BATCH ACTIVE EVIDENCE ACQUISITION VIA SUBMODULAR UTILITY COVERAGE ACROSS HYPOTHESES

Anonymous authors
Paper under double-blind review

ABSTRACT

Decision systems in clinical diagnosis, fraud review, industrial monitoring, and tool use must choose which costly evidence to acquire before making a prediction. Sequential acquisition selects one action, observes its outcome, and then chooses the next. Test panels, imaging protocols, and related workflows instead require the complete action batch to be selected before any outcome is available. We study this one-shot batch active evidence acquisition setting, where an action can reveal a feature, an image patch, or the answer to a diagnostic question. Without the observation loop, the policy cannot re-estimate the value of the remaining actions after each acquisition, and directly maximizing the expected downstream-loss reduction of a batch is intractable. We introduce hypothesis utility coverage (HUC). Given a partial observation, HUC retrieves similar completed instances to represent plausible completions, transfers their precomputed explanation-derived action utilities, and applies diminishing returns to the utility already covered within each completion. The surrogate is monotone submodular by construction, so greedy selection carries the classical $(1 - 1/e)$ guarantee on the surrogate under a cardinality budget without assuming anything about the data, and a conditional bound separates its gap from the ideal objective into retrieval, saturation-fit, and utility-estimation terms. On a controlled synthetic benchmark, a matched control shows that the gain over ranking actions by posterior-averaged importance comes from this set-dependent saturation rather than from rescoring individual actions. On diagnostic question acquisition (DDXPlus) HUC improves over a fixed acquisition path at every plotted budget, on real image and tabular datasets the gains are small or absent, and under the reported implementations it is competitive with direct batch scoring at a fraction of the selection cost.

1 INTRODUCTION

Decision-making under incomplete information usually requires acquiring evidence, and each acquisition carries a cost. Examples include laboratory tests in clinical diagnosis, expert reviews in fraud investigation, sensor measurements in industrial monitoring, and tool calls in automated systems. The active feature acquisition (AFA) literature studies this tradeoff, with sequential methods acquiring one feature at a time and conditioning the next acquisition on its observed value (Aronsson et al., 2025; Dulac-Arnold et al., 2011; Covert et al., 2023; Guney et al., 2025). Batch acquisition instead selects the complete set of actions before observing any of their outcomes, as when a complete request must be issued at once in laboratory panels and imaging protocols. We study one-shot batch active evidence acquisition, where the policy selects the full set of actions from the initial observation and receives their outcomes after committing to the batch. This one-shot constraint requires the policy to evaluate the complete action set under unobserved outcomes and account for how the actions jointly influence the downstream decision. Direct optimization requires integration over joint outcomes and combinatorial search over action sets, making exact optimization intractable.

The natural one-shot policy ranks candidate actions by their importance averaged over plausible completions of the observation and acquires the top of the list. Once the first selected actions have covered what one completion needs, that completion should weigh less in the choice of the next action. Linear averaging keeps paying it full weight. The closest prior method for the batch setting, GenEx (Asgaonkar et al., 2024), obtains a guarantee under weak-submodularity conditions on the data. We take the opposite route and build a surrogate that is submodular by construction.

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.

The method approximates unknown action outcomes with a nonparametric retrieval posterior over completed instances, estimates the ideal utilities with precomputed explanation-derived importance (Lundberg & Lee, 2017; Guney et al., 2025), and applies a concave saturating function to the utility a batch covers within each completion. Greedy selection then carries the classical $(1 - 1/e)$ guarantee on the surrogate, and the gap between the surrogate and the ideal objective decomposes into three additive terms that measure retrieval error, saturation-fit error, and the discrepancy between explanation-derived and ideal utilities. The transfer to the ideal objective is conditional on those terms, which contain latent quantities.

Our contributions are:

- **A one-shot batch objective that is submodular by construction.** We formulate one-shot evidence selection through expected downstream-loss reduction and derive the HUC surrogate from retrieval, explanation-derived utilities, and concave saturation. Greedy selection attains the $(1 - 1/e)$ guarantee on the surrogate under a cardinality constraint, and a conditional bound decomposes the gap to the ideal objective across the three approximations.
- **Where the gain comes from, and where it appears.** On a controlled synthetic benchmark a matched control that shares every input with HUC and differs only in set dependence shows that the gain over ranking is set-dependent saturation. Across five datasets and three action types the gain over a fixed acquisition path is large on the synthetic benchmark and on DDXPlus and small or absent on real tabular and image data, and HUC is competitive with direct batch scoring at a fraction of its selection cost.

2 Related work

Sequential active feature acquisition. Sequential AFA acquires one feature, observes its value, and updates the next decision. Reinforcement-learning approaches model this process as an MDP and learn policies that trade predictive performance against acquisition cost (Shim et al., 2018; Janisch et al., 2019). Generative-surrogate approaches augment reinforcement-learning policies with auxiliary information and intermediate rewards (Li & Oliva, 2021). Other methods rank the next feature using explanation-derived importance (Guney et al., 2025). SEFA instead trains a supervised latent-variable model with stochastic encoders and selects features by evaluating possible unobserved realizations in the latent space (Norcliffe et al., 2025). The approximate acquisition-conditioned oracle (AACO) estimates conditional losses with nearest neighbors and searches over feature subsets to account for jointly informative acquisitions, but its published policy selects one feature from the chosen subset, observes its value, and repeats (Valancius et al., 2024). Each of these uses intermediate observations to guide subsequent acquisitions.

One-shot batch acquisition. GenEx directly studies batch feature acquisition, where the policy selects the complete feature set before receiving any new values (Asgaonkar et al., 2024). It partitions heterogeneous partial observations using random-hyperplane hashing, trains a generator-assisted mixture of experts, and optimizes a cardinality-constrained surrogate with a greedy algorithm. Its analysis derives an approximation bound under partial-monotonicity and weak-submodularity conditions. Although published AACO is sequential, its subset-scoring objective also yields a one-shot adaptation that returns the selected subset as the batch instead of using it to choose one next feature (Valancius et al., 2024). We use this adaptation as a batch baseline. GenEx is the closest prior method, and we do not compare against it. Its default inference can request further features after inspecting the confidence of a first batch. Its published $V = \varnothing$ ablation queries the whole selected set at once and is the natural one-shot comparator, but running it requires its generator-assisted mixture-of-experts pipeline, which we leave to future work. Our method instead combines retrieval, concave saturation, and explanation-derived utilities, yielding a monotone submodular objective with a $(1 - 1/e)$ greedy guarantee under a cardinality constraint and a conditional fidelity bound.

Submodular selection. For monotone submodular maximization under a cardinality constraint, greedy selection attains a $(1 - 1/e)$ approximation (Nemhauser et al., 1978). Under a knapsack constraint, the same guarantee requires partial enumeration (Sviridenko, 2004). Plain cost-benefit greedy does not attain this guarantee. Concave functions of modular coverage are the standard construction for submodularity by design (Lin & Bilmes, 2011; Krause & Golovin, 2014). Our sur-

rogate is this construction with retrieved completions as coverage targets and transferred attributions as the covered mass. We do not assume that the ideal predictive-loss objective is submodular. We quantify the discrepancy between the surrogate and that ideal objective through separate retrieval, saturation-fit, and utility-estimation terms.

Explanation-derived acquisition utilities. Feature-attribution methods estimate the contribution of each feature to a prediction (Lundberg & Lee, 2017; Jethani et al., 2022), instance-wise selectors learn related predictive rankings (Yoon et al., 2019; Chen et al., 2018), and SAGE estimates global feature importance (Covert et al., 2020). Prior work demonstrates the utility of explanation-derived rankings for sequential acquisition (Guney et al., 2025). We use precomputed rankings to estimate the latent utility of each action within a set objective, and retain the discrepancy between explanation-derived and ideal utilities as an explicit term in the fidelity bound.

Cost-aware evidence and tool selection. Budgeted sequential classifiers and cost-sensitive cascades learn staged prediction policies that balance predictive performance against feature cost (Trapeznikov & Saligrama, 2013; Xu et al., 2012). Cost-aware tool-selection methods trade task quality against execution cost (Wu et al., 2025), while CostBench studies multi-turn planning with intermediate feedback (Liu et al., 2026). Our experiments cover tabular features, image patches, and diagnostic questions selected in one shot. Tool calls fit the same abstraction and are not evaluated here.

3 Problem Formulation

Let $[d] = \{1, \dots, d\}$ index the features, and for an instance x let $M \subseteq [d]$ be the set of features observed so far, x_M the corresponding partial observation, and $A(M) = [d] \setminus M$ the candidate set of unobserved features. Each feature j carries an acquisition cost $c_j > 0$, a batch $S \subseteq A(M)$ has cost $c(S) = \sum_{j \in S} c_j$, and S is feasible under a budget B if $c(S) \leq B$. Prediction is performed by a fixed predictor f, the label is denoted y, and a bounded loss by ℓ. We study the *one-shot* batch setting: starting from M, a policy q selects a single feasible batch $S = q(x_M)$, all of S is acquired at once, and prediction is then made from $x_{M \cup S}$; since no value is seen during selection, the batch is chosen from x_M alone. Its quality depends on the joint value of its members rather than on their individual contributions. We write $\mathcal{F}_B = \{S \subseteq A(M) : c(S) \leq B\}$ for the feasible family. We measure this value by the predictive improvement the batch yields, defining the ideal batch value

$$G(S \mid x_M) \;=\; \mathbb{E}_{(x,y)\sim p(\cdot \mid x_M)}\big[\, \ell(f(x_M), y) - \ell(f(x_{M\cup S}), y) \,\big], \tag{1}$$

where $x_{M \cup S}$ is x_M together with the acquired values x_S. The best feasible batch is

$$S^\star \;\in \underset{S \subseteq A(M)\,:\, c(S) \leq B}{\arg\max} \; G(S \mid x_M). \tag{2}$$

Since $S^\star$ is defined through G, which is intractable, it is a target we aim for rather than a batch we can compute directly. We therefore seek a policy q whose feasible batch $q(x_M)$ attains value close to the optimum $G(S^\star \mid x_M)$. The function G in equation 1 is defined by the loss, the predictor, the budget, and the conditional distribution of the unobserved features. It is the quantity we want, yet it cannot be optimized directly, for three reasons. It is an expectation over the unknown conditional $p(\cdot \mid x_M)$; it is a set function over subsets of $A(M)$, so its maximization is combinatorial; and it is in general neither concave nor submodular, since features can interact. The remainder of the paper builds a tractable, monotone submodular surrogate for G, from which the policy q produces its batch, and bounds the gap between $G(q(x_M) \mid x_M)$ and $G(S^\star \mid x_M)$.

4 Methodology

The three obstacles named at the end of Section 3 are addressed by three approximations, each introducing a lemma that bounds its error, and Section 4.4 combines them with greedy selection. The end-to-end guarantee depends only on these three bounds. Detailed proofs of all results are in Appendix A.7.

4.1 Step 1: A retrieval-based posterior

We approximate the expectation in equation 1 nonparametrically, without a generative model. Let g be a fixed encoder mapping a partial observation to an embedding, and let $\mathcal{D} = \{(x^k, y^k)\}_k$ be a

reference set. For a query x_M we retrieve the K nearest neighbors $\mathcal{N}(x_M) = \mathrm{NN}_K(g(x_M))$ with weights

$$\alpha_k \;=\; \frac{\kappa\big(g(x_M), g(x_M^k)\big)}{\sum_{k'} \kappa\big(g(x_M), g(x_M^{k'})\big)}, \qquad \kappa(u,v) = \exp\big(-\|u-v\|_2^2/\vartheta\big), \tag{3}$$

requiring only $\alpha_k \geq 0$ and $\sum_k \alpha_k = 1$. The weighted neighbors form an empirical posterior $p(\cdot \mid x_M) \approx \sum_k \alpha_k \, \delta_{(x^k, y^k)}$, each neighbor a hypothesis about how the query completes. With $x_M \oplus x_S^k$ the input agreeing with the query on M and imputing block S from neighbor k, the induced batch value is

$$\widehat{G}(S \mid x_M) = \sum_k \alpha_k \, \gamma_k(S), \qquad \gamma_k(S) = \ell\big(f(x_M), y^k\big) - \ell\big(f(x_M \oplus x_S^k), y^k\big). \tag{4}$$

The gain $\gamma_k(S)$ is hybrid, combining the query context x_M with the neighbor's block x_S^k and label y^k. Selecting a batch that scores well across these hypotheses hedges over the retrieval posterior.

Lemma 1 (Posterior fidelity, bounding ε_1)**.** *Let (X, Y) denote a random complete instance and define the true-completion gain*

$$\gamma_S^{\mathrm{true}}(X,Y) = \ell(f(X_M), Y) - \ell(f(X_{M\cup S}), Y), \qquad m_S(v) = \mathbb{E}[\gamma_S^{\mathrm{true}}(X,Y) \mid g(X_M) = v].$$

For a fixed query x_M, let $v_0 = g(x_M)$, let $v_k = g(x_M^k)$ for the K retrieved neighbors, and let $\mathcal{V} = (v_1, \ldots, v_K)$. Write $\mu_{k,S} = \mathbb{E}[\gamma_k(S) \mid \mathcal{V}]$. Let $\mathcal{F}_B$ be the feasible family, $N_B = |\mathcal{F}_B|$, and fix $\eta \in (0,1)$. Assume that, uniformly over $S \in \mathcal{F}_B$, there exist $\delta_{\mathrm{repr}}, \delta_{\mathrm{imp}}, L_m, h, \Gamma \geq 0$ such that: (i) $|G(S \mid x_M) - m_S(v_0)| \leq \delta_{\mathrm{repr}}$, (ii) m_S is L_m-Lipschitz, (iii) $\max_k \|v_k - v_0\|_2 \leq h$, and (iv) conditional on $\mathcal{V}$, the hybrid gains $\{\gamma_k(S)\}_k$ are independent, each has conditional range of width at most Γ, and $|\mu_{k,S} - m_S(v_k)| \leq \delta_{\mathrm{imp}}$. Then, with probability at least $1-\eta$, simultaneously for all $S \in \mathcal{F}_B$,

$$\big|\widehat{G}(S \mid x_M) - G(S \mid x_M)\big| \leq \underbrace{\delta_{\mathrm{repr}}}_{\textit{representation}} + \underbrace{L_m h}_{\textit{retrieval bias}} + \underbrace{\delta_{\mathrm{imp}}}_{\textit{hybrid imputation}} + \underbrace{\Gamma\sqrt{\Big(\textstyle\sum_k \alpha_k^2\Big)\tfrac{1}{2}\ln\tfrac{2N_B}{\eta}}}_{\textit{finite-}K\textit{ variance}} =: \varepsilon_1.$$

For uniform weights $\alpha_k = 1/K$, the variance term is $\Gamma\sqrt{\frac{1}{2K}\ln\frac{2N_B}{\eta}}$. Under a cardinality budget $|S| \leq b$ with $b \geq 1$, $N_B = \sum_{r=0}^{b}\binom{|A(M)|}{r} \leq (b+1)d^b$, hence $\ln N_B = O(b \ln d)$.

The representation term records batch-gain information lost by the encoder. The hybrid-imputation term records the discrepancy introduced by grafting a neighbor's unobserved block onto the query context. The retrieval-radius term and finite-K term retain the standard bias-variance trade-off. The observed radius, weights, and feasible-family size are computable, while L_m, δ_{repr}, and δ_{imp} require assumptions or empirical upper bounds. We do not verify the conditional independence in (iv). The retrieved neighbors share the query context and the predictor, which works against it.

4.2 Step 2: A submodular saturation surrogate

The estimate $\widehat{G}$ still lacks submodular structure because each γ_k encodes unrestricted interactions. For each hypothesis k, let $u_k^\star(j) \geq 0$ be a normalized ideal per-feature importance with $\sum_{j \in A(M)} u_k^\star(j) = 1$, and let $z_k^\star(S) = \sum_{j \in S} u_k^\star(j)$ be the mass a batch covers. Define

$$F^\star(S \mid x_M) = \sum_k \alpha_k \, \psi\big(z_k^\star(S)\big) \tag{5}$$

for a nondecreasing concave $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $\psi(0) = 0$.

Lemma 2 (Saturation fidelity)**.** *Assume the ideal importances $\{u_k^\star\}$ and ψ satisfy $\big|\gamma_k(S) - \psi(z_k^\star(S))\big| \leq \eta_\psi$ for all k and feasible S. Then*

$$\big|\widehat{G}(S \mid x_M) - F^\star(S \mid x_M)\big| \;\leq\; \eta_\psi \;=:\; \varepsilon_2.$$

Concavity encodes diminishing returns, and η_ψ measures the envelope fit over feasible batches. Normalization makes $z_k^\star(S) \in [0,1]$ comparable across hypotheses. One shared envelope acts on normalized mass while the raw gains γ_k have hypothesis-dependent scale, so η_ψ absorbs this scale mismatch and grows with it. Nonnegative per-hypothesis scales or concave envelopes preserve submodularity, but they define a different surrogate and require corresponding fidelity bounds.

Proposition 1 (Submodular by construction)**.** *For any nonnegative importances, $\alpha_k \geq 0$, and nondecreasing concave ψ with $\psi(0) = 0$, the surrogate is normalized, monotone, and submodular.*

The covered mass is modular, a nondecreasing concave function of it is monotone submodular, and nonnegative sums preserve both. Unlike G, the surrogate is submodular by design rather than by assumption on the data. Proposition 1 needs only nonnegative importances. Lemmas 2 and 3 are stated for row-normalized importances, whereas the reported runs on Synthetic, METABRIC, CIFAR-10 and CIFAR-100 use unnormalized transformed scores (Section 5), so on those datasets the structural guarantee applies and the fidelity lemmas do not.

Choice of envelope. We use $\psi_a(z) = \log(1+az)$, $a > 0$, which satisfies Proposition 1 for every a. As $a \to 0$, $\psi_a(z) \approx az$ is linear and the surrogate reduces to ranking features by posterior-averaged importance and taking the top of the list. Larger a favors spreading coverage across hypotheses. If every u_k is proportional to one vector v, then $F(S \mid x_M) = h\big(\sum_{j\in S} v(j)\big)$ for a nondecreasing h, and under a cardinality budget greedy returns the TOPK batch at every a. Saturation acts only on differences between hypotheses. An exact nesting claim during tuning requires including TOPK as an explicit candidate. This inclusion guarantees a score no lower than TOPK only on the validation split and does not imply the same ordering on test data.

4.3 STEP 3: PER-FEATURE IMPORTANCE

The ideal importance $u_k^\star$ is not observable. We estimate it from precomputed explanations. Let $\chi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a fixed transformation applied coordinatewise to the attribution scores. For a local attribution ϕ^k of neighbor k, define $w_k(j) = \chi(\phi_j^k)$, let $Z_k = \sum_{\ell \in A(M)} w_k(\ell)$, and set

$$u_k(j) = \begin{cases} w_k(j)/Z_k, & Z_k > 0, \\ 1/|A(M)|, & Z_k = 0, \end{cases} \tag{6}$$

for a nonempty candidate set. The structural result requires only nonnegative transformed scores. A strictly positive transformation guarantees $Z_k > 0$. A nonnegative transformation also satisfies Proposition 1 when paired with the zero-mass fallback in equation 6. The reported runs use the magnitude $\chi(t) = |t|$, the exponential $\chi_\beta(t) = \exp(t/\beta)$, and the exponential of the magnitude $\chi_\beta(t) = \exp(|t|/\beta)$, with $\beta > 0$. On DDXPlus a zero attribution is kept at zero before row normalization, so there the exponential transforms are nonnegative rather than strictly positive. Attributions are computed once per reference instance and reused, so once the predictor, encoder and any attribution model are available, online selection adds no training. This yields the deployed surrogate $F(S \mid x_M) = \sum_k \alpha_k \, \psi(z_k(S))$ with $z_k(S) = \sum_{j\in S} u_k(j)$.

Lemma 3 (Explanation fidelity)**.** *Let ψ be L_ψ-Lipschitz and assume $\max_k \mathrm{TV}(u_k, u_k^\star) \leq \tau$ over $A(M)$. Then*

$$\big|F(S \mid x_M) - F^\star(S \mid x_M)\big| \;\leq\; 2\, L_\psi\, \tau \;=:\; \varepsilon_3.$$

For the envelope ψ_a, $L_\psi = a$, so $\varepsilon_3 \leq 2a\tau$.

The constant is loose. For probability vectors $\sup_S |\sum_{j\in S}(u_k(j) - u_k^\star(j))| = \mathrm{TV}(u_k, u_k^\star)$, so $L_\psi \tau$ suffices, and for ψ_a the bound $|\log(1+ax) - \log(1+ay)| \leq \log(1+a|x-y|)$ on $x, y \geq 0$ gives $|F - F^\star| \leq \log(1+a\tau)$. We keep the looser constant in Theorem 1. The quantity τ separates utility-estimation error from the other approximations, but it is not directly observable because $u_k^\star$ is latent. The analysis accepts any normalized nonnegative importance signal, and the choice of χ enters the theorem through τ. Proposition 4 analyzes the identity transform applied to the nonnegative Shapley vector of the same loss-reduction game. A nonlinear transform such as an exponential requires a corresponding bound on the transformed utilities. The proposition supplies the τ in this lemma only when its reference choice also satisfies Lemma 2. The utilities used in our experiments are transformed attributions of the predictor's output rather than Shapley values of the loss-reduction game (Section 5), so the proposition does not supply their τ, and for them τ is an assumption.

4.4 STEP 4: GREEDY SELECTION AND END-TO-END GUARANTEE

The policy q maximizes the deployed surrogate F greedily. From $S_0 = \varnothing$, at step r it adds

$$j_r \in \arg \max_{\substack{j \in A(M) \setminus S_{r-1} \\ c(S_{r-1}) + c_j \le B}} \frac{\Delta(j \mid S_{r-1})}{c_j}, \quad \Delta(j \mid S) = \sum_k \alpha_k \big[\psi(z_k(S) + u_k(j)) - \psi(z_k(S))\big], \quad (7)$$

stopping when no feature fits the budget, and returns $q(x_M) = S$. No value in S is seen during selection, so the procedure is one-shot, with per-query cost $O(b\,|A(M)|\,K)$. Combining the greedy guarantee for monotone submodular maximization with the three fidelity bounds gives our main result.

Theorem 1 (End-to-end guarantee)**.** *Assume the hypotheses of Lemmas 1–3, and set $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$. Let $S^\star$ be an optimal feasible batch for G and let $q(x_M)$ be the batch returned by greedy maximization of F. Under a cardinality budget, with probability at least $1 - \eta$ over the draw of the reference set,*

$$G\big(q(x_M) \mid x_M\big) \;\ge\; \Big(1 - \tfrac{1}{e}\Big)\, G(S^\star \mid x_M) \;-\; \Big(2 - \tfrac{1}{e}\Big)\, \varepsilon.$$

The same bound holds under a cost budget when the cost-benefit rule is combined with partial enumeration over seed sets of cardinality at most three.

As the three errors vanish, greedy recovers the classical $\big(1 - \frac{1}{e}\big)$ guarantee against the true objective G. The probability $1 - \eta$ comes from Lemma 1 alone, since ε_2 and ε_3 are deterministic given their assumptions. Because F is submodular by design, the classical greedy factor applies without any assumption on the true objective, and the three approximation errors enter additively rather than multiplied by that factor. The bound is not a numerical certificate. The error terms contain latent quantities that we do not estimate, and ε_3 scales with the curvature a, whose validation-selected values span several orders of magnitude across datasets (Table 3), so we report the theorem as a conditional transfer result. When the acquired actions are strongly complementary, the saturation-fit term ε_2 can be large; the guarantee exposes that failure rather than removing it. The multi-channel variant keeps the structural guarantee, and the redundancy-penalized variant keeps it for penalty weights below the threshold of Appendix A.5.

5 EXPERIMENTAL SETTING

Datasets. We evaluate on five datasets and three action types. On **Synthetic** and **METABRIC** an action acquires one feature: Synthetic (SYNAEA-2) has 104 features and eight classes, and the three informative features of an instance depend on its class and on a latent context among four that is observed for free, so no fixed feature set suits every instance. METABRIC is a breast cancer cohort of 1,898 samples with 489 mRNA expression z-scores and six subtypes. On **CIFAR-10** and **CIFAR-100** each 32×32 image is split into 64 patches of 4×4 pixels and an action reveals one patch. On **DDXPlus** an action asks one of 223 evidence questions and the label is one of 49 pathologies. Details are in Appendix A.1.

Models. On Synthetic, METABRIC, CIFAR-10 and CIFAR-100 the predictor f and encoder g are one masked network, trained once per dataset and seed with uniformly random masks and then frozen. The mask is concatenated to the input so an unacquired action differs from an acquired zero, and the retrieval embedding $g(x_M)$ is the network's penultimate activation under the context-only mask. We use a ResNet-18 with a linear head on the image datasets and multilayer perceptrons on the tabular ones. On DDXPlus the predictor is a Laplace-smoothed naive Bayes classifier fitted on the reference sample, with each question encoded as its own block sized by its answer type and "not asked" as a separate state. We report top-1 accuracy throughout.

Compared methods. All methods start from the same observed context and are scored by the same frozen predictor. Each is tuned on validation under its own selection rule. **Fixed acquisition path** (TOPK-POOLED) ranks actions once by importance pooled over the reference set and acquires the same batch for every instance. It is the zero line of Figure 1. **TOPK** ranks $A(M)$ by the posterior-averaged importance $\sum_k \alpha_k u_k(j)$, the $a \to 0$ limit of the surrogate. On Synthetic and METABRIC it shares the neighborhood size and utility transform with HUC, so there the two differ

Table 1: Selection time per instance (ms) at the largest reported budget, excluding retrieval on the feature and image datasets and including it on DDXPlus. CIFAR-10 and CIFAR-100: one NVIDIA A10G (one CIFAR-10 seed on an NVIDIA L4), 4 vCPU, every method on the GPU. Synthetic, METABRIC and DDXPlus: 48 vCPU Intel Xeon Platinum 8259CL at 2.5 GHz; Top-K, HUC and its extensions run in NumPy on the CPU, AACO-greedy in PyTorch on one NVIDIA T4; DDXPlus timings are from one seed. Combined was timed but not accuracy-evaluated on DDXPlus. Devices and retrieval accounting differ across columns, so ratios are comparable within a column only.

ms / instance	Synthetic B=25	DDXPlus B=7	METABRIC B=25	CIFAR-10 B=20	CIFAR-100 B=20
Top-K	0.06	1.67	0.37	0.05	0.05
HUC	1.94	2.55	8.03	1.18	1.24
+ multi-index	3.08	2.58	17	1.58	1.97
+ redundancy	1.93	4.58	9.92	3.19	3.22
+ multi-index + redundancy	3.12	4.63	11	3.56	3.93
Combined	13	9.36	75	33	35
AACO-greedy	24	7.43	359	296	312

only in the envelope. On the image datasets each selects its own utility transform, and on DDXPlus the two share the transform but HUC selects its neighborhood size jointly with the curvature, so those comparisons are between tuned selectors. **HUC** is greedy maximization of the deployed surrogate F with $R = 1$ and $\lambda = 0$. **AACO-greedy** is a one-shot adaptation of the acquisition-conditioned oracle of Valancius et al. (2024): it imputes each candidate batch from ten uniformly weighted retrieved neighbors, scores it by the predictive loss of the frozen predictor on the imputed completions, commits one action at a time, and re-scores every remaining candidate after each commitment. Each score is a forward pass per neighbor and candidate, so on the image datasets every greedy step scores c candidates drawn uniformly at random from the remaining patches ($c = 16$ on CIFAR-10 and $c = 32$ on CIFAR-100, out of 61 and 55), which makes it a sampled approximation of the exact greedy there. A variant that scores $1{,}000$ random batches was worse on Synthetic and is omitted; the published AACO policy is sequential and not a one-shot competitor. **Combined** selects, per instance, the best of the four HUC variants' batches under the AACO objective; Appendix A.5 reports its accuracy on the four datasets other than DDXPlus and Table 1 its selection time on all five. **Random** samples a feasible batch uniformly.

Protocol. Every instance starts from a nonempty context M, excluded from the candidate set and the budget: the two context features on Synthetic, a fixed set of six features on METABRIC, the three central patches on CIFAR-10, the 3×3 central block on CIFAR-100, and on DDXPlus the released presenting complaint with age and sex. We report budgets of 5 to 25 features, 5 to 20 patches, and 1 to 7 questions. Every statement below refers to these plotted budgets. Utilities are precomputed once per reference instance and serve as proxies for the ideal importances of Section 4: tree-SHAP for METABRIC, Shapley values of the encoder's log-probability for Synthetic, a FastSHAP explainer conditioned on M for the image datasets, and truncated permutation Shapley values for DDXPlus. The selected nonnegative transform is the magnitude on Synthetic and METABRIC and is selected per budget on validation on the image datasets and DDXPlus; the transformed scores are row-normalized on DDXPlus only. Hyperparameters are selected on validation, per budget and in stages, with the grids of the extensions containing $R = 1$ and $\lambda = 0$ so that an extension is used only when it improves validation accuracy. The reference set is the training split, joined by the validation split on Synthetic and the image datasets. On DDXPlus both the reference sample and the hyperparameter-selection patients are drawn from the released validation split and overlap by about seven percent. The test splits were evaluated repeatedly during development, so the reported numbers are benchmark estimates on the final configurations rather than single-use held-out estimates. Accuracies are means over three seeds that redraw the split and retrain the predictor. Differences are paired within seed and error bars are the half range across seeds. Split sizes, overlap counts, grids, and the selected hyperparameters are in Appendix A.2.

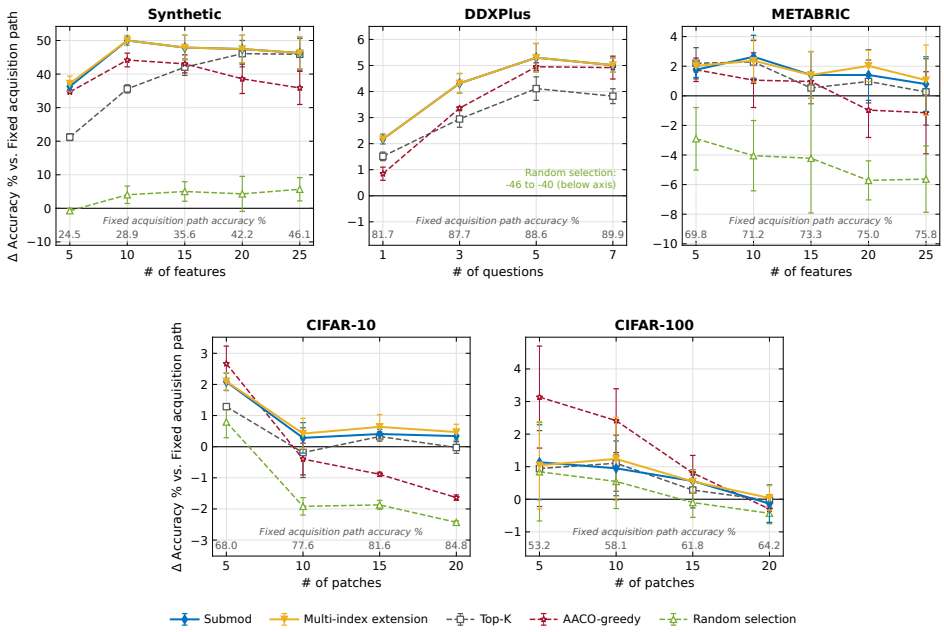


Figure 1: Paired accuracy margin over the fixed acquisition path (zero line) versus budget, mean over three seeds; error bars are the half range across seeds. The fixed path's absolute accuracy is printed along the bottom of each panel. HUC is the line labelled "Submod" and the multi-index extension is shown for reference. On DDXPlus, random selection lies 40 to 46 points below the axis and is omitted.

6 Results and Discussion

Figure 1 plots each method as its paired margin over the fixed acquisition path, whose absolute accuracy is printed along the bottom of each panel (tabulated in Appendix A.3), and Table 1 reports selection time per instance. Numbers below are percentage points of top-1 accuracy, mean over seeds, at the plotted budgets.

Instance-conditional selection gives large gains on Synthetic and DDXPlus and small or absent gains elsewhere. On Synthetic no single path can track an informative set that changes with the instance, so the fixed path is no better than random, and HUC adds 36 to 50 points. On DDXPlus the fixed path reaches 82% at one question and 88 to 90% from three on, and HUC adds 2 points at one question, 4 at three, and 5 from five on. On METABRIC the gain is one to three points through twenty features and under one at twenty-five, on CIFAR-10 it is two points at five patches and under half a point beyond, and on CIFAR-100 it is about one point at five patches and gone at twenty, where HUC sits 0.1 points below the fixed path. Random selection is above the fixed path at five patches on both image datasets and below it at the largest plotted budget on every dataset except Synthetic, where the fixed path itself falls below random from ten features on.

On Synthetic, set-dependent saturation accounts for the gain over ranking. TopK and HUC differ in two ways: the concave ψ_a rescores individual actions across hypotheses, and the greedy on F discounts an action whose hypotheses earlier picks have already covered. To separate them we add a third selector, CS, that ranks actions by the concave singleton score $\sum_k \alpha_k \psi_a(u_k(j))$ and takes the top b. It shares α, u, K and a with the other two and coincides with HUC at $b = 1$. We evaluate all three on 1,600 freshly generated Synthetic queries for each of the three model seeds, with the validation-selected (K, a) of the record runs (Appendix A.4, Table 5). Rescoring alone does not help: CS is within a point of TopK at five and ten features and one to two points below it at larger budgets. The saturating greedy is what moves the accuracy: HUC exceeds CS by 16 points at five and ten features, by 9 at fifteen, by 4 at twenty and by 2 at twenty-five, with paired bootstrap intervals excluding zero in every seed through twenty features. The picture is the same

when the curvature is fixed at $a \in \{1, 10, 100\}$ instead of selected (Table 6), and the record test-split gap between HUC and TopK is reproduced exactly by the same code, with the same direction and scale on the fresh queries.

On METABRIC, where the two selectors share every input, they are within noise of each other. On DDXPlus and the image datasets they are tuned separately, so their gap compares tuned selectors rather than isolating the envelope: HUC is above TopK by 0.7 to 1.4 points on DDXPlus, positive at every budget and seed, and by 0.1 to 0.8 on CIFAR-10, and within 0.3 points on CIFAR-100, where TopK is itself within a point of random, so the utilities carry little usable selection signal.

HUC replaces predictor-based batch rescoring at lower online selection cost. AACO-greedy is the only baseline that consults the predictor during selection. It re-scores every remaining candidate after each commitment, whereas HUC commits to precomputed utilities. HUC is ahead at every plotted budget on Synthetic (2 points at five features, 10 at twenty-five) and METABRIC (0 to 2), and on average at every budget on DDXPlus (1.3 points at one question, 0.1 at seven, with one seed marginally below at five and seven). On the image datasets the two cross: AACO-greedy leads by 0.6 (CIFAR-10) and 2.0 (CIFAR-100) at five patches, and HUC leads at twenty by 2.0 on CIFAR-10 and by 0.2 on CIFAR-100, where it trails at ten and fifteen as well. Re-scoring is worth most when the batch is small and each patch changes the imputed completion. Neither TopK nor HUC calls the predictor after retrieval, while AACO-greedy evaluates it on every imputed completion of every candidate at every step (Table 1): a factor of 12 (Synthetic) and 45 (METABRIC) in selection time, about 250 on the image datasets with the candidate set already sampled down, and about 3 on DDXPlus, where the naive Bayes predictor makes re-scoring cheap.

Extensions. The multi-channel and redundancy-penalized variants (Appendix A.5) stay within 0.3 points of HUC on the image datasets and DDXPlus and within a point on METABRIC; on Synthetic the redundancy penalty adds one to two points at five to fifteen features and nothing beyond.

Summary. On the controlled benchmark the large gain over ranking is the set-dependent term. On the other datasets the gains over a fixed path and over tuned ranking are consistent on DDXPlus and small or absent on the real tabular and image data. Against direct batch rescoring HUC is ahead on the non-image datasets and at the larger image budgets and behind at the smallest ones, at a selection cost lower by one to two orders of magnitude except on DDXPlus. It is a one-shot selector with a surrogate guarantee that is competitive with direct batch scoring and does not dominate it.

7 Conclusion

We studied one-shot batch active evidence acquisition, where a complete set of costly actions is committed before any outcome is observed, and introduced HUC, a surrogate that represents the unknown completion by retrieved completed instances, transfers their precomputed explanation-derived utilities, and applies diminishing returns to the utility a batch covers within each completion. The surrogate is monotone submodular by construction, so greedy selection carries the classical $(1 - 1/e)$ guarantee on the surrogate, and a conditional bound separates its gap from the ideal objective into retrieval, saturation-fit, and utility-estimation terms. On a controlled synthetic benchmark a matched control showed that the gain over ranking is set-dependent saturation and that rescoring individual actions with the same concave function contributes nothing. On DDXPlus the method improves on a fixed acquisition path and on tuned ranking at every plotted budget; on the real tabular and image data the gains are small or absent; and against direct batch rescoring it trades accuracy at some image budgets for a large reduction in selection cost.

Sequential policies, which observe each outcome before choosing the next, lie outside the setting and should be compared under measured waiting time and execution cost. The transferable idea is the control. When a coverage-style objective is credited with a gain, ranking by the same concave singleton score separates set dependence from rescoring, and it costs one extra sort. The open question is whether the completion-specific utilities with substitutable coverage that make saturation pay on the synthetic benchmark appear on real data under a shared protocol; the present real-data comparisons are between tuned selectors and do not settle it.

REPRODUCIBILITY STATEMENT

The synthetic benchmark is fully specified by its generator (Appendix A.1); the public datasets, splits, masking, and predictor training are summarized in Appendices A.1 and A.2, the selection procedure in Algorithm 1, and the selected hyperparameters in Table 3. The result numbers in Section 6 and Appendices A.3–A.4 are generated or checked against the experiment result files by scripts, and the mechanism experiment of Appendix A.4 reproduces the record test-split accuracies of TOPK and HUC exactly before evaluating the fresh queries. Code and result files are not released with this submission.

STATEMENT ON THE USE OF LARGE LANGUAGE MODELS

Large language model assistants were used in this work for literature retrieval, for drafting and editing text, for writing and auditing analysis code, and for checking proof steps and the consistency between the theorem statements and the implementation. All mathematical statements and proofs were reviewed by the authors, and every reported number was regenerated from the experiment artifacts by scripts that the authors ran and inspected. The authors take full responsibility for the content.

REFERENCES

Linus Aronsson, Arman Rahbar, and Morteza Haghir Chehreghani. A survey on active feature acquisition strategies. *arXiv preprint arXiv:2502.11067*, 2025.

Vedang Asgaonkar, Aditya Jain, and Abir De. Generator assisted mixture of experts for feature acquisition in batch. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(10):10927–10934, Mar. 2024. doi: 10.1609/aaai.v38i10.28967.

Jianbo Chen, Le Song, Martin Wainwright, and Michael Jordan. Learning to explain: An information-theoretic perspective on model interpretation. In Jennifer Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 883–892. PMLR, 10–15 Jul 2018.

Ian Covert, Scott M Lundberg, and Su-In Lee. Understanding global feature contributions with additive importance measures. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 17212–17223. Curran Associates, Inc., 2020.

Ian Connick Covert, Wei Qiu, Mingyu Lu, Na Yoon Kim, Nathan J White, and Su-In Lee. Learning to maximize mutual information for dynamic feature selection. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 6424–6447. PMLR, 23–29 Jul 2023.

Gabriel Dulac-Arnold, Ludovic Denoyer, Philippe Preux, and Patrick Gallinari. Datum-wise classification: A sequential approach to sparsity. In Dimitrios Gunopulos, Thomas Hofmann, Donato Malerba, and Michalis Vazirgiannis (eds.), *Machine Learning and Knowledge Discovery in Databases*, pp. 375–390, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

Arsene Fansi Tchango, Rishab Goel, Zhi Wen, Julien Martel, and Joumana Ghosn. DDXPlus: A new dataset for automatic medical diagnosis. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 31306–31318. Curran Associates, Inc., 2022. doi: 10.52202/068431-2270.

Osman Berke Guney, Ketan Suhaas Saichandran, Karim Elzokm, Ziming Zhang, and Vijaya B Kolachalama. Active feature acquisition via explainability-driven ranking. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 20748–20765. PMLR, 13–19 Jul 2025.

Jaromír Janisch, Tomáš Pevný, and Viliam Lisý. Classification with costly features using deep reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01): 3959–3966, Jul. 2019. doi: 10.1609/aaai.v33i01.33013959.

Neil Jethani, Mukund Sudarshan, Ian Connick Covert, Su-In Lee, and Rajesh Ranganath. FastSHAP: Real-time Shapley value estimation. In *International Conference on Learning Representations*, 2022.

Andreas Krause and Daniel Golovin. Submodular function maximization. In Lucas Bordeaux, Youssef Hamadi, and Pushmeet Kohli (eds.), *Tractability: Practical Approaches to Hard Problems*, pp. 71–104. Cambridge University Press, 2014. doi: 10.1017/CBO9781139177801.004.

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. PMLR, 18–24 Jul 2021.

Hui Lin and Jeff Bilmes. A class of submodular functions for document summarization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 510–520, Portland, Oregon, USA, 2011. Association for Computational Linguistics. URL https://aclanthology.org/P11-1052/.

Jiayu Liu, Cheng Qian, Zhaochen Su, Qing Zong, Shijue Huang, Bingxiang He, and Yi R. Fung. CostBench: Evaluating multi-turn cost-optimal planning and adaptation in dynamic environments for LLM tool-use agents. In Maria Liakata, Viviane P. Moreira, Jiajun Zhang, and David Jurgens (eds.), *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 12826–12858, San Diego, California, United States, July 2026. Association for Computational Linguistics. doi: 10.18653/v1/2026.acl-long.584.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, 1978.

Alexander Luke Ian Norcliffe, Changhee Lee, Fergus Imrie, Mihaela Van Der Schaar, and Pietro Lio. Stochastic encodings for active feature acquisition. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 46784–46814. PMLR, 13–19 Jul 2025.

Hajin Shim, Sung Ju Hwang, and Eunho Yang. Joint active feature acquisition and classification with variable-size set encoding. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Maxim Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, 2004. doi: 10.1016/S0167-6377(03)00062-2.

Kirill Trapeznikov and Venkatesh Saligrama. Supervised sequential classification under budget constraints. In Carlos M. Carvalho and Pradeep Ravikumar (eds.), *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, volume 31 of *Proceedings of Machine Learning Research*, pp. 581–589, Scottsdale, Arizona, USA, 29 Apr–01 May 2013. PMLR.

Michael Valancius, Maxwell Lennon, and Junier Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 48957–48975. PMLR, 21–27 Jul 2024.

Duo Wu, Jinghe Wang, Yuan Meng, Yanning Zhang, Le Sun, and Zhi Wang. CATP-LLM: Empowering large language models for cost-aware tool planning. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 8699–8709, 2025. doi: 10.1109/ICCV51701.2025.00814.

Zhixiang Xu, Kilian Q. Weinberger, and Olivier Chapelle. The greedy miser: learning under test-time budgets. In *Proceedings of the 29th International Conference on Machine Learning*, pp. 1299–1306. Omnipress, 2012.

Jinsung Yoon, James Jordon, and Mihaela van der Schaar. INVASE: Instance-wise variable selection using neural networks. In *International Conference on Learning Representations*, 2019.

A Appendix

A.1 Dataset details

A.1.1 Tabular datasets

METABRIC is a breast cancer cohort of $1{,}898$ samples described by 489 mRNA expression z-scores, with six molecular subtypes as the label. We use the data and the accompanying tree-SHAP attribution arrays released with the benchmark of Guney et al. (2025), so that the utilities entering equation 6 are the published ones rather than a reimplementation. Preprocessing follows the source publication: identifier columns are dropped, categorical columns that are neither ordinal nor binary are dropped, columns with more than ten percent missing values are dropped, and the remaining missing values are mean imputed. Features are standardized with the mean and variance of the training split.

Synthetic benchmark (SynAEA-2). We construct a controlled acquisition problem in which the informative features depend on a latent context that is revealed for free, the remaining features are either distractors or partially redundant copies, and actions carry heterogeneous costs. Each instance has a context $t \in \{1, \dots, 4\}$ drawn uniformly and a class $c \in \{1, \dots, 8\}$ drawn from a mildly skewed prior $\pi_k \propto \exp(-0.08\,k)$ (realized majority share 0.168 in the generated data, chance 0.125). Two *context* features, $x_0 = t + \epsilon$ and $x_1 = -t + \epsilon$ with $\epsilon \sim \mathcal{N}(0, 0.15^2)$, identify the context nearly perfectly but carry no class information; they are observed by every instance and form the warm start (a random forest on these two features alone reaches 0.138 accuracy). For every (context, class) pair (t', k) there is a *question* with $M = 3$ correlated copies,

$$x_{t',k,m} \;=\; 3.0 \cdot \mathbb{1}[t' = t,\; k = c] \;+\; 0.7\, g_{t',k} \;+\; \sigma_{\tau(t',k)}\, \eta_{t',k,m}, \qquad m = 1, 2, 3,$$

where $g_{t',k} \sim \mathcal{N}(0, 1)$ is a latent factor shared by the three copies and $\eta_{t',k,m} \sim \mathcal{N}(0, 1)$ is independent noise. Only the three copies of question (t, c) carry signal for an instance with context t and class c; the 93 remaining question features are distractors. Because the copies share $g_{t',k}$, acquiring additional copies reduces the independent noise but never the shared component, yielding saturating rather than additive returns within a question. Six further features are pure $\mathcal{N}(0, 1)$ noise. In total each instance has $d = 2 + 4 \cdot 8 \cdot 3 + 6 = 104$ features. The 32 questions are assigned round-robin to three tiers with per-feature costs $\{1, 2, 3\}$ and noise scales $\sigma \in \{1.6, 1.0, 0.6\}$ respectively, so that cheaper questions are noisier and cost buys precision; context and noise features cost 1. Under a cardinality budget the costs are ignored; under a price budget the problem is a knapsack. We generate $N = 8{,}000$ instances (seed 0) and use $4{,}800$ for training and $1{,}600$ for testing, holding out the remainder for validation. A random forest on all 104 raw features attains 0.86 test accuracy, a reference for the headroom above the 0.138 warm start rather than an upper bound (the masked network reaches 0.87 with every feature observed). The reported runs use attributions from the trained encoder; exact tree path contributions (Saabas) from a 100-tree random forest ($\min$ leaf size 3) were also computed and agree with them at a flat correlation of 0.71 and a mean per-instance correlation of 0.85.

Each dataset is split into disjoint train, validation, and test partitions in proportions $60/20/20$, drawn before standardization. The train partition supplies the predictor's training data; the reference set $\mathcal{D}$ is the train partition on METABRIC and the train and validation partitions on Synthetic; the validation partition is used for all hyperparameter selection. The test partition was evaluated repeatedly during development, so the reported numbers are benchmark estimates on the final configurations rather than single-use held-out estimates. The observed context M is fixed and identical for every instance and every method: the two context features on Synthetic and six features on METABRIC. Budgets run over eleven cardinalities from one to 25 features, so that the small-budget regime, where selection matters most, is densely sampled.

The test splits hold a few hundred instances, so a single split cannot resolve differences of the size the methods produce: one percentage point on the METABRIC split is a handful of instances. We therefore repeat the entire pipeline, split included, over three seeds and report differences paired within seed, which removes the between-seed variation the arms share and isolates the effect of the acquisition rule.

Table 2: Dataset and action-space summary. For DDXPlus, the sample counts are the reference, validation, and test patients used by the protocol. The action count is the full action universe before removing the initially observed context.

Dataset	Domain	Samples	Actions	Targets	Acquired unit
Synthetic	Tabular	8,000	104	8	Feature
METABRIC	Tabular	1,898	489	6	Feature
CIFAR-10	Image	60,000	64	10	4×4 patch
CIFAR-100	Image	60,000	64	100	4×4 patch
DDXPlus	Diagnosis	10,000/4,096/134,529	223	49	Evidence question

A.1.2 IMAGE DATASETS

Each 32×32 image of **CIFAR-10** and **CIFAR-100** is divided into a regular grid of 64 non-overlapping 4×4 patches and an action reveals one patch; unrevealed patches are replaced by a constant fill value, and the patch mask is supplied to the network alongside the image so that a hidden patch is distinguishable from a uniform one. The observed context is fixed and identical for every image and is excluded from both the candidate set and the budget: the three central patches on CIFAR-10 and the 3×3 block of nine central patches on CIFAR-100. Per-patch attributions are produced by FastSHAP surrogate and explainer networks trained on the training split, conditioned on the observed context, and then run in inference only, so no attribution is computed during selection. The reference set $\mathcal{D}$ is the union of the training and validation splits, the validation split supplies the hyperparameters, and the query instances come from the test split, so a query is never its own neighbor.

The acquisition-conditioned baseline is treated differently here than on the tabular datasets. Because the masking layer feeds a convolutional network, the value of a candidate batch cannot be decomposed into per-patch contributions, so each candidate requires a complete forward pass and the cost of one budget grows with the product of the batch size, the number of candidates, and the neighborhood size. Each of its greedy steps therefore scores c candidates drawn uniformly at random from the remaining patches, $c = 16$ on CIFAR-10 and $c = 32$ on CIFAR-100, so the baseline is a sampled approximation of the exact greedy on these two datasets; every method is evaluated on the full test split.

A.1.3 DDXPLUS

DDXPlus (Fansi Tchango et al., 2022) is a synthetic differential diagnosis corpus generated from a proprietary medical knowledge base; it contains no real patient records. An action asks one of 223 evidence questions, and the answer type varies by question: binary, single-choice categorical, or multi-select. Each question occupies its own block in the predictor's input, sized by its answer type, with "not asked" encoded as its own state rather than as a zero, so that an unasked question is never confused with a negative answer.

The observed context is instance-specific rather than fixed. Every patient record carries a presenting complaint, and we take that single evidence item as the initial observation, together with age and sex, which are always available and are never acquirable. This is closer to a clinical workup than a fixed context, and it also means the candidate set varies slightly across patients. We sample $10{,}000$ patients from the released validation split as the reference set and $4{,}096$ patients from the same split for hyperparameter selection; the two samples overlap in 309, 311 and 319 patients across the three seeds (about 7.5% of the selection sample). We evaluate on the full test split of $134{,}529$ patients at budgets of one, three, five, and seven questions.

Attributions here are computed rather than downloaded, as truncated permutation Shapley values of the naive Bayes predictor's log-probability over the 223 questions, with $2{,}048$ antithetic permutations walked in segments of 32 questions; each estimate is unbiased for the marginal contribution averaged over preceding coalitions whose size is uniform on 0 to 31, a truncated form of the Shapley value and not the full Shapley value. Because the test split is large, differences of a fraction of a percentage point are resolvable; the three seeds differ in the reference sample and the trained predictor, not in the test instances.

A.2 PROTOCOL DETAILS

Reference sets and queries. The reference set is the training split on METABRIC, the training plus validation splits on Synthetic, CIFAR-10 and CIFAR-100, and a 10,000-patient sample of the released validation split on DDXPlus. Queries come from the test split, so a query is never its own neighbor. On CIFAR-100 the context is the 3×3 central block because with three patches a 100-way predictor is near chance.

Hyperparameter selection. All selection is on the validation split, per budget and in stages. On Synthetic and METABRIC, TOPK selects the neighborhood size $K \in \{5, 7, 11, 21, 51, 101\}$ and HUC inherits it and selects the curvature a from $\{a_{\text{auto}}, 1, 10, 100\}$, where a_{auto} is the reciprocal of the mean nonzero utility; the utility transform is the magnitude. On the image datasets TOPK selects $K \in \{11, 51, 101\}$ and a transform from $|t|$, $\exp(t/\beta)$ and $\exp(|t|/\beta)$ with $\beta \in \{0.1, 0.5, 1\}$, and HUC selects its own transform and a at that K, using a fixed subsample of 512 (CIFAR-10) or 1,024 (CIFAR-100) validation instances. On DDXPlus the transform is selected per budget from the same grid for TOPK and shared by HUC (the exponential with $\beta = 0.1$ at every plotted budget, except $\beta = 1$ at one question in two seeds); HUC then selects K (up to 800) and a jointly, with ties resolved to the smallest K and then the smallest a. The multi-index and penalized variants then select R and λ from grids containing $R = 1$ and $\lambda = 0$. Table 3 lists the selected (K, a) for HUC at the plotted budgets; a is on the scale of the utilities, which differ across datasets. The transformed scores enter unnormalized on Synthetic, METABRIC and the image datasets and row-normalized on DDXPlus, where a zero attribution stays zero after the transform.

Test-set use. The test splits were evaluated many times during development, under configurations that differed from the final ones, before the reported runs were fixed. Within each reported run hyperparameters were selected on validation rather than test accuracy; repeated test diagnostics could nevertheless influence development, so the reported test numbers are benchmark estimates on the final configurations rather than single-use held-out estimates. On DDXPlus a seed redraws the reference sample and retrains the predictor and every seed is evaluated on the full test split of 134,529 patients; on the tabular datasets one percentage point is a handful of test instances and is not treated as real.

Table 3: Hyperparameters selected on validation for HUC at the plotted budgets, shown as seed $0/1/2$: neighborhood size K and curvature a. On Synthetic and METABRIC K is inherited from TOPK; on the image datasets TOPK and HUC may select different utility transforms; on DDXPlus (K, a) are selected jointly for HUC.

Dataset	b	K	a
Synthetic	5	51/101/101	100/100/21.2
Synthetic	10	101/101/101	100/100/100
Synthetic	15	51/101/51	20.9/100/100
Synthetic	20	101/101/101	100/10/21.2
Synthetic	25	101/101/101	100/1/1
METABRIC	5	5/11/11	100/1/1
METABRIC	10	7/5/5	1/100/10
METABRIC	15	101/11/5	1/10/10
METABRIC	20	5/21/101	1/10/1
METABRIC	25	5/11/101	1/1/1
CIFAR-10	5	51/11/11	12.5/0.1/1
CIFAR-10	10	11/51/51	1/0.1/0.2
CIFAR-10	15	11/101/51	0.441/0.05/0.1
CIFAR-10	20	51/101/11	0.05/0.05/0.1
CIFAR-100	5	101/101/51	10/0.2/0.00196
CIFAR-100	10	101/11/11	0.05/0.1/7.41e-06
CIFAR-100	15	51/11/101	0.5/0.05/7.41e-06
CIFAR-100	20	101/11/51	1/10/0.05
DDXPlus	1	200/200/100	5.68e+04/5.68e+04/5.68e+04
DDXPlus	3	200/50/200	5.68e+04/5.68e+04/5.68e+04
DDXPlus	5	200/200/400	3.55e+03/5.68e+04/1.42e+04
DDXPlus	7	400/400/200	1.42e+04/1.42e+04/3.55e+03

A.3 MARGINS BEHIND FIGURE 1

Table 4 tabulates the paired margins plotted in Figure 1 at the plotted budgets, together with HUC's margin over AACO-greedy.

Table 4: Paired accuracy margins of HUC at the plotted budgets, in percentage points: over the fixed acquisition path (TOPK-POOLED), over tuned TOPK, and over AACO-greedy. Mean over three seeds with the half range across seeds in parentheses. The fixed path's absolute accuracy is given for scale. These are the numbers plotted in Figure 1.

Dataset	b	Fixed path (%)	HUC−fixed	HUC−TOPK	HUC−AACO
Synthetic	5	24.5	+36.3 (0.9)	+15.1 (0.1)	+1.6 (1.2)
Synthetic	10	28.9	+50.0 (1.5)	+14.4 (0.9)	+5.8 (0.8)
Synthetic	15	35.6	+47.9 (3.7)	+5.9 (1.6)	+4.9 (2.3)
Synthetic	20	42.2	+47.5 (4.1)	+1.4 (0.5)	+8.9 (0.5)
Synthetic	25	46.1	+46.3 (4.7)	+0.4 (0.2)	+10.4 (0.8)
METABRIC	5	69.8	+1.8 (0.5)	−0.4 (0.5)	+0.0 (1.3)
METABRIC	10	71.2	+2.6 (1.5)	+0.4 (0.3)	+1.6 (0.4)
METABRIC	15	73.3	+1.4 (1.6)	+0.9 (0.5)	+0.4 (1.7)
METABRIC	20	75.0	+1.4 (1.7)	+0.4 (0.4)	+2.4 (1.5)
METABRIC	25	75.8	+0.8 (1.8)	+0.5 (0.9)	+1.9 (3.0)
CIFAR-10	5	68.0	+2.1 (0.3)	+0.8 (0.2)	−0.6 (0.4)
CIFAR-10	10	77.6	+0.3 (0.5)	+0.5 (0.4)	+0.7 (0.1)
CIFAR-10	15	81.6	+0.4 (0.2)	+0.1 (0.1)	+1.3 (0.2)
CIFAR-10	20	84.8	+0.3 (0.2)	+0.4 (0.1)	+2.0 (0.1)
CIFAR-100	5	53.2	+1.1 (1.1)	+0.2 (0.4)	−2.0 (0.4)
CIFAR-100	10	58.1	+0.9 (0.8)	−0.2 (0.3)	−1.5 (0.3)
CIFAR-100	15	61.8	+0.6 (0.4)	+0.3 (0.3)	−0.2 (0.4)
CIFAR-100	20	64.2	−0.1 (0.6)	−0.1 (0.6)	+0.2 (0.9)
DDXPlus	1	81.7	+2.2 (0.2)	+0.7 (0.1)	+1.3 (0.1)
DDXPlus	3	87.7	+4.3 (0.4)	+1.4 (0.4)	+1.0 (0.3)
DDXPlus	5	88.6	+5.3 (0.5)	+1.2 (0.4)	+0.3 (0.4)
DDXPlus	7	89.9	+5.0 (0.3)	+1.2 (0.2)	+0.1 (0.3)

A.4 MECHANISM DECOMPOSITION ON SYNTHETIC

Design. Three selectors share the retrieval weights α, the utilities u, the neighborhood size K and the curvature a of the record runs and differ only in how they turn them into a batch: TOPK ranks by $\sum_k \alpha_k u_k(j)$, CS ranks by $\sum_k \alpha_k \psi_a(u_k(j))$, and HUC is the cardinality greedy of Section 4.4. The difference CS−TOPK measures the effect of rescoring individual actions with the concave function; HUC−CS measures the effect of set dependence, the discount an action receives once its hypotheses are covered. At $b = 1$ the two concave selectors coincide, which we verify on every query. Queries are 1,600 instances drawn from the generator with a seed disjoint from every training, validation, reference and test split; the frozen predictors, reference pools, attributions and selected (K, a) of the three record seeds are used unchanged, and the same 1,600 queries are evaluated for each seed. Before the fresh queries are scored, the same code reproduces the record test-split accuracies of TOPK and HUC at every plotted budget and seed to the last digit.

Prespecified predictions. Four predictions were fixed before the fresh queries were scored. Two hold: HUC−CS is positive with the paired interval excluding zero at five and ten features in every seed, and it exceeds CS−TOPK at those budgets in every seed. Two fail as registered: at twenty-five features one seed still shows differences of four points in each direction (the prediction was that all differences fall below one point), and the fresh-query gap between HUC and TOPK at five features differs from the record test-split gap by 3.06 points in one seed against a registered window of three points (the other two seeds differ by -2.4 and -0.8 points). The window was narrower than the sampling variability of two 1,600-query estimates; the direction and scale of the gap replicate.

Table 5: Selector decomposition on Synthetic, 1,600 freshly generated queries, three model seeds. TopK ranks by $\sum_k \alpha_k u_k(j)$; CS (concave singleton) ranks by $\sum_k \alpha_k \psi_a(u_k(j))$; HUC is the greedy of Section 4.4. All three share α, u, K and a (the values selected on validation for the record runs). Accuracy in percent, mean over seeds; differences in percentage points, mean over seeds with the half range across seeds in parentheses; "CI" counts seeds whose 95% paired query bootstrap interval lies above / below zero. At $b = 1$ CS and HUC coincide by construction (verified on every query).

b	TopK	CS	HUC	CS−TopK	CI	HUC−CS	CI	HUC−TopK
5	42.5	41.4	57.6	−1.1 (1.0)	0/0	+16.1 (1.7)	3/0	+15.0 (2.7)
10	60.5	60.5	76.6	+0.1 (0.4)	0/0	+16.1 (1.2)	3/0	+16.1 (0.9)
15	75.2	73.8	82.6	−1.4 (0.5)	0/1	+8.7 (0.5)	3/0	+7.3 (1.0)
20	86.1	84.1	88.0	−2.0 (0.9)	0/3	+3.9 (1.7)	3/0	+1.9 (0.8)
25	90.4	88.9	90.6	−1.5 (1.9)	0/2	+1.8 (2.0)	2/0	+0.3 (0.3)

Table 6: The decomposition of Table 5 with the curvature fixed at $a \in \{1, 10, 100\}$ instead of the validation-selected value (percentage points, mean over three seeds, fresh queries).

	$a = 1$		$a = 10$		$a = 100$	
b	HUC−CS	CS−TopK	HUC−CS	CS−TopK	HUC−CS	CS−TopK
5	+10.7	−1.3	+16.5	−1.8	+16.4	−0.8
10	+11.1	−0.7	+16.7	−1.1	+16.1	+0.1
15	+4.9	−0.2	+7.8	−0.6	+9.3	−2.0
20	+1.3	+0.0	+2.8	−1.0	+4.3	−2.4
25	+0.5	−0.1	+1.3	−1.0	+4.1	−3.8

A.5 Extensions: multi-channel utilities and a redundancy penalty

The scalar construction extends to multiple utility channels when the aggregator is separable. For channel $r \in [R]$, let $z_k^{(r)}(S) = \sum_{j \in S} u_k^{(r)}(j)$ with $u_k^{(r)}(j) \geq 0$, and let each ψ_r be nondecreasing and concave with $\psi_r(0) = 0$. Define

$$F_{\text{multi}}(S \mid x_M) = \sum_k \alpha_k \sum_{r=1}^{R} \psi_r\big(z_k^{(r)}(S)\big).$$

For a pairwise redundancy penalty, let

$$P(S) = \sum_{j<\ell \in S} \rho(j, \ell), \qquad F_\lambda(S \mid x_M) = F(S \mid x_M) - \lambda P(S), \qquad \rho(j, \ell) = \rho(\ell, j) \geq 0.$$

F_{multi} is normalized, monotone, and submodular, so greedy keeps the $(1 - 1/e)$ surrogate guarantee; F_λ is normalized and submodular for every $\lambda \geq 0$ and monotone for λ below a data-dependent threshold $\lambda_{\max}$, and the end-to-end bound of Theorem 1 transfers with an additive $\varepsilon_\rho = \lambda \max_{S \in \mathcal{F}_B} P(S)$ (Propositions 2 and 3 below). Separability is what makes the multi-index form work: joint concavity alone does not give submodularity. Under a cardinality budget the penalized greedy completes the batch in decreasing posterior-averaged importance so that every method uses the same number of actions.

Guarantees.

Proposition 2 (Separable multi-index surrogate)**.** *The function F_{multi} is normalized, monotone, and submodular. Therefore, greedy maximization attains the usual $(1 - 1/e)$ surrogate guarantee under a cardinality constraint. An end-to-end guarantee against G additionally requires saturation and utility-fidelity bounds for F_{multi} analogous to Lemmas 2 and 3.*

The separable form is a sufficient condition for the stated structural result. Joint concavity and coordinatewise monotonicity alone do not imply submodularity. For example, $\Psi(z_1, z_2) = \min(z_1, z_2)$

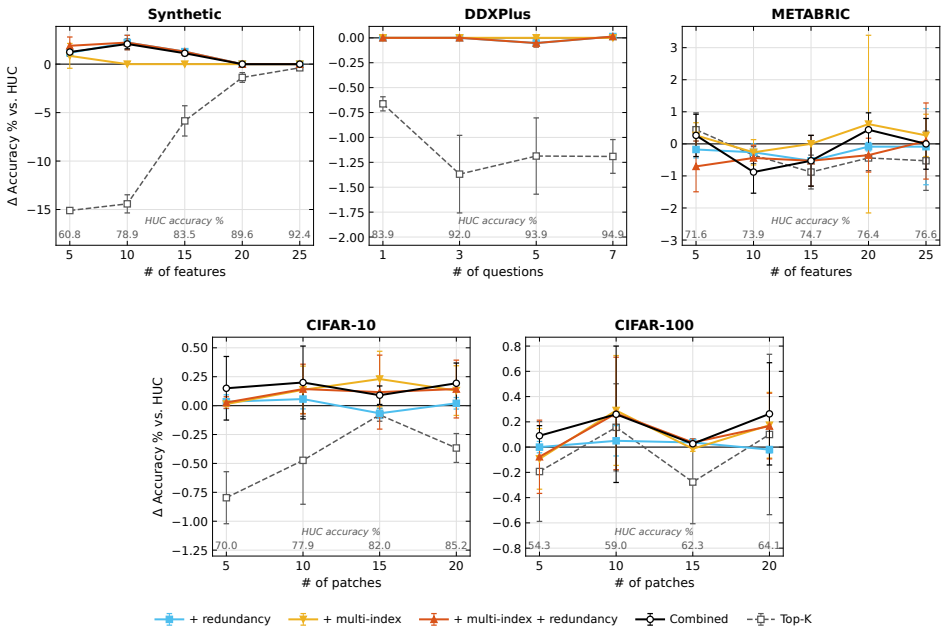


Figure 2: Paired accuracy margin of the extensions over HUC (zero line), with TopK for scale; mean and half range over three seeds. HUC's absolute accuracy is printed along the bottom. The combined variant was not run on DDXPlus.

and two actions with utility vectors $(1, 0)$ and $(0, 1)$ produce complementary, not diminishing, returns.

For the penalty F_λ, write $D(j \mid S) = \sum_{\ell \in S} \rho(j, \ell)$ and define

$$\lambda_{\max} := \min_{\substack{S \subseteq A(M),\ j \notin S \\ D(j|S) > 0}} \frac{\Delta(j \mid S)}{D(j \mid S)}, \tag{8}$$

with the minimum of the empty set defined as $+\infty$.

Proposition 3 (Pairwise redundancy penalty)**.** *For every $\lambda \geq 0$, F_λ is normalized and submodular. If $0 \leq \lambda \leq \lambda_{\max}$, it is also monotone. Let $\varepsilon_\rho = \lambda \max_{S \in \mathcal{F}_B} P(S)$. For $0 \leq \lambda \leq \lambda_{\max}$, on the event of Theorem 1, greedy selection for F_λ under a cardinality budget returns $q_\lambda(x_M)$ satisfying*

$$G(q_\lambda(x_M) \mid x_M) \geq \left(1 - \tfrac{1}{e}\right) G(S^\star \mid x_M) - \left(2 - \tfrac{1}{e}\right)(\varepsilon + \varepsilon_\rho).$$

The same transfer holds for a cost budget when partial enumeration supplies the $(1-1/e)$ surrogate guarantee.

If a positive redundancy marginal coincides with zero coverage gain, then $\lambda_{\max} = 0$. For $\lambda > \lambda_{\max}$, the penalized objective remains submodular but need not be monotone. Under a cost budget, selection stops at the first nonpositive penalized marginal. Under a cardinality budget, the reported implementation instead completes the batch in decreasing posterior-averaged importance so that every method uses the same number of actions. This completion rule is an empirical extension and has no $(1 - 1/e)$ guarantee.

A.6 One-shot batch-selection algorithm

Algorithm 1 gives the cardinality and non-uniform-cost selectors used by the method. Retrieval, weighting, and set selection depend on x_M and the reference set. The algorithm does not observe any value from the selected batch before returning it.

Algorithm 1 One-shot submodular batch selection with partial enumeration

Require: Partial observation x_M, candidate actions $A(M)$, costs $\{c_j\}$, budget B, cardinality b when costs are uniform, reference set $\mathcal{D}$, encoder g, precomputed utilities $\{u_k\}$, envelope ψ, and neighborhood size K

Ensure: A feasible batch S

Retrieve $\mathcal{N}(x_M) \leftarrow \mathrm{NN}_K(g(x_M))$ from $\mathcal{D}$
Compute the weights $\{\alpha_k\}$ using equation 3
Define $F(S) = \sum_k \alpha_k \psi(\sum_{j \in S} u_k(j))$ and marginals $\Delta(j \mid S) = F(S \cup \{j\}) - F(S)$
if the constraint is $|S| \leq b$ **then**
    $S \leftarrow \varnothing$
    while $|S| < b$ and $A(M) \setminus S \neq \varnothing$ **do**
        $j^\star \leftarrow \arg\max_{j \in A(M) \setminus S} \Delta(j \mid S)$
        $S \leftarrow S \cup \{j^\star\}$
    end while
    return S
end if
$S_{\mathrm{best}} \leftarrow \varnothing$
for all $T \subseteq A(M)$ such that $|T| \leq 3$ and $c(T) \leq B$ **do**
    $S \leftarrow T$
    while $\{j \in A(M) \setminus S : c(S) + c_j \leq B\} \neq \varnothing$ **do**
        $j^\star \leftarrow \arg\max_{j \notin S,\ c(S)+c_j \leq B} \Delta(j \mid S)/c_j$
        $S \leftarrow S \cup \{j^\star\}$
    end while
    if $F(S) > F(S_{\mathrm{best}})$ **then**
        $S_{\mathrm{best}} \leftarrow S$
    end if
end for
return S_{best}

Under a cardinality budget, the loop performs at most $b|A(M)|K$ marginal evaluations up to constant factors. Under non-uniform costs, enumeration considers $\sum_{r=0}^{3} \binom{|A(M)|}{r}$ seeds and returns the completion with the largest surrogate value.

A.7 PROOFS

PROOF OF LEMMA 1 (POSTERIOR FIDELITY)

Fix a feasible S. Let $v_0 = g(x_M)$, let $v_k = g(x_M^k)$, and condition on $\mathcal{V} = (v_1, \dots, v_K)$. The weights α_k are then fixed. Define

$$\bar{\mu}_S = \sum_k \alpha_k \mu_{k,S}, \qquad \bar{m}_S = \sum_k \alpha_k m_S(v_k).$$

By assumption (iv), $\bar{\mu}_S$ is the conditional mean of $\widehat{G}(S \mid x_M)$. Insert $\bar{\mu}_S$, $\bar{m}_S$, and $m_S(v_0)$ between $\widehat{G}(S \mid x_M)$ and $G(S \mid x_M)$ to obtain

$$\begin{aligned}\left|\widehat{G}(S \mid x_M) - G(S \mid x_M)\right| \le \left|\widehat{G}(S \mid x_M) - \bar{\mu}_S\right| + \left|\bar{\mu}_S - \bar{m}_S\right| \\ + \left|\bar{m}_S - m_S(v_0)\right| + \left|m_S(v_0) - G(S \mid x_M)\right|. \end{aligned} \tag{9}$$

The imputation-bias assumption and $\sum_k \alpha_k = 1$ give

$$|\bar{\mu}_S - \bar{m}_S| \le \sum_k \alpha_k |\mu_{k,S} - m_S(v_k)| \le \delta_{\mathrm{imp}}.$$

The Lipschitz and radius assumptions give

$$|\bar{m}_S - m_S(v_0)| \le \sum_k \alpha_k L_m \|v_k - v_0\|_2 \le L_m h.$$

Assumption (i) bounds the last term of equation 9 by δ_{repr}.

Conditionally on $\mathcal{V}$, the hybrid gains are independent and each has range width at most Γ. Weighted Hoeffding therefore yields, for every $t > 0$,

$$\Pr\left(\left|\widehat{G}(S \mid x_M) - \bar{\mu}_S\right| \ge t \,\middle|\, \mathcal{V}\right) \le 2\exp\left(-\frac{2t^2}{\Gamma^2 \sum_k \alpha_k^2}\right).$$

Set the right-hand side to η/N_B and take a union bound over the N_B feasible batches. With probability at least $1 - \eta$, simultaneously for every feasible S,

$$\left|\widehat{G}(S \mid x_M) - \bar{\mu}_S\right| \le \Gamma \sqrt{\left(\sum_k \alpha_k^2\right) \tfrac{1}{2} \ln \tfrac{2N_B}{\eta}}.$$

Combining the four terms in equation 9 proves the stated bound. For uniform weights, $\sum_k \alpha_k^2 = 1/K$. Under a cardinality budget $|S| \le b$,

$$N_B = \sum_{r=0}^{b} \binom{|A(M)|}{r} \le (b+1)|A(M)|^b \le (b+1)d^b,$$

which gives $\ln N_B = O(b \ln d)$. □

PROOF OF LEMMA 2 (SATURATION FIDELITY)

Fix a feasible S. Recall $\widehat{G}(S \mid x_M) = \sum_k \alpha_k \gamma_k(S)$ and $F^\star(S \mid x_M) = \sum_k \alpha_k \psi(z_k^\star(S))$, with $\alpha_k \ge 0$ and $\sum_k \alpha_k = 1$. Subtracting and applying the triangle inequality,

$$\left|\widehat{G}(S \mid x_M) - F^\star(S \mid x_M)\right| = \left|\sum_k \alpha_k \big(\gamma_k(S) - \psi(z_k^\star(S))\big)\right| \le \sum_k \alpha_k \left|\gamma_k(S) - \psi(z_k^\star(S))\right|.$$

By assumption $\left|\gamma_k(S) - \psi(z_k^\star(S))\right| \le \eta_\psi$ for every k, so

$$\left|\widehat{G}(S \mid x_M) - F^\star(S \mid x_M)\right| \le \sum_k \alpha_k \eta_\psi = \eta_\psi = \varepsilon_2,$$

using $\sum_k \alpha_k = 1$. Since S was arbitrary, the bound holds for all feasible S. □

PROOF OF PROPOSITION 1 (SUBMODULAR BY CONSTRUCTION)

Write the surrogate as $F(S \mid x_M) = \sum_k \alpha_k \, \psi(z_k(S))$ with $z_k(S) = \sum_{j \in S} u_k(j)$, $u_k(j) \geq 0$, $\alpha_k \geq 0$, and ψ nondecreasing and concave with $\psi(0) = 0$. (The argument applies verbatim to $F^\star$, replacing u_k by $u_k^\star$.) We treat the three properties in turn.

Normalization. Since $z_k(\varnothing) = 0$ for every k, we have $F(\varnothing \mid x_M) = \sum_k \alpha_k \, \psi(0) = 0$.

Monotonicity. For $S \subseteq T \subseteq A(M)$, nonnegativity of u_k gives $z_k(S) = \sum_{j \in S} u_k(j) \leq \sum_{j \in T} u_k(j) = z_k(T)$, and ψ nondecreasing gives $\psi(z_k(S)) \leq \psi(z_k(T))$. A nonnegative weighted sum preserves the inequality, so $F(S \mid x_M) \leq F(T \mid x_M)$.

Submodularity. Fix k and consider $S \mapsto \psi(z_k(S))$. Because z_k is modular, $z_k(S \cup \{j\}) = z_k(S) + u_k(j)$ for $j \notin S$, so the marginal gain of j is

$$\psi\big(z_k(S) + u_k(j)\big) - \psi\big(z_k(S)\big).$$

Take $S \subseteq T \subseteq A(M)$ and $j \notin T$, and set $a = z_k(S)$, $b = z_k(T)$, and $\delta = u_k(j) \geq 0$; monotonicity of z_k gives $a \leq b$. Concavity of ψ makes the increment $x \mapsto \psi(x + \delta) - \psi(x)$ nonincreasing in x, hence

$$\begin{aligned} \psi\big(z_k(S) + u_k(j)\big) - \psi\big(z_k(S)\big) &= \psi(a + \delta) - \psi(a) \\ &\geq \psi(b + \delta) - \psi(b) = \psi\big(z_k(T) + u_k(j)\big) - \psi\big(z_k(T)\big). \end{aligned}$$

which is the diminishing-marginal-returns condition for submodularity. Thus each $S \mapsto \psi(z_k(S))$ is submodular. A nonnegative combination of submodular functions is submodular, so $F(\cdot \mid x_M) = \sum_k \alpha_k \, \psi(z_k(\cdot))$ is submodular. Together with the previous two parts, F is normalized, monotone, and submodular. □

PROOF OF LEMMA 3 (EXPLANATION FIDELITY)

Fix a feasible S. Since $F(S \mid x_M) = \sum_k \alpha_k \, \psi(z_k(S))$ and $F^\star(S \mid x_M) = \sum_k \alpha_k \, \psi(z_k^\star(S))$ with $\alpha_k \geq 0$, $\sum_k \alpha_k = 1$, the triangle inequality gives

$$\big|F(S \mid x_M) - F^\star(S \mid x_M)\big| \leq \sum_k \alpha_k \, \big|\psi(z_k(S)) - \psi(z_k^\star(S))\big|.$$

By L_ψ-Lipschitzness of ψ, $\big|\psi(z_k(S)) - \psi(z_k^\star(S))\big| \leq L_\psi \, \big|z_k(S) - z_k^\star(S)\big|$. Since $z_k(S) = \sum_{j \in S} u_k(j)$ and $z_k^\star(S) = \sum_{j \in S} u_k^\star(j)$,

$$\begin{aligned} \big|z_k(S) - z_k^\star(S)\big| &= \Big|\sum_{j \in S} \big(u_k(j) - u_k^\star(j)\big)\Big| \\ &\leq \sum_{j \in S} \big|u_k(j) - u_k^\star(j)\big| \leq \sum_{j \in A(M)} \big|u_k(j) - u_k^\star(j)\big| \\ &= 2 \, \mathrm{TV}(u_k, u_k^\star) \leq 2\tau. \end{aligned}$$

where the last two steps enlarge the sum from S to $A(M)$ (all terms nonnegative) and use the definition of total variation together with the assumption $\max_k \mathrm{TV}(u_k, u_k^\star) \leq \tau$. Substituting and using $\sum_k \alpha_k = 1$,

$$\big|F(S \mid x_M) - F^\star(S \mid x_M)\big| \leq \sum_k \alpha_k \, L_\psi \, (2\tau) = 2 \, L_\psi \, \tau = \varepsilon_3.$$

For $\psi_a(z) = \log(1 + az)$ we have $\psi_a'(z) = a/(1 + az) \leq a$ for $z \geq 0$, so $L_\psi = a$ and $\varepsilon_3 \leq 2a\tau$. As S was arbitrary, the bound holds for all feasible S. □

Proposition 4 (SHAP versus singleton utility)**.** *Fix a hypothesis k and regard $\gamma_k(S) = \ell(f(x_M), y^k) - \ell(f(x_M \oplus x_S^k), y^k)$ as a cooperative game on $A(M)$, with $\gamma_k(\varnothing) = 0$ and Harsanyi dividends $m_k(T) = \sum_{W \subseteq T} (-1)^{|T| - |W|} \gamma_k(W)$. Let ϕ^k be the Shapley vector of this*

game. Assume that the singleton gains and Shapley values are nonnegative, $\|\phi^k\|_1 > 0$*, and* $s_k := \sum_j \gamma_k(\{j\}) > 0$*. Define*

$$u_k(j) = \frac{\phi_j^k}{\|\phi^k\|_1}, \qquad u_k^{\mathrm{sg}}(j) = \frac{\gamma_k(\{j\})}{s_k}.$$

Then

$$\mathrm{TV}(u_k, u_k^{\mathrm{sg}}) \le \frac{\mathcal{I}(\gamma_k)}{\|\phi^k\|_1} = \frac{\sum_{|T|\ge 2} |m_k(T)|}{\gamma_k(A(M))}, \qquad \mathcal{I}(\gamma_k) := \sum_{T\subseteq A(M):\,|T|\ge 2} |m_k(T)|.$$

Consequently, $\tau_{\mathrm{sg}} := \max_k \mathrm{TV}(u_k, u_k^{\mathrm{sg}})$ *obeys the corresponding maximum of these bounds. If every* γ_k *is additive, then* $\tau_{\mathrm{sg}} = 0$*.*

Proof. The dividend identities give $\gamma_k(\{j\}) = m_k(\{j\})$ and $\phi_j^k = \sum_{T\ni j} m_k(T)/|T|$. Hence

$$\phi_j^k - \gamma_k(\{j\}) = \sum_{T\ni j:\,|T|\ge 2} \frac{m_k(T)}{|T|},$$

and therefore

$$\sum_j |\phi_j^k - \gamma_k(\{j\})| \le \sum_{T:\,|T|\ge 2} |m_k(T)| \sum_{j\in T} \frac{1}{|T|} = \mathcal{I}(\gamma_k).$$

By nonnegativity and Shapley efficiency, $\|\phi^k\|_1 = \sum_j \phi_j^k = \gamma_k(A(M)) > 0$.

For nonnegative vectors a, b with positive ℓ_1 norms,

$$\left\| \tfrac{a}{\|a\|_1} - \tfrac{b}{\|b\|_1} \right\|_1 = \left\| \tfrac{a-b}{\|a\|_1} + b\,\tfrac{\|b\|_1 - \|a\|_1}{\|a\|_1\|b\|_1} \right\|_1 \le \frac{\|a-b\|_1}{\|a\|_1} + \frac{\big|\|b\|_1 - \|a\|_1\big|}{\|a\|_1} \le \frac{2\|a-b\|_1}{\|a\|_1},$$

using $\big|\|b\|_1 - \|a\|_1\big| \le \|a-b\|_1$. Apply this inequality with $a = \phi^k$ and $b = (\gamma_k(\{j\}))_j$, then divide by two. This yields

$$\mathrm{TV}(u_k, u_k^{\mathrm{sg}}) \le \frac{\|\phi^k - (\gamma_k(\{j\}))_j\|_1}{\|\phi^k\|_1} \le \frac{\mathcal{I}(\gamma_k)}{\gamma_k(A(M))}.$$

Taking the maximum over k proves the stated bound. Additivity makes every dividend of order at least two zero, giving $\tau_{\mathrm{sg}} = 0$. □

This proposition analyzes one concrete reference utility. It supplies the τ in Lemma 3 only after setting $u_k^\star = u_k^{\mathrm{sg}}$, establishing the saturation assumption of Lemma 2 for that choice, computing ϕ^k as the Shapley vector of the same loss-reduction game γ_k, and using the identity transform on its nonnegative coordinates. Generic precomputed prediction attributions and nonlinear transforms are not covered unless a separate fidelity bound connects their normalized utilities to $u_k^\star$.

PROOF OF THEOREM 1 (END-TO-END GUARANTEE)

Write $S_g = q(x_M)$ for the batch the policy returns, $S^\star$ for an optimal feasible batch for G, and $\mathcal{F} = \mathcal{F}_B$ for the feasible family. All three fidelity lemmas are invoked on the event of Lemma 1, which has probability at least $1-\eta$ and on which that lemma's bound holds simultaneously for every $S \in \mathcal{F}$; this is the event referred to in the statement.

Step 1: uniform approximation of G by F. The three lemmas bound consecutive links of the chain $G \to \widehat{G} \to F^\star \to F$, each uniformly over $\mathcal{F}$. By the triangle inequality, for every $S \in \mathcal{F}$,

$$\big|G(S \mid x_M) - F(S \mid x_M)\big| \;\le\; \underbrace{\big|G - \widehat{G}\big|}_{\le \varepsilon_1} + \underbrace{\big|\widehat{G} - F^\star\big|}_{\le \varepsilon_2} + \underbrace{\big|F^\star - F\big|}_{\le \varepsilon_3} \;=\; \varepsilon. \tag{10}$$

Two properties of equation 10 carry the whole argument, and it is worth isolating them because nothing else about the three lemmas is used below. It is *two-sided*, so it can be applied in one direction at S_g and in the other at $S^\star$. And it is *uniform* over $\mathcal{F}$, so a single ε serves both, even though S_g is produced by the algorithm and is therefore data-dependent. A pointwise bound would not do: S_g is not fixed before the reference set is drawn, and the union bound over the N_B feasible batches inside ε_1 is precisely what licenses evaluating the bound at a selected batch.

Step 2: greedy is near-optimal for the surrogate. By Proposition 1, $F(\cdot \mid x_M)$ is normalized, monotone and submodular on $2^{A(M)}$. Under a cardinality budget every c_j is equal, so the rule in equation 7 reduces to taking the largest marginal gain, and the classical analysis of greedy maximization of a normalized monotone submodular function subject to a cardinality constraint (Nemhauser et al., 1978) gives

$$F(S_g \mid x_M) \;\geq\; \Big(1 - \tfrac{1}{e}\Big) \max_{S \in \mathcal{F}} F(S \mid x_M). \tag{11}$$

This step concerns F alone and introduces no approximation error. It is also where the construction pays off: because F is submodular by design rather than by an assumption on the data, the hypothesis of equation 11 is met exactly rather than approximately, and the two sources of error stay separated.

Step 3: transfer of optimality. Let $S_F \in \arg\max_{S \in \mathcal{F}} F(S \mid x_M)$. Since $S^\star$ is feasible and S_F maximizes F over $\mathcal{F}$,

$$F(S_F \mid x_M) \;\geq\; F(S^\star \mid x_M). \tag{12}$$

This is the only appearance of the ideal optimum, and it enters through feasibility alone. In particular we never require $S^\star$ to be close to optimal for F, which would be a substantive assumption; we require only that it is a candidate the surrogate's maximizer was free to choose.

Step 4: chaining. Combining the three steps,

$$\begin{aligned}
G(S_g \mid x_M) \;&\geq\; F(S_g \mid x_M) - \varepsilon && \text{by equation 10 at } S_g \\
&\geq\; \Big(1 - \tfrac{1}{e}\Big) F(S_F \mid x_M) - \varepsilon && \text{by equation 11} \\
&\geq\; \Big(1 - \tfrac{1}{e}\Big) F(S^\star \mid x_M) - \varepsilon && \text{by equation 12} \\
&\geq\; \Big(1 - \tfrac{1}{e}\Big) \big(G(S^\star \mid x_M) - \varepsilon\big) - \varepsilon && \text{by equation 10 at } S^\star \\
&=\; \Big(1 - \tfrac{1}{e}\Big) G(S^\star \mid x_M) - \Big(2 - \tfrac{1}{e}\Big)\varepsilon,
\end{aligned}$$

because $\big(1 - \frac{1}{e}\big)\varepsilon + \varepsilon = \big(2 - \frac{1}{e}\big)\varepsilon$. Since every inequality holds on the event of Lemma 1, so does the conclusion, which proves the cardinality case. □

The constant. The factor $2 - \frac{1}{e}$ decomposes as $1 + \big(1 - \frac{1}{e}\big)$: one unit of ε is paid for scoring the returned batch under the wrong objective, and a further $\big(1 - \frac{1}{e}\big)\varepsilon$ for measuring the optimum under the wrong objective. The approximation error therefore enters additively rather than being amplified by the greedy factor, which is what makes the bound useful: halving ε halves the penalty, whereas a multiplicative coupling would leave it governed by the larger of the two error sources.

Cost budgets. Only Step 2 uses the structure of the budget; Steps 1, 3 and 4 use feasibility and equation 10 alone. With non-uniform costs the ratio rule has no constant-factor guarantee by itself, since a cheap item of modest value can be taken early and exhaust the budget before a valuable expensive one becomes reachable, and the ratio of greedy to optimal can be made arbitrarily small. The guarantee is recovered by partial enumeration: take every feasible subset of cardinality at most three as a seed, complete each by the cost-benefit rule, and return the best completion. This attains $\big(1 - \frac{1}{e}\big)$ for a monotone submodular objective under a knapsack constraint (Sviridenko, 2004). Substituting that bound for equation 11 leaves the rest of the chain untouched and yields the second claim of the theorem verbatim. The enumeration adds a factor $O(|A(M)|^3)$ to the per-query cost, which is why we use it only where costs are non-uniform.

The separable multi-index extension. Each map $S \mapsto z_k^{(r)}(S)$ is nonnegative and modular. Proposition 1 applied to $\psi_r \circ z_k^{(r)}$ shows that each channel is normalized, monotone, and submodular. Summing over k and r with nonnegative weights preserves all three properties, which proves Proposition 2. The cardinality-constrained greedy guarantee then follows from the same argument as equation 11. This structural proof does not establish fidelity to G. Such a transfer requires analogues of Lemmas 2 and 3 for the multi-index surrogate.

The pairwise redundancy penalty. The marginal increase of P from adding j to S is $D(j \mid S) = \sum_{\ell \in S} \rho(j, \ell)$. This marginal is nondecreasing as S grows, so P is supermodular and $-\lambda P$ is submodular. Since $P(\varnothing) = 0$, F_λ is normalized. Its marginal is

$$\Delta_\lambda(j \mid S) = \Delta(j \mid S) - \lambda D(j \mid S).$$

If $D(j \mid S) = 0$, this marginal is nonnegative because F is monotone. If $D(j \mid S) > 0$, equation 8 and $\lambda \le \lambda_{\max}$ give $\lambda D(j \mid S) \le \Delta(j \mid S)$. Thus every marginal is nonnegative and F_λ is monotone.

For a cardinality budget b, global monotonicity is stronger than the greedy proof requires. It is enough to have nonnegative marginals for $|S| \le 2(b-1)$ because the proof compares each greedy iterate with its union with an optimal batch. For the cardinality case, the corresponding threshold is

$$\lambda^{(b)}_{\max} := \min_{\substack{|S| \le 2(b-1),\ j \notin S \\ D(j|S) > 0}} \frac{\Delta(j \mid S)}{D(j \mid S)},$$

with the minimum of the empty set defined as $+\infty$. This budget-local threshold is no smaller than the global threshold. Proposition 3 uses the global threshold because it does not depend on the budget and also covers the non-uniform-cost statement.

On the event of Theorem 1, equation 10 gives, for every feasible S,

$$|G(S \mid x_M) - F_\lambda(S \mid x_M)| \le |G(S \mid x_M) - F(S \mid x_M)| + \lambda P(S) \le \varepsilon + \varepsilon_\rho.$$

Repeating Steps 2 through 4 of the theorem proof with F_λ and the uniform error $\varepsilon + \varepsilon_\rho$ proves Proposition 3. Under non-uniform costs, partial enumeration replaces only the surrogate-optimization step, as in the proof of Theorem 1.