# KNOWING WHAT TO ADAPT: CONTINUAL TEST-TIME ADAPTATION OF WORLD MODELS

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

## ABSTRACT

A world model that adapts during deployment can repair today's forecast while learning the wrong dynamics. A pushed object that moves less than expected may reflect a shifted camera or increased friction: both corrections fix the prediction, but only the friction should be remembered. We cast continual test-time adaptation as a question of *correction lifetime*: what should the model use now, and what should it keep? *Conditionally Resolved Dynamics* (CoRD) splits every correction into a temporary cache and a candidate dynamics memory. The candidate holds only the part of the correction that no observation change can produce, and an exact least-squares identity shows that the split itself costs no local fit. Because equally good fits can call for different memories, unresolved directions never persist, and a candidate persists only if the model left after cache expiry beats the incumbent on fresh outcomes under anytime-valid comparisons. Across pushing, navigation, and three control tasks with appearance shifts, overlapping dynamics shifts, and returns to nominal, CoRD lowers 10-step forecast error by 12–27% over the strongest baseline, cuts first-return error by 34–47% versus carried adapters while recovering faster than episode-reset ones, and retains nothing before the dynamics change, whereas validating only the full fit commits spurious corrections in 42% of streams.

## 1 INTRODUCTION

A better prediction today can produce a worse world model tomorrow. Suppose a pushed object moves less than the model expects (Figure 1). A shifted camera and increased friction both explain the error, and correcting either one repairs the forecast. Only the friction belongs in memory: once the camera returns, a remembered "slower object" becomes a new error. A model that adapts continually must therefore decide, for every correction, what to use now and what to keep.

We study action-conditioned world models (Hafner et al., 2019; Hansen et al., 2024; Zhou et al., 2025) that adapt during deployment from their own observations and actions, without source data or rewards. Continual test-time adaptation stabilizes such updates with teachers, restoration, reliability filtering, and memory (Wang et al., 2022; Döbler et al., 2023; Niu et al., 2022; 2023; Yuan et al., 2023), and world-model methods adapt representations or lightweight residuals around a frozen predictor (Wang et al., 2026; Soni et al., 2026). These methods decide how much to update, but treat persistence uniformly: a correction is carried forward, reset, or restored toward the source, regardless of what it explains. Carrying writes observation effects into dynamics memory; resetting discards genuine dynamics changes and relearns them. Restricting updates to identified directions (Simchowitz et al., 2018; Jhawar & Wang, 2026) avoids stale memory but withholds corrections the forecast needs when appearance and dynamics change together. We call this the *fit–memory dilemma*.

We resolve it by making **correction lifetime a design variable**. **Conditionally Resolved Dynamics (CoRD)** splits each correction into a temporary cache that serves the forecast now and a candidate memory that persists only if it earns its place. The split answers two questions in turn.

*What can be remembered?* Only the part of a correction that no observation change could have produced. CoRD removes from the dynamics response everything the observation adapter can explain; what remains is candidate memory, and the rest stays in the cache. A least-squares identity shows that

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

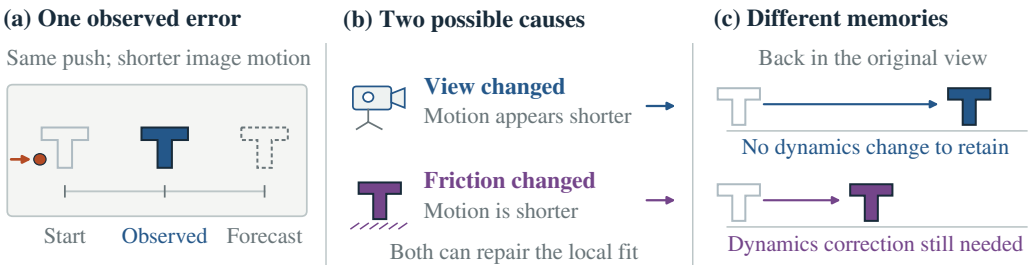


Figure 1: **The same local fit can require different memories.** (a) A push produces less image motion than forecast. (b) A changed view or increased friction can explain this local error. (c) Back in the original view, only the physical change still requires a dynamics correction. Fitting the present cannot, by itself, decide what to remember; that takes new evidence.

the split itself costs nothing: cache and candidate together can reproduce the full joint fit, so memory can be selective while the predictor keeps its full capacity.

*What should be remembered?* Fitting alone cannot say: two worlds can produce identical data yet require different memories, so any rule that only fits the past errs in one of them. CoRD therefore keeps what the data cannot resolve in the cache and commits the rest only after a test: it freezes the exact model left after cache expiry and requires it to beat the incumbent on fresh outcomes, using anytime-valid forecast comparisons (Howard et al., 2021; Choe & Ramdas, 2024; Arnold et al., 2026) that hold under adaptive actions. Where the data cannot separate explanations, an active variant spends a few actions on probes that can. Memory is earned by evidence, not granted by fit.

Across pushing, navigation, and three continuous-control tasks with appearance shifts, dynamics shifts, and returns to nominal conditions, CoRD gives the most accurate long-horizon forecasts among methods sharing its adapter interface, lowers first-return error by 34–47% relative to carried adapters while recovering faster than episode-reset ones, and retains nothing before the dynamics change. Without probing, CoRD raises Walker return by 79 points over the best comparator; spending up to 10% of its actions on informative probes raises PushT success to 80.3%, against 69.5% for the strongest baseline. Our contributions are:

- **Correction lifetime.** We identify correction lifetime as a missing design axis: splitting a correction into cache and memory costs no local fit, while fit alone cannot choose the memory.
- **CoRD.** A cache–memory architecture derived from conditional geometry, with prospective certification of the exact post-cache model, evidence refresh, and optional active probing.
- **Evidence.** Gains in forecasting, control, and return-to-nominal error on five visual-control tasks, and matched lifetime ablations that show how each check shapes what the model keeps.

## 2 Related Work

**World models and robust representations.** World models support planning in latent spaces (Ha & Schmidhuber, 2018; Hafner et al., 2019; 2025; Hansen et al., 2022; 2024; Zhou et al., 2025). Unified features (Li et al., 2025), nuisance-aware abstraction (Higgins et al., 2017; Zhang et al., 2021; Fu et al., 2021), augmentation (Laskin et al., 2020; Yarats et al., 2021; Hansen et al., 2021b), and replay (Chen et al., 2022) make them robust before deployment; CoRD adapts them online.

**Test-time and continual adaptation.** Test-time learning uses self-supervision, entropy, reconstruction, and source hypotheses (Sun et al., 2020; Wang et al., 2021; Zhang et al., 2022; Gandelsman et al., 2022; Liang et al., 2020). Continual methods stabilize updates with teachers, restoration, reliability filtering, and memory (Wang et al., 2022; Döbler et al., 2023; Niu et al., 2022; 2023; Yuan et al., 2023; Niloy et al., 2024), and identifiability analysis asks when unlabeled evidence determines the right update (Jhawar & Wang, 2026). Periodic resets (Press et al., 2023) and persistence under recurring shifts (Hoang et al., 2024) show that what an adapter keeps matters as much as how it updates; CoRD decides how long each part of an update should last.

**Fast and slow memory.** Fast weights (Hinton & Plaut, 1987; Ba et al., 2016), complementary learning systems (McClelland et al., 1995; Kumaran et al., 2016; Arani et al., 2022), and two-timescale adapters (Pham et al., 2023; Liu et al., 2024) separate quickly adapted knowledge from slowly consolidated knowledge. CoRD gives this separation an explicit promotion rule for world models: a correction moves from the fast cache to slow memory only when conditional geometry allows it and prospective prediction confirms it.

**Adaptive dynamics and inference-time control.** Adaptive dynamics and physical-context inference enable rapid adjustment (Nagabandi et al., 2018; 2019; Kumar et al., 2021; Ball et al., 2021; Hansen et al., 2021a), and recent methods adapt latent predictors, transferred world models, or residuals around frozen predictors (Wang et al., 2026; Wan et al., 2026; Lanier et al., 2025; Levy et al., 2026; Orsula et al., 2026; Soni et al., 2026); CoRD gives them distinct local and persistent roles.

**Identification, testing, and sequential evidence.** Excitation determines finite-trajectory and controlled-representation recovery (Simchowitz et al., 2018; Zhang et al., 2026), and active identification selects informative inputs (Wagenmaker & Jamieson, 2020). Variable projection (Golub & Pereyra, 2003), spectral regularization (Hansen, 1987), and multiple and structural tests (Benjamini & Yekutieli, 2001; Chow, 1960) support estimation and decisions; excitation-set RLS combines informative memory, forgetting, and change detection (Leung et al., 2024). Prequential evaluation (Dawid, 1984) and sequential forecast comparison (Choe & Ramdas, 2024; Arnold et al., 2026) judge models by forecasts issued before outcomes arrive, and high-confidence improvement and risk monitoring certify updates or flag harmful shifts (Thomas et al., 2015; Podkopaev & Ramdas, 2022); CoRD uses these principles to decide what a world model remembers.

## 3 CoRD: Adaptation with Selective Memory

CoRD answers two questions for every correction: what *can* be remembered (Sections 3.2–3.3) and what *should* be remembered (Sections 3.4–3.5). Figure 2 traces a correction from the local fit to the model that remains; the backbone stays frozen throughout.

### 3.1 Source-Free Continual Adaptation

Dynamics $s_{t+1} \sim \mathcal{T}_{\eta_t}(\cdot \mid s_t, u_t)$ and observations $x_t \sim \mathcal{O}_{\xi_t}(\cdot \mid s_t)$ can change independently. Deployment reveals observations and executed actions, but neither hidden states nor change boundaries. A frozen source encoder and transition support adapted predictions in a $d_z$-dimensional latent space

$$z_t = E_{\alpha_t}(x_t), \qquad \widehat{z}_{t+h} = F^{(h)}_{\beta_t}(z_t, u_{t:t+h-1}), \tag{1}$$

and a frozen task readout $R_0$ or goal metric scores plans. The observation adapter is $E_\alpha(x) = D_\alpha E_0(x) + \alpha_b$ with $D_\alpha = \mathrm{diag}(\exp \alpha_s)$. The transition adapter is $F_\beta(z,u) = F_0(z,u) + W_\beta \phi_0(z,u)$: a frozen, nonlinear state–action feature map feeds trainable output weights. With $n_\phi$ features, the correction has rank at most $\min(d_z, n_\phi)$; an affine alternative $F_\beta = \mathrm{diag}(\exp \beta_s)F_0 + \beta_b$ cannot represent contact-dependent changes in general. Checkpoint weights, coordinate bounds, and source-calibrated statistics remain available; source trajectories and deployment rewards do not.

The deployed transition can also include a context-local increment $b_U$; only the separately validated increment $b_R$ (Section 3.4) enters persistent $\beta$. Episode resets end incomplete trajectories without revealing regimes, and forecasts precede outcomes. Native proprioception is allowed, whereas physical states, canonical images, and shift labels are evaluation-only (Appendix B).

### 3.2 Conditional Evidence for Dynamics

An observation change and a transition change can produce the same residual. For a completed endpoint $k$ and horizon $h$, the pulled-back residual is

$$r_{k,h} = D_\alpha^{-1}\big[E_\alpha(x_k) - F^{(h)}_\beta\big(E_\alpha(x_{k-h}), u_{k-h:k-1}\big)\big]. \tag{2}$$

The pullback expresses error in frozen feature units, while the adapted encoding of the input state still enters through the transition. Let $J_o$ and $J_d$ denote the Jacobians of the stacked residuals with respect to the observation and dynamics adapter parameters; Appendix A.1 differentiates both encoder paths.

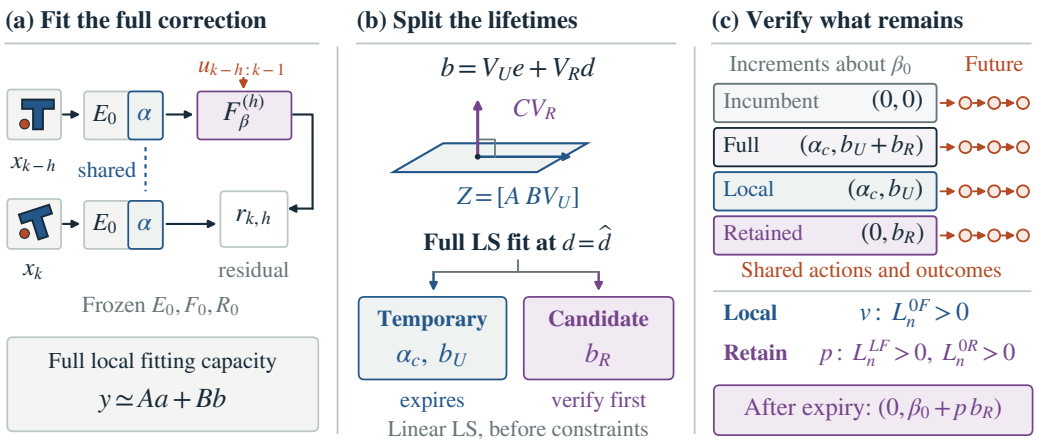


Figure 2: **Fit the present; verify what remains.** (a) Both encoder paths define the local residual. (b) Conditional geometry separates temporary compensation from candidate memory without losing the linear least-squares fit. (c) Incumbent, full, local, and retained snapshots share future actions and predict before outcomes. Local permission $v$ and retention permission $p$ determine deployment; cache expiry leaves exactly $(0, \beta_0 + pb_R)$, the model carried forward.

Source-fixed coordinate scales $D_o, D_d$ normalize increments, and the source covariance $\Omega_0 \succ 0$ weights the fit:

$$y = -\Omega_0^{-1/2} r, \quad A = \Omega_0^{-1/2} J_o D_o, \quad B = \Omega_0^{-1/2} J_d D_d, \qquad y \simeq Aa + Bb + \varepsilon. \tag{3}$$

A *context* is one completed fitting block $j$ with endpoint–horizon pairs $\mathcal{R}_j$, so $y$ stacks $m = d_z \sum_j |\mathcal{R}_j|$ coordinates; the weighting serves fitting only, and validation scores fresh errors.

Compatible contexts have separate observation coefficients $a_j$ and shared dynamics $b$. With block-diagonal $A$ and stacked $B, y$, profiling gives

$$\min_a \|y - Aa - Bb\|^2 = \|P(y - Bb)\|^2, \qquad P = I - AA^\dagger, \quad C = PB. \tag{4}$$

Columns of $C$ are dynamics responses that observation corrections cannot cancel. The Gram matrix $G = C^\top C$ defines the conditional geometry; under correctly whitened Gaussian noise it is also the profiled information matrix (Golub & Pereyra, 2003). Weakly resolved directions can still help prediction, so CoRD gives them a temporary lifetime instead of removing them.

### 3.3 A Fit–Memory Factorization

Let $C = U\Sigma V^\top$ with singular triplets $(\sigma_i, \mathrm{u}_i, v_i)$. Choose nonzero singular directions $V_R$ as candidate persistent coordinates, with matching $U_R, \Sigma_R$, and complete them with an orthonormal complement $V_U$, which contains the unresolved directions. Write

$$b = V_R d + V_U e, \qquad Z = [A\ BV_U], \qquad q = (a, e). \tag{5}$$

The increment $V_R d$ is candidate memory; $q$ is the temporary prediction cache. Like $b$, the dynamics coefficient $e$ is shared across fitted contexts, preserving the parameterization of the original correction.

**Proposition 1** (Local fit, selective lifetime)**.** *For any realized weighted design and fixed split $(V_R, V_U)$, let $\widehat{d} = \Sigma_R^{-1} U_R^\top Py$ and $\mathrm{RSS} = \min_{a,b} \|y - Aa - Bb\|^2$. Then, for every $d$,*

$$\boxed{\min_q \|y - BV_R d - Zq\|^2 = \mathrm{RSS} + \|\Sigma_R(d - \widehat{d})\|^2.} \tag{6}$$

*Moreover $P_Z BV_R = PBV_R$, where $P_Z = I - ZZ^\dagger$.*

*Proof sketch.* The SVD makes $CV_R$ orthogonal to both $A$ and $CV_U$, so $\mathrm{col}([A, B]) = \mathrm{col}(Z) \oplus \mathrm{col}(CV_R)$. Profiling $q$ and applying Pythagoras gives this variable-projection (Frisch–Waugh–Lovell) identity (Golub & Pereyra, 2003; Frisch & Waugh, 1933; Lovell, 1963); see Appendix A.

**Algorithm 1** CoRD: local prediction and verified persistence

**Require:** Frozen $E_0, F_0, R_0$; source scales; budgets $\delta_j$; decision times $\mathcal{N}$, $N = \max \mathcal{N}$

$\beta \leftarrow 0$; $\mathcal{M} \leftarrow \varnothing$
**for** each completed fitting block $\mathcal{D}_j$, ending at step $t_j$ **do**
  $\mathcal{M} \leftarrow \mathcal{H}(\mathcal{M} \cup \{\mathcal{D}_j\})$; $\beta_0 \leftarrow \beta$
  $(y, A, B) \leftarrow$ Eq. (3); $PB = U\Sigma V^\top$
  $(\mathcal{S}, d, q) \leftarrow$ Eqs. (7)–(8); form admissible $\chi_0, \chi_F, \chi_L, \chi_R$
  **for** $n = 1, \ldots, N$ **do**
    $t \leftarrow t_j + n$; $u_t \leftarrow \mathrm{Plan}(x_t, \chi_0, V_U)$; store all four forecasts
    Observe $x_{t+1}$; score Eq. (10); update $L_n^{0F}, L_n^{LF}, L_n^{0R}$
    **if** $n \in \mathcal{N}$ **then**
      $v \leftarrow \mathbf{1}[L_n^{0F} > 0]$; $p \leftarrow \mathbf{1}[\min(L_n^{LF}, L_n^{0R}) > 0]$
      **if** $v \vee p$ **then break**
  $\beta \leftarrow \beta_0 + p b_R$; deploy $\chi_F$ if $v = 1$, else $\chi_R$ if $p = 1$, else $\chi_0$
  At episode end (cache expiry): deploy $(0, \beta)$; add scored data to $\mathcal{M}$

**A smaller memory can support a full local fit.** At $d = \widehat{d}$ the cache recovers the full joint least-squares fit and leaves the conditional information about the selected coordinates unchanged, while shrinking $d$ costs exactly the second term. The identity therefore dictates the architecture: keep unresolved effects available for prediction, but give them a temporary lifetime.

**Identical histories can require different memories.** Suppose $PB\bar{b} = 0$ for a dynamics direction $\bar{b} \neq 0$. Then some $\bar{a}$ satisfies $A\bar{a} = B\bar{b}$, and the corrections $(0, 0)$ and $(-\bar{a}, \bar{b})$ produce identical local data. Once observation compensation expires, a future readout $\Gamma$ separates their dynamics by $\Gamma\bar{b}$, and every history-based retained estimate has worst-case squared error at least $\|\Gamma\bar{b}\|^2/4$ across the two cases (Appendix A). Local agreement cannot authorize persistence; only new evidence can.

### 3.4 Learning a Correction and Its Lifetime

**Construct the local and persistent components.** Stable modes satisfy $\sigma_i \geq \tau$. Among them, Benjamini–Yekutieli (BY) selection on the scores $T_i$ nominates an index set $\mathcal{S}$ whose right singular vectors form $V_R$, and positive-part shrinkage sets their coefficients:

$$\widehat{\nu}^2 = \frac{\|(I - XX^\dagger)y\|^2}{m - \operatorname{rank} X}, \quad c_i = \frac{\mathrm{u}_i^\top P y}{\sigma_i}, \quad T_i = \frac{\sigma_i c_i}{\widehat{\nu}}, \qquad X = [A\ B], \tag{7}$$

$$d_i = \widehat{w}_i c_i, \quad \widehat{w}_i = \left(1 - \frac{\widehat{\nu}^2}{\sigma_i^2 c_i^2}\right)_+, \qquad q = \arg\min_q \|y - BV_R d - Zq\|^2 + \lambda_q \|q\|^2. \tag{8}$$

Zero coefficients receive zero weight, and $\lambda_q$ is a fixed ridge weight for the cache. Spectral regularization and multiple testing (Hansen, 1987; Benjamini & Yekutieli, 2001) construct the candidate; the predictive comparison below decides its lifetime. With insufficient residual degrees of freedom, $\mathcal{S}$ is empty and the candidate is a regularized local fit. A common step size $\gamma$ converts the coefficients, in the source scales $D_o, D_d$, into the current observation correction $\alpha_c$ and dynamics increments $b_R$ (from $V_R d$) and $b_U$ (from $V_U e$), keeping the four snapshots below within source bounds.

**Verify the exact model that will remain.** For incumbent dynamics $\beta_0$, freeze four snapshots $\chi = (\alpha, \beta)$ of observation and dynamics parameters:

$$\begin{aligned} \chi_0 &= (0, \beta_0), & \chi_F &= (\alpha_c, \beta_0 + b_R + b_U), \\ \chi_L &= (\alpha_c, \beta_0 + b_U), & \chi_R &= (0, \beta_0 + b_R). \end{aligned} \tag{9}$$

$F$ is the full local predictor, $L$ removes only the proposed persistent increment, and $R$ is the retained predictor without the cache. The cache expires at the next episode boundary, which turns $F$ into $R$; the two are therefore distinct objects of evaluation. Accepted snapshots are deployed unchanged.

For scalar target 1, let the incumbent predict 0, local compensation add 2, and the proposed persistent increment be $-1$. The full prediction 1 beats both incumbent 0 and local control 2; after expiry, prediction $-1$ is worse than the incumbent. CoRD therefore evaluates $R$ explicitly.

Each snapshot predicts the outcome of the same executed action before that outcome arrives. For fixed source scale $M > 0$, use the observable common-target loss

$$\ell_t(\chi) = \min\left\{ \frac{1}{M} \left\| E_0(x_{t+1}) - D_\alpha^{-1}[F_\beta(E_\alpha(x_t), u_t) - \alpha_b] \right\|^2, 1 \right\}. \tag{10}$$

For snapshots $\chi, \chi'$, the paired gain $g_t^{\chi\chi'} = \ell_t(\chi) - \ell_t(\chi')$ lies in $[-1, 1]$; superscripts such as $0F$ abbreviate $(\chi_0, \chi_F)$. Let $\mathcal{F}_t$ contain the observations, forecasts, and action before $x_{t+1}$ arrives. With an error budget $\delta_j$ for the candidate from block $j$, split over three comparisons, a fixed positive grid $\Lambda$ gives the standard time-uniform lower bound (Howard et al., 2021; Choe & Ramdas, 2024)

$$L_n^{\chi\chi'} = \frac{\sum_{t=1}^n g_t^{\chi\chi'} - \mathcal{B}(n, \delta_j/3)}{n}, \qquad \mathcal{B}(n, \delta) = \min_{\lambda \in \Lambda} \left\{ \frac{\log(|\Lambda|/\delta)}{\lambda} + \frac{\lambda n}{2} \right\}. \tag{11}$$

Simultaneously over $n$, $L_n^{\chi\chi'}$ bounds $n^{-1} \sum_t \mathbb{E}[g_t^{\chi\chi'} \mid \mathcal{F}_t]$ with the stated confidence. Predictable forecasts and bounded scores suffice, including under adaptive actions and nonstationary noise.

**Allocate lifetime from predictive evidence.** At a decision time, $L_n^{0F} > 0$ permits local use of $F$. Persistence requires both $L_n^{LF} > 0$ and $L_n^{0R} > 0$: the increment helps with the cache present, and the retained model beats the incumbent without it. If only local use passes, CoRD deploys $F$ until the cache expires and then returns to the incumbent. If only persistence passes, it deploys $R$; if both pass, it deploys $F$, whose cache later expires to the already scored $R$. Otherwise it keeps $\chi_0$. For $\sum_j \delta_j \le \delta_{\text{total}}$, a union bound controls all erroneous gain certificates across attempts (Arnold et al., 2026). The certified quantity is the clipped one-step prediction gain on the verification actions; Appendix A gives the proof and a finite-budget power condition.

### 3.5 Accumulating and Refreshing Evidence

Evidence memory $\mathcal{M}$ stores raw transitions, so every fit relinearizes the retained contexts. Accumulated conditional information $G = \sum_j C_j^\top C_j$ combines complementary directions, while only $\text{row}(C_{\text{old}}) \cap \text{row}(C_{\text{new}})$ can test a historical change (Appendix A.4). A shared–split regression (Chow, 1960) nominates a history cutoff; frozen recent-only and old-history predictors then undergo the same prospective gain test, with a separate error budget, before any evidence is deleted. The history operator $\mathcal{H}$ keeps the newest contexts after the latest verified cutoff, up to an age limit. Accepted parameters persist through both operations, and episode boundaries cancel pending candidates.

CoRD-P plans with task cost only (Plan in Algorithm 1); CoRD-A adds conditional information in unresolved directions to task-feasible action search (Wagenmaker & Jamieson, 2020) (Appendix A.6). Only executed outcomes enter memory, so probes add evidence exactly as task actions do.

## 4 Experiments

We evaluate the complete adaptation loop through three questions: **(1)** Does separating correction lifetimes improve prediction and control? **(2)** Does the model remember dynamics changes without remembering appearance changes? **(3)** Which lifetime decisions drive this behavior over repeated updates? Main comparisons use five visual-control tasks; lifetime ablations follow the same streams, and controlled experiments isolate estimation and verification.

### 4.1 Tasks, Comparisons, and Metrics

**Tasks and continual deployment.** We use PushT, PointMaze-Medium, Reacher-easy, Cartpole-swingup, and Walker-walk from the DeepMind Control Suite (DMC; Zhou et al., 2025; Tassa et al., 2018). Each task has 20 paired, 40-episode streams: ten nominal, ten appearance-shifted, ten with overlapping appearance and dynamics shifts, and ten nominal-return episodes. Each episode allows 500 actions. A 32-transition fitting block initiates one candidate, followed by at most 256 fresh verification outcomes and deployment for the remaining steps. Across a stream, conditional selection, four-snapshot verification, cache expiry, and history refresh operate together without shift labels.

**Models and comparisons.** A frozen encoder and action-conditioned predictor use 32-dimensional features. The observation adapter is affine; the dynamics adapter has 16 frozen nonlinear state–action

Table 1: **Prediction under distribution shift.** Frozen-probe state nMSE ($\times 10^{-2}$, ↓) on matched histories. CoRD-P/A use identical fixed-history updates. DMC averages three tasks.

| | PushT | | PointMaze | | DMC | |
|---|---|---|---|---|---|---|
| Method | $h = 5$ | $h = 10$ | $h = 5$ | $h = 10$ | $h = 5$ | $h = 10$ |
| Frozen | 7.8±0.6 | 16.5±1.3 | 6.5±0.4 | 13.6±0.9 | 6.4±0.4 | 12.5±0.9 |
| AdaJEPA-style E | 5.7±0.4 | 12.1±0.8 | 4.4±0.3 | 9.2±0.6 | 4.5±0.3 | 8.8±0.5 |
| AdaJEPA-style C | 4.8±0.4 | 10.1±0.9 | 4.0±0.3 | 8.4±0.7 | 4.2±0.3 | 8.3±0.5 |
| Sandwich-style E | 4.7±0.3 | 10.0±0.7 | 3.6±0.2 | 7.4±0.4 | 4.1±0.3 | 8.0±0.6 |
| Sandwich-style C | 4.4±0.3 | 9.3±0.5 | 3.4±0.2 | 7.1±0.4 | 3.9±0.2 | 7.7±0.5 |
| CoRD (P/A) | **3.2±0.3** | **6.8±0.7** | **3.0±0.2** | **6.2±0.4** | **3.5±0.3** | **6.8±0.5** |

Table 2: **Closed-loop performance at a matched interaction budget.** PushT and PointMaze report success (%, ↑); DMC reports task return (↑). Task and probe actions share one budget.

| | Success (%) | | DMC return | | |
|---|---|---|---|---|---|
| Method | PushT | PointMaze | Reacher | Cartpole | Walker |
| Frozen | 49.2±3.5 | 57.1±3.2 | 573±25 | 687±18 | 407±21 |
| AdaJEPA-style E | 59.0±3.9 | 71.7±3.1 | 708±19 | 785±23 | 508±23 |
| AdaJEPA-style C | 63.0±3.7 | 69.0±3.2 | 744±22 | 811±19 | 571±21 |
| Sandwich-style E | 68.5±3.3 | 72.1±2.9 | 809±21 | 855±18 | 613±24 |
| Sandwich-style C | 69.5±4.1 | 72.7±2.7 | 780±24 | 867±18 | 615±19 |
| CoRD-P | 73.2±3.4 | 73.5±2.5 | 804±18 | **872±20** | 695±21 |
| CoRD-A | **80.3±3.8** | **80.6±3.1** | **837±21** | 861±19 | **716±23** |

Table 3: **Adaptation speed and stale-memory cost.** Nominal-normalized post-change error area and first-return forecast error (↓). Return forecasts are scored before adaptation.

| | PushT | | PointMaze | | DMC | |
|---|---|---|---|---|---|---|
| Method | Area (ep.) | Revisit | Area (ep.) | Revisit | Area (ep.) | Revisit |
| Frozen | 14.8±1.6 | **0.96±0.03** | 15.7±1.4 | **1.00±0.03** | 14.7±1.1 | **1.01±0.02** |
| AdaJEPA-style E | 8.8±1.1 | 1.14±0.03 | 9.5±1.0 | 1.07±0.04 | 9.3±0.8 | 1.13±0.03 |
| AdaJEPA-style C | 7.3±1.0 | 2.39±0.20 | 7.8±0.9 | 1.98±0.11 | 7.8±0.8 | 2.32±0.12 |
| Sandwich-style E | 6.8±1.0 | 1.06±0.03 | 7.2±0.7 | 1.09±0.03 | 7.4±0.7 | 1.10±0.02 |
| Sandwich-style C | 5.6±0.8 | 2.39±0.17 | 6.0±0.6 | 1.87±0.14 | 6.1±0.7 | 2.17±0.10 |
| CoRD-P | 6.1±1.0 | 1.27±0.05 | 6.5±0.8 | 1.23±0.05 | 6.6±0.7 | 1.24±0.03 |
| CoRD-A | **5.1±0.7** | 1.29±0.06 | **5.3±0.6** | 1.22±0.04 | **5.5±0.6** | 1.28±0.04 |

features and 512 output weights. Frozen, AdaJEPA-style E/C, and Sandwich-style E/C share this interface, source model, readout, and planner (Wang et al., 2026; Soni et al., 2026). E resets its correction at each episode; C carries it forward. CoRD-P uses task actions, and CoRD-A allocates up to 10% of the same action budget to conditional-information probes. Lifetime ablations change one memory decision while preserving fitting data and verification opportunities.

**Metrics and uncertainty.** Forecast nMSE uses source physical-coordinate scales and a frozen evaluation probe. Control uses success or DMC return normalized to a 1,000-step scale. Recovery Area adds the positive normalized excess error over the five episodes after each change, averaged over the appearance and dynamics changes; Revisit scores the first nominal-return forecast before renewed fitting. Unless noted, ± values are 95% intervals over 20 paired streams, and paired contrasts use Holm correction over task–baseline pairs. Candidate events are aggregated within streams, so each stream counts once. Appendix B specifies the budgets and metric definitions.

### 4.2 Prediction, Control, and Recovery

**Selective memory costs no forecast accuracy.** Table 1 evaluates every method on the same action history, so CoRD-P and CoRD-A coincide. At horizon 10, CoRD reduces nMSE from 9.3 to 6.8 on PushT, 7.1 to 6.2 on PointMaze, and 7.7 to 6.8 on DMC relative to Sandwich-style C. The cache keeps corrections available for prediction even when their evidence is too weak for persistence: selection determines the memory, not the capacity of the current predictor.

**Better forecasts translate into task-dependent control gains.** Table 2 evaluates each method's own actions. Relative to the strongest shared-interface comparator, CoRD-P raises Walker return

Table 4: **Paired control differences.** CoRD-P minus the strongest shared-interface baseline per task, followed by planner contrasts. PushT/Maze differences are percentage points; DMC differences are return. Intervals are paired 95% $t$-intervals over 20 streams, $p$ is the two-sided paired $t$-test, and ✓ marks contrasts significant at $0.05$ after Holm correction over all 25 main or six planner contrasts.

| Task | Comparator | Difference | Paired interval | $p$ | Holm |
|---|---|---|---|---|---|
| PushT | Sandwich-style C | $+3.70$ | $[0.80, 6.60]$ | $0.015$ | – |
| PointMaze | Sandwich-style C | $+0.82$ | $[-1.54, 3.19]$ | $0.48$ | – |
| Reacher | Sandwich-style E | $-4.63$ | $[-20.25, 10.99]$ | $0.54$ | – |
| Cartpole | Sandwich-style C | $+5.65$ | $[-13.33, 24.63]$ | $0.54$ | – |
| Walker | Sandwich-style C | $+79.07$ | $[61.92, 96.23]$ | $< 0.001$ | ✓ |
| CoRD-A minus raw information (six-contrast planner family) | | | | | |
| PushT | Raw information | $+6.04$ | $[3.26, 8.82]$ | $< 0.001$ | ✓ |
| PointMaze | Raw information | $+6.09$ | $[3.36, 8.82]$ | $< 0.001$ | ✓ |
| DMC | Raw information | $+6.38$ | $[-2.39, 15.15]$ | $0.14$ | – |

Table 5: **Prediction use and persistent memory across complete streams.** Per-stream counts over 100 streams (five tasks, 20 each; mean $\pm$ 95% interval). Cache uses count temporary-correction deployments. Spurious and Resolved@5 are percentages of streams: spurious streams retain a correction during the initial nominal or appearance-only phases; resolved streams hold a beneficial retained correction by the fifth dynamics episode.

| Method | Cache uses | Retentions | Spurious (%) | Resolved@5 (%) | Refreshes |
|---|---|---|---|---|---|
| CoRD-P | $17.9 \pm 0.6$ | $1.8 \pm 0.1$ | $0.0$ | $77.0$ | $2.0 \pm 0.1$ |
| CoRD-A | $13.8 \pm 0.1$ | $2.0 \pm 0.1$ | $0.0$ | $100.0$ | $2.0 \pm 0.1$ |
| No temporary cache | $0.0 \pm 0.0$ | $2.0 \pm 0.2$ | $0.0$ | $100.0$ | $2.0 \pm 0.2$ |
| Full-fit validation | $12.0 \pm 0.7$ | $2.5 \pm 0.1$ | $42.0$ | $100.0$ | $2.0 \pm 0.1$ |
| No history refresh | $21.3 \pm 0.5$ | $1.8 \pm 0.1$ | $0.0$ | $44.0$ | $0.0 \pm 0.0$ |

by $79.1$ (significant after Holm correction) and PushT success by $3.7$ percentage points (Table 4); on PointMaze ($+0.8$ points), Reacher ($-4.6$ return), and Cartpole ($+5.7$ return) the paired intervals include zero. Closed-loop forecast curves follow these method-dependent actions (Figure 6).

**Correction lifetime sets the cost of returning to nominal.** Carried adapters (C) recover quickly but return with stale memory, at 1.9–2.4 times nominal error; episode-reset adapters (E) return cleanly but relearn after every reset and recover more slowly (Table 3). CoRD-P lowers PushT Revisit error from $2.39$ to $1.27$ versus Sandwich-style C at a comparable Area ($6.1$ vs. $5.6$), and CoRD-A recovers fastest on every benchmark: CoRD keeps verified dynamics and lets compensation expire.

**Targeted probing buys evidence.** CoRD-A improves PushT recovery Area to $5.1$ and success to $80.3\%$, and raises success or return over CoRD-P on four of the five tasks. At a matched interaction budget, conditional-information probing outperforms raw-information probing by about six success points on PushT and PointMaze after Holm correction, whereas the DMC difference is not significant (Tables 4 and 6). What a probe measures matters, not only how many actions it spends.

### 4.3 The Complete Loop across Repeated Updates

**Local use and persistence are distinct outcomes.** Across 4,000 CoRD-P attempts, 1,648 authorize local use only, 40 authorize persistence only, 140 authorize both, and 2,172 authorize neither (Table 12): the model often improves its current prediction without changing its lasting dynamics.

**Each lifetime check guards a different failure.** Full-fit validation retains the selected increment whenever the full local prediction improves, without testing it after cache expiry; the no-cache ablation removes all compensation before scoring (Table 5). One transfers a local benefit to an untested memory; the other gives up capacity before retention is settled. CoRD accepts local use without granting persistence, so a low persistent-update rate still leaves the prediction corrected.

**Evidence refresh and parameter expiry play different roles.** Figure 3c separates confirmed history rejection from age-based removal; neither resets verified dynamics. The no-refresh control resolves $44\%$ of streams within five dynamics episodes, versus $77\%$ for CoRD-P, linking evidence renewal to useful retention (task-level counts in Appendix C.4).

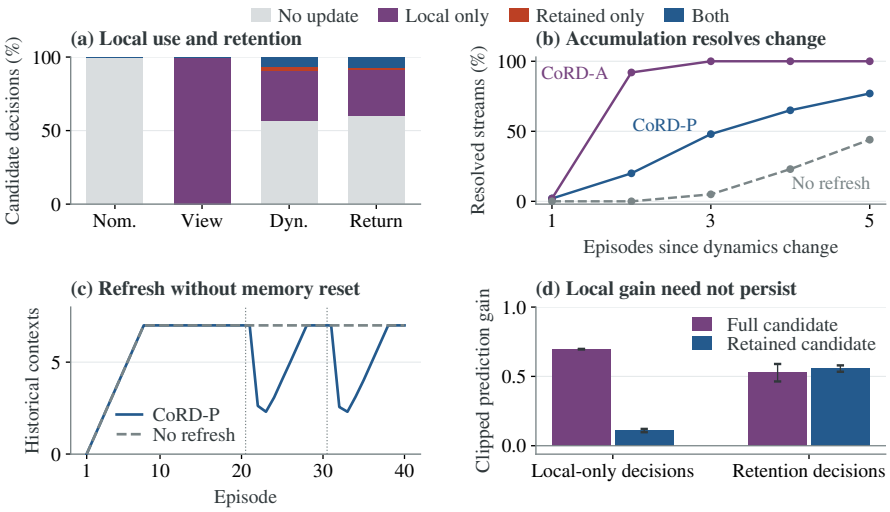


Figure 3: **From fitted correction to changing memory.** (a) CoRD-P decision outcomes by phase; (b) streams holding a beneficial retained correction; (c) stored evidence contexts; (d) full versus retained-only gains, with intervals over complete streams.

**Every lifetime decision earns its place.** Ablating any single component raises PushT $h = 10$ nMSE from 6.8 to between 7.2 and 8.3 (Table 8). Without the change test, nominal drift grows from 0.65 to 2.59, and accepting on fitting data raises it to 1.98; keeping observation corrections across contexts raises hidden-shift error from 1.36 to 1.89, and dropping the history audit raises it to 2.01. Dropping the source-fixed coordinate scales, which set the units of the split, costs the most (8.3), and the full method takes at most 5.6 ms more per update than any ablation.

### 4.4 Why the Lifetime Split Matters

**Removing ambiguous directions can destroy a useful fit.** In 1,000 paired attribution trials at overlap $\rho = 0.75$ (Table 10), conditional selection lowers wrong-component drift from 0.482 to 0.179 under appearance change and from 7.539 to 0.611 under dynamics change ($\times 10^{-3}$), but raises mixed-change prediction error from 5.660 to 9.460. Proposition 1 shows why the cache removes this conflict: at the least-squares solution the cache restores the joint fit, so ambiguity restricts what persists rather than what helps predict, and the local forecast keeps the full correction.

**The factorization separates capacity from memory.** Eq. (6) prices shrinkage exactly. Ridge is strongest on dense changes, testing and shrinkage help on sparse or null ones, and feature adapters fit contact-dependent changes better than affine ones (Tables 9–11).

**Future evidence must support the model left after expiry.** In the visual benchmark, validating only the full local predictor retains corrections before any dynamics change in 42% of streams; testing the retained predictor reduces this to 0% (Table 5). In a Pendulum feedback calibration with 120 streams per condition, the rule promotes $0/120$ candidates under nominal dynamics and under correlated sensor noise, but $66/120$ under changed gravity and torque (Appendix B.1).

## 5 Conclusion

Continual adaptation must decide what to remember as well as what to change. CoRD answers with two tests. Conditional geometry decides what *can* be remembered: only the part of a correction that observation changes cannot produce. Prospective comparison decides what *should* be remembered: only what still predicts better once the cache is gone. Choosing what to remember does not limit the current fit, and verified increments carry dynamics changes forward without appearance changes. The same question arises wherever a deployed model learns from its own experience, from a robot meeting new materials to an agent adapting to a new user: what helps now need not deserve to last. Explain the present, then verify what survives: adaptation may be fast, but memory must be earned.

## AI USE STATEMENT

We used AI-based tools to assist with language editing, LaTeX typesetting, and figure scripting. All research ideas, methods, experiments, analyses, and conclusions are the authors' own, and every AI-assisted passage was checked by the authors.

## REFERENCES


Elahe Arani, Fahad Sarfraz, and Bahram Zonooz. Learning fast, learning slow: A general continual learning method based on complementary learning system. In *International Conference on Learning Representations*, 2022.

Sebastian Arnold, Georgios Gavrilopoulos, Benedikt Schulz, and Johanna Ziegel. Sequential model confidence sets. *arXiv preprint arXiv:2404.18678v4*, 2026.

Jimmy Ba, Geoffrey E. Hinton, Volodymyr Mnih, Joel Z. Leibo, and Catalin Ionescu. Using fast weights to attend to the recent past. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Philip J Ball, Cong Lu, Jack Parker-Holder, and Stephen Roberts. Augmented world models facilitate zero-shot dynamics generalization from a single offline environment. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 619–629. PMLR, 2021.

Yoav Benjamini and Daniel Yekutieli. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4):1165–1188, 2001. doi: 10.1214/aos/1013699998.

Geng Chen, Wendong Zhang, Han Lu, Siyu Gao, Yunbo Wang, Mingsheng Long, and Xiaokang Yang. Continual predictive learning from videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

Yo Joong Choe and Aaditya Ramdas. Comparing sequential forecasters. *Operations Research*, 72(4):1368–1387, 2024. doi: 10.1287/opre.2021.0792.

Gregory C. Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3): 591–605, 1960. doi: 10.2307/1910133.

A. Philip Dawid. Present position and potential developments: Some personal views. Statistical theory: The prequential approach. *Journal of the Royal Statistical Society. Series A (General)*, 147(2):278–292, 1984.

Mario Döbler, Robert A. Marsden, and Bin Yang. Robust mean teacher for continual and gradual test-time adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7704–7714, 2023.

Ragnar Frisch and Frederick V. Waugh. Partial time regressions as compared with individual trends. *Econometrica*, 1(4):387–401, 1933.

Xiang Fu, Ge Yang, Pulkit Agrawal, and Tommi Jaakkola. Learning task informed abstractions. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3480–3491. PMLR, 2021.

Yossi Gandelsman, Yu Sun, Xinlei Chen, and Alexei A. Efros. Test-time training with masked autoencoders. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Gene Golub and Victor Pereyra. Separable nonlinear least squares: The variable projection method and its applications. *Inverse Problems*, 19(2):R1–R26, 2003. doi: 10.1088/0266-5611/19/2/201.

David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2555–2565. PMLR, 2019.

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.

Nathan Halko, Per-Gunnar Martinsson, and Joel A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011. doi: 10.1137/090771806.

Nicklas Hansen, Rishabh Jangir, Yu Sun, Guillem Alenyà, Pieter Abbeel, Alexei A. Efros, Lerrel Pinto, and Xiaolong Wang. Self-supervised policy adaptation during deployment. In *International Conference on Learning Representations*, 2021a.

Nicklas Hansen, Hao Su, and Xiaolong Wang. Stabilizing deep Q-learning with ConvNets and vision transformers under data augmentation. In *Advances in Neural Information Processing Systems*, volume 34, 2021b.

Nicklas Hansen, Hao Su, and Xiaolong Wang. TD-MPC2: Scalable, robust world models for continuous control. In *International Conference on Learning Representations*, 2024.

Nicklas A Hansen, Hao Su, and Xiaolong Wang. Temporal difference learning for model predictive control. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8387–8406. PMLR, 2022.

Per Christian Hansen. The truncated SVD as a method for regularization. *BIT Numerical Mathematics*, 27: 534–553, 1987. doi: 10.1007/BF01937276.

Irina Higgins, Arka Pal, Andrei Rusu, Loic Matthey, Christopher Burgess, Alexander Pritzel, Matthew Botvinick, Charles Blundell, and Alexander Lerchner. DARLA: Improving zero-shot transfer in reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1480–1490. PMLR, 2017.

Geoffrey E. Hinton and David C. Plaut. Using fast weights to deblur old memories. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, pp. 177–186, 1987.

Trung-Hieu Hoang, Duc Minh Vo, and Minh N. Do. Persistent test-time adaptation in recurring testing scenarios. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. doi: 10.1214/20-AOS1991.

Kartik Jhawar and Lipo Wang. When is test-time adaptation identifiable from unlabeled evidence? *arXiv preprint arXiv:2609.11235*, 2026.

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

Dharshan Kumaran, Demis Hassabis, and James L. McClelland. What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, 20(7):512–534, 2016.

JB Lanier, Kyungmin Kim, Armin Karamzade, Yifei Liu, Ankita Sinha, Kat He, Davide Corsi, and Roy Fox. Adapting world models with latent-state dynamics residuals. *arXiv preprint arXiv:2504.02252*, 2025.

Misha Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srinivas. Reinforcement learning with augmented data. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Chi Ho Leung, Ashish R. Hota, and Philip E. Paré. Online identification of time-varying systems using excitation sets and change point detection. *arXiv preprint arXiv:2406.10349*, 2024.

Jacob Levy, Tyler Westenbroek, Kevin Huang, Fernando Palafox, Patrick Yin, Shayegan Omidshafiei, Dong-Ki Kim, Abhishek Gupta, and David Fridovich-Keil. Simulation distillation: Pretraining world models in simulation for rapid real-world adaptation. In *Proceedings of Robotics: Science and Systems*, 2026. doi: 10.15607/RSS.2026.XXII.017.

Shuang Li, Yihuai Gao, Dorsa Sadigh, and Shuran Song. Unified video action model. In *Proceedings of Robotics: Science and Systems*, 2025. doi: 10.15607/RSS.2025.XXI.074.

Jian Liang, Dapeng Hu, and Jiashi Feng. Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6028–6039. PMLR, 2020.

Jiaming Liu, Senqiao Yang, Peidong Jia, Renrui Zhang, Ming Lu, Yandong Guo, Wei Xue, and Shanghang Zhang. ViDA: Homeostatic visual domain adapter for continual test time adaptation. In *International Conference on Learning Representations*, 2024.

Michael C. Lovell. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, 58(304):993–1010, 1963.

James L. McClelland, Bruce L. McNaughton, and Randall C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457, 1995.

Anusha Nagabandi, Chelsea Finn, and Sergey Levine. Deep online learning via meta-learning: Continual adaptation for model-based RL. *arXiv preprint arXiv:1812.07671*, 2018.

Anusha Nagabandi, Ignasi Clavera, Simin Liu, Ronald S. Fearing, Pieter Abbeel, Sergey Levine, and Chelsea Finn. Learning to adapt in dynamic, real-world environments through meta-reinforcement learning. In *International Conference on Learning Representations*, 2019.

Fahim Faisal Niloy, Sk Miraj Ahmed, Dripta S. Raychaudhuri, Samet Oymak, and Amit K. Roy-Chowdhury. Effective restoration of source knowledge in continual test time adaptation. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2091–2100, 2024.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofo Chen, Shijian Zheng, Peilin Zhao, and Mingkui Tan. Efficient test-time model adaptation without forgetting. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 16888–16905. PMLR, 2022.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Zhiquan Wen, Yaofo Chen, Peilin Zhao, and Mingkui Tan. Towards stable test-time adaptation in dynamic wild world. In *International Conference on Learning Representations*, 2023.

Andrej Orsula, Miguel Olivares-Mendez, and Carol Martinez. Reward-free continual adaptation for resilient space robots. *arXiv preprint arXiv:2608.23452*, 2026.

Quang Pham, Chenghao Liu, Doyen Sahoo, and Steven C. H. Hoi. Learning fast and slow for online time series forecasting. In *International Conference on Learning Representations*, 2023.

Aleksandr Podkopaev and Aaditya Ramdas. Tracking the risk of a deployed model and detecting harmful distribution shifts. In *International Conference on Learning Representations*, 2022.

Ori Press, Steffen Schneider, Matthias Kümmerer, and Matthias Bethge. RDumb: A simple approach that questions our progress in continual test-time adaptation. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Max Simchowitz, Horia Mania, Stephen Tu, Michael I. Jordan, and Benjamin Recht. Learning without mixing: Towards a sharp analysis of linear system identification. In *Proceedings of the 31st Conference on Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pp. 439–473. PMLR, 2018.

Krishnam Soni, Aditya Sehgal, Vedant Dave, and Elmar Rueckert. Sandwich-residuals: Parameter-efficient test-time adaptation of world models. *arXiv preprint arXiv:2609.21740*, 2026.

Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei A. Efros, and Moritz Hardt. Test-time training with self-supervision for generalization under distribution shifts. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9229–9248. PMLR, 2020.

Yuval Tassa, Yotam Doron, Alistair Muldal, Tom Erez, Yazhe Li, Diego de Las Casas, David Budden, Abbas Abdolmaleki, Josh Merel, Andrew Lefrancq, Timothy Lillicrap, and Martin Riedmiller. DeepMind control suite. *arXiv preprint arXiv:1801.00690*, 2018.

Philip S. Thomas, Georgios Theocharous, and Mohammad Ghavamzadeh. High confidence policy improvement. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2380–2388, 2015.

Andrew Wagenmaker and Kevin Jamieson. Active learning for identification of linear dynamical systems. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 3487–3582. PMLR, 2020.

Chi Wan, Kangrui Wang, Yuan Si, Pingyue Zhang, and Manling Li. WorldAgen: Unified state-action prediction with test-time world model training. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(22): 18584–18592, 2026. doi: 10.1609/aaai.v40i22.38925.

Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *International Conference on Learning Representations*, 2021.

Qin Wang, Olga Fink, Luc Van Gool, and Dengxin Dai. Continual test-time domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7201–7211, 2022.

Ying Wang, Oumayma Bounou, Yann LeCun, and Mengye Ren. AdaJEPA: An adaptive latent world model. *arXiv preprint arXiv:2606.32026*, 2026.

Denis Yarats, Ilya Kostrikov, and Rob Fergus. Image augmentation is all you need: Regularizing deep reinforcement learning from pixels. In *International Conference on Learning Representations*, 2021.

Longhui Yuan, Binhui Xie, and Shuang Li. Robust test-time adaptation in dynamic scenarios. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15922–15932, 2023.

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.

Marvin Zhang, Sergey Levine, and Chelsea Finn. MEMO: Test time robustness via adaptation and augmentation. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Xiangteng Zhang, Yang Guan, Bo Zhang, Hongyang Li, Ya-Qin Zhang, and Shengbo Eben Li. On the identifiability of controlled world models. *arXiv preprint arXiv:2607.22430*, 2026.

Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. DINO-WM: World models on pre-trained visual features enable zero-shot planning. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 79115–79135. PMLR, 2025.

# A Conditional Geometry and Predictive Certificates

## A.1 Coordinates and Explicit Jacobians

For one endpoint–horizon pair, write $\tilde{z}_0 = E_\alpha(x_{k-h})$, $\tilde{z}_\ell = F_\beta(\tilde{z}_{\ell-1}, u_{k-h+\ell-1})$, $J_\ell^z = \partial_z F_\beta$ and $J_\ell^\beta = \partial_\beta F_\beta$ at $(\tilde{z}_{\ell-1}, u_{k-h+\ell-1})$, and $\Phi_{h:\ell} = J_h^z \cdots J_\ell^z$, with empty products equal to $I$. For $\Psi_\alpha(x) = [\mathrm{diag}(D_\alpha E_0(x)), I]$, direct differentiation gives

$$\begin{aligned} J_o^{k,h} &= D_\alpha^{-1}[\Psi_\alpha(x_k) - \Phi_{h:1}\Psi_\alpha(x_{k-h})] - [\mathrm{diag}(r_{k,h}), 0], \\ J_d^{k,h} &= -D_\alpha^{-1} \sum_{\ell=1}^{h} \Phi_{h:\ell+1} J_\ell^\beta. \end{aligned} \tag{12}$$

Stack these blocks over completed endpoint–horizon pairs; $m = d_z \sum_j |\mathcal{R}_j|$ counts scalar residuals, not independent observations. Source-fixed coordinate scales $D_o, D_d$ and a positive-definite weighting $\Omega_0$ define

$$y = -\Omega_0^{-1/2} r, \quad A = \Omega_0^{-1/2} J_o D_o, \quad B = \Omega_0^{-1/2} J_d D_d, \quad P = I - AA^\dagger, \quad C = PB. \tag{13}$$

The geometry below is deterministic for the realized weighted matrices. It does not require $\Omega_0$ to remain the target covariance. Whitening and coordinate normalization determine which directions are selected.

Classical variable projection gives (Golub & Pereyra, 2003)

$$\min_a \|y - Aa - Bb\|^2 = \|P(y - Bb)\|^2, \qquad G = B^\top PB = B^\top B - B^\top A(A^\top A)^\dagger A^\top B. \tag{14}$$

Under a correct fixed-design Gaussian model with variance $\nu^2 I$, $G/\nu^2$ is also the profiled Fisher information.

## A.2 Local Fit and Persistent Memory

**Proof of Proposition 1.** Let $C = U\Sigma V^\top$, let $V_R$ contain a subset of nonzero right singular vectors, and let $V_U$ complete the orthonormal right basis. Set $C_R = CV_R$, $Z = [A, BV_U]$, and $\widehat{d} = \Sigma_R^{-1} U_R^\top Py$. Since $C_R$ is orthogonal to $A$ and $CV_U$, and $BV_U = AA^\dagger BV_U + CV_U$, it is orthogonal to $Z$. Also $BV_R = AA^\dagger BV_R + C_R$. Hence

$$\mathrm{col}([A, B]) = \mathrm{col}(Z) \oplus \mathrm{col}(C_R), \qquad P_Z BV_R = C_R, \quad P_Z = I - ZZ^\dagger.$$

Minimizing over the local coefficients gives $\|P_Z y - C_R d\|^2$. Its orthogonal decomposition into the unrestricted least-squares residual and $C_R(\widehat{d} - d)$ proves Eq. (6). In particular, the local completion $q = Z^\dagger(y - BV_R\widehat{d})$ attains unrestricted joint-LS fit, while $\widehat{d}(y + Zq') = \widehat{d}(y)$ for every temporary compensation $Zq'$.

Projection geometry thus justifies assigning lifetimes without deleting locally useful model directions. The statement is pathwise for each realized design and frozen split, even when either was selected from earlier data. Within this stacked fit the unresolved coefficient is shared; allowing independent unresolved coefficients in every historical block would enlarge $Z$ and need not preserve the identity. Ridge, clipping, and nonlinear steps can change the exact fit equality; the actual parameter snapshots are therefore checked prospectively.

**Why local fit does not identify persistent effects.** For $\bar{b} \in \ker(PB)$, choose $\bar{a}$ satisfying $A\bar{a} = B\bar{b}$. The parameter pairs $(a, b) = (0, 0)$ and $(-\bar{a}, \bar{b})$ induce identical local observation distributions under the same additive noise. If a later nuisance-free response is $\Gamma b$, any retained estimator obeys

$$\max_{j \in \{0,1\}} \mathbb{E}_j \|\Gamma\widehat{b} - \Gamma b_j\|^2 \;\geq\; \tfrac{1}{4}\|\Gamma\bar{b}\|^2. \tag{15}$$

Indeed, its distribution is identical in the two worlds, and the average squared distance to 0 and $\Gamma\bar{b}$ is at least $\|\Gamma\bar{b}\|^2/4$. An informative executed action can break this equivalence. Thus the bound motivates temporary compensation, not permanent abstention after new evidence becomes available.

### A.3 Certificates for Adaptive Nonlinear Forecasts

All four parameter snapshots in Eq. (9) are frozen before verification. For an executed action $u_t$, their common observable target is $E_0(x_{t+1})$. The pulled-back prediction and bounded loss for snapshot $\chi = (\alpha, \beta)$ are

$$\widehat{z}^{\chi}_{t+1} = D_\alpha^{-1}[F_\beta(E_\alpha(x_t), u_t) - \alpha_b], \qquad \ell^{\chi}_t = \min\{\|E_0(x_{t+1}) - \widehat{z}^{\chi}_{t+1}\|^2/M, 1\}. \tag{16}$$

The positive scale $M$ is fixed before verification; it defines the metric rather than an assumed target noise variance. Forecasts and actions are measurable with respect to $\mathcal{F}_t$ before the outcome arrives. For each pair $(\chi, \chi')$, let $g_t = \ell^{\chi}_t - \ell^{\chi'}_t \in [-1, 1]$, $\mu_t = \mathbb{E}[g_t \mid \mathcal{F}_t]$, $S_n = \sum_{t=1}^n g_t$, and $\bar{S}_n = \sum_{t=1}^n \mu_t$.

For $\Lambda = \{1/16, 1/8, 1/4, 1/2, 1\}$ with $K = 5$, the cumulative-gain lower confidence bound is

$$\mathrm{LCB}_{\delta,n} = S_n - \mathcal{B}(n, \delta), \qquad \mathcal{B}(n, \delta) = \min_{\lambda \in \Lambda} \left\{ \frac{\log(K/\delta)}{\lambda} + \frac{\lambda n}{2} \right\}. \tag{17}$$

Dividing by $n$ gives the average-gain bound. This is the same boundary as Eq. (11). Conditional Hoeffding's lemma gives $\mathbb{E}[\exp\{\lambda(g_t - \mu_t)\} \mid \mathcal{F}_t] \le \exp(\lambda^2/2)$. Therefore $\exp\{\lambda(S_n - \bar{S}_n) - \lambda^2 n/2\}$ is a nonnegative supermartingale. Ville's inequality bounds a crossing of $\log(K/\delta)$ by $\delta/K$. A union bound over the finite grid proves

$$\Pr\{\exists n : \bar{S}_n < \mathrm{LCB}_{\delta,n}\} \le \delta. \tag{18}$$

This standard confidence-sequence construction tolerates adaptive actions, stopping, and conditional variances (Choe & Ramdas, 2024; Howard et al., 2021; Arnold et al., 2026). It does not use Jacobian exogeneity or correct source whitening.

**Deployment and retention.** For the candidate from block $j$, allocate $\delta_j/3$ to each contrast $0 - F$, $L - F$, and $0 - R$. Both retention bounds must be positive at the same initial decision; separate past crossings cannot be combined. A jointly accepted correction can subsequently remain deployed. A positive $0 - F$ bound permits local deployment of the full fit. Positive $L - F$ and $0 - R$ bounds jointly permit persistence: the retained increment helps both with the local cache present and on its own. If only local deployment passes, every new correction expires. If only persistence passes, deploy $R$; if both pass, deploy $F$ and then expire its local cache to recover exactly $R$. One common step must make $F, L, R$ individually admissible, not merely their summed correction.

For $\sum_j \delta_j \le \delta_{\mathrm{total}}$, all three contrasts, every attempt, and every decision time satisfy their bounds simultaneously with probability at least $1 - \delta_{\mathrm{total}}$. On this event, every newly deployed snapshot has positive cumulative conditional gain over the incumbent on its verification prefix: $F$ by $0 - F$, or $R$ by $0 - R$. A retained increment additionally improves $F$ over the cache-matched control $L$. The coefficients deployed or retained must be exactly those scored: post-validation filtering or refitting creates another candidate. The certificate concerns cumulative conditional predictive advantage on the executed verification actions.

**Finite-budget power.** At a fixed verification horizon $n$, suppose $\bar{S}_n \ge n\underline{\mu}$ almost surely for a contrast. Martingale Hoeffding gives $\Pr\{S_n < \bar{S}_n - \sqrt{2n\log(1/\eta)}\} \le \eta$. Consequently a positive certificate is obtained by $n$ with probability at least $1 - \eta$ whenever

$$n\underline{\mu} > \mathcal{B}(n, \delta) + \sqrt{2n\log(1/\eta)}. \tag{19}$$

For the two retention contrasts use $\eta/2$ and their smaller advantage. Weakly informative actions can hide a genuine dynamics change; CoRD-A spends part of the action budget on conditional information for this reason.

### A.4 Historical Coverage and Predictive Conflict

With context-specific observation corrections and shared dynamics,

$$\min_{\{a_j\}, b} \sum_j \lambda_j \|y_j - A_j a_j - B_j b\|^2 = \min_b \sum_j \lambda_j \|P_j(y_j - B_j b)\|^2, \qquad G_t = \sum_j \lambda_j C_j^\top C_j. \tag{20}$$

For positive weights, $\ker G_t = \bigcap_j \ker C_j$; therefore the row-space span measures accumulated coverage. Common coordinates and relinearization are necessary. Complementary coverage need not permit revalidation: with $C_1 = [1, 0]$ and $C_2 = [0, 1]$, neither block checks the other's response.

A shared–split residual statistic can nominate a candidate historical cutoff. If $X_s$ shares dynamics and $X_f$ fits separate dynamics while both allow separate observation nuisance, its reference statistic is

$$F = \frac{(\mathrm{RSS}_s - \mathrm{RSS}_f)/d_1}{\mathrm{RSS}_f/d_2}, \qquad d_1 = \operatorname{rank} X_f - \operatorname{rank} X_s, \quad d_2 = m - \operatorname{rank} X_f. \tag{21}$$

Here $d_1 = \dim[\mathrm{row}(C_1) \cap \mathrm{row}(C_2)]$, by subtracting the shared and split model ranks. The $F_{d_1,d_2}$ law holds under the fixed-design Gaussian model (Chow, 1960); $d_1 = 0$ supplies no structural test. In deployment the statistic nominates a comparison, and the prospective certificate below decides deletion.

Freeze otherwise matched old-history and recent-only predictors after nomination. A fresh bounded-gain certificate for old-history loss minus recent-only loss verifies a predictive conflict before evidence-driven deletion. Allocate a separate history-audit error budget using Eq. (18). Let $j_\star$ be the newest verified cutoff, or $-\infty$ if no comparison passes. With source-fixed age limit $N_{\mathrm{age}}$ and capacity $K_H$, the retained history is

$$\mathcal{H}(\mathcal{M}) = \mathrm{newest}_{K_H}\{\mathcal{D}_j \in \mathcal{M} : j \geq j_\star,\ t - t_j \leq N_{\mathrm{age}}\}. \tag{22}$$

Remove horizons crossing the cutoff. Scheduled expiry manages resources, while the certificate governs evidence-driven deletion. Accepted parameters persist independently of evidence expiry.

### A.5 Fixed-Design Reference Calculations

The following classical calculations define the fixed-design diagnostics. For $X = [A, B]$, $d = m - \operatorname{rank} X > 0$, and a correct fixed-design model with isotropic Gaussian noise,

$$\widehat{\nu}^2 = \|(I - XX^\dagger)y\|^2/d, \qquad T_i = \mathrm{u}_i^\top P y/\widehat{\nu}, \qquad p_i = 2\{1 - F_{t_d}(|T_i|)\}. \tag{23}$$

Under the directional null, numerator and residual variance are independent; hence $T_i \sim t_d$. BY correction then controls within-fit FDR (Benjamini & Yekutieli, 2001); in deployment, the prospective certificate controls errors. For fixed directional weights and $\theta_i = v_i^\top b^\star$,

$$\mathbb{E}\|\widehat{b} - b^\star\|^2 = \sum_{\sigma_i > 0} \left[(1 - w_i)^2\theta_i^2 + w_i^2\nu^2/\sigma_i^2\right] + \|\Pi_{\ker C} b^\star\|^2. \tag{24}$$

### A.6 Excitation, Action Choice, and Numerical Approximation

For unresolved directions $V_U$ and candidate-conditioned predicted Jacobians, the conditional information bonus is

$$\mathcal{J}_{\mathrm{info}}(u) = \log\det\left[I + \kappa^{-1}V_U^\top \widetilde{B}(u)^\top \widetilde{P}(u)\widetilde{B}(u)V_U\right]. \tag{25}$$

For fixed-count CEM elites $\mathcal{E}$ and task cost $\mathcal{C}$, define

$$u^0 = \arg\min_{u \in \mathcal{E}} \mathcal{C}(u), \qquad u^A = \arg \max_{\substack{u \in \mathcal{E}: \\ \mathcal{C}(u) \leq \mathcal{C}(u^0) + \epsilon}} [\mathcal{J}_{\mathrm{info}}(u) - \zeta\|u - u^0\|^2]. \tag{26}$$

Plan executes the first action of $u^0$ (CoRD-P) or $u^A$ (CoRD-A), with source-fixed $\epsilon, \zeta$. The candidate Jacobians describe predicted information, not newly observed evidence, and are recomputed for every proposed action sequence.

For $s_{j+1} = as_j + bu_j$ and $x_j = s_j + c$ with known $a$, the $(c, b)$ sensitivities are

$$A = \begin{bmatrix} a - 1 \\ a^2 - 1 \end{bmatrix}, \quad B = \begin{bmatrix} -u_1 \\ -(au_0 + u_1) \end{bmatrix}, \quad \det[A\ B] = a(a - 1)(u_1 - u_0).$$

For $a \notin \{0, 1\}$, varied actions distinguish gain from offset (Simchowitz et al., 2018). The same logic holds for gains: for $s_{t+1} = s_t + \mathrm{diag}(g)u_t$, disjoint data conceal a gain change from $(1, 2)$ to $(5, 2)$, and the probe $u = (0.4, 0)$ separates displacements $(0.4, 0)$ and $(2, 0)$ (Figure 4).

Products $Cw = P(Bw)$ and $C^\top w' = B^\top(Pw')$ admit Jacobian-vector and vector-Jacobian implementations. Krylov or randomized decompositions (Halko et al., 2011) avoid storing the full Jacobian.

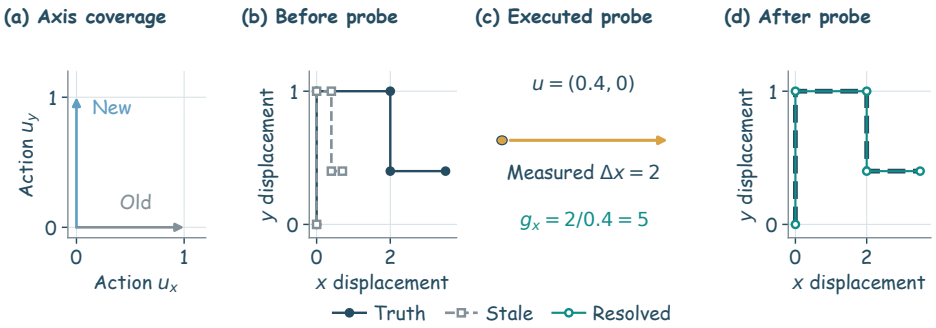


Figure 4: **A horizontal probe detects the changed horizontal gain.** Exact gain dynamics: disjoint coverage (a), stale rollout (b), discriminating probe (c), and corrected rollout (d). Forecasts share initial states and actions.

## B Evaluation Protocol and Metric Definitions

**Tasks and episode order.** PushT, PointMaze-Medium, Reacher-easy, Cartpole-swingup, and Walker-walk each have 20 paired streams. The primary order is fixed: episodes 1–10 nominal, 11–20 appearance shift, 21–30 a dynamics shift added while the appearance shift persists, and 31–40 nominal return. Nominal environment and rendering parameters are both restored at return. Stream indices pair starts, goals, intervention strengths, and task randomness across methods. The within-episode hidden-change comparison in Table 8 is a separate stress condition. All methods use the same fixed order and 500-action episode allowance. CoRD completes one fitting–verification–deployment cycle per episode, giving 40 candidates per stream. Its local cache expires at each episode boundary; persistent dynamics carry over. The history retains at most eight completed fitting contexts, with an age limit of eight episodes. Conflict-based refresh and age expiry remove evidence, not learned dynamics parameters.

**Interventions and available information.** Appearance changes affect rendering; dynamics changes affect transition response. PushT tests controller/contact response, PointMaze actuation/damping, and DMC actuation or body parameters. Adaptation receives images and executed actions, with native proprioception where provided. Physical states, intervention labels, canonical images, success, and target rewards are evaluator-only. A source-defined goal or frozen reward head supplies task cost, not adaptation supervision. Coordinate bounds and covariances are fixed source assets; deployment retains no source examples.

**Source model and matched update rules.** Tables 1–3 compare update rules on one source model, adapter family, loss normalization, and planner. The frozen visual encoder produces 32-dimensional latents; the dynamics residual uses 16 frozen nonlinear state–action features and 512 trainable output weights. The observation adapter has 32 log-scales and 32 biases. Frozen is an unchanged checkpoint; the other four baseline configurations are two transferred rules with episode-reset (E) and persistent (C) states. AdaJEPA-style controls adapt the representation-side interface; Sandwich-style controls jointly adapt the observation and transition interfaces (Wang et al., 2026; Soni et al., 2026). The shared parameterization isolates update and retention rules. The regression comparisons below additionally evaluate Joint-LS, ridge, pure testing, shrinkage, sequential fitting, and BIC.

**Fitting and four-snapshot verification.** Algorithm 1 uses the history operator in Eq. (22), planner in Eq. (26), and paired lower bound in Eq. (17). Only completed residuals enter fitting. The incumbent 0, full fit $F$, local-only control $L$, and retained-only fit $R$ remain frozen throughout verification; each predicts the same executed action before its outcome arrives. A common step must make all four parameter snapshots admissible. No coefficient is filtered or refitted after its validation.

Each candidate uses 32 new fitting transitions. Verification uses up to 256 one-step outcomes, with decisions at $n \in \{32, 64, 128, 256\}$. The five-element grid $\Lambda = \{1/16, 1/8, 1/4, 1/2, 1\}$ defines the confidence sequence; each of the three contrasts receives $\delta_j/3$, where $\delta_j = 0.05/40$. A separate $0.05$ stream budget is split uniformly across the 40 history comparisons. Stable modes use the singular-value threshold $\tau = 0.05$ in normalized coordinates, and the step size $\gamma$ is the first factor in $\{1, 1/2, 1/4, 1/8, 1/16\}$ that gives a bounded step with smaller nonlinear context loss. We select

Table 6: Planner controls at the same interaction budget. PushT and PointMaze report success (%); DMC reports the mean of three task returns. Latency includes adaptation and planning on PushT. Values are mean ± 95% interval over 20 paired streams.

| Planner | PushT ↑ | PointMaze ↑ | DMC ↑ | Latency (ms) ↓ |
|---|---|---|---|---|
| CoRD-P | $73.2\pm3.4$ | $73.5\pm2.5$ | $790\pm14$ | $25.7\pm1.1$ |
| Random probing | $72.3\pm3.6$ | $72.6\pm2.7$ | $783\pm16$ | $31.0\pm1.1$ |
| Raw-information | $74.2\pm2.9$ | $74.5\pm3.1$ | $798\pm15$ | $33.9\pm1.1$ |
| CoRD-A | $80.3\pm3.8$ | $80.6\pm3.1$ | $805\pm17$ | $35.4\pm1.6$ |
| Robust planner | $72.9\pm3.4$ | $73.2\pm3.1$ | $788\pm14$ | $38.3\pm1.1$ |

$\lambda_q = 1.0$ on an independent source-calibration split from $\{10^{-3}, 10^{-2}, 10^{-1}, 1\}$ and keep it fixed for all target streams. The loss scale $M$ is the 99.5th percentile of squared source prediction errors on separate calibration episodes and remains fixed at deployment. Local deployment requires a positive $0 - F$ bound; persistence requires positive $L - F$ and $0 - R$ bounds at the same decision time. The first decision with either permission ends verification. If only local deployment passes, all new coefficients expire; if only persistence passes, deploy $R$; if both pass, deploy $F$ and expire its cache to $R$. Resets cancel pending candidates. Fitting outcomes cannot verify their own candidate, and verification outcomes enter subsequent fits only after that candidate's decision. A history cutoff verified on the same fresh outcomes is queued for the next fit; it cannot alter an ongoing candidate or its four frozen forecasts. CoRD-A probes occupy at most 50 of the 500 action slots, including actions during verification. Multi-step evaluation at $h \in \{5, 10\}$ is separate from one-step verification.

**Forecast error and direct attribution.** For evaluator-only physical coordinate $\pi_j$, source standard deviation $\varsigma_j$, and frozen-probe forecast $\widehat{\pi}_j$, define

$$\mathrm{nMSE}(h) = \frac{1}{d_\pi} \sum_{j=1}^{d_\pi} \left( \frac{\widehat{\pi}_{t+h,j} - \pi_{t+h,j}}{\varsigma_j} \right)^2 . \tag{27}$$

PushT uses position and periodic orientation, PointMaze position and velocity, and DMC fixed task-state coordinates. Source scaling and angular encoding remain fixed. Table 1 uses a shared action history; Figure 6a uses each method's closed-loop history at $h = 10$, so their forecast averages need not coincide. The linear attribution experiment measures each head's function on a separate fixed bank (Appendix C).

**Control and recovery.** PushT and PointMaze report success percentages. For the common 500-action allowance, DMC reports 1000 times mean per-step reward, preserving the standard $[0, 1000]$ return scale. DMC averages its three tasks within each stream before cross-stream inference. Action counts and planner search budgets are paired; probes consume the same interaction budget as task actions. For stream $s$, let $\mathrm{err}_{s,k}$ be the closed-loop 10-step error in episode $k$ and $\mathrm{err}_s^0$ its nominal reference. Define

$$\mathrm{Area}_s = \frac{1}{2} \sum_{c \in \{11,21\}} \sum_{i=0}^{4} [\mathrm{err}_{s,c+i}/\mathrm{err}_s^0 - 1]_+, \qquad \mathrm{Revisit}_s = \mathrm{err}_{s,31}/\mathrm{err}_s^0. \tag{28}$$

$\mathrm{Area}_s$ is measured in episodes of normalized excess error; $\mathrm{Revisit}_s$ is dimensionless and scores the first return forecast before renewed fitting. Latency averages adaptation and planning per decision.

**Paired comparisons and multiplicity.** The reporting unit is a complete stream, not an episode. Marginal intervals use $t_{19,0.975}$ over 20 stream summaries. For methods 1 and 2, $\Delta_s = Y_s^{(1)} - Y_s^{(2)}$ gives the paired interval $\overline{\Delta} \pm t_{19,0.975} s_\Delta / \sqrt{20}$ and a two-sided paired test. The main analysis applies Holm correction to all 25 task-by-baseline contrasts. A separate six-contrast family compares CoRD-A with CoRD-P and raw information on PushT, PointMaze, and within-stream DMC means. Table 4 gives paired intervals, $p$-values, and Holm decisions. Marginal overlap does not determine paired significance.

**Fixed-design diagnostics.** Gaussian noise with standard deviation $0.25$ generates 1,000 trials per condition. The null experiment uses conditional singular values $(4, 2, 1)$, threshold $0.5$, and 43 residual degrees of freedom, with BY level $0.05$. Compatibility uses two 38-row blocks with column norms 3 and 2, testing a coefficient change of $0.6$ with $F_{2,72}$ at $0.05$. Prospective validation fits on 16 samples and evaluates 32 independent new samples. The extended comparisons examine null, sparse, and dense changes, bidirectional attribution, and capacity.

Table 7: **State-input feedback calibration.** Rates (%) over 120 independent Pendulum streams per condition. CS columns report Wilson 95% intervals at 256 future transitions. Delay is the median number of future transitions among promoted streams; 54 dynamics-change streams remain unpromoted. BY and Chow are fitting-context diagnostics; Chow permits separate nuisance intercepts in both hypotheses.

| Condition | BY | Chow | CS update [95% CI] | CS discard [95% CI] | Delay |
|---|---|---|---|---|---|
| Nominal | 0.0 | 0.0 | 0.0 [0.0, 3.1] | 0.0 [0.0, 3.1] | – |
| Serial-noise shift | 4.2 | 2.5 | 0.0 [0.0, 3.1] | 0.0 [0.0, 3.1] | – |
| Dynamics shift | 100.0 | 97.5 | 55.0 [46.1, 63.6] | 69.2 [60.4, 76.7] | 128 |

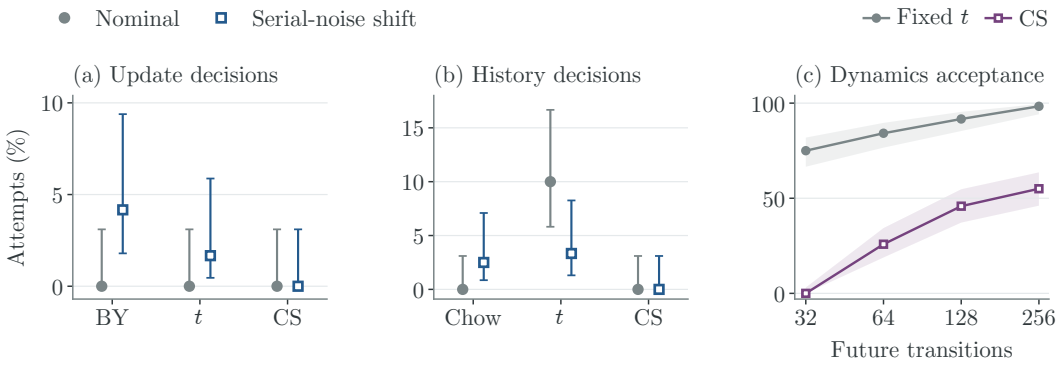


Figure 5: **Predictive evidence determines retention under feedback.** Pendulum calibration: (a) updating and (b) history rejection at 256 steps under unchanged physical dynamics; (c) acceptance under gravity/torque change as verification increases. Intervals use 120 independent streams. Fixed-window $t$ decisions use the same trajectories as the sequential comparisons.

### B.1 Feedback Calibration

**Environment and source model.** The complementary calibration uses noisy state inputs in Gymnasium Pendulum-v1. The model predicts normalized angular-velocity increments from four raw state–action inputs, a constant, and 16 frozen tanh features, giving 21 fitted source coefficients. A residual candidate adds a local intercept and four state–action coefficients. To hold candidate geometry constant across the three noise and dynamics conditions, the calibration fixes $V_R = I$ and an empty $V_U$. This fixes candidate geometry while varying feedback, noise, and dynamics.

One hundred nominal source episodes supply 10,000 transitions: 6,500 for fitting, 2,000 for regularization selection, and 1,500 for loss-scale calibration, separated by episode. Source and adapter ridge penalties are $0.001$ and $1.0$; $M = 2.8318 \times 10^{-5}$ is fixed from the calibration split. The shooting planner evaluates 17 candidate action sequences over eight steps and replans after each noisy observation. Its actions depend on the deployed model; unnoised simulator states and future observations are not available to the predictor.

**Streams and interventions.** There are 120 independently reset streams per condition, paired across conditions by initial seed: 360 streams in total. Each stream contains 32 nominal historical transitions, 32 fitting transitions, and 256 fresh verification transitions. Conditions are unchanged dynamics; sensor noise with fourfold marginal standard deviation and AR coefficient $0.85$; and gravity changing from 10 to 20 with torque gain changing from 1 to $0.6$. The environment time limit is extended to 10,000 steps so no time-limit reset interrupts a stream.

**Snapshots, decisions, and denominators.** For this calibration, the four snapshots are $(0, 0)$, $(\alpha, \beta)$, $(\alpha, 0)$, and $(0, \beta)$. Each stream fits one candidate and uses $\delta = 0.05$ across the three promotion contrasts. The exact retained-only snapshot controls subsequent actions at the first passing decision; all frozen snapshots continue to predict those same actions. The retained model $R$ supplies feedback after promotion; the other snapshots supply paired forecasts. CS promotion counts are cumulative through each reported horizon; fixed-window tests evaluate that horizon alone. A separate old-versus-recent retained-model comparison has its own $0.05$ history budget; its decision is logged without refitting the candidate. The two budgets form separate error families. Fixed-window Student-$t$

Table 8: Component ablations on PushT. Forecast error is state nMSE ($\times 10^{-2}$); hidden-shift error is nominal-normalized. Nominal drift is squared dynamics displacement ($\times 10^{-3}$). All columns are lower-is-better; values are mean $\pm$ 95% interval over 20 paired streams.

| Variant | $h = 10$ | Hidden shift | Nominal drift | Latency (ms) |
|---|---|---|---|---|
| Full CoRD-P | $6.8 \pm 0.7$ | $1.36 \pm 0.07$ | $0.65 \pm 0.04$ | $25.7 \pm 1.1$ |
| No change test | $8.0 \pm 0.9$ | $1.41 \pm 0.09$ | $2.59 \pm 0.17$ | $23.5 \pm 0.4$ |
| In-sample acceptance | $7.7 \pm 0.9$ | $1.43 \pm 0.09$ | $1.98 \pm 0.13$ | $21.8 \pm 0.4$ |
| Observation-first fit | $8.1 \pm 0.8$ | $1.57 \pm 0.08$ | $1.36 \pm 0.07$ | $22.2 \pm 0.3$ |
| Persistent observation | $7.5 \pm 0.9$ | $1.89 \pm 0.11$ | $1.26 \pm 0.08$ | $25.2 \pm 0.5$ |
| Current context only | $7.8 \pm 0.7$ | $1.33 \pm 0.06$ | $0.70 \pm 0.03$ | $20.1 \pm 0.2$ |
| No history audit | $7.2 \pm 0.7$ | $2.01 \pm 0.11$ | $0.76 \pm 0.04$ | $22.7 \pm 0.4$ |
| Unscaled coordinates | $8.3 \pm 0.9$ | $1.55 \pm 0.09$ | $1.45 \pm 0.09$ | $23.8 \pm 0.4$ |

decisions are recorded on the same trajectories, not deployed as separate policies or treated as anytime tests. The Chow null and alternative both allow separate historical and current nuisance intercepts; only the null shares the four dynamics coefficients.

**Acceptance, conflict, and power.** Table 7 and Figure 5 summarize the rates. By 256 outcomes, neither nominal nor noise-shift streams promote a correction or declare a predictive history conflict: each rate is $0/120$, with Wilson 95% interval $[0, 3.10]\%$. Under the dynamics intervention, promotion reaches $66/120 = 55.0\%$ ($[46.08, 63.61]\%$) and history conflict $83/120 = 69.17\%$ ($[60.42, 76.73]\%$). Median promotion delay is 128 transitions among accepted streams; the other 54 streams are right-censored at 256 and remain in the power denominator.

The incumbent/retained loss-clipping fractions are respectively $0.365\%/0.348\%$ under nominal dynamics, $6.475\%/7.376\%$ under serial-noise shift, and $97.30\%/70.13\%$ under the large dynamics intervention. The clipping scale remains fixed throughout all target experiments.

**Known predictive null.** A separate diagnostic uses 10,000 independent bounded martingale streams with three contrasts and 256 outcomes. Gain amplitudes and signs depend on previous outcomes, while fresh Rademacher innovations give exactly zero conditional means. Checking every time point yields 118 streams with any false crossing: $1.18\%$ with Wilson 95% interval $[0.986, 1.411]\%$, at the simultaneous $5\%$ budget. The calibration records all 92,160 future transitions, frozen predictions, gains, and decisions.

# C Additional Experiments and Ablations

**Paired task comparisons.** Table 4 reports paired effect sizes. The strongest comparator is shown for display; Holm adjustment covers all 25 main comparisons. In the six-comparison planner family, the conditional-information bonus outperforms raw-information probing on PushT and PointMaze.

**Prediction, memory, and action budgets.** Figure 6 summarizes closed-loop 10-step forecasts, first-return error, PushT component ablations, and planner controls at a matched interaction budget; Figure 7 illustrates forecasts under single changes, and Figure 8 shows rendered rollouts in which the local cache and the retained correction diverge.

**Matched component removals.** Removing the change test raises nominal dynamics displacement from $0.65$ to $2.59$, and accepting on fitting data raises it to $1.98$. Keeping observation corrections across contexts increases hidden-change error from $1.36$ to $1.89$, and removing the history audit raises it to $2.01$. Current-context fitting has slightly less hidden-change error but worse aggregate forecasts. These controls distinguish parameter lifetime, history retention, and acceptance within the shared model and adaptation interface.

## C.1 Estimator Risk under Dense, Sparse, and Null Changes

Each block has 64 rows, four context-local nuisance columns, and six dynamics columns. Profiling leaves 60 rows per block; the independent variance estimate has $60N - 6$ degrees of freedom. The truths are $(1, -0.75, 0.4, 0.25, -0.15, 0.08)$, $(1, 0, 0.4, 0, 0, 0)$, and zero. All methods receive the same coefficient statistics and residual sum of squares. A separate 4,000-problem source calibration selects the ridge penalty $0.2511886$ under a fixed isotropic coefficient prior; this value is unchanged across target conditions.

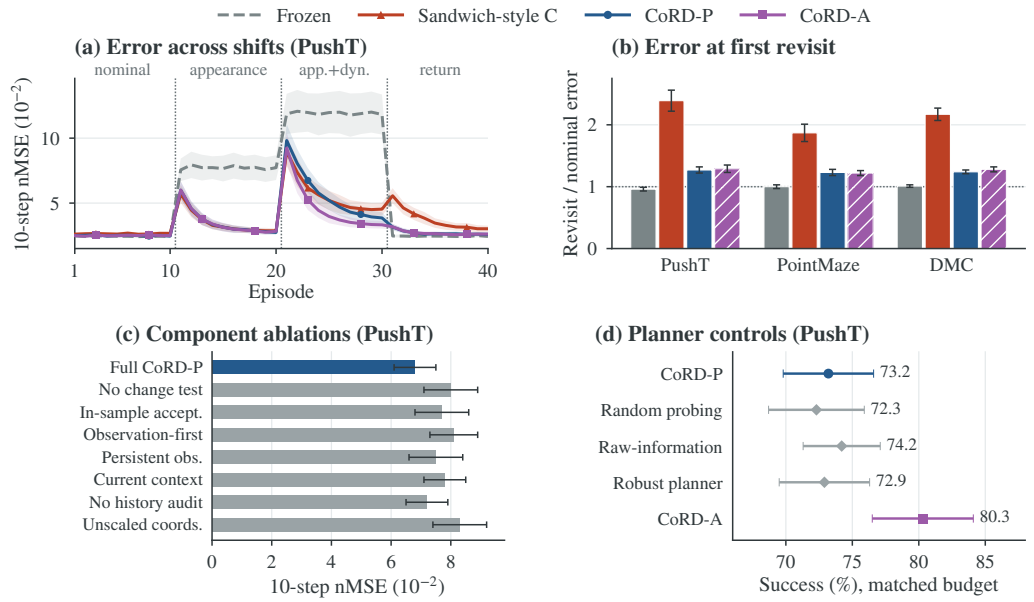


Figure 6: **Adaptation, memory, and probing.** (a) Closed-loop PushT $h = 10$ error across the four ten-episode phases, where episodes 21–30 combine both shifts; Table 1 instead fixes the action history. (b) First-return error relative to nominal (Table 3). (c) PushT component ablations at $h = 10$ (Table 8). (d) PushT success of planners at a matched interaction budget (Table 6). Uncertainty: 95% intervals over 20 paired streams.

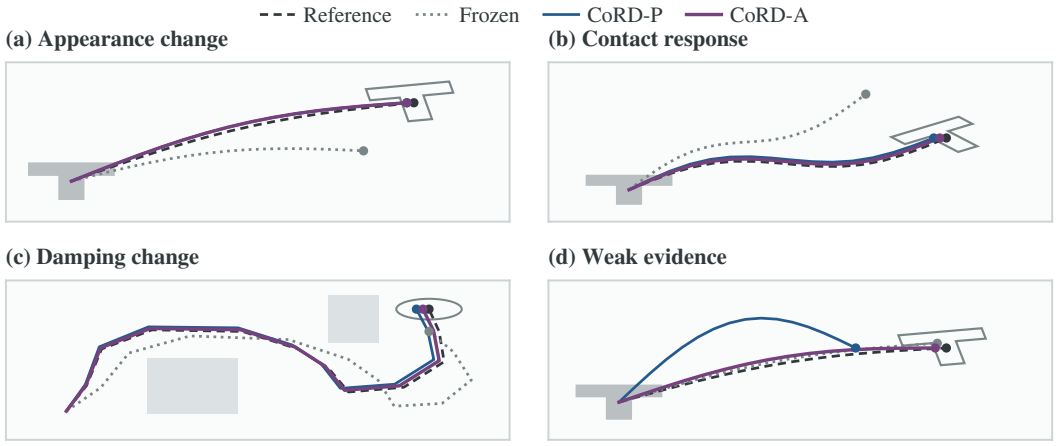


Figure 7: **Motion forecasts under different shifts.** Appearance, contact-response, and damping changes produce different prediction errors; weak evidence leaves an incomplete correction. Paths share their origin and reference request. These four single-change cases illustrate failure modes; mixed attribution is quantified separately in Table 10.

Table 9 and Figure 9 report the resulting risks. Pure testing uses two-sided Student probabilities and BY level 0.05. The shrinkage comparison changes only the retained coefficient magnitudes. BIC examines all $2^6$ supports using the same residual likelihood and parameter-count penalty. Dense, sparse, and null conditions test different bias–variance regimes, and all-zero changes separate proposal frequency from proposal size.

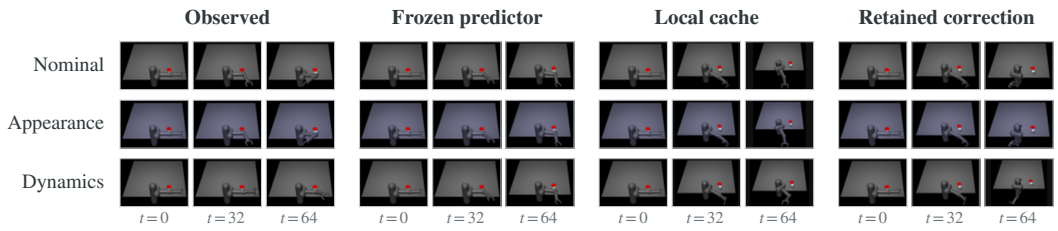


Figure 8: **Illustrative rendered rollouts.** Frames at $t = 0, 32, 64$ in a simulated arm-pushing scene outside the benchmark, under nominal, appearance-shifted, and dynamics-shifted conditions. Columns show the observed environment and rollouts from the frozen predictor, the predictor with its local cache, and the retained correction alone. In this demonstration the cache is released after $t = 32$, while the retained correction persists.

Table 9: **Estimation under dense, sparse, and absent changes.** Parameter MSE ($\times 10^{-3}$, $\downarrow$), mean $\pm$ 95% Monte Carlo interval over 1,000 paired draws. Ridge is source-calibrated; Test uses BY without shrinkage; Shrink adds positive-part shrinkage. All six directions are estimable.

| Blocks | Change | LS | Ridge | Test | Shrink | BIC |
|---|---|---|---|---|---|---|
| 1 | Dense | 200.23±10.69 | 42.07±1.64 | 64.13±6.78 | 58.77±5.38 | 91.14±9.31 |
| 1 | Sparse | 191.07±10.57 | 37.17±1.45 | 37.66±5.43 | 34.43±4.27 | 66.04±8.94 |
| 1 | Null | 202.52±11.29 | 37.42±1.53 | 8.34±4.49 | 6.87±3.70 | 57.58±10.20 |
| 16 | Dense | 12.75±0.72 | 10.21±0.54 | 11.94±0.58 | 11.17±0.48 | 12.16±0.57 |
| 16 | Sparse | 13.08±0.75 | 10.47±0.57 | 1.92±0.45 | 1.69±0.36 | 1.83±0.44 |
| 16 | Null | 12.29±0.69 | 9.80±0.52 | 0.36±0.27 | 0.29±0.22 | 0.89±0.39 |
| 64 | Dense | 3.16±0.19 | 2.97±0.17 | 4.48±0.15 | 4.25±0.14 | 5.11±0.13 |
| 64 | Sparse | 3.05±0.16 | 2.88±0.15 | 0.35±0.08 | 0.31±0.06 | 0.24±0.05 |
| 64 | Null | 3.09±0.18 | 2.91±0.16 | 0.11±0.08 | 0.09±0.07 | 0.14±0.09 |

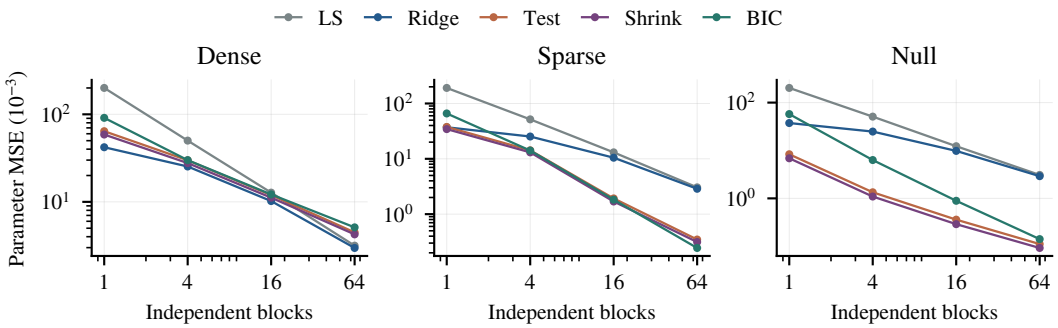


Figure 9: **Selection trades dense-change bias for sparse-change stability.** Parameter MSE for five estimators on identical Gaussian draws. Dense, sparse, and null changes share the same design and noise. Each point averages 1,000 paired trials; intervals appear in Table 9.

### C.2 Direct Attribution and the Value of Validation

For 64-row fitting and future blocks, draw mutually orthonormal $Q_1, Q_2 \in \mathbb{R}^{64\times 3}$ and set $A = 4Q_1$, $B = 4(\rho Q_1 + \sqrt{1-\rho^2}Q_2)$. True appearance change is $a = (0.6, -0.4, 0.3)$; true dynamics change is $b = (0.8, -0.5, 0.35)$. Each may independently be zero. Separate fixed 128-row Gaussian banks $\Xi_o, \Xi_d$ define $\|\Xi_o(\widehat{a} - a)\|^2/128$ and $\|\Xi_d(\widehat{b} - b)\|^2/128$. Bank entries have variance $1/3$. The banks never fit or accept a candidate. In a single-change condition, the unchanged head's error is exactly its wrong-head displacement. Under mixed changes, displacement does not identify a wrong head.

Joint-LS, source ridge, one-pass observation-then-dynamics fitting, ten alternating sweeps, and conditional testing/shrinkage are each crossed with the same validation switch. The conditional diagnostic applies BY and shrinkage to all six coefficients, isolating the attribution principle in a

Table 10: **Direct attribution with identical future validation.** Fixed-bank errors ($\times 10^{-3}$, $\downarrow$), overlap $\rho = 0.75$, 1,000 paired trials. Displacement measures wrong-head drift only for an unchanged head. Mixed rows compare each head with its own truth; Cond. denotes conditional estimation.

| Condition / metric | LS | Ridge | Sequential | Alternating | Cond. |
|---|---|---|---|---|---|
| View: dynamics drift | $0.48 \pm 0.16$ | $0.43 \pm 0.15$ | $0.02 \pm 0.01$ | $0.48 \pm 0.16$ | $0.18 \pm 0.18$ |
| Dynamics: view drift | $7.54 \pm 0.37$ | $6.95 \pm 0.35$ | $151.52 \pm 1.50$ | $7.50 \pm 0.37$ | $0.61 \pm 0.27$ |
| Mixed: view MSE | $7.33 \pm 0.36$ | $6.59 \pm 0.33$ | $150.96 \pm 1.53$ | $7.29 \pm 0.36$ | $15.36 \pm 0.96$ |
| Mixed: dynamics MSE | $8.33 \pm 0.41$ | $7.54 \pm 0.37$ | $150.81 \pm 5.43$ | $8.28 \pm 0.40$ | $15.39 \pm 0.98$ |
| Mixed: future MSE | $5.66 \pm 0.20$ | $5.35 \pm 0.19$ | $72.44 \pm 1.24$ | $5.64 \pm 0.20$ | $9.46 \pm 0.52$ |

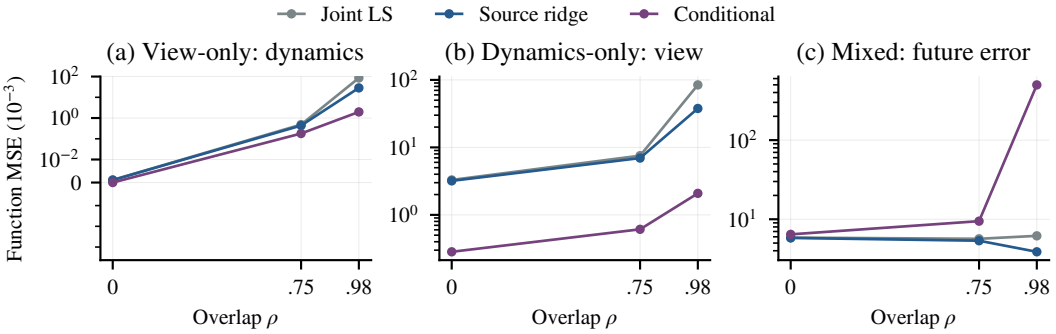


Figure 10: **Better attribution need not imply lower prediction error.** Identical fresh-data validation is applied to Joint-LS, ridge, and symmetric conditional estimation. (a,b) Drift of the unchanged component. (c) Future error when both components change. Higher overlap makes attribution harder; each point averages 1,000 trials.

controlled linear setting. For frozen future predictions $f_0, f_1$, independent Gaussian noise of known source variance gives the paired mean-loss difference variance

$$\mathrm{Var}(\overline{\ell_0 - \ell_1} \mid f_0, f_1) = 4\nu^2 \|f_1 - f_0\|^2 / m_{\mathrm{future}}^2.$$

The three one-sided bounds each use $0.05/3$. All comparisons use the same future noise, and no candidate is refitted on those outcomes. This supplies an exact controlled acceptance rule, distinct from source-calibrated block validation on a correlated deployment stream.

The results expose both effects of restriction (Table 10 and Figure 10). At $\rho = 0.75$, conditional estimation suppresses cross-head leakage but increases mixed prediction error relative to validated LS and ridge. At $\rho = 0.98$, conditional mixed excess risk is $0.4990$, versus $0.00617$ for validated LS. Jointly predictive directions can remain weakly attributable. The lifetime experiments below track the two decisions separately: local deployment and persistent memory.

### C.3 Adapter Capacity Is Separate from Update Allocation

The capacity audit uses state $(x, y, v_x, v_y)$, action $(u_x, u_y)$, and a frozen nominal integrator with damped, action-driven velocity. A smooth contact-region perturbation depends jointly on position, velocity, and action. Twelve independent replicates each have 1,024 fitting pairs and 16,384 test pairs; fitting targets have noise standard deviation $0.002$. All methods see identical evaluator targets. This supervised audit isolates approximation capacity.

A disjoint source sample of 4,096 pairs fixes feature centering, scaling, and PCA of six inputs plus 96 seeded tanh features. Widths 4 and 16 include an intercept; the full last layer uses 103 features. The four-output correction has matrix rank at most four, irrespective of feature width. Affine fitting solves its bounded least-squares problem; other fits use source-noise/prior ridge $0.0016$. Affine correction is strongest on the representable affine change. More features improve contact-region approximation but increase nominal noise fitting; even the full layer leaves approximation error.

Table 11: **Adapter approximation error.** Evaluator-only regression on a four-state, two-action contact surrogate. Held-out MSE ($\times 10^{-6}$), mean $\pm$ 95% interval over 12 seeds. Width denotes fixed feature dimension, not matrix rank. These are capacity fits, not adaptation returns.

| Adapter | Parameters | Nominal | Affine change | Contact change |
|---|---|---|---|---|
| Frozen | 0 | $0.000 \pm 0.000$ | $955.809 \pm 2.990$ | $332.493 \pm 3.225$ |
| Affine | 8 | $0.006 \pm 0.002$ | $0.006 \pm 0.002$ | $276.485 \pm 2.449$ |
| Residual width 4 | 16 | $0.016 \pm 0.003$ | $390.082 \pm 1.988$ | $272.328 \pm 2.306$ |
| Residual width 16 | 64 | $0.065 \pm 0.008$ | $7.253 \pm 0.060$ | $225.827 \pm 1.639$ |
| Full last layer | 412 | $0.454 \pm 0.026$ | $0.454 \pm 0.026$ | $191.219 \pm 2.454$ |

Table 12: **Correction lifetimes and evidence refresh by task.** CoRD-P counts per 40-episode stream (20 streams per task). Local-only and retained-only decisions are disjoint; Both authorizes local use followed by retention. History columns count deleted and age-expired evidence contexts.

| Task | Local-only | Retained-only | Both | Neither | Deleted | Expired |
|---|---|---|---|---|---|---|
| PushT | $16.4 \pm 1.9$ | $0.2 \pm 0.2$ | $1.8 \pm 0.2$ | $21.6 \pm 1.9$ | $13.3 \pm 0.4$ | $18.7 \pm 0.4$ |
| PointMaze | $15.8 \pm 1.4$ | $0.4 \pm 0.3$ | $1.5 \pm 0.4$ | $22.2 \pm 1.3$ | $13.5 \pm 0.3$ | $18.5 \pm 0.3$ |
| Reacher | $17.0 \pm 1.8$ | $0.5 \pm 0.3$ | $1.2 \pm 0.4$ | $21.4 \pm 1.6$ | $13.2 \pm 0.5$ | $18.8 \pm 0.5$ |
| Cartpole | $17.1 \pm 1.9$ | $0.4 \pm 0.3$ | $1.4 \pm 0.4$ | $21.1 \pm 1.8$ | $13.4 \pm 0.5$ | $18.6 \pm 0.5$ |
| Walker | $15.9 \pm 1.4$ | $0.6 \pm 0.4$ | $1.2 \pm 0.4$ | $22.3 \pm 1.2$ | $12.6 \pm 1.5$ | $19.4 \pm 1.5$ |

### C.4 Repeated Local Deployment and Persistent Memory

Table 12 expands the complete-stream comparison by task. Every method starts from the same source predictor and receives the same initial conditions and intervention schedule. The deployed correction and raw-transition history carry different state: a history cutoff changes the evidence available to a later fit, whereas cache expiry changes the current predictor. Verified dynamics remain in both cases.

**Candidate outcomes.** For each task, 20 streams contribute 800 candidate attempts. The four outcome categories are mutually exclusive. Local only means $v = 1, p = 0$ at the first passing decision; persistent only means $v = 0, p = 1$; both means $v = p = 1$; neither means no permission through 256 outcomes. A local-only decision closes the candidate and cannot subsequently become a persistent decision. Cache expiry returns a local-only correction to the incumbent and a both-pass correction to its frozen retained snapshot. Separate counts track nominations, confirmed history refreshes, and age-based removals. Evidence removal leaves persistent dynamics intact.

**Matched lifetime controls.** Full-fit validation uses the local-benefit decision to retain the selected dynamics increment without the incremental and cache-free comparisons. The no-cache control removes observation and unresolved compensation before prospective scoring. No-refresh retains historical contexts until age expiry without an evidence-driven cutoff. All controls share 40 fitting opportunities, the 500-action episode allowance, and the same source scaling. These interventions separate the benefit of the current fit, the effect of its lifetime, and the cost of stale evidence.

**History scheduling.** Each completed fitting context has a timestamp and remains eligible for relinearization until confirmed conflict or age expiry. A prospective history comparison starts from frozen recent-only and historical predictors; its decision is queued for the next fitting cycle. Updating a history record never changes the forecasts already being scored. Age expiry is applied before fitting, using the same eight-episode limit in every variant. The active variant spends at most 10% of the common action budget on probing; only executed actions can add evidence to subsequent fits.

### C.5 Fixed-Design Checks of Validation and History Tests

Table 13 summarizes these checks. Under the null, conditional BY selection proposes an update in $2.4\%$ of fits. All nominal scalar fits reduce fitting loss, while only $1.5\%$ pass future validation (Figure 11): mean MSE rises from $0.0591$ on fitting samples to $0.0674$ on fresh samples. Compatibility separately detects $99.7\%$ of the specified changes and rejects $5.5\%$ of unchanged pairs. Without row-space overlap its test has zero degrees of freedom: a small pooled residual then cannot verify old dynamics.

Table 13: **Nominal updates and detected changes.** Fixed-design diagnostics; 1,000 trials per condition, with 95% Wilson intervals.

| Condition / event | Rule | Rate (%) | 95% interval |
|---|---|---|---|
| Nominal / proposal | Conditional BY test | 2.4 | [1.6, 3.5] |
| Nominal / acceptance | In-sample decrease | 100.0 | [99.6, 100.0] |
| Nominal / acceptance | Future validation | 1.5 | [0.9, 2.5] |
| Unchanged dynamics / split | Shared–split $F$ test | 5.5 | [4.2, 7.1] |
| Dynamics change / split | Shared–split $F$ test | 99.7 | [99.1, 99.9] |

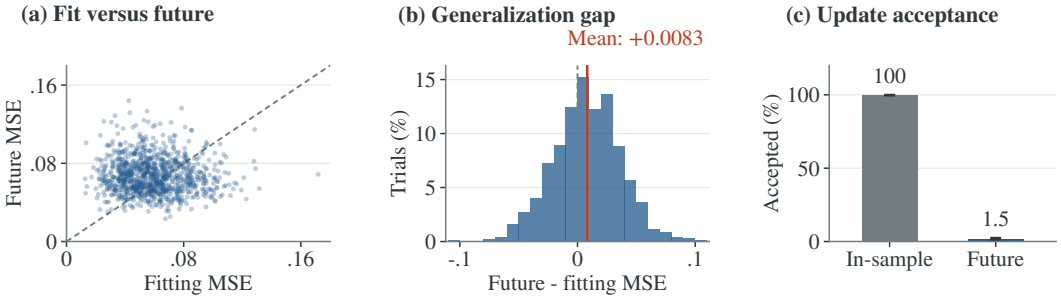


Figure 11: **Fitting error versus future usefulness.** (a) Paired candidate MSE. (b) Future-minus-fitting MSE. (c) Acceptance rates with Wilson intervals. The 1,000 nominal trials use 16 fitting and 32 independent future samples.