BEFORE THE NEXT OUTCOME: HISTORY-AWARE TEST-TIME ADAPTATION FOR LATENT WORLD MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Test-time adaptation of latent world models typically relies on real transitions observed after the agent acts, leaving the model unadapted at the beginning of deployment. We study *pre-outcome test-time adaptation*: adapting a pretrained world model using only the target-domain history that is already available, before the next real state is observed. This setting is challenging because historical variation mixes appearance nuisance with genuine dynamics changes, while representation and dynamics adaptation are coupled through a shared latent space. We propose **HATTA** (**H**istory-**A**ware **T**est-**T**ime **A**daptation), which exploits the two forms of evidence already present in the observed history. HATTA uses historical transition compatibility to select counterfactual visual hypotheses and construct a dynamics-aware representation correction, while observed local motion is used to calibrate the action-conditioned dynamics predictor. A latent-space coordination step controls the accepted representation update and re-encodes the original history before dynamics adaptation. HATTA requires no new future outcome, source-domain data, reward, goal, or shift label. Experiments on PushT visual shifts and PointMaze dynamics shifts show that short histories provide useful pre-outcome signals for both visual and dynamics adaptation. The unified method retains most of the gains of the two specialized branches and achieves overall point estimates close to a feedback-based AdaJEPA reference without using new real transition feedback. These results suggest that target-domain history can support meaningful world-model adaptation before additional interaction feedback becomes available.

1 INTRODUCTION

World models provide internal models of the environment for robot planning and decision-making by predicting how actions change future states (Hafner et al., 2019; 2020; Zhou et al., 2025). However, when the test environment differs from the training distribution, a frozen world model can accumulate prediction errors and produce unreliable plans (Wang et al., 2026a). Recent work therefore adapts world models during deployment (Wang et al., 2026a; Soni et al., 2026). For example, AdaJEPA performs test-time adaptation inside closed-loop MPC: it executes an action, observes the real successor state, updates the world model with latent prediction error, and replans (Wang et al., 2026a).

In this feedback-based schedule, adaptation follows the collection of a new real transition (Wang et al., 2026a). Early in deployment, however, the model has the least target-domain feedback while already needing to plan with its current predictions. We therefore study an earlier adaptation window: *can a pretrained world model adapt using only its already observed target-domain history, before the next real state becomes available?* We call this setting *pre-outcome test-time adaptation*. At test time, the model can only use the current observation-action history. It does not execute extra probe actions, and it does not access source-domain samples, rewards, or task goals. Figure 1 contrasts this setting with feedback-based world-model adaptation.

Standard image and video TTA provides useful mechanisms for visual robustness, including statistic alignment and augmentation-based objectives (Schneider et al., 2020; Zhang et al., 2022; Lin et al.,

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.

2023). However, these objectives alone do not specify how to adapt an action-conditioned world model. In such a model, temporal variation also contains the dynamics that the model is designed to predict. Appearance changes such as color, noise, and lighting should be corrected, while real dynamics changes, such as those caused by density or damping, should be preserved and used to recalibrate the model. The distinction between nuisance and control-relevant information is also central to world-model representation learning and robustness evaluation (Wang et al., 2022b; Maes et al., 2026). The problem is further coupled by the latent world-model structure: the visual encoder and the dynamics predictor share the same latent space, so changing the representation also changes the coordinates in which the predictor operates, whose geometry can affect planning (Wang et al., 2026b). World-model TTA must therefore address three related challenges: how to construct supervision without a new future target, how to avoid removing real dynamics evidence as visual nuisance, and how to coordinate representation and dynamics adaptation in a shared latent space.

Our key observation is that, although the next future outcome is unavailable, the observed short history already contains real transitions. HATTA uses these transitions to address the three challenges above. First, the frozen world model predicts already observed successors from historical prefixes. These predictions define a *historical compatibility* score for counterfactual visual hypotheses and provide a dynamics-aware correction signal for the encoder. Second, HATTA preserves the latent motion already observed in the history and uses it as target-local dynamics evidence to calibrate the action-conditioned responses of the predictor. Third, because visual adaptation changes the latent coordinates used by the predictor, HATTA controls the accepted representation displacement with a fixed strength parameter, re-encodes the original history, and performs dynamics adaptation in the resulting latent space.

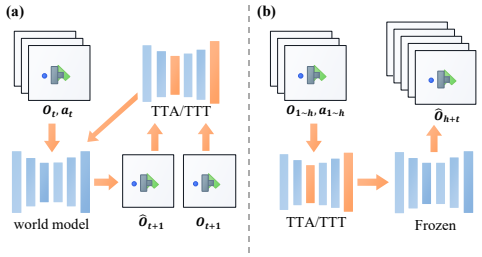


Figure 1: **Adaptation timing.** (a) Feedback-based world-model adaptation updates after a newly observed successor. (b) HATTA adapts from the already observed target history before the next outcome arrives.

Based on this design, we propose **HATTA** (**H**istory-**A**ware **T**est-**T**ime **A**daptation). HATTA uses only the target-domain history available before adaptation. It does not construct an unobserved future state as a pseudo-target and does not require source data, rewards, goals, or shift labels. The method follows one unified pipeline: historical compatibility drives visual correction, latent coordination controls the representation update, and observed local motion drives dynamics calibration.

We evaluate HATTA on the PushT visual-shift and PointMaze dynamics-shift settings derived from AdaJEPA. The results show that short observed histories already provide useful pre-outcome adaptation signals. Visual adaptation improves planning under visual OOD shifts, while dynamics adaptation improves performance under dynamics OOD shifts. Applying both updates at full strength introduces a cross-factor trade-off, whereas latent coordination retains most of the gains from both specialists. In the current evaluation, HATTA also achieves overall point estimates close to a feedback-based AdaJEPA reference without using new real transition feedback. We do not interpret this result as showing that pre-outcome adaptation can replace later real feedback. Instead, it shows that the target-domain history already available before the next feedback arrives can itself support meaningful world-model adaptation.

Our main contributions are:

1. **Pre-outcome world-model TTA.** We study how a trained latent world model can adapt from its existing target-domain observation-action history before the next real outcome becomes available.

2. **HATTA.** We construct representation and dynamics adaptation signals from historical compatibility and observed local motion, and use latent-space coordination to handle their interaction.

3. **Systematic evaluation.** Experiments on PushT and PointMaze show that short histories support both visual and dynamics pre-outcome adaptation, and that the unified method retains most of the gains from both adaptation specialists without new future feedback.

2 Related Work

2.1 Test-Time Adaptation

Test-time adaptation (TTA) updates pretrained models using unlabeled deployment data. Early approaches adapt normalization statistics or feature extractors under distribution shift (Schneider et al., 2020; Liang et al., 2020); Tent and MEMO instead construct test-time objectives from entropy minimization and augmentation consistency (Wang et al., 2021; Zhang et al., 2022). Continual variants further address instability, forgetting, and temporally correlated shifts through teacher updates, sample filtering, sharpness-aware optimization, memory, or restoration mechanisms (Wang et al., 2022a; Niu et al., 2022; 2023; Yuan et al., 2023; Döbler et al., 2023). Test-Time Training (TTT) follows a related but different protocol: source training prepares a self-supervised objective that remains optimizable at deployment, including feature alignment, masked reconstruction, and short-window video adaptation (Sun et al., 2020; Liu et al., 2021; Gandelsman et al., 2022; Wang et al., 2025b; Lin et al., 2023).

These methods establish that unlabeled inputs and short temporal context can provide useful adaptation signals. Their objectives, however, are mainly designed for recognition, reconstruction, or prediction consistency. In an action-conditioned world model, temporal variation may instead reflect a real change in the transition dynamics. HATTA therefore does not enforce generic temporal consistency; it uses the pretrained predictor and already observed transitions to construct separate representation and dynamics correction signals, without source-domain statistics or a separately trained test-time recognition head.

2.2 Test-Time Adaptation of World Models

The most closely related work adapts the world model itself during deployment. AdaJEPA updates a latent world model from newly observed transitions inside closed-loop MPC, while Sandwich-Residuals restricts the same online prediction-error adaptation to lightweight residual modules (Wang et al., 2026a; Soni et al., 2026). AdaWM instead distinguishes policy and dynamics-model mismatch and selects which component to fine-tune, whereas ReOI intervenes on visual observations to suppress distractors before planning (Wang et al., 2025a; Chen et al., 2026). HATTA differs in protocol and objective: it completes the current update using only the already available history, without waiting for an additional outcome, explicitly classifying the shift source, or replacing the deployment observation.

Using recent interaction history for adaptation also has earlier precedents. Meta-learned dynamics priors, RMA, and PAD extract adaptation signals from recent transitions or self-supervised deployment objectives, but rely on meta-trained priors, dedicated adaptation modules, or auxiliary objectives prepared during training (Nagabandi et al., 2019; Kumar et al., 2021; Hansen et al., 2021). Recent world-model methods likewise prepare dedicated test-time mechanisms: AdaPower adds temporal–spatial TTT modules, while WAM-TTT meta-trains adaptive memory from paired human–robot data (Huang et al., 2025; Feng et al., 2026). HATTA instead assumes an already trained JEPA-style latent world model and uses only the currently available target-domain observation–action history to adapt both representation and dynamics.

3 Method

3.1 Pre-Outcome Adaptation

We follow the latent world-model and MPC framework of AdaJEPA (Wang et al., 2026a). Given an observation o_t, the encoder E_ϕ produces a latent state $z_t = E_\phi(o_t)$, and the predictor P_θ predicts the next latent state from a recent latent-state/action history. We use the causal one-step prediction interface of the AdaJEPA-style backbone: at valid historical positions, the predictor produces one-step-ahead outputs from the observed context, and HATTA uses only positions whose successor

observation is already contained in the current history. This use is consistent with causal predictor formulations in AdaJEPA-based world models (Wang et al., 2026a; Soni et al., 2026). For each such valid position,

$$\hat{z}_{i+1} = P_\theta\left(z_{t-K+1:i}, a_{t-K+1:i}\right). \tag{1}$$

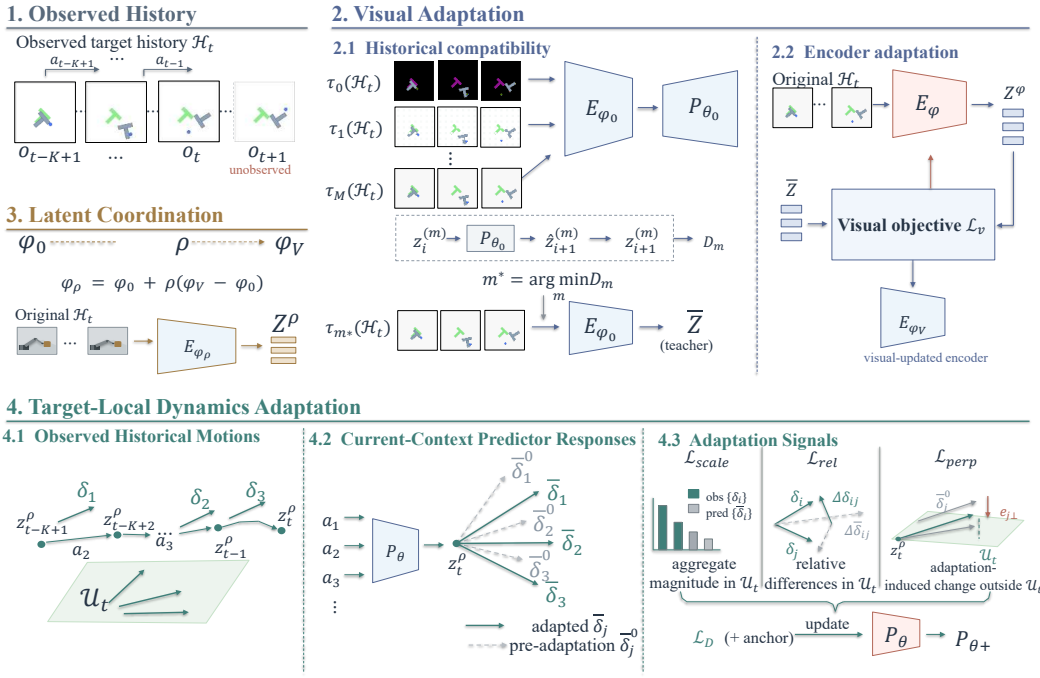


Figure 2: **Overview of HATTA.** HATTA scores counterfactual visual hypotheses by historical compatibility and uses the selected hypothesis for prediction-sensitive encoder adaptation. A fixed strength ρ controls the accepted representation displacement before the original history is re-encoded. Observed motions then define a target-local subspace U_t, while historical actions serve only as non-executed probes at the current context; scale, relative-response, and orthogonal-preservation signals adapt the predictor before MPC/CEM resumes.

At test time, the model only observes the target-domain history $\mathcal{H}_t = (o_{t-K+1:t}, a_{t-K+1:t-1})$. It cannot access o_{t+1} or any later real observation, and it does not execute extra probe actions to construct an adaptation target. Although the current future outcome is unavailable, $\mathcal{H}_t$ already contains $K-1$ real transitions (o_i, a_i, o_{i+1}).

This information boundary directly motivates HATTA. First, because no new future target is available, test-time supervision must come from transitions that have already occurred inside the observed history. Second, temporal variation cannot be treated uniformly as nuisance: appearance variation should mainly correct the representation, whereas real motion should be preserved as target-domain dynamics evidence. Third, the encoder and predictor share the same latent space, so representation adaptation changes the coordinate system used by subsequent dynamics calibration. HATTA therefore performs visual correction, latent-space coordination, and target-local dynamics adaptation in sequence. Figure 2 summarizes the complete adaptation pipeline.

3.2 Visual Adaptation from Historical Compatibility

Historical compatibility. At test time, no clean source view is paired with the target observation. We therefore do not attempt to directly invert the unknown appearance shift. Instead, we construct a fixed set of counterfactual visual hypotheses $\mathcal{T} = \{\tau_0, \ldots, \tau_M\}$, where τ_0 is the identity and each τ_m is applied consistently across the full visual history to preserve temporal correspondence.

For candidate m, the frozen encoder produces $z_i^{(m)} = E_{\phi_0}(\tau_m(o_i))$. Because the corresponding successors have already been observed, the frozen predictor can re-predict them from each valid historical prefix. We define

$$\begin{aligned} \hat{z}_{i+1}^{(m)} &= P_{\theta_0}\left(z_{t-K+1:i}^{(m)}, a_{t-K+1:i}\right), \\ D_m &= \frac{1}{K-1}\sum_{i=t-K+1}^{t-1} \ell_{\text{hist}}\left(\hat{z}_{i+1}^{(m)}, \text{sg}[z_{i+1}^{(m)}]\right), \\ m^\star &= \arg\min_m D_m. \end{aligned} \tag{2}$$

Thus, the supervision in Eq. (2) comes entirely from historical transitions that have already occurred, rather than from the unseen o_{t+1}. We interpret $\tau_{m^\star}$ only as the visual hypothesis most compatible with the frozen source dynamics, not as the true inverse of the unknown corruption.

Prediction-sensitive representation correction. The selected hypothesis defines a frozen pseudo-canonical teacher $\bar{z}_i = \text{sg}[E_{\phi_0}(\tau_{m^\star}(o_i))]$, while the student encoder always processes the original target observation, $z_i^\phi = E_\phi(o_i)$. The counterfactual view therefore provides only a correction direction and never replaces the actual deployment input.

A uniform feature alignment would treat all latent dimensions equally, although only part of the representation may be important for action-conditioned prediction. We therefore use the frozen predictor to estimate dimension-wise prediction sensitivity. For historical probe action a_j, let $\bar{r}_j = P_{\theta_0}(\bar{Z}; a_j) - \bar{z}_t$ denote the teacher response. The response-separation energy E_{act} induces the sensitivity weight

$$w_{i,d} \propto \left|\frac{\partial E_{\text{act}}}{\partial \bar{z}_{i,d}}\right|. \tag{3}$$

A larger $w_{i,d}$ indicates that the corresponding teacher dimension has a stronger influence on distinguishing action-conditioned responses. The gradient in Eq. (3) is used only as a sensitivity estimator and is detached before optimization. The exact construction is given in Appendix A.3.

The visual objective combines prediction-sensitive feature correction, action-response consistency, and a parameter anchor:

$$L_V = L_{\text{feat}} + L_{\text{resp}} + \lambda_V L_{\text{anchor}}^V. \tag{4}$$

Here, L_{feat} aligns the original target representation with prediction-relevant teacher dimensions, while L_{resp} preserves the teacher's action-conditioned behavior under the frozen predictor. The visual stage freezes the predictor and updates only the encoder parameter subset Θ_V, producing the visual-adapted encoder E_{ϕ_V}. Exact loss definitions and optimization details are given in Appendix A.3.

3.3 Latent-Space Coordination

Visual adaptation reduces observation-side mismatch, but it also moves the latent coordinates in which the predictor operates. Moreover, ϕ_V is obtained under frozen source dynamics. When the target transition law itself has changed, accepting the full visual correction can suppress target-motion evidence that should instead be preserved.

We therefore accept only a controlled fraction of the visual parameter displacement and re-encode the *original* history:

$$\phi_\rho = \phi_0 + \rho(\phi_V - \phi_0), \qquad Z^\rho = \left(E_{\phi_\rho}(o_{t-K+1}), \ldots, E_{\phi_\rho}(o_t)\right). \tag{5}$$

Here, $\rho \in [0, 1]$ is a fixed coordination strength rather than a shift classifier or sample-dependent router. From this point onward, the counterfactual teacher is no longer used. All observed motions and predictor responses are reconstructed in Z^ρ, so representation correction and dynamics adaptation share the same final latent coordinate system.

3.4 TARGET-LOCAL DYNAMICS ADAPTATION

Generic temporal consistency tends to suppress historical change. For a world model, however, change already observed in the target history is direct evidence about the target dynamics. HATTA therefore uses the latent motion $\delta_i = z^\rho_{i+1} - z^\rho_i$ to calibrate the predictor. These observed motions and their relative differences define a target-local motion subspace U_t that contains only the latent directions directly supported by the current short history. Its exact construction is given in Appendix A.4.

Local action-response calibration. The historical pairs (a_i, δ_i) come from nearby but different states, whereas predictor adaptation is applied at the current context. We assume local smoothness: within the short state region covered by the history, the main scale and relative structure of action-conditioned responses vary only moderately. Historical actions can therefore be reused as *non-executed probes* in the current context:

$$\hat{\delta}_j(\theta) = P_\theta(Z^\rho; a_j) - z^\rho_t, \qquad \hat{\delta}^0_j = P_{\theta_0}(Z^\rho; a_j) - z^\rho_t. \tag{6}$$

These actions only query the model and are not executed again in the environment. All predicted responses are queried from the same current context Z^ρ, whereas the observed motions δ_i come from different historical transitions.

Dynamics adaptation follows a simple principle: *adapt where the target history provides evidence, and preserve pretrained behavior elsewhere*. Specifically, L_{scale} matches the aggregate motion scale of observed and predicted responses inside U_t; L_{rel} matches relative action-response differences; and $L_\perp$ suppresses adaptation-induced response changes outside the target-supported subspace. The final objective is

$$L_D = L_{\text{scale}} + L_{\text{rel}} + L_\perp + \lambda_D L^D_{\text{anchor}}. \tag{7}$$

The dynamics stage freezes E_{ϕ_ρ} and updates only the predictor parameter subset Θ_D. Thus, the first two terms use target-domain motion evidence already present in the history, while $L_\perp$ and the parameter anchor preserve pretrained behavior in directions not supported by that history. Exact objectives and projection details are given in Appendix A.5.

After adaptation, both the encoder and predictor are frozen and used with the original MPC/CEM planner.

4 EXPERIMENTS

We evaluate HATTA around three questions: whether a short observed history can provide useful adaptation signals before the next future outcome becomes available; whether the visual and dynamics branches exploit the intended target-domain evidence; and whether latent coordination can reduce interference between the two updates.

4.1 EXPERIMENTAL SETUP

We evaluate two local benchmark suites derived from the AdaJEPA setup (Wang et al., 2026a). For visual shifts, we keep the original T-shaped PushT task and apply seven appearance perturbations: Blur, SNP1, SNP5, Dark, Red Block, Red Agent, and Red Anchor. We therefore refer to this suite as *PushT Visual OOD*; it should not be confused with AdaJEPA's PushObj shape-variation benchmark. For PointMaze, we use the reference dynamics together with Low Density and High Damping. *Low Density* is the name used by our environment configuration (`density_scale`=0.2); because the maze geometry is fixed, this realizes the same $0.2\times$ mass scaling referred to as the low-mass condition in AdaJEPA. High Damping uses `damping_scale`=20. Each condition is evaluated with three random seeds and 50 episodes per seed. All methods use the same pretrained checkpoint, episode initialization, and MPC/CEM protocol, and are compared on matched episode keys.

We compare the frozen model (Frozen), visual adaptation only (Visual-only), dynamics adaptation only (Dynamics-only), and the full HATTA pipeline. As a feedback-based reference, we report *V0-X1*, a local AdaJEPA-derived comparison protocol rather than a named variant from Wang et al. (Wang et al., 2026a). V0-X1 uses the same local V0 checkpoint and MPC/CEM implementation, performs one AdaJEPA feedback update after each of the first three feedback events, and then

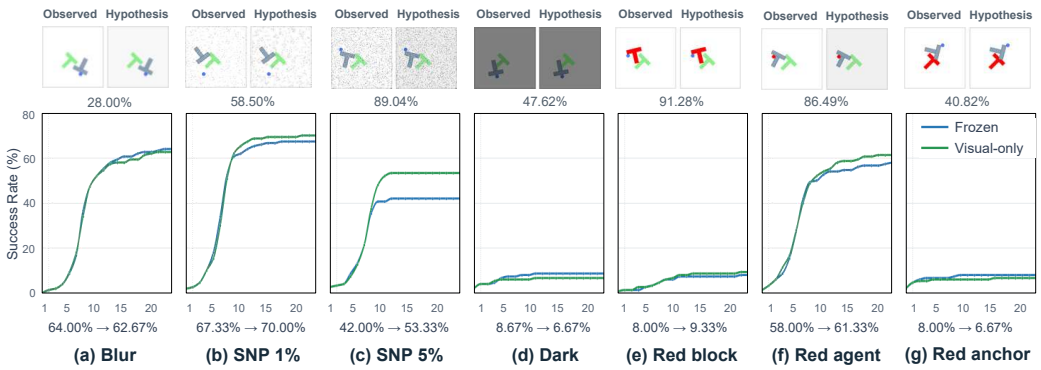


Figure 3: Visual-only adaptation under PushT appearance shifts. For each shift, the top row shows an observed frame and the most frequently selected visual hypothesis; the percentage reports how often that hypothesis is selected among legal adaptation cases. The bottom row shows cumulative planning success versus MPC step for Frozen and Visual-only, with the final success rates reported below each panel. The selected hypothesis is determined by historical compatibility and is not assumed to be the true clean view or the inverse of the test corruption.

freezes the adapted model for the remainder of the episode; Appendix A.1 gives the implementation correspondence. Our primary metric is planning success rate. For PushT, we additionally report final distance. For PointMaze, we report both final distance and true trajectory cost.

4.2 Main Results

Table 1 summarizes the main results.

Table 1: Main comparison. HATTA adapts without using new future feedback.

Method	New future feedback	Overall ↑	PushT Visual OOD ↑	Maze Dynamics OOD ↑
Frozen	No	49.27%	36.57%	70.67%
Visual-only	No	50.30%	**38.57%**	70.00%
Dynamics-only	No	49.45%	36.00%	**73.00%**
Local AdaJEPA ref. (V0-X1)	Yes	49.94%	37.43%	72.67%
HATTA ($\rho = 0.4$)	**No**	**50.48%**	38.00%	72.00%

HATTA improves overall success from 49.27% to 50.48%, while reaching 38.00% on PushT Visual OOD and 72.00% on Maze Dynamics OOD. The two isolated branches show the expected factor preference: Visual-only reaches 38.57% under visual shifts, while Dynamics-only reaches 73.00% under dynamics shifts. The full HATTA pipeline retains most of these gains without using shift labels.

The local V0-X1 reference reaches 49.94% overall success, 37.43% on PushT Visual OOD, and 72.67% on Maze Dynamics OOD after receiving additional real transition feedback. HATTA achieves similar point estimates under a different information budget: it completes adaptation before the next future observation becomes available, whereas V0-X1 uses later real transitions. We therefore use this comparison to assess how much useful adaptation signal is already present before new feedback arrives, rather than as a direct ranking between the two protocols.

4.3 Factor-Specific Analysis

Visual adaptation. We first isolate the visual branch. The transformation operators used by the visual hypothesis bank do not directly overlap with the formal test corruptions: the benchmark includes Blur, salt-and-pepper noise, Dark, and color shifts, whereas the candidate bank uses a separate set of counterfactual transformations. The Visual-only results therefore cannot be explained simply as searching for the inverse of a known test corruption.

Table 2: Target-local dynamics adaptation on PointMaze. Δ is Dynamics-only minus Frozen. Positive Δ favors success; negative Δ favors distance and cost.

	Success ↑			Final Distance ↓			True Cost ↓		
Condition	Frozen	Dynamics	Δ	Frozen	Dynamics	Δ	Frozen	Dynamics	Δ
Default	77.3	**78.0**	**+0.7**	5.210	**5.062**	**-0.148**	1.057	**1.032**	**-0.025**
Low Density (0.2×)	77.3	**80.0**	**+2.7**	**6.872**	7.345	+0.473	1.087	**1.066**	**-0.021**
High Damping	64.0	**66.0**	**+2.0**	4.649	**4.426**	**-0.223**	1.582	**1.342**	**-0.240**
Dynamics OOD Macro	70.7	**73.0**	**+2.3**	**5.761**	5.885	+0.124	1.334	**1.204**	**-0.130**

Figure 3 shows the selected visual hypotheses and the corresponding closed-loop success curves. Visual-only improves All Visual OOD success from 36.6% to 38.6%, while reducing aggregate final distance from 120.933 to 119.464. The largest gain appears on SNP5, with additional improvements on SNP1 and Red Agent; Blur, Dark, and Red Anchor show small degradations.

This non-monotonic behavior is consistent with the role of historical compatibility. A smaller D_m indicates that a counterfactual representation is more compatible with the frozen source dynamics, but it does not guarantee that the same representation is optimal for the final MPC objective. The visual branch therefore provides a test-time representation correction signal from the pretrained world model, rather than recovering a predefined clean image.

Dynamics adaptation. We next isolate the target-local dynamics branch.

Dynamics-only improves Dynamics OOD macro success from 70.7% to 73.0% and reduces true trajectory cost from 1.334 to 1.204. Under High Damping, success, final distance, and true cost all improve. Under Low Density (`density_scale` $= 0.2$, corresponding to the low-mass setting), success and true cost improve while final distance increases slightly.

These results show that latent motion already present in the observed history can provide an effective target-local adaptation signal for the predictor before the next future state becomes available. The trajectory-level metrics need not move in the same direction because HATTA calibrates local action responses rather than directly minimizing the episode-level final-position error.

4.4 Effect of Latent Coordination

Visual-only and Dynamics-only perform best within their corresponding shift families, but applying the full visual update before dynamics adaptation does not simply combine their gains. This behavior is consistent with the latent-space coupling described in Sec. 3.3: visual adaptation changes the representation coordinates used by the predictor. We therefore study the coordination strength ρ.

When $\rho = 0$, the visual parameter displacement is discarded and the method reduces to Dynamics-only. When $\rho = 1$, the full visual correction is retained before dynamics adaptation.

Figure 4 makes this trade-off explicit. As ρ increases, PushT Visual OOD performance generally improves and reaches 38.9% at $\rho = 0.6$, while Maze Dynamics OOD performance decreases from 73.0%. This continuous trend exposes the cross-factor trade-off between representation correction and dynamics calibration.

We use $\rho = 0.4$ in the main results. This setting reaches 38.0% on PushT Visual OOD and 72.0% on Maze Dynamics OOD, with 50.5% overall success. The corresponding bootstrap intervals include zero. We therefore use the sweep to characterize the trade-off captured by latent coordination and to select an operating point for the unified method, rather than to claim that any individual ρ is statistically superior to Frozen.

5 Conclusion

We introduced *pre-outcome test-time adaptation* for latent world models: adapting from target-domain history that is already available before an additional real successor is observed. HATTA turns this history into two complementary signals. Historical compatibility provides a dynamics-aware

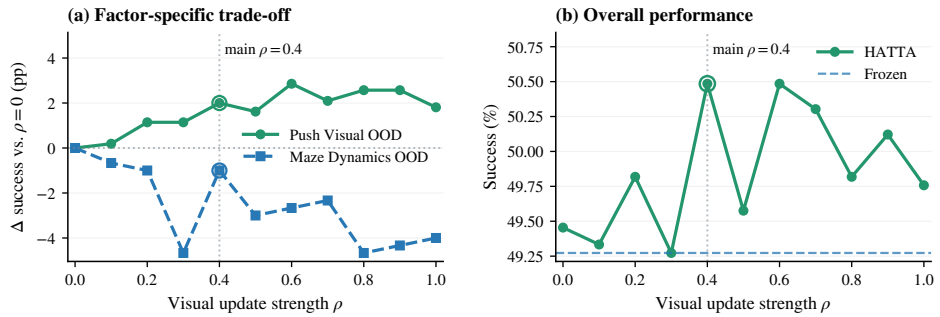


Figure 4: Effect of latent coordination strength ρ. (a) Change in success relative to $\rho = 0$, the Dynamics-only endpoint, for PushT Visual OOD and Maze Dynamics OOD. Increasing ρ strengthens the accepted visual correction but gradually reduces dynamics-shift performance, exposing the cross-factor trade-off. (b) Overall success across the same measured sweep, with Frozen shown as a reference. The main experiments use $\rho = 0.4$. Lines connect evaluated sweep points only and do not imply a statistically optimal value of ρ.

direction for visual representation correction, while observed local motion calibrates the action-conditioned predictor. Because the two updates share a latent coordinate system, HATTA explicitly controls the accepted representation displacement before dynamics adaptation.

Experiments on PushT and PointMaze show that short histories can support useful visual and dynamics adaptation before new interaction feedback arrives, and that latent coordination exposes and manages the trade-off between the two corrections. These results position pre-outcome adaptation as complementary to, rather than a replacement for, later feedback-based updates. A natural next step is to combine both stages in a single deployment protocol and to study richer histories and larger world models without introducing shift labels or additional probe interactions.

AI Usage Disclosure

In this work, we used generative AI tools to assist with research ideation and execution, drafting parts of the paper, and language editing and polishing. We reviewed all AI-assisted work: methodological suggestions and generated materials were checked and validated by the authors, and all drafted and edited text was read, verified, and revised by the authors. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

References

Yuxin Chen, Jianglan Wei, Chenfeng Xu, Boyi Li, Masayoshi Tomizuka, Andrea Bajcsy, and Ran Tian. Reimagination with Test-time Observation Interventions: Distractor-Robust World Model Predictions for Visual Model Predictive Control. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2026. URL https://arxiv.org/abs/2506.16565.

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*, 2023. URL https://arxiv.org/abs/2211.13081.

Yusen Feng, Bingchen Han, Jiangran Lyu, Kai Liu, Yixin Zheng, Yuxuan Wan, Weiheng Liu, Sun Han, Ruiqin Li, Yulong Zhang, Fangfu Liu, Xuesong Shi, Libin Liu, Yizhou Wang, Zhizheng Zhang, and He Wang. WAM-TTT: Steering World-Action Models by Watching Human Play at Test Time. arXiv:2607.06988, 2026.

Yossi Gandelsman, Yu Sun, Xinlei Chen, and Alexei A. Efros. Test-Time Training with Masked Autoencoders. In *Advances in Neural Information Processing Systems*, 2022. URL https://arxiv.org/abs/2209.07522.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning Latent Dynamics for Planning from Pixels. In *International Conference on Machine Learning*, pp. 2555–2565, 2019.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to Control: Learning Behaviors by Latent Imagination. In *International Conference on Learning Representations*, 2020.

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*, 2021.

Yuhang Huang, Shilong Zou, Jiazhao Zhang, Xinwang Liu, Ruizhen Hu, and Kai Xu. AdaPower: Specializing World Foundation Models for Predictive Manipulation. arXiv:2512.03538, 2025.

Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. RMA: Rapid Motor Adaptation for Legged Robots. In *Robotics: Science and Systems*, 2021.

Jian Liang, Dapeng Hu, and Jiashi Feng. Do We Really Need to Access the Source Data? Source Hypothesis Transfer for Unsupervised Domain Adaptation. In *International Conference on Machine Learning*, pp. 6028–6039, 2020.

Wei Lin, Muhammad Jehanzeb Mirza, Mateusz Kozinski, Horst Possegger, Hilde Kuehne, and Horst Bischof. Video Test-Time Adaptation for Action Recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

Yuejiang Liu, Parth Kothari, Bastien van Delft, Baptiste Bellot-Gurlet, Taylor Mordan, and Alexandre Alahi. TTT++: When Does Self-Supervised Test-Time Training Fail or Thrive? In *Advances in Neural Information Processing Systems*, 34:21808–21820, 2021.

Lucas Maes, Quentin Le Lidec, Dan Haramati, Nassim Massaudi, Damien Scieur, Yann LeCun, and Randall Balestriero. stable-worldmodel-v1: Reproducible World Modeling Research and Evaluation. arXiv:2602.08968, 2026.

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.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofo Chen, Shijian Zheng, Peilin Zhao, and Mingkui Tan. Efficient Test-Time Model Adaptation without Forgetting. In *International Conference on Machine Learning*, pp. 16888–16905, 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.

Steffen Schneider, Evgenia Rusak, Luisa Eck, Oliver Bringmann, Wieland Brendel, and Matthias Bethge. Improving Robustness against Common Corruptions by Covariate Shift Adaptation. In *Advances in Neural Information Processing Systems*, 2020.

Krishnam Soni, Aditya Sehgal, Vedant Dave, and Elmar Rueckert. Sandwich-Residuals: Parameter-Efficient Test-time Adaptation of World Models. 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 *International Conference on Machine Learning*, pp. 9229–9248, 2020.

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.

Hang Wang, Xin Ye, Feng Tao, Chenbin Pan, Abhirup Mallik, Burhaneddin Yaman, Liu Ren, and Junshan Zhang. AdaWM: Adaptive World Model based Planning for Autonomous Driving. In *International Conference on Learning Representations*, 2025.

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.

Renhao Wang, Yu Sun, Arnuv Tandon, Yossi Gandelsman, Xinlei Chen, Alexei A. Efros, and Xiaolong Wang. Test-Time Training on Video Streams. *Journal of Machine Learning Research*, 26(9):1–29, 2025.

Tongzhou Wang, Simon S. Du, Antonio Torralba, Phillip Isola, Amy Zhang, and Yuandong Tian. Denoised MDPs: Learning World Models Better Than the World Itself. In *International Conference on Machine Learning*, pp. 22591–22612, 2022.

Ying Wang, Oumayma Bounou, Yann LeCun, and Mengye Ren. AdaJEPA: An Adaptive Latent World Model. arXiv:2606.32026, 2026.

Ying Wang, Oumayma Bounou, Gaoyue Zhou, Randall Balestriero, Tim G. J. Rudner, Yann LeCun, and Mengye Ren. Temporal Straightening for Latent Planning. arXiv:2603.12231, 2026.

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.

Marvin Zhang, Sergey Levine, and Chelsea Finn. MEMO: Test Time Robustness via Adaptation and Augmentation. In *Advances in Neural Information Processing Systems*, 2022.

Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. DINO-WM: World Models on Pre-trained Visual Features Enable Zero-shot Planning. In *International Conference on Machine Learning*, pp. 79115–79135, 2025.

A Adaptation Objectives and Implementation Details

A.1 Backbone Interface, Benchmark Correspondence, and V0-X1

Historical prefix predictions. The main text uses a model-agnostic prefix notation, but the implementation evaluates only historical one-step predictions whose successor observations are already present in the legal pre-outcome context. In our experiments, $K = 3$, so historical compatibility is computed from the two already observed transitions inside the three-frame context. The scoring path uses the same pretrained causal predictor/checkpoint as the other evaluation branches and does not expose o_{t+1} or any later observation when constructing D_m. HATTA therefore requires a prefix-prediction interface rather than a particular predictor architecture.

Benchmark naming. Our visual-shift suite uses the original T-shaped PushT task and applies Blur, salt-and-pepper noise, Dark, and red/color appearance perturbations. We call this suite *PushT Visual OOD*; it is distinct from the PushObj shape-variation benchmark in AdaJEPA. In PointMaze, the local condition name *Low Density* corresponds to `density_scale`=0.2. Because the geometry is fixed, this scales the simulated mass by the same factor and realizes the $0.2\times$ low-mass perturbation described by AdaJEPA. The High Damping condition uses `damping_scale`=20.

Local feedback reference. V0-X1 is a local budgeted AdaJEPA reference, not a named variant from the AdaJEPA paper. It uses the same local V0 checkpoint and the current AdaJEPA MPC/CEM feedback implementation. The X1 wrapper changes the update budget: it permits one feedback update after each of the first three feedback events and then disables further model updates while planning continues. With the raw-frame indexing used by our archive, these three updates occur after frames 6, 11, and 16 become available. The third update still uses only already observed adjacent transitions in the recent-history loss. Thus, V0-X1 is included only as a feedback-based information-budget reference to HATTA's pre-outcome protocol.

A.2 Historical Compatibility and Visual Hypotheses

The main text describes historical compatibility in a model-agnostic form. In our implementation, the legal pre-outcome context has $K = 3$. The fixed candidate bank contains 16 visual hypotheses: identity; JPEG with qualities $85/65/45/25$; Posterize with $7/6/5/4$ bits; Solarize with thresholds $0.90/0.75/0.60/0.45$; and deterministic GridMask with ratios $0.05/0.10/0.15$. Each transform is applied consistently to the full three-frame RGB history before the fixed checkpoint normalization.

Candidate scoring is performed entirely with the frozen model. The causal predictor evaluates the two valid historical transitions, and the action suffix is excluded from the prediction discrepancy. Let d_s denote the state dimension used by the discrepancy. The normalized historical score is

$$D_m = \mathrm{mean}_i \frac{\sum_{d=1}^{d_s} \left(\hat{z}_{i+1,d}^{(m)} - \mathrm{sg}[z_{i+1,d}^{(m)}] \right)^2}{d_s^{-1} \sum_{d=1}^{d_s} \mathrm{sg}[z_{i+1,d}^{(m)}]^2 + \epsilon}. \tag{8}$$

All candidate scores are computed without gradients, including the identity candidate, and the final selection is

$$m^\star = \arg\min_m D_m. \tag{9}$$

The final visual branch uses this D-only selection rule; no additional candidate action-discrimination score or shift classifier is used.

A.3 Prediction-Sensitive Visual Objective

The selected candidate provides the frozen pseudo-canonical teacher. The student always processes the original target observations. Historical actions are reused only as internal predictor probes. Let

$$\bar{r}_j = \mathrm{sg}\left[P_{\theta_0}(\bar{Z}; a_j) - \bar{z}_t \right]$$

denote the frozen teacher response. We define the action-response separation

$$E_{\mathrm{act}} = \frac{1}{|\mathcal{P}|} \sum_{j<k} \|\bar{r}_j - \bar{r}_k\|_2^2, \tag{10}$$

and compute dimension-wise sensitivity

$$g_{i,d} = \left| \frac{\partial E_{\mathrm{act}}}{\partial \bar{z}_{i,d}} \right|, \qquad w_{i,d} = \frac{g_{i,d}}{\mathrm{mean}_{d'}\, g_{i,d'} + \epsilon}. \tag{11}$$

Only a detached copy of the teacher representation is made differentiable for this calculation. The resulting w is detached before the adaptation loss is optimized, so no model parameter receives the sensitivity gradient.

The weighted feature objective is

$$L_{\mathrm{feat}} = \mathrm{mean}_i \frac{\sum_d w_{i,d} (z_{i,d}^{\phi} - \bar{z}_{i,d})^2}{\sum_d w_{i,d} + \epsilon}. \tag{12}$$

For the same historical probe actions, the student response is $r_j^{\phi} = P_{\theta_0}(Z^{\phi}; a_j) - z_t^{\phi}$, and the response-consistency loss is

$$L_{\mathrm{resp}} = \frac{1}{|\mathcal{A}_H|} \sum_j \frac{\|r_j^{\phi} - \bar{r}_j\|_2^2}{\|\bar{r}_j\|_2^2 + \epsilon}. \tag{13}$$

We use the parameter anchor

$$L_{\mathrm{anchor}}^V = \frac{1}{|\Theta_V|} \sum_{p \in \Theta_V} \|p - p_0\|_2^2, \tag{14}$$

and the complete visual objective

$$L_V = L_{\mathrm{feat}} + L_{\mathrm{resp}} + 10^{-3} L_{\mathrm{anchor}}^V. \tag{15}$$

The visual stage updates only the terminal encoder projection/output-normalization parameters; the predictor is fully frozen. We use Adam with learning rate 10^{-4}, weight decay 0, four adaptation steps, and gradient-norm clipping at 1.0. The model remains in evaluation mode and frozen buffers are not updated.

A.4 Target-Local Motion Subspace and Probe Construction

After latent-space coordination, the original three-frame history is re-encoded by E_{ϕ_ρ}. With $K = 3$, the observed motions are

$$\delta_1 = z_2^\rho - z_1^\rho, \qquad \delta_2 = z_3^\rho - z_2^\rho. \tag{16}$$

We form

$$R = [\delta_1, \delta_2, \delta_1 - \delta_2] \tag{17}$$

and compute the thin SVD

$$R = U\Sigma V^\top. \tag{18}$$

We retain singular directions satisfying $\sigma_i \geq 10^{-6}\sigma_1$, subject to the maximum rank supported by the short history. If $\sigma_1 \leq 10^{-12}$, the local rank is set to zero. In the zero-rank case, the evidence-dependent dynamics terms are defined as differentiable zeros, avoiding unsupported predictor updates.

Historical actions are reused only as non-executed probes. For $K = 3$, the two probes reuse a_1 and a_2, but both query the predictor at the same latest context Z^ρ. We define

$$\hat{\delta}_j(\theta) = P_\theta(Z^\rho; a_j) - z_3^\rho, \qquad \hat{\delta}_j^0 = P_{\theta_0}(Z^\rho; a_j) - z_3^\rho. \tag{19}$$

Observed motions, the local subspace, probe inputs, the latest state, and the step-zero responses are detached. Only predictor parameters are updated in this stage.

A.5 Exact Dynamics Objective

Inside the valid local subspace U, we define the observed and predicted motion scales as

$$s_{\text{obs}} = \sqrt{\text{mean}_j \, \|U^\top \delta_j\|_2^2 + \epsilon}, \qquad s_{\text{pred}} = \sqrt{\text{mean}_j \, \|U^\top \hat{\delta}_j\|_2^2 + \epsilon}. \tag{20}$$

The motion-scale consistency loss is

$$L_{\text{scale}} = \left[\log(s_{\text{pred}} + \epsilon) - \log(\text{sg}[s_{\text{obs}}] + \epsilon)\right]^2. \tag{21}$$

For an action pair (a_i, a_j), define $\Delta\delta_{ij} = \delta_i - \delta_j$ and $\Delta\hat{\delta}_{ij} = \hat{\delta}_i - \hat{\delta}_j$, with action-distance weight

$$\omega_{ij} = \frac{\|a_i - a_j\|_2^2}{\|a_i - a_j\|_2^2 + 10^{-6}}. \tag{22}$$

The relative-response loss is

$$L_{\text{rel}} = \text{mean}_{i<j} \, \omega_{ij} \frac{\left\| U^\top \left[\Delta\hat{\delta}_{ij} - \text{sg}(\Delta\delta_{ij})\right] \right\|_2^2}{\|U^\top \Delta\delta_{ij}\|_2^2 + \epsilon}. \tag{23}$$

To preserve pretrained behavior in directions unsupported by the short history, define the adaptation-induced response change

$$e_j = \hat{\delta}_j - \text{sg}[\hat{\delta}_j^0].$$

The orthogonal-preservation loss is

$$L_\perp = \text{mean}_j \left\| (I - UU^\top) e_j \right\|_2^2. \tag{24}$$

With the parameter anchor

$$L_{\text{anchor}}^D = \frac{1}{|\Theta_D|} \sum_{p \in \Theta_D} \|p - p_0\|_2^2, \tag{25}$$

the complete dynamics objective is

$$L_D = L_{\text{scale}} + L_{\text{rel}} + L_\perp + 10^{-3} L_{\text{anchor}}^D. \tag{26}$$

The dynamics stage updates all affine LayerNorm weights and biases in the predictor, corresponding to 26 tensors and 10,504 scalar parameters in the current implementation. The encoder is frozen. We use Adam with learning rate 10^{-3}, weight decay 0, four adaptation steps, and gradient-norm clipping at 1.0. The model remains in evaluation mode.

A.6 Deployment Protocol

For each legal target history, HATTA performs the following steps:

1. Score the fixed counterfactual visual candidates with the frozen world model and select $m^\star$ by historical compatibility.
2. Construct the selected frozen teacher and update the visual encoder.
3. Interpolate the visual parameter displacement using the fixed coordination strength ρ.
4. Re-encode the same original target history with E_{ϕ_ρ}.
5. Construct target-local motion evidence in the final latent space and update the predictor.
6. Freeze E_{ϕ_ρ} and P_{θ^+} and continue with the original MPC/CEM planner.

The procedure does not acquire an additional real successor for HATTA, does not use source-domain samples, rewards, goals, or shift labels, and does not repeatedly perform the same pre-outcome adaptation as later feedback arrives.