# EQUIVARIANT COUNTERFACTUAL SUPERVISION FOR HIDDEN-DYNAMICS REPRESENTATION LEARNING

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

We introduce Equivariant Counterfactual Residual (ECR) supervision for learning physical context from a robot's interaction history. The method compares simulated responses under controlled changes to an environmental property, using comparison rules appropriate to material and geometric changes. Subtracting the mean response across the compared conditions produces residual targets. A reservoir feature bank summarizes the interaction history, and a learned code predicts these targets at a query from another rollout in the same environment. We evaluate what physical information the code makes accessible and how useful it is to new prediction tasks. Relative to Next-Step Prediction (NSP), ECR raises steepness linear-probe $R^2$ from 0.555 to 0.907 and improves mean exact-cell retrieval in all ten factor-level holdouts. A paired target-centering ablation corroborates the accessibility difference under matched heads and scales. Fresh predictors show a modest in-distribution next-step prediction benefit, whereas NSP has lower mean error on all four held-out factor-response tasks. The results establish a supervision-dependent accessibility effect whose predictive value depends on the receiving task.

## 1 INTRODUCTION

Robots often interact with environments whose physical properties are not directly measured. Friction, compliance, and terrain geometry affect the response to an action, while recent interaction provides evidence about these conditions. History-conditioned models and policies use this evidence to estimate physical parameters or learn environment representations (Yu et al., 2017; Kumar et al., 2021; Lee et al., 2020). Our goal is to learn a representation that captures this physical context and remains useful beyond the trajectory from which it was inferred.

A common approach trains a history code through Next-Step Prediction (NSP), asking it to predict the robot's next transition. This connects the representation to physical dynamics, but prediction accuracy alone does not tell us which environmental properties are easy to recover from the code. We investigate a different supervision signal: the changes in response produced by differences in physical conditions.

We propose *equivariant counterfactual supervision* to learn from these differences. In simulation, we compare responses to the same action while changing one environmental property. The other factors remain fixed, and the initial robot–environment configuration is kept comparable according to the factor being varied. Material changes and geometric changes therefore use different comparison rules. Subtracting the mean response across the compared conditions produces residual targets that preserve their pairwise differences. We hypothesize that predicting these targets makes environmental variation more accessible in the learned code.

To represent interaction history, we use reservoir computing (RC) with a temporal feature bank (Jaeger, 2001). It combines reservoir states sampled throughout a history window with their mean and standard deviation, rather than relying only on the final state. A learned mapping converts these features into an environment code. We train the code to predict residual responses at a state–action query from another rollout under the same environmental conditions. This pairing encourages context that is useful beyond one local movement. We then freeze the representation and assess both its physical-factor information and its usefulness to new predictors.



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.

Our main contributions are:

1. We introduce ECR supervision that uses controlled physical-response differences to learn environment representations across rollouts.
2. We compare supervision objectives under common reservoir features and show improved steepness accessibility relative to NSP training, supported by history and target-centering controls.
3. We evaluate representation transfer and downstream prediction. ECR improves exact-cell retrieval at unseen factor levels and modestly improves in-distribution next-step prediction, while NSP performs better on the four held-out factor-response tasks.

## 2 Related work

**Dynamics context from interaction.** UP-OSI estimates physical parameters from recent states and actions (Yu et al., 2017), while RMA learns an environment representation for locomotion (Kumar et al., 2021). CaDM learns trajectory context through forward and backward dynamics prediction (Lee et al., 2020); LDG structures a dynamics latent through predictive and contrastive learning (Xu et al., 2026). We compare the supervision principles behind these approaches under common recurrent features, rather than their complete control systems.

**Counterfactual physical learning.** CoPhy infers physical context from observed dynamics to predict intervened outcomes (Baradel et al., 2020); Filtered-CoPhy predicts such outcomes in pixel space without object-state supervision (Janny et al., 2022). CoDA and MoCoDA generate alternative transitions through local causal structure or learned dynamics (Pitis et al., 2020; 2022). We instead supervise representations with simulator-derived response contrasts, extending a related friction-only construction (Anonymous, 2026) to geometry-dependent matching and multiple factors. Unlike residual control (Johannink et al., 2019), our residuals are training targets, not action corrections.

**Representations from structured changes.** Paired observations can support weakly supervised disentanglement (Locatello et al., 2020). CITRIS identifies causal factors from temporal observations and known intervention targets under explicit assumptions (Lippe et al., 2022); BISCUIT infers causal variables and binary interactions with unknown targets (Lippe et al., 2023). We use specified simulator factors to construct supervision and assess accessibility, rather than identify a causal graph. CausalWorld also uses robotic interventions to define generalization distributions (Ahmed et al., 2020); here we study history representations across contact conditions.

**Temporal representations and task utility.** Contrastive Predictive Coding predicts in latent space (van den Oord et al., 2018), while TS2Vec learns hierarchically from augmented temporal contexts (Yue et al., 2022). Reservoir theory connects echo-state maps to dynamical-system embeddings under suitable assumptions (Hart et al., 2020); it motivates temporal features, not an embedding guarantee for our finite-window code. Task-oriented representations instead preserve behavioural similarity (Zhang et al., 2021) or jointly learn dynamics and value for planning, as in TD-MPC (Hansen et al., 2022). Our objective supervises physical contrasts without policy or value learning.

**Geometric consistency and assessment.** External fields determine which transformations preserve physical dynamics (Han et al., 2022). We distinguish a joint scene-and-gravity consistency check from a slope intervention with fixed gravity; the history network is not constrained to be equivariant. Information estimates and disentanglement measures (Ross, 2014; Eastwood & Williams, 2018) complement our retrieval and predictive evaluations rather than define a single ranking of representations.

## 3 Method

Our method uses controlled simulator comparisons to define factor-specific response differences, then trains a reservoir-based history representation to predict them across rollouts (Figure 1).

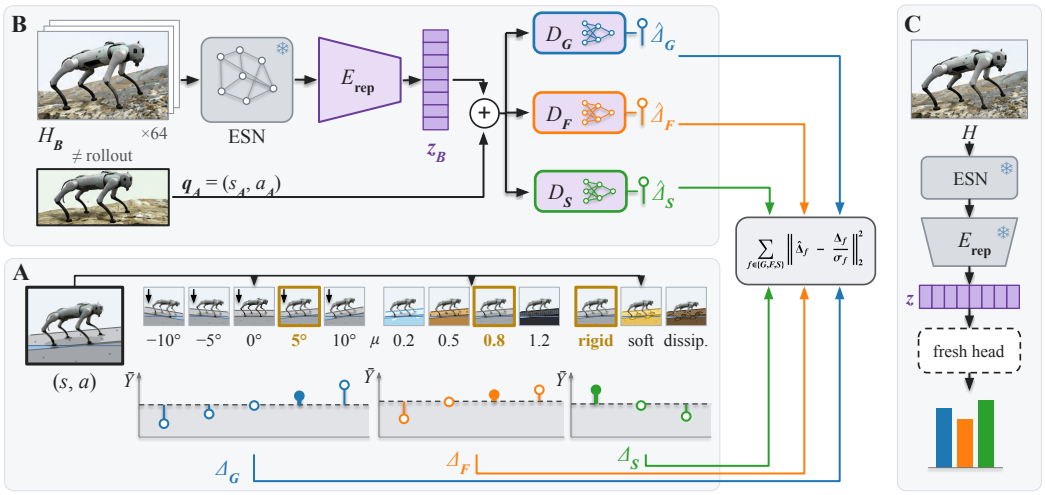


Figure 1: Counterfactual supervision and evaluation. (A) Factor-dependent response contrasts. (B) Cross-rollout prediction: the junction denotes query–code concatenation, and outputs are normalized as in Eq. (5). (C) Frozen-code factor probes; response predictors also receive a query. Targets and environment labels are not encoder inputs.

### 3.1 Factor-dependent counterfactual interventions

Let $\xi = (\alpha, \mu, \nu)$ denote signed terrain slope, friction coefficient, and contact-mechanics profile. An anchor contains state $s$ and action $a$. A matched intervention $\mathcal{I}_\xi(s, a)$ preserves the action and appropriate initial robot–environment configuration while varying one factor. Its one-step response $Y(s, a; \xi) \in \mathbb{R}^{36}$ uses body-frame coordinates (Appendix A).

**Material interventions.** We compare contact laws by changing friction or contact-profile parameters while preserving the robot state, supporting geometry, and action.

**Geometry interventions.** For a slope change from $\alpha_0$ to $\alpha$, let $R_T(\alpha)$ be the terrain rotation and define $Q = R_T(\alpha)R_T(\alpha_0)^\top$. With pivot $c$ on the source support plane, the branch applies

$$p' = c + Q(p - c), \qquad R'_B = QR_B, \qquad v'_W = Qv_W, \tag{1}$$

to the robot root and transforms the floor by the same rigid motion. We retain joint states, joint actions, and body-local angular velocity. This preserves the relative robot–terrain configuration while fixed world gravity changes the loading.

Rotating gravity with the scene instead gives an equivariance check: body-frame responses should agree up to simulator tolerance. This differs from the fixed-gravity slope intervention (Appendix A).

### 3.2 Factor-conditioned residual supervision

At each anchor, we collect responses over the slope, friction, and contact-profile sets $\mathcal{A}, \mathcal{M}, \mathcal{S}$. Holding two factors fixed gives the contrasts below; anchor arguments are omitted.

$$\Delta_G(\alpha, \mu, \nu) = Y(\alpha, \mu, \nu) - \frac{1}{|\mathcal{A}|} \sum_{\alpha' \in \mathcal{A}} Y(\alpha', \mu, \nu), \tag{2}$$

$$\Delta_F(\alpha, \mu, \nu) = Y(\alpha, \mu, \nu) - \frac{1}{|\mathcal{M}|} \sum_{\mu' \in \mathcal{M}} Y(\alpha, \mu', \nu), \tag{3}$$

$$\Delta_S(\alpha, \mu, \nu) = Y(\alpha, \mu, \nu) - \frac{1}{|\mathcal{S}|} \sum_{\nu' \in \mathcal{S}} Y(\alpha, \mu, \nu'). \tag{4}$$

For one anchor and fixed remaining factors, write the branch responses as $Y_i$, their mean as $\bar{Y}$, and their contrasts as $\Delta_i = Y_i - \bar{Y}$. Then

$$\frac{1}{M}\sum_{i=1}^{M}\Delta_i = 0, \qquad \Delta_i - \Delta_j = Y_i - Y_j.$$

Centering removes the branch-shared component while preserving pairwise differences. The contrasts can remain query-dependent, and factor interactions can enter several contrasts; they are not disjoint additive components. Held-out values are excluded from training branches and reference means. Their test contrasts use the retained training reference.

Removing shared structure changes a finite learner's approximation problem, while normalization reweights coordinates. This may help the code retain environmental variation, but creates no new physical information: a predictor with the exact reference $\bar{Y}$ can convert between full and residual responses. We test the combined design and examine centering separately with matched heads and scales.

### 3.3 Reservoir-based history representation

A donor history $H_B$ consists of $K = 64$ completed proprioceptive/action transitions. A receiving query $q_A \in \mathbb{R}^{49}$ describes terrain-relative body orientation, joint positions and velocities, body-frame velocities, the action, and forward-speed command. It omits explicit friction, surface labels, and world-frame gravity direction. Terrain metadata is used to form this canonical *prediction query*; it is not supplied as a numerical environment label to the history encoder.

A 512-unit echo-state network processes each history from a zero initial state. We concatenate its states at steps 16, 32, 48, and 64 with their temporal mean and standard deviation to form the feature bank $F(H_B) \in \mathbb{R}^{3072}$ (Eq. (12)). This retains several views of the history rather than only its final state. A nonlinear representation head and the response decoders are trained as

$$z_B = E_\theta(F(H_B)) \in \mathbb{R}^{32}, \qquad \widehat{\widetilde{\Delta}}_f = D_{\phi_f}(\widetilde{q}_A, z_B), \quad f \in \{G, F, S\}. \tag{5}$$

We draw donor histories from another rollout in the same environment cell, preferring a different anchor time when available and excluding pairs without an eligible donor. The encoder therefore does not receive the query's trajectory. Branch and pairing labels are training metadata, not inputs to $E_\theta$.

Here $\widetilde{q}_A$ is standardized and $\widetilde{\Delta}_f = \Delta_f/\sigma_f$ uses per-coordinate training scales. Representation training minimizes

$$\mathcal{L}_{\text{res}}(\theta, \phi) = \mathbb{E}_{(A,\xi,B)}\left[\sum_{f\in\{G,F,S\}} \frac{1}{36}\left\| D_{\phi_f}(\widetilde{q}_A, E_\theta(F(H_B))) - \frac{\Delta_f(A,\xi)}{\sigma_f}\right\|_2^2\right]. \tag{6}$$

We fit scales on permitted training data and separate validation and test rollouts. During evaluation, the representation weights remain frozen while each new code is computed from completed interaction history.

The recurrent weights are fixed; the learned mappings use the architecture in Appendix B. Appendix I reports separate development diagnostics motivating temporal aggregation.

### 3.4 Geometry information and predictive sufficiency

The code and query must together supply the target's required information. Positive and negative slopes can share a canonical query but yield different responses $y_+, y_-$. If the remaining input gives only steepness magnitude, both complete inputs coincide. Any common prediction $u$ satisfies

$$\frac{\|u - y_+\|^2 + \|u - y_-\|^2}{2} = \left\|u - \frac{y_+ + y_-}{2}\right\|^2 + \frac{\|y_+ - y_-\|^2}{4}. \tag{7}$$

The second term remains unless the inputs distinguish the responses. A code or query supplying direction is outside this restriction. Our codes are not constrained to magnitude alone; Section 5.5 examines this requirement alongside factor accessibility.

### 3.5 COMPARISON OBJECTIVES

*NSP (Next-Step Prediction)* predicts the entire 36-dimensional response with its own training normalization. *CaDM-style* combines forward prediction with a coordinate-consistent backward-response objective (Lee et al., 2020). *LDG-style* combines a variational transition objective, KL regularization, and a multi-positive contrastive term for histories from the same environment cell (Xu et al., 2026). *RMA-style* adapts privileged environment supervision (Kumar et al., 2021), using steepness magnitude, friction, and surface one-hot labels. *Random* uses the fixed reservoir and a randomly initialized, untrained representation head.

The original objectives share splits, sampled triples, update budget, and a 32-dimensional interface. Different heads and normalizations, including LDG's variance branch, make this a comparison of supervision designs rather than subtraction alone or equal computation.

The *NSP-Matched* control uses ECR's three heads and scales but predicts $Y(A,\xi)/\sigma_f$ rather than $[Y(A,\xi) - \bar{Y}_f(A,\xi)]/\sigma_f$, where $\bar{Y}_f$ is the sibling-factor mean. A fresh *ECR-Paired* model matches initialization, batches, preprocessing, optimizer, and budget. No additional target centering, target clipping, or arm-specific loss rescaling is used. Section 5.2.1 and Appendix H report the comparison.

## 4 EXPERIMENTAL PROTOCOL

We test factor accessibility, unseen-level organization, predictive utility, and the contribution of history.

**Simulation and data.** We use Unitree Go1 in MuJoCo Playground (Zakka et al., 2025), with MJX/Warp simulation. Slopes $\{-10, -5, 0, 5, 10\}^\circ$, friction coefficients $\{0.2, 0.5, 0.8, 1.2\}$, and contact profiles Rigid, Soft, and Soft+Dissipative define 60 environment cells. We split 10,950 anchors by rollout into 7,660/1,642/1,648 training/validation/test anchors. Factual cross-rollout prediction uses the 7,660/1,639/1,643 eligible anchors; representation diagnostics retain the full test set. All rollouts use the same robot.

**Training.** We train for at most 6000 updates with batches of 256 and AdamW (Loshchilov & Hutter, 2019), using initial rate $10^{-3}$ and cosine decay. Checkpoints minimize validation loss. Main factor-response and paired next-step prediction probes use 5000 updates. Sample SD describes three runs sharing one reservoir and split. Appendices B and H specify the protocols and the separate aggregate-only diagnostic's unverified budget.

**Information and retrieval.** We estimate normalized mutual information (NMI) with a continuous–discrete nearest-neighbour estimator (Ross, 2014). Ridge probes measure linear accessibility by $R^2$ for steepness and friction and accuracy for surface class. Nearest-neighbour retrieval excludes candidates from the query's rollout and tests exact-cell identity or a selected factor. Exact cells retain signed slope, while steepness metrics use $|\alpha|$. One-anchor-per-rollout and rollout-mean estimates assess sensitivity to repeated observations within trajectories.

**Unseen levels.** We hold out each of three steepness magnitudes, four friction levels, and three surface profiles in ten separate tests. Both signs are removed for a nonzero steepness holdout. Training anchors, donors, branches, reference means, preprocessing, and model selection use only permitted conditions. NMI is evaluated over *seen plus held-out* test histories. Held-out retrieval queries use a gallery containing seen and held-out histories from other test rollouts.

**Downstream prediction and history.** Fresh heads predict full transitions in-distribution and factor responses at unseen $5^\circ$, $10^\circ$, $\mu = 0.5$, and Soft+Dissipative conditions. Appendix E gives test populations. Factual probes predict the receiving rollout's branch; all-branch probes use sampled counterfactual branches. Query-only and parameter-conditioned heads are learned references. An even-response target tests geometry compatibility. One-transition and 64-step representation heads have 1,644,192 parameters, with different architectures and input processing.

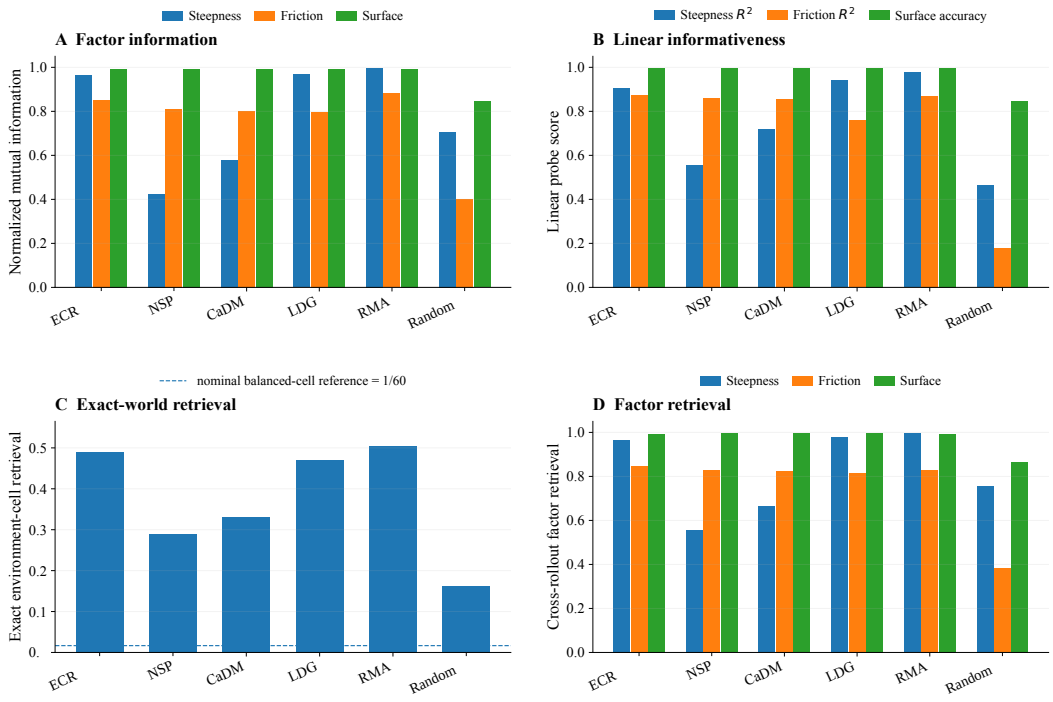


Figure 2: Representation diagnostics over three learned-head runs with one fixed reservoir. Steepness uses $|\alpha|$; exact cells retain signed slope. Panel C includes the nominal balanced-cell reference $1/60$. Table 10 reports rollout-aware NMI.

**Reporting.** We report MSE on normalized response coordinates and compare methods within each task, since different holdouts use different normalizations. Some in-distribution diagnostics report

$$\mathrm{Gap}(z) = 100\,\frac{\mathrm{MSE}(q) - \mathrm{MSE}(q, z)}{\mathrm{MSE}(q) - \mathrm{MSE}(q, \xi)}. \tag{8}$$

Values above 100% exceed a trained reference, not a theoretical bound. We also report raw MSE and available sample SD.

## 5 Results

### 5.1 Supervision changes accessible factor information

ECR makes terrain steepness more accessible than NSP or CaDM-style training under the common reservoir features (Figure 2). Steepness NMI is 0.962 for ECR, 0.425 for NSP, and 0.578 for CaDM. Linear-probe $R^2$ gives the same ordering: 0.907, 0.555, and 0.721. ECR also retains strong friction information, with $R^2 = 0.875$ versus 0.859 for NSP. Surface classification is close to saturation for all trained representations.

Reducing repeated observations within trajectories preserves this difference. We select one anchor from each of 414 test rollouts and average 100 selections per encoder run. Steepness NMI is 0.900 for ECR, 0.148 for NSP, and 0.348 for CaDM; rollout-mean codes corroborate the separation (Appendix C.2).

LDG has stronger steepness accessibility ($R^2 = 0.940$) but weaker friction accessibility ($R^2 = 0.758$); the privileged reference reaches steepness $R^2 = 0.977$. Random heads also expose information from temporal features. We test whether the predictor uses steepness context by replacing a matched donor with one from a different steepness. ECR's geometry-head error rises by $1.60\times$, versus $1.00\times$ for NSP and CaDM.

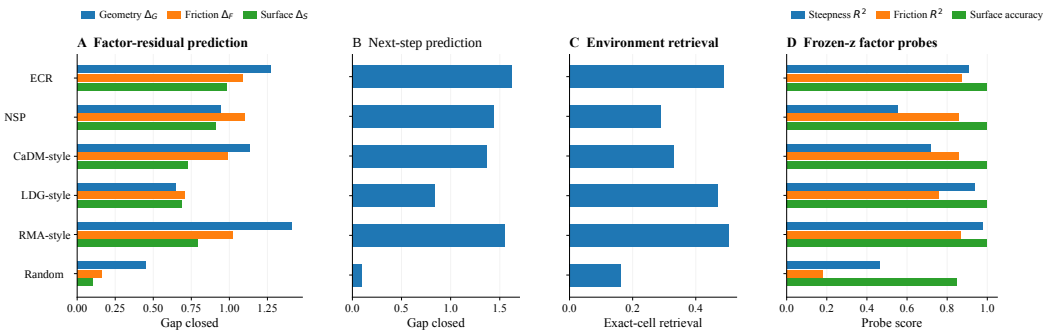


Figure 3: Cross-objective evaluation. (A,B) Factor-residual and next-step-prediction reference-gap fractions ($1 = 100\%$ in Eq. (8)). (C) Exact-cell retrieval. (D) Factor probes. Colors identify factors. Panel B uses the aggregate-only Table 9 diagnostic, separate from the paired study in Table 1.

Table 1: Paired centering ablation; mean $\pm$ sample SD over three runs. The first two rows pair encoder training; †NSP/CaDM reuse original encoders with new probes. Accessibility/prediction use 1,648/1,643 test anchors; fresh prediction heads train for 5,000 updates.

| Representation | Steepness $R^2$ | Cell retrieval | NSP MSE ↓ |
|---|---|---|---|
| ECR-Paired | $0.900 \pm 0.014$ | $0.498 \pm 0.015$ | $0.026610 \pm 0.000354$ |
| NSP-Matched | $0.380 \pm 0.017$ | $0.290 \pm 0.001$ | $0.028762 \pm 0.000178$ |
| NSP† | $0.555 \pm 0.081$ | $0.289 \pm 0.012$ | $0.027102 \pm 0.000089$ |
| CaDM† | $0.721 \pm 0.016$ | $0.331 \pm 0.021$ | $0.027291 \pm 0.000143$ |

### 5.2 Downstream prediction across objectives

The prediction differences are more modest. On factual residual prediction, ECR attains geometry MSE $0.2533 \pm 0.0128$ versus $0.2599 \pm 0.0130$ for NSP, a 2.6% difference in aggregate means. Friction errors are close ($0.2437$ versus $0.2417$), and ECR's mean surface error is slightly lower ($0.0410$ versus $0.0418$). Table 6 gives the complete comparison.

The paired study below tests centering with a fresh next-step-prediction task; Figure 3B retains the separate aggregate diagnostic in Table 9.

#### 5.2.1 Paired target-centering ablation

The accessibility gap persists with matched heads, scales, initialization, batches, and budget. ECR-Paired has higher steepness $R^2$, NMI, and exact-cell retrieval than NSP-Matched in all three pairs (Table 1). Fresh next-step prediction error is lower in each pair, with a 7.48% reduction in aggregate mean error.

Fresh 5000-update heads on the existing NSP and CaDM encoders provide context; their encoder training is not paired with the new ECR run. ECR's mean error is 1.8% lower than standard NSP's, with one run nearly tied. This is a modest in-distribution benefit beyond the residual training targets.

The uncentered targets are much larger, and gradient clipping occurs on every training update of the control versus 0.6–11.9% across ECR runs. The control's validation loss decreases, but optimization differs substantially. The comparison therefore includes centering's target-scale and finite-budget effects, not information filtering independently of optimization (Appendix H).

### 5.3 Representation transfer to unseen factor levels

The ten holdouts test how the representations organize histories from factor levels excluded during training (Figure 4). At unseen steepness magnitudes $0°$, $5°$, and $10°$, ECR's factor retrieval is 86.5%, 86.1%, and 94.2%, versus 52.6%, 58.9%, and 79.9% for NSP. ECR exceeds NSP and CaDM

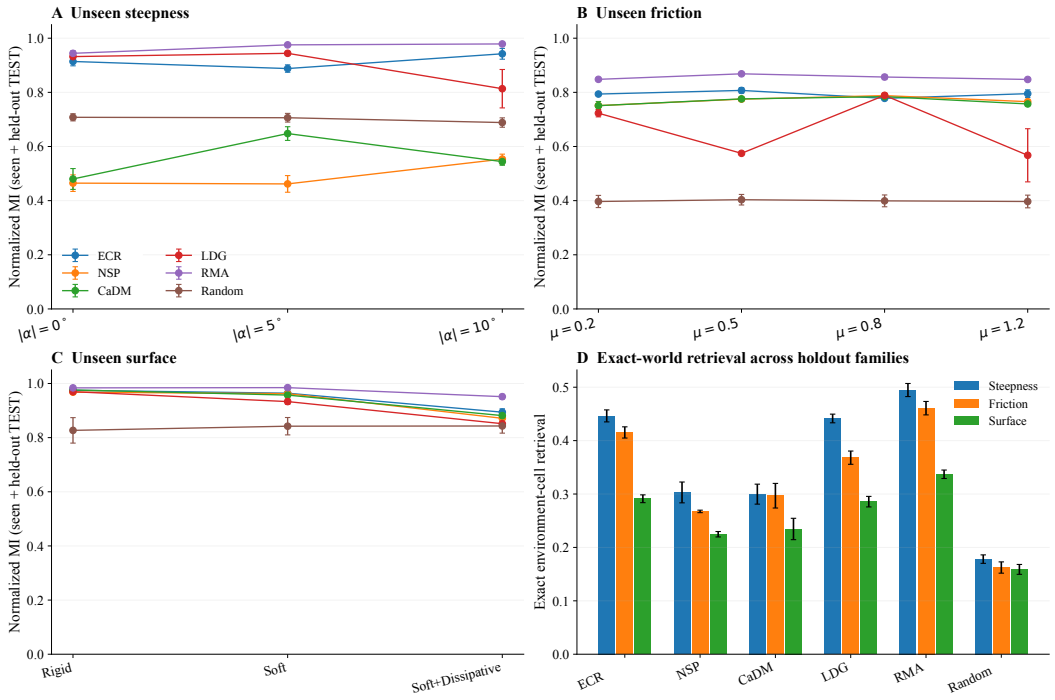


Figure 4: Leave-one-level-out evaluation. (A–C) NMI over seen-plus-held-out test histories. (D) Exact-cell retrieval aggregated within holdout families; held-out queries retrieve from other seen and held-out test rollouts. Error bars show sample SD over three learned-head runs.

Table 2: History versus local-information control. Retrieval is mean ± sample SD; physical errors are mean normalized MSE (full uncertainties in Appendix F). Both representation heads have the same trainable parameter count.

| Holdout | Objective | Retrieval: 1 step | Retrieval: 64 steps | MSE: 1 step | MSE: 64 steps |
|---|---|---|---|---|---|
| 5° | ECR | $0.550 \pm 0.022$ | $0.861 \pm 0.015$ | 0.1877 | 0.1601 |
| 5° | NSP | $0.450 \pm 0.012$ | $0.589 \pm 0.002$ | 0.1320 | 0.1241 |
| 10° | ECR | $0.706 \pm 0.025$ | $0.942 \pm 0.017$ | 2.4890 | 2.4578 |
| 10° | NSP | $0.517 \pm 0.012$ | $0.799 \pm 0.004$ | 1.9256 | 1.9356 |

in mean exact-cell retrieval in every holdout. LDG is stronger on steepness retrieval at 0° and 5°; the privileged-label reference is stronger on exact-cell retrieval throughout.

At $\mu = 0.5$, ECR and NSP both attain 75.6% friction retrieval, but exact-cell retrieval is 39.1% versus 20.5%. The advantage is consistent with retaining additional environmental information within unseen-friction histories. Factor-specific rankings remain mixed: NSP has higher friction retrieval at $\mu = 0.8$ and higher surface retrieval when Soft is withheld. Appendix D reports the complete sweep.

### 5.4 History contributes beyond a local transition

To test the contribution of history, we compare parameter-count-matched representation heads (Table 2). Extending ECR's input from one transition to 64 steps raises steepness retrieval from 55.0% to 86.1% at 5° and from 70.6% to 94.2% at 10°. Both improvements hold in all three runs. NSP also benefits, but has lower steepness retrieval.

History reduces ECR's mean directional-geometry MSE by 14.7% at 5° and 1.3% at 10°. NSP remains the better 64-step predictor. The different upstream features mean that this comparison tests information access and processing together, rather than isolating temporal order or RC.

Table 3: Held-out physical-response prediction. Entries are mean normalized MSE over three runs; lower is better. NSP and CaDM refer to matched objectives. All methods, references, and sample SDs appear in Appendix E.

| Held-out condition | Query only | ECR | NSP | CaDM |
|---|---|---|---|---|
| $\|\alpha\| = 5^\circ$ (geometry) | 0.1394 | 0.1601 | 0.1241 | 0.1235 |
| $\mu = 0.5$ (friction) | 0.3672 | 0.2836 | 0.2766 | 0.2654 |
| Soft+Dissipative | 1.4910 | 1.3678 | 1.3119 | 1.3992 |
| $\|\alpha\| = 10^\circ$ (geometry) | 2.0897 | 2.4578 | 1.9356 | 1.9450 |

### 5.5 Physical transfer and geometry compatibility

The held-out prediction tests give a different ordering (Table 3). ECR improves over query-only for friction and contact profile, but not directional geometry. NSP has lower mean error on all four factor-response tasks.

The paired-slope diagnostic illustrates a specific information requirement. Opposite slopes give a canonical-query RMS difference of $1.32 \times 10^{-8}$ but a normalized response difference of 1.2166, exposing the magnitude-only restriction in Section 3.4. A separate probe benefits from signed rather than magnitude-only conditioning (Figure 7; Appendix G). This does not establish that the learned codes omit direction.

The even-response diagnostic, $Y_{\text{even}}(r) = [Y(+r)+Y(-r)]/2$, tests a magnitude-compatible target with the encoders frozen. At $10^\circ$, ECR reaches MSE 14.699 versus 16.517 for NSP, with the random reference close at 14.916. At $5^\circ$, query-only is better than every conditioned model. These probes use different targets and populations from Table 3.

The contrast between in-distribution, held-out, and even-response results shows task-dependent utility. Tables 21–22 report both even-response settings; the geometry restriction alone does not explain their rankings.

## 6 Discussion and conclusion

Factor-dependent physical comparisons and a temporal reservoir feature bank serve complementary roles: the bank summarizes interaction history, while the targets determine what the learned code must predict. ECR improves steepness accessibility, including in the paired study, but predictive rankings depend on the task. Accessible factor information and downstream prediction must therefore be assessed together.

The main study covers one robot, engineered contacts, one reservoir realization, and a controlled grid. Transfer is across rollouts, not morphologies; earlier collection covers the broader domain. Architecture-development diagnostics use separate protocols, and evaluation concerns prediction rather than control. Centering's numerical and approximation effects remain intertwined, so superiority to an independently tuned uncentered learner is not established.

The results support counterfactual target construction as a way to shape accessible physical context. Its value must be assessed together with the temporal features supplied to the learner and the information required by the downstream task.

## AI USE STATEMENT

OpenAI ChatGPT assisted with drafting and organizing the manuscript, identifying and checking literature, discussing experimental controls and interpretation, formatting the result tables, and developing numerical-consistency checks and document-generation code. The experimental results reported here were supplied by the authors; no synthetic measurements were introduced during manuscript preparation. The authors are responsible for verification of the method, experimental implementation, analysis, references, and final claims.

## REPRODUCIBILITY STATEMENT

The main text defines the intervention construction, residual targets, and evaluation questions. Appendices A–B specify coordinate conventions, models, losses, training settings, and cross-rollout eligibility. Appendices C–H provide expanded numerical results, rollout-aware information checks, geometry diagnostics, and the paired ablation with optimization records. Appendix I reports separate reservoir-feature development diagnostics. Source-to-result records retain configuration, preprocessing, and seed identifiers. Full experimental reproduction additionally requires the simulator environment, raw rollout data, split manifests, and model checkpoints.

## ETHICS STATEMENT

The reported experiments use simulated robots and contain no human-subject data. Contact profiles are engineered simulation conditions rather than calibrated models of real terrain materials. Representation or prediction results in this setting should not be interpreted as evidence of safe hardware deployment. A physical deployment would require independent validation of sensing, dynamics shifts, and closed-loop failure modes.

## REFERENCES

Ossama Ahmed, Frederik Träuble, Anirudh Goyal, Alexander Neitz, Yoshua Bengio, Bernhard Schölkopf, Manuel Wüthrich, and Stefan Bauer. CausalWorld: A robotic manipulation benchmark for causal structure and transfer learning. arXiv:2010.04296, 2020. URL https://arxiv.org/abs/2010.04296.

Anonymous. Learning contact-dynamics representations through counterfactual residual supervision, 2026. Related manuscript under review.

Fabien Baradel, Natalia Neverova, Julien Mille, Greg Mori, and Christian Wolf. CoPhy: Counterfactual learning of physical dynamics. In *International Conference on Learning Representations*, 2020. URL https://arxiv.org/abs/1909.12000.

Cian Eastwood and Christopher K. I. Williams. A framework for the quantitative evaluation of disentangled representations. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=By-7dz-AZ.

Jiaqi Han, Wenbing Huang, Hengbo Ma, Jiachen Li, Joshua B. Tenenbaum, and Chuang Gan. Learning physical dynamics with subequivariant graph neural networks. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://arxiv.org/abs/2210.06876.

Nicklas A. Hansen, Hao Su, and Xiaolong Wang. Temporal difference learning for model predictive control. In *International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8387–8406, 2022. URL https://proceedings.mlr.press/v162/hansen22a.html.

Allen Hart, James Hook, and Jonathan Dawes. Embedding and approximation theorems for echo state networks. arXiv:1908.05202v2, 2020. URL https://arxiv.org/abs/1908.05202v2.

Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks. Technical Report 148, German National Research Center for Information Technology, 2001.

Steeven Janny, Fabien Baradel, Natalia Neverova, Madiha Nadri, Greg Mori, and Christian Wolf. Filtered-CoPhy: Unsupervised learning of counterfactual physics in pixel space. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2202.00368.

Tobias Johannink, Shikhar Bahl, Ashvin Nair, Jianlan Luo, Avinash Kumar, Matthias Loskyll, Juan Aparicio Ojea, Eugen Solowjow, and Sergey Levine. Residual reinforcement learning for robot control. In *IEEE International Conference on Robotics and Automation*, pp. 6023–6029, 2019.

Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems*, 2021. doi: 10.15607/RSS.2021.XVII.011.

Kimin Lee, Younggyo Seo, Seunghyun Lee, Honglak Lee, and Jinwoo Shin. Context-aware dynamics model for generalization in model-based reinforcement learning. In *International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5757–5766, 2020. URL https://proceedings.mlr.press/v119/lee20g.html.

Phillip Lippe, Sara Magliacane, Sindy Löwe, Yuki M. Asano, Taco Cohen, and Stratis Gavves. CITRIS: Causal identifiability from temporal intervened sequences. In *International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 13557–13603, 2022. URL https://proceedings.mlr.press/v162/lippe22a.html.

Phillip Lippe, Sara Magliacane, Sindy Löwe, Yuki M. Asano, Taco Cohen, and Efstratios Gavves. BISCUIT: Causal representation learning from binary interactions. In *Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 1263–1273, 2023. URL https://proceedings.mlr.press/v216/lippe23a.html.

Francesco Locatello, Ben Poole, Gunnar Raetsch, Bernhard Schölkopf, Olivier Bachem, and Michael Tschannen. Weakly-supervised disentanglement without compromises. In *International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6348–6359, 2020. URL https://proceedings.mlr.press/v119/locatello20a.html.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1711.05101.

Silviu Pitis, Elliot Creager, and Animesh Garg. Counterfactual data augmentation using locally factored dynamics. In *Advances in Neural Information Processing Systems*, volume 33, pp. 3976–3990, 2020.

Silviu Pitis, Elliot Creager, Ajay Mandlekar, and Animesh Garg. MoCoDA: Model-based counterfactual data augmentation. In *Advances in Neural Information Processing Systems*, volume 35, pp. 18143–18156, 2022.

Brian C. Ross. Mutual information between discrete and continuous data sets. *PLOS ONE*, 9(2): e87357, 2014. doi: 10.1371/journal.pone.0087357.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. arXiv:1807.03748, 2018. URL https://arxiv.org/abs/1807.03748.

Zhiming Xu, Weitao Zhou, Xianghui Pan, Nanshan Deng, Chengju Liu, Qijun Chen, and Chenpeng Yao. Dynamics are learned, not told: Semi-supervised discovery of latent dynamics geometries for zero-shot policy adaptation. arXiv:2606.02280, 2026. URL https://arxiv.org/abs/2606.02280.

Wenhao Yu, Jie Tan, C. Karen Liu, and Greg Turk. Preparing for the unknown: Learning a universal policy with online system identification. In *Robotics: Science and Systems*, 2017. doi: 10.15607/RSS.2017.XIII.048.

Zhihan Yue, Yujing Wang, Juanyong Duan, Tianmeng Yang, Congrui Huang, Yunhai Tong, and Bixiong Xu. TS2Vec: Towards universal representation of time series. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 8980–8987, 2022. doi: 10.1609/aaai.v36i8.20881. URL https://ojs.aaai.org/index.php/AAAI/article/view/20881.

Kevin Zakka, Baruch Tabanpour, Qiayuan Liao, Mustafa Haiderbhai, Samuel Holt, Jing Yuan Luo, Arthur Allshire, Erik Frey, Koushil Sreenath, Lueder A. Kahrs, Carmelo Sferrazza, Yuval Tassa, and Pieter Abbeel. MuJoCo playground. arXiv:2502.08844, 2025. URL https://arxiv.org/abs/2502.08844.

Amy Zhang, Rowan McAllister, Roberto Calandra, Yarin Gal, and Sergey Levine. Learning invariant representations for reinforcement learning without reconstruction. In *International Conference on Learning Representations*, 2021. URL https://arxiv.org/abs/2006.10742.

# A Coordinate conventions and counterfactual construction

## A.1 Terrain-relative matching

The simulator stores free-joint root position $p$, root orientation $R_B$, world-frame linear velocity $v_W$, body-local angular velocity $\omega_B$, and joint positions and velocities. To change terrain orientation, the source root is projected onto the source plane to obtain pivot $c$. The same rotation $Q = R_T(\alpha)R_T(\alpha_0)^\top$ is applied to the root position about $c$, the root orientation, the world-frame linear velocity, and the floor pose. The joint positions, joint velocities, and actuator command remain unchanged. Body-local angular velocity remains unchanged because its coordinate frame rotates with the body. The world-linear component of the acceleration warm start is rotated consistently.

The static floor pose is synchronized from the randomized model into the runtime geometry pose used for collision calculations. This is part of the branch construction rather than a learned correction. The configuration uses an elliptic friction cone, 100 solver iterations, 50 line-search iterations, and tolerance $10^{-10}$. The validity checks distinguish a coordinate change from a physical intervention: joint rotation of robot, terrain, and gravity tests response consistency; joint rotation of robot and terrain with world gravity fixed changes the physical loading. This distinction follows from the role of external fields in physical symmetry (Han et al., 2022).

## A.2 Query and response coordinates

Write $R = R_{WB}(t)$, $R^+ = R_{WB}(t+1)$, and $U = R^\top R^+$. The canonical query concatenates

$$q = \left[\mathrm{vec}_6(R_T^\top R),\ q_J,\ \dot{q}_J,\ R^\top v_W,\ \omega_B,\ a,\ c_x\right] \in \mathbb{R}^{49}, \tag{9}$$

where $\mathrm{vec}_6$ stacks the first two columns of the relative orientation. The component dimensions are $6, 12, 12, 3, 3, 12, 1$. The scalar $c_x$ is the forward-speed command. The query contains neither explicit environment labels nor a separate gravity-direction vector, although natural posture and motion may correlate with the environment. Forming the terrain-relative query uses simulator terrain information in the training and prediction-evaluation pipeline.

The response convention is

$$Y = \left[R^\top(p^+ - p),\ \log(U)^\vee,\ q_J^+ - q_J,\ R^\top(v_W^+ - v_W),\ U\omega_B^+ - \omega_B,\ \dot{q}_J^+ - \dot{q}_J\right] \in \mathbb{R}^{36}. \tag{10}$$

The six blocks have dimensions $3, 3, 12, 3, 3, 12$. A global rigid rotation of the complete physical experiment leaves these body-frame quantities unchanged. The backward response used by CaDM-style training must be expressed in the next body frame: vector increments are rotated by $U^\top$ and negated, while joint increments are negated. It is therefore not generally the negative of the forward 36-vector in the original frame.

## A.3 Contact profiles

Rigid retains the base model's floor contact parameters. Let its positive time constant and damping ratio be $t_0, d_0$, and let its five impedance parameters be $b_0, \ldots, b_4$. Soft uses time constant $\max(3t_0, 0.06)$ and damping ratio $d_0$. Soft+Dissipative uses the same time constant and damping ratio $\max(2d_0, 2)$. Both soft profiles use impedance entries $\min(b_0, 0.70)$, $\min(b_1, 0.85)$, and $\max(b_2, 0.003)$, retaining the base midpoint and power. The contact profiles are therefore solref $= (0.02, 1)$, $(0.06, 1)$, and $(0.06, 2)$ for Rigid, Soft, and Soft+Dissipative, respectively. Their impedance vectors are $(0.9, 0.95, 0.001, 0.5, 2)$ for Rigid and $(0.7, 0.85, 0.003, 0.5, 2)$ for both soft profiles. These are engineered simulator profiles, not measurements of particular real materials.

# B Representation and optimization details

## B.1 Fixed recurrent features

For standardized history token $x_k$, the leaky reservoir updates as

$$r_k = (1-\ell)r_{k-1} + \ell\tanh(W_{\mathrm{in}}x_k + W_{\mathrm{res}}r_{k-1} + b), \qquad r_0 = 0. \tag{11}$$

Table 4: Shared recurrent features and deterministic representation/prediction networks.

| Component | Configuration |
|---|---|
| History | 64 completed transitions |
| Reservoir | 512 units; tanh; seed 777 |
| Leak / spectral-radius setting | 0.05 / 0.9 |
| Input scale / recurrent density | 0.5 / 0.1 |
| Bias scale | 0.1 |
| Temporal taps | 16, 32, 48, 64 |
| Feature dimension | 3072 (four taps, mean, standard deviation) |
| Representation head | $3072 \to 512 \to 128 \to 32$ |
| Nonlinearities | SiLU; LayerNorm after first hidden activation |
| Trainable representation parameters | 1,644,192 |
| Response head | $(49 + 32) \to 256 \to 256 \to 256 \to 36$ |
| Response-head activations | SiLU between hidden layers |
| Local-information head | $79 \to 1228 \to 1224 \to 32$ |
| Local-head normalization | SiLU; LayerNorm after first hidden activation |
| Local trainable parameters | 1,644,192 |

The reservoir parameters are fixed throughout training. The implementation processes each 64-step window separately and concatenates

$$F(H) = [r_{16}, r_{32}, r_{48}, r_{64}, \mathrm{mean}_k\, r_k, \mathrm{std}_k\, r_k] \in \mathbb{R}^{3072}. \tag{12}$$

The temporal standard deviation uses the sample convention. Feature means and scales are fitted on training anchors. A numerical offset of $10^{-5}$ is added to feature scales. Fixed recurrent weights avoid optimizing the recurrence, but the implemented window-restarted feature calculation is not the same as a benchmark of one streaming update per control step.

The one-transition control takes the most recent completed 79-dimensional transition token and has no access to older tokens or reservoir features. Its wider hidden layers match the deterministic representation head's parameter count. The comparison controls trainable capacity, not computation or identical architectural inductive bias.

### B.2 Learning objectives and normalizations

ECR training sums the three normalized mean-squared errors in Eq. (6). Each contrast uses a per-coordinate training standard deviation with floor $10^{-6}$. NSP targets use their own training mean and scale. Query preprocessing uses training query statistics; histories and ESN features are normalized using permitted training anchors. When a factor level is held out, every reference mean is formed over retained levels. For example,

$$\Delta^*_G(A, \alpha^*, \mu, \nu) = Y(A, \alpha^*, \mu, \nu) - \frac{1}{|\mathcal{A}_{\mathrm{train}}|} \sum_{\alpha \in \mathcal{A}_{\mathrm{train}}} Y(A, \alpha, \mu, \nu). \tag{13}$$

This fixes the reference distribution independently of the unseen target value.

CaDM-style training uses equal weights for its forward and backward normalized MSEs. LDG-style training uses a Gaussian latent with log standard deviation clamped to $[-6, 2]$. The transition decoder receives a LayerNorm-normalized sampled latent during training and the mean latent during evaluation. Its objective is

$$\mathcal{L}_{\mathrm{LDG}} = 5\,(\mathcal{L}_{\mathrm{transition}} + \beta(u)\mathcal{L}_{\mathrm{KL}}) + \mathcal{L}_{\mathrm{contrast}}, \qquad \beta(u) = 0.07 \min(1, 2u), \tag{14}$$

where $u$ is the fraction of completed updates. The contrastive term uses normalized latent means, temperature 1, and multiple positive histories from the same environment cell. An additional same-cell history is drawn from a different rollout for its positive pool. The Gaussian head predicts both mean and scale while the saved evaluation representation retains only the 32-dimensional mean. This extra variance branch is not present in the deterministic objectives. Policy optimization from the full LDG system is not part of this comparison.

The privileged-label reference predicts $[|\alpha|/10, \mu, \mathrm{onehot}(\nu)]$ from $z$ using a five-output regression head and MSE. The decoder is trained, so this reference is not an information-theoretic ceiling. The

Table 5: Optimization settings for the main objectives, holdouts, and paired study. The earlier aggregate-only next-step-prediction diagnostic (Table 9) has an unverified executed probe budget.

| Setting | Value |
|---|---|
| Maximum representation updates | 6000 |
| Main factor / paired next-step-prediction probe updates | 5000 |
| Batch size | 256 |
| Optimizer / initial learning rate | AdamW / $10^{-3}$ |
| Learning-rate schedule | Cosine decay over the update budget |
| Representation weight decay | $10^{-4}$ |
| Prediction-head weight decay | $10^{-5}$ |
| Gradient norm clip | 1.0 |
| Validation selection | Lowest own-objective validation loss |
| Validation frequency | Twelve intervals over training |
| Independent learned-head runs | 3 (same ESN realization and split) |
| Holdout training seeds | 67001, 67102, 67203 |

RMA-style name denotes the use of privileged environment supervision; its labels and architecture are those specified here, not a reproduction of the original RMA teacher–student controller.

### B.3 Pairing, evaluation populations, and uncertainty

Training samples consist of an anchor, an eligible environmental branch, and a donor history from that branch's environment cell. Donor and receiving rollout identities must differ; a different anchor time is preferred when available. The sampler raises an error if no cross-rollout donor can be formed rather than using the receiving rollout. Branches without eligible donor cells are excluded or resampled. Complete rollout identities define all data partitions, and corresponding objectives use matching recorded training-triplet fingerprints.

The main split contains 7,660 training, 1,642 validation, and 1,648 test anchors. Factual donor-dependent prediction uses 7,660, 1,639, and 1,643 eligible anchors, respectively. The remaining validation/test anchors have no same-cell donor from a different rollout in their partition. Representation-information diagnostics do not require a receiving-query donor and retain all 1,648 test anchors. These populations are reported separately throughout.

A factual probe predicts the receiving rollout's own branch response. An all-branch probe is separately fitted on sampled counterfactual branches. The two protocols have different target populations and are not interchangeable. Holdout probes fit only permitted training levels and use held-out test rollouts for evaluation. No held-out example is used to optimize the frozen representation or the prediction head.

Sample standard deviations in the tables use $n - 1$ in the denominator. Three learned-head runs do not provide three independent reservoir realizations. Pairwise differences use corresponding runs where reported. Aggregate NMI and ridge values in the main diagnostics are retained as descriptive estimates; their estimation populations and representation objective matter when comparing their magnitude.

## C Matched prediction and representation diagnostics

### C.1 Response prediction and history substitution

The common test evaluates each learned representation with fresh response heads. The next-step prediction head and factor-residual heads are different prediction tasks. Their raw errors are not measured on the same normalized scale. For in-distribution geometry, gap closure uses the magnitude-conditioned reference; for friction and surface it uses the corresponding parameter-conditioned reference. A gap exceeding 100% indicates a better trained predictor than this finite-capacity reference, not recovery of more than all the environmental information.

Table 6: Factual residual prediction on test rollouts. Entries are normalized MSE, mean ± sample SD over three learned-head runs. Geometry refers to the directional residual, not the even-target diagnostic.

| Objective | Geometry ↓ | Friction ↓ | Surface ↓ |
|---|---|---|---|
| ECR | $0.2533 \pm 0.0128$ | $0.2437 \pm 0.0096$ | $0.0410 \pm 0.0019$ |
| NSP | $0.2599 \pm 0.0130$ | $0.2417 \pm 0.0099$ | $0.0418 \pm 0.0032$ |
| CaDM | $0.2561 \pm 0.0055$ | $0.2588 \pm 0.0110$ | $0.0437 \pm 0.0033$ |
| LDG | $0.2658 \pm 0.0171$ | $0.3019 \pm 0.0092$ | $0.0442 \pm 0.0032$ |
| RMA | $0.2506 \pm 0.0058$ | $0.2543 \pm 0.0069$ | $0.0431 \pm 0.0014$ |
| Random | $0.2698 \pm 0.0035$ | $0.3860 \pm 0.0041$ | $0.0505 \pm 0.0012$ |

Table 7: All-branch counterfactual response prediction with separately fitted probes. Normalized MSE, mean ± sample SD over three runs. Signed geometry requires information not specified by the magnitude-only reference.

| Input / representation | Geometry | Friction | Surface |
|---|---|---|---|
| Query only | $0.9690 \pm 0.0000$ | $0.9223 \pm 0.0000$ | $0.9681 \pm 0.0000$ |
| Signed-parameter reference | $0.1672 \pm 0.0004$ | $0.2532 \pm 0.0031$ | $0.0723 \pm 0.0009$ |
| Magnitude-based reference | $0.9516 \pm 0.0001$ | $0.2527 \pm 0.0026$ | $0.0721 \pm 0.0010$ |
| ECR | $0.9781 \pm 0.0287$ | $0.3185 \pm 0.0030$ | $0.0702 \pm 0.0010$ |
| NSP | $0.9626 \pm 0.0010$ | $0.3210 \pm 0.0017$ | $0.0701 \pm 0.0011$ |
| CaDM | $0.9573 \pm 0.0004$ | $0.3151 \pm 0.0010$ | $0.0703 \pm 0.0006$ |
| LDG | $0.9520 \pm 0.0002$ | $0.4188 \pm 0.0083$ | $0.0858 \pm 0.0013$ |
| RMA | $0.9518 \pm 0.0002$ | $0.3201 \pm 0.0045$ | $0.0710 \pm 0.0005$ |
| Random | $0.9580 \pm 0.0014$ | $0.7357 \pm 0.0405$ | $0.2381 \pm 0.0190$ |

Table 7 reports directional geometry for the all-branch population. Query-only, magnitude-based conditioning, and the learned codes have much higher geometry error than the signed-parameter reference. This contrasts with factual prediction and is one reason to examine the information contract of the receiving task rather than compare normalized scores from different populations.

Table 8: Named-factor history substitution. Mean ratio of wrong-factor-donor error to a same-environment cross-rollout control. Values above one show dependence on the substituted history information.

| Objective | Steepness to geometry | Friction to friction | Surface to surface |
|---|---|---|---|
| ECR | 1.60 | 5.51 | 20.92 |
| NSP | 1.00 | 5.16 | 20.96 |
| CaDM | 1.00 | 4.68 | 19.88 |
| LDG | 1.35 | 3.02 | 16.19 |
| RMA | 2.29 | 5.05 | 20.67 |
| Random | 1.02 | 1.05 | 1.27 |

For history substitution, the denominator is the error using a same-environment donor from another rollout. A donor differing in one named factor is substituted while the receiving query is held fixed. The diagonal ratio is the error on that factor's head divided by the control error. It tests dependence on the substituted factor. Friction and surface substitutions have strong effects for several learned representations; the geometry contrast separates ECR from NSP and CaDM more clearly.

Table 9 is a separate aggregate-only next-step prediction evaluation, with its own normalization and fitted references. ECR has the lowest recorded mean error, but the per-run errors and executed head-update budget are unavailable; the surviving 3000-update code default does not establish that budget. Its gap percentages are not percentage error reductions. The documented 5000-update study in Table 1 provides separate evidence, not retrospective uncertainty for this fit.

Table 9: Earlier aggregate-only next-step-prediction diagnostic on frozen representations. Entries are normalized MSE and reference-gap percentages; per-run uncertainty and the executed head-update budget are unavailable. This fit is separate from Table 1.

| Objective | Normalized MSE | Gap |
|---|---|---|
| ECR | 0.030802 | 162.0% |
| NSP | 0.031352 | 144.4% |
| CaDM | 0.031576 | 137.3% |
| LDG | 0.033260 | 83.6% |
| RMA | 0.031021 | 155.0% |
| Random | 0.035589 | 9.3% |

### C.2 Rollout-aware information estimation

The main diagnostic uses 1,648 anchors from 414 test rollouts. To assess sensitivity to repeated observations within a trajectory, we evaluate each frozen code in three ways: all anchors, one randomly selected anchor per rollout, and one code averaged within each rollout. The same continuous–discrete estimator uses $k = 3$ in each case. All factors are constant within the evaluated rollouts.

For the one-anchor test, we make 100 random selections for each of the three learned-head runs. The entry in Table 10 first averages these selections within each run, then reports the mean and sample SD across the three run means. The 100 selections measure sample-selection sensitivity; they are not additional training replications. The all-anchor and rollout-mean columns similarly summarize the three frozen representations. ECR's steepness-information advantage over NSP and CaDM persists

Table 10: Steepness NMI under rollout-aware populations. Mean $\pm$ sample SD across three learned-head runs. For each run, the one-anchor entry first averages 100 random selections, each using one anchor from every test rollout.

| Objective | All anchors | One anchor/rollout | Rollout-mean code |
|---|---|---|---|
| ECR | $0.9625 \pm 0.0150$ | $0.8998 \pm 0.0350$ | $0.9590 \pm 0.0224$ |
| NSP | $0.4246 \pm 0.0165$ | $0.1478 \pm 0.0225$ | $0.2407 \pm 0.0200$ |
| CaDM | $0.5777 \pm 0.0462$ | $0.3481 \pm 0.0376$ | $0.4710 \pm 0.0301$ |
| LDG | $0.9675 \pm 0.0104$ | $0.9555 \pm 0.0101$ | $0.9845 \pm 0.0110$ |
| RMA | $0.9973 \pm 0.0010$ | $0.9891 \pm 0.0032$ | $0.9999 \pm 0.0011$ |
| Random | $0.7061 \pm 0.0148$ | $0.4861 \pm 0.0077$ | $0.7041 \pm 0.0311$ |

under both rollout-aware populations. LDG and the privileged-label reference remain strong on steepness. These estimates complement the independently fitted linear probes and the cross-rollout retrieval tests; their numerical values describe estimated dependence in the specified populations.

### C.3 Additional representation diagnostics

DCI diagnostics do not give ECR the strongest coordinate-wise disentanglement score: reported mean disentanglement is 0.239 for ECR, 0.438 for NSP, 0.326 for CaDM, 0.245 for LDG, 0.167 for the privileged-label reference, and 0.127 for the random head. Likewise, global ordering is close to perfect for many objectives, including random features. We therefore distinguish factor accessibility and local ordering from a claim of coordinate-wise disentanglement.

## D Complete leave-one-level-out results

The following tables contain all ten designated holdouts. Steepness holdouts remove both signs of a nonzero magnitude. An interior numerical level tests interpolation, whereas an edge level tests extension beyond the retained levels. Surface profiles are categorical: their numeric identifiers do not impose a meaningful physical order.

The NMI population combines seen and held-out test histories. For retrieval, held-out query histories use a gallery of seen and held-out test histories, excluding the query's own rollout. The exact-

cell metric retains signed slope. Factor retrieval collapses slope to magnitude when steepness is evaluated. Thus an unseen cell may be matched to an example of the same cell in another test rollout without that cell having been available during training. The metric measures transferable organization under this gallery protocol, not a label predictor trained without examples of its test category.

Held-out test-anchor counts are 364, 672, and 612 for $0^\circ$, $5^\circ$, and $10^\circ$; 344, 412, 448, and 444 for friction 0.2, 0.5, 0.8, and 1.2; and 532, 639, and 477 for the three surface profiles. The eligible gallery determines the chance reference; missing cells and exclusion of same-rollout neighbours make a single $1/60$ value only a nominal balanced-cell reference.

Table 11: Steepness holdouts. Mean $\pm$ sample SD. NMI uses seen-plus-held-out test histories; retrieval uses held-out queries and a mixed test gallery. Ordering is not defined for categorical surfaces.

| Held out | Objective | NMI | Factor retrieval | Cell retrieval | Ordering |
|---|---|---|---|---|---|
| $\lvert\alpha\rvert = 0^\circ$ | ECR | $0.914 \pm 0.016$ | $0.865 \pm 0.037$ | $0.727 \pm 0.040$ | $1.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 0^\circ$ | NSP | $0.465 \pm 0.031$ | $0.526 \pm 0.066$ | $0.466 \pm 0.040$ | $0.273 \pm 0.158$ |
| $\lvert\alpha\rvert = 0^\circ$ | CaDM | $0.480 \pm 0.038$ | $0.521 \pm 0.035$ | $0.452 \pm 0.033$ | $0.424 \pm 0.278$ |
| $\lvert\alpha\rvert = 0^\circ$ | LDG | $0.932 \pm 0.009$ | $0.880 \pm 0.025$ | $0.710 \pm 0.004$ | $1.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 0^\circ$ | RMA | $0.944 \pm 0.009$ | $0.909 \pm 0.021$ | $0.791 \pm 0.011$ | $1.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 0^\circ$ | Random | $0.708 \pm 0.013$ | $0.766 \pm 0.021$ | $0.286 \pm 0.017$ | $0.970 \pm 0.052$ |
| $\lvert\alpha\rvert = 5^\circ$ | ECR | $0.888 \pm 0.014$ | $0.861 \pm 0.015$ | $0.360 \pm 0.015$ | $1.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 5^\circ$ | NSP | $0.462 \pm 0.031$ | $0.589 \pm 0.002$ | $0.256 \pm 0.022$ | $0.273 \pm 0.091$ |
| $\lvert\alpha\rvert = 5^\circ$ | CaDM | $0.648 \pm 0.025$ | $0.646 \pm 0.007$ | $0.255 \pm 0.014$ | $0.394 \pm 0.139$ |
| $\lvert\alpha\rvert = 5^\circ$ | LDG | $0.944 \pm 0.007$ | $0.946 \pm 0.012$ | $0.380 \pm 0.011$ | $0.879 \pm 0.139$ |
| $\lvert\alpha\rvert = 5^\circ$ | RMA | $0.975 \pm 0.002$ | $0.976 \pm 0.003$ | $0.426 \pm 0.006$ | $1.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 5^\circ$ | Random | $0.706 \pm 0.016$ | $0.660 \pm 0.034$ | $0.126 \pm 0.002$ | $0.939 \pm 0.105$ |
| $\lvert\alpha\rvert = 10^\circ$ | ECR | $0.942 \pm 0.020$ | $0.942 \pm 0.017$ | $0.252 \pm 0.028$ | $0.909 \pm 0.000$ |
| $\lvert\alpha\rvert = 10^\circ$ | NSP | $0.553 \pm 0.018$ | $0.799 \pm 0.004$ | $0.187 \pm 0.003$ | $0.000 \pm 0.000$ |
| $\lvert\alpha\rvert = 10^\circ$ | CaDM | $0.544 \pm 0.013$ | $0.795 \pm 0.030$ | $0.193 \pm 0.016$ | $0.061 \pm 0.052$ |
| $\lvert\alpha\rvert = 10^\circ$ | LDG | $0.814 \pm 0.071$ | $0.860 \pm 0.033$ | $0.235 \pm 0.011$ | $0.151 \pm 0.052$ |
| $\lvert\alpha\rvert = 10^\circ$ | RMA | $0.979 \pm 0.004$ | $0.972 \pm 0.010$ | $0.267 \pm 0.026$ | $0.667 \pm 0.501$ |
| $\lvert\alpha\rvert = 10^\circ$ | Random | $0.689 \pm 0.017$ | $0.883 \pm 0.022$ | $0.123 \pm 0.009$ | $0.909 \pm 0.091$ |

Table 12: Friction holdouts. Mean ± sample SD. NMI uses seen-plus-held-out test histories; retrieval uses held-out queries and a mixed test gallery. Ordering is not defined for categorical surfaces.

| Held out | Objective | NMI | Factor retrieval | Cell retrieval | Ordering |
|---|---|---|---|---|---|
| $\mu = 0.2$ | ECR | $0.794 \pm 0.004$ | $0.884 \pm 0.006$ | $0.448 \pm 0.011$ | $0.795 \pm 0.089$ |
| $\mu = 0.2$ | NSP | $0.751 \pm 0.004$ | $0.894 \pm 0.015$ | $0.355 \pm 0.026$ | $0.846 \pm 0.000$ |
| $\mu = 0.2$ | CaDM | $0.751 \pm 0.015$ | $0.877 \pm 0.024$ | $0.354 \pm 0.009$ | $0.846 \pm 0.000$ |
| $\mu = 0.2$ | LDG | $0.723 \pm 0.013$ | $0.776 \pm 0.008$ | $0.368 \pm 0.029$ | $0.436 \pm 0.160$ |
| $\mu = 0.2$ | RMA | $0.848 \pm 0.002$ | $0.910 \pm 0.013$ | $0.485 \pm 0.019$ | $0.846 \pm 0.000$ |
| $\mu = 0.2$ | Random | $0.397 \pm 0.022$ | $0.532 \pm 0.029$ | $0.208 \pm 0.020$ | $0.179 \pm 0.044$ |
| $\mu = 0.5$ | ECR | $0.807 \pm 0.010$ | $0.756 \pm 0.011$ | $0.391 \pm 0.007$ | $0.974 \pm 0.044$ |
| $\mu = 0.5$ | NSP | $0.775 \pm 0.008$ | $0.756 \pm 0.018$ | $0.205 \pm 0.020$ | $0.923 \pm 0.000$ |
| $\mu = 0.5$ | CaDM | $0.776 \pm 0.003$ | $0.752 \pm 0.007$ | $0.242 \pm 0.028$ | $0.897 \pm 0.044$ |
| $\mu = 0.5$ | LDG | $0.575 \pm 0.002$ | $0.594 \pm 0.028$ | $0.353 \pm 0.020$ | $0.667 \pm 0.044$ |
| $\mu = 0.5$ | RMA | $0.868 \pm 0.005$ | $0.765 \pm 0.009$ | $0.408 \pm 0.013$ | $0.923 \pm 0.000$ |
| $\mu = 0.5$ | Random | $0.403 \pm 0.019$ | $0.288 \pm 0.021$ | $0.098 \pm 0.011$ | $0.282 \pm 0.117$ |
| $\mu = 0.8$ | ECR | $0.778 \pm 0.002$ | $0.652 \pm 0.020$ | $0.413 \pm 0.028$ | $0.923 \pm 0.000$ |
| $\mu = 0.8$ | NSP | $0.788 \pm 0.007$ | $0.703 \pm 0.032$ | $0.265 \pm 0.013$ | $0.923 \pm 0.000$ |
| $\mu = 0.8$ | CaDM | $0.784 \pm 0.005$ | $0.669 \pm 0.006$ | $0.316 \pm 0.044$ | $0.923 \pm 0.000$ |
| $\mu = 0.8$ | LDG | $0.789 \pm 0.008$ | $0.681 \pm 0.036$ | $0.437 \pm 0.012$ | $0.846 \pm 0.077$ |
| $\mu = 0.8$ | RMA | $0.857 \pm 0.009$ | $0.741 \pm 0.024$ | $0.487 \pm 0.014$ | $0.923 \pm 0.000$ |
| $\mu = 0.8$ | Random | $0.399 \pm 0.022$ | $0.314 \pm 0.022$ | $0.158 \pm 0.008$ | $0.333 \pm 0.194$ |
| $\mu = 1.2$ | ECR | $0.795 \pm 0.014$ | $0.740 \pm 0.011$ | $0.410 \pm 0.026$ | $0.923 \pm 0.000$ |
| $\mu = 1.2$ | NSP | $0.766 \pm 0.006$ | $0.723 \pm 0.011$ | $0.245 \pm 0.036$ | $0.923 \pm 0.000$ |
| $\mu = 1.2$ | CaDM | $0.757 \pm 0.006$ | $0.722 \pm 0.009$ | $0.276 \pm 0.020$ | $0.923 \pm 0.000$ |
| $\mu = 1.2$ | LDG | $0.568 \pm 0.098$ | $0.531 \pm 0.102$ | $0.315 \pm 0.059$ | $0.385 \pm 0.077$ |
| $\mu = 1.2$ | RMA | $0.848 \pm 0.004$ | $0.719 \pm 0.023$ | $0.463 \pm 0.039$ | $0.923 \pm 0.000$ |
| $\mu = 1.2$ | Random | $0.397 \pm 0.023$ | $0.429 \pm 0.006$ | $0.185 \pm 0.007$ | $0.282 \pm 0.089$ |

Table 13: Surface holdouts. Mean ± sample SD. NMI uses seen-plus-held-out test histories; retrieval uses held-out queries and a mixed test gallery. Ordering is not defined for categorical surfaces.

| Held out | Objective | NMI | Factor retrieval | Cell retrieval | Ordering |
|---|---|---|---|---|---|
| Rigid | ECR | $0.975 \pm 0.005$ | $0.979 \pm 0.006$ | $0.227 \pm 0.013$ | – |
| Rigid | NSP | $0.968 \pm 0.007$ | $0.965 \pm 0.007$ | $0.195 \pm 0.013$ | – |
| Rigid | CaDM | $0.976 \pm 0.001$ | $0.975 \pm 0.006$ | $0.208 \pm 0.027$ | – |
| Rigid | LDG | $0.970 \pm 0.006$ | $0.954 \pm 0.008$ | $0.221 \pm 0.012$ | – |
| Rigid | RMA | $0.984 \pm 0.001$ | $0.977 \pm 0.006$ | $0.234 \pm 0.003$ | – |
| Rigid | Random | $0.827 \pm 0.047$ | $0.948 \pm 0.005$ | $0.173 \pm 0.007$ | – |
| Soft | ECR | $0.964 \pm 0.001$ | $0.953 \pm 0.004$ | $0.335 \pm 0.021$ | – |
| Soft | NSP | $0.963 \pm 0.004$ | $0.977 \pm 0.002$ | $0.258 \pm 0.016$ | – |
| Soft | CaDM | $0.957 \pm 0.003$ | $0.963 \pm 0.003$ | $0.265 \pm 0.031$ | – |
| Soft | LDG | $0.933 \pm 0.010$ | $0.940 \pm 0.003$ | $0.366 \pm 0.011$ | – |
| Soft | RMA | $0.985 \pm 0.001$ | $0.976 \pm 0.002$ | $0.413 \pm 0.010$ | – |
| Soft | Random | $0.842 \pm 0.032$ | $0.848 \pm 0.012$ | $0.188 \pm 0.008$ | – |
| Soft+Diss. | ECR | $0.894 \pm 0.013$ | $0.854 \pm 0.030$ | $0.312 \pm 0.015$ | – |
| Soft+Diss. | NSP | $0.871 \pm 0.013$ | $0.814 \pm 0.023$ | $0.222 \pm 0.005$ | – |
| Soft+Diss. | CaDM | $0.881 \pm 0.017$ | $0.834 \pm 0.024$ | $0.231 \pm 0.038$ | – |
| Soft+Diss. | LDG | $0.852 \pm 0.006$ | $0.769 \pm 0.027$ | $0.270 \pm 0.013$ | – |
| Soft+Diss. | RMA | $0.951 \pm 0.009$ | $0.884 \pm 0.023$ | $0.364 \pm 0.016$ | – |
| Soft+Diss. | Random | $0.843 \pm 0.026$ | $0.827 \pm 0.062$ | $0.115 \pm 0.015$ | – |

## E FOCUSED PHYSICAL-RESPONSE TRANSFER

The encoders are the frozen checkpoints from the corresponding holdout experiments. Fresh factual prediction heads use retained training conditions and are scored on held-out source-condition test anchors. The target is the directional geometry residual at the two steepness holdouts, the friction residual at $\mu = 0.5$, and the surface residual at Soft+Dissipative. Eligible donor-dependent test

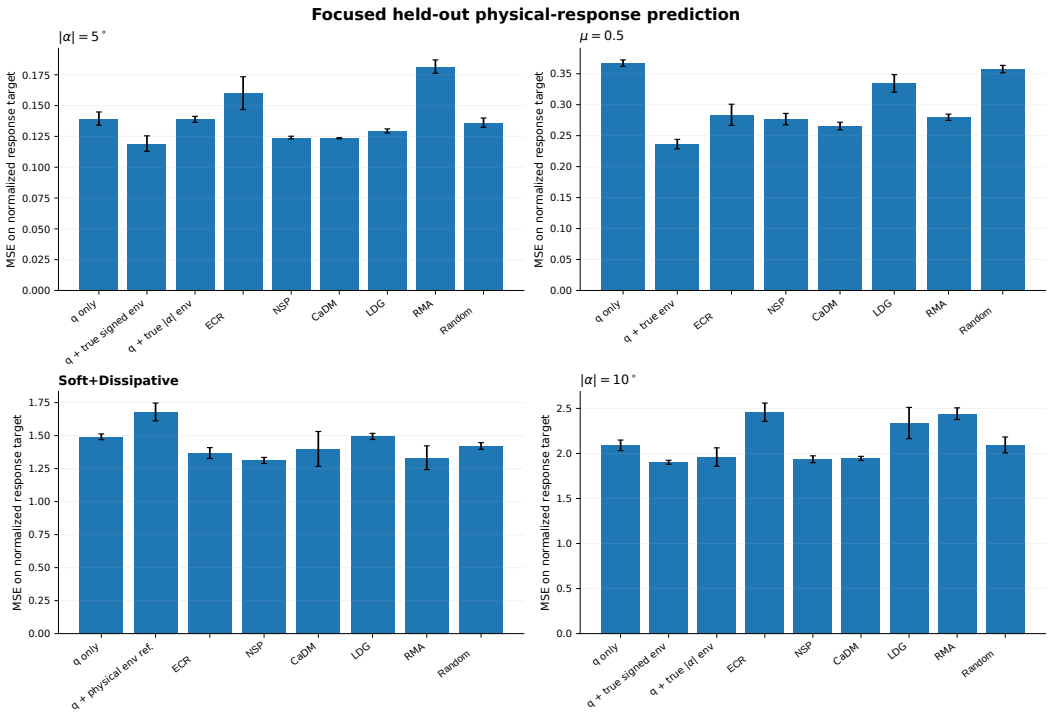


Figure 5: Focused unseen physical-response prediction. The four targets and normalizations differ; compare methods within a panel. Error bars show sample SD. The surface physical-parameter reference is a separate 477-anchor diagnostic; the learned-code comparison uses 476 eligible anchors. It is not a bound on other methods.

counts are 672, 611, 412, and 476, respectively. The representation sweep retains its full 612-anchor extreme-steepness and 477-anchor surface populations. MSE units depend on each target's training normalization.

Table 14: Held-out $|\alpha| = 5^\circ$ (geometry). Normalized response MSE, mean $\pm$ sample SD over three runs.

| Input / representation | MSE ↓ |
|---|---|
| $q$ only | $0.139436 \pm 0.005380$ |
| $q+$ signed parameters | $0.119189 \pm 0.006265$ |
| $q+$ magnitude-based parameters | $0.138837 \pm 0.002433$ |
| ECR | $0.160105 \pm 0.013288$ |
| NSP | $0.124054 \pm 0.001091$ |
| CaDM | $0.123505 \pm 0.000451$ |
| LDG | $0.129441 \pm 0.001672$ |
| RMA | $0.181633 \pm 0.005442$ |
| Random | $0.136154 \pm 0.003740$ |

Table 15: Held-out $\mu = 0.5$ (friction). Normalized response MSE, mean $\pm$ sample SD over three runs.

| Input / representation | MSE ↓ |
|---|---|
| $q$ only | $0.367182 \pm 0.005173$ |
| $q+$ parameters | $0.236348 \pm 0.007681$ |
| ECR | $0.283647 \pm 0.017139$ |
| NSP | $0.276613 \pm 0.009260$ |
| CaDM | $0.265400 \pm 0.006255$ |
| LDG | $0.334357 \pm 0.014173$ |
| RMA | $0.279730 \pm 0.005076$ |
| Random | $0.357450 \pm 0.006023$ |

Table 16: Held-out Soft+Dissipative. Normalized response MSE, mean $\pm$ sample SD over three runs. †The physical-parameter reference is a separate fit; its 477-anchor population differs from the 476-anchor learned-code comparison. See Table 18.

| Input / representation | MSE ↓ |
|---|---|
| $q$ only | $1.490992 \pm 0.021787$ |
| $q+$ physical-parameter reference† | $1.679088 \pm 0.067564$ |
| ECR | $1.367759 \pm 0.041412$ |
| NSP | $1.311938 \pm 0.022757$ |
| CaDM | $1.399196 \pm 0.132051$ |
| LDG | $1.493848 \pm 0.023255$ |
| RMA | $1.332678 \pm 0.089857$ |
| Random | $1.421438 \pm 0.025153$ |

Table 17: Held-out $|\alpha| = 10^\circ$ (geometry). Normalized response MSE, mean $\pm$ sample SD over three runs.

| Input / representation | MSE ↓ |
|---|---|
| $q$ only | $2.089727 \pm 0.058947$ |
| $q+$ signed parameters | $1.902110 \pm 0.021751$ |
| $q+$ magnitude-based parameters | $1.960235 \pm 0.102144$ |
| ECR | $2.457831 \pm 0.101832$ |
| NSP | $1.935598 \pm 0.038329$ |
| CaDM | $1.945042 \pm 0.022842$ |
| LDG | $2.337729 \pm 0.173521$ |
| RMA | $2.441353 \pm 0.065037$ |
| Random | $2.094124 \pm 0.088847$ |

Table 18: Separate surface-reference diagnostic at the held-out Soft+Dissipative profile. Each entry is a mean over three runs with sample SD on 477 test anchors. The one-hot input contains an unseen categorical coordinate and is retained only as a diagnostic. These fits are distinct from the primary 476-anchor comparison.

| Input | Normalized MSE |
|---|---|
| Query only | $1.557247 \pm 0.047848$ |
| Legacy one-hot reference | $4.068609 \pm 0.226269$ |
| Physical-parameter reference | $1.679088 \pm 0.067564$ |

The physical reference augments the canonical query with nine known-parameter features: signed slope divided by 10, friction divided by 1.2, and the two `solref` plus five `solimp` entries of the contact profile. Each contact-parameter coordinate is scaled by its maximum absolute value over the three specified profiles. This is a known-parameter diagnostic that uses the profile catalogue for descriptor scaling, not a label-free history representation.

Prediction heads are fitted on 5,174 training and 1,127 validation anchors from the retained surface profiles, then evaluated on 477 Soft+Dissipative anchors. Query and target statistics use permitted training data. The heads have three 256-wide SiLU hidden layers and use 5,000 AdamW updates, batch size 256, learning rate $10^{-3}$, weight decay $10^{-5}$, and cosine decay. The separate query-only and one-hot fits are shown with the physical descriptor in Table 18. None is a theoretical generalization ceiling.

## F TEMPORAL-INFORMATION CONTROL

The local-information head receives one completed transition, whereas the full-history head receives the fixed features of Eq. (12). Both use 32-dimensional output codes and 1,644,192 trainable parameters. Their architecture and input processing differ, so this is a capacity-matched information-access control rather than a temporal-order ablation. The comparison is repeated for ECR and NSP at $5^\circ$ and $10^\circ$ holdouts.

Table 19: Temporal-control information and retrieval. Mean $\pm$ sample SD over three runs.

| Holdout | Objective | Input | NMI | Factor retrieval | Cell retrieval |
|---|---|---|---|---|---|
| $5^\circ$ | ECR | 64 steps | $0.888 \pm 0.014$ | $0.861 \pm 0.015$ | $0.360 \pm 0.015$ |
| $5^\circ$ | ECR | 1 transition | $0.433 \pm 0.024$ | $0.550 \pm 0.022$ | $0.107 \pm 0.015$ |
| $5^\circ$ | NSP | 64 steps | $0.462 \pm 0.031$ | $0.589 \pm 0.002$ | $0.256 \pm 0.022$ |
| $5^\circ$ | NSP | 1 transition | $0.089 \pm 0.012$ | $0.450 \pm 0.012$ | $0.112 \pm 0.014$ |
| $10^\circ$ | ECR | 64 steps | $0.942 \pm 0.020$ | $0.942 \pm 0.017$ | $0.252 \pm 0.028$ |
| $10^\circ$ | ECR | 1 transition | $0.420 \pm 0.007$ | $0.706 \pm 0.025$ | $0.122 \pm 0.021$ |
| $10^\circ$ | NSP | 64 steps | $0.553 \pm 0.018$ | $0.799 \pm 0.004$ | $0.187 \pm 0.003$ |
| $10^\circ$ | NSP | 1 transition | $0.195 \pm 0.003$ | $0.517 \pm 0.012$ | $0.098 \pm 0.019$ |

Table 20: Temporal-control ordering and physical prediction. Mean $\pm$ sample SD over three runs.

| Holdout | Objective | Input | Group ordering | Geometry MSE |
|---|---|---|---|---|
| $5^\circ$ | ECR | 64 steps | $1.000 \pm 0.000$ | $0.1601 \pm 0.0133$ |
| $5^\circ$ | ECR | 1 transition | $0.970 \pm 0.052$ | $0.1877 \pm 0.0131$ |
| $5^\circ$ | NSP | 64 steps | $0.273 \pm 0.091$ | $0.1241 \pm 0.0011$ |
| $5^\circ$ | NSP | 1 transition | $0.303 \pm 0.052$ | $0.1320 \pm 0.0016$ |
| $10^\circ$ | ECR | 64 steps | $0.909 \pm 0.000$ | $2.4578 \pm 0.1018$ |
| $10^\circ$ | ECR | 1 transition | $0.970 \pm 0.052$ | $2.4890 \pm 0.0610$ |
| $10^\circ$ | NSP | 64 steps | $0.000 \pm 0.000$ | $1.9356 \pm 0.0383$ |
| $10^\circ$ | NSP | 1 transition | $0.030 \pm 0.052$ | $1.9256 \pm 0.0254$ |

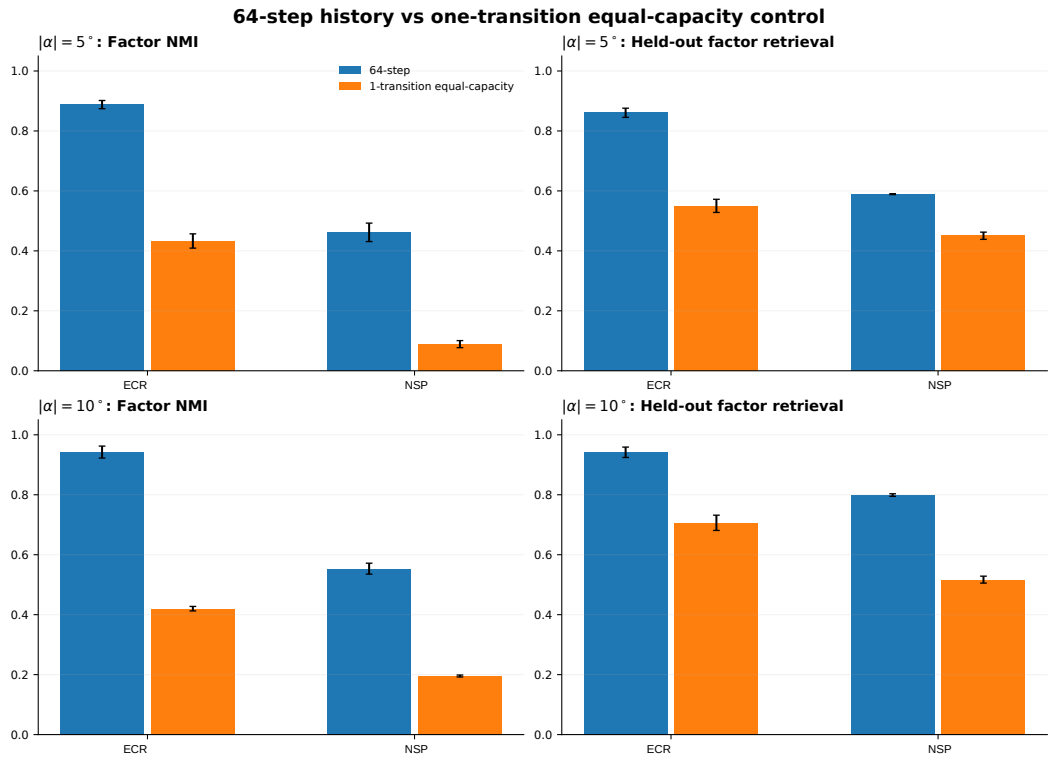


Figure 6: Parameter-count-matched history control. The information and retrieval tests improve substantially with 64-step history. Physical prediction follows a different, condition-dependent pattern. Both models use completed interaction only.

# G GEOMETRY INFORMATION AND PREDICTION TARGETS

## G.1 MAGNITUDE AND DIRECTIONAL RESPONSES

The paired-branch diagnostic compares $+5^\circ$ and $-5^\circ$ responses at equivalent terrain-relative configurations. Canonical-query RMS mismatch is $1.32103 \times 10^{-8}$ and its maximum absolute mismatch is $3.05474 \times 10^{-7}$. The normalized response RMS mismatch is 1.21663. Consequently, a decoder using the canonical query and magnitude alone receives essentially identical inputs for physically different directional targets.

For a balanced pair with identical available input, Eq. (7) separates the error of a common prediction from the squared difference between the targets. Averaging its second term per coordinate and pair gives 0.37005 in the balanced all-anchor diagnostic. This is not a lower bound for a different factual-test distribution, nor for an encoder that retains additional directional information.

A separately fitted factual diagnostic at held-out $5^\circ$ gives query-only MSE $0.140165 \pm 0.003835$, query plus magnitude $0.131551 \pm 0.000975$, and query plus signed slope $0.124971 \pm 0.001544$. These are separate probe fits from the focused physical comparison and should not be combined as though all entries shared an identical predictor initialization.

## G.2 EVEN-RESPONSE DIAGNOSTIC

For $r = |\alpha|$, decompose the paired response into

$$Y_{\text{even}}(r) = \tfrac{1}{2}[Y(+r) + Y(-r)], \qquad Y_{\text{odd}}(r) = \tfrac{1}{2}[Y(+r) - Y(-r)]. \tag{15}$$

The even target is centered using the permitted training magnitudes and normalized with its own training statistics. Existing encoders trained on the original objectives remain frozen; fresh heads

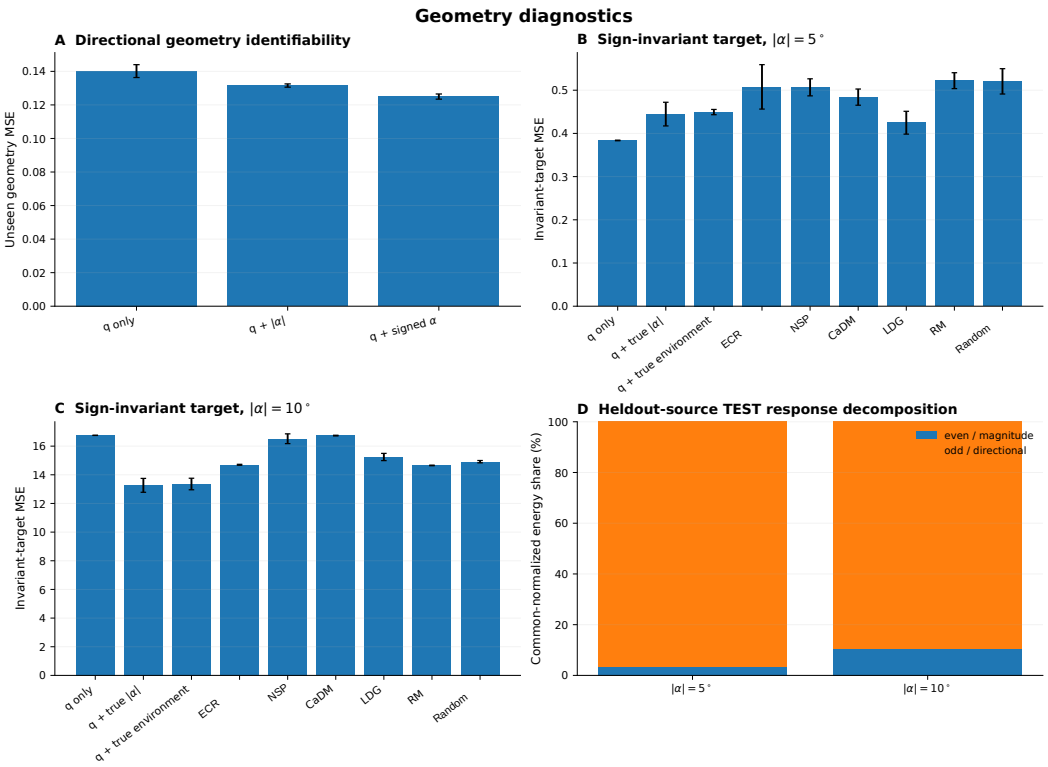


Figure 7: Geometry diagnostics: query/response compatibility, directional and magnitude response components, and prediction of an even target. The decomposition statistics and prediction errors refer to the populations and normalizations defined in this appendix; they are not interchangeable performance measures.

predict this alternative target. Receiving queries use the full test anchor population; magnitude-matched donors come from separate held-out test rollouts. The fresh even-response probes explicitly seed both PyTorch parameter initialization and NumPy sampling. They do not retrain the encoders or use the same target/population as the factual response test.

The energy diagnostic uses a common per-coordinate scale: the training standard deviation of signed responses centered by the retained-magnitude even-response mean. On heldout-source test anchors (672 at $5^\circ$, 612 at $10^\circ$), it gives even/odd shares of 3.28%/96.72% at the $5^\circ$ holdout and 10.38%/89.62% at $10^\circ$. These percentages characterize the tested response distribution, not a universal decomposition of contact mechanics. Tables 21 and 22 report both holdouts: ECR improves over NSP at $10^\circ$, whereas query-only is best at $5^\circ$.

Table 21: Sign-invariant geometry response at held-out $5^\circ$. Frozen encoders with new probes; normalized MSE, mean $\pm$ sample SD.

| Input / representation | MSE ↓ |
|---|---|
| q only | $0.383700 \pm 0.000705$ |
| q + true $\lvert\alpha\rvert$ | $0.444705 \pm 0.027400$ |
| q + true environment | $0.449385 \pm 0.006090$ |
| ECR | $0.507730 \pm 0.051419$ |
| NSP | $0.506658 \pm 0.019736$ |
| CaDM | $0.483988 \pm 0.018754$ |
| LDG | $0.424768 \pm 0.026422$ |
| RMA | $0.522103 \pm 0.018395$ |
| Random | $0.520503 \pm 0.029331$ |

Table 22: Sign-invariant geometry response at held-out $10^\circ$. Frozen encoders with new probes; normalized MSE, mean $\pm$ sample SD.

| Input / representation | MSE ↓ |
|---|---|
| q only | $16.752207 \pm 0.006097$ |
| q + true $\lvert\alpha\rvert$ | $13.263787 \pm 0.484427$ |
| q + true environment | $13.357233 \pm 0.402455$ |
| ECR | $14.698549 \pm 0.032791$ |
| NSP | $16.516724 \pm 0.340719$ |
| CaDM | $16.736066 \pm 0.023573$ |
| LDG | $15.241901 \pm 0.251868$ |
| RMA | $14.653011 \pm 0.020579$ |
| Random | $14.916250 \pm 0.081026$ |

# H PAIRED TARGET-CENTERING ABLATION

## H.1 TARGETS, MATCHING, AND REPRESENTATION RUNS

For the fixed-head, fixed-scale control, each of the three heads predicts the complete response using the corresponding ECR scale vector. Its loss is

$$\mathcal{L}_{\text{NSP-Matched}}(\theta,\phi)=\mathbb{E}_{(A,\xi,B)}\left[\sum_{f\in\{G,F,S\}}\frac{1}{36}\left\|D_{\phi_f}(\widetilde{q}_A,E_\theta(F(H_B)))-\frac{Y(A,\xi)}{\sigma_f}\right\|_2^2\right]. \quad (16)$$

The paired ECR loss is Eq. (6). The control does not subtract the sibling-factor mean or introduce global target centering, target clipping, or arm-specific loss rescaling. Both arms use the same sum over three heads and the same stored scale vectors. Thus their target difference in head $f$ is exactly $\bar{Y}_f(A,\xi)/\sigma_f$.

The pair uses the strict cross-rollout sampler, common features and 32-dimensional interface, and a maximum of 6,000 updates. Recorded initialization, sampled-batch, preprocessing, and scale-vector fingerprints match within each of the three runs (encoder seeds 67001, 67102, and 67203). Checkpoints minimize each arm's validation objective at the 500-update selection intervals. The original NSP/CaDM encoders are reused as contextual references, not as members of this new encoder-training pair. All runs share the same fixed reservoir realization.

Table 1 gives the main metrics; Table 23 gives the remaining factor and predictive diagnostics. Table 24 retains the individual run values. The mean paired differences, with sample SD, are $0.520559 \pm 0.030571$ for steepness $R^2$ and $0.208333 \pm 0.013961$ for exact-cell retrieval (ECR minus control). For next-step prediction MSE, control minus ECR is $0.00215201 \pm 0.00018277$. Each difference favors ECR in all three pairs. These are descriptive comparisons across three runs, not uncertainty over independently sampled reservoirs.

Table 23: Additional paired-study diagnostics. Mean ± sample SD over three runs; NSP/CaDM are contextual encoder references. The reference-gap values use the fresh fits specified in Section H.2.

| Representation | Steepness NMI | Friction $R^2$ | Surface accuracy | Gap (%) |
|---|---|---|---|---|
| ECR-Paired | $0.9588 \pm 0.0025$ | $0.8767 \pm 0.0060$ | $0.9951 \pm 0.0018$ | $159.25 \pm 8.84$ |
| NSP-Matched | $0.4589 \pm 0.0262$ | $0.8356 \pm 0.0019$ | $0.9958 \pm 0.0011$ | $105.55 \pm 4.44$ |
| NSP$^\dagger$ | $0.4246 \pm 0.0165$ | $0.8595 \pm 0.0029$ | $0.9958 \pm 0.0012$ | $146.96 \pm 2.22$ |
| CaDM$^\dagger$ | $0.5777 \pm 0.0462$ | $0.8577 \pm 0.0017$ | $0.9964 \pm 0.0011$ | $142.24 \pm 3.57$ |

## H.2 FRESH PREDICTION PROTOCOL AND EARLIER AGGREGATE DIAGNOSTIC

In the paired study, each encoder is frozen before fitting a fresh next-step prediction head for 5,000 updates with the common prediction task, normalization, and fitting budget. The head seed is the encoder seed plus seven. Factual cross-rollout prediction uses 7,660 training, 1,639 validation, and 1,643 test anchors. Representation diagnostics use the full 1,648 test histories instead. The receiving query and donor history satisfy the strict rollout constraint. This is an in-distribution evaluation on unseen rollouts, not a repetition of the leave-one-factor-level-out experiments.

The query-only and known-parameter reference MSEs are 0.032991689 and 0.028984126, respectively. Each reference averages two fresh fitted heads; separate reference-fit errors are not available, so no reference-fit uncertainty is inferred. Using these references in Eq. (8) gives mean gap closure 159.25%, 105.55%, 146.96%, and 142.24% for ECR-Paired, NSP-Matched, NSP, and CaDM. The corresponding raw MSEs are the primary predictive results.

The six-objective diagnostic in Table 9 uses different encoders, fits, and references, with the provenance limits stated in Appendix C. It remains in Figure 3B; the present four-representation comparison does not replace its original ECR encoder or supply uncertainty for its aggregates.

Table 24: Individual encoder/probe runs. Seed identifies the encoder; prediction-head seed is encoder seed plus seven. The NSP/CaDM encoder training is not part of the new ECR/NSP-Matched pairing.

| Representation | Seed | Steepness $R^2$ | NMI | Cell retrieval | NSP MSE |
|---|---|---|---|---|---|
| ECR-Paired | 67001 | 0.8852 | 0.9587 | 0.4830 | 0.02632378 |
| ECR-Paired | 67102 | 0.9129 | 0.9614 | 0.5121 | 0.02700596 |
| ECR-Paired | 67203 | 0.9024 | 0.9564 | 0.5000 | 0.02649960 |
| NSP-Matched | 67001 | 0.3941 | 0.4533 | 0.2888 | 0.02864870 |
| NSP-Matched | 67102 | 0.3607 | 0.4874 | 0.2900 | 0.02896674 |
| NSP-Matched | 67203 | 0.3840 | 0.4360 | 0.2913 | 0.02866993 |
| NSP$^\dagger$ | 67001 | 0.6349 | 0.4210 | 0.2967 | 0.02719982 |
| NSP$^\dagger$ | 67102 | 0.4732 | 0.4426 | 0.2749 | 0.02702542 |
| NSP$^\dagger$ | 67203 | 0.5582 | 0.4102 | 0.2949 | 0.02708187 |
| CaDM$^\dagger$ | 67001 | 0.7028 | 0.5256 | 0.3064 | 0.02720075 |
| CaDM$^\dagger$ | 67102 | 0.7316 | 0.6136 | 0.3374 | 0.02745589 |
| CaDM$^\dagger$ | 67203 | 0.7287 | 0.5940 | 0.3477 | 0.02721676 |

### H.3 Target scales and optimization

Dividing the uncentered response by ECR scale vectors produces a largest coordinate-wise NSP/ECR target RMS ratio of 339.48. The geometry-head scaled full-target absolute 99th percentile reaches 605.09. These measure target-scale imbalance, not a Hessian condition number. Table 25 gives gradient and selection diagnostics.

Table 25: Optimization diagnostics over all 6,000 representation updates. The clipping threshold is 1.0. Selected checkpoints minimize each arm's own validation loss; clipping statistics cover the full training run, not only the selected prefix.

| Arm | Seed | Clipped (%) | Mean pre-clip norm | Selected step |
|---|---|---|---|---|
| NSP-Matched | 67001 | 100.000 | 3207.610 | 6000 |
| NSP-Matched | 67102 | 100.000 | 3131.807 | 6000 |
| NSP-Matched | 67203 | 100.000 | 3154.722 | 6000 |
| ECR-Paired | 67001 | 0.583 | 0.512 | 6000 |
| ECR-Paired | 67102 | 11.883 | 0.679 | 4000 |
| ECR-Paired | 67203 | 7.650 | 0.644 | 3000 |

The uncentered control activates norm clipping on every update, with mean pre-clipping norms above 3,100; ECR clips on 0.583%, 11.883%, and 7.650% of updates. These summaries span all 6,000 updates, even where an earlier checkpoint was selected. The control's validation objective decreases from approximately 24,248 to 1,124–1,219, and its friction and surface diagnostics remain informative. Finite losses do not establish equally effective optimization, but the observed run is not numerical divergence. All control runs select the final checkpoint; two ECR runs select earlier checkpoints because later validation loss worsens.

These observations bound the mechanism interpretation. At matched architecture, scales, data, initialization, and maximum budget, removing the branch-shared response improves accessibility and fresh prediction relative to the specified NSP-Matched control. The intervention also changes target magnitudes, gradient clipping, and optimization. The experiment does not separate a representation-allocation effect from those numerical consequences or establish superiority to an independently tuned uncentered learner. Clipping frequency alone is not a measure of effective AdamW parameter-step magnitude. The smaller mean gain over standard NSP (about 1.8%) is a contextual predictive comparison, not the fully paired encoder-training intervention; one corresponding run differs by only 0.00001946 MSE.

# I TEMPORAL RESERVOIR FEATURE CONSTRUCTION

These architecture-development diagnostics examine how a fixed reservoir trajectory is summarized before learning the final shared code. They use a preceding dataset and target definitions, and are not additional replications of the supervision comparison or its factor-level holdouts. In particular, local terrain-gradient regression and material classification below differ from the main steepness probes and physical-response targets. We report the feature comparisons as exploratory development evidence; their test scores were inspected when choosing the feature views.

## I.1 INFORMATION IN TEMPORAL FEATURES

A single 512-unit reservoir (seed 15101, leak 0.05, spectral-radius setting 0.9) processes each 64-step history from zero state. Six feature views compare the final state, temporal taps, temporal statistics, and a coordinate-wise squared expansion. Here $r_k$ denotes the state after processing $k$ observations, not a readout of dimension $k$. Let $T = [r_{16}, r_{32}, r_{48}, r_{64}]$ and let $m$ and $s$ be the temporal mean and standard deviation. This diagnostic computes $s$ from the population second moment with a $10^{-8}$ offset inside the square root; the main representation uses the sample convention in Appendix B.

Independent linear heads predict the two-dimensional local terrain gradient, three friction classes, three contact-profile classes, and their nine joint material classes. These labels supervise the diagnostic heads only and are not reservoir inputs. Features and geometry targets are standardized with training statistics. The fitting code minimizes geometry MSE plus classification cross-entropies and selects a checkpoint using validation geometry error plus one minus joint-material accuracy. There is no learned shared bottleneck in this diagnostic.

Table 26: Reservoir-feature development diagnostic. One reservoir realization; geometry is normalized MSE and classification values are percentages. These are separate tasks from the main supervision study.

| Features | Dimension | Geometry ↓ | Material ↑ | Friction ↑ | Profile ↑ |
|---|---|---|---|---|---|
| $r_{64}$ | 512 | 0.02749 | 82.14 | 86.73 | 82.65 |
| $[r_{64}, m]$ | 1,024 | 0.02422 | 83.42 | 88.52 | 83.16 |
| $[r_{64}, m, s]$ | 1,536 | 0.02468 | 86.10 | 91.71 | 85.59 |
| $T$ | 2,048 | 0.02594 | 82.14 | 87.88 | 82.02 |
| $[T, m, s]$ | 3,072 | 0.02509 | 86.61 | 92.73 | 86.99 |
| $[T, m, s, (T, m, s)^{\odot 2}]$ | 6,144 | 0.02605 | 85.84 | 91.58 | 86.73 |

Adding temporal statistics improves friction and contact-profile classification relative to the final state alone. Final state plus mean gives the lowest geometry error, whereas taps plus moments give the highest material, friction, and profile accuracies. The squared expansion does not improve those results. The ordering supports complementary feature summaries, not a claim that each physical parameter has a uniquely identified timescale. Increasing feature dimension also increases linear-probe capacity, so this is not a parameter-count-matched ablation. The contact-profile labels do not constitute calibrated stiffness measurements.

### I.2 USE BY PHYSICAL-RESPONSE PREDICTORS

A separate development comparison fits independent geometry and material-response predictors using three reservoir realizations (seeds 15101, 15102, and 15103). The material target combines friction and contact-profile effects in this earlier formulation. Final-state predictors use $r_{64}$ for both tasks; rich views use $[r_{64}, m]$ for geometry and $[T, m, s]$ for material. All views are standardized on training data. Unlike the main method, these predictors have no shared 32-dimensional representation bottleneck.

The additive decoder separately maps the query and reservoir features to 256-dimensional vectors, sums them, and applies a nonlinear output network. The concatenated decoder instead applies a three-hidden-layer, width-256 SiLU network to their concatenation. A validation screen compares additive, concatenated, FiLM, and bilinear decoders. The final comparison retains the two additive baselines and the selected concatenated decoder. Recorded training logs span 4000 updates, and each task selects its own checkpoint by validation MSE. Training uses a smooth-$L_1$ loss; the reported evaluation metric is normalized MSE.

Table 27: Development prediction tasks with independent geometry and material heads. Normalized MSE, mean $\pm$ sample SD over three reservoir realizations. Rich views and decoder definitions are given in the text.

| Features and decoder | Geometry ↓ | Material ↓ |
|---|---|---|
| Final state, additive | $0.246970 \pm 0.002956$ | $0.576442 \pm 0.002810$ |
| Rich views, additive | $0.246924 \pm 0.001273$ | $0.529660 \pm 0.004345$ |
| Rich views, concatenated MLP | $0.240152 \pm 0.002649$ | $0.522350 \pm 0.001044$ |

Within the additive decoder family, richer features reduce mean material error by 8.12%, with improvement in all three corresponding runs; mean geometry error is nearly unchanged. Changing to the concatenated decoder improves both aggregate means. This supports using richer reservoir views for these development tasks, but includes their increased input dimension and decoder capacity. It does not isolate the effect of temporal aggregation in the main learned-code architecture.

The main study therefore uses the full temporal feature bank as common input to a nonlinear representation head, then compares supervision objectives under that fixed feature construction. The development diagnostics motivate this design choice without establishing a new reservoir architecture, a factor-specific timescale law, or superiority over all trainable temporal encoders. Likewise, the embedding results of Hart et al. (2020) concern suitable dynamical systems and reservoirs; they do not establish an embedding guarantee for this finite-window, action-driven experiment or its learned 32-dimensional code.