200. † Protocol-adapted criterion; ‡ mechanism-inspired proxy (Appendix G). CAL and ARSL-IMVC are unavailable (–) on the two 30k datasets because their official formulations use $N \times N$ variables.

percentage points, and exceeds the CAL-based acquisition criterion's 82.09% by 1.70 points. On FashionMNIST-30k and EMNIST-30k, NMI reaches 67.00% and 64.53%, improving over Random by 2.70 and 2.58 points and leading the reported scalable acquisition baselines. The advantage is not uniform: another policy leads NMI on COIL20, and ACC gains are less consistent. With the same decoder and budget, EAA achieves stronger clustering results on both standard and large-scale datasets.

4.3 Analysis of the acquisition mechanism

EAA couples three design choices: the observed-support check that assigns branches, branch-specific acquisition ranking, and a type-aware likelihood. Figure 2 compares Full EAA with active-component ablations (w/o), using five paired seeds and shared initial masks, budgets, three rounds, and decoder. Removing the support check forces ORL and Yale into typed CIG; ORL NMI drops by 5.08 percentage points. Randomizing CIG-based entry ranking reduces BBCSport NMI by 3.68 points; randomizing residual-based view ranking reduces ORL NMI by 5.28 points. Both retain fitted models, branch assignments, and selection constraints. Replacing the directional likelihood with a Gaussian, with cosine graphs and observed-only initialization fixed, lowers BBCSport NMI from 90.37% to 86.15%. The effect is smaller on Reuters-1200 (27.22% to 26.48%), where the directional model wins in three of five paired seeds. These results support the residual branch under limited dense support, informative ranking in both branches, and a type-aware likelihood for typed CIG.

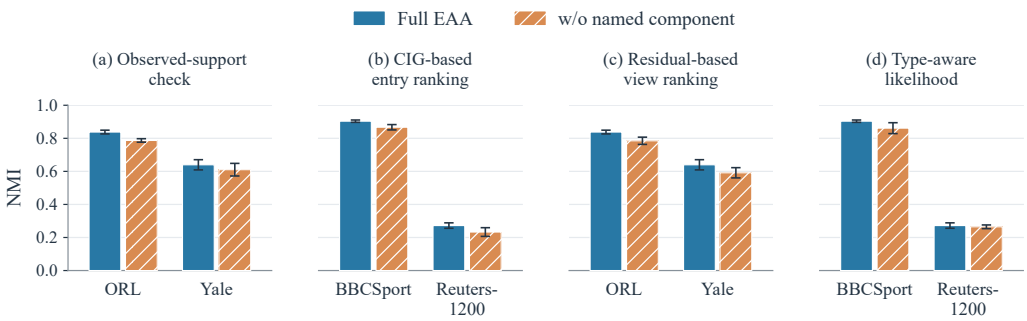

Figure 2: Acquisition-mechanism ablations. Bar heights denote mean NMI; error bars show one sample standard deviation over five paired seeds.

4.4 Performance under varying missingness and budgets

Beyond the default protocol, performance depends on initial missingness and the fraction of missing entries that may be acquired. Figure 3 varies these separately on Mfeat, ORL, and Reuters-1200. At 50% initial missingness, the budget ratio $B/|\Omega(\mathbf{M}_0)|$ takes values $\{10, 20, 30, 40, 50\}\%$; at a 30% budget ratio, missingness takes values $\{30, 50, 65\}\%$. Here $|\Omega(\mathbf{M}_0)|$ counts initially missing sample–view entries. EAA leads mean NMI in 19 of 21 distinct dataset–condition pairs, except Reuters-1200 at 10% and 20% budgets. On ORL, raising the budget from 10% to 30% increases NMI from 70.85% to 83.79%, while 40% to 50% adds only 0.12 percentage points. With increasing missingness, Mfeat remains nearly flat, ORL declines, and Reuters-1200 rises. Despite these different trends, EAA leads at every tested missingness level.

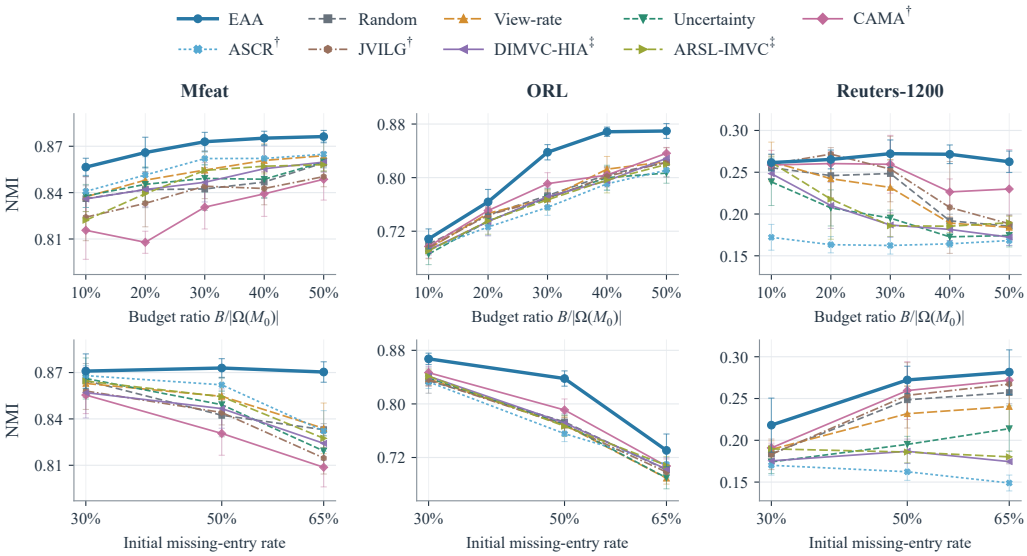

Figure 3: Performance under varying budgets (top) and missingness (bottom). Error bars: five-seed sample standard deviations. † Protocol-adapted criterion; ‡ mechanism-inspired proxy.

4.5 Hyperparameter sensitivity

We jointly vary assignment temperature γ and typed-mixture projection cap $r_{\max}$ on four dense typed-CIG datasets. Figure 4 reports five-seed paired NMI retention relative to $(\gamma, r_{\max}) = (0.5, 32)$ on a shared vertical scale. Across 100 dataset–parameter combinations, mean retention ranges from 98.21% to 101.36%. Mfeat is nearly flat along $r_{\max}$ because its view dimensions bound the working dimension below the projection cap. The broad plateau supports the same default parameters across these datasets.

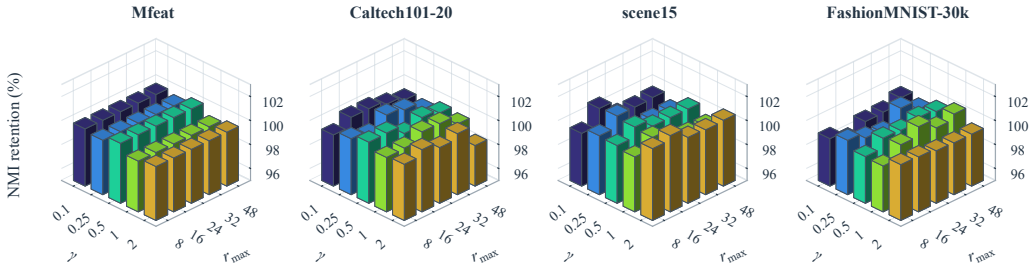

Figure 4: Joint hyperparameter sensitivity.

4.6 Clustering quality across acquisition rounds

Each acquisition round adds observations that inform subsequent selections. Figure 5 tracks three representative datasets over three rounds (all six in Appendix Figure A5). The horizontal axis is cumulative acquisitions divided by initially missing sample–view entries; bands show one standard error over five seeds. Mfeat improves most in the first round. Scene15 NMI rises from 34.12% to 41.85%, exceeding Random's final 38.55%; FashionMNIST-30k approaches a plateau (67.03% after round two, 67.00% after round three). EAA leads Random in mean NMI from the first round onward on all three datasets, showing an early and sustained advantage. Across ten datasets, final NMI exceeds Random in all 50 paired seeds; ARI and ACC improve in 41 and 34 pairs, respectively (Appendix Figure A6).

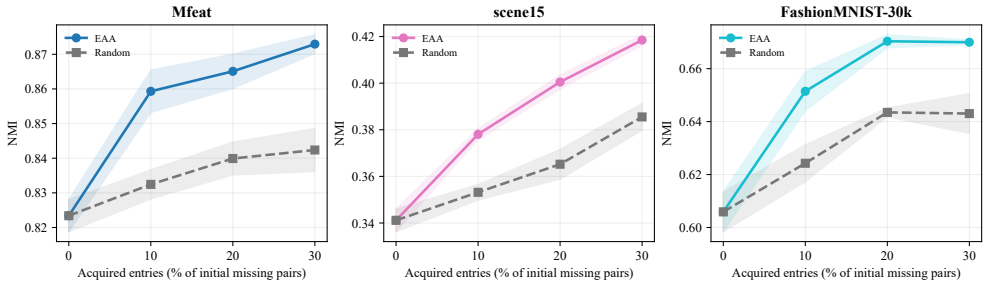

Figure 5: Clustering quality over acquisition rounds.

5 Discussion and Limitations

EAA avoids fine-grained scoring for dense views whose latent models have limited observed support. Its observed-support check routes these cases to the residual branch; sparse count views use a directional model. The check does not establish formal identifiability or guarantee agreement with external labels.

We assume known K, equal acquisition costs, noiseless retrieval, and correctly specified feature types. The objective does not enforce cluster-size balance. Future work includes cost-aware, size-regularized information gain and handling acquisition failures.

6 Conclusion

We presented EAA for exact-budget acquisition in IMVC. It uses typed CIG for sparse count views and dense views with sufficient observed support, and residual-based view selection for other dense views. Both branches refit over three rounds before consensus-graph clustering. Across ten benchmarks, including two 30,000-sample datasets, EAA achieves the highest mean NMI on nine, ARI on eight, and ACC on six. These results support incorporating observed-support assessment into unsupervised acquisition policy design.

Reproducibility statement

The supplementary materials provide the source code of EAA and the scripts for reproducing the experiments. The datasets and experimental settings are described in Section 4.1, with the mask-generation and budget-allocation procedures detailed in Appendix E.

AI use statement

We used generative AI tools to assist with code implementation, manuscript language editing and polishing, and table formatting. Generative AI was not used to generate research ideas or design the proposed method. All AI-assisted outputs were reviewed by the authors. We take full responsibility for the final content of this work, including all text, claims, and artifacts produced with the assistance of generative AI.

References

Hassan Ashtiani, Shrinu Kushagra, and Shai Ben-David. Clustering with same-cluster queries. In *Advances in Neural Information Processing Systems*, volume 29, pp. 3216–3224, 2016.

Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and Suvrit Sra. Clustering on the unit hypersphere using von Mises-Fisher distributions. *Journal of Machine Learning Research*, 6: 1345–1382, 2005.

Sugato Basu, Arindam Banerjee, and Raymond J. Mooney. Active semi-supervision for pairwise constrained clustering. In *Proceedings of the 2004 SIAM International Conference on Data Mining*, pp. 333–344, 2004.

Yiming Du, Ziyu Wang, Jian Li, Rui Ning, and Lusi Li. Deep incomplete multi-view clustering via hierarchical imputation and alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 20941–20949, 2026.

Ran Eisenberg, Jonathan Svirsky, and Ofir Lindenbaum. COPER: Correlation-based permutations for multi-view clustering. In *International Conference on Learning Representations*, pp. 46154–46179, 2025.

Brian Eriksson, Gautam Dasarathy, Aarti Singh, and Rob Nowak. Active clustering: Robust and efficient hierarchical clustering using adaptively selected similarities. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 260–268, 2011.

Chuanxing Geng, Aiyang Han, and Songcan Chen. Explicit view-labels matter: A multifacet complementarity study of multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(4):2520–2532, 2025.

Menglei Hu and Songcan Chen. Doubly aligned incomplete multi-view clustering. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 2262–2268, 2018.

Bingbing Jiang, Chenglong Zhang, Xinyan Liang, Peng Zhou, Jie Yang, Xingyu Wu, Junyi Guan, Weiping Ding, and Weiguo Sheng. Collaborative similarity fusion and consistency recovery for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 17617–17625, 2025.

Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, and En Zhu. Deep incomplete multi-view clustering with cross-view partial sample and prototype alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11600–11609, 2023.

Jiaqi Jin, Siwei Wang, Zhibin Dong, Xihong Yang, Xinwang Liu, En Zhu, and Kunlun He. Deep incomplete multi-view clustering with distribution dual-consistency recovery guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1016–1026, 2025.

Yang Li and Junier Oliva. Active feature acquisition with generative surrogate models. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 6450–6459, 2021.

Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. COMPLETER: Incomplete multi-view clustering via contrastive prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11174–11183, 2021.

Qian Liu, Huibing Wang, Jinjia Peng, Yawei Chen, Mingze Yao, Xianping Fu, and Yang Wang. Consensus-guided incomplete multi-view clustering via cross-view affinities learning. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 5761–5769, 2025.

Xinwang Liu, Xinzhong Zhu, Miaomiao Li, Chang Tang, En Zhu, Jianping Yin, and Wen Gao. Efficient and effective incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 4392–4399, 2019.

Xinwang Liu, Ke Liang, Jun Wang, Suyuan Liu, Xiangke Wang, and Huaimin Wang. Two decades of multi-view clustering: Taxonomy, application, and challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(3):3744–3764, 2026.

Chao Ma, Sebastian Tschiatschek, Konstantina Palla, José Miguel Hernández-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient dynamic discovery of high-value information with partial VAE. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4234–4243, 2019.

Andrew Y. Ng, Michael I. Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, volume 14, pp. 849–856, 2001.

Tillmann Rheude, Roland Eils, and Benjamin Wild. Cohort-based active modality acquisition. *arXiv preprint arXiv:2505.16791*, 2025.

Maytal Saar-Tsechansky, Prem Melville, and Foster Provost. Active feature-value acquisition. *Management Science*, 55(4):664–684, 2009.

Huayi Tang and Yong Liu. Deep safe incomplete multi-view clustering: Theorem and algorithm. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 21090–21110, 2022.

Siwei Wang, Xinwang Liu, Li Liu, Wenxuan Tu, Xinzhong Zhu, Jiyuan Liu, Sihang Zhou, and En Zhu. Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9776–9785, 2022.

Xinxin Wang, Yongshan Zhang, and Yicong Zhou. Learn multi-task anchor: Joint view imputation and label generation for incomplete multi-view clustering. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 6451–6459, 2025.

Youqing Wang, Yu Cao, Jinlu Wang, Xiang Xu, Jiapu Wang, Tengfei Liu, Junbin Gao, and Jipeng Guo. Aligning collaborative view recovery and tensorial subspace learning via latent representation for incomplete multi-view clustering. In *International Conference on Learning Representations*, pp. 79726–79740, 2026.

Jie Wen, Zheng Zhang, Lunke Fei, Bob Zhang, Yong Xu, Zhao Zhang, and Jinxing Li. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2):1136–1149, 2023.

Tingting Wu, Songhe Feng, and Jiazheng Yuan. Low-rank kernel tensor learning for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 15952–15960, 2024.

Deng Xu, Chao Zhang, Cong Guo, Chunlin Chen, and Huaxiong Li. Fast incomplete multi-view clustering with adaptive similarity completion and reconstruction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21734–21742, 2025.

Gehui Xu, Jie Wen, Chengliang Liu, Bing Hu, Yicheng Liu, Lunke Fei, and Wei Wang. Deep variational incomplete multi-view clustering: Exploring shared clustering structures. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 16147–16155, 2024.

Shengju Yu, Zhibin Dong, Siwei Wang, Xinhang Wan, Yue Liu, Weixuan Liang, Pei Zhang, Wenxuan Tu, and Xinwang Liu. Towards resource-friendly, extensible and stable incomplete multi-view clustering. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 57415–57440, 2024.

Shengju Yu, Yiu-Ming Cheung, Siwei Wang, Xinwang Liu, and En Zhu. Bifurcate then alienate: Incomplete multi-view clustering via coupled distribution learning with linear overhead. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 72853–72894, 2025.

Helin Zhao, Wei Chen, and Peng Zhou. Active deep multi-view clustering. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 5554–5562, 2024.

Shuping Zhao, Jie Wen, Lunke Fei, and Bob Zhang. Tensorized incomplete multi-view clustering with intrinsic graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 11327–11335, 2023.

A INNER-MODEL CONVERGENCE

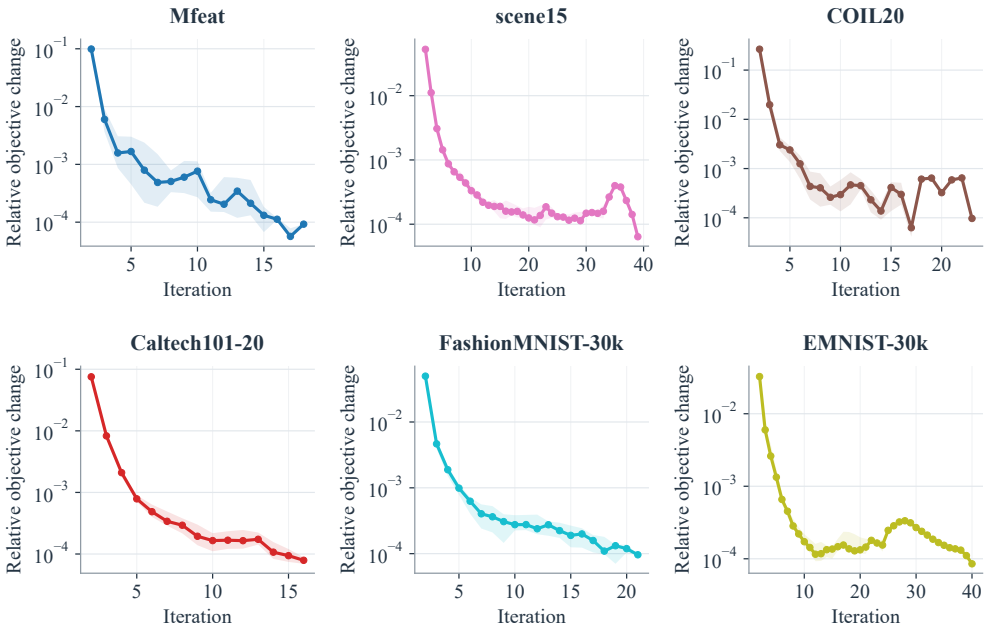

Figure A1: Inner-model convergence.

Figure A1 separates numerical convergence of each refit from the downstream acquisition trajectory in Figure 5. On six datasets over five seeds, the ordinate is the per-iteration relative objective change on a logarithmic scale; the stopping tolerance is 10^{-4}. Curves show the median across rounds and seeds and end at the longest executed trajectory; shading shows the interquartile range where multiple runs remain active. Relative changes decrease by several orders of magnitude and reach the tolerance, despite small non-monotone fluctuations caused by alternating latent assignments.

B OBSERVED-ONLY BRANCH ASSIGNMENT

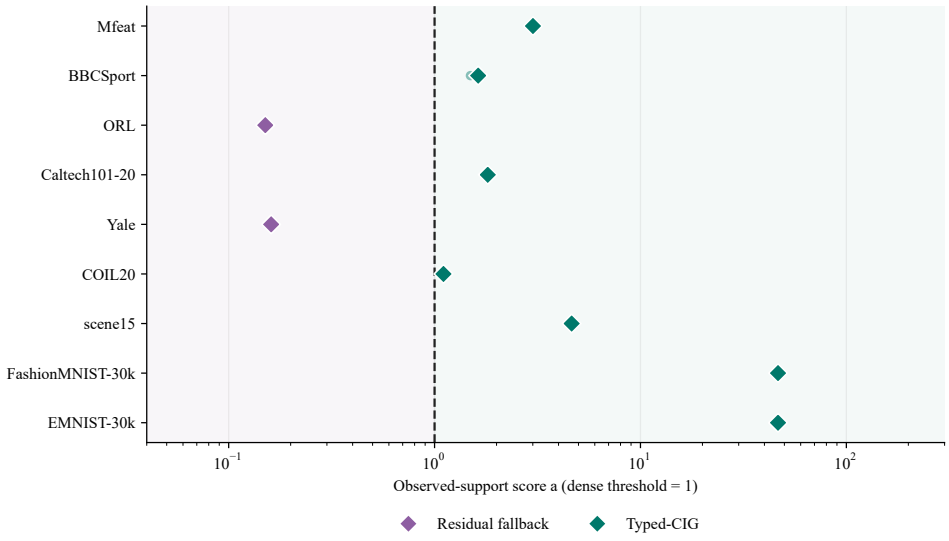

Figure A2: Observed-support values and acquisition branches.

Table A1: Ten-dataset acquisition-ranking ablations (NMI, mean ± sample standard deviation over five seeds).

Dataset	Full EAA	w/o CIG-based entry ranking
BBCSport	0.9037 ± 0.0071	0.8669 ± 0.0161
Reuters-1200	0.2722 ± 0.0164	0.2330 ± 0.0263
COIL20	0.8343 ± 0.0085	0.8160 ± 0.0078
Mfeat	0.8729 ± 0.0061	0.8480 ± 0.0122
Caltech101-20	0.5912 ± 0.0097	0.5660 ± 0.0081
scene15	0.4185 ± 0.0051	0.3883 ± 0.0057
FashionMNIST-30k	0.6700 ± 0.0019	0.6486 ± 0.0074
EMNIST-30k	0.6453 ± 0.0116	0.6063 ± 0.0158
Dataset	**Full EAA**	**w/o residual-based view ranking**
Yale	0.6398 ± 0.0309	0.5915 ± 0.0310
ORL	0.8379 ± 0.0114	0.7851 ± 0.0219

Figure A2 reports the observed-support score from Equation 1 over five mask seeds. The dense ORL and Yale datasets fall below the preset threshold of one and use the residual branch; the other dense datasets exceed it, while sparse BBCSport uses directional typed CIG by feature type. The plot documents the observed-only branch decisions; threshold separation follows from the rule itself and is not independent validation of the check's benefit. That benefit is evaluated by the check-removal comparison in Figure 2(a).

C Round-wise view allocation

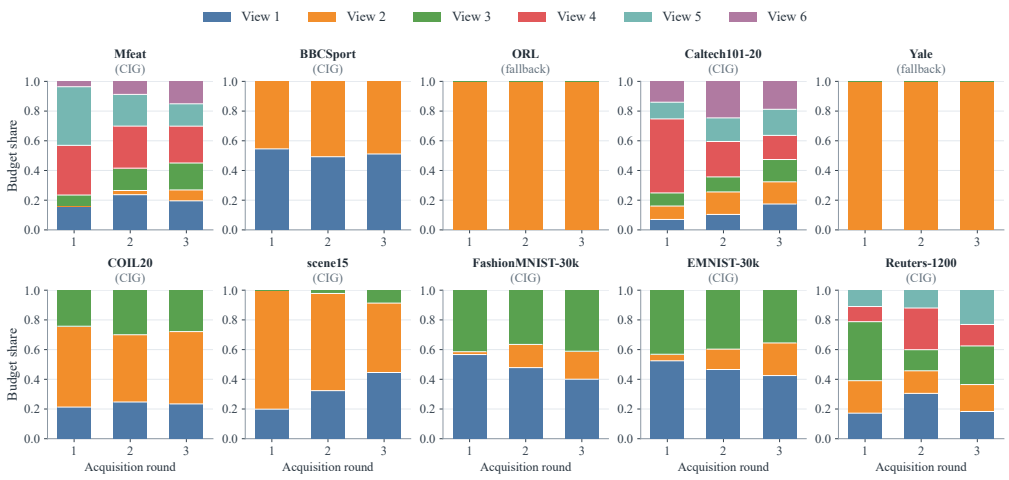

Figure A3: View-budget shares across acquisition rounds.

Figure A3 shows five-seed mean shares of each round's exact budget across all ten datasets. Typed CIG can change view allocation after model refitting; the residual branch concentrates view-level commitment on one view for ORL and Yale. Within a committed view, eligible samples are drawn uniformly. View indices identify dataset-specific views.

D Ten-dataset acquisition-component comparisons

Table A1 extends the two ranking-component removals to all ten datasets. Removing CIG-based entry ranking substitutes random entry ranks on the eight typed-CIG datasets; removing residual-based view ranking substitutes random view order on the two residual-branch datasets. Each branch

otherwise remains unchanged. Full EAA has higher NMI in 40 of 40 typed-CIG seed pairs and 9 of 10 residual-branch seed pairs. Reuters-1200 uses typed CIG for sparse text regardless of its support score; its row therefore tests entry ranking, not residual-based view ranking. The five-seed mean improves for every dataset, although individual seeds and the ACC and ARI metrics need not follow the same pattern. The likelihood ablation has a smaller effect on Reuters-1200: the directional model exceeds the Gaussian in three of five paired seeds. These comparisons test acquisition components under fixed budgets and round counts.

E EXACT MASK AND BUDGET PROTOCOL

For seed s, a deterministic permutation of all NV entries is traversed. An observed entry is removed only if the corresponding sample retains another observed view. Traversal stops exactly at $\lfloor 0.5NV \rceil$ removals. The budget is $B = \lfloor 0.3|\Omega(\mathbf{M}_0)| \rceil$. Round quotas sum exactly to B, and every arm is checked by

$$\|\mathbf{M}_T\|_0 = \|\mathbf{M}_0\|_0 + B, \qquad \mathcal{S} \subseteq \Omega(\mathbf{M}_0), \qquad |\mathcal{S}| = B. \tag{8}$$

These assertions prevent silent budget differences between acquisition policies.

F ADDITIONAL IMPLEMENTATION DETAILS

Dense views are cast to single precision after observed-only standardization. Principal components are bounded by the feature dimension, observed sample count, and rank cap. Sparse views remain in compressed sparse row format and use brute-force cosine nearest neighbors. For the directional likelihood, observed rows are ℓ_2-normalized, projected by an observed-only truncated SVD, and normalized again in the working space. Consensus spectral embeddings are computed from sparse normalized affinities. On the eight standard datasets, candidate partitions comprise 32 seeded single-initialization K-means runs, QR-based discretization, and iterative spectral discretization; the partition with the highest normalized association is retained. On the two large datasets, K-means uses three initializations. The same decoder settings apply to every policy within a dataset. On large datasets, empirical CIG outcomes are coarsened once per fitted model, not independently for each candidate.

Dense typed-CIG models use assignment temperature $\gamma = 0.5$ for initialization. For declared sparse count views, the optimized implementation calibrates the initialization temperature using observed graph structure: QR-based spectral assignments define a target equal to the weighted fraction of consensus edges crossing provisional clusters, and bisection matches the mean normalized assignment entropy to this target. This calculation uses no ground-truth labels. All variants in the sparse-likelihood comparison retain this initialization and differ only in the specified likelihood or final graph metric.

For empirical CIG, the fitted likelihood is normalized separately for each cluster over the observed outcomes of a view: $\widetilde{p}_v(j \mid k) = p_v(\mathbf{x}_j^{(v)} \mid k) / \sum_{\ell \in \mathcal{O}_v} p_v(\mathbf{x}_\ell^{(v)} \mid k)$. On large datasets, fixed-seed grouping of likelihood signatures yields groups $\mathcal{G}_{vg}$ with mass $p_v(g \mid k) = \sum_{j \in \mathcal{G}_{vg}} \widetilde{p}_v(j \mid k)$. The candidate-specific joint distribution is $q_i(k)p_v(g \mid k)$; its marginals and conditional posteriors give the entropy reduction in Equation 6 for the grouped outcome. Grouping retains all conditional mass but coarsens the outcome resolution.

The residual branch recomputes shared prototypes and residual scores after each acquisition round. Views are traversed in descending residual order. Eligible samples in each view are shuffled with a fixed seed and selected without replacement until the exact round quota is reached, with at most one new view per sample per round.

G PROTOCOL-ADAPTED CRITERIA AND MECHANISM-INSPIRED PROXIES

The comparison isolates entry acquisition from downstream clustering. Native acquisition heuristics are used directly; protocol-adapted criteria derive entry scores from method-specific outputs; and mechanism-inspired proxies translate published IMVC principles into label-free entry scores when no entry-selection rule is defined. Table A2 summarizes the resulting policies.

Table A2: Construction of native, protocol-adapted, and proxy acquisition policies. Every score is computed from the current observed state without class labels.

Method	Status	Retained output or principle	Entry-score construction	Selection
Random	Native acquisition	–	Uniform random score over feasible missing sample-view pairs.	Exact quota; at most one view per sample.
View-rate	Native acquisition	Current observation mask	Rank candidates by the current number of missing entries in their view, with seeded tie-breaking.	Exact quota; at most one view per sample.
Uncertainty	Native acquisition	Consensus affinity and cluster posterior	Marginal graph-coverage gain weighted by posterior uncertainty, $1 - \max_k q_{ik}$.	Greedy coverage; exact quota.
CAMA†	Adapted from active modality acquisition	Predicted change in output probabilities	KL divergence between the current cluster posterior and the posterior predicted after inserting the candidate view's estimated graph contribution.	Top score; exact quota.
CAL†	Adapted using released code	View self-representation L_v and consensus L_c	Normalized L_v–L_c residual on observed samples, transferred to missing samples by neighbors in the consensus embedding.	Top score; exact quota.
ASCR†	Adapted using released code	Completed similarity structure and embedding	Normalized view-structure residual from the ASCR output, transferred to missing samples by neighbors in the learned embedding.	Top score; exact quota.
JVILG†	Adapted using released code	Joint imputation and cluster assignment	Target cluster-view dispersion multiplied by one plus the sample's prototype residual over its observed views.	Top score; exact quota.
DIMVC-HIA‡	Proxy inspired by published mechanism	Hierarchical assignment and imputation	Jensen–Shannon disagreement between assignments predicted from available views and from observed neighbors of the target view, weighted by one plus the normalized target-view entropy.	Top score; exact quota.
ARSL-IMVC‡	Proxy inspired by published mechanism	Collaborative recovery from a shared representation	Normalized error of predicting an observed view representation from the consensus embedding, transferred by neighbors to missing candidates.	Top score; exact quota.

† Protocol-adapted acquisition criterion. ‡ Mechanism-inspired proxy policy. The released CAL, ASCR, and JVILG optimizers are used inside their adapters. Official implementations of DIMVC-HIA and ARSL-IMVC exist, but their proxy policies do not call the original optimizers.

CAMA†. *Original:* CAMA allocates additional modalities across a test cohort for supervised prediction and studies several acquisition criteria, including expected KL divergence between predictions before and after acquisition. *Retained:* We retain the prediction-change principle, not the complete supervised CAMA procedure. *Modification:* For each missing sample-view pair, we estimate the view-induced update of the observed-only consensus graph and rank the pair by $D_{\mathrm{KL}}(q_{\mathrm{after}} \| q_{\mathrm{before}})$ between cluster posteriors. This substitutes an observed-only clustering model and graph-based update for supervised predictions and generative modality imputation.

CAL†. *Original:* CAL learns view-specific self-representations L_v and a consensus representation L_c under a fixed observation mask. *Retained:* We run the released optimizer and retain these representation matrices. *Modification:* We compute the normalized L_v–L_c residual on observed samples and transfer it to missing candidates using nearest neighbors in the CAL consensus embedding.

ASCR†. *Original:* ASCR completes view-specific similarity structure and learns a shared embedding under a fixed mask. *Retained:* We run the released optimizer and retain its embedding and view-specific structural residuals. *Modification:* Observed residuals are normalized within each view and transferred to missing candidates by inverse-distance neighbors in the learned embedding.

JVILG†. *Original:* JVILG jointly models incomplete views and cluster assignments. *Retained:* We run the released optimizer and retain its current assignments together with view-specific prototype dispersion. *Modification:* A missing pair is scored by the dispersion of its assigned cluster in the target view multiplied by one plus the sample's mean prototype residual over its observed views.

DIMVC-HIA‡. *Original:* DIMVC-HIA uses hierarchical assignment and feature imputation under a fixed mask. *Retained:* We retain the principle that assignment disagreement indicates failure of hierarchical imputation. *Modification:* The proxy compares the cluster distribution predicted from a sample's observed views with the distribution interpolated from observed neighbors of the target view, using Jensen–Shannon divergence weighted by one plus the normalized target-view entropy.

ARSL-IMVC‡. *Original:* ARSL-IMVC recovers view-specific structure from a shared representation. *Retained:* We retain collaborative recovery error as the structural signal. *Modification:* The proxy regresses an observed view representation on the consensus embedding, normalizes the observed recovery residual, and transfers that residual to missing candidates by nearest neighbors.

H COMPLETE CLUSTERING PERFORMANCE

Tables A3 and A4 provide the five-seed means and sample standard deviations for every method, dataset, and metric summarized in Table 2. Datasets are split across the two tables in increasing order of sample count, following Tables 1 and 2. Abbreviations and baseline markers have the same meanings as in Table 2.

Table A3: Complete clustering performance: first five datasets in sample-count order (mean ± sample standard deviation over five seeds).

Metric	Method	Yale	ORL	BBCSport	Reuters	COIL20
ACC	EAA	**0.6255 ± 0.0372**	**0.7260 ± 0.0240**	**0.9691 ± 0.0030**	**0.3932 ± 0.0218**	0.7435 ± 0.0278
	Random	0.5588 ± 0.0391	0.6355 ± 0.0212	0.9640 ± 0.0044	0.3700 ± 0.0180	0.7339 ± 0.0089
	View-rate	0.5297 ± 0.0195	0.6290 ± 0.0282	0.9474 ± 0.0215	0.3438 ± 0.0698	0.7371 ± 0.0094
	Uncertainty	0.5612 ± 0.0396	0.6400 ± 0.0188	0.9574 ± 0.0015	0.2770 ± 0.0217	**0.7464 ± 0.0106**
	CAMA†	0.5515 ± 0.0148	0.6580 ± 0.0369	0.8860 ± 0.0098	0.3898 ± 0.0479	0.6879 ± 0.0213
	CAL†	0.5261 ± 0.0607	0.7105 ± 0.0237	0.7982 ± 0.1154	0.2747 ± 0.0130	0.6889 ± 0.0304
	ASCR†	0.4945 ± 0.0175	0.5950 ± 0.0179	0.9452 ± 0.0102	0.2625 ± 0.0093	0.7360 ± 0.0205
	JVILG†	0.5248 ± 0.0185	0.6225 ± 0.0272	0.9353 ± 0.0372	0.3880 ± 0.0612	0.7264 ± 0.0312
	DIMVC-HIA‡	0.5564 ± 0.0248	0.6400 ± 0.0244	0.9537 ± 0.0122	0.2730 ± 0.0068	0.7174 ± 0.0219
	ARSL-IMVC‡	0.5164 ± 0.0658	0.6240 ± 0.0311	0.6618 ± 0.0882	0.2717 ± 0.0127	0.7186 ± 0.0259
NMI	EAA	**0.6398 ± 0.0309**	**0.8379 ± 0.0114**	**0.9037 ± 0.0071**	**0.2722 ± 0.0164**	0.8343 ± 0.0085
	Random	0.5746 ± 0.0328	0.7731 ± 0.0101	0.8835 ± 0.0124	0.2485 ± 0.0160	0.8129 ± 0.0084
	View-rate	0.5619 ± 0.0157	0.7687 ± 0.0113	0.8627 ± 0.0163	0.2318 ± 0.0334	0.8154 ± 0.0039
	Uncertainty	0.5785 ± 0.0231	0.7698 ± 0.0136	0.8677 ± 0.0077	0.1951 ± 0.0096	0.8238 ± 0.0044
	CAMA†	0.5763 ± 0.0119	0.7910 ± 0.0164	0.7861 ± 0.0199	0.2595 ± 0.0346	0.8112 ± 0.0053
	CAL†	0.5532 ± 0.0548	0.8209 ± 0.0178	0.7366 ± 0.0635	0.1822 ± 0.0100	0.7942 ± 0.0062
	ASCR†	0.5153 ± 0.0238	0.7556 ± 0.0119	0.8454 ± 0.0164	0.1624 ± 0.0103	0.8275 ± 0.0047
	JVILG†	0.5541 ± 0.0173	0.7685 ± 0.0117	0.8461 ± 0.0480	0.2538 ± 0.0392	0.8356 ± 0.0079
	DIMVC-HIA‡	0.5783 ± 0.0295	0.7717 ± 0.0140	0.8604 ± 0.0295	0.1868 ± 0.0146	0.8324 ± 0.0095
	ARSL-IMVC‡	0.5524 ± 0.0554	0.7669 ± 0.0152	0.5907 ± 0.0562	0.1860 ± 0.0128	**0.8359 ± 0.0119**
ARI	EAA	**0.4383 ± 0.0386**	**0.6076 ± 0.0255**	**0.9176 ± 0.0094**	**0.1598 ± 0.0136**	**0.6989 ± 0.0309**
	Random	0.3404 ± 0.0425	0.4849 ± 0.0155	0.9058 ± 0.0117	0.1383 ± 0.0100	0.6747 ± 0.0150
	View-rate	0.3190 ± 0.0180	0.4788 ± 0.0297	0.8829 ± 0.0177	0.1092 ± 0.0520	0.6811 ± 0.0086
	Uncertainty	0.3478 ± 0.0259	0.4893 ± 0.0265	0.8873 ± 0.0069	0.0653 ± 0.0199	0.6963 ± 0.0097
	CAMA†	0.3458 ± 0.0238	0.5290 ± 0.0374	0.8173 ± 0.0213	0.1577 ± 0.0383	0.6308 ± 0.0181
	CAL†	0.3146 ± 0.0660	0.5792 ± 0.0317	0.6712 ± 0.1433	0.0411 ± 0.0088	0.6180 ± 0.0294
	ASCR†	0.2716 ± 0.0177	0.4467 ± 0.0216	0.8577 ± 0.0313	0.0310 ± 0.0061	0.6880 ± 0.0134
	JVILG†	0.3108 ± 0.0219	0.4746 ± 0.0311	0.8743 ± 0.0444	0.1377 ± 0.0523	0.6813 ± 0.0296
	DIMVC-HIA‡	0.3498 ± 0.0334	0.4914 ± 0.0313	0.8808 ± 0.0291	0.0460 ± 0.0172	0.6740 ± 0.0225
	ARSL-IMVC‡	0.3026 ± 0.0770	0.4784 ± 0.0311	0.4563 ± 0.0926	0.0406 ± 0.0080	0.6639 ± 0.0274

Best means are **bold**; second-best means are underlined. † Protocol-adapted criterion; ‡ mechanism-inspired proxy. Abbreviations and unavailable results follow Table 2.

Table A4: Complete clustering performance: remaining five datasets in sample-count order (mean ± sample standard deviation over five seeds).

Metric	Method	Mfeat	Caltech20	scene15	FMNIST	EMNIST
ACC	EAA	0.8272 ± 0.0044	**0.4545 ± 0.0194**	0.4164 ± 0.0118	0.5945 ± 0.0068	**0.6781 ± 0.0165**
	Random	0.8366 ± 0.0245	0.4394 ± 0.0352	0.4088 ± 0.0056	0.6231 ± 0.0319	0.6595 ± 0.0503
	View-rate	0.8327 ± 0.0234	0.4353 ± 0.0078	0.4077 ± 0.0148	0.6156 ± 0.0028	0.6730 ± 0.0592
	Uncertainty	0.8254 ± 0.0201	0.4303 ± 0.0015	0.4154 ± 0.0090	**0.6407 ± 0.0439**	0.6731 ± 0.0555
	CAMA†	0.8325 ± 0.0166	0.4334 ± 0.0179	0.3644 ± 0.0152	0.5989 ± 0.0449	0.5860 ± 0.0361
	CAL†	0.8387 ± 0.0233	0.4189 ± 0.0219	0.4083 ± 0.0135	–	–
	ASCR†	0.8284 ± 0.0069	0.4347 ± 0.0176	**0.4173 ± 0.0187**	0.5830 ± 0.0096	0.6110 ± 0.0272
	JVILG†	0.8253 ± 0.0183	0.4340 ± 0.0136	0.4094 ± 0.0144	0.6000 ± 0.0291	0.6327 ± 0.0309
	DIMVC-HIA‡	0.8552 ± 0.0112	0.4483 ± 0.0457	0.3667 ± 0.0089	0.5763 ± 0.0369	0.5645 ± 0.0454
	ARSL-IMVC‡	**0.8655 ± 0.0030**	0.4267 ± 0.0180	0.4136 ± 0.0072	–	–
NMI	EAA	**0.8729 ± 0.0061**	**0.5912 ± 0.0097**	**0.4185 ± 0.0051**	**0.6700 ± 0.0019**	**0.6453 ± 0.0116**
	Random	0.8424 ± 0.0139	0.5670 ± 0.0056	0.3855 ± 0.0129	0.6430 ± 0.0168	0.6195 ± 0.0215
	View-rate	0.8546 ± 0.0120	0.5642 ± 0.0144	0.3858 ± 0.0051	0.6492 ± 0.0141	0.6235 ± 0.0224
	Uncertainty	0.8492 ± 0.0088	0.5590 ± 0.0052	0.3790 ± 0.0062	0.6298 ± 0.0137	0.6143 ± 0.0173
	CAMA†	0.8306 ± 0.0141	0.5606 ± 0.0091	0.3470 ± 0.0102	0.5889 ± 0.0219	0.5269 ± 0.0387
	CAL†	0.8695 ± 0.0092	0.5772 ± 0.0139	0.3902 ± 0.0273	–	–
	ASCR†	0.8620 ± 0.0041	0.5772 ± 0.0155	0.3941 ± 0.0096	0.6084 ± 0.0062	0.5454 ± 0.0103
	JVILG†	0.8441 ± 0.0080	0.5580 ± 0.0070	0.3755 ± 0.0090	0.6310 ± 0.0129	0.5934 ± 0.0206
	DIMVC-HIA‡	0.8467 ± 0.0127	0.5628 ± 0.0097	0.3387 ± 0.0131	0.6097 ± 0.0185	0.5188 ± 0.0185
	ARSL-IMVC‡	0.8544 ± 0.0078	0.5682 ± 0.0058	0.3972 ± 0.0126	–	–
ARI	EAA	0.8076 ± 0.0054	**0.3230 ± 0.0195**	**0.2383 ± 0.0083**	0.4790 ± 0.0064	**0.5313 ± 0.0200**
	Random	0.7859 ± 0.0171	0.3178 ± 0.0322	0.2342 ± 0.0096	0.4830 ± 0.0249	0.5085 ± 0.0433
	View-rate	0.7976 ± 0.0150	0.3047 ± 0.0106	0.2318 ± 0.0094	0.4845 ± 0.0054	0.5213 ± 0.0475
	Uncertainty	0.7905 ± 0.0143	0.2985 ± 0.0044	0.2340 ± 0.0027	**0.4894 ± 0.0304**	0.5159 ± 0.0386
	CAMA†	0.7717 ± 0.0104	0.3050 ± 0.0249	0.1937 ± 0.0134	0.4448 ± 0.0394	0.3994 ± 0.0450
	CAL†	**0.8104 ± 0.0157**	0.2973 ± 0.0217	0.2331 ± 0.0183	–	–
	ASCR†	0.8003 ± 0.0037	0.3051 ± 0.0165	0.2345 ± 0.0103	0.4303 ± 0.0093	0.4391 ± 0.0233
	JVILG†	0.7860 ± 0.0064	0.3034 ± 0.0099	0.2244 ± 0.0114	0.4645 ± 0.0198	0.4795 ± 0.0377
	DIMVC-HIA‡	0.7925 ± 0.0128	0.3227 ± 0.0454	0.1883 ± 0.0108	0.4348 ± 0.0293	0.3876 ± 0.0403
	ARSL-IMVC‡	0.8078 ± 0.0093	0.3034 ± 0.0080	0.2356 ± 0.0075	–	–

Best means are **bold**; second-best means are underlined. † Protocol-adapted criterion; ‡ mechanism-inspired proxy. Abbreviations and unavailable results follow Table 2.

I ADDITIONAL BUDGET AND MISSINGNESS RESULTS

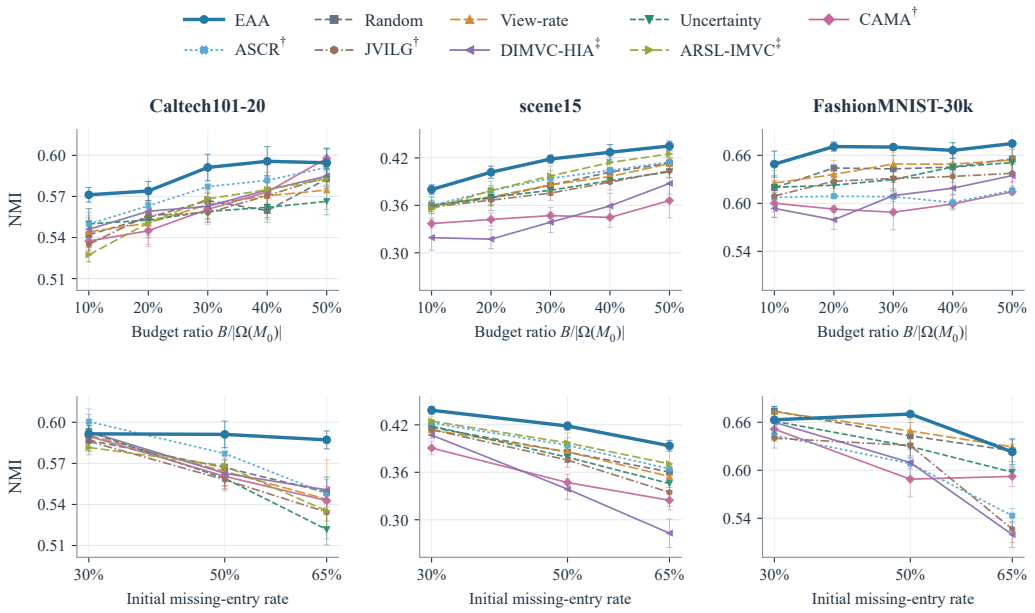

Figure A4: Additional budget (top) and missingness (bottom) results. Error bars: five-seed sample standard deviations. † Protocol-adapted criterion; ‡ mechanism-inspired proxy.

Figure A4 completes the six-dataset sweep with Caltech101-20, scene15, and FashionMNIST-30k, using the same settings as Figure 3. Together, the figures contain all seven distinct conditions per dataset: five budget ratios at 50% initial missingness and three missingness levels at a 30% budget ratio, with the shared default condition counted once. CAL remains in Table 2 but is omitted from this sweep because repeated $N \times N$ optimizations make a complete curve prohibitively expensive; ARSL-IMVC retains its scalability exclusion on FashionMNIST-30k.

Among the plotted policies, EAA has the highest mean NMI in 19 of the 21 main-text dataset–condition pairs and 17 of the 21 pairs here, totaling 36 of 42. Other policies lead at the 10% and 20% budget points on Reuters-1200. Reuters-1200 has a non-monotonic budget response: EAA NMI increases from 0.2615 at a 10% budget to 0.2722 at 30%, then decreases to 0.2625 at 50%. At a 30% budget ratio, increasing initial missingness from 30% to 65% reduces scene15 NMI from 0.4384 to 0.3935, consistent with the decline observed on ORL. Neither increasing the budget nor changing initial missingness yields a universal monotonic response across these datasets.

J ADDITIONAL ACQUISITION TRAJECTORIES AND PAIRED GAINS

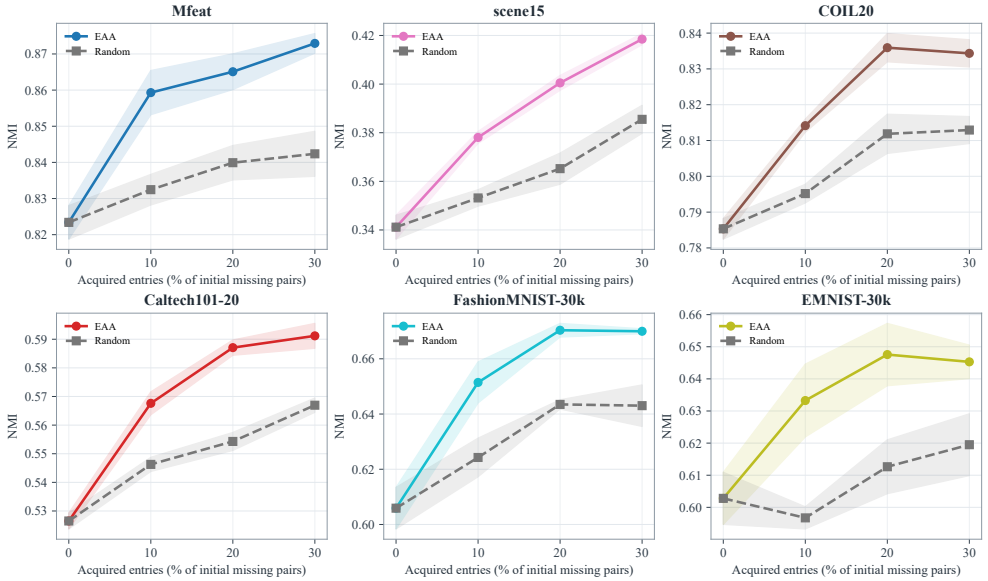

Figure A5: Complete six-dataset acquisition trajectories.

Figure A5 retains the complete six-dataset trajectories underlying Figure 5. Mfeat, scene15, and FashionMNIST-30k are shown in the main text to illustrate predominantly early gains, sustained improvement, and a late-round plateau, respectively, including a 30,000-sample setting. The horizontal axis is cumulative acquisitions divided by initially missing sample–view entries. Bands show one standard error over five seeds along the same three-round trajectory, not independent runs with different round counts.

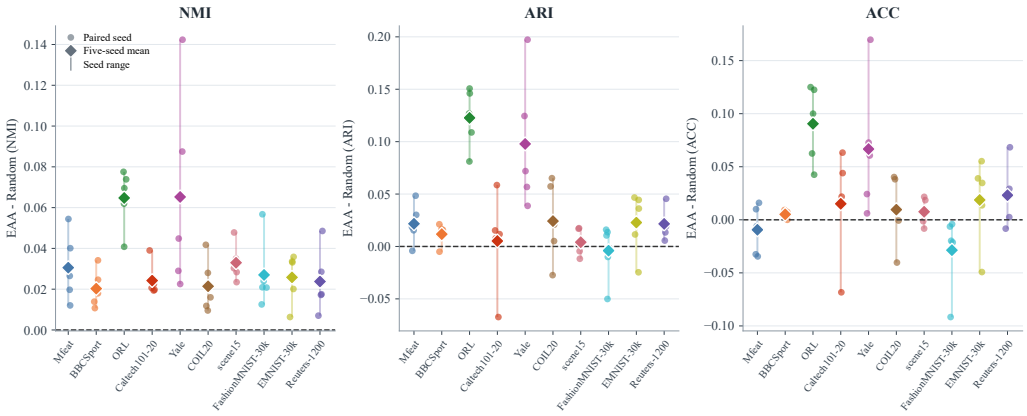

Figure A6: Five-seed paired gains over Random.

Figure A6 compares final-round EAA and Random under matched masks and seeds on ten datasets. Dots show paired differences; diamonds show dataset means. NMI improves in all 50 pairs (mean gain 3.36 percentage points), ARI in 41 of 50 (3.28 points), and ACC in 34 of 50 (1.98 points). The NMI advantage persists across all tested seeds, while the less consistent ACC gains show that improvements need not be uniform across metrics.