EPISTEMIC FORWARD MODELS: LEARNING PREDICTIVE STATES FOR ACTION-CONDITIONED FUTURE DISTINGUISHABILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

A robot choosing its next action needs to predict not only what it may observe, but also which task-state directions that observation will constrain. We study this post-action identifiability geometry as an explicit predictive target. We characterize decision-sufficient representations, showing that orthogonally invariant scalar summaries are insufficient in general, deriving dimension lower bounds for continuous sufficient targets, and establishing that marginal observation prediction need not determine the local information increment. This motivates an epistemic predictive state: a history summary sufficient to determine the distribution of post-action geometry under every candidate action. Our epistemic forward model (EFO) learns a finite-dimensional history representation from which it predicts structured post-action geometry by composing accumulated information inferred from history with an action-conditioned increment. Training uses finite-difference Jacobians of a fixed noise-normalized RGB-D measurement map as offline geometric supervision, whereas inference requires only observed history and candidate actions. In a controlled Habitat layout, candidate actions induce different effective post-action subspaces at 53.7% of 285 decision checkpoints. EFO improves effective-rank Macro-F1 from 0.204 to 0.329 over a no-action predictor. Moreover, among same-rank, near-equal-trace action pairs with distinct subspaces, retaining the full predicted geometry reduces directional regret from 0.222 to 0.083 relative to scalarizing the same predictions. These results establish action-conditioned future identifiability as a predictable representation target whose directional structure can matter for sensing decisions.

1 INTRODUCTION

An embodied agent's actions change what it can perceive: moving around an occlusion or changing viewpoint can reveal different aspects of a target state. Active perception therefore requires anticipating perceptual consequences before the resulting observation is available (Bajcsy, 1988; Placed et al., 2023; Chen et al., 2024). World models predict future observations, latent states, dynamics, or rewards (Wu et al., 2023; Hafner et al., 2025; Assran et al., 2025), but need not explicitly reveal which target-state directions the resulting observation will constrain; one view may resolve depth while another resolves lateral position.

Active sensing instead ranks candidate actions using information gain, uncertainty reduction, reconstruction gain, or estimation quality (Tchuiev & Indelman, 2023; Xue et al., 2024; Chen et al., 2025). Similar scalar scores can conceal different directional sensitivities, while belief-space and active-estimation methods retain richer uncertainty (Nagami & Schwager, 2024; Kitanov & Indelman, 2024; Xue et al., 2026). We ask: **Can interaction history predict, before execution, which task-state directions each candidate action will make identifiable? Our work makes three contributions.**

First, we characterize **what a predictive representation must preserve for directional epistemic decisions**. Theorem 1 shows that a target is decision-sufficient exactly when it determines the relevant geometry kernel; hence orthogonally invariant summaries such as trace, effective rank, and regularized log-determinant are insufficient in general, and continuous sufficient targets obey dimen-

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.

sion lower bounds. The same theorem further separates marginal observation prediction from local information prediction by showing that the former need not determine the information increment. These results motivate an **epistemic predictive state**: a task-relative summary of history sufficient for predicting the action-conditioned distribution of future identifiability geometry.

Second, we propose **epistemic forward models (EFO)**, which predict post-action identifiability geometry from interaction history and a candidate action. We define this geometry through the sensitivity of a fixed RGB-D encoder to local perturbations of a four-dimensional target pose. Effective rank describes how many pose directions receive support, while the identifiable subspace specifies which directions they are; principal angles compare predicted and reference subspaces (Björck & Golub, 1973). At inference time, EFO requires neither the action's future observation nor simulator access.

Third, we establish an **empirical evidence chain** for action-conditioned epistemic prediction within one controlled Habitat layout. We test whether actions change identifiable directions, whether persistent sensing history can disambiguate future geometry hidden from the current view, and whether scalar summaries conceal distinctions relevant to action choice. A feature-predictive RSSM probe provides a predictive-target comparison under a shared geometry-scoring protocol, and controlled sensing changes expose where the observed gains persist and where prediction error increases.

2 Related Work

Active perception and epistemic action selection. Active perception makes sensing part of action selection (Bajcsy, 1988; Chen et al., 2011). Recent systems learn viewpoint policies for reconstruction, localization, and manipulation (Chen et al., 2024; Li et al., 2025; Kerr et al., 2025; Xiong et al., 2025), while others evaluate candidate views through predicted visibility, uncertainty, or information gain (Dhami et al., 2023; Xue et al., 2024; Wilson et al., 2025). Belief-space and active-estimation methods also retain covariances or viewpoint-dependent uncertainty (Placed et al., 2023; Kitanov & Indelman, 2024; Tchuiev & Indelman, 2023; Nagami & Schwager, 2024). EFO focuses on a narrower predictive object: the effective rank and identifiable directions of a continuous target pose under each candidate action.

World models and predictive representations. World models predict future observations, representations, or rewards (Wu et al., 2023; Hafner et al., 2025; Assran et al., 2025); predictive-state and belief representations summarize history under partial observability (Singh et al., 2004; Boots et al., 2011; Igl et al., 2018). Related visual-planning methods also predict what an action may reveal or use predicted object structure to select a view. These lines of work motivate history-conditioned prediction, but EFO makes the future observation's local pose identifiability, rather than its visual content, the explicit target.

Observability, identifiability, and structured information. Active estimation uses covariance or information structure to assess sensing actions (Costante et al., 2016; Tchuiev & Indelman, 2023; Kitanov & Indelman, 2024); mapping methods model directional uncertainty for view selection (Jiang et al., 2025; Wilson et al., 2025; Xue et al., 2026). Here a Jacobian-induced information matrix measures encoder-relative local pose sensitivity, with effective rank and identifiable subspace defining the geometry predicted by EFO.

3 Predicting What an Action Will Make Identifiable

Action selection may depend on which target-state direction a future observation constrains. Marginal forecasts need not reveal this sensitivity, and query-independent scalar summaries can discard its orientation. We therefore ask what action-conditioned object is sufficient for such decisions.

Let $x \in \mathbb{R}^d$ be the task state in fixed normalized coordinates, where $d \geq 2$, and let ξ collect the remaining scene and sensing variables. The state x is fixed within an episode. At step t, the agent has received RGB-D observations $O_{0:t}$ and executed actions $A_{0:t-1}$. Its available history is defined as

$$H_t = (O_{0:t}, A_{0:t-1}). \tag{1}$$

Let $\mathcal{A}_t$ be a finite, nonempty set of candidate actions. A unit vector $g \in \mathbb{R}^d$ specifies the state direction of interest at decision time. We write $\mathbb{S}^d_+$ and $\mathbb{S}^d_{++}$ for the real symmetric positive semidefinite and positive definite $d \times d$ matrices, I_d for the identity matrix, and $\|\cdot\|_{\mathrm{op}}$ for the spectral norm.

Let $Y_k = \psi_k(O_k) \in \mathbb{R}^{m_k}$ be the RGB-D feature vector produced by a preprocessing map ψ_k that is fixed when local state perturbations are compared. For a fixed realized action path, its conditional law under a local Gaussian working model is given by

$$Y_k \mid x, \xi, H_{k-1}, A_{k-1} \sim \mathcal{N}\big(\mu_k(x,\xi), W_k^{-1}\big)\,, \tag{2}$$

where $\mu_k(\cdot,\xi) : \mathbb{R}^d \to \mathbb{R}^{m_k}$ is differentiable at x and $W_k \in \mathbb{S}^{m_k}_{++}$ is independent of x. The history and action terms are omitted at $k = 0$. Observation noise is conditionally independent across steps along the fixed path. The local Jacobian and information contribution are defined as

$$J_k = \nabla_x \mu_k(x,\xi) \in \mathbb{R}^{m_k \times d}, \qquad \Delta\Gamma_k = J_k^\top W_k J_k \in \mathbb{S}^d_+. \tag{3}$$

The current geometry and the geometry following candidate action $a \in \mathcal{A}_t$ are defined as

$$\Gamma_t = \sum_{k=0}^{t} \Delta\Gamma_k, \qquad \Gamma^a_{t+1} = \Gamma_t + \Delta\Gamma^a_{t+1}. \tag{4}$$

Under the working model, Γ^a_{t+1} determines local Hellinger distinguishability along the fixed action path. The experimental teacher constructs positive semidefinite geometry from finite differences of noise-scaled RGB-D features. Appendix B.1 specifies this construction and the scope of its Gaussian interpretation.

Let $\lambda_1(\Gamma) \geq \cdots \geq \lambda_d(\Gamma) \geq 0$ be the ordered eigenvalues of $\Gamma \in \mathbb{S}^d_+$, with corresponding orthonormal eigenvectors $v_1, \ldots, v_d$. For a relative threshold $\epsilon \in (0,1)$, the effective identifiable rank and ambiguous-subspace projector are defined as

$$r_\epsilon(\Gamma) = \sum_{i=1}^{d} \mathbf{1}\{\lambda_i(\Gamma) > \epsilon\lambda_1(\Gamma)\}\,, \qquad P^\epsilon_{\mathrm{A}}(\Gamma) = I_d - \sum_{i=1}^{r_\epsilon(\Gamma)} v_i v_i^\top. \tag{5}$$

For an action a and a unit query direction g, the directional cost is defined as

$$c_g(\Gamma^a_{t+1}) = g^\top K(\Gamma^a_{t+1}) g. \tag{6}$$

We consider the ridge and null-exposure kernels, defined as

$$K_\rho(\Gamma) = (\Gamma + \rho I_d)^{-1}, \qquad K_\epsilon(\Gamma) = P^\epsilon_{\mathrm{A}}(\Gamma), \tag{7}$$

where $\rho > 0$ is shared by all candidate actions at a decision point and ϵ is the threshold in Eq. (5). The costs measure ridge-regularized directional ambiguity and exposure to the ambiguous subspace, respectively. Both kernels satisfy $K(R\Gamma R^\top) = RK(\Gamma)R^\top$ for every orthogonal $R \in \mathbb{R}^{d\times d}$.

Definition 1 (Decision sufficiency)**.** A target $T : \mathbb{S}^d_+ \to \mathcal{T}$, where $\mathcal{T}$ is an arbitrary set, is decision-sufficient for a kernel K if there exists a rule δ such that, for every finite nonempty action set $\mathcal{A}$, every assignment $\Gamma^a \in \mathbb{S}^d_+$ for $a \in \mathcal{A}$, and every unit vector $g \in \mathbb{R}^d$, the selected action satisfies

$$\delta\big(g, (T(\Gamma^a))_{a\in\mathcal{A}}\big) \in \arg\min_{a\in\mathcal{A}} g^\top K(\Gamma^a) g. \tag{8}$$

Theorem 1 (What a directional decision requires)**.** *Let $d \geq 2$ and $K \in \{K_\rho, K_\epsilon\}$.*

(i) A target T is decision-sufficient for K if and only if there exists a map F on $T(\mathbb{S}^d_+)$ such that $K = F \circ T$.

(ii) No orthogonally invariant target satisfying $T(R\Gamma R^\top) = T(\Gamma)$ for every orthogonal R is decision-sufficient. Examples include trace, effective rank, and regularized log-determinant scores.

(iii) If $T : \mathbb{S}^d_+ \to \mathbb{R}^q$ is continuous and decision-sufficient, then $q \geq d(d+1)/2$ for K_ρ and $q \geq \lfloor d^2/4 \rfloor + 1$ for K_ϵ.

(iv) The marginal predictive law $p(Y_{t+1} \mid H_t, A_t = a)$ does not generally determine the local information increment. Linear Gaussian models can have the same marginal predictive law but different information increments with respect to x.

The intuition behind part (ii) is that an orthogonally invariant target cannot distinguish Γ from $R\Gamma R^\top$, whereas the directional decision kernel transforms equivariantly as $K(R\Gamma R^\top) = RK(\Gamma)R^\top$. Thus a rotation can leave the target unchanged while changing the query-dependent costs $g^\top K(\Gamma)g$.

Part (iv) concerns marginal forecasts. A state-conditioned observation model that retains derivatives with respect to x can recover the information increment. Proofs are given in Appendix A.2.

Because (x, ξ) need not be fully known when an action is selected, each candidate action induces a conditional law of post-action geometry. The one-step epistemic predictive state is defined as

$$\mathsf{E}_t = \big(\mathsf{E}_t(a)\big)_{a\in\mathcal{A}_t}, \qquad \mathsf{E}_t(a) = \mathcal{L}\big(\Gamma^a_{t+1} \mid H_t\big)\,, \tag{9}$$

where $\mathcal{L}(\cdot \mid H_t)$ denotes the conditional law under the data-generating process. This state can differ between histories with the same current observation.

Corollary 1 (Risk-neutral decision target)**.** *For each candidate action, the conditional cost matrix is defined as*

$$M^a_t = \mathbb{E}\big[K(\Gamma^a_{t+1}) \mid H_t\big]\,. \tag{10}$$

For every unit query direction g, the expected directional cost is given by

$$\mathbb{E}\big[c_g(\Gamma^a_{t+1}) \mid H_t\big] = g^\top M^a_t g. \tag{11}$$

Thus $(M^a_t)_{a\in\mathcal{A}_t}$ suffices for risk-neutral one-step decisions under Eq. (6).

EFO predicts a geometry matrix $\widehat{\Gamma}^a_{t+1}$ and uses $K(\widehat{\Gamma}^a_{t+1})$ to approximate M^a_t. These matrices need not be equal. The following proposition bounds the resulting decision regret for the ridge kernel.

Proposition 1 (Directional regret from geometry error)**.** *Fix a unit query direction g and define M^a_t in Eq. (10) using $K = K_\rho$. For each $a \in \mathcal{A}_t$, let $\widehat{\Gamma}^a_{t+1} \in \mathbb{S}^d_+$ be measurable with respect to H_t. The predicted and optimal actions are selected as*

$$\widehat{a} \in \arg\min_{a\in\mathcal{A}_t} g^\top K_\rho(\widehat{\Gamma}^a_{t+1})g, \qquad a^\star \in \arg\min_{a\in\mathcal{A}_t} g^\top M^a_t g. \tag{12}$$

Their conditional excess cost satisfies

$$g^\top M^{\widehat{a}}_t g - g^\top M^{a^\star}_t g \le \frac{2}{\rho^2} \max_{a\in\mathcal{A}_t} \mathbb{E}\left[\left\|\widehat{\Gamma}^a_{t+1} - \Gamma^a_{t+1}\right\|_{\mathrm{op}} \mid H_t\right]. \tag{13}$$

For K_ϵ, no bound that vanishes uniformly with matrix error holds at eigenvalue threshold crossings. This motivates direct rank and subspace supervision. Appendix A.2 gives the proof and a threshold counterexample.

4 Epistemic Forward Model

The EFO predicts post-action geometry $\widehat{\Gamma}^a_{t+1}$ from an observed history and a candidate action. Its deterministic output is a point prediction derived from a learned history representation; it does not parameterize the full distribution-valued epistemic predictive state in Eq. (9). Figure 1 shows the prediction pipeline and its training-only teacher path.

4.1 History-Conditioned Prediction

EFO encodes H_t using frozen ResNet-18 RGB features, pooled depth features, a train-fold PCA projection, and a GRU. The history and action-conditioned representations are

$$z^E_t = \phi_\theta(H_t), \qquad u^a_t = f_\theta\big([z^E_t, e_\theta(a)]\big)\,, \quad a \in \mathcal{A}_t, \tag{14}$$

where θ denotes trainable EFO parameters, $e_\theta(a)$ is the embedding of action a, and brackets denote vector concatenation.

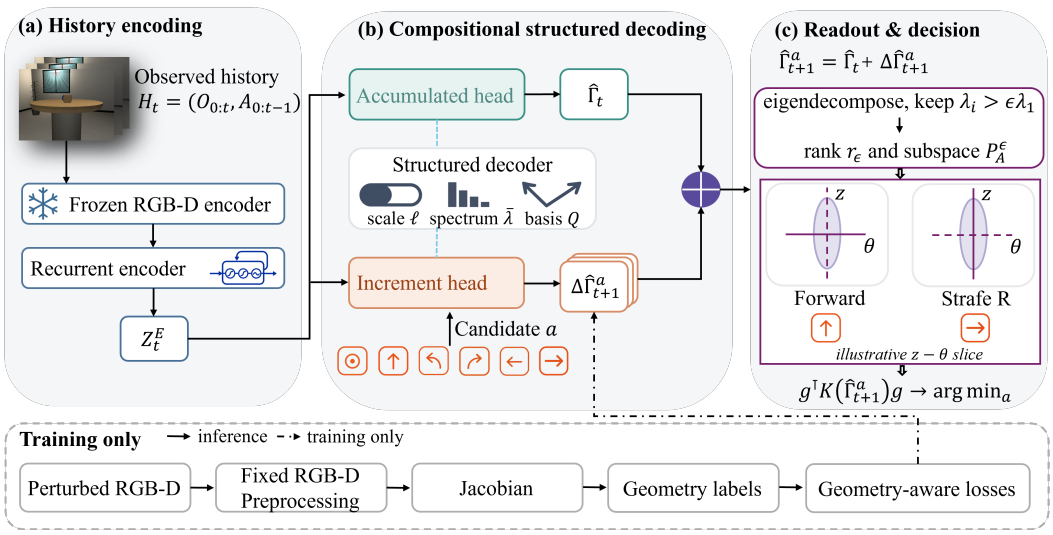


Figure 1: EFO framework. (a) A recurrent encoder summarizes the observed history. (b) Structured heads predict current geometry and an action-conditioned increment, which are composed into post-action geometry. (c) Geometry readouts support directional action selection.

4.2 Compositional Geometry Decoder

Each branch maps an input representation u to an orthogonal matrix $Q_\omega(u) \in \mathbb{R}^{d\times d}$, a positive spectrum $\bar{\lambda}_\omega(u) \in \mathbb{R}^d_{++}$ whose entries sum to one, and a log-trace $\ell_\omega(u) \in \mathbb{R}$. Its predicted geometry is defined as

$$\mathcal{G}_\omega(u) = \exp(\ell_\omega(u))\, Q_\omega(u)\, \mathrm{diag}\big(\bar{\lambda}_\omega(u)\big)\, Q_\omega(u)^\top \in \mathbb{S}^d_{++}. \tag{15}$$

Here ω denotes branch-specific decoder parameters. QR factorization constructs Q_ω, while sorted positive spectral outputs are normalized to construct $\bar{\lambda}_\omega$. The current, incremental, and post-action predictions are defined as

$$\hat{\Gamma}_t = \mathcal{G}_{\omega_{\mathrm{cur}}}(z_t^E), \qquad \widehat{\Delta\Gamma}^a_{t+1} = \mathcal{G}_{\omega_{\mathrm{inc}}}(u_t^a), \qquad \hat{\Gamma}^a_{t+1} = \hat{\Gamma}_t + \widehat{\Delta\Gamma}^a_{t+1}. \tag{16}$$

All matrix, spectrum, rank, and subspace readouts use the same composed prediction; auxiliary rank heads provide training signals only.

4.3 Supervision and Decision Use

At each checkpoint, the teacher supplies current, incremental, and post-action geometries. The corresponding losses compare matrix shape, spectrum, trace scale, effective rank, and identifiable subspace, while auxiliary rank probabilities are aligned with a differentiable spectral activity readout. Appendix B.2 gives the complete definitions and coefficients. The objective is

$$\mathcal{L}_{\mathrm{EFO}} = \mathcal{L}_{\mathrm{cur}} + \mathcal{L}_{\mathrm{inc}} + 2\mathcal{L}_{\mathrm{post}} + 0.75\mathcal{L}_{\mathrm{act}}. \tag{17}$$

The action loss $\mathcal{L}_{\mathrm{act}}$ matches projector distances between candidate actions at the same checkpoint and separates pairs with sufficiently different teacher projectors. Appendix B.2 defines the loss components and their coefficients.

At inference, EFO composes its predicted current geometry and action-dependent increment without teacher matrices. Each resulting $\hat{\Gamma}^a_{t+1}$ can be scored for any unit query direction g using either kernel in Eq. (7). The same post-action prediction supports different query directions without retraining the model.

5 EXPERIMENTS

Experiments test action dependence, hidden-history dependence, scalar collisions, comparison with a feature-predictive RSSM probe (Hafner et al., 2019; 2025), and controlled sensing changes. Appendix C gives the acquisition, teacher, and evaluation protocols.

5.1 ACTION CHOICE CHANGES IDENTIFIABLE DIRECTIONS

At each checkpoint, we branch all six candidate actions from the same saved state. We measure differences between their information increments and between the accumulated post-action geometries used for subsequent decisions. Identifiable subspaces use a relative eigenvalue threshold of $\epsilon = 0.01$. A projector distance of at least 0.25 defines a subspace switch.

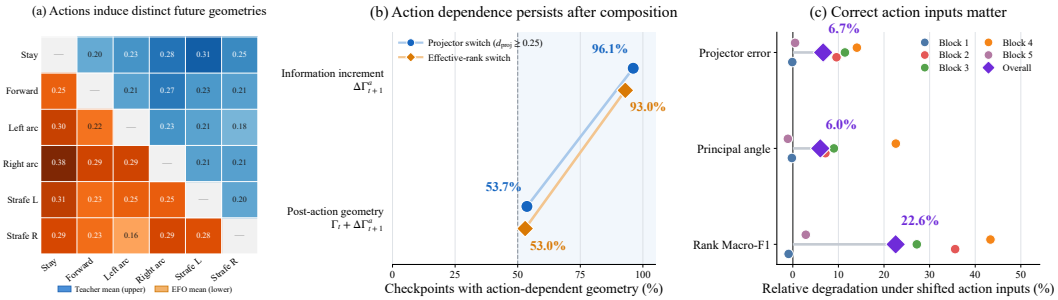


Figure 2: Action-dependent identifiability. (a) Mean pairwise post-action projector distance. (b) Subspace and rank switch rates for increments and post-action geometry. (c) EFO errors with correct and shifted action inputs. $\epsilon = 0.01$; projector-switch threshold 0.25.

Candidate actions change the identifiable subspace of the teacher information increment at 96.1% of checkpoints (Figure 2(b)). Some differences disappear after the increment is combined with accumulated information. Yet the decision-relevant post-action subspace still differs across actions at 53.7% of checkpoints, with distances varying by action pair (Figure 2(a)).

EFO predicts its rank and directional structure more accurately than the tested alternatives on average (Table 1). Its effective-rank Macro-F1 rises from 0.2036 without action input to 0.3294 with action input. The gain is positive in all five held-out blocks, with a block-bootstrap 95% interval of $[0.046, 0.176]$. EFO also has the lowest average projector, principal-angle, and spectrum errors, although its projector gain over the action-free model varies across blocks. The step-and-action mean has a slightly lower Frobenius error but worse directional and rank scores. Cyclically shifting the trained EFO's action inputs increases projector and principal-angle errors by 6.7% and 6.0%, while reducing rank Macro-F1 by 22.6% (Figure 2(c)). Further diagnostics appear in Appendix D.

Table 1: Post-action geometry prediction. Learned results are mean ± standard deviation over three seeds. Macro-F1 uses ranks 1–4. Shifted action denotes cyclically permuted test-time action inputs.

Method	$e_F \downarrow$	$d_{proj} \downarrow$	θ_{max} (°) $\downarrow$	$e_\lambda \downarrow$	Macro-F1↑
Step+action mean	**0.5955**	0.5309	64.80	0.7347	0.1445
EFO w/o action	$0.6114_{\pm.0189}$	$0.4281_{\pm.0252}$	$49.86_{\pm3.15}$	$0.8172_{\pm.0052}$	$0.2036_{\pm.0321}$
EFO (shifted action)	$0.6637_{\pm.0110}$	$0.4299_{\pm.0226}$	$49.36_{\pm3.05}$	$0.7172_{\pm.0365}$	$0.2551_{\pm.0285}$
Unstructured EFO	$0.8884_{\pm.0635}$	$0.5708_{\pm.1051}$	$60.55_{\pm8.89}$	$0.7204_{\pm.0167}$	$0.2741_{\pm.0127}$
EFO	$0.6008_{\pm.0181}$	$\mathbf{0.4031_{\pm.0253}}$	$\mathbf{46.55_{\pm3.61}}$	$\mathbf{0.6500_{\pm.0307}}$	$\mathbf{0.3294_{\pm.0211}}$

5.2 PERSISTENT SENSING HISTORY CAN DISAMBIGUATE FUTURE GEOMETRY

We test whether a persistent sensing condition hidden from the matched current view can make history necessary for predicting action-conditioned geometry. Each pair shares the same current RGB-D observation and six candidate actions, but differs in earlier depth sensing; that sensing condition persists into the candidate-action observations. The experiment therefore isolates whether prior sensing history helps predict a future geometry difference that is not visible in the current observation.

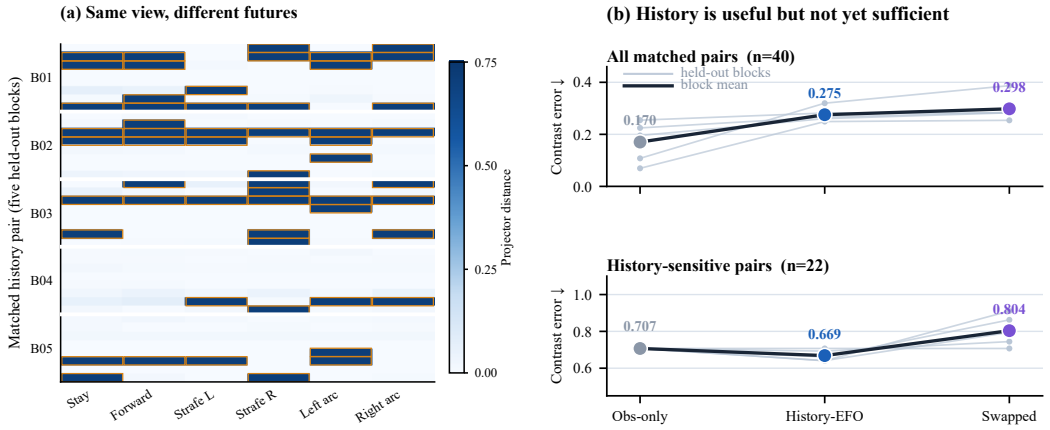


Figure 3: Future geometry under matched current views. (a) Teacher projector distances between paired histories; orange outlines mark ≥ 0.25. (b) Signed projector-contrast error on history-sensitive comparisons. Thin lines: held-out blocks; bold: block means.

Despite identical current views, 55.0% of matched pairs have different teacher post-action identifiable subspaces under at least one shared action (Figure 3(a)).

On teacher-identified history-sensitive comparisons, History-EFO reduces block-mean signed projector-contrast error from 0.707 for the observation-only model to 0.669 (Figure 3(b)). The paired reduction is 0.039, with a descriptive block-bootstrap 95% interval of $[0.015, 0.062]$. Swapping the paired histories raises the error to 0.804. This result concerns the selected comparisons. Across all matched comparisons, History-EFO has higher absolute projector error than the observation-only model. Appendix E reports the pairing and full-set results.

5.3 Directional Structure Matters When Scalar Scores Collide

This experiment is an empirical stress test of Theorem 1(ii). We isolate actions with similar trace and equal effective rank but different identifiable subspaces and ask whether retaining predicted orientation improves directional action selection. For a unit query direction q, residual ambiguity is $\rho_A(q;\Gamma) = q^\top P_A(\Gamma)q$. Null-exposure regret measures the excess teacher ambiguity of a selected action relative to the teacher-optimal action for q.

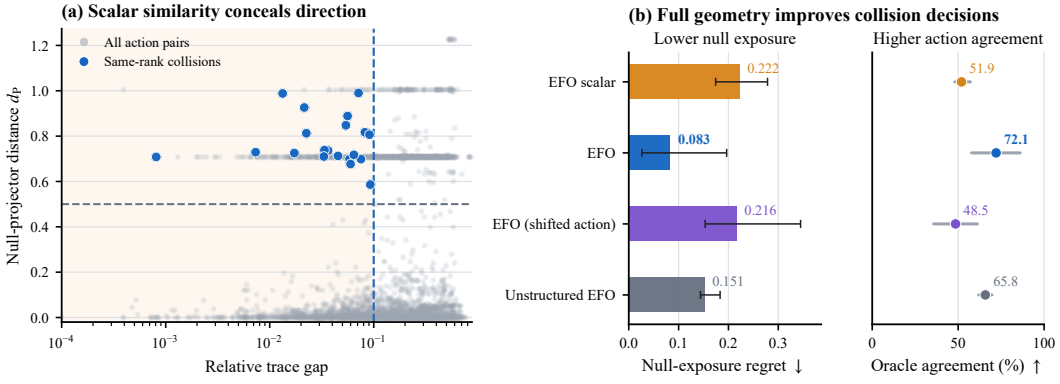


Figure 4: Scalar collisions and directional action selection. (a) Relative trace gap and ambiguous-subspace distance for within-checkpoint action pairs. Dashed lines mark the 10% trace-gap and 0.5 distance thresholds. Blue points mark same-rank collisions. (b) Null-exposure regret and teacher-oracle agreement on the diagnostic set. Error bars show descriptive block-bootstrap 95% intervals.

We isolate this decision problem using teacher-defined pairs with similar trace, equal effective rank, separated ambiguous subspaces, and a crossover in the preferred action across query directions (Figure 4(a)). Both policies use the same held-out EFO predictions. The full-matrix policy scores actions for the given q. The scalar policy uses a spectral rule selected on other blocks and ranks actions without receiving q.

On the diagnostic set, full-matrix selection reduces pooled null-exposure regret from 0.2219 to 0.0831 and raises teacher-oracle agreement from 51.9% to 72.1% (Figure 4(b)). The equally weighted block regret reduction is 0.118, with a descriptive 95% block-bootstrap interval of $[0.066, 0.205]$. Regret and agreement favor the full matrix in each of the three represented blocks.

The result is specific to these teacher-defined collisions. Across all same-rank pairs, the average regret gain is close to zero and agreement does not improve consistently. The experiment identifies a limitation of query-independent spectral rankings. A scalar utility computed from the full matrix for a known q can still express directional preference. Appendix F gives the selection criteria, alternative scores, and broader-population results.

5.4 Direct Geometry Prediction and Feature-Predictive Representation

We compare EFO with a frozen feature-predictive RSSM followed by a geometry probe. Both predict action-conditioned information increments; for decision scoring, each increment is added to the same teacher current geometry. This is a predictive-target comparison, not a parameter-matched end-to-end world-model comparison.

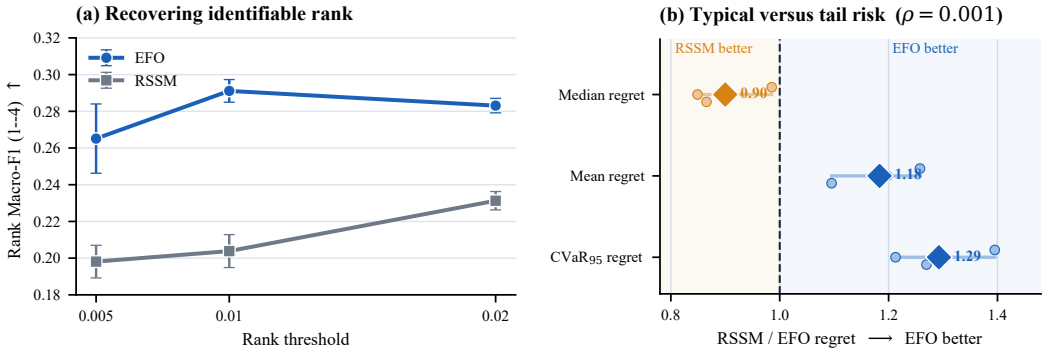


Figure 5: Increment geometry and directional decisions. (a) Rank Macro-F1 across thresholds. (b) RSSM/EFO ratios of median, mean, and CVaR95 normalized regret at $\rho = 10^{-3}$; ratios above one favor EFO.

At the primary rank threshold of 0.01, EFO achieves higher information-increment rank Macro-F1 than the RSSM probe, 0.291 versus 0.204 (Figure 5(a)). It also has lower projector error, 0.534 versus 0.552, and lower padded principal-angle error, $55.4°$ versus $61.3°$. The RSSM probe has lower relative Frobenius error, 0.756 versus 0.792. Both analyses use rank classes 1–4, but these increment results are not directly comparable with the post-action results in Table 1.

We next test whether these geometric differences affect directional decisions. At ridge factor 10^{-3}, EFO reduces mean normalized regret from 3.14 to 2.65 and CVaR95 from 27.49 to 21.28 (Figure 5(b)). The RSSM probe has lower median regret, 0.569 versus 0.633, and higher teacher-oracle agreement, 0.202 versus 0.188. Appendix G reports the full metrics and sensitivity analyses. These results support direct geometry prediction as a useful learning target under the evaluated protocol. They do not establish a general limitation of feature-predictive world models.

5.5 Sensitivity to Controlled Sensing Changes

Finally, we examine how controlled sensing changes affect geometry prediction and whether EFO's advantage over the tested controls persists. Changes to target configuration, camera pose, occlusion, lighting, and depth quality are evaluated against paired baselines. Each shifted branch shares its

candidate action, model, and seed with its baseline branch. We reuse the held-out models without retraining and estimate descriptive uncertainty across acquisition blocks.

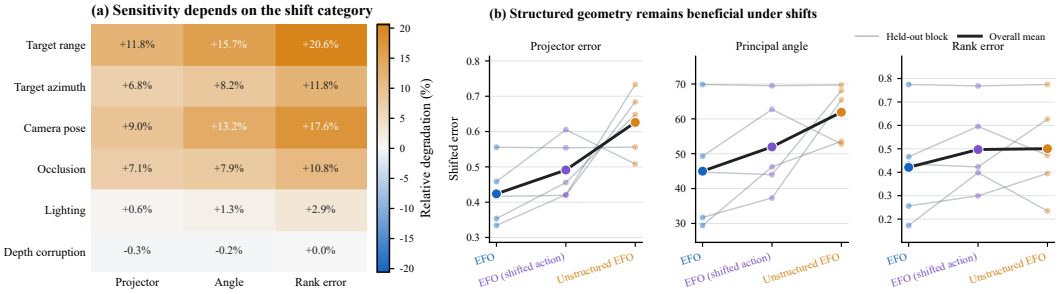


Figure 6: Geometry prediction under controlled sensing changes. (a) Paired changes in EFO error by shift category. (b) Errors on shifted branches for EFO, shifted-action EFO, and Unstructured EFO. Thin lines show held-out blocks, and large markers show block means. Lower is better.

Prediction sensitivity varies by condition (Figure 6(a)). Target-range changes yield the largest average increases in projector, principal-angle, and rank errors, at 11.8%, 15.7%, and 20.6%. Camera-pose changes also increase these errors, while lighting and depth corruption have smaller average effects. Across all paired shifts, projector error rises from 0.398 to 0.424, although its block-bootstrap interval includes zero. Matrix and spectrum errors increase on average with intervals above zero.

EFO nevertheless retains lower block-mean projector error on shifted branches than shifted-action EFO (0.424 versus 0.491) and Unstructured EFO (0.424 versus 0.626), as shown in Figure 6(b). It also has the lowest block-mean principal-angle and rank errors among these methods. Its projector advantage holds in all five blocks against Unstructured EFO and in four against shifted-action EFO. The independently trained action-free model has similar projector error, but EFO has lower rank error in every block (Appendix H). These tests use controlled variations within one scene layout and do not establish transfer to new environments or closed-loop navigation.

6 Discussion and Conclusion

Action-conditioned future identifiability geometry is richer than a query-independent scalar yet narrower than a full world model. On selected matched-view cases, history helps recover future-geometry differences despite higher overall projector error. Relative to an RSSM geometry probe, EFO improves increment rank and mean/tail regret, while RSSM has lower Frobenius error, median regret, and higher oracle agreement. The RSSM study is a predictive-target comparison under a shared teacher-current-geometry scoring condition; it is neither parameter-matched nor an end-to-end comparison, and it does not establish a general limitation of feature-predictive world models. A sufficiently informative world-model belief could support the same epistemic prediction.

The empirical study is limited to one simulated indoor layout and encoder-relative geometry estimated from controlled perturbations. Effective rank depends on a spectral threshold, and subspaces may be unstable near eigenvalue crossings. Cross-scene transfer, cheaper supervision, and closed-loop epistemic planning remain open. Within these limits, the theory and experiments support EFO as a direct task-specific route from interaction history and candidate actions to the future distinctions needed for directional sensing decisions.

AI USE STATEMENT

We used generative AI tools for drafting, formatting, and language refinement. We take responsibility for the final content.

REPRODUCIBILITY STATEMENT

To support reproducibility, Appendix C details Habitat data collection, teacher construction, and evaluation protocols. The other appendices provide proofs, model details, and supplementary analyses.

REFERENCES

Mido Assran, Adrien Bardes, David Fan, Quentin Garrido, Russell Howes, Mojtaba Komeili, Matthew Muckley, Ammar Rizvi, Claire Roberts, Koustuv Sinha, Artem Zholus, Sergio Arnaud, Abha Gejji, Ada Martin, Francois Robert Hogan, Daniel Dugas, Piotr Bojanowski, Vasil Khalidov, Patrick Labatut, Francisco Massa, Marc Szafraniec, Kapil Krishnakumar, Yong Li, Xiaodong Ma, Sarath Chandar, Franziska Meier, Yann LeCun, Michael Rabbat, and Nicolas Ballas. V-JEPA 2: Self-supervised video models enable understanding, prediction and planning, 2025. URL https://arxiv.org/abs/2506.09985.

Ruzena Bajcsy. Active perception. *Proceedings of the IEEE*, 76(8):966–1005, 1988. doi: 10.1109/5.5968.

Åke Björck and Gene H. Golub. Numerical methods for computing angles between linear subspaces. *Mathematics of Computation*, 27(123):579–594, 1973.

Byron Boots, Sajid M Siddiqi, and Geoffrey J Gordon. Closing the learning-planning loop with predictive state representations. *The International Journal of Robotics Research*, 30(7):954–966, 2011.

Liyan Chen, Huangying Zhan, Kevin Chen, Xiangyu Xu, Qingan Yan, Changjiang Cai, and Yi Xu. ActiveGAMER: Active gaussian mapping through efficient rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16486–16497, 2025. URL https://openaccess.thecvf.com/content/CVPR2025/html/Chen_ActiveGAMER_Active_GAussian_Mapping_through_Efficient_Rendering_CVPR_2025_paper.html.

Shengyong Chen, Youfu Li, and Ngai Ming Kwok. Active vision in robotic systems: A survey of recent developments. *The International Journal of Robotics Research*, 30(11):1343–1377, 2011.

Xiao Chen, Quanyi Li, Tai Wang, Tianfan Xue, and Jiangmiao Pang. GenNBV: Generalizable next-best-view policy for active 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16436–16445, 2024. URL https://openaccess.thecvf.com/content/CVPR2024/html/Chen_GenNBV_Generalizable_Next-Best-View_Policy_for_Active_3D_Reconstruction_CVPR_2024_paper.html.

Gabriele Costante, Christian Forster, Jeffrey Delmerico, Paolo Valigi, and Davide Scaramuzza. Perception-aware path planning. *Arxiv*, 2016. doi: 10.48550/arXiv.1605.04151.

Harnaik Dhami, Vishnu D. Sharma, and Pratap Tokekar. Pred-nbv: Prediction-guided next-best-view planning for 3d object reconstruction. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 7149–7154, 2023. doi: 10.1109/IROS55552.2023.10341650.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In Kamalika Chaudhuri and Ruslan Salakhutdinov (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2555–2565, 2019. URL https://proceedings.mlr.press/v97/hafner19a.html.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse control tasks through world models. *Nature*, 640:647–653, 2025. doi: 10.1038/s41586-025-08744-2.

Maximilian Igl, Luisa Zintgraf, Tuan Anh Le, Frank Wood, and Shimon Whiteson. Deep variational reinforcement learning for POMDPs. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2117–2126, 2018. URL https://proceedings.mlr.press/v80/igl18a.html.

Wen Jiang, Boshu Lei, Katrina Ashton, and Kostas Daniilidis. Multimodal LLM guided exploration and active mapping using fisher information. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5392–5404, 2025. URL https://openaccess.thecvf.com/content/ICCV2025/html/Jiang_Multimodal_LLM_Guided_Exploration_and_Active_Mapping_using_Fisher_Information_ICCV_2025_paper.html.

Justin Kerr, Kush Hari, Ethan Weber, Chung Min Kim, Brent Yi, Tyler Bonnen, Ken Goldberg, and Angjoo Kanazawa. Eye, robot: Learning to look to act with a BC-RL perception-action loop. In *Proceedings of the 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 3647–3664, 2025. URL https://proceedings.mlr.press/v305/kerr25a.html.

Andrej Kitanov and Vadim Indelman. Topological belief space planning for active slam with pairwise gaussian potentials and performance guarantees. *The International Journal of Robotics Research*, 43(1):69–97, 2024. doi: 10.1177/02783649231204898.

Jiajie Li, Boyang Sun, Luca Di Giammarino, Hermann Blum, and Marc Pollefeys. Actloc: Learning to localize on the move via active viewpoint selection. In *Proceedings of The 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 1225–1245, 2025. URL https://proceedings.mlr.press/v305/li25b.html.

Keiko Nagami and Mac Schwager. State estimation and belief space planning under epistemic uncertainty for learning-based perception systems. *IEEE Robotics and Automation Letters*, 9(6): 5118–5125, 2024. doi: 10.1109/LRA.2024.3387139.

Julio A. Placed, Jared Strader, Henry Carrillo, Nikolay Atanasov, Vadim Indelman, Luca Carlone, and José A. Castellanos. A survey on active simultaneous localization and mapping: State of the art and new frontiers. *IEEE Transactions on Robotics*, 39(3):1686–1705, 2023. doi: 10.1109/TRO.2023.3248510.

Satinder Singh, Michael R. James, and Matthew R. Rudary. Predictive state representations: a new theory for modeling dynamical systems. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pp. 512–519, 2004. ISBN 0974903906.

Vladimir Tchuiev and Vadim Indelman. Epistemic uncertainty aware semantic localization and mapping for inference and belief space planning. *Artificial Intelligence*, 319: 103903, 2023. URL https://www.sciencedirect.com/science/article/pii/S0004370223000498.

Joey Wilson, Marcelino Almeida, Sachit Mahajan, Martin Labrie, Maani Ghaffari, Omid Ghasemalizadeh, Min Sun, Cheng-Hao Kuo, and Arnab Sen. POp-GS: Next best view in 3d-gaussian splatting with p-optimality. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3646–3655, 2025. URL https://openaccess.thecvf.com/content/CVPR2025/html/Wilson_POp-GS_Next_Best_View_in_3D-Gaussian_Splatting_with_P-Optimality_CVPR_2025_paper.html.

Philipp Wu, Alejandro Escontrela, Danijar Hafner, Pieter Abbeel, and Ken Goldberg. DayDreamer: World models for physical robot learning. In *Proceedings of the 6th Conference on Robot Learning*, volume 205 of *Proceedings of Machine Learning Research*, pp. 2226–2240, 2023. URL https://proceedings.mlr.press/v205/wu23c.html.

Haoyu Xiong, Xiaomeng Xu, Jimmy Wu, Yifan Hou, Jeannette Bohg, and Shuran Song. Vision in action: Learning active perception from human demonstrations. In *Proceedings of the 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 5450–5463, 2025. URL https://proceedings.mlr.press/v305/xiong25a.html.

Shangjie Xue, Jesse Dill, Pranay Mathur, Frank Dellaert, Panagiotis Tsiotra, and Danfei Xu. Neural visibility field for uncertainty-driven active mapping. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18122–18132, 2024. URL https://openaccess.thecvf.com/content/CVPR2024/html/Xue_Neural_Visibility_Field_for_Uncertainty-Driven_Active_Mapping_CVPR_2024_paper.html.

Shangjie Xue, Jesse Dill, Dhruv Ahuja, Frank Dellaert, Panagiotis Tsiotras, and Danfei Xu. Uncertainty-driven 3d gaussian splatting active mapping via anisotropic visibility field. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5014–5026, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Xue_Uncertainty-driven_3D_Gaussian_Splatting_Active_Mapping_via_Anisotropic_Visibility_Field_CVPR_2026_paper.html.

A Identifiability Geometry and Predictive-State Details

A.1 Gaussian Working Model and Local Distinguishability

Fix a realized action path, a task state x, and nuisance variables ξ. At step k, let $Y_k \in \mathbb{R}^{m_k}$ have the local Gaussian density in Eq. (2). Its score with respect to x, holding the realized action path and ξ fixed, is given by

$$s_k = \nabla_x \log p(Y_k \mid x, \xi, H_{k-1}, A_{k-1}) = J_k^\top W_k \big(Y_k - \mu_k(x, \xi)\big). \tag{18}$$

Because the conditional noise has mean zero and covariance W_k^{-1}, the score covariance is given by

$$\mathbb{E}[s_k s_k^\top \mid x, \xi, H_{k-1}, A_{k-1}] = J_k^\top W_k J_k = \Delta\Gamma_k. \tag{19}$$

For $k = 0$, the expectation is conditioned only on (x, ξ). Summing independent conditional contributions along the fixed path yields Eq. (4). This is a local information statement for that path. It does not hold the robot policy fixed while allowing a perturbation in x to induce a different realized action sequence.

For probability measures with densities p and q on the same feature space, define squared Hellinger distance by $H^2(p, q) = 1 - \int \sqrt{pq}$. Let $p^a_{x,\xi}$ denote the joint Gaussian law of the observations on the fixed path ending with candidate action a. For a small displacement $h \in \mathbb{R}^d$, common covariances and conditional independence give

$$\begin{aligned} H^2\big(p^a_{x,\xi}, p^a_{x+h,\xi}\big) &= 1 - \exp\left[-\frac{1}{8}\sum_{k=0}^{t+1} \|\mu_k(x+h,\xi) - \mu_k(x,\xi)\|^2_{W_k}\right] \\ &= \frac{1}{8} h^\top \Gamma^a_{t+1} h + o(\|h\|_2^2), \end{aligned} \tag{20}$$

where $\|v\|^2_{W_k} = v^\top W_k v$, and the final mean is evaluated under a. The last equality uses $\mu_k(x + h, \xi) - \mu_k(x, \xi) = J_k h + o(\|h\|_2)$. It explains the directional meaning of Eq. (4). The finite-difference matrix used in experiments is an operational approximation to these local derivatives. The approximation need not equal the Fisher information of a globally valid pixel likelihood when rendering, cropping, or depth-validity masks change discontinuously.

A.2 Proof of Decision Sufficiency

Proof of Theorem 1. For part (i), suppose first that $K = F \circ T$. The rule can evaluate $g^\top F(T(\Gamma^a))g$ for each action and choose a minimizer, so T is decision-sufficient. Conversely, suppose there are $G_1, G_2 \in \mathbb{S}^d_+$ such that $T(G_1) = T(G_2)$ but $K(G_1) \neq K(G_2)$. Since $K(G_1) - K(G_2)$ is symmetric and nonzero, some unit vector g has $g^\top K(G_1)g \neq g^\top K(G_2)g$. Consider a two-action instance with the ordered matrices (G_1, G_2) and another instance with (G_2, G_1). Both present the identical ordered targets $(T(G_1), T(G_1))$ to δ. The unique minimizing index changes between the instances. No one rule can be correct in both. Consequently, $T(G_1) = T(G_2)$ implies $K(G_1) = K(G_2)$, and $F(T(G)) = K(G)$ is well defined on the image of T.

For part (ii), let $G = \mathrm{diag}(1, 0, \ldots, 0)$ and let R exchange the first and second coordinate axes. Orthogonal invariance gives $T(G) = T(RGR^\top)$. For the ridge kernel, the two kernel matrices differ because their first two diagonal entries are interchanged. For the null-exposure kernel, the two ambiguous-subspace projectors also differ, because $\epsilon \in (0, 1)$. Part (i) then excludes decision sufficiency. Trace, effective rank, and $\log\det(G + \rho I_d)$ are orthogonally invariant examples.

For the ridge claim in part (iii), the map $G \mapsto (G + \rho I_d)^{-1}$ is injective on $\mathbb{S}^d_+$. Part (i) therefore forces T to be injective. The interior $\mathbb{S}^d_{++}$ is an open subset of the $D = d(d+1)/2$ dimensional vector space of real symmetric matrices. A continuous injective map from an open subset of $\mathbb{R}^D$ into $\mathbb{R}^q$ requires $q \geq D$ by invariance of dimension.

For the null-exposure claim, fix an integer k with $1 \leq k \leq d-1$. For each rank-k orthogonal projector P, put $G_P = I_d - P$. Its largest eigenvalue equals one, and $K_\epsilon(G_P) = P$. Part (i) makes $P \mapsto T(G_P)$ continuous and injective on the real Grassmann manifold $\mathrm{Gr}(k, d)$, represented by its orthogonal projectors. This manifold is compact, has no boundary, and has dimension $n = k(d-k)$. If $q \leq n$, composing an alleged embedding into $\mathbb{R}^q$ with the standard inclusion $\mathbb{R}^q \hookrightarrow \mathbb{R}^n$ would give a continuous injection of this compact n-manifold into $\mathbb{R}^n$. Invariance of domain would make its image open. Compactness would make that image compact, which is impossible for a nonempty open subset of $\mathbb{R}^n$. Hence $q \geq k(d-k) + 1$. Maximizing over k gives $q \geq \lfloor d^2/4 \rfloor + 1$.

For part (iv), let the conditional distribution of x given a fixed history be $\mathcal{N}(0, I_d)$. Choose any invertible diagonal matrix B with unequal positive diagonal entries and any orthogonal $Q \in \mathbb{R}^{d\times d}$. Consider two possible state-conditioned observation models, $Y \mid x \sim \mathcal{N}(Bx, I_d)$ and $Y \mid x \sim \mathcal{N}(BQx, I_d)$. Integrating over the same conditional distribution of x gives the same marginal observation law $\mathcal{N}(0, BB^\top + I_d)$ for both models. Their information increments with respect to the same coordinates x are $B^\top B$ and $Q^\top B^\top BQ$. For rotations that change the eigendirections of $B^\top B$, these increments differ. Some Q may leave the increment unchanged, which does not affect the counterexample. The claim concerns the marginal law after integrating out x. If the entire family of conditional laws indexed by x is known, its derivatives also determine the increments. □

Proof of Corollary 1. For any fixed history and unit g, conditional linearity of expectation gives $\mathbb{E}[g^\top K(\Gamma^a_{t+1})g \mid H_t] = g^\top \mathbb{E}[K(\Gamma^a_{t+1}) \mid H_t]g$. The assumed integrability makes every entry of the conditional matrix expectation well defined. Comparing these quadratic forms over the finite candidate set yields the stated decision rule. □

Proof of Proposition 1. Write $q_a = g^\top M^a_t g$ and $\widehat{q}_a = g^\top K_\rho(\widehat{\Gamma}^a_{t+1})g$. Because $\widehat{q}_{\widehat{a}} \leq \widehat{q}_{a^\star}$, the excess cost obeys $q_{\widehat{a}} - q_{a^\star} \leq 2\max_a |q_a - \widehat{q}_a|$. For positive semidefinite matrices G and $\widehat{G}$, the resolvent identity and $\|(G + \rho I_d)^{-1}\|_{\mathrm{op}} \leq \rho^{-1}$ give

$$\left\| K_\rho(\widehat{G}) - K_\rho(G) \right\|_{\mathrm{op}} \leq \rho^{-2} \|\widehat{G} - G\|_{\mathrm{op}}. \tag{21}$$

Since $\widehat{\Gamma}^a_{t+1}$ is fixed given (H_t, a), conditional Jensen's inequality and Eq. (21) imply

$$|q_a - \widehat{q}_a| \leq \left\| \mathbb{E}\Big[K_\rho(\Gamma^a_{t+1}) - K_\rho(\widehat{\Gamma}^a_{t+1}) \mid H_t \Big] \right\|_{\mathrm{op}} \leq \rho^{-2} \mathbb{E}\Big[\|\Gamma^a_{t+1} - \widehat{\Gamma}^a_{t+1}\|_{\mathrm{op}} \mid H_t \Big]. \tag{22}$$

Taking the maximum over actions proves Eq. (13). □

The projector kernel is discontinuous at its relative threshold. For example, let $d = 2$, fix $\epsilon \in (0, 1)$, and choose $\eta > 0$ small enough that $\epsilon + \eta < 1$. The matrices $G_- = \mathrm{diag}(1, \epsilon - \eta)$ and $G_+ = \mathrm{diag}(1, \epsilon + \eta)$ differ by 2η in spectral norm. Their ambiguous projectors are $\mathrm{diag}(0, 1)$ and 0. Thus matrix error can tend to zero while null exposure changes by one for $g = (0, 1)^\top$. An eigengap relative to the threshold is needed for a local projector-error guarantee.

A.3 History Dependence of the Predictive State

Here is an explicit model in which the last observation does not determine the future geometry. Let $\xi \in \{0, 1\}$ be a persistent sensing condition with prior probabilities $1/2$. Before the decision, observe $U \mid \xi \sim \mathcal{N}(\xi, 1)$ and then a current observation $V \sim \mathcal{N}(0, 1)$ independent of (U, ξ, x). The

two histories $H = (U = u_1, V = v)$ and $H' = (U = u_2, V = v)$ have the same last observation v. Bayes' rule gives

$$\Pr(\xi = 1 \mid U = u, V = v) = \frac{1}{1 + \exp[-(u - 1/2)]}. \tag{23}$$

Let each candidate action a produce $Y^a \mid x, \xi \sim \mathcal{N}(b_{a,\xi} x, 1)$ for a scalar state x. Earlier observations contain no information about x, while the action increment equals $b_{a,\xi}^2$. Choose two actions a and b with $(b_{a,0}, b_{a,1}) = (2, 1)$ and $(b_{b,0}, b_{b,1}) = (1, 2)$. Their expected ridge costs depend on the posterior probability in Eq. (23). The preferred action is a when this probability is below $1/2$ and b when it is above $1/2$. Histories with $u_1 < 1/2 < u_2$ therefore produce different predictive-state families and opposite optimal actions despite the same current observation. The scalar example embeds in any larger state dimension by making the remaining coordinates independent of the observation.

B Teacher Construction and EFO Training Details

B.1 Finite-Difference RGB-D Teacher

The implemented teacher is constructed from fixed RGB-D preprocessing, not from the trainable EFO recurrent model or its frozen ResNet-18 input backbone. Let the physical target pose be $p = (p_1, p_2, p_3, p_4)$, with the first three coordinates in meters and p_4 in degrees. Define normalized task coordinates by $x_j = (p_j - p_{j,0})/s_j$, where $p_{j,0}$ is the nominal pose coordinate and the fixed scales are $(s_1, s_2, s_3, s_4) = (0.01\,\mathrm{m}, 0.01\,\mathrm{m}, 0.01\,\mathrm{m}, 10^\circ)$. The central perturbation magnitudes are $(h_1, h_2, h_3, h_4) = (0.001\,\mathrm{m}, 0.001\,\mathrm{m}, 0.001\,\mathrm{m}, 1^\circ)$.

For each action branch and each pose coordinate j, the simulator renders observations at $p + h_j e_j$ and $p - h_j e_j$, where e_j is the jth coordinate vector. All eight perturbations of a branch share its nominal action-conditioned region of interest. Let $r_j^+, r_j^- \in [0,1]^{n_R}$ be the flattened resized RGB arrays, let $d_j^+, d_j^- \in \mathbb{R}^{n_D}$ be the flattened resized depth arrays in meters, and let $v_j \in \{0,1\}^{n_D}$ mark depth pixels valid in both perturbed observations. The image size is 64×64, so $n_R = 3 \cdot 64^2$ and $n_D = 64^2$. RGB and depth use bilinear resizing. The validity mask uses nearest-neighbor resizing. The jth numerical Jacobian column is defined as

$$\widehat{J}_j = \frac{s_j}{2h_j} \begin{bmatrix} \dfrac{w_R (r_j^+ - r_j^-)}{\sigma_R \sqrt{n_R}} \\ \dfrac{w_D\, v_j \odot (d_j^+ - d_j^-)}{\sigma_D \sqrt{n_D}} \end{bmatrix}, \tag{24}$$

where $\odot$ denotes entrywise multiplication, $(w_R, w_D) = (1, 1)$, $\sigma_R = 2/255$ in RGB units, and $\sigma_D = 0.01\,\mathrm{m}$. Invalid depth pixels contribute zero to that coordinate's column. The block sizes n_R and n_D stay fixed as validity changes. With $\widehat{J} = [\widehat{J}_1, \ldots, \widehat{J}_4]$, the branch teacher is defined as

$$\Delta\Gamma_{t+1}^a = \widehat{J}^\top \widehat{J}, \qquad \Gamma_{t+1}^a = \Gamma_t + \Delta\Gamma_{t+1}^a. \tag{25}$$

The executed observation path supplies the accumulated Γ_t, including its initial observation. Candidate branches at a checkpoint share this same matrix. Every branch matrix is positive semidefinite by construction. If one common fixed preprocessing map is differentiable in x and its normalized feature noise is locally Gaussian with identity covariance, Eq. (25) approximates the Gaussian information in Eq. (3). Per-coordinate validity intersections and rendering discontinuities need not satisfy that differentiability condition. The experimental target is therefore the stated Jacobian-induced RGB-D geometry, whose operational definition does not rely on a globally exact likelihood.

Numerical evaluation symmetrizes each teacher matrix and clips roundoff-scale negative eigenvalues to zero. It declares effective rank zero when the largest eigenvalue is at most 10^{-15}. Otherwise, it counts eigenvalues whose ratio to the largest is strictly greater than $\epsilon = 0.01$. These implementation tolerances do not change the abstract definition in Eq. (5).

B.2 Decoder Readouts and Component Losses

Here $d = 4$. The history input uses frozen ResNet-18 RGB features concatenated with pooled depth features. A 32-dimensional principal component projection is fitted on the training blocks of each

fold. The recurrent hidden dimension is 64, and the action embedding has dimension 16. The current and incremental structured branches each form a 4×4 orthogonal basis by QR factorization of a learned matrix. Their four spectral outputs are passed through a positive softplus transform, sorted, and normalized. Each branch's predicted log-trace is clipped to $[-20, 20]$ before exponentiation. Three auxiliary heads produce rank logits for the current, incremental, and composed matrices. These heads do not alter the matrices used for evaluation.

Let $b \in \{\mathrm{cur}, \mathrm{inc}, \mathrm{post}\}$ index the current, incremental, and post-action branches. Let $X^b \in \mathbb{S}^4_+$ and $\widehat{X}^b \in \mathbb{S}^4_+$ denote their teacher and predicted matrices. For a matrix X, define the stabilized trace $\tau(X) = \max\{\mathrm{tr}(X), 10^{-12}\}$ and the shape $S(X) = X/\tau(X)$. Define the log-trace $\ell(X) = \log \tau(X)$ and the normalized ordered spectrum $\bar{\lambda}_i(X) = \lambda_i(X)/\tau(X)$. A hat denotes the corresponding prediction. Define $\mathrm{MSE}(U, V)$ as the arithmetic mean of squared entrywise differences. Define $\mathrm{SL1}(u, v)$ as the arithmetic mean of $h(u_i - v_i)$ over entries. The function $h(z)$ equals $z^2/2$ for $|z| < 1$ and $|z| - 1/2$ otherwise. The matrix-shape and trace components are defined as

$$L^b_{\mathrm{shape}} = \mathrm{MSE}\Big(S(\widehat{X}^b), S(X^b)\Big), \qquad L^b_{\mathrm{trace}} = \mathrm{SL1}\Big(\ell(\widehat{X}^b), \ell(X^b)\Big). \tag{26}$$

The spectrum component uses the teacher effective rank $r^b = r_\epsilon(X^b)$. The stabilized relative eigenvalues are defined as

$$\begin{aligned} u^b_i &= \max\left\{ \frac{\bar{\lambda}_i(X^b)}{\max\{\bar{\lambda}_1(X^b), 10^{-8}\}}, 10^{-8} \right\}, \\ \widehat{u}^b_i &= \max\left\{ \frac{\bar{\lambda}_i(\widehat{X}^b)}{\max\{\bar{\lambda}_1(\widehat{X}^b), 10^{-8}\}}, 10^{-8} \right\}. \end{aligned} \tag{27}$$

Let $\mathrm{mean}_{i \in I}$ denote an arithmetic mean over an index set I, with value zero if I is empty. The rank-aware spectral loss is defined as

$$L^b_{\mathrm{spec}} = \mathrm{mean}_{i \le r^b}\big(\log \widehat{u}^b_i - \log u^b_i\big)^2 + \frac{1}{4}\,\mathrm{mean}_{i > r^b}\big[\log \widehat{u}^b_i - \log(\epsilon/2)\big]^2_+, \tag{28}$$

where $[z]_+ = \max(z, 0)$ and $\epsilon = 0.01$. The first term matches identifiable relative eigenvalues. The second penalizes excessive predicted energy in teacher-inactive modes.

Let $\widehat{X}^b = \sum_{i=1}^4 \widehat{\lambda}^b_i \widehat{v}^b_i (\widehat{v}^b_i)^\top$ be its eigendecomposition in descending order. The differentiable spectral activity and projector used in training are defined as

$$\widehat{\alpha}^b_i = \sigma\left(\frac{\log \max\{\widehat{\lambda}^b_i / \max(\widehat{\lambda}^b_1, 10^{-8}), 10^{-8}\} - \log \epsilon}{0.5} \right), \qquad \widetilde{P}^b_{\mathrm{I}} = \sum_{i=1}^4 \widehat{\alpha}^b_i \widehat{v}^b_i (\widehat{v}^b_i)^\top, \tag{29}$$

where $\sigma(z) = (1 + e^{-z})^{-1}$. The teacher identifiable projector is $P^b_{\mathrm{I}} = I_4 - P^\epsilon_{\mathrm{A}}(X^b)$. Projector supervision is defined as

$$L^b_{\mathrm{proj}} = \mathrm{MSE}\Big(\widetilde{P}^b_{\mathrm{I}}, P^b_{\mathrm{I}}\Big). \tag{30}$$

Let z^b_r be the auxiliary rank logit for rank $r \in \{0, 1, 2, 3, 4\}$, and let $\pi^b_r = \exp(z^b_r)/\sum_{j=0}^4 \exp(z^b_j)$ be its probability. For rank class counts n^b_r in the training fold, class weights have $w^b_0 = 0$ and, for $r \ge 1$, start from $\max(n^b_r, 1)^{-1/2}$. The latter four weights are divided by their mean among classes present in training and clipped to $[0.25, 4]$. Define label-smoothed weighted cross-entropy by

$$\mathrm{CE}_{w,\alpha}(\pi, r) = -\sum_{j=0}^4 w_j \left((1 - \alpha)\mathbf{1}\{j = r\} + \frac{\alpha}{5} \right) \log \pi_j, \qquad \alpha = 0.02. \tag{31}$$

For a minibatch $\mathcal{B}$ of branch examples and focal exponent $\gamma = 1$, the focal rank component is defined as

$$L^b_{\mathrm{rank}} = \frac{1}{|\mathcal{B}|} \sum_{i \in \mathcal{B}} \Big(1 - \pi^b_{r^b_i, i}\Big)^\gamma \mathrm{CE}_{w^b,\alpha}\big(\pi^b_i, r^b_i\big). \tag{32}$$

The rank-zero weight convention is part of this auxiliary loss. Shape, trace, spectrum, and projector terms still supervise zero-rank teacher examples.

For $c \in \{2,3,4\}$, let $q^b_{c,i} = \sum_{r=c}^4 \pi^b_{r,i}$ and $y^b_{c,i} = \mathbf{1}\{r^b_i \geq c\}$. Let $\beta^b_i = w^b_{r^b_i} / \max\{|\mathcal{B}|^{-1} \sum_{j\in\mathcal{B}} w^b_{r^b_j}, 10^{-6}\}$. With $\mathrm{BCE}(q,y) = -y \log q - (1-y)\log(1-q)$, evaluated after clipping q to $[10^{-6}, 1-10^{-6}]$, ordinal supervision is defined as

$$L^b_{\mathrm{ord}} = \frac{1}{|\mathcal{B}|} \sum_{i\in\mathcal{B}} \frac{\beta^b_i}{3} \sum_{c=2}^4 \mathrm{BCE}\big(q^b_{c,i}, y^b_{c,i}\big)\,. \tag{33}$$

The auxiliary probability that mode j is active is $a^b_{j,i} = \sum_{r=j}^4 \pi^b_{r,i}$. Spectral and auxiliary rank readouts are aligned through

$$L^b_{\mathrm{cons}} = \frac{1}{4|\mathcal{B}|} \sum_{i\in\mathcal{B}} \sum_{j=1}^4 \big(\widehat{\alpha}^b_{j,i} - a^b_{j,i}\big)^2\,. \tag{34}$$

The branch loss, averaged over its minibatch, is defined as

$$\mathcal{L}_b = 2L^b_{\mathrm{shape}} + 0.35L^b_{\mathrm{spec}} + 0.50L^b_{\mathrm{trace}} + L^b_{\mathrm{rank}} + 0.25L^b_{\mathrm{ord}} + 1.50L^b_{\mathrm{proj}} + 0.25L^b_{\mathrm{cons}}. \tag{35}$$

The same component coefficients apply to the current, incremental, and post-action branches. Equation equation 17 then assigns the branch coefficients 1, 1, and 2.

B.3 Within-Checkpoint Action Loss and Optimization

Consider a minibatch containing checkpoint groups t with at least two candidate actions. For $b \in \{\mathrm{inc}, \mathrm{post}\}$ and an unordered pair (a, a') in group t, let the teacher and predicted projector distances be

$$d^b_t(a,a') = \frac{\|P^{b,a}_{\mathrm{I},t} - P^{b,a'}_{\mathrm{I},t}\|_{\mathrm{F}}}{\sqrt{2}}, \qquad \widehat{d}^b_t(a,a') = \frac{\|\widetilde{P}^{b,a}_{\mathrm{I},t} - \widetilde{P}^{b,a'}_{\mathrm{I},t}\|_{\mathrm{F}}}{\sqrt{2}}. \tag{36}$$

Let $\mathcal{U}_t$ denote all unordered action pairs in that group, and let $\mathcal{C}^b_t = \{(a,a') \in \mathcal{U}_t : d^b_t(a,a') \geq \delta\}$ with $\delta = 0.25$. Define $\mathrm{avg}_S f = |S|^{-1} \sum_{s\in S} f(s)$ when S is nonempty and zero otherwise. The separation margin is $m = 0.15$. The checkpoint contribution is defined as

$$L^b_t = \mathrm{avg}_{(a,a')\in\mathcal{U}_t} \mathrm{SL1}\Big(\widehat{d}^b_t(a,a'), d^b_t(a,a')\Big) + \mathrm{avg}_{(a,a')\in\mathcal{C}^b_t} \Big[m - \widehat{d}^b_t(a,a')\Big]^2_+\,. \tag{37}$$

The action loss is defined as

$$\mathcal{L}_{\mathrm{act}} = \frac{1}{2} \sum_{b\in\{\mathrm{inc},\mathrm{post}\}} \mathrm{avg}_t\, L^b_t. \tag{38}$$

The second average is taken over switching pairs only. It contributes zero for a checkpoint without such pairs. This matches the training implementation and avoids changing the hinge weight when the number of nonswitching pairs varies.

The primary run uses five held-out acquisition folds. In each fold, three blocks train the model, one selects the epoch and fits no model parameter, and one is held out for testing. The training schedule allows at most 150 epochs with patience 20, AdamW learning rate 3×10^{-4}, and weight decay 10^{-4}. Results average three seeds. No teacher matrix, target perturbation image, or simulator state enters standalone EFO inference. The teacher current geometry used as a common scoring condition in the RSSM comparison is confined to that controlled evaluation protocol.

C Experimental Setup

C.1 Habitat Scene and Counterfactual Acquisition

We use one fixed Habitat indoor layout containing a target object and controllable occluders. The robot observes RGB-D images and chooses among six sensing actions: stay, move forward, left arc, right arc, strafe left, and strafe right. At each decision checkpoint, all six actions branch independently from the same saved state and share the same pre-action history.

The benchmark contains five acquisition blocks. Each block has one baseline trial and 18 one-factor variants. The variants change nine factors at two prescribed settings each: target azimuth,

target range, occluder tangential position, occluder radial position, occluder yaw, lighting, camera lateral position, camera yaw, and depth missingness. Each trial contains three decision checkpoints connected by two executed actions. In total, 95 trials yield 285 checkpoints and 1,710 counterfactual action branches. Figure 7 shows representative nominal observations.

C.2 Teacher Geometry and Evaluation Splits

The target pose is (x, y, z, θ). For each action branch, we collect one nominal future RGB-D observation and eight additional observations. These perturb one pose coordinate at a time in both directions. Position perturbations are $\pm 0.001\,\mathrm{m}$, and orientation perturbations are $\pm 1^\circ$. The robot and camera remain fixed during these perturbations. Symmetric finite differences of the fixed noise-normalized RGB-D measurement map provide Jacobians. We use them to construct positive semidefinite information increments and accumulated post-action geometry, as defined in Sections 3 and 4. Perturbed observations and teacher matrices are used for offline supervision. **Standalone EFO inference uses nominal observed history and a candidate action.**

We evaluate five held-out acquisition blocks. In each fold, three blocks provide 57 training trials, one provides 19 validation trials, and one provides 19 test trials. Every branch of a trial remains in the same split. Data-fitted feature reductions use training data only. We aggregate results across held-out blocks and report seed variation for retrained models. Multiple directional queries at one checkpoint share geometry targets and are not independent trials.

C.3 Matched-History Benchmark

The matched-history analysis uses a separate paired benchmark based on the baseline configurations. Each block contributes eight pairs, giving 40 pairs and 80 trials. Within a pair, the robot trajectory, scene configuration, executed actions, and final current RGB-D observation are identical. One history retains stable depth sensing. The other applies 15% depth dropout to earlier observations and candidate-action future observations while leaving the final current observation unchanged. Within each future branch, nominal and pose-perturbed observations use the same dropout mask.

Each trial has three checkpoints and six candidate branches per checkpoint, giving 1,440 branches. We evaluate the final checkpoint, which yields 240 matched pair–action comparisons. **The benchmark tests whether history helps predict a persistent sensing condition and its future geometry.** Because the condition also persists into future observations, the benchmark does not isolate a memory effect from an independently randomized future condition.

C.4 Comparison Protocols and Feature-Predictive Baseline

The action-dependence analysis compares EFO with a separately trained predictor without action input. A shifted-action control instead holds the trained EFO fixed and cyclically replaces candidate-action inputs at evaluation. The scalar decision analysis applies full-matrix and scalar action rules to the same held-out EFO predictions. The matched-history analysis compares history-conditioned and current-observation predictors. Other analyses examine a feature-predictive baseline and controlled sensing changes.

For the feature-predictive comparison, we train an RSSM. This baseline probes a feature-predictive representation. It does not reproduce the full DreamerV3 system. The RSSM uses nominal RGB-D histories and actions, ResNet18 RGB features, pooled depth features, and a 32-dimensional PCA projection fitted on the training split. It has a 128-dimensional deterministic recurrent state and a 32-dimensional Gaussian stochastic state. Training combines future-feature reconstruction with a Gaussian prior–posterior KL term weighted by 0.01. It has no actor, critic, reward head, or imagined policy optimization.

After RSSM training, we freeze its representation. A separately trained probe maps the predicted 160-dimensional future-prior latent to a positive semidefinite 4×4 information increment using teacher increments as supervision. For the matched decision analysis, both EFO and the RSSM probe add their predicted increments to the same teacher current geometry. **Teacher current geometry is a shared scoring condition, not an input to either increment predictor.** Standalone EFO instead estimates current geometry from observed history. The comparison is neither parameter-

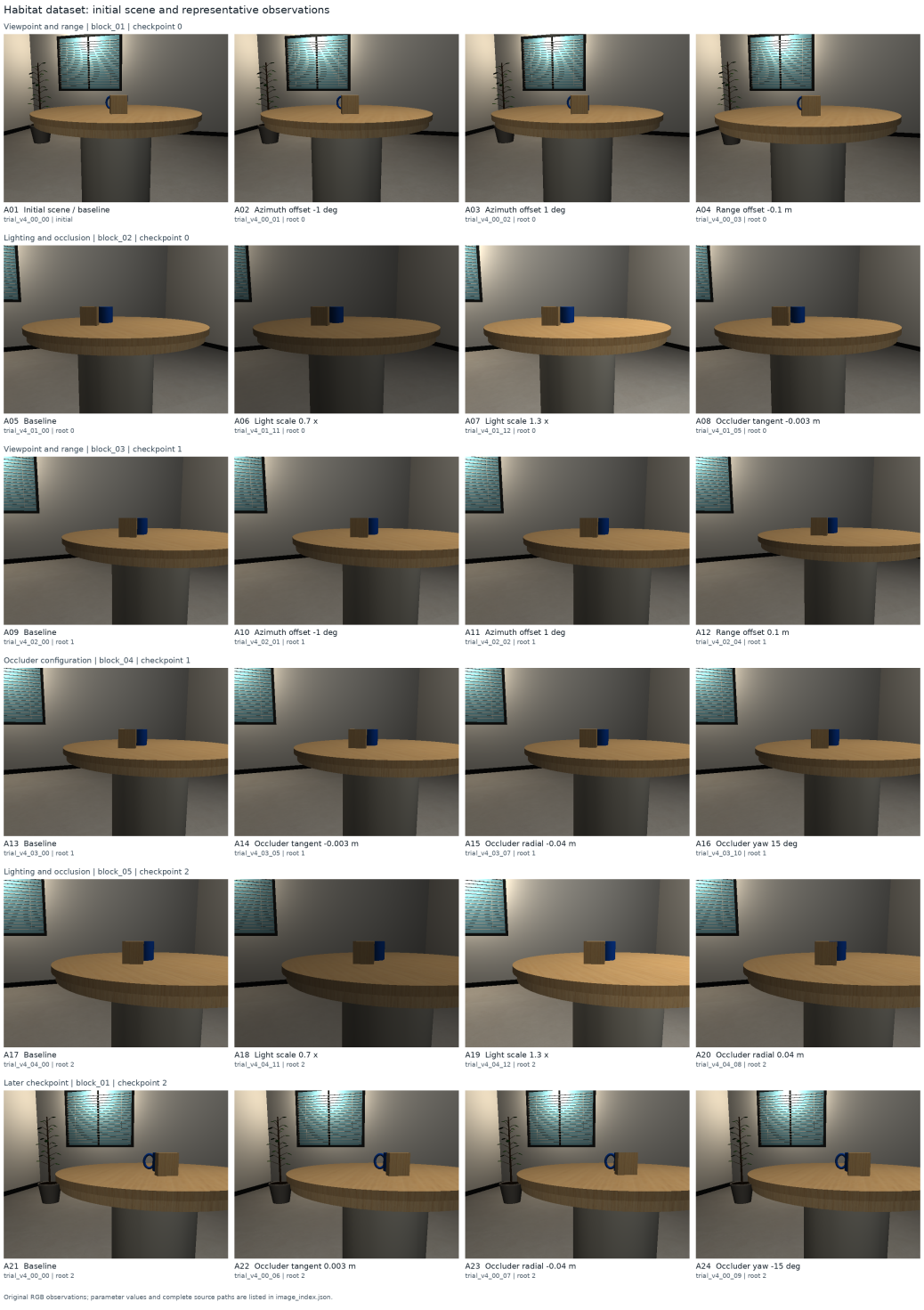


Figure 7: Representative RGB observations across acquisition blocks and controlled conditions. Each tile shows a nominal observation. Separate RGB-D observations of perturbed target states construct the geometry targets.

matched nor inference-equivalent. The RSSM has 182,432 trainable world-model parameters, followed by a 22,672-parameter probe.

D Additional Results for Action-Dependent Identifiability

D.1 Metrics and Action-Switch Criteria

Each learned model in this analysis is trained with three random seeds. Unless specified otherwise, prediction metrics use accumulated post-action geometry and a relative eigenvalue threshold of $\epsilon = 0.01$.

For a teacher matrix Γ and its prediction $\widehat{\Gamma}$, relative Frobenius error is $e_{\mathrm{F}}(\widehat{\Gamma}, \Gamma) = \|\widehat{\Gamma} - \Gamma\|_{\mathrm{F}} / \max(\|\Gamma\|_{\mathrm{F}}, 10^{-12})$. Let P_{I} and $\widehat{P}_{\mathrm{I}}$ be the thresholded identifiable projectors. Projector error is $d_{\mathrm{proj}}(\widehat{\Gamma}, \Gamma) = \|\widehat{P}_{\mathrm{I}} - P_{\mathrm{I}}\|_{\mathrm{F}} / \sqrt{2}$. We report the maximum principal angle $\theta_{\max}$ and assign 90° when teacher and predicted ranks differ. Positive-spectrum error is $e_\lambda(\widehat{\Gamma}, \Gamma) = r_\epsilon(\Gamma)^{-1} \sum_{i \le r_\epsilon(\Gamma)} |\log \max(\widehat{\lambda}_i, \delta) - \log \max(\lambda_i, \delta)|$, where $\delta = 10^{-8} \max(\lambda_1, 10^{-12})$ depends on the teacher spectrum. Rank Macro-F1 averages F1 over the fixed classes 1–4, including classes absent from an individual held-out block. Macro-F1 values using a different class set are not directly comparable.

We assess action dependence separately for the information increment $X^a_{t+1} = \Delta\Gamma^a_{t+1}$ and the accumulated post-action geometry $X^a_{t+1} = \Gamma_t + \Delta\Gamma^a_{t+1}$. For either choice, pairwise action distance is $d^X_{\mathrm{pair}}(a, a') = \|P_{\mathrm{I}}(X^a_{t+1}) - P_{\mathrm{I}}(X^{a'}_{t+1})\|_{\mathrm{F}} / \sqrt{2}$. A projector switch occurs when $\max_{a,a'} d^X_{\mathrm{pair}}(a, a') \ge 0.25$. A rank switch occurs when candidate actions have different effective ranks. A rotation-only switch satisfies the projector criterion while all actions retain the same effective rank.

D.2 Effective-Rank Diagnostics

Figure 8 reports row-normalized confusion matrices for post-action effective rank. EFO without action input concentrates predictions in ranks three and four. With cyclically shifted test-time actions, the trained EFO concentrates near rank three. Correct action inputs produce a broader predicted rank distribution and improve recall for ranks two and four. Rank one accounts for only 0.8% of teacher labels, so its recall is based on few examples.

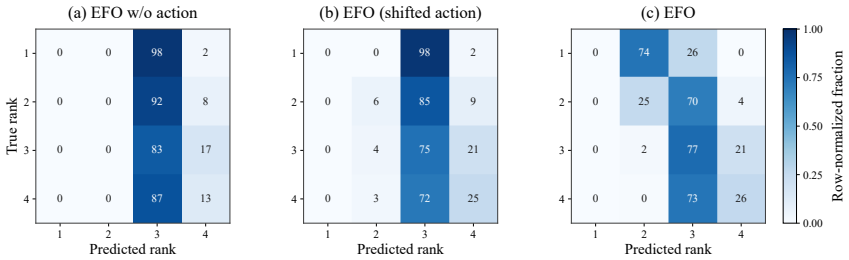


Figure 8: Post-action effective-rank diagnostics. Row-normalized confusion matrices for (a) EFO without action input, (b) EFO with cyclically shifted test-time actions, and (c) EFO with correct actions. Entries are averaged over held-out folds and seeds.

D.3 Threshold and Held-Out-Block Results

Figure 9(a,b) evaluates four relative eigenvalue thresholds from 0.005 to 0.05. EFO has the highest rank Macro-F1 through $\epsilon = 0.02$. At $\epsilon = 0.05$, EFO and EFO without action input are nearly tied, at 0.2967 and 0.2977. EFO has the lowest projector error at the evaluated thresholds from 0.01 to 0.05. The size of its rank advantage therefore depends on the threshold.

Figure 9(c,d) separates the held-out blocks. EFO has the lowest projector error in three of five blocks. Its block-bootstrap interval for projector-error improvement over EFO without action input includes zero. Rank Macro-F1 improves in all five blocks. The equally weighted block gain is 0.111, with a 95% block-bootstrap interval of $[0.046, 0.176]$. **The effective-rank gain is more consistent across blocks than the projector-error gain.**

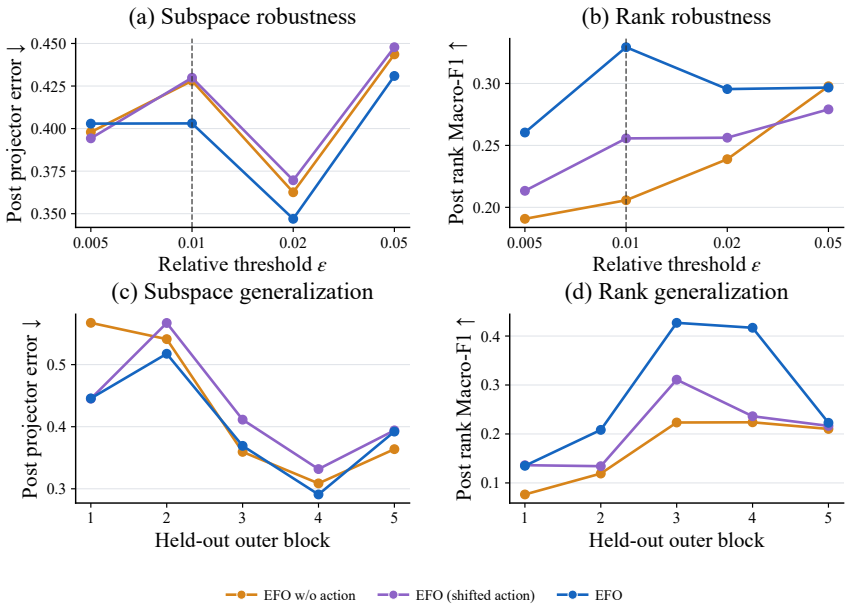


Figure 9: Threshold and held-out-block results. (a,b) Post-action projector error and rank Macro-F1 at four relative eigenvalue thresholds. The dashed line marks $\epsilon = 0.01$. (c,d) The same metrics for each held-out block.

D.4 Teacher Statistics and Decoder Ablation

Table 2 reports teacher statistics, a decoder ablation, and one-factor teacher sensitivity. Panel A separates the rank distribution of individual post-action matrices from switch regimes at checkpoints. Rank three accounts for 52.2% of post-action labels, which motivates Macro-F1 alongside accuracy. Rank switches are common across candidate information increments but less common after accumulation.

Panel B compares direct decoding of post-action geometry with composition of predicted current information and an action-conditioned increment. Composition improves every listed metric. It reduces projector error by 18.5% and positive-spectrum error by 69.9%. **This ablation supports EFO's compositional decoder.**

Panel C compares each one-factor variant with its paired baseline at the initial checkpoint under the same action. Each factor has 60 action-level comparisons. Target range, camera yaw, and camera lateral position yield the largest median teacher projector changes. These teacher-target changes are distinct from the prediction-error changes analyzed in Appendix H.

E Additional Results for Matched Histories

E.1 Matched Views and Teacher Geometry

Each acquisition block contributes eight matched history pairs. Within a pair, the two conditions share the current RGB-D observation and all six candidate actions. Equality is verified in the stored current images and frozen visual features. The maximum current-view RGB pixel MSE and depth MSE are both zero. Earlier observations differ because one condition retains stable depth while the other applies 15% depth dropout. The same sensing condition continues into its candidate-action observations. Neither model receives a teacher matrix at inference.

For pair i and action a, let $P^a_{i,\mathrm{s}}$ and $P^a_{i,\mathrm{d}}$ denote the teacher post-action identifiable projectors for the stable and dropout conditions. We define

$$D^{\mathrm{post}}_{i,a} = \frac{\|P^a_{i,\mathrm{s}} - P^a_{i,\mathrm{d}}\|_{\mathrm{F}}}{\sqrt{2}}. \tag{39}$$

Table 2: Teacher statistics, decoder ablation, and controlled teacher sensitivity.

Panel A: Teacher statistics

Statistic	Categories	Fraction (%)
True post-action rank ($n = 1{,}710$)	1 / 2 / 3 / 4	0.8 / 7.9 / 52.2 / 39.1
Increment regime ($n = 285$)	no switch / rotation-only / rank	3.9 / 3.2 / 93.0
Post-action regime ($n = 285$)	no switch / rotation-only / rank	46.3 / 0.7 / 53.0

Panel B: Post-action decoder ablation

Metric	Direct	Compositional	Improvement (%)
Relative Frobenius error ↓	0.7287	**0.6008**	17.5
Projector error ↓	0.4944	**0.4031**	18.5
Maximum principal angle (°) ↓	56.50	**46.55**	17.6
Positive-spectrum error ↓	2.1608	**0.6500**	69.9
Rank accuracy (%) ↑	42.2	**52.6**	24.6

Panel C: One-factor teacher sensitivity

Perturbed factor	Rank-change rate (%)	Median projector distance
Target range	53.3	0.7071
Camera yaw	36.7	0.2632
Camera lateral position	31.7	0.1794
Target azimuth	31.7	0.0064
Occluder radial position	18.3	0.0004
Occluder yaw	13.3	0.0001
Occluder tangential position	6.7	0.0001
Lighting intensity	1.7	0.0007
Depth dropout	1.7	0.0001

A comparison is history-sensitive when $D^{\text{post}}_{i,a} \geq 0.25$. This teacher-defined criterion selects 56 of 240 pair–action comparisons. They arise from 22 of 40 pairs across all five blocks. The block counts are 17, 13, 15, 4, and 7. All 56 comparisons also differ in teacher effective rank. **The same current view and candidate action can precede different post-action identifiable subspaces in this benchmark.** These counts do not estimate prevalence in unrestricted Habitat trajectories.

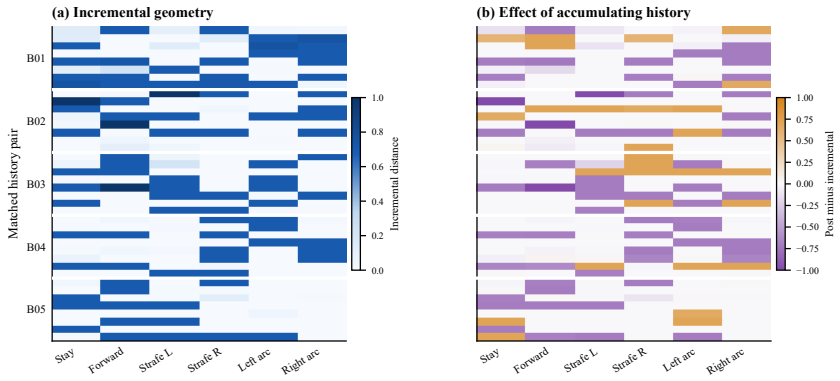


Figure 10: Teacher geometry for matched histories. (a) Distance between identifiable projectors of the next-observation information increments. (b) Post-action projector distance minus increment projector distance for the same pair and action. Purple indicates a decrease after accumulation, and orange indicates an increase. Rows are grouped by held-out block.

Figure 10 shows that composition with accumulated information can increase or decrease the distance between the paired identifiable subspaces. This does not contradict positive semidefinite information accumulation. The distance compares thresholded subspaces between two conditions, rather than information gained within one condition.

E.2 Recovering History-Dependent Geometry

Obs-only receives the current observation and candidate action. History-EFO additionally receives preceding observations and actions. Both use matched recurrent architectures and the same frozen visual features. Principal-component reduction is fitted on training data, and model selection uses a separate validation block. In the swapped-history control, the trained History-EFO receives the other history in a matched pair while the current observation and candidate action remain fixed.

Let $\widehat{P}^a_{i,\mathrm{s}}$ and $\widehat{P}^a_{i,\mathrm{d}}$ be the predicted post-action projectors. Signed projector-contrast error is

$$e_{\Delta P}(i,a) = \frac{\|(\widehat{P}^a_{i,\mathrm{s}} - \widehat{P}^a_{i,\mathrm{d}}) - (P^a_{i,\mathrm{s}} - P^a_{i,\mathrm{d}})\|_{\mathrm{F}}}{\sqrt{2}}. \tag{40}$$

This metric compares the signed projector difference, not merely its magnitude. Obs-only receives identical inputs for both conditions, so its predicted contrast is zero. For Figure 3(b), we average selected actions within each pair and seed, then average pairs within each held-out block. The bold line gives the mean across blocks.

On history-sensitive comparisons, History-EFO reduces block-mean signed projector-contrast error from 0.707 for Obs-only to 0.669. The paired reduction is 0.039, with a descriptive block-bootstrap 95% interval of $[0.015, 0.062]$. Swapping paired histories increases the error to 0.804. **Correct history helps recover the difference between the paired future geometries.**

Table 3: Absolute post-action prediction for matched histories. Entries pool evaluated targets within each seed and report mean $\pm$ standard deviation over five seeds. The history-sensitive subset is defined by teacher geometry.

Method	Projector error ↓	Principal angle (°) ↓	Rank accuracy ↑
All matched comparisons			
Obs-only	0.2824 ± 0.0042	32.82 ± 0.80	0.6638 ± 0.0123
History-EFO	0.3318 ± 0.0171	37.80 ± 1.81	0.6217 ± 0.0187
Swapped history	0.3429 ± 0.0168	39.11 ± 1.70	0.6075 ± 0.0171
History-sensitive comparisons			
Obs-only	0.3844 ± 0.0049	46.83 ± 0.38	0.4982 ± 0.0040
History-EFO	0.3850 ± 0.0150	45.62 ± 1.30	0.5250 ± 0.0132
Swapped history	0.4320 ± 0.0157	51.19 ± 1.14	0.4643 ± 0.0109

Table 3 evaluates absolute prediction error, which differs from signed contrast recovery. On the history-sensitive subset, History-EFO improves rank accuracy and principal angle. Its pooled projector error is nearly the same as Obs-only. Its block-balanced paired projector-error benefit is 0.0029, with a 95% block-bootstrap interval of $[-0.0161, 0.0227]$. Swapping history increases block-balanced projector error by 0.0639, with an interval of $[0.0249, 0.1025]$. Paired effects need not equal differences between table entries because the paired analysis averages within pairs before weighting blocks equally. Across all matched comparisons, Obs-only has lower absolute errors.

E.3 Held-Out Blocks and Uncertainty

History-EFO has higher overall projector error in four of five blocks (Figure 11(a)). On selected comparisons, its signed-contrast benefit is positive in four blocks and zero in one. Swapping history increases projector error in four blocks and leaves one unchanged. Seeds are averaged within each pair before paired block effects are calculated. Descriptive confidence intervals resample the five blocks. Exact one-sided block sign-flip values are 0.0625 for both sensitive-subset contrast recovery and swapped-history projector damage. **The supported gain is recovery of history-associated differences on teacher-identified comparisons, not an overall post-action accuracy gain.**

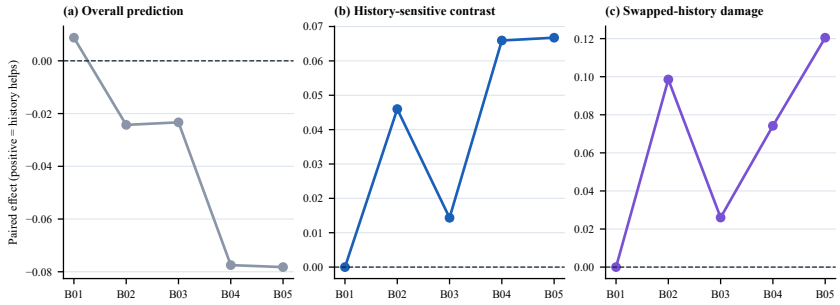


Figure 11: Paired effects across held-out blocks. (a) Obs-only minus History-EFO projector error on all comparisons. (b) Obs-only minus History-EFO signed projector-contrast error on history-sensitive comparisons. (c) Swapped-history minus correct-history projector error on history-sensitive comparisons. Positive values favor History-EFO or correctly matched history. Connecting lines do not imply an ordered trend.

F Additional Results for Directional Action Selection under Scalar Collisions

The scalar comparison uses held-out EFO matrices. For each test block and model seed, we select the EFO scalar scoring rule using held-out predictions and teacher regret from the other blocks. We then apply it to the same test-block predictions used by the full-matrix policy.

F.1 Collision Definition and Decision Metrics

For actions a and b, let Γ^a_{t+1} and Γ^b_{t+1} be teacher post-action information matrices. Their relative trace gap is $\delta_{\mathrm{tr}}(a,b) = |\operatorname{tr}(\Gamma^a_{t+1}) - \operatorname{tr}(\Gamma^b_{t+1})| / \max\{|\operatorname{tr}(\Gamma^a_{t+1})|, |\operatorname{tr}(\Gamma^b_{t+1})|, 10^{-12}\}$. Their ambiguous-subspace distance is $d_{\mathrm{P}}(a,b) = \|P_{\mathrm{A}}(\Gamma^a_{t+1}) - P_{\mathrm{A}}(\Gamma^b_{t+1})\|_{\mathrm{F}}/\sqrt{2}$.

The primary diagnostic set requires $\delta_{\mathrm{tr}}(a,b) \le 0.10$, $d_{\mathrm{P}}(a,b) \ge 0.50$, equal effective rank, and a maximum principal angle of at least 30°. It also requires a teacher action-preference crossover: one evaluated direction favors a, while another favors b. Evaluated directions include the four normalized pose-coordinate axes, bases of both teacher ambiguous subspaces, and a direction of maximum projector disagreement. Directions equivalent up to sign are deduplicated. These teacher-derived directions define the same diagnostic task for every policy. They are neither independent trials nor selected by EFO.

For a unit direction q, residual ambiguity is $\rho_{\mathrm{A}}(q;\Gamma) = q^\top P_{\mathrm{A}}(\Gamma) q$. The teacher-optimal action minimizes this value within the pair. For a selected action $\widehat{a}$, null-exposure regret is $R_{\mathrm{null}}(q,\widehat{a}) = \rho_{\mathrm{A}}(q;\Gamma^{\widehat{a}}_{t+1}) - \min_{a\in\{a_1,a_2\}} \rho_{\mathrm{A}}(q;\Gamma^a_{t+1})$. Oracle agreement records whether the selected action reaches the teacher minimum. The primary null-exposure metric uses no numerical ridge. Posterior regret with a ridge is a secondary analysis.

F.2 Directional Crossovers and Scalar Summaries

Figure 12(a) gives a directional crossover. Forward and Strafe R have traces of 68.6k and 63.7k, a relative gap of 7.1%. Their ambiguous-projector distance is 0.99, and their maximum principal angle is 82°. Strafe R is preferred for the z direction, while Forward is preferred for the θ direction. A single action ranking that does not receive the query direction cannot express both preferences.

Table 4 examines other spectral summaries at a 10% score tolerance and ambiguous-projector threshold of 0.5. Ordinary scalar scores use relative score gaps. D-optimality uses $1 - \exp(-|\bar{\ell}_a - \bar{\ell}_b|)$, where $\bar{\ell}$ is the mean regularized log eigenvalue. The full-spectrum gap uses $1 - \exp(-\|\ell_a - \ell_b\|_\infty)$ for the regularized log-eigenvalue vectors. The effective-rank gap is zero for equal ranks and one otherwise. Regularization uses the checkpoint-specific ridge for spectral scores.

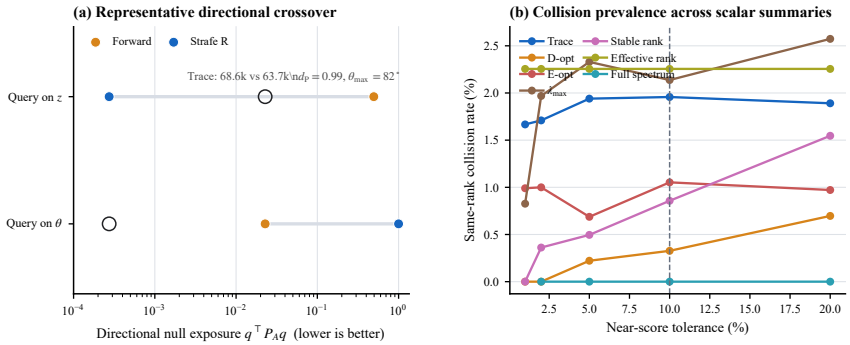


Figure 12: Directional collision diagnostics. (a) A same-rank action pair whose teacher-preferred action changes between the z and θ directions. (b) Same-rank collision prevalence across scalar-score tolerances. The dashed line marks the primary 10% tolerance.

Table 4: Collision prevalence at 10% score tolerance and ambiguous-projector threshold 0.5. Near rates use all 4,275 within-checkpoint action pairs. Percentages in the final column use near pairs of equal effective rank.

Summary	Near pairs	Near rate	Collisions among near	Same-rank collisions
Trace	**1,467**	**34.3%**	**24.9%**	**22 (2.0%)**
D-optimality	966	22.6%	5.3%	3 (0.3%)
E-optimality	1,128	26.4%	16.7%	10 (1.1%)
Largest eigenvalue	1,493	34.9%	26.4%	24 (2.1%)
Stable rank	2,556	59.8%	23.1%	17 (0.9%)
Effective rank	3,104	72.6%	2.3%	70 (2.3%)
Full eigenvalue spectrum	104	2.4%	0.0%	0 (0.0%)

Several scalar summaries have same-rank collisions, although their prevalence differs. No pair satisfies the near-spectrum and projector-distance criteria for the complete eigenvalue vector at the primary tolerance. **The observed collision result concerns scalar spectral summaries.** It does not establish insufficiency of the complete eigenvalue spectrum.

F.3 Collision-Conditioned Action Selection

The primary diagnostic set has 22 action pairs from three held-out blocks. Their teacher-defined directions give 154 pair–query decisions. EFO scores each action using its full predicted post-action matrix. EFO scalar uses a cross-block-selected spectral scoring rule on the same predictions without receiving the query direction. EFO with shifted action input uses the trained EFO with cyclically replaced actions at evaluation.

Table 5 reports pooled absolute outcomes and equal-block paired effects. Absolute outcomes pool pair–query decisions after averaging predictions across seeds. Paired effects first compare policies on the same pair and query, then weight the represented blocks equally. These aggregation rules produce different numerical contrasts.

EFO has the lowest pooled regret and highest pooled agreement among the learned policies. Its paired comparison with EFO scalar tests the value of retaining orientation in the same predicted matrices. The shifted-action control tests correct action correspondence. Paired intervals against Unstructured EFO cross zero. **On this diagnostic set, the clearest improvement is over direction-independent scalarization of EFO's own predictions.**

F.4 Blockwise and Threshold Sensitivity

Figure 13(a,b) shows positive EFO-versus-scalar regret and agreement effects in each of the three represented blocks. Effects against shifted-action and Unstructured EFO are less uniform. Figure 13(c,d) varies trace tolerance over $\{1, 2, 5, 10, 20\}\%$ and projector threshold over

Table 5: Action selection on the teacher-defined collision set. Absolute results cover 22 pairs, 154 pair–query decisions, and three represented blocks. Confidence intervals resample blocks and are descriptive. Positive paired effects favor EFO.

Policy	$R_{\text{null}} \downarrow$	95% CI	Agreement↑	95% CI
Teacher geometry	0.000	$[0.000, 0.000]$	100.0%	$[100.0, 100.0]$
EFO scalar	0.222	$[0.174, 0.278]$	51.9%	$[48.1, 57.1]$
EFO (shifted action)	0.216	$[0.153, 0.345]$	48.5%	$[35.7, 61.2]$
Unstructured EFO	0.151	$[0.144, 0.183]$	65.8%	$[61.9, 69.7]$
EFO	**0.083**	$[0.026, 0.196]$	**72.1**%	$[57.8, 86.1]$
Paired equal-block effect of EFO relative to each comparator				
Comparator	$\Delta R_{\text{null}} \uparrow$	95% CI	ΔAgreement (pp)↑	95% CI
EFO scalar	**0.118**	$[\mathbf{0.066}, \mathbf{0.205}]$	**14.0**	$[\mathbf{1.2}, \mathbf{38.1}]$
EFO (shifted action)	0.126	$[0.045, 0.183]$	20.1	$[-3.4, 41.1]$
Unstructured EFO	0.047	$[-0.013, 0.118]$	2.9	$[-4.1, 16.5]$

$\{0.25, 0.50, 0.75\}$ while retaining the other diagnostic criteria. Every displayed nonempty cell has a positive mean effect relative to EFO scalar. The primary thresholds give a mean equal-block regret reduction of 0.118 and agreement gain of 14.0 percentage points.

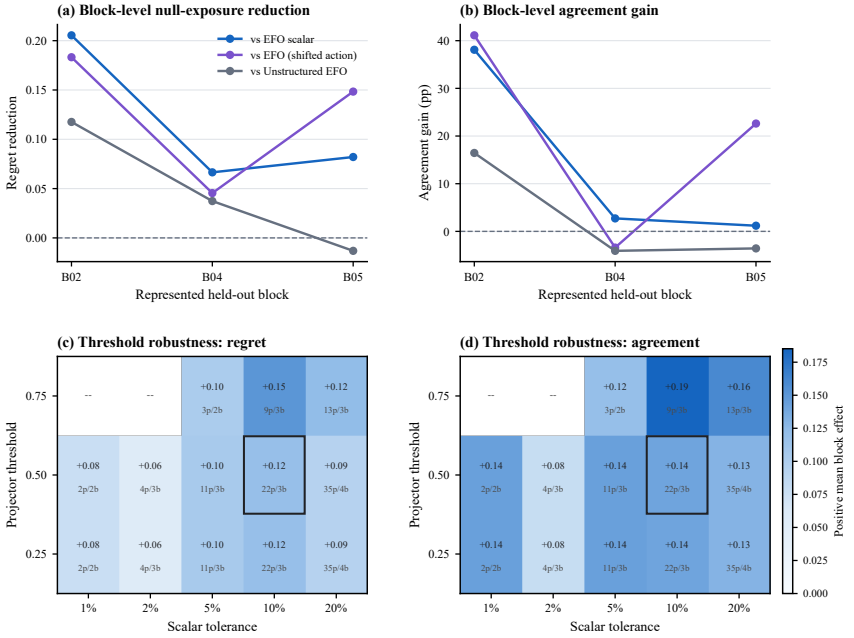


Figure 13: Blockwise effects and collision-threshold sensitivity. (a,b) Paired block effects of EFO relative to alternative policies on the primary diagnostic set. (c,d) Equal-block effects relative to EFO scalar across trace-gap and projector-distance thresholds. Cells report effects and represented pair and block counts. The outlined cell marks the primary setting. Positive values favor EFO.

Only three independent blocks contribute to the primary set. The smallest possible one-sided value in the exact block sign-flip test is 0.125. EFO-versus-scalar regret and agreement effects attain that value. Descriptive block-bootstrap intervals summarize variation among represented blocks. They do not establish conventional $p < 0.05$ significance.

F.5 Scope and Query-Direction Sensitivity

Figure 14(a) reports secondary posterior regret across ridge factors. The primary null-exposure metric remains ridge-free. Figure 14(b) groups all 3,104 same-rank pairs by teacher ambiguous-projector distance. Regret reduction is near zero for similar ambiguous subspaces and reaches 0.065

in the largest-distance bin. Agreement gains are negative in several intermediate bins and positive in the largest bin.

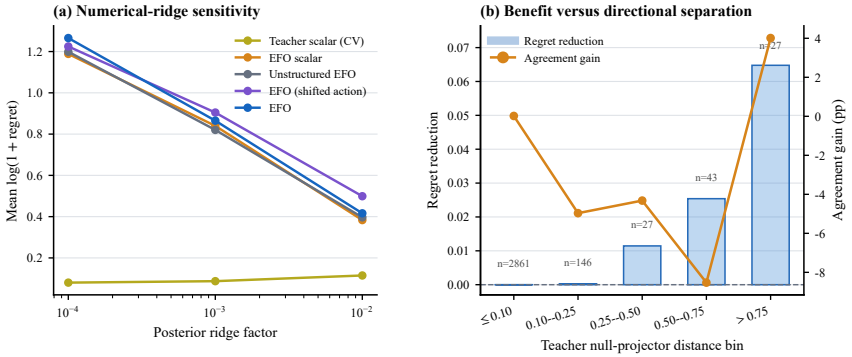


Figure 14: Decision sensitivity beyond the primary diagnostic set. (a) Secondary posterior regret across numerical ridge factors. (b) EFO's effect relative to EFO scalar across teacher ambiguous-projector-distance bins for all same-rank pairs. Bars show regret reduction, the line shows agreement gain, and counts denote action pairs.

Table 6 separates the four physical pose-coordinate axes from directions constructed using teacher geometry. EFO reduces null-exposure regret relative to EFO scalar in both groups. On physical axes, Unstructured EFO has higher agreement than EFO, despite EFO's lower regret. The result does not give one architectural ranking across metrics or query groups.

Table 6: Query groups and broader-population results. Panel A reports absolute outcomes on the primary diagnostic set. Panel B reports equally weighted block effects across all 3,104 same-rank pairs from five blocks. Positive paired effects favor EFO.

Query group	Policy	$R_{\text{null}} \downarrow$	Agreement↑
Physical axes	EFO scalar	0.143	52.7%
	EFO (shifted action)	0.138	45.8%
	Unstructured EFO	0.069	**70.5%**
	EFO	**0.060**	64.0%
Constructed directions	EFO scalar	0.327	51.0%
	EFO (shifted action)	0.320	52.0%
	Unstructured EFO	0.261	59.6%
	EFO	**0.114**	**82.8%**

Comparator	ΔR_{null}	95% CI	ΔAgreement (pp)	95% CI
EFO scalar	0.0011	$[-0.0014, 0.0044]$	-0.35	$[-0.86, 0.07]$
EFO (shifted action)	0.0019	$[-0.0015, 0.0064]$	0.20	$[-1.27, 1.29]$
Unstructured EFO	-0.0002	$[-0.0028, 0.0022]$	-0.35	$[-0.79, 0.09]$

Across all same-rank pairs, the block-mean EFO-versus-scalar regret reduction is 0.0011, with an interval crossing zero. Agreement changes by -0.35 percentage points, also with an interval crossing zero. Exact one-sided block sign-flip values are 0.3125 and 0.875. **The observed decision benefit is concentrated in the teacher-defined cases with similar scalar scores and separated ambiguous directions.**

G Additional Results for Epistemic Geometry and Directional Decisions

G.1 Evaluation Protocol and Incremental Geometry

EFO and the feature-predictive RSSM probe each predict an action-conditioned information increment from observed history. Geometry metrics compare $\widehat{\Delta\Gamma}^{a}_{t+1}$ with the same teacher increment

$\Delta\Gamma^a_{t+1}$. For decision scoring, each predicted increment is added to the teacher current geometry Γ^{teach}_t. This teacher matrix is shared by both methods and is not an input to either increment predictor. Standalone EFO estimates current geometry from observed history, as described in Section 4.

We use three seeds and relative eigenvalue thresholds of 0.005, 0.01, and 0.02. Table 7 reports the primary threshold of 0.01. Neither teacher nor predicted increments have effective rank zero at these thresholds. Rank Macro-F1 therefore averages F1 over the fixed classes 1–4 while retaining all evaluated actions. Projector error and padded maximum principal-angle error assess identifiable subspaces. Relative Frobenius error assesses the full increment. Metrics pool held-out actions within each seed and report mean $\pm$ standard deviation across seeds. Both methods have 100% valid prediction coverage.

Table 7: Information-increment prediction at relative rank threshold 0.01. Entries are mean $\pm$ standard deviation over three seeds. Macro-F1 uses fixed ranks 1–4.

Method	Rank accuracy↑	Macro-F1↑	$d_{\text{proj}}\downarrow$	$\theta_{\text{pad}}(^\circ)\downarrow$	$e_{\text{F}}\downarrow$
RSSM probe	0.383 ± 0.019	0.204 ± 0.009	0.552 ± 0.011	61.32 ± 1.34	$\mathbf{0.756 \pm 0.016}$
EFO	$\mathbf{0.490 \pm 0.015}$	$\mathbf{0.291 \pm 0.006}$	$\mathbf{0.534 \pm 0.003}$	$\mathbf{55.37 \pm 0.87}$	0.792 ± 0.046

EFO has higher rank accuracy and Macro-F1 and lower projector and padded-angle errors. The RSSM probe has lower Frobenius error. **Lower full-matrix error does not guarantee more accurate effective rank and identifiable subspace.** Both this table and Table 1 use rank classes 1–4. This table evaluates information increments, while Table 1 evaluates accumulated post-action geometry. Their Macro-F1 values are therefore not directly comparable.

G.2 Directional Decision Risk Across Ridge Factors

We use the directional action rule and teacher oracle described in Appendix F. There are 68 fixed query directions at each of 285 checkpoints, giving 19,380 query–checkpoint pairs per seed. Queries at one checkpoint share geometry targets and are not independent replicates. Normalized regret measures the cost of the selected action relative to the teacher-optimal action. CVaR95 averages the highest 5% of regrets. Oracle agreement is the fraction of decisions that attain a teacher-optimal action.

The reported ρ is a ridge factor. It sets a checkpoint-specific regularizer rather than a constant added to all matrices:

$$\rho^{\text{reg}}_t = \max\left\{\rho \ \text{median}_a \ \lambda_{\max}\left(\Gamma^{\text{teach}}_t + \Delta\Gamma^a_{t+1}\right), 10^{-12}\right\}. \tag{41}$$

Both methods and all actions at checkpoint t use the same ρ^{reg}_t. We report variation across seeds without treating queries as independent training runs.

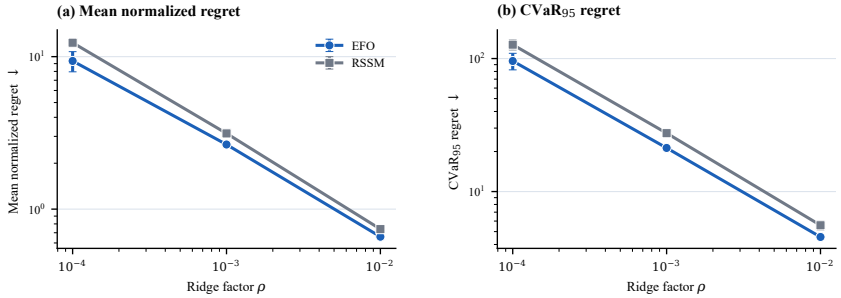


Figure 15: Directional decision regret across ridge factors. (a) Mean normalized regret. (b) CVaR95, the mean regret of the worst 5% of query–checkpoint decisions. Points show three-seed means. Both axes use logarithmic scales.

EFO has lower mean regret and CVaR95 at all three ridge factors (Figure 15 and Table 8). At the primary factor of 10^{-3}, mean regret is 2.65 for EFO and 3.14 for the RSSM probe. CVaR95 is 21.28

Table 8: Directional decisions across ridge factors using all 68 query directions. Entries are mean ± standard deviation over three seeds.

Ridge factor ρ	Method	Median regret↓	Mean regret↓	CVaR95↓	Oracle agreement↑
10^{-4}	RSSM probe	$\mathbf{0.667 \pm 0.055}$	12.34 ± 0.55	127.19 ± 11.07	$\mathbf{0.199 \pm 0.011}$
	EFO	0.741 ± 0.035	$\mathbf{9.37 \pm 1.42}$	$\mathbf{95.86 \pm 13.63}$	0.184 ± 0.003
10^{-3}	RSSM probe	$\mathbf{0.569 \pm 0.042}$	3.14 ± 0.18	27.49 ± 1.75	$\mathbf{0.202 \pm 0.011}$
	EFO	0.633 ± 0.021	$\mathbf{2.65 \pm 0.06}$	$\mathbf{21.28 \pm 0.18}$	0.188 ± 0.002
10^{-2}	RSSM probe	$\mathbf{0.313 \pm 0.021}$	0.739 ± 0.043	5.58 ± 0.43	$\mathbf{0.222 \pm 0.013}$
	EFO	0.329 ± 0.012	$\mathbf{0.659 \pm 0.017}$	$\mathbf{4.56 \pm 0.12}$	0.196 ± 0.006

and 27.49, respectively. The RSSM probe retains lower median regret and higher oracle agreement at every factor. **EFO's advantage under this shared-current-geometry protocol concerns average and severe-loss regret, rather than every decision metric.**

G.3 Query-Set Sensitivity

The 68 directions comprise four normalized physical-coordinate axes and 64 fixed mixtures. Table 9 reports results separately at the primary ridge factor. EFO has lower mean regret and CVaR95 in both groups. The RSSM probe has higher oracle agreement in both, with the larger gap on physical-coordinate queries.

Table 9: Directional decisions by query group at ridge factor $\rho = 10^{-3}$. Entries are mean ± standard deviation over three seeds. N counts query–checkpoint pairs per seed.

Query set	Method	N	Mean regret↓	CVaR95↓	Oracle agreement↑
Physical (4)	RSSM probe	1,140	2.022 ± 0.066	26.12 ± 0.75	$\mathbf{0.280 \pm 0.008}$
	EFO	1,140	$\mathbf{1.651 \pm 0.098}$	$\mathbf{18.99 \pm 2.24}$	0.199 ± 0.010
Mixed (64)	RSSM probe	18,240	3.207 ± 0.185	27.53 ± 1.82	$\mathbf{0.197 \pm 0.012}$
	EFO	18,240	$\mathbf{2.716 \pm 0.075}$	$\mathbf{21.38 \pm 0.18}$	0.188 ± 0.003

H Additional Results for Controlled Sensing Changes

H.1 Pairing Protocol and Aggregate Sensitivity

This analysis reuses predictions from the five held-out acquisition blocks without retraining. Nine factors have two shifted conditions each. At the first checkpoint, the 18 conditions yield 540 distinct shifted action branches. Each shifted branch is paired with its baseline trial under the same candidate action, method, and seed. All 8,640 shifted prediction rows have matched baselines. Acquisition blocks are the units for uncertainty estimation.

For metric j, branch i, and method M, let $e_j(i, M)$ denote post-action prediction error and $b(i)$ its paired baseline branch. After averaging seeds within a branch, paired degradation is $d_j(i, M) = e_j(i, M) - e_j(b(i), M)$. We average paired differences within each block and then weight the five blocks equally. Positive values indicate higher shifted error. Descriptive confidence intervals resample blocks 5,000 times. Effective ranks and identifiable projectors use $\epsilon = 0.01$. Rank error averages $1 - \mathbb{1}\{\widehat{r} = r\}$ over branches and differs from rank Macro-F1.

Table 10 shows positive average changes in all five errors. Block-bootstrap intervals include zero for projector, principal-angle, and rank errors. Intervals for matrix and spectrum errors lie above zero. These intervals describe variation across five blocks. They are not equivalence or non-inferiority tests.

H.2 Shift Categories and Individual Conditions

Figure 6(a) groups changes by target range, target azimuth, camera pose, occlusion, lighting, and depth corruption. Target-range changes raise projector, principal-angle, and rank errors by 11.8%,

Table 10: Paired EFO post-action prediction errors at the first checkpoint. Changes compare shifted branches with matched baselines. Confidence intervals apply to absolute paired changes.

Metric	Baseline	Shifted	Change	Relative	95% block CI
Projector error	0.3978	0.4235	+0.0257	+6.5%	[−0.0036, 0.0508]
Principal angle (°)	41.52	45.00	+3.48	+8.4%	[−0.60, 7.45]
Rank error	0.3778	0.4210	+0.0432	+11.4%	[−0.0056, 0.0920]
Matrix error	0.5311	0.5676	+0.0364	+6.9%	[0.0028, 0.0808]
Spectrum error	0.5969	0.6588	+0.0619	+10.4%	[0.0065, 0.1090]

15.7%, and 20.6% on average. Camera-pose changes raise them by 9.0%, 13.2%, and 17.6%. Lighting and depth corruption have smaller average effects on these metrics. **Prediction sensitivity depends on which part of the sensing configuration changes.**

Opposite perturbations within a category can have different effects. Figure 16 retains the signed changes for all 18 conditions. Increasing target range by 0.1 m raises projector error by 0.109, while decreasing it by 0.1 m changes projector error by −0.015. A camera-yaw change of −1° raises projector error by 0.100 and principal-angle error by 16.1°. These are outcomes at the tested perturbations, not an estimated response curve for larger shifts.

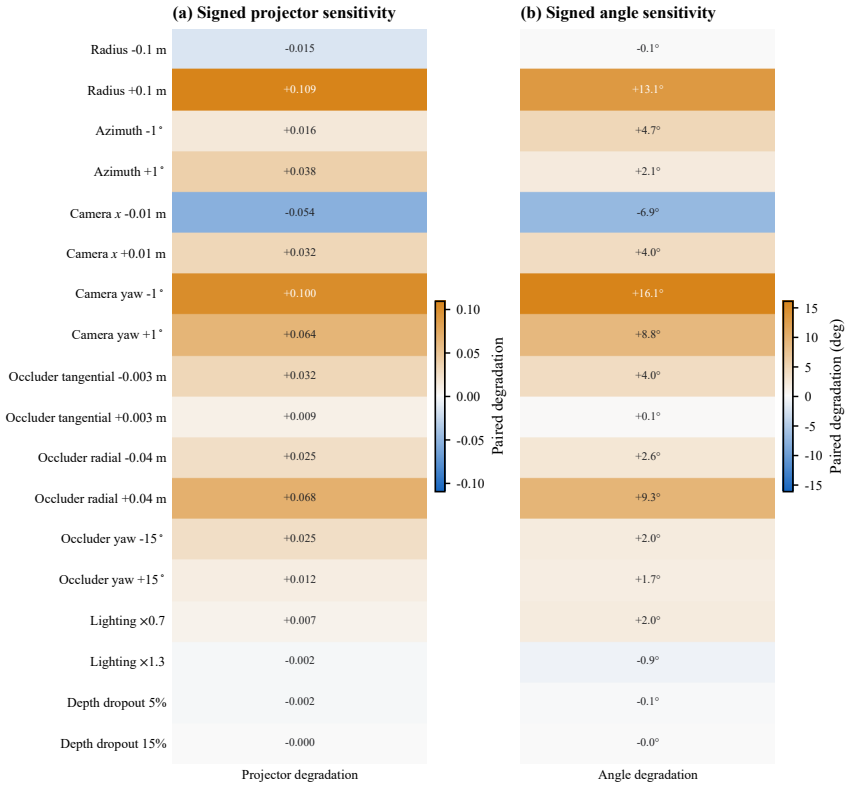


Figure 16: Signed prediction-error sensitivity at the first checkpoint for all 18 conditions. Cells show paired changes in EFO post-action projector and principal-angle errors. Positive values indicate higher error under the shift. Panels use separate color scales because the metrics have different units.

H.3 Method Comparisons Under Shifts

Table 11 compares methods on the same shifted first-checkpoint branches. EFO without action input is separately trained. EFO with shifted actions holds its trained parameters fixed and cyclically

replaces test-time action inputs. The upper panel reports equally weighted block means. The lower panel reports paired advantages of EFO.

Table 11: Post-action prediction on shifted first-checkpoint branches. The upper panel gives equally weighted block means. The lower panel gives paired EFO advantages, descriptive block-bootstrap intervals, and exact one-sided block sign-flip values. The p values are unadjusted for multiple comparisons.

Method	Projector	Angle (°)	Rank error	Matrix	Spectrum
EFO w/o action	0.4288	47.41	0.4698	**0.5656**	0.7590
EFO (shifted action)	0.4911	51.96	0.4969	0.6424	0.7829
Unstructured EFO	0.6260	61.92	0.5006	0.8593	0.6686
EFO	**0.4235**	**45.00**	**0.4210**	0.5676	**0.6588**

Comparator	Metric	EFO advantage [95% CI]	Positive blocks	Exact p
EFO w/o action	Projector	0.0053 [−0.0304, 0.0379]	3/5	0.4375
EFO w/o action	Angle (°)	2.41 [−0.33, 5.49]	4/5	0.0938
EFO w/o action	Rank error	0.0488 [0.0235, 0.0821]	5/5	0.0313
EFO (shifted action)	Projector	0.0676 [0.0189, 0.1171]	4/5	0.0625
EFO (shifted action)	Angle (°)	6.96 [0.76, 13.23]	3/5	0.1250
EFO (shifted action)	Rank error	0.0759 [0.0012, 0.1599]	3/5	0.1250
Unstructured EFO	Projector	0.2025 [0.0692, 0.3313]	5/5	0.0313
Unstructured EFO	Angle (°)	16.92 [6.03, 27.82]	4/5	0.0625
Unstructured EFO	Rank error	0.0796 [0.0160, 0.1444]	4/5	0.0625

EFO has the lowest block-mean projector, principal-angle, and rank errors on these branches. Its projector advantage over Unstructured EFO is positive in all five blocks. The advantage over shifted-action EFO is positive in four. EFO without action input has a slightly lower mean matrix error, while EFO has lower rank error in all five blocks. **The clearest shifted-condition results support structured prediction and correct action correspondence on the evaluated variations.**

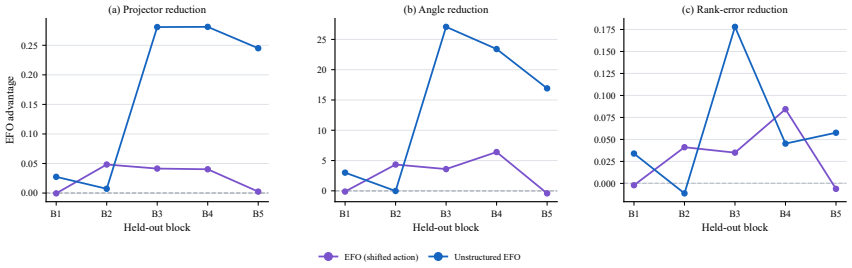


Figure 17: EFO's blockwise advantage over EFO with shifted action input and Unstructured EFO. This diagnostic includes all shifted checkpoints. Figure 6(b) and Table 11 use only the first checkpoint. Positive values favor EFO. Connecting lines do not indicate a temporal sequence.

Figure 17 extends the method comparison to all shifted checkpoints. EFO's projector advantage over Unstructured EFO remains positive in every held-out block. Its advantage over shifted-action EFO varies more across blocks, especially for rank error. These results concern controlled variations in the evaluated layout. They do not establish transfer to new environments or closed-loop navigation.