Imagined Rollouts Are Kinematic, Not Dynamic: A Diagnosis of Long-Horizon World-Model Failure

Finn Rasmus Schäfer[1], Korbinian Moller[1], Yuan Gao[1], Christian Oefinger[1],
Sebastian Schmidt[2] and Johannes Betz[1]
[1]Autonomous Vehicle Systems Lab
Technical University of Munich, Garching b. München, Germany
Email: finn.schaefer@tum.de
[2]Data Analytics and Machine Learning Group
Technical University of Munich, Garching b. München, Germany

***Abstract*—Long-horizon failure in world models is conventionally attributed to compounding error, a generic framing that does not distinguish what kind of error compounds. We propose a kinematic-vs-dynamic reframing: world models tend to imagine kinematically rather than dynamically. We operationalize this as the imagined Kinematic-Consistency Error, a per-step diagnostic that measures how far a rollout departs from a closed-form kinematic null, paired with a perturbation protocol that tests whether iKCE responds when physical conditions cross a regime boundary. We instantiate the diagnostic on a released DreamerV3 checkpoint trained on DMC walker-walk, where imagined iKCE runs roughly two orders of magnitude above that of matched real-physics rollouts. Across a friction sweep that crosses the gait-collapse boundary, the model's iKCE stays statistically flat even as the trained policy's reward collapses through the same range, providing the kinematic-not-dynamic signature. The diagnostic distinguishes kinematic from dynamic imagination at horizons longer than the embodiment's gait period.**

I. Introduction

World models have become a load-bearing component of recent embodied AI, serving as latent simulators for planning [3, 4, 5] and as generative environments for self-supervised learning [6, 7]. A widely-noted failure mode is the deterioration of imagined rollouts over long horizons, conventionally attributed to *compounding error* [8]. This framing is accurate but underspecified: it does not distinguish what kind of error compounds or which feature dimensions deteriorate. Across four recent Observations in driving VLM/VLA and trajectory-prediction benchmarks indicate that the deterioration carries a specific structural signature that the compounding-error framing obscures.

We adopt the classical mechanics distinction between kinematic and dynamic motion. We define *kinematic* as motion described purely through position, velocity, and acceleration time series, without invoking the forces or physical constraints that produced it. As *dynamic*, we define motion that requires those constraints (e.g., mass, friction, contact) to be reproduced correctly.

Central Claim

*Current world models imagine **kinematically** rather than **dynamically**:* they extrapolate position-velocity-acceleration trajectories that are internally consistent with linear kinematic update rules, but inconsistent with the physical constraints that produce real motion.

Kinematic fallback is a third, structurally distinct account of long-horizon world-model failure, alongside the two positions that dominate the literature: *predictable-representation engineering* (the Dreamer line [5]) and *error-compounding bounds* (MBPO [8]). Predictable-representation engineering attributes long-horizon reliability to stable, predictable latent representations and pursues it through normalization and balancing techniques; a single algorithm with fixed hyperparameters across 150+ tasks empirically validates the position, yet it is silent on *what* those representations should contain. Error-compounding bounds derive an explicit quadratic-in-horizon bound on the gap between model-based and true returns under policy distribution shift, and respond pragmatically by limiting model trust to short branched rollouts from real states. A world model whose latent contains rich kinematic features but no dynamic features satisfies both: stable enough for Dreamer-line predictability, accurate enough inside the training regime for Janner-line short-rollout bounds, yet still biased toward kinematic continuation once conditioning pushes its rollouts across a physical-regime boundary. The three accounts are therefore not mutually exclusive but different layers of the same failure surface; because they predict distinct empirical signatures, the protocol of Section III can target the kinematic-fallback layer specifically.

We make three contributions. We (i) **recast long-horizon world-model failure in kinematic-vs-dynamic terms**, distinguishing a structural-content layer from the variance-engineering and error-compounding layers studied in prior work; (ii) **introduce imagined kinematic-consistency error (iKCE) together with a conditioning-perturbation protocol** that operationalize this account as a falsifiable evaluation diagnostic; and (iii) **instantiate the diagnostic on an open-weight checkpoint** (DreamerV3 on DMC walker-walk), where

it exhibits both predicted signatures of kinematic imagination, with controls ruling out the principal confounds. The two signatures are a kinematic-null residual $\sim$180$\times$ above matched physics at T=16, and statistical invariance of the imagined rollouts to a friction sweep that crosses the empirical gait-collapse boundary. The diagnostic signature is this regime-invariance, not the absolute iKCE magnitude: a trivially kinematic predictor would produce zero iKCE.

II. Evidence

Our diagnosis is motivated by four existing observations, each inconclusive alone and explained only in isolation by its original authors, but jointly forming a coherent structural signature.

(i) Representational diagnostic on driving VLMs and VLAs. Schäfer et al. [11] present EgoDyn-Bench, a video-QA diagnostic that decouples physical reasoning from visual perception. (i), the weighted physics consistency rate (WPCR) saturates with a single static frame: rising from $\sim$20 with no visual input to $\sim$97 with one frame and remaining essentially flat as additional frames are added or temporally shuffled. (ii), reintroducing video to a text-only baseline recovers only $\sim$2.6pp on balanced accuracy under the best encoding, while text-only input already achieves 59.6% BAcc. The authors characterize this as a "functional decoupling between vision and language": ego-motion understanding is derived almost exclusively from the language modality, with visual observations **Implication for world models is structural.** *Contributing static context rather than temporal evidence. Imagined rollouts that depend on such encoders for motion-conditional features extrapolate from representations that under-encode the temporal dynamics they would need to imagine correctly.*

(ii) Sensor-degraded behavioral diagnostic. Priyadershi and Frtunikj [10] stress-test Alpamayo R1, a 10B-parameter driving VLA, across 1,996 scenarios under eight sensor perturbations. Under heavy Gaussian noise ($\sigma = 70$), the authors characterize the failure mode as one where the trajectory decoder "fails via collapsing kinematic priors while the language branch continues producing coherent but safety-irrelevant explanations." **Independent observation.** *This is an independent observation of the same kinematic-fallback failure mode, in a third-party VLA under naturalistic sensor degradation rather than controlled diagnostic stimuli.*

(iii) Open-loop trajectory-prediction baselines. Zhai et al. [13] train a 3-layer MLP that consumes only the ego vehicle's kinematic state and matches perception-based end-to-end planners on nuScenes open-loop L2 (0.29 m vs. 0.37 m for VAD-Base), with no camera, LiDAR, or HD-map input. The authors read this as a benchmark artifact, attributing it to the trajectory distribution of nuScenes and to a coarse collision-evaluation grid, and call for rethinking the open-loop evaluation scheme. We accept the empirical finding but read its significance differently. **Third independent signature.** *We read the same result as a third independent signature of kinematic imagination: when an ego-state-only predictor saturates the dominant open-loop metric, perception-based planners on this benchmark are not doing substantially more than kinematic extrapolation.*

(iv) Physics-consistency scoring on fine-tuned VLAs. Gao et al. [2] introduce a Kinematic Consistency Error (KCE) that scores a trajectory by checking each predicted next position against a closed-form kinematic extrapolation of the current state, and reuse it as a training loss for a fine-tuned 4B VLA. On their style-conditioned benchmark, KCE shows no monotonic relationship to model scale or modality: the strongest generalist (Gemini-3-Pro, 0.06–0.11 m) and the smaller fine-tuned models (0.08–0.12 m) overlap, with no ordering by parameter count or sensor richness. **A data/training deficit, not a capacity one.** *A deficit that is invariant to model scale and modality is unlikely to reflect a capacity limitation. It points instead to the training signal and data distribution rather than to model size.*

Conclusion. The four observations are concentrated in the driving and VLA setting but employ heterogeneous methodological registers, from representational probing to behavioral perturbation to physics-consistency scoring, and converge on a single structural deficit: the learned representation is dominated by kinematic features, and the dynamic features required for physical-regime-conditional behavior are systematically under-represented.

III. Diagnostic Protocol

A. iKCE: Imagined Kinematic Consistency Error

Definition (Imagined Kinematic-Consistency Error)**.** For an imagined rollout $\{\hat{x}_t^{\mathrm{WM}}\}_{t=0}^{T}$ produced by a world model, with $\hat{x}_t$ a chosen kinematic state vector (e.g. $[x, y, v, a, \theta]^\top$), the *imagined kinematic-consistency error* is

$$\mathrm{iKCE} \doteq \frac{1}{T}\sum_{t=0}^{T-1} \left\| \hat{x}_{t+1}^{\mathrm{WM}} - \mathrm{kin}\big(\hat{x}_t^{\mathrm{WM}}\big) \right\|^2, \tag{1}$$

where $\mathrm{kin}(\cdot)$ is any closed-form kinematic predictor (e.g., constant-velocity or constant-acceleration) chosen to match the WM's underlying embodiment and output space.

Equation 1 follows the mathematical form of the kinematic-consistency loss in Gao et al. [2], which supervises a fine-tuned VLA at training time. We repurpose it as a test-time diagnostic on imagined rollouts (the prefixed "i" denotes *imagined*), measuring kinematic inconsistency rather than reducing it. We repurpose it as an *evaluation diagnostic*: applied at test time to imagined rollouts, iKCE measures how far each predicted next state *departs* from the kinematic extrapolation of its predecessor.

Counter-intuitively, **low iKCE does not indicate dynamic imagination**: a world model with near-zero iKCE predicts, by construction, next states that coincide with the kinematic continuation of their predecessors: it imagines kinematically. The signature of dynamic imagination is the opposite: iKCE positive, growing with horizon, and responsive to physical-regime conditioning (friction transients, contact events, regime-boundary crossings). iKCE is therefore necessary but not sufficient to certify dynamic imagination.

B. Conditioning perturbations

Static iKCE measures internal kinematic self-consistency along a single rollout. To turn this into a diagnostic that *separates* kinematic from dynamic imagination, we drive iKCE through a dose-response curve over the conditioning state. For each base rollout, the world model generates K imagined rollouts under controlled perturbations of physically meaningful conditioning parameters, initial velocity v_0, friction coefficient μ, or lateral-acceleration limit $a_{\text{lat,max}}$ for driving, terrain compliance, or payload for legged locomotion, analogous to physically-grounded parameters for other embodiments. The perturbation set is embodiment-specific. The protocol is not.

Two diagnostic signatures emerge from the resulting $\{\text{iKCE}_k\}_{k=1}^{K}$ ensemble. First, the *shape* of iKCE($\|\Delta\|$) as a function of perturbation magnitude: a kinematic imaginer produces a curve that scales smoothly with $\|\Delta\|$ regardless of physical regime, because it is extrapolating the same linear update structure in every case. iKCE measures per-step kinematic-null deviation. Physical-regime perturbations whose effects are slow relative to the per-step timescale (e.g., friction-driven slipping in legged locomotion, which accumulates over multiple footfalls) are not visible at horizons shorter than their characteristic accumulation time. The diagnostic protocol should be applied at horizons longer than the embodiment's gait period: $\sim$25 ms $\times$ 64 steps $\approx$ 1.6 s is sufficient for walker-class locomotion. Driving trajectories are usually planned over longer horizons. Second, *rollout-pair monotone consistency*: physics predicts monotone responses to specific perturbation pairs (higher initial velocity implies longer stopping distance under braking; heavier payload implies slower acceleration), and for each such pair (Δ_a, Δ_b) we check whether the imagined rollouts respect the monotonicity. A kinematic world model trivially respects *linear* monotonicities but fails *physical-regime-conditional* ones, where the monotonicity only holds above or below a physical threshold.

This is the closest defensible analog to EgoDyn-Bench's weighted pairwise consistency rate [11]: rather than checking answer consistency across tagged question pairs on a single observation, we check rollout consistency across controlled conditioning perturbations on a single base scenario. The diagnostic is reframed rather than contrived, and produces falsifiable curves on any embodiment whose state admits a kinematic predictor.

IV. Experiments and Results

The experimental setup is documented in Appendix A. **Hypothesis 1 (H1): iKCE is non-degenerate on a published checkpoint.** At both measurement horizons, the trained DreamerV3 walker-walk world model produces an imagined iKCE at least an order of magnitude above matched policy-driven real-physics rollouts on the same conditioning (see Table I: $\sim$180$\times$ at T=16, $\sim$30$\times$ at T=64). The narrowing at the longer horizon is consistent with the per-step dilution noted in §V, limitation (iii): the WM's smooth post-transient tail averages down the integrated metric. iKCE is therefore non-degenerate at both horizons (see the trivial-WM scale anchor in Appendix A4 for the analytic lower bound). The WM's imagination carries a substantial residual against any constant-velocity null, with the H1 being a lower bound on the magnitude separation.

TABLE I

iKCE on the (z, v_z) view at μ=1.0 baseline, K=20 rollouts, 95% bootstrap CIs. WM iKCE exceeds physics by at least an order of magnitude at both measurement horizons. The ratio narrows at T=64 due to per-step dilution from the WM's smooth long-horizon tail (see §V, limitation (iii)).

Source	Mean iKCE	95% CI
Horizon T=16		
Real physics (matched policy)	4.2×10^{-5}	$[3.2, 5.3]\times10^{-5}$
DreamerV3 WM (imagined)	7.7×10^{-3}	$[5.2, 10.3]\times10^{-3}$
Ratio (WM / physics)	$\sim 180\times$	—
Horizon T=64		
Real physics (matched policy)	8.6×10^{-5}	$[6.4, 11.0]\times10^{-5}$
DreamerV3 WM (imagined)	2.6×10^{-3}	$[2.0, 3.2]\times10^{-3}$
Ratio (WM / physics)	$\sim 30\times$	—

Hypothesis 2 (H2): WM imagination is friction-invariant across a regime boundary. We sweep surface friction across 13 magnitudes in $[0.1, 1.7]$, spanning the regime boundary at μ=0.20 where the trained gait first drops below 50% of baseline episodic reward (empirically determined, see Fig. 1). For each μ, we compute iKCE on (a) real-physics rollouts under the trained actor and (b) WM-imagined rollouts conditioned on the first 5 perturbed observations.

At T=64, WM iKCE across the sweep is statistically flat (max/min spread 1.32$\times$, 95% CIs overlap at every μ, see Fig. 2). A log-log regression on the same sweep makes this falsifiable (Appendix A): the WM slope's 95% bootstrap CI ($\beta_{\text{WM}} = -0.009$, $[-0.096, +0.082]$) contains zero, while the physics slope's ($\beta_{\text{phys}} = -0.220$, $[-0.301, -0.142]$) does not. The trained policy's reward, by contrast, collapses through the same range (from $\sim$650 at μ=0.5 to $\sim$200 at μ=0.10), a real behavioral regime change to which the WM's imagined rollouts are blind. Real-physics iKCE under the trained actor shows more cell-to-cell variability than the WM, with elevated values in the low-μ region (means 1.04-1.30 $\times$ 10^{-4} across $\mu \in [0.1, 0.3]$ vs. 0.70–0.92 $\times$ 10^{-4} across $\mu \in [0.5, 1.7]$). This confirms that the iKCE metric is not degenerate, while the WM's invariance to the same perturbation provides the kinematic, not dynamic, signature. The structural difference is reinforced by appendix controls: a per-step decomposition (Fig. 6) shows that the physics and WM residuals differ qualitatively in temporal structure, not just in magnitude, and the same H2 signature reappears under a richer kinematic slice (Fig. 8). At the shorter T=16 horizon, neither side shows friction sensitivity. The dynamic signature emerges only as slip accumulates into per-step deviation at the longer horizon (quantified in Appendix Fig. 3). Controls supporting H2 are reported in detail in the appendix: per-step structure under the actor ablation (Fig. 7), robustness to the kinematic-state choice (Fig. 8), and a joint-noise positive control (Fig. 2, right panel)

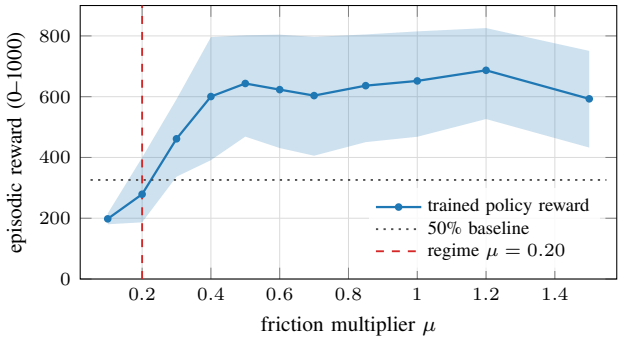

Fig. 1. **Empirical regime boundary.** Mean episodic reward of the trained DreamerV3 walker policy across friction (K=10, 95% bootstrap CI). The regime boundary at μ=0.20 (red dashed) is the friction at which mean reward first drops below 50% of the μ=1.0 baseline (dotted). This boundary anchors the dashed reference line in Fig. 2.

confirming the WM responds to kinematic-axis perturbations.

Actor training horizon ablation. An actor-training-horizon ablation (see Appendix Fig. 4) rules out the most concerning confound: retraining at imag horizon = 64 produces identical WM iKCE friction spread (1.32×), confirming H2 is not an artifact of the default actor's 15-step training horizon.

Domain-randomization control. A domain-randomization control (Appendix B2, Fig. 5) bounds the policy-OOD confound: under a policy trained with $\mu \sim \mathcal{U}(0.1, 1.7)$, the WM slope still contains zero ($\beta^{\mathrm{DR}}_{\mathrm{WM}} = -0.026$) and the physics slope still excludes it ($\beta^{\mathrm{DR}}_{\mathrm{phys}} = -0.114$), so the kinematic-not-dynamic contrast survives with the policy in-distribution at every swept friction.

V. Discussion & open directions

iKCE is a per-step kinematic null fit integrated over a horizon. It diagnoses kinematic imagination by the absence of regime sensitivity, not by the presence of dynamic prediction quality. A WM that achieves low iKCE everywhere has been correctly identified as kinematic by our protocol, but has not been certified as a useful predictor. A high-iKCE WM with friction sensitivity has been certified as dynamic but not as accurate. The diagnostic is structural, not predictive, by design.

A downstream behavioral prediction. If imagined rollouts are kinematically structured but not dynamically faithful, then policy gradients propagated through long imagined rollouts optimize the actor against a trajectory distribution that diverges from real dynamics in directions iKCE itself does not capture (rotational drift, contact-event timing, accumulated absolute-state error). Long-horizon actor training should therefore be unstable and yield a weaker deployed policy. The h=64 ablation is consistent with this prediction: under matched hyperparameters, the long-horizon actor converged to ∼400 episodic reward versus ∼955 for the default h=15 checkpoint, and exhibited training instability throughout. We do not claim a causal link. Long-horizon Dreamer training is known to be sensitive to multiple factors, but the observation is what one would predict from the kinematic-not-dynamic hypothesis, and pre-empts the natural counterfactual that scaling the actor's imagination horizon would have closed the gap.

Limitations of the present measurement. Several limitations bound the result. (i) The empirical result rests on a single embodiment (DMC walker-walk, a 2D 9-DOF system) restricted to the (z, v_z) sub-slice, and on a single open-weight WM family. Extending the flatness test for H2 to quadruped, humanoid, and driving embodiments, as well as to other WM families such as GAIA-1 [7] or R2Dreamer [9], would further broaden the evidence base. (ii) The policy-OOD confound of the physics-side signature is bounded, not eliminated, by the domain-randomization control (Appendix B2): the H2 contrast survives under a policy in-distribution at every swept μ, while the partition of the original low-μ elevation holds at the point-estimate level only. Two residuals remain: the DR policy adapts its gait to friction in closed loop, so the physics response is not policy-free; and imagined rollouts condition on only five observations (under one gait period), so the WM-side flatness is informative only up to the regime evidence the prefix can carry. (iii) Per-step displacement decays over the rollout horizon, so the WM's low long-horizon iKCE reflects in part reduced motion magnitude rather than purely cleaner kinematic imagination.

Open directions. Embodiment extension to quadruped-walk would test the diagnostic on richer contact dynamics than the planar walker provides. Fixed open-loop action sequences, applied identically across physics and WM rollouts, would remove the closed-loop policy adaptation left in place by the domain-randomization control. An explicit-conditioning experiment, in which friction or contact indicators are appended to the WM's observation, together with a conditioning-prefix-length sweep (5–64 observed steps), would distinguish representational absence from architectural insensitivity – and from regime evidence the prefix simply cannot carry. A driving-WM cross-anchor on a model exposing an ego-pose head (Vista [1], DriveDreamer [12]) would connect the measurement to the autonomous-driving evidence of §II.

References

[1] Shenyuan Gao, Jiazhi Yang, Li Chen, Kashyap Chitta, Yihang Qiu, Andreas Geiger, Jun Zhang, and Hongyang Li. Vista: A generalizable driving world model with high fidelity and versatile controllability, 2024. URL https://arxiv.org/abs/2405.17398.

[2] Yuan Gao, Dengyuan Hua, Mattia Piccinini, Finn Rasmus Schäfer, Korbinian Moller, Lin Li, and Johannes Betz. StyleVLA: Driving style-aware vision language action model for autonomous driving. *arXiv preprint arXiv:2603.09482*, 2026.

[3] David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[4] Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In

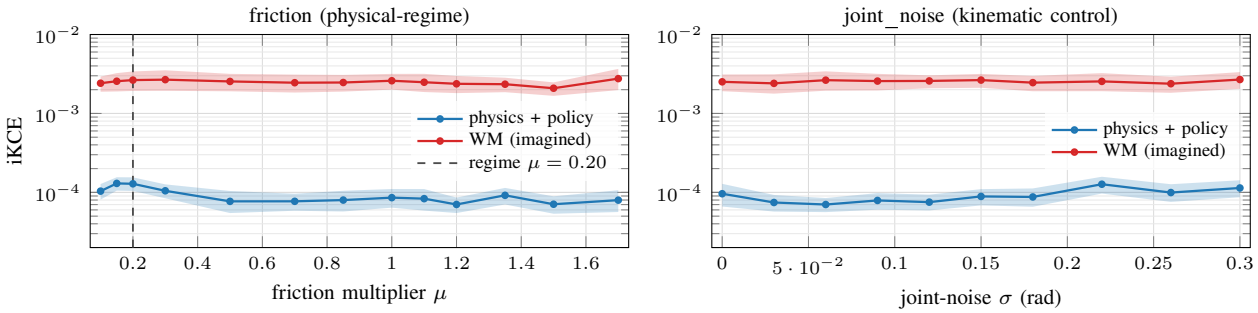

Fig. 2. **iKCE diverges in physics, stays flat in imagination.** Identity kinematic view (z, v_z) at horizon $T = 64$. *Left:* friction sweep $\mu \in [0.1, 1.7]$ (physical regime axis). Real-physics rollouts under the trained policy (blue) show modestly elevated iKCE in the low-μ region near the empirical regime boundary (μ=0.20, dashed line (see Fig. 1)). WM-imagined rollouts (red) are statistically flat across the entire 17× sweep. *Right:* joint-noise sweep $\sigma \in [0, 0.3]$ rad (kinematic control axis). Both channels respond similarly, confirming that the WM is not insensitive to all perturbations, only to dynamic ones. Shaded bands: 95% bootstrap CI from $K = 20$ rollouts per cell.

International Conference on Machine Learning (ICML), 2019.

[5] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *Nature*, 640:647–653, 2025.

[6] Anthony Hu, Gianluca Corrado, Nicolas Griffiths, Zak Murez, Corina Gurau, Hudson Yeo, Alex Kendall, Roberto Cipolla, and Jamie Shotton. Model-based imitation learning for urban driving. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[7] Anthony Hu, Lloyd Russell, Hudson Yeo, Zak Murez, George Fedoseev, Alex Kendall, Jamie Shotton, and Gianluca Corrado. Gaia-1: A generative world model for autonomous driving, 2023. URL https://arxiv.org/abs/2309.17080.

[8] Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[9] Naoki Morihira, Amal Nahar, Kartik Bharadwaj, Yasuhiro Kato, Akinobu Hayashi, and Tatsuya Harada. R2-dreamer: Redundancy-reduced world models without decoders or augmentation, 2026. URL https://arxiv.org/abs/2603.18202.

[10] Abhinaw Priyadershi and Jelena Frtunikj. Lost in fog: Sensor perturbations expose reasoning fragility in driving VLAs. *arXiv preprint arXiv:2605.21446*, 2026.

[11] Finn Rasmus Schäfer, Yuan Gao, Dingrui Wang, Thomas Stauner, Stephan Günnemann, Mattia Piccinini, Sebastian Schmidt, and Johannes Betz. Egodyn-bench: Evaluating ego-motion understanding in vision-centric foundation models for autonomous driving, 2026. URL https://arxiv.org/abs/2604.22851.

[12] Xiaofeng Wang, Zheng Zhu, Guan Huang, Xinze Chen, Jiagang Zhu, and Jiwen Lu. Drivedreamer: Towards real-world-driven world models for autonomous driving, 2023. URL https://arxiv.org/abs/2309.09777.

[13] Jiang-Tian Zhai, Ze Feng, Jinhao Du, Yongqiang Mao, Jiang-Jiang Liu, Zichang Tan, Yifu Zhang, Xiaoqing Ye, and Jingdong Wang. Rethinking the open-loop evaluation of end-to-end autonomous driving in nuscenes, 2023. URL https://arxiv.org/abs/2305.10430.

APPENDIX

This appendix documents (i) the experimental configuration (Table II), (ii) the methodological details behind the headline numbers (§A1-A4: regime-boundary determination, the flatness and horizon-emergence tests, and the trivial-WM scale anchor), (iii) the controls supporting H2 (§B1-B6: actor-training-horizon ablation, domain-randomization control, per-step structure decomposition, per-step structure under the actor ablation, robustness to the kinematic-state choice, and the joint-noise positive control), and (iv) implementation specifics needed to reproduce the measurement (§C1-C2). Code and data will be released upon acceptance.

TABLE II
EXPERIMENTAL SETUP. HORIZONS h=16 AND h=64 REFER TO THE iKCE ROLLOUT LENGTH, EVALUATED ON THE SAME TRAINED POLICY.

Field	Value
World model	DreamerV3 (NM512 PyTorch port, commit `6ef8646`)
Task	DMC walker-walk, `dmc_proprio` config
Training	1M env steps, seed 0, RTX 5090
Final reward	955 ± 30 (mean over last 100k steps)
Evaluation policy	Trained actor (same checkpoint on both physics and WM sides)
Backend	`dm_control` 1.0.20, `mujoco` 3.1.6
K	20 rollouts per perturbation cell
Kinematic spec	(z, v_z) root-vertical-motion (1D)
Extrapolation	constant velocity

A. Methodological Details

1) Regime-boundary determination.: The boundary μ=0.20 referenced in Fig. 2 (dashed line) and Fig. 1 is not chosen a priori. It is determined empirically from the policy's reward collapse. We roll out the trained actor for K=10 episodes at each of 12 friction multipliers $\mu \in [0.1, 1.5]$, compute the mean episodic reward and a 95% bootstrap CI per cell, and define the regime boundary as the *largest* μ at which mean reward has dropped below 50% of its μ=1.0 baseline. The μ=1.0 mean over this sweep is ∼650 (K=10 episodes), giving a threshold of ∼325. On our checkpoint, the boundary lies at μ=0.20. (The 955 ± 30 final reward in Table II averages over the last 100k training steps and is not directly comparable.)

2) Flatness test for H2.: We make the "statistically flat" claim falsifiable by regressing $\log(\text{iKCE})$ on $\log(\mu)$ across the T=64 friction sweep, with each of the K=20 rollouts at each of the 13 μ values contributing one observation (n=260 per seed). To rule out a seed-dependent artifact, we repeat the regression for three independently-trained DreamerV3 walker-walk checkpoints (seeds $\{0, 1, 2\}$, matched hyperparameters and step budget; see Appendix A). A 95% percentile bootstrap (1000 resamples over rollouts) on each slope gives: $\beta^{(0)}_{\text{WM}} = -0.009$, CI $[-0.096, +0.082]$; $\beta^{(1)}_{\text{WM}} = +0.031$, CI $[-0.072, +0.129]$; $\beta^{(2)}_{\text{WM}} = +0.038$, CI $[-0.039, +0.123]$, all three WM slope CIs contain zero, so H2's flatness claim survives a falsifiable statistical test across seeds, not just on the original training run. The same procedure on the physics-side sweep (seed 0 only, as physics is not a learned model) gives $\beta_{\text{phys}} = -0.220$, CI $[-0.301, -0.142]$, comfortably excluding zero. The low-μ elevation reported in §IV is statistically real. The quantitative form of the kinematic-not-dynamic signature is therefore: across three seeds, $|\beta_{\text{WM}}|$ is bounded above by ∼0.13 (one decade of μ changes WM iKCE by at most ∼13%), while $|\beta_{\text{phys}}| \approx 0.22$ (∼25% per decade), with non-overlapping confidence intervals.

3) Horizon-emergence test.: Section III-B argues that the diagnostic should be applied at horizons longer than the embodiment's gait period. We sharpen this claim quantitatively by repeating the flatness regression of the preceding paragraph at four sub-horizons $T \in \{8, 16, 32, 64\}$, re-integrating each rollout's saved per-step iKCE trace from the existing T=64 sweep (no new rollouts, see Fig. 3). The dynamic signature in physics emerges with horizon: the slope β_{phys} grows in magnitude from $+0.012$ (CI $[-0.121, +0.146]$) at T=8 to -0.221 (CI $[-0.301, -0.142]$) at T=64, crossing out of the CI-contains-zero region between T=32 and T=64 – consistent with friction effects accumulating over multiple footfalls before becoming detectable in the per-step kinematic-null residual. The WM-side slope β_{WM} is statistically indistinguishable from zero at every horizon tested (-0.028, -0.012, -0.013, -0.009 at T=8, 16, 32, 64, with CIs of width ≤ 0.32 all straddling zero). The contrast is the H2-emergence claim stated quantitatively: the dynamic signature emerges with horizon in physics but not in the WM. Note that long-horizon iKCE in both channels is in part dilated by reduced per-step motion magnitude (§V, limitation (iii)). The present result is robust to that effect because the WM-physics contrast *widens* with horizon rather than shrinks, which is the opposite of what a horizon-degenerate metric would produce.

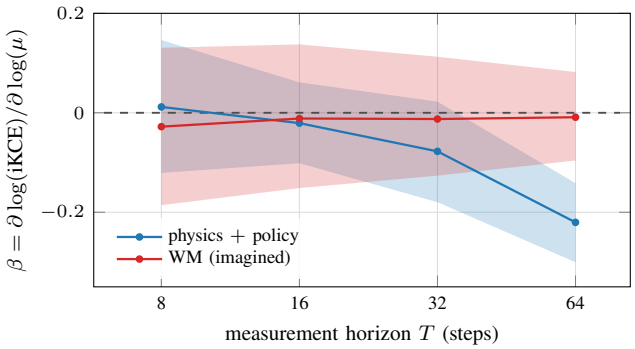

Fig. 3. **Horizon-emergence test.** Slope $\beta = \partial \log(\text{iKCE}) / \partial \log(\mu)$ of the friction sweep at four measurement horizons $T \in \{8, 16, 32, 64\}$, computed by re-integrating the saved T=64 per-step iKCE traces (no new rollouts). Physics slope (blue) grows in magnitude with T and crosses out of the CI-contains-zero region by T=64. WM slope (red) is statistically indistinguishable from zero at every horizon. Dashed line marks the H2 flatness target ($\beta = 0$). Shaded bands: 95% percentile bootstrap CI over 1000 resamples at the rollout level.

4) Trivial-WM scale anchor.: The iKCE scale has an analytic lower bound that anchors the magnitudes in Table I. A trivial "WM" that imagines by applying the kinematic

predictor to its own current state, $\hat{x}_{t+1}^{\mathrm{WM}} = \mathrm{kin}(\hat{x}_t^{\mathrm{WM}})$, produces iKCE $= 0$ by construction: each predicted next state is identically the kinematic continuation of its predecessor, so every residual in Eq. 1 is zero. The measured ordering on walker-walk is therefore

$$\underbrace{0}_{\text{trivial kinematic}} \quad < \quad \underbrace{4.2 \times 10^{-5}}_{\text{matched real physics}} \quad \ll \quad \underbrace{7.7 \times 10^{-3}}_{\text{DreamerV3 WM}} \quad (T{=}16).$$

The WM lies further from the trivial-kinematic baseline than real physics does, ruling out a naive reading in which "imagining kinematically" would imply small absolute iKCE. The diagnostic signature, per Section III, is friction-invariance under perturbation, not low absolute magnitude. The order holds for $T{=}64$ as well.

B. Controls

1) Actor-training-horizon control.: A natural concern is that the WM's friction-invariance at $T{=}64$ reflects the default actor operating out-of-distribution from its training horizon (`imag horizon` $= 15$) rather than a structural property of WM imagination. To rule this out, we retrain an identical checkpoint at `imag horizon` $= 64$, matching the measurement horizon, with the same seed, hyperparameters, and total training budget. Fig. 4 shows the result: WM iKCE friction spread under the $h{=}64$-trained actor is identical to the default-actor headline ($1.32\times$ in both cases, CIs overlapping at every μ), confirming that friction-invariance is not an artifact of actor training horizon.

2) Domain-randomization control.: Limitation (ii) in §V names the most consequential confound of the physics-side H2 signature: the evaluation policy acted only at $\mu{=}1.0$ during training, so the elevated low-μ physics iKCE may reflect an in-distribution policy slipping under out-of-distribution friction rather than a genuine friction response of the contact dynamics. To quantify this confound, we train a fourth, otherwise identical DreamerV3 checkpoint with per-episode domain randomization of friction, $\mu \sim \mathcal{U}(0.1, 1.7)$ drawn at every episode reset (matched hyperparameters and step budget; evaluation reward 930 ± 36 across the full sweep range). Both the DR policy and the DR world model are therefore in-distribution at every friction value of the H2 sweep. The DR policy exhibits no reward collapse anywhere in the tested range, so the regime boundary of Fig. 1 is a property of the fixed-μ policy and does not transfer to this control.

Table III reports the flatness regression of Appendix A2 under the DR checkpoint, alongside the fixed-μ results; Fig. 5 shows the underlying sweeps. *WM side:* the DR-trained world model is statistically flat ($\beta_{\mathrm{WM}}^{\mathrm{DR}} = -0.026$, CI $[-0.123, +0.076]$), indistinguishable from the three fixed-μ seeds. This closes a data-coverage loophole in the seed-level flatness test: a world model trained only at $\mu{=}1.0$ has never observed friction variation and cannot have learned friction-conditional latent dynamics, so its flatness is partially guaranteed by construction. The DR world model was trained on transitions spanning the full sweep range, could in principle infer the friction regime from its conditioning prefix and imagine regime-conditional rollouts – and it remains friction-invariant (Fig. 5, left). The flatness signature is therefore not an artifact of the training-time friction distribution.

Physics side: under the DR policy, the physics slope remains strictly negative ($\beta_{\mathrm{phys}}^{\mathrm{DR}} = -0.114$, CI $[-0.201, -0.024]$), at roughly half the default-policy magnitude (-0.220). Read at the point-estimate level, this partitions the original low-μ elevation into comparable parts: about half attributable to an out-of-distribution policy slipping, about half a genuine friction response of the contact dynamics that persists when the policy is in-distribution everywhere. We state this partition at the point-estimate level only: a percentile bootstrap on the difference of the two physics slopes does not resolve it at $K{=}20$ ($\beta_{\mathrm{phys}}^{\mathrm{default}} - \beta_{\mathrm{phys}}^{\mathrm{DR}} = -0.107$, CI $[-0.221, +0.004]$).

Under matched DR conditions on both sides, the H2 contrast retains its falsifiable form: the physics slope's CI excludes zero while the WM slope's contains it. The slope difference itself is directionally consistent but not resolved at this sample size ($\beta_{\mathrm{phys}}^{\mathrm{DR}} - \beta_{\mathrm{WM}}^{\mathrm{DR}} = -0.087$, CI $[-0.211, +0.049]$).

This control addresses the training-distribution side of limitation (ii). Two residual caveats remain, both deferred to the open directions: the DR policy still adapts its gait to friction in closed loop, so the fully policy-free variant (fixed open-loop action sequences applied identically to both channels) remains the cleaner disambiguation; and the imagined rollouts condition on only five observed steps (125 ms, under one gait period), which bounds how much regime evidence even a dynamically capable imaginer could extract from the prefix.

TABLE III
DOMAIN-RANDOMIZATION CONTROL: FLATNESS REGRESSION $\beta = \partial \log(\mathrm{iKCE}) / \partial \log(\mu)$ AT $T{=}64$, 95% PERCENTILE BOOTSTRAP CIS (1000 RESAMPLES AT THE ROLLOUT LEVEL, $K{=}20$ PER CELL). ALL FOUR WM CHECKPOINTS ARE STATISTICALLY FLAT REGARDLESS OF THE FRICTION DISTRIBUTION SEEN AT TRAINING TIME; BOTH PHYSICS SLOPES ARE STRICTLY NEGATIVE.

Side	Checkpoint / policy	β	95% CI
WM	seed 0 (fixed $\mu{=}1.0$)	-0.009	$[-0.096, +0.082]$
WM	seed 1 (fixed $\mu{=}1.0$)	$+0.031$	$[-0.072, +0.129]$
WM	seed 2 (fixed $\mu{=}1.0$)	$+0.038$	$[-0.039, +0.123]$
WM	DR ($\mu \sim \mathcal{U}(0.1, 1.7)$)	-0.026	$[-0.123, +0.076]$
Physics	default policy (seed 0)	-0.220	$[-0.301, -0.142]$
Physics	DR policy	-0.114	$[-0.201, -0.024]$

3) Per-step structure decomposition.: The integrated iKCE of Table I and Fig. 2 averages over the rollout horizon and so does not reveal whether the WM's imagined residual has the same temporal structure as physics. Fig. 6 decomposes per-step iKCE at three friction values $\mu \in \{0.15, 1.0, 1.5\}$: physics exhibits sparse contact-event spikes whose positions shift with μ (consistent with footfall dynamics driving the kinematic-null residual), while the WM shows a one-to-two-step encoder-decoder transient followed by a smooth, friction-invariant tail. The two residuals are not merely different in magnitude (the H1 ratio) but qualitatively different in temporal structure, the WM's imagined rollouts do not reproduce the contact-event signature that defines the physics-side residual.

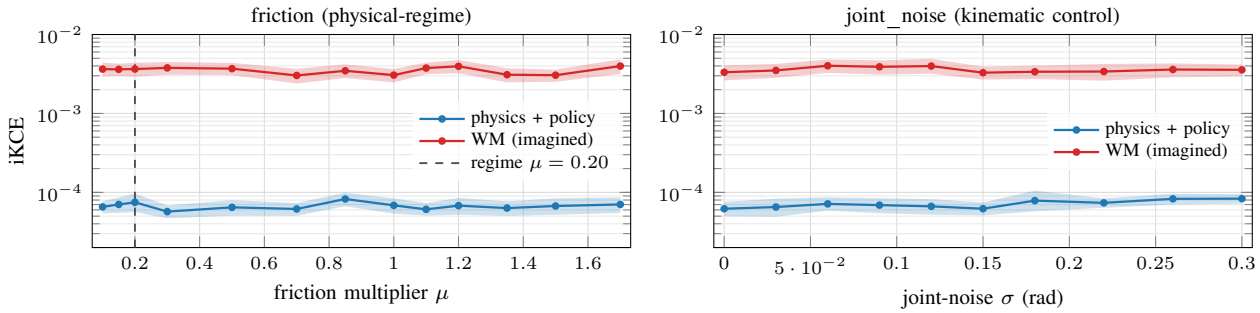

Fig. 4. **Actor-training-horizon ablation.** Identity view (z, v_z) at T=64, with the actor retrained at `imag horizon` $= 64$ (matching the measurement horizon). WM iKCE friction spread is identical to the default-actor headline ($1.32\times$ in both cases), confirming the friction-insensitivity of imagined rollouts is not an artifact of the default actor's h=15 training horizon. Shaded bands: 95% bootstrap CI from $K = 20$ rollouts per cell.

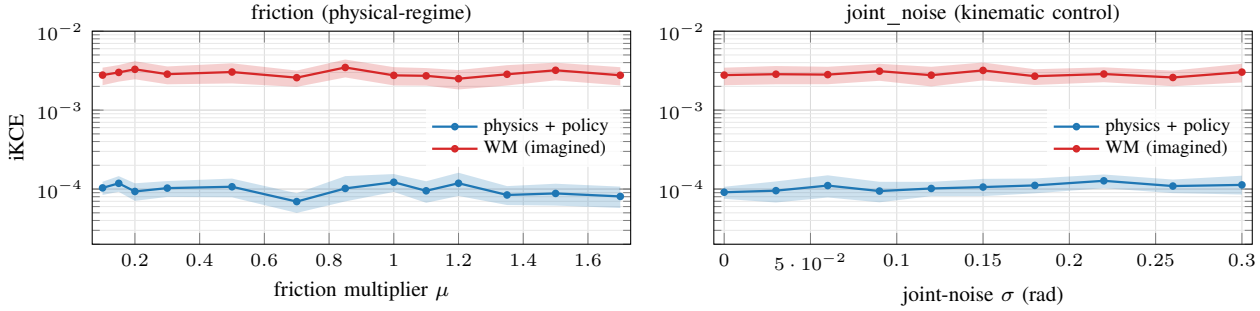

Fig. 5. **Domain-randomization control.** Identity view (z, v_z) at T=64 with friction domain-randomized at training time ($\mu \sim \mathcal{U}(0.1, 1.7)$ per episode). Left: friction sweep $\mu \in [0.1, 1.7]$ (physical-regime axis). WM-imagined rollouts remain statistically flat despite the world model having been trained on the full friction range; physics-side iKCE under the DR policy retains a strictly negative slope, with the point estimate at roughly half the default-policy magnitude. Right: joint-noise sweep $\sigma \in [0, 0.3]$ rad (kinematic control axis). The DR-trained WM responds to kinematic-axis perturbations, replicating the positive control of Fig. 2 under the DR checkpoint. Shaded bands: 95% bootstrap CI from K=20 rollouts per cell.

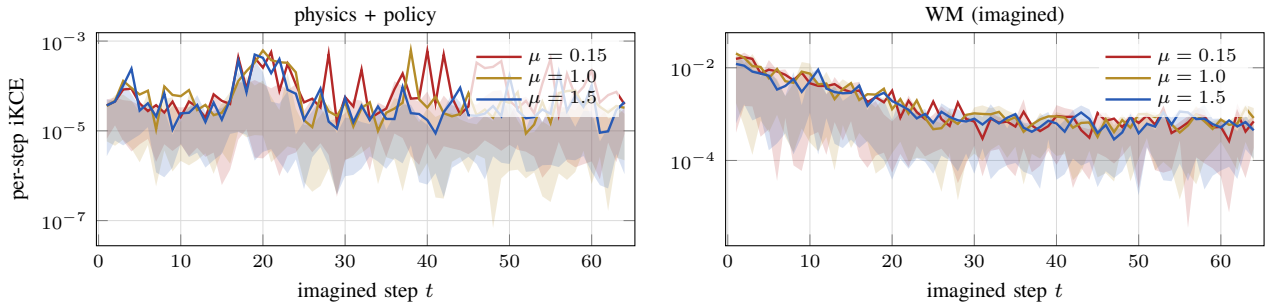

Fig. 6. **Per-step iKCE structure.** Log-y. Physics (left) has discrete contact-event spikes whose positions shift with μ. WM (right) shows a one- to two-step encoder-decoder transient followed by a smooth tail. The per-step structure is essentially constant across friction.

4) Per-step structure under the actor-ablation checkpoint.: Figure 6 reports the per-step decomposition under the default actor. Figure 4 reports the integrated iKCE under the retrained h=64 actor. Figure 7 combines the two controls: per-step WM iKCE under both the default (h=15) and retrained (h=64) actor, at the same three friction values $\mu \in \{0.15, 1.0, 1.5\}$. The transient-plus-smooth-tail structure is unchanged across actor training horizons, ruling out the joint concern that the per-step signature reflects an actor-out-of-distribution artifact rather than a property of the WM's imagination.

5) Robustness to the kinematic-state choice.: iKCE depends on a chosen kinematic state vector $\hat{x}_t$ (Definition 1). The main result uses the root-vertical-motion slice (z, v_z). Figure 8 repeats the protocol with a richer representation, the walker's gait degrees of freedom. The qualitative pattern of Fig. 2 reappears: WM iKCE is flat across friction, while physics shows low-μ elevation, confirming that H2 reflects a property of the WM's imagination rather than the particular state slice used to evaluate it. This generalizes the diagnostic claim: any kinematic state extraction whose dynamics are sensitive to the perturbation axis can serve as a probe, with no special status accorded to the identity slice.

6) Joint-noise as a kinematic positive control.: The friction sweep is a dynamic perturbation (physically-grounded, regime-

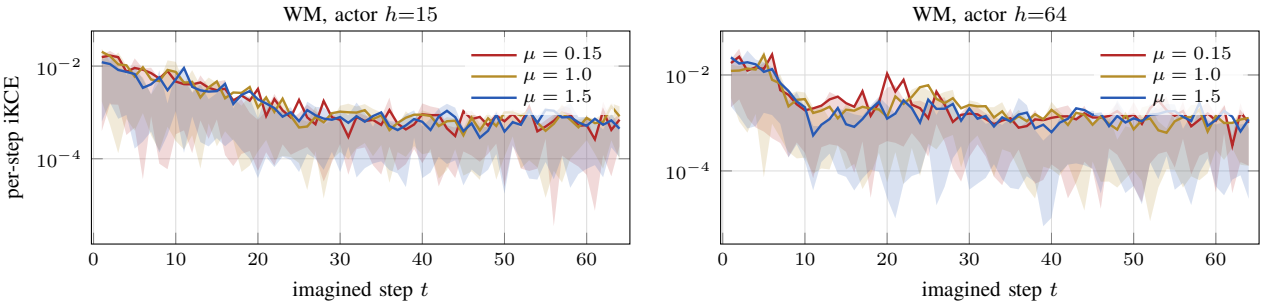

Fig. 7. **Per-step actor ablation.** WM per-step iKCE at three friction values for both actor checkpoints. The transient-plus-smooth-tail structure is independent of the actor training horizon.

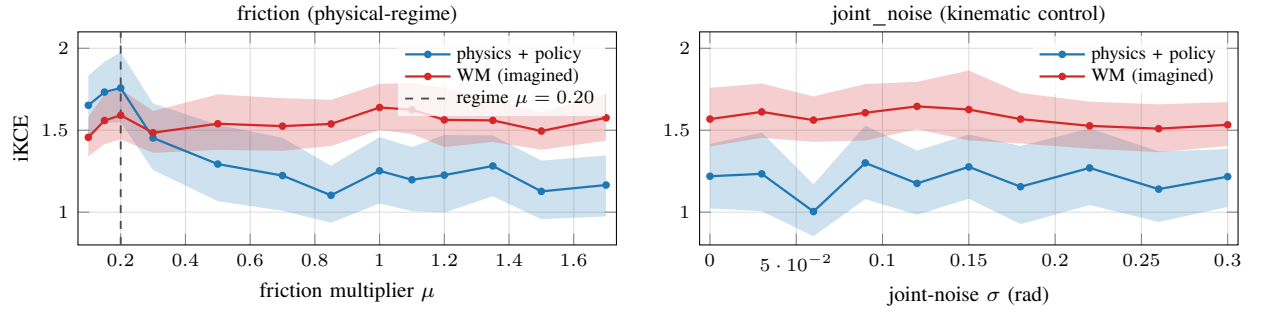

Fig. 8. **Gait DOFs view.** Same protocol as Fig. 2 but using the walker_gait kinematic slice. Physics still shows low-μ elevation, WM is still flat, the finding is not an artifact of the (z, v_z) identity view.

crossing). The joint-noise sweep is its kinematic counterpart (zero-mean Gaussian noise added to the joint-position channel of every observation before the WM encoder, with the physics state left unperturbed). A kinematic imaginer should respond to the joint-noise sweep. The perturbation directly modifies the kinematic state $\hat{x}_t$ that drives the extrapolation, while a dynamic imaginer should respond to the friction sweep. The right panel of Fig. 2 confirms that both channels respond to joint noise, ruling out the alternative explanation that the WM's iKCE is simply insensitive to all perturbations. The contrast (response to joint noise, non-response to friction) is the diagnostic signature named in the paper's title.

C. Reproducibility

1) Code and data availability.: The diagnostic pipeline, trained checkpoints, perturbation-sweep CSVs, and PGFPlots figure sources for this paper are released at https://github.com/TUM-AVS/iKCE. The DreamerV3 implementation is the NM512 PyTorch port at commit `6ef8646`. The upstream algorithm is [5]. Checkpoints are released at https://huggingface.co/fnc1901/ikce-walker-walk-artifacts. All experiments ran on a single RTX 5090 (training: ∼24 h per checkpoint. The full perturbation sweep, including the actor-horizon ablation: ∼2 h).

2) Implementation specifics.: For the identity view, the kinematic predictor is constant-velocity on the root vertical state: $\mathrm{kin}([z_t, \dot{z}_t]) = [z_t + \Delta t\, \dot{z}_t, \dot{z}_t]$ with $\Delta t = 25\,\mathrm{ms}$ (the DMC physics timestep). For the gait view, $\hat{x}_t$ stacks the unit-circle embedding $(\cos\theta_j, \sin\theta_j)$ of each walker joint angle, and $\mathrm{kin}(\cdot)$ applies one-step extrapolation per joint. Friction perturbations are applied by scaling the MuJoCo friction tuple of every geom by μ at episode reset (`model.geom_friction[:, 0] *= mu`). Joint-noise perturbations add zero-mean Gaussian noise with standard deviation σ (rad) to the joint-position channel of every observation before the WM encoder, with the physics state itself left unperturbed. For WM-imagined rollouts, we condition on the first 5 perturbed observations (encoder unroll), then roll out free imagination for the remaining $T - 5$ steps. This matches the standard Dreamer evaluation protocol. All 95% confidence intervals are computed using a percentile bootstrap with 1000 resamples across the K=20 rollouts per cell.