SPLIT THE WORLD: ARE REGIONAL DYNAMICS BETTER THAN GLOBAL DYNAMICS FOR LATENT WORLD MODELS?

Anonymous authors
Paper under double-blind review

ABSTRACT

Latent world models, due to their efficiency, have become one of the main paradigms for world modeling. In latent space, they learn from data the dynamics governing how an agent's actions affect the environment and use these dynamics to predict the environment state after executing actions. Traditionally, most existing latent world models learn a single global dynamics model from the full task dataset, despite possible differences in dynamics across different regions of the task state space. Existing approaches have shown that regional dynamics can be useful. However, either global or regional dynamics are used by default, without a clear and general theory for selecting between them. Motivated by this, we study when regional dynamics are better than global dynamics and formulate this question as a dynamics-selection problem. To answer this question, we propose the Worst-Case Partition Gain criterion, which is based on Wasserstein distances between regional response distributions. Using this criterion, we establish a necessary-and-sufficient condition for selecting the better dynamics under reasonable assumptions. Furthermore, we develop a dynamics-selection algorithm based on this condition, with a calibrated threshold held fixed across different modeling configurations. We evaluate our method using, to our knowledge, the first systematic experimental framework for dynamics selection, comparing 22 methods on transfer across different numbers of regions and world-model backbones. Experimental results demonstrate state-of-the-art selection performance: our method correctly selects the better dynamics in all 8 region-count transfer configurations and in 18 of 20 experimental configurations overall.

1 INTRODUCTION

Latent world models have recently achieved remarkable progress across a diverse set of applications, including continuous control in robotics (Hansen et al., 2022), planning in video games and board games (Schrittwieser et al., 2020), and autonomous driving (Hu et al., 2022).By encoding high-dimensional observations into compact latent vectors, these models efficiently learn the dynamics governing how actions affect the environment (Ha and Schmidhuber, 2018; Hafner et al., 2019).

Most existing approaches train a single dynamic model on the full task dataset (Hafner et al., 2019; 2020; Maes et al., 2026), which we refer to as `global` dynamics modeling. Yet, this `global` modeling paradigm often fails to capture nuanced variations in system dynamics, across different regions of the task state space. For example, free motion and contact-rich interaction can exhibit substantially different dynamics within the same task (Abdul Khader et al., 2020). Consequently, a `global` model may fail to capture some local accurately transitions, which can critically undermine long-horizon planning and control performance.

To overcome this challenge, recent research has explored region-specific modeling of task space (Abdul Khader et al., 2020; Ahn and Sentis, 2021), leading to the concept of `regional` dynamics—state-transition equations tailored to distinct partitions of the environment. However, these studies typically assume that `regional` dynamics are uniformly preferable without systematically comparing `regional` and `global` dynamics. In fact, previous study has shown that the regional partition can be inconsistent with the true environment dynamics. As an extreme case, when the `regional`

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

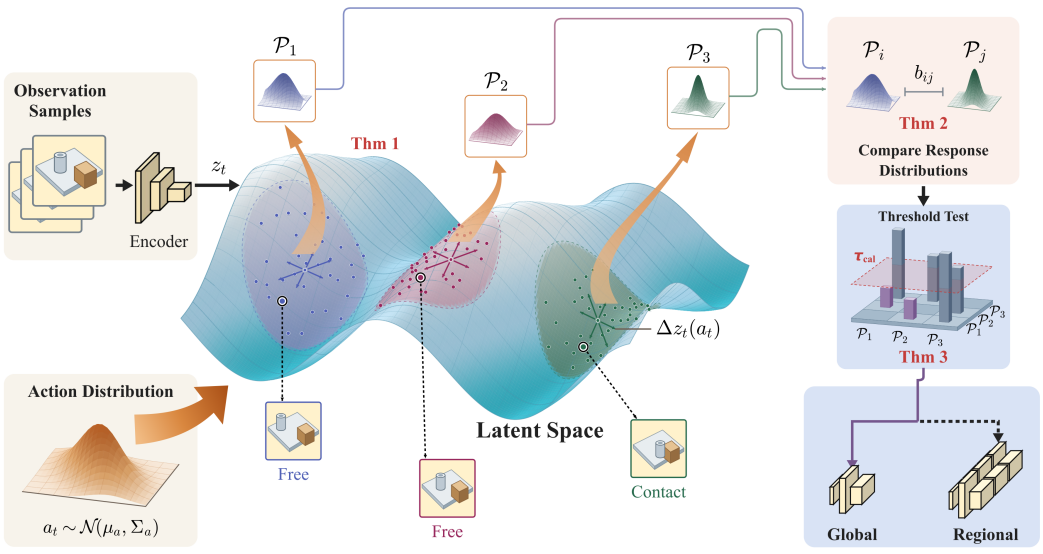


Figure 1: **Overview of the dynamics-selection framework.** Given a fixed partition of latent space, we summarize each region's dynamics by its response distribution $\mathcal{P}_i$. Comparing $\mathcal{P}_i$ and $\mathcal{P}_j$ produces a pairwise score b_{ij} that measures how different the dynamics of the two regions are. Finally, we compare the weakest-pair score (purple) with the calibrated threshold τ_{cal} (red plane), selecting `regional` dynamics modeling when the score exceeds the threshold and `global` dynamics modeling otherwise. The detailed procedure and related formulas are provided in Section 3.

dynamics throughout different partitions are identical or sufficiently similar, fitting separate models may offer little reduction in approximation bias while sacrificing statistical effectiveness (Cerviño et al., 2023).

This leads to a fundamental and previously underexplored question: **Under what conditions do `regional` dynamics outperform `global` dynamics for latent world models?** We rigorously formalize this as a dynamics-selection problem: Given a state-space partition and an encoder, how can we select—without retraining—between `regional` and `global` dynamics models? Recognizing that planning unfolds in latent space, we derive a principled selection criterion directly within this representation. Theoretically, we establish that under reasonable assumptions, sufficiently different local dynamics remain separated in the latent space (Theorem 1).

To construct this criterion, we start from the principle that regions exhibiting sufficiently similar dynamics should share a single dynamics model, whereas separate models are warranted only when their dynamics differ sufficiently. Since Markovian planning proceeds by sequential application of one-step transitions, we characterize each region's local dynamics by its response distribution, that is, the distribution of action-induced one-step displacements in latent space. We quantitatively assess the dynamical similarity between any two regions by comparing their response distributions: a small distributional distance favors sharing a model, whereas a large distance supports separate modeling. After traversing all regions in the partition, we obtain a comparison for every pair of regions.

A central challenge then becomes **how to aggregate these pairwise assessments to select for the full partition**. Under reasonable assumptions, the pair with the smallest loss acts as the bottleneck: this critical pair determines which modeling paradigm yields superior long-horizon planning success, leading to our Worst-Case Partition Gain criterion. Notably, our theory provides a sharp characterization: `Regional` models provably outperform `Global` model exactly when the Worst-Case Partition Gain score exceeds a threshold shared across different world-model architectures (Theorem 3).

Figure 1 summarizes the resulting framework, from estimating `regional` response distributions to computing pairwise response distribution distances and performing threshold-based dynamics selection.

We make three contributions. **We formulate `regional`-versus-`global` dynamics modeling as a dynamics-selection problem and provide a fair evaluation framework.** To our knowledge, this is the first framework that isolates its effect by holding all other components fixed. **We establish a necessary-and-sufficient condition for dynamics selection.** Under reasonable assumptions, its theoretical threshold $\tau_\star$ is shared across modeling configurations. The resulting algorithm selects between `regional` and `global` dynamics from latent transitions, given a partition and an encoder, without first training both models. **We validate the proposed method experimentally.** Across 22 methods under a shared calibration and evaluation protocol, Worst-Case Partition Gain achieves state-of-the-art selection performance, correctly selecting the better dynamics in all 8 region-count transfer configurations and in 18 of 20 test configurations overall.

2 Background and Related Work

Existing dynamics-modeling approaches take both sides of the choice studied here. Most latent world models use a single dynamics model over the task data (Ha and Schmidhuber, 2018; Hafner et al., 2019; 2020; Hansen et al., 2022), while hybrid, modular, and expert-based approaches specialize dynamics across modes, mechanisms, or tasks (Abdul Khader et al., 2020; Ahn and Sentis, 2021; Lei et al., 2023; 2024; Zhao et al., 2025; Li et al., 2025; Zhang et al., 2026). These works establish that specialized dynamics can be useful and provide several ways to represent or exploit heterogeneous dynamics.

A further difficulty is that independently trained world models may learn different representations, and representation geometry can itself affect planning outcomes (Hansen et al., 2024; Zhou et al., 2025; Hu et al., 2026). Consequently, a performance difference between separately trained `Regional` and `Global` dynamics models need not be caused by their dynamics parameterization alone. Joint-Embedding Predictive Architecture (JEPA) and recent action-conditioned variants make this distinction explicit: an encoder defines the latent representation, while a predictor models the action-conditioned latent dynamics used for planning (LeCun, 2022; Klindt et al., 2026; Maes et al., 2026; Zhao et al., 2026). By sharing a frozen encoder, the two branches can therefore be compared in the same planning space while varying only whether dynamics parameters are shared across regions.

The following works are most closely related to ours and address different forms of specialization. VCD and COMET discover and reuse modular causal or compositional mechanisms across environments (Lei et al., 2023; 2024); their question is mechanism discovery and reuse rather than whether specialization should be used. M3W and Mixture-of-World Models employ expert dynamics to handle multi-task heterogeneity (Zhao et al., 2025; Zhang et al., 2026), but study task-conditioned modularity and architecture-level specialization rather than a matched `Regional`-versus-`Global` selection for a fixed partition. PRISM-WM is closest to our setting: it compares a gated, orthogonalized MoE world model with a monolithic baseline (Li et al., 2025). However, its comparison changes both the dynamics and reward models, so differences in performance cannot be attributed to dynamics specialization alone. Thus, prior work studies how to construct or exploit specialized dynamics, whereas we study whether specialization should be used at all. Given a partition and a fixed planning representation, we ask when its dynamics should be modeled regionally or globally.

3 A Geometric Theory of Dynamics Selection

We consider a latent world model with an encoder h, an actual latent state-transition function f, and a dynamics model $\widehat{f}$. The encoder maps an environment state $\mathbf{s}_t$ to a latent vector $\mathbf{z}_t = h(\mathbf{s}_t)$. Given action $\mathbf{a}_t$, the actual next latent state is $\mathbf{z}_{t+1} = f(\mathbf{z}_t, \mathbf{a}_t)$, while the dynamics model predicts $\widehat{\mathbf{z}}_{t+1} = \widehat{f}(\mathbf{z}_t, \mathbf{a}_t)$. For a partition Π into $K \geq 2$ regions, `Regional` uses one dynamics model per region, whereas `Global` shares one model across all regions.

Our goal is to choose which approach achieves higher long-horizon planning success before training dynamics models. Given the encoder, partition, and available latent transitions $(\mathbf{z}_t, \mathbf{a}_t, \mathbf{z}_{t+1})$, we seek a criterion for comparing the two approaches. Formal assumptions and proofs are provided in Appendix G.

3.1 Dynamics Differences in Latent Space

We summarize the local dynamics around an environment state by its action response, defined by

$$\Delta \mathbf{z}_t(\mathbf{a}_t) := f(\mathbf{z}_t, \mathbf{a}_t) - \mathbf{z}_t = \mathbf{z}_{t+1} - \mathbf{z}_t, \tag{1}$$

If the latent state is Markov, its one-step transition kernel determines the conditional next-state distribution, while a multi-step rollout is obtained by repeatedly composing one-step transitions. The action-dependent one-step displacement is therefore a natural local object for studying dynamics relevant to planning.

Because planning takes place in latent space, the dynamics differences used for selection need to remain distinguishable there. Since action responses may have poor continuity in latent space, we assume that they can be approximated by a Lipschitz-continuous function with bounded error (Assumption 1).

Theorem 1 (Latent response separation, informal)**.** *Under Assumption 1, states with sufficiently different action responses remain separated in latent space.*

Unfortunately, Theorem 1 alone cannot determine whether modeling a given partition separately is beneficial. We therefore compare the distributions of action responses across the given regions.

3.2 The Distance of Response Distributions

We assume that the reference action distribution follows $\mathbf{a}_t \sim \mathcal{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$. Within region k, the first-order approximation around the action distribution mean $\boldsymbol{\mu}_a$ gives

$$\Delta \mathbf{z}_t^k(\mathbf{a}_t) \approx \Delta \mathbf{z}_t^k(\boldsymbol{\mu}_a) + \mathbf{J}_k \cdot (\mathbf{a}_t - \boldsymbol{\mu}_a) \sim \mathcal{N}\big(\Delta \mathbf{z}_t^k(\boldsymbol{\mu}_a), \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a \cdot \mathbf{J}_k^\top\big)\,. \tag{2}$$

Here, $\mathbf{J}_k := \partial \Delta \mathbf{z}_t^k(\mathbf{a}) / \partial \mathbf{a}\big|_{\mathbf{a}=\boldsymbol{\mu}_a}$ is the Jacobian of the action response at $\boldsymbol{\mu}_a$. Assumption 2 bounds the first-order approximation error within each region.

The term $\mathbf{J}_k \cdot (\mathbf{a}_t - \boldsymbol{\mu}_a)$ is the first-order approximation to the latent-displacement difference $\Delta \mathbf{z}_t^k(\mathbf{a}_t) - \Delta \mathbf{z}_t^k(\boldsymbol{\mu}_a)$. Under the shared reference action distribution, its distribution is

$$\mathcal{P}_k := \mathcal{N}(\mathbf{0}, \mathbf{C}_k), \qquad \mathbf{C}_k := \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a \cdot \mathbf{J}_k^\top. \tag{3}$$

The covariance $\mathbf{C}_k$ captures the directions and strength of the action-induced latent variations, while $e_k := \mathrm{tr}(\mathbf{C}_k)$ measures their total variance. Assumption 3 guarantees that the effect of dynamics information omitted from this description is bounded.

Region-specific modeling is warranted when the dynamics differ sufficiently across regions. To quantify these differences, we define the response distribution distance by

$$d_{ij}^2 := \min_{\mathbf{Q} \in \mathrm{O}(d)} \left\| \mathbf{C}_i^{1/2} - \mathbf{C}_j^{1/2} \cdot \mathbf{Q} \right\|_F^2, \tag{4}$$

Here, $\mathrm{O}(d)$ denotes the set of $d \times d$ orthogonal matrices. Each covariance factor transforms a standard Gaussian variable into the corresponding regional response distribution. Since orthogonal transformations preserve the standard Gaussian distribution, minimizing over $\mathbf{Q}$ removes the ambiguity in the factor representation.

Theorem 2 (Wasserstein representation of the response distribution distance)**.** *For the centered Gaussian response distributions $\mathcal{P}_i$ and $\mathcal{P}_j$, the following identity holds:*

$$d_{ij}^2 = W_2^2(\mathcal{P}_i, \mathcal{P}_j). \tag{5}$$

This identity relates the response distribution distance to the 2-Wasserstein distance W_2 between the regional responses. For centered Gaussian distributions, its derivation is given in Appendix G.2.

Now that we have quantified the pairwise response-distribution difference, we consider the full partition. We seek a conservative measure that is large only when every pair is sufficiently different. To this end, we define the *Worst-Case Partition Gain* criterion as

$$b_{ij} = \frac{d_{ij}^2}{e_i + e_j} = \frac{W_2^2(\mathcal{P}_i, \mathcal{P}_j)}{e_i + e_j}, \qquad b(\Pi) = \min_{i<j} b_{ij}. \tag{6}$$

For $e_i + e_j > 0$, we have $0 \leq b_{ij} \leq 1$. The score b_{ij} compares the squared Wasserstein distance between the regional response distributions with their total response energy. The minimum identifies the pair with the smallest distributional difference.

This worst-case distributional difference provides a lower bound on the aggregate geometric difference in the affine comparison. To determine whether separate modeling is preferable overall, we also account for the gap between the minimum and aggregate differences, together with the additional fitting cost of training `Regional` models on fewer samples each. Assumption 4 guarantees that these two effects are bounded. The aggregation gap is small when pairwise benefits are comparable; when they are not, the lower bound remains useful for a one-sided guarantee, as shown in Appendix G.3.

3.3 From Response Distribution Distance to Dynamics Selection

Better one-step predictions can improve planning when their advantage persists over multi-step rollouts and affects task success. We compare success probabilities $p_{\mathrm{R},H}$ and $p_{\mathrm{G},H}$ over a fixed execution horizon H, averaged over the same initial-state distribution. The shared encoder and planner make this comparison meaningful.

Under Assumption 5, we have that models with lower prediction error are more likely to plan successfully. This conclusion holds after accounting for bounded effects from multi-step error propagation and task relevance. Combining this connection with the response-distribution comparison above yields the selection result below. The quantitative decomposition and bounds are provided in Appendix G.5.

Theorem 3 (Dynamics selection with a controlled margin)**.** *Under the shared comparison setting and Assumptions 2–5, let $\tau_\star$ be the theoretical threshold and $\epsilon(c)$ the total deviation bound in score units for configuration c. Whenever $|b(\Pi) - \tau_\star| > \epsilon(c)$,*

$$p_{\mathrm{R},H} > p_{\mathrm{G},H} \quad \Longleftrightarrow \quad b(\Pi) > \tau_\star. \tag{7}$$

The theoretical threshold $\tau_\star$ represents the cost that specialization needs to overcome. The margin requires its benefit to be clear enough to withstand the remaining modeling and planning effects. This threshold is shared within the stated comparison class; the required margin may vary. Theorem 4 in Appendix G.6 shows that selection remains correct when the total deviation is smaller than the margin to $\tau_\star$. Appendix G.6 also gives an exact zero-remainder case and a one-sided `Regional` guarantee under weaker conditions.

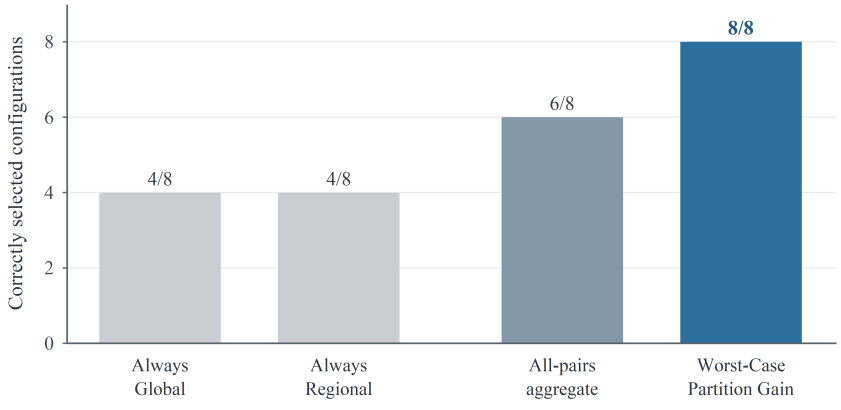


Figure 2: **Worst-case ablation.** We reuse the experimental protocol from Section 4.2 to test whether the all-pairs aggregate and Worst-Case Partition Gain select the better dynamics branch across eight configurations, compared with always selecting the same branch. Details of the ablation and the all-pairs aggregate are provided in Appendix E.2.

3.4 Finite-Sample Dynamics Selection

The following Algorithm 1 summarizes the finite-sample selection procedure.

Algorithm 1: Finite-Sample Dynamics Selection

Input: encoder h of a latent world model; training transitions $\mathcal{D} = \{(\mathbf{s}_n, \mathbf{a}_n, \mathbf{s}'_n)\}_{n=1}^{N}$; a candidate K-region partitioning method; partition seeds $\mathcal{S}$; ridge coefficient λ; threshold τ_{cal}.

1. **Encode transitions.** Encode the training transitions: $\mathbf{z}_n = h(\mathbf{s}_n)$, $\mathbf{z}'_n = h(\mathbf{s}'_n)$, and $\Delta\mathbf{z}_n = \mathbf{z}'_n - \mathbf{z}_n$.
2. **Assign regions.** For each $\ell \in \mathcal{S}$, obtain a candidate K-region partition and assign each transition a region label $r_n^{(\ell)} \in \{1, \dots, K\}$.
3. **Set the common action reference.** Compute the common action mean $\widehat{\boldsymbol{\mu}}_a$ and covariance $\widehat{\boldsymbol{\Sigma}}_a$ from the full training dataset.
4. **Estimate `regional` responses.** For each partition seed ℓ and region k, fit an affine map from centered actions $\mathbf{a}_n - \widehat{\boldsymbol{\mu}}_a$ to latent displacements $\Delta\mathbf{z}_n$ using ridge regression on that region's transitions. Let $\widehat{\mathbf{J}}_k^{(\ell)}$ be the fitted slope matrix, and construct
$$\widehat{\mathbf{C}}_k^{(\ell)} = \widehat{\mathbf{J}}_k^{(\ell)} \cdot \widehat{\boldsymbol{\Sigma}}_a \cdot (\widehat{\mathbf{J}}_k^{(\ell)})^\top, \qquad \widehat{\mathcal{P}}_k^{(\ell)} := \mathcal{N}\Big(\mathbf{0}, \widehat{\mathbf{C}}_k^{(\ell)}\Big).$$
5. **Compare region pairs.** For every region pair $i < j$, compute the normalized squared 2-Wasserstein score $\widehat{b}_{ij}^{(\ell)}$ between $\widehat{\mathcal{P}}_i^{(\ell)}$ and $\widehat{\mathcal{P}}_j^{(\ell)}$ using Equation (6).
6. **Aggregate the scores.** Compute the Worst-Case Partition Gain score by taking the minimum pairwise score within each partition and averaging across partition seeds:
$$\widehat{b}_K = \frac{1}{|\mathcal{S}|} \cdot \sum_{\ell \in \mathcal{S}} \min_{i<j} \widehat{b}_{ij}^{(\ell)}.$$
7. **Select the dynamics model.** Return `Regional` if $\widehat{b}_K > \tau_{\text{cal}}$; otherwise return `Global`.

4 Experiments on Regional-versus-Global Dynamics Modeling

The theory establishes a shared theoretical threshold $\tau_\star$ under reasonable assumptions, and the finite-sample rule uses the calibrated threshold τ_{cal}. In the following, we evaluate the dynamics-selection rule in Section 3.4 under matched `Regional`-versus-`Global` comparisons. Specifically, we investigate the following research questions:

- Does the calibrated threshold remain effective across different numbers of regions?
- Does the same threshold generalize across world-model backbones?
- How does Worst-Case Partition Gain compare with alternative selection methods on the full test benchmark?

4.1 Experimental Setup

For each task, we freeze the encoder of the pretrained model and post-train only the predictor. The `Regional` branch divides the data into regions and post-trains a separate predictor for each region, whereas the `Global` branch post-trains one predictor on all the data. Both branches share all other training and evaluation settings. For each experimental configuration, defined by a task, a world-model backbone, and a region count K, the branch with the higher mean 50-step long-horizon success rate over three post-training seeds provides the fixed point-estimate winner label used to evaluate every selection method. Long-horizon planning generally passes through more regions than short-horizon planning. We use five evaluation sampling seeds for the 50-step tasks; all other training and evaluation settings follow LeWM (Maes et al., 2026). Appendix Table E.6 summarizes the complete recipe.

We use four datasets—TwoRoom (Sobal et al., 2025), PushT (Chi et al., 2023), Reacher (Tassa et al., 2018), and OGBench–Cube (Cube) (Park et al., 2025)—that are commonly used with JEPA and its variants (Maes et al., 2026). They cover different types of planning tasks. To test whether the theory applies across models, we use LeWM and Sub-JEPA, which have different latent-space

Table 1: **Accuracy of all frozen dynamics-selection methods.** Accuracy is the fraction of configurations in which a method selects the better branch. "Calib." denotes calibration.

Method type	Method	Calib.	Region-count transfer			Backbone transfer				Test
		$K=4$	$K=2$	$K=3$	mean	$K=2$	$K=3$	$K=4$	mean	
Spectral stability	Retained safety fraction	75.0	25.0	0.0	12.5	50.0	50.0	75.0	58.3	40.0
	Prominence ratio	75.0	50.0	**100.0**	75.0	50.0	**100.0**	**100.0**	83.3	80.0
Manifold geometry	Curvature tail	**100.0**	75.0	75.0	75.0	0.0	0.0	0.0	0.0	30.0
	Tangent contrast	**100.0**	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Geometry–response alignment	Response–centroid Spearman	50.0	0.0	50.0	25.0	0.0	25.0	50.0	25.0	25.0
	Response–boundary Spearman	75.0	0.0	**100.0**	50.0	0.0	**100.0**	75.0	58.3	55.0
Jacobian geometry	Jacobian log-scale distance	50.0	**100.0**	25.0	62.5	**100.0**	**100.0**	**100.0**	**100.0**	85.0
	Jacobian cosine distance	75.0	50.0	50.0	50.0	25.0	0.0	25.0	16.7	30.0
	Jacobian subspace chordal distance	**100.0**	75.0	75.0	75.0	**100.0**	**100.0**	75.0	91.7	85.0
Regional response	Affine response contrast	**100.0**	25.0	50.0	37.5	25.0	0.0	0.0	8.3	20.0
	Minimum pairwise response	**100.0**	25.0	50.0	37.5	0.0	25.0	25.0	16.7	25.0
	Pairwise response uniformity	**100.0**	50.0	75.0	62.5	0.0	75.0	75.0	50.0	55.0
	Boundary-local Jacobian Bures	**100.0**	50.0	50.0	50.0	0.0	0.0	25.0	8.3	25.0
Temporal dynamics	Flow persistence	50.0	50.0	75.0	62.5	**100.0**	75.0	75.0	83.3	75.0
	Latent velocity eta-squared	**100.0**	25.0	50.0	37.5	50.0	25.0	25.0	33.3	35.0
	Action-residual velocity eta-squared	**100.0**	25.0	75.0	50.0	25.0	25.0	25.0	25.0	35.0
Neighborhood geometry	k-NN purity	75.0	50.0	25.0	37.5	**100.0**	50.0	50.0	66.7	55.0
	Margin–radius ratio	75.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Spectral structure	Eigengap after K	75.0	25.0	75.0	50.0	25.0	25.0	50.0	33.3	40.0
Partition structure	Normalized cluster entropy	**100.0**	25.0	25.0	25.0	75.0	25.0	25.0	41.7	35.0
	Prototype distance ratio	50.0	50.0	50.0	50.0	50.0	**100.0**	75.0	75.0	65.0
Our method	**Worst-Case Partition Gain**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	75.0	75.0	83.3	**90.0**

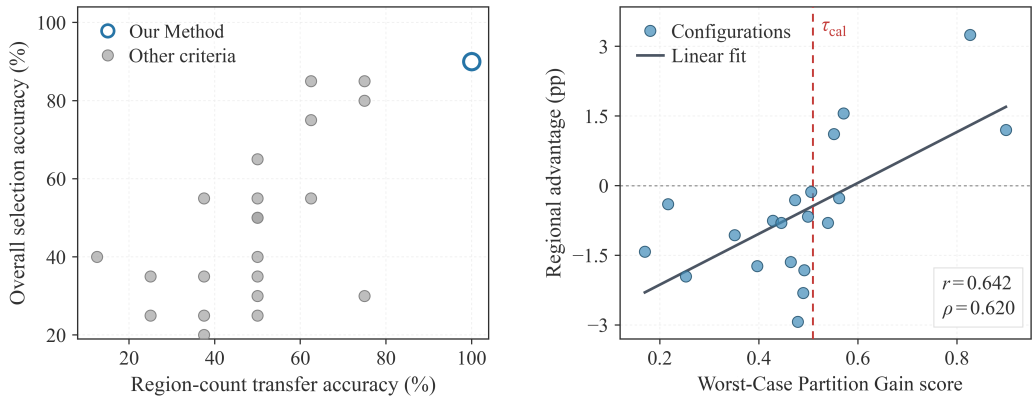
geometries. Because environment dynamics remain fixed across partition seeds, a partition intended to capture local dynamics should also remain stable. The spectral partitions achieve a mean adjusted Rand index (ARI) of 0.744 across seeds, compared with at most 0.334 for the other four methods, including a dynamics-informed method. We therefore use spectral partitions with two, three, and four regions in our main experiments. Implementation and stability details are provided in Appendices E.6 and E.1. We benchmark Worst-Case Partition Gain against 21 selection methods (Table 1). Because this is, to our knowledge, the first systematic study of dynamics selection for latent world models, no established baselines are available; we therefore construct these methods for this study from representative partition, transition, and regional-response statistics. Full method definitions and frozen policies are provided in Appendix A and Table A.1.

Calibration and evaluation protocol. To determine the thresholds and hyperparameters of different methods fairly, we calibrate all 22 methods on the four LeWM $K = 4$ configurations using exactly the same calibration procedure, detailed in Appendix B. For Worst-Case Partition Gain, this gives the calibrated threshold $\tau_{\text{cal}} \approx 0.509$ used in Section 3.4. Our method and ten other methods correctly select the better branch in all four calibration configurations. For the evaluation, we freeze each method's calibrated threshold, selection direction, and hyperparameters before examining the LeWM or Sub-JEPA test results. The eight LeWM configurations test transfer across region counts, and the twelve Sub-JEPA configurations test transfer across world-model backbones. Together, these twenty configurations form the test benchmark for the overall comparison.

Table 1 summarizes all 22 methods; Figure 3 shows the leaderboard and the correlation between the Worst-Case Partition Gain score and the `Regional−Global` planning-success difference.

4.2 Generalization across Region Counts

To examine threshold transfer across region counts, we apply the selection rules calibrated at $K = 4$ to the eight LeWM $K = 2, 3$ configurations, keeping all method-specific choices fixed. The test configurations contain four `Regional` and four `Global` winners. As shown in Figure 3(a), Worst-Case Partition Gain is the only method with 100% selection accuracy across these eight configurations, whereas the strongest alternatives reach 75%. It therefore achieves state-of-the-



(a) Criterion leaderboard **(b) Score–advantage correlation**

Figure 3: **Method leaderboard and score–advantage correlation.** **(a)** Selection accuracy of 22 methods on region-count transfer and the full test set. **(b)** Worst-Case Partition Gain score versus `Regional–Global` planning success difference on the 20 test configurations. red dashed: calibration threshold τ_{cal}. Pearson r and Spearman ρ use rounded values from Appendix Tables C.1 and D.1.

art selection performance on the region-count transfer benchmark. Its correct selection of both `Regional` and `Global` winners agrees with the necessary-and-sufficient condition in Theorem 3 and the shared threshold across region counts in Corollary 3. Replacing weakest-pair aggregation with an all-pairs aggregate reduces selection accuracy from 100% to 75% after separately calibrating both rules with the same $K = 4$ protocol (Appendix E.2). Complete results are provided in Appendix C.

4.3 Generalization across World-Model Backbones

We compute the method scores using the Sub-JEPA representation and apply the directions and thresholds calibrated on LeWM, with all method hyperparameters fixed. This comparison covers the four tasks at $K = 2, 3, 4$, giving twelve test configurations. Under the matched long-horizon evaluation, all twelve configurations favor `Global`. Worst-Case Partition Gain correctly identifies ten of these configurations; its two false `Regional` predictions occur on TwoRoom at $K = 3$ and $K = 4$, where `Global` leads by 0.27 and 0.80 percentage points, respectively. In Figure 3(b), these two configurations lie to the right of τ_{cal} but below zero advantage. Jacobian log-scale distance reaches 100% selection accuracy, while Jacobian subspace chordal distance (Appendix A, Equation 25) reaches 91.67%. The 83.33% selection accuracy of our method across these twelve configurations provides empirical support for threshold transfer across world-model backbones, as motivated by Theorem 3. Complete Sub-JEPA results are given in Appendix D. Figure 4 illustrates dynamics selection and paired planning executions on Cube with $K = 3$ for LeWM and Sub-JEPA.

4.4 Overall Selection Performance

Finally, we compare all 22 methods on the 20 test configurations. We report both selection accuracy and balanced accuracy, where balanced accuracy is the mean of the fractions of `Regional` and `Global` winners correctly identified. As summarized in Table 1 and Figure 3(a), Worst-Case Partition Gain achieves 90% selection accuracy. Jacobian log-scale and subspace chordal distance each obtain 85% accuracy, while the Always-`Global` rule obtains 80%. Our method also recovers all four `Regional` winners and 14 of 16 `Global` winners, yielding 93.75% balanced accuracy; prominence ratio has the strongest balanced accuracy among the other methods at 87.5%. It therefore achieves state-of-the-art selection performance on this benchmark, with the highest accuracy and balanced accuracy among the 22 frozen methods. The full test leaderboard and reproducibility metadata are provided in Appendix E. Table 2 compares selection performance with and without near-tie configurations (Appendix E.3).

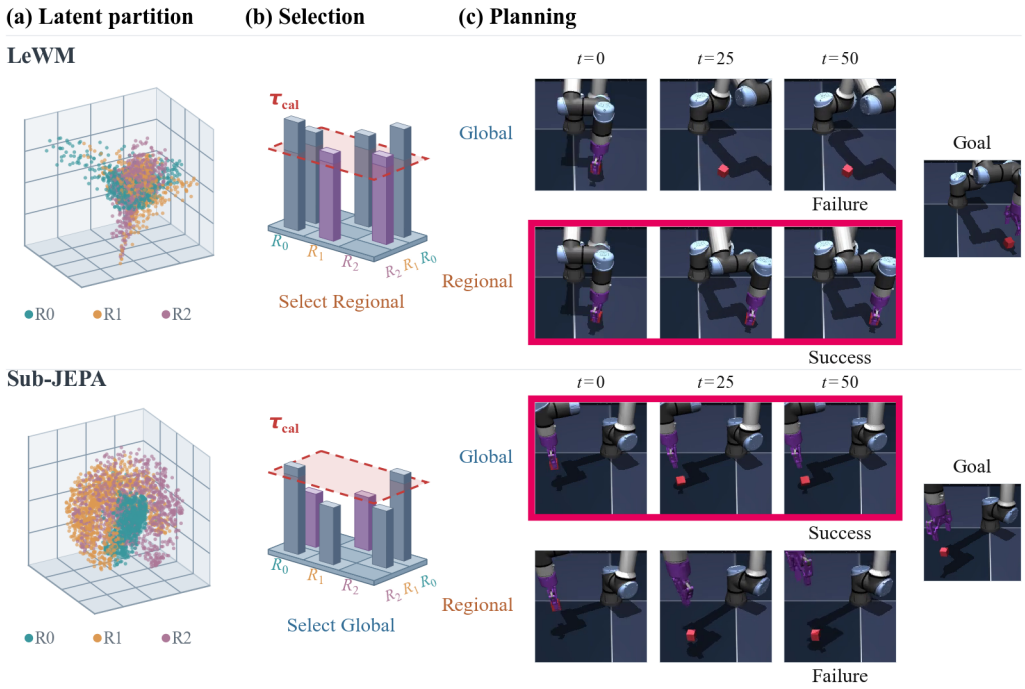


Figure 4: **Cube case visualization.** **(a)** Latent-space partition visualized by 3D PCA. **(b)** branch selection using normalized squared Wasserstein scores; purple marks the weakest pair. **(c)** Planning executions comparing branches; outlines mark the selected branch, and t indexes video frames.

Table 2: **Sensitivity to near ties.** We exclude configurations with Gap at or below each column's threshold and report selection accuracy / balanced accuracy (%). Gap is the absolute `Regional`–`Global` success-rate difference in percentage points.

Method	**Gap > 0.0**	**Gap > 0.4**	**Gap > 1.0**
Jacobian log-scale	85.00 / 71.88	81.25 / 70.83	83.33 / 75.00
Subspace chordal	85.00 / 71.88	87.50 / 75.00	83.33 / 75.00
Prominence ratio	80.00 / 87.50	75.00 / 83.33	83.33 / 87.50
Our Method	**90.00 / 93.75**	**93.75 / 95.83**	**100.00 / 100.00**

5 Conclusion

In this work, we formulate the choice between `regional` and `global` dynamics as a dynamics-selection problem for latent world models. To address this problem, we introduce Worst-Case Partition Gain, which captures the weakest distance between regional response distributions. Using this criterion, we show that, under reasonable assumptions, `Regional` achieves higher long-horizon planning success exactly when its score exceeds a theoretical threshold $\tau_\star$ shared across modeling configurations. This characterization yields a finite-sample algorithm that uses latent transitions to select between `Regional` and `Global` dynamics before training. We evaluate its effectiveness and generalizability through a systematic experimental framework comparing 22 methods while holding τ_{cal} fixed. Worst-Case Partition Gain selects the better dynamics in all eight region-count transfer configurations and in 18 of 20 test configurations overall, achieving the highest selection accuracy and balanced accuracy among the evaluated methods. These results also show that the two dynamics modeling approaches should be selected rather than adopted by default.

Reproducibility Statement

An anonymized supplementary code archive accompanies this submission. It contains the dynamics-selection implementation, experiment and evaluation drivers, configurations, and scripts needed to reproduce the reported results, together with a reusable dynamics-selection interface for adapting the method to other latent world models. All pretrained models and datasets used in our experiments were downloaded from their official project websites or release pages. Section 3.4 specifies the algorithm and deployment procedure, and Section 4 gives the main evaluation protocol. Appendix A defines the methods and frozen policies; Appendices B–E report calibration, transfer, robustness, implementation, and provenance; Appendix F studies manually defined regions; and Appendix G provides the assumptions and proofs.

AI Use Statement

ChatGPT assisted with reviewing and revising the authors' existing theoretical framework, assumptions, theorem/proposition statements, derivations, and proofs; reviewing the existing method design; and implementing, reviewing, refactoring, and standardizing dynamics-selection and experimental code and reproducibility interfaces. It was not used for data generation, cleaning, splitting, preprocessing, or result interpretation. The authors independently verified all AI-assisted text, theory, and code and take full responsibility for the final paper.

References

Shahbaz Abdul Khader, Hang Yin, Pietro Falco, and Danica Kragic. Data-efficient model learning and prediction for contact-rich manipulation tasks. *IEEE Robotics and Automation Letters*, 5(3): 4321–4328, 2020. doi: 10.1109/LRA.2020.2996067.

Junhyeok Ahn and Luis Sentis. Nested mixture of experts: Cooperative and competitive learning of hybrid dynamical system. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pages 779–790. PMLR, 2021.

David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.

Alberto Bemporad. A piecewise linear regression and classification algorithm with application to learning and model predictive control of hybrid systems. *IEEE Transactions on Automatic Control*, 68(6):3194–3209, 2023. doi: 10.1109/TAC.2022.3183036.

Juan Cerviño, Juan Andrés Bazerque, Miguel Calvo-Fullana, and Alejandro Ribeiro. Multi-task bias-variance trade-off through functional constraints. In *2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1–5. IEEE, 2023. doi: 10.1109/ICASSP49357.2023.10094811.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Robotics: Science and Systems*, 2023.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 2555–2565. PMLR, 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, Hao Su, and Xiaolong Wang. TD-MPC2: Scalable, robust world models for continuous control. In *International Conference on Learning Representations*, 2024.

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

Anthony Hu, Gianluca Corrado, Nicolas Griffiths, Zachary 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*, volume 35, pages 20703–20716, 2022. doi: 10.52202/068431-1505.

Jiaming Hu, Yan Zheng, and Tian Wang. SCALE: State-calibrated latent embeddings for JEPA planning in the right geometry. *arXiv preprint arXiv:2608.16287*, 2026.

Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2:193–218, 1985. doi: 10.1007/BF01908075.

David Klindt, Yann LeCun, and Randall Balestriero. When does LeJEPA learn a world model? *arXiv preprint arXiv:2605.26379*, 2026.

Yann LeCun. A path towards autonomous machine intelligence. *OpenReview*, 2022. Version 0.9.2.

Anson Lei, Bernhard Schölkopf, and Ingmar Posner. Variational causal dynamics: Discovering modular world models from interventions. *Transactions on Machine Learning Research*, 2023.

Anson Lei, Frederik Nolte, Bernhard Schölkopf, and Ingmar Posner. Compete and compose: Learning independent mechanisms for modular world models. *arXiv preprint arXiv:2404.15109*, 2024.

Mingwei Li, Xiaoyuan Zhang, Chengwei Yang, Zilong Zheng, and Yaodong Yang. Prismatic world model: Learning compositional dynamics for planning in hybrid systems. *arXiv preprint arXiv:2512.08411*, 2025.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Lucas Maes, Quentin Le Lidec, Damien Scieur, Yann LeCun, and Randall Balestriero. LeWorldModel: Stable end-to-end joint-embedding predictive architecture from pixels. *arXiv preprint arXiv:2603.19312*, 2026.

Geoffrey J. McLachlan and S. U. Chang. Mixture modelling for cluster analysis. *Statistical Methods in Medical Research*, 13(5):347–361, 2004. doi: 10.1191/0962280204sm372ra.

Andrew Y. Ng, Michael I. Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, volume 14, 2001.

Seohong Park, Kevin Frans, Benjamin Eysenbach, and Sergey Levine. OGBench: Benchmarking offline goal-conditioned RL. In *International Conference on Learning Representations*, 2025.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588: 604–609, 2020. doi: 10.1038/s41586-020-03051-4.

Uladzislau Sobal, Wancong Zhang, Kyunghyun Cho, Randall Balestriero, Tim G. J. Rudner, and Yann LeCun. Learning from reward-free offline data: A case for planning with latent dynamics models. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-1465.

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

Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems*, volume 17, 2004.

Boxuan Zhang, Weipu Zhang, Zhaohan Feng, Wei Xiao, Jian Sun, Jie Chen, and Gang Wang. Mixture-of-world models: Scaling multi-task reinforcement learning with modular latent dynamics. In *International Conference on Learning Representations*, 2026.

Kai Zhao, Dongliang Nie, Yuchen Lin, Zhehan Luo, Yixiao Gu, Deng-Ping Fan, and Dan Zeng. Sub-JEPA: Subspace gaussian regularization for stable end-to-end world models. *arXiv preprint arXiv:2605.09241*, 2026.

Zijie Zhao, Zhongyue Zhao, Kaixuan Xu, Yuqian Fu, Jiajun Chai, Yuanheng Zhu, and Dongbin Zhao. Learning and planning multi-agent tasks via an MoE-based world model. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

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

A Method Definitions and Frozen Policies

This appendix defines the 22 candidate methods and records the complete policies frozen after calibration. No direction, threshold, or hyperparameter was changed after inspecting the region-count transfer or backbone transfer outcomes.

A.1 Method definitions

The benchmark contains Worst-Case Partition Gain and 21 comparison methods constructed for this study because no established dynamics-selection baselines are available. For each task, region count, and partition seed, let $r_n \in [K]$ be the spectral-region label assigned from the normalized latent coordinate $\mathbf{x}_n$ defined in Appendix E.6, let $\widehat{\pi}_k$ be the fraction of points in region k, and let $\Delta\mathbf{z}_n := \mathbf{z}'_n - \mathbf{z}_n$ be the one-step displacement of the frozen-encoder latent $\mathbf{z}_n$. After standardizing the action coordinates using the full transition dataset, let $\widehat{\boldsymbol{\mu}}_a$ and $\widehat{\boldsymbol{\Sigma}}_a$ denote their common mean and covariance. For readability, we suppress the partition-seed superscript (b) on region-dependent fitted quantities throughout these definitions. We fit the `regional` affine responses

$$\Delta\mathbf{z}_n \approx \widehat{\mathbf{b}}_k + \widehat{\mathbf{J}}_k \cdot (\mathbf{a}_n - \widehat{\boldsymbol{\mu}}_a) \qquad \text{when } r_n = k. \tag{8}$$

Write

$$\widehat{\mathbf{A}}_k := \widehat{\mathbf{J}}_k \cdot \widehat{\boldsymbol{\Sigma}}_a^{1/2}, \qquad \widehat{\mathbf{C}}_k := \widehat{\mathbf{A}}_k \cdot \widehat{\mathbf{A}}_k^{\top}, \qquad \widehat{\mathcal{P}}_k := \mathcal{N}(\mathbf{0}, \widehat{\mathbf{C}}_k), \qquad \widehat{e}_k := \operatorname{tr}(\widehat{\mathbf{C}}_k), \tag{9}$$

and let $\widehat{e}_{\text{res}}$ be the mean squared residual energy of one `global` affine response fit. For each pair of regions, define the normalized affine response contrast

$$\widehat{\rho}_{ij} := \frac{\|\widehat{\mathbf{b}}_i - \widehat{\mathbf{b}}_j\|_2^2 + \|(\widehat{\mathbf{J}}_i - \widehat{\mathbf{J}}_j) \cdot \widehat{\boldsymbol{\Sigma}}_a^{1/2}\|_F^2}{\widehat{e}_{\text{res}}}. \tag{10}$$

The implemented methods are:

1. **Normalized cluster entropy.** The balance of the region sizes is
$$h_{\text{norm}} := -\frac{\sum_{k=1}^{K} \widehat{\pi}_k \cdot \log \widehat{\pi}_k}{\log K}. \tag{11}$$
2. **Eigengap after K.** For the normalized Laplacian in Equation (37), the score is
$$g_K := \lambda_{K+1}(\mathbf{L}_{\text{n}}) - \lambda_K(\mathbf{L}_{\text{n}}). \tag{12}$$
3. **Prototype distance ratio.** Let $\mathbf{q}_p$ be a normalized routing prototype and ℓ_p its region label. With cosine distance $d_{\cos}(\mathbf{u}, \mathbf{v}) := 1 - \mathbf{u}^{\top} \cdot \mathbf{v}$, the score is
$$\frac{\sum_p \min_{q:\ell_q \neq \ell_p} d_{\cos}(\mathbf{q}_p, \mathbf{q}_q)}{\sum_p \min_{q \neq p:\ell_q = \ell_p} d_{\cos}(\mathbf{q}_p, \mathbf{q}_q)}. \tag{13}$$
4. **k-NN purity.** Using the 30 nearest neighbors $\mathcal{N}_{30}(n)$, the score is the average fraction that share the query label:
$$\frac{1}{30 \cdot N} \cdot \sum_{n=1}^{N} \sum_{m \in \mathcal{N}_{30}(n)} \mathbb{I}\{r_m = r_n\}. \tag{14}$$
5. **Margin–radius ratio.** For each point, the radius is its cosine distance to the 30th nearest point in the same region, and the margin is its cosine distance to the nearest point in another region. The method averages $\text{margin}_n / \max\{\text{radius}_n, 10^{-12}\}$ over points.
6. **Tangent contrast.** A local tangent basis of rank $r_{\tan} = 8$ is fitted by singular value decomposition of the offsets to 30 same-region neighbors. For two orthonormal bases $\mathbf{U}$ and $\mathbf{V}$, define
$$d_{\tan}(\mathbf{U}, \mathbf{V}) := \sqrt{\max\left\{0,\, 1 - \frac{\|\mathbf{U}^{\top} \cdot \mathbf{V}\|_F^2}{r_{\tan}}\right\}}. \tag{15}$$
The score averages the tangent distance to the nearest cross-region point minus the tangent distance to a same-region neighbor.

7. **Curvature tail.** If $s_{n,\ell}$ are the singular values of the same local-offset matrix, the local tail fraction is
$$\frac{\sum_{\ell > r_{\tan}} s_{n,\ell}^2}{\sum_{\ell} s_{n,\ell}^2}. \tag{16}$$
The score is its average over the sampled query points.

8. **Flow persistence.** With lag $\ell_{\text{flow}} = 10$, among valid within-episode pairs separated by ℓ_{flow} latent time indices, the score is
$$\mathbb{P}(r_{n+\ell_{\text{flow}}} = r_n). \tag{17}$$

9. **Latent velocity eta-squared.** Let $\overline{\Delta \mathbf{z}}_k$ and $\overline{\Delta \mathbf{z}}$ be the `regional` and `global` means of $\Delta \mathbf{z}_n$. The explained fraction of latent-velocity variation is
$$\eta_v^2 := \frac{\sum_k n_k \cdot \|\overline{\Delta \mathbf{z}}_k - \overline{\Delta \mathbf{z}}\|_2^2}{\sum_n \|\Delta \mathbf{z}_n - \overline{\Delta \mathbf{z}}\|_2^2}. \tag{18}$$

10. **Action-residual velocity eta-squared.** A `global` affine action-response model is first fitted to $\Delta \mathbf{z}_n$. The eta-squared formula in Equation (18) is then applied to its residual vectors, grouped by spectral region.

11. **Affine response contrast.** This score is the mean of $\widehat{\rho}_{ij}$ in Equation (10) over all region pairs.

12. **Minimum pairwise response.** This score is
$$\min_{i<j} \widehat{\rho}_{ij}. \tag{19}$$

13. **Pairwise response uniformity.** This method compares the weakest pair with the average pair:
$$\frac{\min_{i<j} \widehat{\rho}_{ij}}{\text{mean}_{i<j} \widehat{\rho}_{ij}}. \tag{20}$$

14. **Response–centroid Spearman correlation.** For every pair, let $d_{ij}^{\text{cent}} := \|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|_2$, where $\bar{\mathbf{x}}_k$ is the latent centroid of region k. The score is the Spearman correlation between $\{\widehat{\rho}_{ij}\}_{i<j}$ and $\{d_{ij}^{\text{cent}}\}_{i<j}$.

15. **Response–boundary Spearman correlation.** Define the symmetric cross-region boundary distance
$$d_{ij}^{\text{bdry}} := \frac{1}{2} \cdot \left[\frac{1}{n_i} \cdot \sum_{n:r_n=i} \min_{m:r_m=j} d_{\cos}(\mathbf{x}_n, \mathbf{x}_m) + \frac{1}{n_j} \cdot \sum_{m:r_m=j} \min_{n:r_n=i} d_{\cos}(\mathbf{x}_m, \mathbf{x}_n) \right]. \tag{21}$$
The score is the Spearman correlation between $\widehat{\rho}_{ij}$ and d_{ij}^{bdry} over region pairs. The two Spearman scores are undefined for $K = 2$, which supplies only one region pair.

16. **Boundary-local Jacobian Bures.** A point is a boundary point when at least one of its 30 nearest neighbors has another region label. `Regional` affine action-response models are refitted using transitions beginning at these boundary points. The score is the minimum normalized squared Bures distance of the resulting boundary-local response matrices.

17. **Worst-Case Partition Gain.** This method uses the weakest-pair normalized squared 2-Wasserstein distance between the centered Gaussian response distributions defined above. Its pairwise score and score for a single partition draw are
$$\widehat{b}_{ij} := \frac{W_2^2(\widehat{\mathcal{P}}_i, \widehat{\mathcal{P}}_j)}{\widehat{e}_i + \widehat{e}_j}, \qquad \min_{i<j} \widehat{b}_{ij}. \tag{22}$$

18. **Jacobian cosine distance.** This method takes the weakest pair under Frobenius cosine distance:
$$\min_{i<j} \left(1 - \frac{\langle \widehat{\mathbf{A}}_i, \widehat{\mathbf{A}}_j \rangle_F}{\|\widehat{\mathbf{A}}_i\|_F \cdot \|\widehat{\mathbf{A}}_j\|_F} \right). \tag{23}$$

19. **Jacobian log-scale distance.** This method retains only pairwise response magnitude:

$$\min_{i<j} \left| \log \frac{\|\widehat{\mathbf{A}}_i\|_F}{\|\widehat{\mathbf{A}}_j\|_F} \right|. \tag{24}$$

20. **Jacobian subspace chordal distance.** Let $\widehat{\mathbf{Q}}_k$ contain an orthonormal basis for the response subspace spanned by $\widehat{\mathbf{A}}_k$, and let r_{ij} be the smaller of the two basis ranks. The score is

$$\min_{i<j} \frac{\|\widehat{\mathbf{Q}}_i \cdot \widehat{\mathbf{Q}}_i^\top - \widehat{\mathbf{Q}}_j \cdot \widehat{\mathbf{Q}}_j^\top\|_F}{\sqrt{2 \cdot r_{ij}}}. \tag{25}$$

21. **Retained safety fraction.** For landmark draw ℓ, neighborhood size $r \in \{27, 30, 33\}$, and a small $\xi > 0$, define

$$g_{K,\ell}^{(r)} := \frac{\lambda_{K+1}(\mathbf{L}_{\mathrm{n},\ell}^{(r)}) - \lambda_K(\mathbf{L}_{\mathrm{n},\ell}^{(r)})}{\lambda_{K+1}(\mathbf{L}_{\mathrm{n},\ell}^{(r)}) + \xi}. \tag{26}$$

Let

$$\begin{aligned} \Delta g_{K,\ell} &:= \max_{r\in\{27,30,33\}} |g_{K,\ell}^{(r)} - g_{K,\ell}^{(30)}|, \\ g_{K,\min} &:= \min_{\ell} g_{K,\ell}^{(30)}, \qquad t_{K,\max}^{g} := c_{\mathrm{pert}} \cdot \max_{\ell} \Delta g_{K,\ell}, \\ r_K &:= g_{K,\min} - t_{K,\max}^{g}. \end{aligned} \tag{27}$$

The score is the retained fraction

$$s_{K,\mathrm{task}} := \frac{r_K}{g_{K,\min}}. \tag{28}$$

22. **Prominence ratio.** For each graph, the next ten relative eigengaps after the candidate gap form the background set $\mathcal{E}_{b,r}$. Its graph-level threshold and the common background threshold are

$$t_{b,r} := \mathrm{median}(\mathcal{E}_{b,r}) + 3 \cdot \mathrm{sMAD}(\mathcal{E}_{b,r}), \qquad t_{\mathrm{bg}} := \max_{b,r} t_{b,r}. \tag{29}$$

The score is the prominence ratio

$$\frac{r_K}{t_{\mathrm{bg}}}. \tag{30}$$

Each task-level score is averaged over the prespecified partition seeds $\{0, 1, 2\}$. The calibration protocol uses LeWM at four regions to fix one orientation and one threshold for every method. Before fitting, the common practical indifference rule in Appendix B removes calibration tasks whose observed branch-gap magnitude is below 0.4 percentage points. The same automatic fitting routine is then applied to all 22 methods, and all resulting policies are frozen for the evaluations.

Table A.1: The 22 frozen selection policies. "High" means that a score above the threshold predicts `Regional`; "Low" reverses the inequality. Every policy is fitted on the same retained subset of the LeWM $K = 4$ calibration configurations.

#	Method	Category	Direction	Threshold	Main hyperparameters	Threshold source
1	Normalized cluster entropy	Partition structure	Low	0.9971145812	None	Calibration
2	Eigengap after K	Spectral structure	High	0.0006008125	Self-tuned 30-NN graph	Calibration
3	Prototype distance ratio	Partition structure	Low	1.0793239934	16 prototypes/region	Calibration
4	k-NN purity	Neighborhood geometry	Low	0.9621138889	30-NN	Calibration
5	Margin–radius ratio	Neighborhood geometry	Low	6.2712070545	30-NN	Calibration
6	Tangent contrast	Manifold geometry	Low	0.4694255837	30-NN; $r_{\tan} = 8$; 128 queries/region	Calibration
7	Curvature tail	Manifold geometry	High	0.0003490457	30-NN; $r_{\tan} = 8$; 128 queries/region	Calibration
8	Flow persistence	Temporal dynamics	Low	0.8095587071	$\ell_{\mathrm{flow}} = 10$	Calibration
9	Latent velocity eta-squared	Temporal dynamics	High	0.0001317047	One-step transitions	Calibration
10	Action-residual velocity eta-squared	Temporal dynamics	High	0.0001342343	Standardized actions; ridge 10^{-8}	Calibration
11	Affine response contrast	`Regional` response	High	0.0162786001	Standardized actions; ridge 10^{-8}	Calibration
12	Minimum pairwise response	`Regional` response	High	0.0133253772	Standardized actions; ridge 10^{-8}	Calibration
13	Pairwise response uniformity	`Regional` response	High	0.7991487261	Standardized actions; ridge 10^{-8}	Calibration
14	Response–centroid Spearman	Geometry–response alignment	High	0.1428571429	Undefined for $K = 2$	Calibration
15	Response–boundary Spearman	Geometry–response alignment	Low	0.0380952381	30-NN boundary; undefined for $K = 2$	Calibration
16	Boundary-local Jacobian Bures Worst-Case	`Regional` response	Low	0.6010348681	30-NN boundary; ridge 10^{-8}	Calibration
17	Partition Gain	Jacobian geometry	High	0.5089478543	Standardized actions; ridge 10^{-8}	Calibration
18	Jacobian cosine distance	Jacobian geometry	Low	0.8265036167	Standardized actions; ridge 10^{-8}	Calibration
19	Jacobian log-scale distance	Jacobian geometry	Low	0.0125081574	Standardized actions; ridge 10^{-8}	Calibration
20	Jacobian subspace chordal distance	Jacobian geometry	High	0.8675955672	Standardized actions; ridge 10^{-8}	Calibration
21	Retained safety fraction	Spectral stability	Low	0.8615306405	20k landmarks; seeds 0–2; $r = 27, 30, 33$	Calibration
22	Prominence ratio	Spectral stability	High	0.2128256289	10 background gaps; $3 \cdot$ sMAD	Calibration

For methods 10–20, the affine fits use standardized actions, one-step transitions, and ridge coefficient 10^{-8}. The spectral-stability methods use 20,000 landmarks, landmark/partition seeds $\{0, 1, 2\}$, neighborhood sizes $\{27, 30, 33\}$, perturbation multiplier $c_{\mathrm{pert}} = 2$, ten background eigengaps, and numerical floor 10^{-8}.

B CALIBRATION RESULTS

The calibration protocol uses only the four LeWM $K = 4$ configurations. Table B.1 reports the calibration labels from the mean 50-step success rates. The same indifference filter and automatic policy-fitting routine described below are applied to every method. Every resulting policy is then frozen. Twelve of the 22 methods fit all three retained calibration labels, so this small calibration set cannot establish which method will generalize. No calibration result enters the reported test accuracy.

Table B.1: LeWM $K = 4$ calibration branch means (%). The difference is `Regional` minus `Global` and the winner uses the unrounded means.

Task	`Global`	`Regional`	Difference	Winner
TwoRoom	58.1333	59.5556	1.4222	`Regional`
PushT	40.4000	40.7556	0.3556	`Regional`
Reacher	78.6667	77.6000	-1.0667	`Global`
OGBench–Cube	48.1333	50.4889	2.3556	`Regional`

Each post-training replicate averages $5 \cdot 50 = 250$ binary evaluations, so one additional success changes its success rate by exactly $100/250 = 0.4$ percentage points. We therefore treat absolute `Regional`–`Global` success-rate gaps below 0.4 percentage points as ties for calibration. This sets the practical tolerance $\delta_{\text{cal}} = 0.4$ percentage points; tied cases are excluded from fitting the direction and threshold of every method. PushT has an observed gap of only $+0.3556$ percentage points and is therefore the only calibration task excluded. Define the common fitting set

$$\mathcal{T}_{\text{fit}} := \left\{ t : \left| \widehat{\Delta p}_t \right| \geq \delta_{\text{cal}} \right\}. \tag{31}$$

PushT remains reported in the calibration tables but does not influence any fitted direction or threshold. No test configuration is removed from the headline region-count transfer or backbone transfer accuracy.

For method m, let $s_{m,t}$ be its partition-seed-averaged score on task $t \in \mathcal{T}_{\text{fit}}$, and let y_t be the observed `Regional`-or-`Global` label. If $u_1 < \cdots < u_M$ are the unique retained scores, define

$$r_m := \max\{u_M - u_1, 1\}, \qquad \mathcal{C}_m := \{u_1 - r_m\} \cup \left\{ \frac{u_j + u_{j+1}}{2} : 1 \leq j < M \right\} \cup \{u_M + r_m\}. \tag{32}$$

The common `fit` routine tests every $\tau \in \mathcal{C}_m$ in both directions: "High" predicts `Regional` when $s_{m,t} > \tau$, while "Low" predicts `Regional` when $s_{m,t} < \tau$. It selects the pair (d_m, τ_m) that lexicographically maximizes balanced accuracy, overall accuracy, the minimum normalized margin $\min_t |s_{m,t} - \tau|/r_m$, preference for the High direction, and finally preference for the smaller threshold. This deterministic rule is applied unchanged to all 22 methods. The common indifference filter and fitting routine were fixed without examining the results.

Table B.2: Raw calibration scores, methods 1–11. Values are means over partition seeds $\{0, 1, 2\}$.

#	Method	Direction	TwoRoom	PushT	Reacher	Cube
1	Normalized cluster entropy	Low	0.994542	0.976834	0.999687	0.979244
2	Eigengap after K	High	0.0000757	0.0015123	0.0001173	0.0010844
3	Prototype distance ratio	Low	1.065924	1.115630	1.092724	1.104009
4	k-NN purity	Low	0.983936	0.931730	0.969977	0.954251
5	Margin–radius ratio	Low	10.072082	1.959822	9.679847	2.862567
6	Tangent contrast	Low	0.461529	0.464093	0.477322	0.454865
7	Curvature tail	High	0.0003865	0.0139766	0.0003116	0.0124351
8	Flow persistence	Low	0.748945	0.877750	0.870172	0.916363
9	Latent velocity eta-squared	High	0.0004019	0.0031785	0.0000605	0.0002029
10	Action-residual eta-squared	High	0.0004375	0.0032274	0.0000605	0.0002080
11	Affine response contrast	High	0.017064	0.060787	0.015493	0.084582

Table B.3: Raw calibration scores, methods 12–22.

#	Method	Direction	TwoRoom	PushT	Reacher	Cube
12	Minimum pairwise response	High	0.015265	0.053902	0.011386	0.073011
13	Pairwise response uniformity	High	0.894541	0.886556	0.734804	0.863493
14	Response–centroid Spearman	High	-0.200000	-0.066667	-0.161905	0.447619
15	Response–boundary Spearman	Low	-0.638095	0.180952	0.714286	-0.695238
16	Boundary-local Jacobian Bures Worst-Case	Low	0.482162	0.400562	0.605612	0.596457
17	Partition Gain	High	0.698369	0.534245	0.382184	0.635712
18	Jacobian cosine distance	Low	0.937769	0.528913	0.898225	0.754782
19	Jacobian log-scale distance	Low	0.008284	0.017682	0.016732	0.021582
20	Jacobian subspace chordal distance	High	0.925106	0.869419	0.852527	0.882664
21	Retained safety fraction	Low	0.820680	0.466828	0.902381	0.959940
22	Prominence ratio	High	0.12168	0.33013	0.13369	0.29196

Table B.4: Calibration Worst-Case Partition Gain scores at the reference ridge coefficient 10^{-8}, together with the single-hyperparameter ridge sensitivity. The last column is the largest absolute change in the task mean over $\{10^{-10}, 10^{-8}, 10^{-6}, 10^{-4}\}$ relative to 10^{-8}.

Task	Seed 0	Seed 1	Seed 2	Mean	Margin to threshold	Max. change
TwoRoom	0.693987	0.696660	0.704459	0.698369	0.189421	$3.25 \cdot 10^{-13}$
PushT	0.517787	0.525483	0.559465	0.534245	0.025297	$3.86 \cdot 10^{-13}$
Reacher	0.381838	0.369421	0.395294	0.382184	-0.126764	$1.61 \cdot 10^{-13}$
Cube	0.600306	0.654861	0.651968	0.635712	0.126764	$4.40 \cdot 10^{-12}$

The partition-seed aggregation used throughout the benchmark is an arithmetic mean over seeds $\{0, 1, 2\}$; policies are applied to that mean rather than to individual seed selections. Table B.4 also shows that the frozen Worst-Case Partition Gain selections are insensitive to the ridge coefficient at the numerical precision of this audit.

C CROSS-K RESULTS

Table C.1 reports every LeWM selection after freezing the policies at $K = 4$. The true winner is determined by the unrounded mean 50-step success rate. Worst-Case Partition Gain is correct in all eight configurations, including four `Regional` and four `Global` winners.

Table C.1: Per-configuration Worst-Case Partition Gain selections for region-count transfer. G and R denote `Global` and `Regional`.

Task	K	Score	`Global` (%)	`Regional` (%)	$R - G$	Prediction	Winner
TwoRoom	2	0.898943	58.133	59.333	1.200	R	R
PushT	2	0.498970	40.400	39.733	-0.667	G	G
Reacher	2	0.169779	78.667	77.244	-1.422	G	G
Cube	2	0.551394	48.133	49.244	1.111	R	R
TwoRoom	3	0.826633	58.133	61.378	3.244	R	R
PushT	3	0.491828	40.400	38.578	-1.822	G	G
Reacher	3	0.428340	78.667	77.911	-0.756	G	G
Cube	3	0.571179	48.133	49.689	1.556	R	R

Table C.2: Complete region-count transfer prediction matrix. Columns combine task initial and region count. A denotes an abstention because a Spearman score is undefined with only one region pair. The true labels are shown last.

#	Method	Dir.	T2	P2	R2	C2	T3	P3	R3	C3
1	Normalized cluster entropy	L	G	R	G	G	G	R	R	R
2	Eigengap after K	H	G	R	R	R	R	R	G	R
3	Prototype distance ratio	L	R	R	R	R	R	R	R	R
4	k-NN purity	L	G	G	G	G	G	R	G	G
5	Margin–radius ratio	L	G	R	G	R	G	R	G	R
6	Tangent contrast	L	G	G	G	G	G	G	G	G
7	Curvature tail	H	R	R	G	R	R	R	G	R
8	Flow persistence	L	G	G	G	G	R	G	G	G
9	Latent velocity eta-squared	H	G	R	G	G	R	R	G	G
10	Action-residual eta-squared	H	G	R	G	G	R	R	G	R
11	Affine response contrast	H	G	R	R	R	G	R	G	R
12	Minimum pairwise response	H	G	R	R	R	G	R	G	R
13	Pairwise response uniformity	H	R	R	R	R	R	R	G	R
14	Response–centroid Spearman	H	A	A	A	A	G	R	G	R
15	Response–boundary Spearman	L	A	A	A	A	R	G	G	R
16	Boundary-local Jacobian Bures Worst-Case	L	R	R	R	R	R	R	R	R
17	Partition Gain	H	R	G	G	R	R	G	G	R
18	Jacobian cosine distance	L	G	R	G	R	G	R	G	R
19	Jacobian log-scale distance	L	R	G	G	R	G	G	R	G
20	Jacobian subspace chordal distance	H	R	G	G	G	R	G	G	G
21	Retained safety fraction	L	G	R	G	G	G	R	R	G
22	Prominence ratio	H	R	R	R	R	R	G	G	R
		True winner	R	G	G	R	R	G	G	R

Table C.3: Worst-Case Partition Gain scores for each partition seed in region-count transfer. Their arithmetic mean is the task-level score used for the selection.

Task	K	Seed 0	Seed 1	Seed 2	Mean
TwoRoom	2	0.897324	0.901657	0.897848	0.898943
PushT	2	0.497525	0.504284	0.495101	0.498970
Reacher	2	0.170555	0.171471	0.167312	0.169779
Cube	2	0.550234	0.550511	0.553437	0.551394
TwoRoom	3	0.822026	0.825331	0.832543	0.826633
PushT	3	0.509323	0.507277	0.458883	0.491828
Reacher	3	0.438167	0.415536	0.431317	0.428340
Cube	3	0.574525	0.569388	0.569624	0.571179

D CROSS-BACKBONE RESULTS

The complete frozen LeWM policy is transferred to Sub-JEPA. All 12 unrounded branch comparisons favor `Global`. Table D.1 shows that Worst-Case Partition Gain identifies ten; its only errors are TwoRoom at $K = 3$ and $K = 4$, where the `Global` advantages are 0.267 and 0.800 percentage points.

Table D.1: Per-configuration Worst-Case Partition Gain selections for backbone transfer to Sub-JEPA.

Task	K	Score	`Global` (%)	`Regional` (%)	$R - G$	Prediction	Winner
TwoRoom	2	0.505810	58.533	58.400	-0.133	G	G
PushT	2	0.464565	46.400	44.756	-1.644	G	G
Reacher	2	0.252717	78.267	76.311	-1.956	G	G
Cube	2	0.445466	49.067	48.267	-0.800	G	G
TwoRoom	3	0.562038	58.533	58.267	-0.267	R	G
PushT	3	0.489578	46.400	44.089	-2.311	G	G
Reacher	3	0.396877	78.267	76.533	-1.733	G	G
Cube	3	0.351332	49.067	48.000	-1.067	G	G
TwoRoom	4	0.539334	58.533	57.733	-0.800	R	G
PushT	4	0.478723	46.400	43.467	-2.933	G	G
Reacher	4	0.473184	78.267	77.956	-0.311	G	G
Cube	4	0.216820	49.067	48.667	-0.400	G	G

Table D.2: Complete backbone transfer prediction matrix. All true winners are `Global`; A denotes an undefined two-region Spearman score.

#	Method	Dir.	T2	P2	R2	C2	T3	P3	R3	C3	T4	P4	R4	C4
1	Normalized cluster entropy	L	G	R	G	G	G	R	R	R	R	R	G	R
2	Eigengap after K	H	G	R	R	R	R	R	G	R	G	R	G	R
3	Prototype distance ratio	L	G	R	R	G	G	G	G	G	G	R	G	G
4	k-NN purity	L	G	G	G	G	G	R	G	R	G	R	G	R
5	Margin–radius ratio	L	G	R	G	R	G	R	G	R	G	R	G	R
6	Tangent contrast	L	R	G	G	R	R	G	G	R	R	G	G	R
7	Curvature tail	H	R	R	R	R	R	R	R	R	R	R	R	R
8	Flow persistence	L	G	G	G	G	R	G	G	G	R	G	G	G
9	Latent velocity eta-squared	H	G	R	G	R	R	R	G	R	R	R	G	R
10	Action-residual eta-squared	H	R	R	G	R	R	R	G	R	R	R	G	R
11	Affine response contrast	H	R	R	G	R	R	R	R	R	R	R	R	R
12	Minimum pairwise response	H	R	R	R	R	R	R	G	R	R	R	G	R
13	Pairwise response uniformity	H	R	R	R	R	G	R	G	G	G	R	G	G
14	Response–centroid Spearman	H	A	A	A	A	R	R	G	R	R	G	R	G
15	Response–boundary Spearman	L	A	A	A	A	G	G	G	G	G	R	G	G
16	Boundary-local Jacobian Bures Worst-Case	L	R	R	R	R	R	R	R	R	R	R	G	R
17	Partition Gain	H	G	G	G	G	R	G	G	G	R	G	G	G
18	Jacobian cosine distance	L	R	R	G	R	R	R	R	R	R	R	G	R
19	Jacobian log-scale distance	L	G	G	G	G	G	G	G	G	G	G	G	G
20	Jacobian subspace chordal distance	H	G	G	G	G	G	G	G	G	G	G	R	G
21	Retained safety fraction	L	R	R	G	G	G	R	R	G	G	R	G	G
22	Prominence ratio	H	G	G	R	R	G	G	G	G	G	G	G	G

Under the same frozen policies, Jacobian log-scale distance predicts `Global` in all 12 backbone transfer configurations and therefore obtains $12/12$; Jacobian subspace chordal distance obtains $11/12$. These results are reported alongside the Worst-Case Partition Gain result rather than used to revise any calibration policy.

Table D.3: Worst-Case Partition Gain scores by partition seed in backbone transfer. All experiments use training seeds $\{0, 42, 625\}$, evaluation seeds $\{0, 1, 2, 3, 4\}$, and the partition seeds shown here.

Task	K	Seed 0	Seed 1	Seed 2	Mean
TwoRoom	2	0.507245	0.505274	0.504911	0.505810
TwoRoom	3	0.553189	0.554623	0.578302	0.562038
TwoRoom	4	0.533532	0.538782	0.545687	0.539334
PushT	2	0.469741	0.463006	0.460949	0.464565
PushT	3	0.532954	0.459529	0.476252	0.489578
PushT	4	0.546004	0.439841	0.450324	0.478723
Reacher	2	0.251057	0.254588	0.252507	0.252717
Reacher	3	0.388813	0.407025	0.394793	0.396877
Reacher	4	0.464106	0.482983	0.472462	0.473184
Cube	2	0.441547	0.445994	0.448859	0.445466
Cube	3	0.351053	0.349359	0.353584	0.351332
Cube	4	0.227924	0.204903	0.217633	0.216820

Seed-level branch means use the same nested aggregation as the main results: evaluation seeds are averaged within each training seed, `Regional` additionally averages the three partition seeds, and the reported branch mean averages the three training seeds. Thus the seed sets and aggregation are explicit even though Table D.3 reports the diagnostic score rather than expanding the full $3 \cdot 5 \cdot 3$ result cube.

D.1 Audit of Identifiable Theory-Control Quantities

We additionally audit the theory-control quantities that can be identified from the saved LeWM and Sub-JEPA latent-transition artifacts. For each backbone, task, region count, and predeclared partition seed, the structural audit computes plug-in values of the intercept, orientation, and pair-aggregation terms r_b, r_{orient}, and r_{pair} in Equation (85). The finite-sample audit estimates $\epsilon_{q,95}$, the 95th percentile absolute deviation of $q = b(\Pi)/2$ under 500 nonparametric bootstrap replicates of 20 episode blocks. It also uses five-fold, episode-disjoint cross-fitting to obtain a one-sided, cluster-robust 95% upper estimate of the affine response residual. We report the normalized residual proxy

$$\widehat{u}_{\mathrm{aff},95} := \sqrt{\widehat{\mathrm{MSE}}_{\mathrm{aff},95}}/s_c,$$

where s_c is the response scale in Equation (85). All columns in Tables D.4 and D.5 take the maximum over the three partition seeds, except that the threshold comparison uses the mean score across those seeds. Across both backbones, the structural decomposition identity is reproduced with maximum numerical error below 2.3×10^{-16}, and none of the bootstrap runs fail.

Table D.4: **Sub-JEPA audit of identifiable theory-control quantities.** Structural terms are plug-in estimates, while $\widehat{\epsilon}_{q,95}$ is an episode-block bootstrap estimate. "Side resolved" tests only whether $|\bar{q} - \tau_{\mathrm{cal}}/2| > \max_\ell \widehat{\epsilon}^{(\ell)}_{q,95}$; it is not a certificate for the complete dynamics-selection condition.

Task	K	$\max_\ell \widehat{r}_b$	$\max_\ell \widehat{r}_{\mathrm{orient}}$	$\max_\ell \widehat{r}_{\mathrm{pair}}$	$\max_\ell \widehat{\epsilon}_{q,95}$	$\max_\ell \widehat{u}_{\mathrm{aff},95}$	Side resolved
TwoRoom	2	0.0063	0.0139	0.0000	0.0046	4.429	No
TwoRoom	3	0.0117	0.1814	0.0254	0.0069	2.672	Yes
TwoRoom	4	0.0098	0.1615	0.0515	0.0032	2.160	Yes
PushT	2	0.0312	0.0001	0.0000	0.0021	5.533	Yes
PushT	3	0.0393	0.0003	0.0203	0.0026	4.472	Yes
PushT	4	0.0531	0.0005	0.0528	0.0031	3.891	Yes
Reacher	2	0.0115	0.7489	0.0000	0.0062	15.106	Yes
Reacher	3	0.0075	0.4660	0.0406	0.0092	10.788	Yes
Reacher	4	0.0047	0.3077	0.0789	0.0094	9.731	Yes
Cube	2	0.0020	0.1226	0.0000	0.0053	3.118	Yes
Cube	3	0.0025	0.1879	0.0371	0.0035	2.217	Yes
Cube	4	0.0022	0.1828	0.0961	0.0042	1.978	Yes

Table D.5: **LeWM audit of identifiable theory-control quantities.** The quantities and threshold-side test follow Table D.4. The candidate partition is treated as fixed, so the audit does not include uncertainty from re-estimating the partition.

Task	K	$\max_\ell \widehat{r}_b$	$\max_\ell \widehat{r}_{\mathrm{orient}}$	$\max_\ell \widehat{r}_{\mathrm{pair}}$	$\max_\ell \widehat{\epsilon}_{q,95}$	$\max_\ell \widehat{u}_{\mathrm{aff},95}$	Side resolved
TwoRoom	2	0.0301	0.0526	0.0000	0.0027	16.181	Yes
TwoRoom	3	0.0670	0.0775	0.0245	0.0032	11.249	Yes
TwoRoom	4	0.0418	0.0819	0.0784	0.0042	9.215	Yes
PushT	2	0.0277	0.0007	0.0000	0.0018	5.580	Yes
PushT	3	0.0390	0.0004	0.0424	0.0028	4.492	Yes
PushT	4	0.0561	0.0005	0.0309	0.0026	3.926	Yes
Reacher	2	0.0140	0.8327	0.0000	0.0059	14.683	Yes
Reacher	3	0.0105	0.4469	0.0523	0.0082	11.913	Yes
Reacher	4	0.0067	0.3076	0.1420	0.0080	10.426	Yes
Cube	2	0.0031	0.0741	0.0000	0.0023	5.655	Yes
Cube	3	0.0038	0.0732	0.0257	0.0027	4.227	Yes
Cube	4	0.0039	0.0550	0.0420	0.0024	3.559	Yes

For Sub-JEPA, orientation mismatch is largest for Reacher, intercept mismatch is largest for PushT, and the largest pair-aggregation remainder occurs on Cube at $K = 4$. Its bootstrap score deviation resolves the calibrated-threshold side in 11 of the 12 configurations; TwoRoom at $K = 2$ remains unresolved. For LeWM, the corresponding maxima occur on Reacher for orientation and pair aggregation, and on TwoRoom for the intercept term. Its bootstrap deviation resolves the threshold side in all 12 configurations. These statements concern score-estimation uncertainty only. The normalized affine-residual proxy ranges from 1.978 to 15.106 for Sub-JEPA and from 3.559 to 16.181 for LeWM. Because this held-out residual also contains conditional transition noise, it does not isolate the conditional-mean approximation error and is too conservative to serve as a tight structural bound.

Most importantly, the available artifacts do not provide a complete and independently justified numerical upper bound on $\epsilon(c)$. They do not identify all terms required to instantiate Theorem 4, including r_2, $\kappa_c - \kappa$, r_{extra}, the planning propagation and task-link deviations, and calibration uncertainty. We therefore treat Tables D.4 and D.5 as partial empirical audits rather than theorem-level certificates, and they are not used to revise the frozen selection policy or any reported decision.

E Combined Benchmark and Provenance

Table E.1: Frozen-policy leaderboard. Balanced accuracy averages recall for the `Regional` and `Global` classes over the combined 20-configuration benchmark.

#	Method	Cross-K	Cross-backbone	Overall	Raw acc. (%)	Balanced acc. (%)
1	Normalized cluster entropy	2/8	5/12	7/20	35	31.25
2	Eigengap after K	4/8	4/12	8/20	40	53.13
3	Prototype distance ratio	4/8	9/12	13/20	65	78.13
4	k-NN purity	3/8	8/12	11/20	55	34.38
5	Margin–radius ratio	4/8	6/12	10/20	50	50.00
6	Tangent contrast	4/8	6/12	10/20	50	31.25
7	Curvature tail	6/8	0/12	6/20	30	56.25
8	Flow persistence	5/8	10/12	15/20	75	56.25
9	Latent velocity eta-squared	3/8	4/12	7/20	35	31.25
10	Action-residual eta-squared	4/8	3/12	7/20	35	40.63
11	Affine response contrast	3/8	1/12	4/20	20	31.25
12	Minimum pairwise response	3/8	2/12	5/20	25	34.38
13	Pairwise response uniformity	5/8	6/12	11/20	55	71.88
14	Response–centroid Spearman	2/8	3/12	5/20	25	25.00
15	Response–boundary Spearman	4/8	7/12	11/20	55	53.13
16	Boundary-local Jacobian Bures Worst-Case	4/8	1/12	5/20	25	53.13
17	Partition Gain	8/8	10/12	18/20	90	93.75
18	Jacobian cosine distance	4/8	2/12	6/20	30	37.50
19	Jacobian log-scale distance	5/8	12/12	17/20	85	71.88
20	Jacobian subspace chordal distance	6/8	11/12	17/20	85	71.88
21	Retained safety fraction	1/8	7/12	8/20	40	25.00
22	Prominence ratio	6/8	10/12	16/20	80	87.50
–	Always `Global`	4/8	12/12	16/20	80	50.00
–	Always `Regional`	4/8	0/12	4/20	20	50.00

Worst-Case Partition Gain is the only method that is perfect on region-count transfer and retains the highest combined raw and balanced accuracy. Jacobian log-scale and subspace chordal distance are the strongest alternatives by raw accuracy at $17/20$, while prominence ratio has the strongest balanced accuracy among the other methods at 87.5%. The Always-`Global` baseline has high raw accuracy because all backbone transfer cases happen to be `Global`, but its balanced accuracy is 50% and it misses all four `Regional` winners.

E.1 Partition stability across seeds

We compare spectral partitioning (Ng et al., 2001; Zelnik-Manor and Perona, 2004), K-means++ (Arthur and Vassilvitskii, 2007), Gaussian mixture modeling (GMM) (McLachlan and Chang, 2004), Controlled PARC (Bemporad, 2023), and a Random Voronoi control at $K = 4$, using the same frozen LeWM latent cache and the same ordered samples within each task. For each method, we form partitions with seeds $\{0, 1, 2\}$. Each task entry in Table E.2 is the mean adjusted Rand index (ARI) (Hubert and Arabie, 1985) over seed pairs 0–1, 0–2, and 1–2; ARI is invariant to permutations of the region labels. The final column averages the four tasks equally. Random Voronoi is included as an unoptimized control, and no predictor training or planning evaluation enters this comparison.

Comparison implementations. All four comparison methods use the same coordinatewise standardization and unit-norm latent preprocessing as the deployed spectral partition. Spherical K-means++ clusters all unique training latents with 50 restarts and at most 1,000 iterations. GMM fits a full-covariance mixture to 20,000 uniformly sampled unique latents, using K-means initialization, one initialization, at most 1,000 iterations, tolerance 10^{-3}, and covariance regularization 10^{-6}; deployment uses maximum-posterior assignment. Controlled PARC fits 20,000 training windows and alternates for at most 15 updates, using the normalized current latent and standardized action embedding to predict the standardized next-latent displacement. Its deployed partition depends only on the current latent through a spherical Voronoi router. Random Voronoi samples K normalized

training latents without replacement and assigns every point to its most similar sampled prototype, without optimization. The partition seed controls sampling or initialization in every method.

Table E.2: Partition stability across seeds for LeWM at $K = 4$. Higher mean adjusted Rand index indicates greater agreement between independently seeded partitions of the same encoded dataset.

Rank	Partition method	TwoRoom	PushT	Reacher	Cube	Four-task mean ARI
1	**Spectral**	**0.934**	**0.613**	**0.814**	**0.615**	**0.744**
2	K-means++	0.255	0.344	0.471	0.266	0.334
3	GMM	0.347	0.246	0.199	0.490	0.321
4	Controlled PARC	0.138	0.235	0.167	0.084	0.156
5	Random Voronoi	0.020	0.027	0.014	0.008	0.018

Spectral partitioning yields by far the most reproducible decomposition across partition seeds. Since the encoded dataset and underlying environment dynamics are fixed across these draws, this result suggests that the spectral partition reflects persistent structure rather than seed-specific clustering artifacts. Notably, this stability exceeds even the tested dynamics-informed partition, Controlled PARC. Stability alone does not establish that the recovered regions are the unique ground-truth dynamics modes; rather, it provides an independent robustness check on the candidate partition used in our experiments.

The supplementary audit at `experiments/control_matrix/assets/lewm_k4_controlled_parc/SEED_STABILITY.md` records the per-seed labels, pairwise ARIs, cache hashes, and analysis command.

E.2 Weakest-pair versus All-pairs Wasserstein Aggregation

We directly probe the weakest-pair bottleneck in Assumption 4 by changing only how the regional pairwise Wasserstein scores are aggregated. For each partition seed ℓ, Worst-Case Partition Gain uses

$$\widehat{b}^{(\ell)}_{\min} := \min_{i<j} \widehat{b}^{(\ell)}_{ij}. \tag{33}$$

The comparison replaces this bottleneck with the energy- and prevalence-weighted all-pairs aggregate

$$\widehat{b}^{(\ell)}_{\mathrm{agg}} := \frac{\sum_{i<j} \widehat{\pi}^{(\ell)}_i \cdot \widehat{\pi}^{(\ell)}_j \cdot W_2^2\left(\widehat{\mathcal{P}}^{(\ell)}_i, \widehat{\mathcal{P}}^{(\ell)}_j\right)}{\sum_{i<j} \widehat{\pi}^{(\ell)}_i \cdot \widehat{\pi}^{(\ell)}_j \cdot \left(\widehat{e}^{(\ell)}_i + \widehat{e}^{(\ell)}_j\right)}, \tag{34}$$

where $\widehat{\mathcal{P}}^{(\ell)}_k = \mathcal{N}(\mathbf{0}, \widehat{\mathbf{C}}^{(\ell)}_k)$. This aggregation allows strongly separated pairs to compensate for a weak pair. Both scores are averaged over the same partition seeds $\{0, 1, 2\}$, use the same `regional` covariances, and are calibrated separately using the same LeWM $K = 4$ fitting protocol in Appendix B. The common 0.4-percentage-point indifference filter retains TwoRoom, Reacher, and Cube; PushT is excluded from both fits. The same candidate search and tie-breaking rule select the High direction for both scores, with thresholds 0.508948 for Worst-Case Partition Gain and 0.668090 for all-pairs. Both fit all three retained calibration labels. We then freeze these policies and evaluate all eight spectral $K = 2, 3$ configurations without further calibration or exclusion.

Table E.3: Region-count transfer after separate, matched $K = 4$ calibration. Worst-Case Partition Gain uses weakest-pair aggregation, while the comparator uses all-pairs aggregation; each rule uses its own frozen threshold.

Task	K	$R - G$ (pp)	Worst-Case Partition Gain	All-pairs	Worst-Case Partition Gain pred.	All-pairs pred.	Winner
TwoRoom	2	1.200	0.898943	0.898943	R	R	R
PushT	2	-0.667	0.498970	0.498970	G	G	G
Reacher	2	-1.422	0.169779	0.169779	G	G	G
Cube	2	1.111	0.551394	0.551394	R	G	R
TwoRoom	3	3.244	0.826633	0.872229	R	R	R
PushT	3	-1.822	0.491828	0.540643	G	G	G
Reacher	3	-0.756	0.428340	0.526067	G	G	G
Cube	3	1.556	0.571179	0.621526	R	G	R

Worst-Case Partition Gain is correct in $8/8$ test configurations, compared with $6/8$ for all-pairs; both raw and balanced accuracy are therefore 100% versus 75%. All-pairs incorrectly selects `Global` for Cube at both $K = 2$ and $K = 3$. Although the two scores coincide at $K = 2$, their predictions can differ because their thresholds were fitted separately at $K = 4$. Thus the performance gap persists after matched calibration and concerns transfer of the frozen selection rules across region counts. In the additional eight $K = 3$ K-means++ and random-Voronoi configurations, both independently calibrated rules achieve $6/8$.

For comparison, applying the Worst-Case Partition Gain threshold to both scores also gives $8/8$ versus $6/8$ on region-count transfer, but all-pairs then misclassifies PushT and Reacher at $K = 3$, rather than Cube. The matched-calibration results provide targeted empirical support for the bottleneck in Assumption 4.

E.3 Sensitivity to near-zero branch gaps

We test whether the combined result is driven by configurations in which the observed `Regional`–`Global` success gap is close to zero. Each reported branch comparison uses three post-training seeds, five evaluation sampling seeds, and 50 paired start–goal cases per evaluation seed, corresponding to $3 \cdot 5 \cdot 50 = 750$ evaluations per fixed partition. The `Regional` result additionally averages over the three partition seeds specified in the evaluation protocol. For an exclusion band δ, we remove configurations with $|R - G| \leq \delta$ percentage points and recompute each frozen method's raw and balanced accuracy without changing any prediction, direction, threshold, or hyperparameter. The bands are sensitivity choices, not statistical significance thresholds.

Table E.4: Sensitivity after excluding near-zero observed branch gaps. Entries for each method are raw accuracy / balanced accuracy (%). The log-scale and subspace chordal distance tie for the highest raw accuracy among the other methods on the full benchmark; prominence ratio has the highest balanced accuracy among the other methods.

Excluded $\lvert R - G \rvert$	Retained	Retained R/G	Worst-Case Partition Gain	Log-scale	Subspace chordal	Prominence ratio	Always `Global`
None	20	4/16	90.00 / 93.75	85.00 / 71.88	85.00 / 71.88	80.00 / 87.50	80.00 / 50.00
≤ 0.4 pp	16	4/12	93.75 / 95.83	81.25 / 70.83	87.50 / 75.00	75.00 / 83.33	75.00 / 50.00
≤ 1.0 pp	12	4/8	100.00 / 100.00	83.33 / 75.00	83.33 / 75.00	83.33 / 87.50	66.67 / 50.00

The per-configuration gaps used for this analysis are reported in Tables C.1 and D.1, and the complete frozen predictions are reported in Tables C.2 and D.2. Worst-Case Partition Gain remains the highest-accuracy method under both exclusion bands and becomes perfect after removing configurations within one percentage point of zero. This sensitivity analysis does not assign a cause to a small observed gap: such a gap may reflect finite-sample variation, limited benefit from the prespecified partition, or both. It instead shows that the main comparative conclusion is not dependent on the near-zero configurations.

E.4 Artifact provenance

The anonymous code supplement is a frozen submission snapshot. For double-blind review, historical Git revision identifiers in its provenance manifests are replaced by `anonymous-submission-v1`; this provenance-only change does not alter the numerical result files.

Compact artifacts for Appendix Tables A.1–F.1 and their SHA-256 hashes are indexed in `experiments/control_matrix/assets/appendix_tables/appendix_artifact_manifest.json`. The LeWM and Sub-JEPA benchmark results are recorded in `experiments/control_matrix/assets/lewm_layer2_22_criteria/benchmark_manifest.json` and `experiments/control_matrix/assets/subjepa_22_criteria/benchmark_manifest.json`, respectively. The separately calibrated all-pairs comparison reported in Table E.3 uses `experiments/control_matrix/assets/lewm_bures_aggregate_validation/validation_manifest.json`. The LeWM and Sub-JEPA theory-control audits in Tables D.5 and D.4 are jointly indexed by `experiments/control_matrix/assets/theory_control_artifact_manifest.json`. Earlier shared-threshold outputs are retained for provenance under

`experiments/control_matrix/assets/lewm_bures_aggregate_validation/legacy_shared_threshold/`. Both aggregation analyses use the same pairwise-value file; their seed-summary and validation files differ because they apply different calibration policies.

Table E.5: SHA-256 hashes of the benchmark, theory-control, and Bures-aggregation artifacts. The matched-calibration files support Table E.3; the historical shared-threshold files are retained for provenance. Both versions share the pairwise-value file. The artifact names retain "Bures"; their pairwise values are identical to the squared Wasserstein distances between the centered Gaussian responses used here.

Evaluation	Artifact	SHA-256
LeWM	Frozen policies	`0253b94efef268837d21c3e72a905b07f51bd00f4010c2161519bc369b632d7c`
LeWM	Task scores	`01233c27337a8361e6e2040c4aac33b0795522dbd3703e12985dfd498c961165`
LeWM	Scores by seed	`99679cb71fa44b5b5336bd19a15394b5bc06f43002b9b0c006fdd02f970a6641`
LeWM	Predictions	`26c6cea7d1c6982e75e0d76b0707ef38267ece0d988cb2fbc282af46a04e1845`
LeWM	Accuracy	`5b7290fb09cf816dec85673269a103ce7931b0a2aa00baa1329b2108df557a94`
Sub-JEPA	Frozen policies	`db7c55e63e3c4f06df954531b240e951783ad44ca548fdd84f4a8938d576aab7`
Sub-JEPA	Task scores	`5d311a725749dcdbec219517ceb33adb292d60af9c9b1d301b02fcfcfefff3cc`
Sub-JEPA	Scores by seed	`a945e761736fe7bbf24ab96343ffb1672de3c8c892f3a9dd0d96a35a1750911d`
Sub-JEPA	Predictions	`d70250782def03e0e39918ebe70c743254c0fdc6edbe6a84fe638b3d7e03744c`
Sub-JEPA	Accuracy	`641a223d0bf69b02eba805640ebae965d1ec0a3e95d8cfc40b8fdeb5c6a34748`
Sub-JEPA theory audit	Structural summary	`8b7f6d85c9f6357b8b45a1aa48022579f6fee0700090e5f12bbe172087f23b7f`
Sub-JEPA theory audit	Finite-sample summary	`2d8c6aabf04ced4bea00d903d3c6b2f274dc1f1792725fe46e67a86014810822`
LeWM theory audit	Finite-sample summary	`b14b20787675b252d874ee9101788a1cfd8fa5d744f627b9d1817e0fce5a685a`
LeWM theory audit	Identifiable-control summary	`e10031927e60eb1a1ef9da245ebf421aa86c2e5b892f4e8ad31c6ee68b27d807`
Theory-control audits	Artifact manifest	`b9f2341ec9ec549b90061624601c61d1ac7aca8c493cf78b1b9e8fdc066fe19f`
Bures aggregation (both versions)	Pairwise values	`1438612cc6d0c1fdddf96c1dbd6fda9cc061377aa6bb1edd89e5e94b67a5039c`
Bures aggregation (matched calibration)	Scores by seed	`c13f647a81f07dca8c13c8710225bd72c98f013a7d4b743691b680e01b89dff7`
Bures aggregation (matched calibration)	Validation	`79e6ad8d7ffa2873b3bf26d8d7169dcbc8c44203d37ee43ccad9e1c6cb6ed022`
Bures aggregation (historical shared threshold)	Scores by seed	`0fdcd850bd22815fe6878db70d900f93cea4a9a35ade7a3dc6ba3fca9d066c3d`
Bures aggregation (historical shared threshold)	Validation	`6232899cb209141fb1955ab8ce3db55061a990d6a81a0a9d4ae04b3e56f38251`

For an archive-only reproduction of the reported comparisons, the commands in `README.md` reaggregate the included per-seed scores using the frozen policies and branch-outcome tables. Figure 2 can be regenerated from the included result tables. These steps require no retraining or reevaluation. Full metric recomputation requires the official datasets and checkpoints, followed by preparation of the latent caches, partitions, and gate manifests; its prerequisites and commands are documented in `REPRODUCIBILITY.md`.

E.5 Training and evaluation recipe

We follow the released LeWM training and MPC recipe (Maes et al., 2026). The only evaluation changes are five sampling seeds and a 50-step start–goal separation for the long-horizon tasks; all other training and evaluation settings are unchanged.

Table E.6: Training and evaluation recipe used by both branches.

Stage	Setting	Value
Training	Initialization	Official pretrained task checkpoint
Training	Frozen modules	Encoder, projector, and action encoder
Training	Trainable modules	Dynamics predictor and prediction projection
Training	Objective	One-step latent mean squared error
Training	Optimizer	AdamW (Loshchilov and Hutter, 2019); learning rate $5 \cdot 10^{-5}$; weight decay 10^{-3}
Training	Schedule	50 epochs; batch size 128; FP32 precision
Training	Temporal input	History size 3; one predicted step; frame skip 5
Training	Post-training seeds	$\{0, 42, 625\}$
Training	`Regional` requirement	At least 256 transition samples in every region
Training	Checkpoint selection	Lowest loss on the corresponding region cache
Partition	Partition seeds	$\{0, 1, 2\}$ for the `Regional` branch
Evaluation	Primary endpoint	Mean 50-step long-horizon task success rate
Evaluation	Sampling seeds	$\{0, 1, 2, 3, 4\}$
Evaluation	Paired set	50 identical start–goal pairs per seed for both branches
Evaluation	Long-horizon setup	Start–goal offset 50; evaluation budget 50
Evaluation	MPC setup	Planning horizon 5; receding horizon 5; action block 5
Evaluation	`Regional` routing	Route from the current latent once per MPC cycle and retain that predictor during each candidate rollout
Reporting	Aggregation	Average evaluation seeds within training seed; `Regional` also averages partition seeds; report mean and sample standard deviation across training seeds

E.6 Spectral partition implementation

For a fixed region count K, let $\{\mathbf{z}_n\}_{n=1}^N \subset \mathbb{R}^d$ be the frozen-encoder latents. We standardize each coordinate and normalize each vector:

$$\bar{\mathbf{z}}_n = (\mathbf{z}_n - \widehat{\boldsymbol{\mu}}_z) \oslash (\widehat{\mathbf{s}}_z + 10^{-6}), \qquad \mathbf{x}_n = \bar{\mathbf{z}}_n / \|\bar{\mathbf{z}}_n\|_2. \tag{35}$$

We sample landmarks with a fixed seed and build a self-tuned 30-nearest-neighbor graph (Zelnik-Manor and Perona, 2004). With $(d_{ij}^{\text{spec}})^2 = \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$ and σ_i the distance to landmark i's 30th neighbor, directed affinities are

$$w_{ij} = \exp[-(d_{ij}^{\text{spec}})^2/(\sigma_i \cdot \sigma_j)]. \tag{36}$$

We form a symmetric affinity matrix $\mathbf{W}$ with entries $[\mathbf{W}]_{ij} = \max\{w_{ij}, w_{ji}\}$, and then compute

$$\mathbf{L}_{\text{n}} = \mathbf{I} - \mathbf{D}^{-1/2} \cdot \mathbf{W} \cdot \mathbf{D}^{-1/2}, \qquad \mathbf{D} = \text{diag}(\mathbf{W} \cdot \mathbf{1}). \tag{37}$$

Let $\mathbf{U}$ contain the eigenvectors associated with the K smallest eigenvalues. Row-normalized spectral coordinates $\mathbf{y}_i = \mathbf{U}_{i,:}/\|\mathbf{U}_{i,:}\|_2$ (Ng et al., 2001) are clustered by Euclidean K-means++ (Arthur and Vassilvitskii, 2007) with 20 initializations and at most 300 Lloyd iterations. This spectral map amplifies latent geometric differences before assigning regions.

For out-of-sample routing, we fit 16 prototype centers in the original normalized coordinates within each spectral region and normalize each prototype. A new latent is assigned to the owner of its most similar prototype:

$$p^*(\mathbf{z}) = \arg\max_p \mathbf{x}(\mathbf{z})^\top \cdot \mathbf{q}_p, \qquad \widehat{k}(\mathbf{z}) = \text{owner}(\mathbf{q}_{p^*(\mathbf{z})}). \tag{38}$$

The stored training mean and scale define $\mathbf{x}(\mathbf{z})$. Multiple prototypes per region permit non-spherical routing regions. This subsection describes only the partition implementation; it does not include the selection methods or post-training procedure.

F Decodability of Manually Defined Regions

We examine whether the frozen LeWM representation retains information about manually defined TwoRoom regions. We train linear and nonlinear probes to predict these region labels from latent vectors while keeping the encoder fixed.

Region labels and data. We consider two geometry-based partitions. `rooms3` distinguishes the left room, doorway corridor, and right room, using the horizontal position boundaries $x = 107$ and $x = 117$ in the environment's coordinate system. `priority5` assigns mutually exclusive labels in priority order: doorway corridor, near-wall region, common interior, remaining right-room region, and remaining left-room region. The near-wall mask uses $x \leq 29$, $x \geq 194$, $y \leq 29$, or $y \geq 194$. The common-interior mask uses $34 \leq y \leq 189$ and either $34 \leq x \leq 87$ or $137 \leq x \leq 189$; higher-priority labels take precedence. Positions define the targets but are not probe inputs. Inputs are single 192-dimensional latent vectors from cached observation sequences, expanded at frame skip 5 and deduplicated by `global` timestep.

Probe protocol. We split episodes into 70% training, 15% validation, and 15% test sets, with no episode shared across these splits. Latent standardization is fitted on the probe-training split only. We compare a linear softmax classifier with a nonlinear classifier using 8,192 fixed random Fourier features approximating an RBF kernel ($\gamma = 1/192$), followed by a linear softmax classifier. Both use cross-entropy loss and AdamW for five epochs, with batch size 16,384, learning rate 10^{-2}, and weight decay 10^{-4}. Within each seed, the probes share the same episode split. Seeds $\{0, 1, 2, 3, 4\}$ control the episode split, probe initialization, minibatch order, and random Fourier features, using the same frozen encoder across all runs.

Results and interpretation. Table F.1 shows high label-recovery performance even with a linear probe. The paired macro-F1 improvement from RFF-RBF is 0.046 ± 0.016 percentage points for `rooms3` and 0.780 ± 0.055 for `priority5`. These results show that the frozen representation retains readily decodable information about the manual regions. Linear probes already recover both partitions with high accuracy, with a larger nonlinear gain for `priority5`.

Table F.1: Recovery of manual TwoRoom region labels from frozen LeWM latents. Test metrics (%) are means ± population standard deviations across five split/probe seeds.

Partition	Probe	Accuracy	Balanced accuracy	Macro-F1
`rooms3`	Linear	99.90 ± 0.01	99.71 ± 0.04	99.67 ± 0.04
	RFF-RBF	99.91 ± 0.01	99.56 ± 0.06	99.72 ± 0.03
`priority5`	Linear	98.70 ± 0.06	98.11 ± 0.05	98.01 ± 0.08
	RFF-RBF	99.21 ± 0.03	98.80 ± 0.11	98.79 ± 0.05

Reproducibility. The implementation is `experiments/tworoom/geometry_latent_probe_multiseed.py`, using the labeling and probe routines in `geometry_latent_svm_rooms3.py` in the same directory. The reported values come from `rooms3_multiseed_summary.csv` and `priority5_multiseed_summary.csv` under `experiments/tworoom/assets/probe_test_multiseed/data/`.

G Proofs and Additional Details for the Geometric Theory

This appendix supplies the supporting definitions, algebraic derivations, and proofs for Section 3, starting with response geometry and then deriving prediction-risk and finite-horizon success guarantees. The original structural assumptions are replaced by explicit budgets; the exact threshold is retained as a restricted special case.

Definition 1 (Matched fair comparison class)**.** *Let u collect the matched task and protocol components: the task, frozen representation, component-model architecture, overall transition dataset, training procedure, evaluation distribution, horizon, search budget, success predicate, environment coupling, and common planner. For a given partition Π with $K \geq 2$ regions, define a configuration $c = (u, \Pi, K)$. The configuration belongs to the fair comparison class $\mathfrak{C}$ when its only branch difference is the choice between `regional` and `global` dynamics modeling; every component collected in u is identical for the two branches. A configuration is called admissible when it also satisfies the assumptions of the selection result under consideration.*

For $c \in \mathfrak{C}$, let $p_{\mathrm{R},H}(c)$ and $p_{\mathrm{G},H}(c)$ denote the population planning-success probabilities at the matched execution horizon H for the `Regional` and `Global` branches, and define

$$\Delta p_H(c) := p_{\mathrm{R},H}(c) - p_{\mathrm{G},H}(c). \tag{39}$$

The selection target is the sign of $\Delta p_H(c)$.

G.1 Latent Responses and Proof of Theorem 1

Let $p(d\mathbf{s}' \mid \mathbf{s}, \mathbf{a})$ be an environment transition kernel on state space $\mathcal{S}$, let $h : \mathcal{S} \to \mathcal{Z} \subseteq \mathbb{R}^d$ be the representation used in the comparison, and let ν_a be a reference probability measure on the action space. Define the expected latent displacement

$$\Gamma_h(\mathbf{s})(\mathbf{a}) := \mathbb{E}[\mathbf{z}_{t+1} - \mathbf{z}_t \mid \mathbf{s}_t = \mathbf{s}, \mathbf{a}_t = \mathbf{a}] \in \mathbb{R}^d, \tag{40}$$

and the response distance

$$d_{\mathrm{dyn}}(\mathbf{s}, \mathbf{t}) := \|\Gamma_h(\mathbf{s}) - \Gamma_h(\mathbf{t})\|_{L^2(\nu_a)}. \tag{41}$$

All response functions below are assumed to belong to $L^2(\nu_a; \mathbb{R}^d)$.

Assumption 1 (Accurate and Lipschitz-continuous auxiliary response map)**.** *There is an idealized auxiliary response map $\widetilde{\Gamma} : \mathcal{Z} \to L^2(\nu_a; \mathbb{R}^d)$ and constants $\varepsilon \geq 0$ and $\ell_\Gamma > 0$ such that, for all $\mathbf{s}, \mathbf{t} \in \mathcal{S}$,*

$$\|\widetilde{\Gamma}(h(\mathbf{s})) - \Gamma_h(\mathbf{s})\|_{L^2(\nu_a)} \leq \varepsilon, \tag{42}$$

$$\|\widetilde{\Gamma}(h(\mathbf{s})) - \widetilde{\Gamma}(h(\mathbf{t}))\|_{L^2(\nu_a)} \leq \ell_\Gamma \cdot \|h(\mathbf{s}) - h(\mathbf{t})\|_2 \tag{43}$$

holds.

Formal statement of Theorem 1. Under Assumption 1, for all $\mathbf{s}, \mathbf{t} \in \mathcal{S}$,

$$\|h(\mathbf{s}) - h(\mathbf{t})\|_2 \geq \frac{[d_{\mathrm{dyn}}(\mathbf{s}, \mathbf{t}) - 2 \cdot \varepsilon]_+}{\ell_\Gamma}. \tag{44}$$

Thus sufficiently different local dynamics remain separated in any latent space that admits an accurate and Lipschitz-continuous auxiliary response map.

Proof of Theorem 1. The triangle inequality and Equations (42)–(43) give

$$\begin{aligned} d_{\mathrm{dyn}}(\mathbf{s}, \mathbf{t}) &\leq \|\Gamma_h(\mathbf{s}) - \widetilde{\Gamma}(h(\mathbf{s}))\|_{L^2(\nu_a)} + \|\widetilde{\Gamma}(h(\mathbf{s})) - \widetilde{\Gamma}(h(\mathbf{t}))\|_{L^2(\nu_a)} \\ &\quad + \|\widetilde{\Gamma}(h(\mathbf{t})) - \Gamma_h(\mathbf{t})\|_{L^2(\nu_a)} \\ &\leq 2 \cdot \varepsilon + \ell_\Gamma \cdot \|h(\mathbf{s}) - h(\mathbf{t})\|_2. \end{aligned}$$

Rearranging and applying the positive part proves Equation (44). □

Fix $K \geq 2$. Let

$$\Pi : \mathcal{Z} \longrightarrow \{1, \ldots, K\} \tag{45}$$

be a measurable partition, and define the region variable $r := \Pi(h(\mathbf{s}))$ for an initial environment state $\mathbf{s} \sim \rho_0$, where ρ_0 is a reference distribution. Its latent pushforward is denoted by $\nu_0 = h_\# \rho_0$. Assume $\pi_k := \mathbb{P}(r = k) > 0$ for every k. Nothing in what follows depends on how Π was produced. In particular, it may be obtained by spectral clustering, K-means, a mixture model, physical labels, or another partitioning procedure.

Let the reference action have mean $\boldsymbol{\mu}_a$ and covariance

$$\boldsymbol{\Sigma}_a := \mathbb{E}_{\mathbf{a} \sim \nu_a} \left[(\mathbf{a} - \boldsymbol{\mu}_a) \cdot (\mathbf{a} - \boldsymbol{\mu}_a)^\top \right] \succeq 0. \tag{46}$$

Assumption 2 (Bounded regionwise first-order response approximation)**.** *For every region k, there are $\mathbf{b}_k \in \mathbb{R}^d$, an action-response slope $\mathbf{J}_k \in \mathbb{R}^{d \times p}$, and an error bound $\delta_k \geq 0$ such that*

$$\|\Gamma_h(\mathbf{s})(\cdot) - [\mathbf{b}_k + \mathbf{J}_k \cdot (\cdot - \boldsymbol{\mu}_a)]\|_{L^2(\nu_a)} \leq \delta_k \tag{47}$$

whenever $\Pi(h(\mathbf{s})) = k$.

The Worst-Case Partition Gain score is constructed from the action-dependent slope $\mathbf{J}_k$. Define the response factor $\mathbf{A}_k$ and response matrix $\mathbf{C}_k$, the covariance of its first-order action-induced latent variation, by

$$\mathbf{A}_k := \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a^{1/2} \in \mathbb{R}^{d \times p}, \qquad \mathbf{C}_k := \mathbf{A}_k \cdot \mathbf{A}_k^\top = \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a \cdot \mathbf{J}_k^\top \succeq 0. \tag{48}$$

and its response energy

$$e_k := \mathrm{tr}(\mathbf{C}_k) = \|\mathbf{A}_k\|_F^2. \tag{49}$$

G.2 Response Distribution Distance and Its Wasserstein Representation

Recall the centered Gaussian response distributions $\mathcal{P}_k = \mathcal{N}(\mathbf{0}, \mathbf{C}_k)$. Their squared 2-Wasserstein distance has the closed form

$$W_2^2(\mathcal{P}_i, \mathcal{P}_j) = \mathrm{tr}(\mathbf{C}_i) + \mathrm{tr}(\mathbf{C}_j) - 2 \cdot \mathrm{tr}\left[\left(\mathbf{C}_i^{1/2} \cdot \mathbf{C}_j \cdot \mathbf{C}_i^{1/2} \right)^{1/2} \right]. \tag{50}$$

This formula applies to positive-semidefinite covariances, including singular and low-rank cases. Matrix square roots are the positive-semidefinite square roots. Write

$$\phi_{ij} := \mathrm{tr}\left[\left(\mathbf{C}_i^{1/2} \cdot \mathbf{C}_j \cdot \mathbf{C}_i^{1/2} \right)^{1/2} \right] \tag{51}$$

for the covariance overlap. For pairs with $e_i + e_j > 0$, the pairwise normalized squared Wasserstein score is

$$b_{ij} := \frac{d_{ij}^2}{e_i + e_j} = \frac{W_2^2(\mathcal{P}_i, \mathcal{P}_j)}{e_i + e_j} = 1 - \frac{2 \cdot \phi_{ij}}{e_i + e_j}. \tag{52}$$

The score is dimensionless. Its numerical value is not itself a universal planning-success threshold.

Lemma 1 (Factor and Procrustes forms)**.** *For the factors in Equation* (48)*,*

$$\phi_{ij} = \|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_*, \tag{53}$$

$$W_2^2(\mathcal{P}_i, \mathcal{P}_j) = \min_{\mathbf{Q}\in \mathrm{O}(p)} \|\mathbf{A}_i - \mathbf{A}_j \cdot \mathbf{Q}\|_F^2. \tag{54}$$

where $\mathrm{O}(p) := \{\mathbf{Q} \in \mathbb{R}^{p\times p} : \mathbf{Q}^\top \cdot \mathbf{Q} = \mathbf{I}_p\}$*. Consequently,*

$$0 \le b_{ij} \le 1, \qquad b_{ij} = 0 \iff \mathbf{C}_i = \mathbf{C}_j. \tag{55}$$

Proof. We first prove the factor form of the covariance overlap. Since $\mathbf{C}_j = \mathbf{A}_j \cdot \mathbf{A}_j^\top$,

$$\mathbf{C}_i^{1/2} \cdot \mathbf{C}_j \cdot \mathbf{C}_i^{1/2} = (\mathbf{C}_i^{1/2} \cdot \mathbf{A}_j) \cdot (\mathbf{C}_i^{1/2} \cdot \mathbf{A}_j)^\top. \tag{56}$$

For any matrix $\mathbf{X}$, the nonzero eigenvalues of $\mathbf{X} \cdot \mathbf{X}^\top$ and $\mathbf{X}^\top \cdot \mathbf{X}$ coincide, and they are the squared singular values of $\mathbf{X}$. Taking $\mathbf{X} = \mathbf{C}_i^{1/2} \cdot \mathbf{A}_j$, we obtain

$$\mathbf{X}^\top \cdot \mathbf{X} = \mathbf{A}_j^\top \cdot \mathbf{C}_i \cdot \mathbf{A}_j = \mathbf{A}_j^\top \cdot \mathbf{A}_i \cdot \mathbf{A}_i^\top \cdot \mathbf{A}_j \tag{57}$$

$$= (\mathbf{A}_i^\top \cdot \mathbf{A}_j)^\top \cdot (\mathbf{A}_i^\top \cdot \mathbf{A}_j). \tag{58}$$

Hence the eigenvalues of $\mathbf{C}_i^{1/2} \cdot \mathbf{C}_j \cdot \mathbf{C}_i^{1/2}$ are the squared singular values of $\mathbf{A}_i^\top \cdot \mathbf{A}_j$, up to additional zeros. Taking the positive-semidefinite square root and then the trace therefore sums those singular values:

$$\phi_{ij} = \sum_{r=1}^{p} \sigma_r(\mathbf{A}_i^\top \cdot \mathbf{A}_j) = \|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_*. \tag{59}$$

This proves Equation (53).

We next prove the Procrustes form. For any $\mathbf{Q} \in \mathrm{O}(p)$, expanding the squared Frobenius norm gives

$$\|\mathbf{A}_i - \mathbf{A}_j \cdot \mathbf{Q}\|_F^2 = \mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_i) + \mathrm{tr}(\mathbf{Q}^\top \cdot \mathbf{A}_j^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}) - 2 \cdot \mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}) \tag{60}$$

$$= e_i + e_j - 2 \cdot \mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}), \tag{61}$$

where orthogonality gives $\mathrm{tr}(\mathbf{Q}^\top \cdot \mathbf{A}_j^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}) = \mathrm{tr}(\mathbf{A}_j^\top \cdot \mathbf{A}_j) = e_j$. Thus minimizing the distance is equivalent to maximizing $\mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_j \cdot \mathbf{Q})$.

Let

$$\mathbf{A}_i^\top \cdot \mathbf{A}_j = \mathbf{U} \cdot \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \cdot \mathbf{V}^\top \tag{62}$$

be a singular value decomposition, with zero singular values included when the matrix is rank deficient. Setting $\mathbf{R} := \mathbf{V}^\top \cdot \mathbf{Q} \cdot \mathbf{U}$, we have $\mathbf{R} \in \mathrm{O}(p)$ and

$$\mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}) = \mathrm{tr}\left(\mathrm{diag}(\sigma_1, \ldots, \sigma_p) \cdot \mathbf{R}\right) \tag{63}$$

$$= \sum_{r=1}^{p} \sigma_r \cdot [\mathbf{R}]_{rr} \le \sum_{r=1}^{p} \sigma_r = \|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_*. \tag{64}$$

The bound is attained by $\mathbf{R} = \mathbf{I}_p$, equivalently by $\mathbf{Q} = \mathbf{V} \cdot \mathbf{U}^\top$. Therefore,

$$\max_{\mathbf{Q}\in \mathrm{O}(p)} \mathrm{tr}(\mathbf{A}_i^\top \cdot \mathbf{A}_j \cdot \mathbf{Q}) = \|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_*. \tag{65}$$

Substituting this identity and Equation (53) into Equation (61) yields

$$\min_{\mathbf{Q}\in \mathrm{O}(p)} \|\mathbf{A}_i - \mathbf{A}_j \cdot \mathbf{Q}\|_F^2 = e_i + e_j - 2 \cdot \|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_* \tag{66}$$

$$= e_i + e_j - 2 \cdot \phi_{ij} = W_2^2(\mathcal{P}_i, \mathcal{P}_j), \tag{67}$$

which proves Equation (54).

The minimum-of-squares representation gives $W_2^2(\mathcal{P}_i, \mathcal{P}_j) \ge 0$. Moreover, $\phi_{ij} \ge 0$, so

$$W_2^2(\mathcal{P}_i, \mathcal{P}_j) = e_i + e_j - 2 \cdot \phi_{ij} \le e_i + e_j. \tag{68}$$

Dividing by $e_i + e_j > 0$ proves $0 \le b_{ij} \le 1$.

Finally, if $W_2^2(\mathcal{P}_i, \mathcal{P}_j) = 0$, the Procrustes form provides an orthogonal $\mathbf{Q}$ such that $\mathbf{A}_i = \mathbf{A}_j \cdot \mathbf{Q}$, and consequently

$$\mathbf{C}_i = \mathbf{A}_i \cdot \mathbf{A}_i^\top = \mathbf{A}_j \cdot \mathbf{Q} \cdot \mathbf{Q}^\top \cdot \mathbf{A}_j^\top = \mathbf{C}_j. \tag{69}$$

Conversely, if $\mathbf{C}_i = \mathbf{C}_j$, two equal-width factors of the same positive-semidefinite matrix are related by a right-orthogonal transformation. Thus $\mathbf{A}_i = \mathbf{A}_j \cdot \mathbf{Q}$ for some $\mathbf{Q} \in \mathrm{O}(p)$, and the Procrustes minimum is zero. Therefore $b_{ij} = 0$ if and only if $\mathbf{C}_i = \mathbf{C}_j$. □

The response distribution distance has the equivalent factor forms

$$\begin{aligned} d_{ij}^2 &= \min_{\mathbf{Q} \in \mathrm{O}(d)} \|\mathbf{C}_i^{1/2} - \mathbf{C}_j^{1/2} \cdot \mathbf{Q}\|_F^2 \\ &= \min_{\mathbf{Q} \in \mathrm{O}(p)} \|\mathbf{A}_i - \mathbf{A}_j \cdot \mathbf{Q}\|_F^2, \qquad \mathbf{A}_k = \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a^{1/2}. \end{aligned} \tag{70}$$

The first form follows by applying the same factor identity to the square-root factors $\mathbf{C}_i^{1/2}$ and $\mathbf{C}_j^{1/2}$. Thus, the $d \times d$ factor form used in the main text and the $d \times p$ factor form used in the subsequent risk analysis are equivalent.

For a probabilistic interpretation, let $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$. For every $\mathbf{Q} \in \mathrm{O}(d)$,

$$\mathbf{C}_i^{1/2} \cdot \boldsymbol{\xi} \sim \mathcal{P}_i, \qquad \mathbf{C}_j^{1/2} \cdot \mathbf{Q} \cdot \boldsymbol{\xi} \sim \mathcal{P}_j. \tag{71}$$

Moreover,

$$d_{ij}^2 = \min_{\mathbf{Q} \in \mathrm{O}(d)} \mathbb{E}\left[\left\|\mathbf{C}_i^{1/2} \cdot \boldsymbol{\xi} - \mathbf{C}_j^{1/2} \cdot \mathbf{Q} \cdot \boldsymbol{\xi}\right\|_2^2\right]. \tag{72}$$

Orthogonal transformations preserve the standard Gaussian distribution, so the two marginal response distributions remain unchanged for every Q. Thus, this quantity is the minimum mean-squared response difference after optimal factor alignment. The variable $\boldsymbol{\xi}$ represents the response distributions and is not a fixed realized action.

Proof of Theorem 2. The definition in Equation (4) and the factor identity above give

$$\begin{aligned} d_{ij}^2 &= \min_{\mathbf{Q} \in \mathrm{O}(d)} \|\mathbf{C}_i^{1/2} - \mathbf{C}_j^{1/2} \cdot \mathbf{Q}\|_F^2 \\ &= W_2^2(\mathcal{P}_i, \mathcal{P}_j). \end{aligned}$$

This proves Equation (5). □

Under the Gaussian reference action distribution used in the main text, $\mathcal{P}_k = \mathcal{N}(\mathbf{0}, \mathbf{C}_k)$ is the distribution of the centered first-order response $\mathbf{J}_k \cdot (\mathbf{a} - \boldsymbol{\mu}_a)$. The squared Wasserstein distance between the full affine-response distributions is

$$W_2^2(\mathcal{N}(\mathbf{b}_i, \mathbf{C}_i), \mathcal{N}(\mathbf{b}_j, \mathbf{C}_j)) = \|\mathbf{b}_i - \mathbf{b}_j\|_2^2 + W_2^2(\mathcal{P}_i, \mathcal{P}_j). \tag{73}$$

The score studied here compares the centered responses. Baseline differences are accounted for by r_b, while the gap between differences in fixed action coordinates and optimal orthogonal alignment is accounted for by r_{orient}, in the later risk decomposition in Equation (85); the regionwise first-order approximation error is controlled by Assumption 2. For a general action reference with finite second moments, $\mathcal{P}_k$ is the Gaussian distribution matching the mean and covariance of the centered first-order response; the Wasserstein identities above concern these Gaussian summaries.

G.2.1 Response Strength and Shape in the Wasserstein Score

The response matrix $\mathbf{C}_k$ records both the overall strength of action-induced latent variation and how that variation is distributed across latent directions. To separate these two contributions, recall the response energy $e_k = \mathrm{tr}(\mathbf{C}_k)$. For a region with $e_k > 0$, define its normalized response matrix by

$$\overline{\mathbf{C}}_k := \frac{\mathbf{C}_k}{e_k}, \qquad \mathrm{tr}(\overline{\mathbf{C}}_k) = 1. \tag{74}$$

Here e_k measures total response strength, while $\overline{\mathbf{C}}_k$ describes response shape: its eigenvectors give the response directions, and its eigenvalues give the fraction of total response energy in each direction.

For two regions with positive response energies, define their shape overlap f_{ij} through their normalized covariance overlap:

$$f_{ij} := \operatorname{tr}\left[\left(\overline{\mathbf{C}}_i^{1/2} \cdot \overline{\mathbf{C}}_j \cdot \overline{\mathbf{C}}_i^{1/2}\right)^{1/2}\right]. \tag{75}$$

This overlap lies in $[0, 1]$, with $f_{ij} = 1$ exactly when the two normalized response matrices are identical. Thus $1 - f_{ij}$ measures their shape discrepancy.

Proposition 1 (Strength–shape decomposition of the Wasserstein score)**.** *For any region pair with $e_i, e_j > 0$, the normalized squared Wasserstein score satisfies*

$$b_{ij} = \underbrace{\frac{(\sqrt{e_i} - \sqrt{e_j})^2}{e_i + e_j}}_{\textit{strength contribution}} + \underbrace{\frac{2 \cdot \sqrt{e_i \cdot e_j}}{e_i + e_j} \cdot (1 - f_{ij})}_{\textit{shape contribution}}. \tag{76}$$

Both contributions are nonnegative. The first measures the difference in root response energy. The second weights the shape discrepancy by $2 \cdot \sqrt{e_i \cdot e_j}/(e_i + e_j)$, which equals 1 for equal energies and decreases as their ratio moves away from 1.

Proof. Substituting $\mathbf{C}_i = e_i \cdot \overline{\mathbf{C}}_i$ and $\mathbf{C}_j = e_j \cdot \overline{\mathbf{C}}_j$ into the covariance overlap defined in Equation (51) gives

$$\phi_{ij} = \sqrt{e_i \cdot e_j} \cdot f_{ij}.$$

The normalized score therefore becomes

$$\begin{aligned} b_{ij} &= 1 - \frac{2 \cdot \sqrt{e_i \cdot e_j}}{e_i + e_j} \cdot f_{ij} \\ &= \frac{e_i + e_j - 2 \cdot \sqrt{e_i \cdot e_j}}{e_i + e_j} + \frac{2 \cdot \sqrt{e_i \cdot e_j}}{e_i + e_j} \cdot (1 - f_{ij}), \end{aligned}$$

which proves the decomposition. To verify the stated range of f_{ij}, apply Lemma 1 to the normalized response matrices, whose factors are $\mathbf{A}_i/\sqrt{e_i}$ and $\mathbf{A}_j/\sqrt{e_j}$. Their total energy is 2, so their normalized squared Wasserstein score is $1 - f_{ij}$. The lemma gives $0 \le 1 - f_{ij} \le 1$, with equality to zero exactly when $\overline{\mathbf{C}}_i = \overline{\mathbf{C}}_j$. Finally, writing the energy ratio as $r := e_i/e_j > 0$, the shape weight is $2 \cdot \sqrt{r}/(1 + r)$. It increases on $(0, 1]$ and decreases on $[1, \infty)$, attaining 1 at $r = 1$. □

Interpretation. When two regions have the same response shape, $f_{ij} = 1$, and the score is entirely determined by their strength difference. When their response energies are equal, the strength contribution is zero and $b_{ij} = 1 - f_{ij}$. When both strength and shape agree, $b_{ij} = 0$. If exactly one region has zero response energy, the Wasserstein formula gives $b_{ij} = 1$, corresponding to the comparison between a zero response and a nonzero response. This decomposition explains how strength and shape differences contribute to the normalized response-distribution discrepancy b_{ij}.

G.3 Prediction Risk and the Worst-Case Partition Gain Lower Bound

Reference distribution and comparison class. Fix the partition and the trained candidates. Let ρ_0 be a distribution of initial environment states and let $\nu_0 = h_{\#}\rho_0$ be its latent pushforward. Actions in the one-step reference experiment are drawn independently from the same ν_a at every start. Thus $\pi_k = \nu_0\{\Pi(\mathbf{z}) = k\}$. A uniform latent-start reference is permitted: for a feasible set V of finite positive measure,

$$\nu_0(d\mathbf{z}) = \frac{\mathbb{I}_V(\mathbf{z}) \cdot \mu(d\mathbf{z})}{\mu(V)}. \tag{77}$$

Use intrinsic surface measure when V lies on a sphere or manifold; uniform starts need not give equal region probabilities. The planning expectations below use this same start distribution and the same conditional distribution of underlying states. For nonuniform evaluation, replace ν_0 by the evaluation distribution throughout. The training-data distribution need not equal this reference distribution.

Write $\mathbf{y} = \mathbf{z}_1 - \mathbf{z}_0$ and $m(\mathbf{z}, \mathbf{a}) = \mathbb{E}[\mathbf{y} \mid \mathbf{z}_0 = \mathbf{z}, \mathbf{a}_0 = \mathbf{a}]$ in this reference experiment. In $\mathcal{H} = L^2(\nu_0 \otimes \nu_a; \mathbb{R}^d)$, let

$$\begin{aligned} \mathcal{F}_{\mathrm{G}} &= \{\mathbf{b} + \mathbf{J} \cdot (\mathbf{a} - \boldsymbol{\mu}_a)\}, \\ \mathcal{F}_{\mathrm{R}} &= \{\mathbf{b}_{\Pi(\mathbf{z})} + \mathbf{J}_{\Pi(\mathbf{z})} \cdot (\mathbf{a} - \boldsymbol{\mu}_a)\}, \qquad \mathcal{F}_{\mathrm{G}} \subseteq \mathcal{F}_{\mathrm{R}}. \end{aligned} \tag{78}$$

These are finite-dimensional closed linear spaces of functions, even if some parameters are not identifiable. They model displacement; the corresponding next-state prediction is $\mathbf{z} + f(\mathbf{z}, \mathbf{a})$. For any fixed predictor define $\mathcal{R}(f) = \mathbb{E}\|\mathbf{y} - f(\mathbf{z}_0, \mathbf{a}_0)\|^2$, assumed finite. For $M \in \{\mathrm{R}, \mathrm{G}\}$, let

$$\mathcal{R}_M^* = \inf_{f \in \mathcal{F}_M} \mathcal{R}(f), \quad \mathcal{E}_M = \mathcal{R}(\widetilde{f}_M) - \mathcal{R}_M^*, \quad \widetilde{f}_M \in \mathcal{F}_M. \tag{79}$$

Excess risk includes estimation and optimization error. It is defined relative to the true oracle, including any non-affine residual in the conditional mean.

Direct link from response geometry to transitions. Set $q_{ij} = b_{ij}/2 = d_{ij}^2/[2 \cdot (e_i + e_j)]$ and $q = \min_{i<j} q_{ij} = b(\Pi)/2$. For $\mathbf{s}, \mathbf{t}$ in regions i, j, action centering gives

$$\|(\mathbf{b}_i - \mathbf{b}_j) + (\mathbf{J}_i - \mathbf{J}_j) \cdot (\mathbf{a} - \boldsymbol{\mu}_a)\|_{L^2(\nu_a)}^2 = \|\mathbf{b}_i - \mathbf{b}_j\|^2 + \|\mathbf{A}_i - \mathbf{A}_j\|_F^2. \tag{80}$$

Since the Procrustes relaxation is no larger than the value at $\mathbf{Q} = \mathbf{I}$, Assumption 2 implies

$$\|\Gamma_h(\mathbf{s}) - \Gamma_h(\mathbf{t})\|_{L^2(\nu_a)} \geq [\sqrt{2 \cdot (e_i + e_j) \cdot q_{ij}} - \delta_i - \delta_j]_+. \tag{81}$$

This inequality uses no completeness, bottleneck, or planning-order assumption. Common action coordinates matter: $\mathbf{A}_j \cdot \mathbf{Q}$ generally describes a different response to a fixed action, even though it has the same $\mathbf{C}_j$.

Exact decomposition for the affine reference. Let $m_0(\mathbf{z}, \mathbf{a}) = \mathbf{b}_{\Pi(\mathbf{z})} + \mathbf{J}_{\Pi(\mathbf{z})} \cdot (\mathbf{a} - \boldsymbol{\mu}_a)$ use the coefficients in Assumption 2. Conditional averaging and Jensen's inequality give

$$\|m - m_0\|_{\mathcal{H}} \leq \delta, \qquad \delta^2 = \sum_k \pi_k \cdot \delta_k^2. \tag{82}$$

This remains true if several environment states have the same latent state. Let $\mathbf{P}_M$ project onto $\mathcal{F}_M$, and set $\mathbf{T} = \mathbf{P}_{\mathrm{R}} - \mathbf{P}_{\mathrm{G}}$. Nestedness gives $\mathbf{T} = \mathbf{T}^* = \mathbf{T}^2$, hence $\|\mathbf{T}\|_{\mathrm{op}} \leq 1$. The conditional-noise risk cancels between the two oracles, so

$$d_* := \mathcal{R}_{\mathrm{G}}^* - \mathcal{R}_{\mathrm{R}}^* = \|\mathbf{T} \cdot m\|^2, \qquad d_0 := \|\mathbf{T} \cdot m_0\|^2. \tag{83}$$

The shared projection of m_0 uses coefficients $\bar{\mathbf{b}} = \sum_k \pi_k \cdot \mathbf{b}_k$, $\bar{\mathbf{A}} = \sum_k \pi_k \cdot \mathbf{A}_k$. The weighted variance identity therefore yields

$$d_0 = \sum_{i<j} \pi_i \cdot \pi_j \cdot (\|\mathbf{b}_i - \mathbf{b}_j\|^2 + \|\mathbf{A}_i - \mathbf{A}_j\|_F^2). \tag{84}$$

Define the following quantities on the same reference distribution:

$$\begin{aligned} s_c &= 2 \cdot \sum_{i<j} \pi_i \cdot \pi_j \cdot (e_i + e_j), & w_{ij} &= \frac{2 \cdot \pi_i \cdot \pi_j \cdot (e_i + e_j)}{s_c}, \\ \omega_{ij} &= \|\mathbf{A}_i - \mathbf{A}_j\|_F^2 - W_2^2(\mathcal{P}_i, \mathcal{P}_j) \geq 0, & r_{\mathrm{pair}} &= \sum_{i<j} w_{ij} \cdot (q_{ij} - q) \geq 0, \\ r_b &= s_c^{-1} \cdot \sum_{i<j} \pi_i \cdot \pi_j \cdot \|\mathbf{b}_i - \mathbf{b}_j\|^2, & r_{\mathrm{orient}} &= s_c^{-1} \cdot \sum_{i<j} \pi_i \cdot \pi_j \cdot \omega_{ij}. \end{aligned} \tag{85}$$

We assume positive pair energies, hence $s_c > 0$ and $\sum w_{ij} = 1$. Substitution gives the exact identity

$$d_0 = s_c \cdot (q + r_{\mathrm{pair}} + r_b + r_{\mathrm{orient}}) \geq s_c \cdot q. \tag{86}$$

The resulting Wasserstein contribution is an energy-weighted pair average. Its difference from the minimum is a nonnegative, explicit term.

Proposition 2 (One-sided prediction-risk guarantee)**.** *Under Assumption 2 and the nested affine comparison in Equation (78),*

$$\mathcal{R}_{\mathrm{G}}^* - \mathcal{R}_{\mathrm{R}}^* \geq [\sqrt{s_c \cdot q} - \delta]_+^2. \tag{87}$$

If $\delta/\sqrt{s_c} \leq \bar{\delta}$ and $(\mathcal{E}_{\mathrm{R}} - \mathcal{E}_{\mathrm{G}})/s_c \leq \bar{\kappa}$, with $\bar{\kappa} \geq 0$, then

$$\frac{\mathcal{R}_{\mathrm{G}}^{\mathrm{aff}} - \mathcal{R}_{\mathrm{R}}^{\mathrm{aff}}}{s_c} \geq \underline{g}(q) := [\sqrt{q} - \bar{\delta}]_+^2 - \bar{\kappa}. \tag{88}$$

Proof of Proposition 2. Projection contraction and Equation (82) imply $\|\mathbf{T}(m - m_0)\| \le \delta$. Therefore

$$[\sqrt{d_0} - \delta]_+^2 \le d_* \le (\sqrt{d_0} + \delta)^2. \tag{89}$$

Using $d_0 \ge s_c \cdot q$ proves the oracle bound. For the trained affine candidates, $\mathcal{R}_{\mathrm{G}}^{\mathrm{aff}} - \mathcal{R}_{\mathrm{R}}^{\mathrm{aff}} = d_* - (\mathcal{E}_{\mathrm{R}} - \mathcal{E}_{\mathrm{G}})$. Divide by s_c and use the assumed bounds on $\delta/\sqrt{s_c}$ and $(\mathcal{E}_{\mathrm{R}} - \mathcal{E}_{\mathrm{G}})/s_c$. □

Two-sided reference decomposition. Define

$$r_2 = \frac{d_* - d_0}{s_c}, \qquad \kappa_c = \frac{\mathcal{E}_{\mathrm{R}} - \mathcal{E}_{\mathrm{G}}}{s_c}. \tag{90}$$

The preceding proof also gives

$$|r_2| \le \epsilon_2 := \frac{2 \cdot \sqrt{d_0} \cdot \delta + \delta^2}{s_c}. \tag{91}$$

For actual neural candidates, including any required routing and risk distribution changes, define r_{extra} by

$$\frac{\mathcal{R}_{\mathrm{G}} - \mathcal{R}_{\mathrm{R}}}{s_c} = \frac{\mathcal{R}_{\mathrm{G}}^{\mathrm{aff}} - \mathcal{R}_{\mathrm{R}}^{\mathrm{aff}}}{s_c} + r_{\mathrm{extra}}. \tag{92}$$

Explicit ways to bound this term are given below. Choose a common reference cost $\kappa \in (0, 1/2)$, specified independently of the target success gaps. Then

$$\begin{aligned} d_{\mathrm{risk}}(c) &:= \frac{\mathcal{R}_{\mathrm{G}} - \mathcal{R}_{\mathrm{R}}}{s_c} = g(q) + r_{\mathrm{risk}}, \quad g(q) = q - \kappa, \\ r_{\mathrm{risk}} &= r_b + r_{\mathrm{orient}} + r_{\mathrm{pair}} + r_2 - (\kappa_c - \kappa) + r_{\mathrm{extra}}. \end{aligned} \tag{93}$$

This is a prediction-risk identity, not a definition of planning advantage. Neither g nor $\underline{g}$ is asserted to be the exact success curve.

G.4 Controlled Replacements for the Structural Assumptions

The following assumptions replace the original completeness and exact bottleneck requirements by quantitative budgets. All bounds are established independently of the success gap being explained. Finite bounds alone do not imply that the selection margins will be satisfied.

Assumption 3 (Controlled response information and model discrepancy)**.** *For the two-sided comparison, the quantities in Equations (85) and (92) satisfy $r_b \le \epsilon_b$, $r_{\mathrm{orient}} \le \epsilon_{\mathrm{orient}}$, and $|r_{\mathrm{extra}}| \le \epsilon_{\mathrm{extra}}$. The one-sided comparison requires only the last bound, since $r_b, r_{\mathrm{orient}} \ge 0$ follows algebraically.*

Assumption 4 (Controlled pair aggregation and estimation cost)**.** *For the two-sided comparison, $r_{\mathrm{pair}} \le \epsilon_{\mathrm{pair}}$ and $|\kappa_c - \kappa| \le \epsilon_{\mathrm{cost}}$. The one-sided comparison requires only an upper bound $\kappa_c \le \bar{\kappa}$, $\bar{\kappa} \ge 0$, without an upper bound on the nonnegative aggregation contribution.*

Approximation error from Assumptions 1 and 2. Let $\mathbf{z}_k$ be a representative of region k, let $r_k = \sup_{\mathbf{z}:\Pi(\mathbf{z})=k} \|\mathbf{z} - \mathbf{z}_k\|$, and suppose the auxiliary response map has representative affine-fit error $\beta_k = \|\widetilde{\Gamma}(\mathbf{z}_k) - [\mathbf{b}_k + \mathbf{J}_k \cdot (\cdot - \boldsymbol{\mu}_a)]\|_{L^2(\nu_a)}$. The triangle inequality gives a valid choice

$$\delta_k = \varepsilon + \ell_\Gamma \cdot r_k + \beta_k \tag{94}$$

in Assumption 2; if a tighter valid bound is available, use it instead. Assumption 1 is not additionally needed for the risk theorem if a bound under Assumption 2 is already available.

Intercept and action-orientation terms. With $d_b = \max_{i<j} \|\mathbf{b}_i - \mathbf{b}_j\|$,

$$0 \le r_b \le \frac{(1 - \sum_k \pi_k^2) \cdot d_b^2}{2 \cdot s_c}. \tag{95}$$

The factor form of the covariance overlap gives

$$\omega_{ij} = 2 \cdot \{\|\mathbf{A}_i^\top \cdot \mathbf{A}_j\|_* - \langle \mathbf{A}_i, \mathbf{A}_j \rangle_F\}, \qquad 0 \le \omega_{ij} \le 4 \cdot \sqrt{e_i \cdot e_j}. \tag{96}$$

Consequently $0 \le r_{\mathrm{orient}} \le 1$; if $\omega_{ij} \le \theta \cdot (e_i + e_j)$ for every pair, then $r_{\mathrm{orient}} \le \theta/2$. The coarse bound of one need not be useful for a two-sided certificate. Estimated intercepts and factors require their own uncertainty bounds when used in these formulas.

Pair aggregation. The weights in Equation (85) give

$$0 \le r_{\text{pair}} \le \max_{i<j} q_{ij} - q \le \tfrac{1}{2}. \tag{97}$$

Thus concentration of pair scores yields a small aggregation remainder. The empirical comparison with all-pairs aggregation evaluates the selection rule; it does not prove that this remainder vanishes.

Corollary 1 (Uniform pairwise irreducibility)**.** *For $\eta \in [0, 1/2]$, $b(\Pi) > 2 \cdot \eta$ if and only if $q_{ij} > \eta$ for every pair.*

Proof. Take the minimum in $q_{ij} = b_{ij}/2$. □

A separate bound for trained-model complexity. Matching the overall data and recipe does not equalize `regional` sample sizes or make κ_c independent of K. One explicit example is balanced fixed-design least squares with $\delta = 0$, known zero intercepts, $\pi_k = 1/K$, $n/K \ge p$, whitened action designs $\mathbf{X}_k^\top \cdot \mathbf{X}_k = (n/K) \cdot \mathbf{I}_p$, and independent training noise of variance σ^2 in each output coordinate. Then

$$\mathbb{E}_{\text{train}}\mathcal{E}_{\text{G}} = \frac{d \cdot p \cdot \sigma^2}{n}, \qquad \mathbb{E}_{\text{train}}\mathcal{E}_{\text{R}} = \frac{K \cdot d \cdot p \cdot \sigma^2}{n}. \tag{98}$$

If the noise is Gaussian and the branches use the same samples,

$$\frac{n}{\sigma^2}(\mathcal{E}_{\text{R}} - \mathcal{E}_{\text{G}}) \sim \chi^2_{(K-1)\cdot d\cdot p}. \tag{99}$$

Indeed, writing each `regional` slope error as $\mathbf{W}_k$, its entries are independent $\mathcal{N}(0, K \cdot \sigma^2/n)$, the shared error is $\bar{\mathbf{W}} = K^{-1} \cdot \sum_k \mathbf{W}_k$, and the excess-risk difference equals $K^{-1} \cdot \sum_k \|\mathbf{W}_k - \bar{\mathbf{W}}\|_F^2$. Thus, with probability at least $1 - \eta$,

$$\kappa_c \le \frac{\sigma^2}{n \cdot s_c} q_{(K-1)\cdot d\cdot p, 1-\eta}, \tag{100}$$

where $q_{r,1-\eta}$ is a chi-square quantile. These are illustrative controls under the stated linear-model conditions, not neural-network excess-risk formulas. An expectation over training runs is not a deterministic bound for each trained model.

Actual-model, routing, and distribution discrepancies. For fixed actual and affine displacement predictors with $\|f_M - \widetilde{f}_M\|_{L^2} \le \zeta_M$, expansion of squared loss gives

$$|\mathcal{R}(f_M) - \mathcal{R}(\widetilde{f}_M)| \le 2 \cdot \sqrt{\mathcal{R}(\widetilde{f}_M)} \cdot \zeta_M + \zeta_M^2. \tag{101}$$

Sum the two branch bounds and divide by s_c to obtain ϵ_{class}. This step is needed because a `Global` neural model can represent distinct `regional` responses internally.

If correct and deployed routing differ only on an event of probability η_Π, and both squared losses lie in $[0, \ell_{\max}]$, their risk difference is at most $\ell_{\max} \cdot \eta_\Pi$. This contributes $\epsilon_{\text{route}} = \ell_{\max} \cdot \eta_\Pi / s_c$. For two risk distributions P, Q and the same two fixed predictors,

$$|\Delta\mathcal{R}_P - \Delta\mathcal{R}_Q| \le 2 \cdot \ell_{\max} \cdot \text{TV}(P, Q), \quad \text{TV}(P, Q) = \sup_A |P(A) - Q(A)|. \tag{102}$$

This contributes ϵ_{dist} after division by s_c. Use $\ell_{\max} = 4$ only when both the true and predicted next states lie on the unit sphere (or in the unit ball). Normalizing only clustering features does not imply this loss bound.

Add budgets only for distinct steps in an explicit comparison chain: $\epsilon_{\text{extra}} = \epsilon_{\text{class}} + \epsilon_{\text{route}} + \epsilon_{\text{dist}}$. If the model-class bound already includes deployed routing, do not add a second routing penalty. Closed-loop occupancy changes are handled below, not automatically charged again as a one-step distribution discrepancy. The resulting bounds are

$$\begin{aligned} |r_{\text{risk}}| &\le \epsilon_{\text{risk}} := \epsilon_b + \epsilon_{\text{orient}} + \epsilon_{\text{pair}} + \epsilon_2 + \epsilon_{\text{cost}} + \epsilon_{\text{extra}}, \\ d_{\text{risk}} &\ge \underline{g}(q) - \epsilon_{\text{extra}}. \end{aligned} \tag{103}$$

The second bound already incorporates δ and does not subtract ϵ_2 again.

G.5 Finite-Horizon Connection to Planning Success

Shared protocol and success event. Let H be the number of actual environment transitions. The internal MPC lookahead is a separate, also matched parameter. Write the common planner as $\mathbf{a}_t^{(M)} = \mathrm{Plan}_u(f_M, \mathsf{h}_t^{(M)}, \mathbf{g}, \zeta_t)$, where $\mathsf{h}_t^{(M)}$ includes all required history. The same algorithm and paired randomness do not imply identical actions after replacing f_M. Conditional on $\mathbf{z}, \mathbf{g}, c$, consider the terminal success event

$$\|\mathbf{z}_H^{(M)} - \mathbf{g}\| \le a(\mathbf{z}, \mathbf{g}), \quad a(\mathbf{z}, \mathbf{g}) > 0, \qquad p_{M,H} = \mathbb{E}_{\mathbf{z}\sim\nu_0,\mathbf{g}|\mathbf{z}} p_{M,H}(\mathbf{z}, \mathbf{g}). \tag{104}$$

Here and below, $M \in \{\mathrm{R}, \mathrm{G}\}$, the configuration argument is suppressed, and $\Delta p_H := p_{\mathrm{R},H} - p_{\mathrm{G},H}$. The derivation below concerns this event. Other task predicates require their own monotonic reference event and boundary control. Shared encoding and normalization provide common coordinates; they do not establish equal error directions, distributions, or closed-loop derivatives.

Common reference and explicit trajectory residual. Fix $\mathbf{z}, \mathbf{g}, c$. Choose a common nominal trajectory $\bar{\mathbf{z}}_t$, with $\bar{\mathbf{z}}_0 = \mathbf{z}$, and write $\mathbf{e}_t^{(M)} = \mathbf{z}_t^{(M)} - \bar{\mathbf{z}}_t$. A local expansion, with all neglected terms retained, takes the form

$$\mathbf{e}_{t+1}^{(M)} = \mathbf{F}_t^{(M)} \cdot \mathbf{e}_t^{(M)} + \boldsymbol{\eta}_t^{(M)} + \mathbf{n}_t^{(M)}. \tag{105}$$

Choose reference matrices $\mathbf{F}_t^0$, deterministic conditional on $\mathbf{z}, \mathbf{g}, c$. Also choose a common joint law for the disturbance sequence $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{H-1}$. The law may be non-Gaussian and temporally correlated. For the actual one-step risk $v_M(\mathbf{z}; c) = \mathbb{E}[\|\mathbf{y} - f_M(\mathbf{z}, \mathbf{a}_0)\|^2 \mid \mathbf{z}_0 = \mathbf{z}]$, define

$$\mathbf{e}_{t+1}^{(M,0)} = \mathbf{F}_t^0 \cdot \mathbf{e}_t^{(M,0)} + \sqrt{v_M(\mathbf{z}; c)} \cdot \boldsymbol{\xi}_t, \qquad \mathbf{e}_0^{(M,0)} = \mathbf{0}. \tag{106}$$

Each actual process is coupled to this reference while preserving both marginal laws. The reference and its error budgets should have an independent modeling or statistical basis; defining a residual alone does not show that it is small. Set

$$\begin{aligned} \boldsymbol{\rho}_t^{(M)} &= (\mathbf{F}_t^{(M)} - \mathbf{F}_t^0) \cdot \mathbf{e}_t^{(M)} + (\boldsymbol{\eta}_t^{(M)} - \sqrt{v_M} \cdot \boldsymbol{\xi}_t) + \mathbf{n}_t^{(M)}, \\ \boldsymbol{\Phi}_{H,t+1}^0 &= \mathbf{F}_{H-1}^0 \cdot \ldots \cdot \mathbf{F}_{t+1}^0, \qquad \mathbf{u}_H = \sum_{t=0}^{H-1} \boldsymbol{\Phi}_{H,t+1}^0 \cdot \boldsymbol{\xi}_t, \\ \mathbf{r}_{M,H} &= (\bar{\mathbf{z}}_H - \mathbf{g}) + \sum_{t=0}^{H-1} \boldsymbol{\Phi}_{H,t+1}^0 \cdot \boldsymbol{\rho}_t^{(M)}. \end{aligned} \tag{107}$$

Empty products equal the identity. Unrolling the recursions proves

$$\mathbf{z}_H^{(M)} - \mathbf{g} = \sqrt{v_M(\mathbf{z}; c)} \cdot \mathbf{u}_H + \mathbf{r}_{M,H}. \tag{108}$$

If $\mathbb{E}\|\boldsymbol{\rho}_t^{(M)}\| \le \epsilon_{M,t}^{\rho}$, then

$$\mathbb{E}\|\mathbf{r}_{M,H}\| \le \epsilon_{M,H}^{\rho} := \|\bar{\mathbf{z}}_H - \mathbf{g}\| + \sum_{t=0}^{H-1} \|\boldsymbol{\Phi}_{H,t+1}^0\|_{\mathrm{op}} \cdot \epsilon_{M,t}^{\rho}. \tag{109}$$

Success-boundary control. The reference success probability is

$$p_{M,H}^0(\mathbf{z}, \mathbf{g}) = \Pr(\|\sqrt{v_M} \cdot \mathbf{u}_H\| \le a(\mathbf{z}, \mathbf{g})).$$

Define its population average and branch gap by

$$p_{M,H}^0 := \mathbb{E}_{\mathbf{z}\sim\nu_0,\mathbf{g}|\mathbf{z}} p_{M,H}^0(\mathbf{z}, \mathbf{g}), \qquad \Delta p_H^0 := p_{\mathrm{R},H}^0 - p_{\mathrm{G},H}^0. \tag{110}$$

Define its boundary mass as

$$\beta_{M,H}(\epsilon; \mathbf{z}, \mathbf{g}) = \Pr\big(\big|\|\sqrt{v_M} \cdot \mathbf{u}_H\| - a(\mathbf{z}, \mathbf{g})\big| \le \epsilon\big). \tag{111}$$

For any $\epsilon > 0$, the success indicators can differ only if the reference lies within ϵ of the boundary or $\|\mathbf{r}_{M,H}\| > \epsilon$. Markov's inequality therefore proves

$$|p_{M,H}(\mathbf{z}, \mathbf{g}) - p_{M,H}^0(\mathbf{z}, \mathbf{g})| \le \beta_{M,H}(\epsilon; \mathbf{z}, \mathbf{g}) + \epsilon_{M,H}^{\rho}(\mathbf{z}, \mathbf{g})/\epsilon. \tag{112}$$

Let $\delta_{M,H}$ be the minimum of one and the infimum of this bound over $\epsilon > 0$, and set

$$\delta_H = \mathbb{E}_{\mathbf{z}\sim\nu_0,\mathbf{g}|\mathbf{z}}[\delta_{\mathrm{R},H}(\mathbf{z}, \mathbf{g}) + \delta_{\mathrm{G},H}(\mathbf{z}, \mathbf{g})]. \tag{113}$$

Then $|\Delta p_H - \Delta p_H^0| \le \delta_H$. When $\epsilon_{M,H}^{\rho} = 0$, the coupled events agree almost surely and we may set $\delta_{M,H} = 0$, even if there is an atom at the boundary.

Averaging local risk sensitivity. The common reference defines a nonincreasing curve

$$\varphi_{H,c,\mathbf{z}}(v) = \mathbb{E}_{\mathbf{g}|\mathbf{z}} \Pr(\|\sqrt{v} \cdot \mathbf{u}_H\| \le a(\mathbf{z}, \mathbf{g})), \qquad v \ge 0. \tag{114}$$

For $v > 0$, its inner probability is $F_{\|\mathbf{u}_H\|^2}(a(\mathbf{z}, \mathbf{g})^2/v)$; for $v = 0$ it equals one. Flat regions and discontinuities are allowed. Define

$$d_c(\mathbf{z}) = v_{\mathrm{G}}(\mathbf{z}; c) - v_{\mathrm{R}}(\mathbf{z}; c), \qquad \ell_{H,c}(\mathbf{z}) = \begin{cases} \dfrac{\varphi_{H,c,\mathbf{z}}(v_{\mathrm{R}}) - \varphi_{H,c,\mathbf{z}}(v_{\mathrm{G}})}{d_c(\mathbf{z})}, & d_c(\mathbf{z}) \ne 0, \\ 0, & d_c(\mathbf{z}) = 0. \end{cases} \tag{115}$$

Thus $\ell_{H,c} \ge 0$. Setting it to zero at tied local risks makes the convention explicit; no pointwise strictly positive derivative is required.

Assumption 5 (Controlled finite-horizon connection and sensitivity)**.** *For the success event under consideration, the common reference coupling has an independently bounded $\epsilon^{\rho}_{M,H}$ and boundary mass as above. The quantities $\ell_{H,c}$, d_c, and $\ell_{H,c} \cdot d_c$ are integrable, $\bar{\ell}_{H,c} := \mathbb{E}_{\nu_0} \ell_{H,c} \in (0, \infty)$, and an independently justified budget satisfies*

$$|\operatorname{Cov}_{\nu_0}(\ell_{H,c}, d_c)| \le s_c \cdot \bar{\ell}_{H,c} \cdot \epsilon_{\mathrm{plan},H}. \tag{116}$$

These conditions apply configuration by configuration; uniform versions should be stated explicitly when claiming a guarantee for a whole class.

The secant definition and covariance identity yield

$$\begin{aligned} \Delta p^0_H &= \mathbb{E}_{\nu_0}[\ell_{H,c} \cdot d_c] \\ &= \bar{\ell}_{H,c} \cdot (\mathcal{R}_{\mathrm{G}} - \mathcal{R}_{\mathrm{R}}) + \operatorname{Cov}_{\nu_0}(\ell_{H,c}, d_c). \end{aligned} \tag{117}$$

Consequently, with

$$\alpha_H = s_c \cdot \bar{\ell}_{H,c}, \quad r_{\mathrm{plan},H} = \frac{\operatorname{Cov}(\ell_{H,c}, d_c)}{\alpha_H}, \quad r_{\mathrm{link},H} = \frac{\Delta p_H - \Delta p^0_H}{\alpha_H}, \tag{118}$$

we obtain the exact decomposition

$$\frac{\Delta p_H}{\alpha_H} = d_{\mathrm{risk}} + r_{\mathrm{plan},H} + r_{\mathrm{link},H} = g(q) + r_{\mathrm{risk}} + r_{\mathrm{plan},H} + r_{\mathrm{link},H}. \tag{119}$$

The bound on $r_{\mathrm{link},H}$ comes from trajectory coupling and boundary mass, rather than estimating the remainder from the target success gap itself.

What the common protocol eliminates. For identical model inputs and paired randomness, $u_{\mathrm{R}} = u_{\mathrm{G}}$ implies

$$\mathrm{Plan}_{u_{\mathrm{R}}}(f_0, \mathbf{x}, \mathbf{g}, \zeta) = \mathrm{Plan}_{u_{\mathrm{G}}}(f_0, \mathbf{x}, \mathbf{g}, \zeta).$$

Thus protocol mismatch is zero. Predictor-induced changes in actions and visited states remain in $\boldsymbol{\rho}^{(M)}_t$. Equal scalar one-step risk does not imply equal error directions, and matched control alone does not imply that a zero risk gap produces a zero success gap.

Propagation, boundary, and covariance budgets. Suppose the propagation matrices and errors satisfy

$$\begin{aligned} \|\mathbf{F}^{(M)}_t - \mathbf{F}^0_t\|_{\mathrm{op}} \le a_{M,t} \quad \text{almost surely}, \qquad & \mathbb{E}\|\mathbf{e}^{(M)}_t\| \le e_{M,t}, \\ \mathbb{E}\|\boldsymbol{\eta}^{(M)}_t - \sqrt{v_M} \cdot \boldsymbol{\xi}_t\| \le s_{M,t}, \qquad & \mathbb{E}\|\mathbf{n}^{(M)}_t\| \le n_{M,t}. \end{aligned}$$

Then

$$\epsilon^{\rho}_{M,t} \le a_{M,t} \cdot e_{M,t} + s_{M,t} + n_{M,t}. \tag{120}$$

If $\|\mathbf{F}^0_t\|_{\mathrm{op}} \le \ell$,

$$\epsilon^{\rho}_{M,H} \le \|\bar{\mathbf{z}}_H - \mathbf{g}\| + \sum_{t=0}^{H-1} \ell^{H-1-t} \cdot (a_{M,t} \cdot e_{M,t} + s_{M,t} + n_{M,t}). \tag{121}$$

Actual and nominal unit-norm states imply $e_{M,t} \le 2$. Fixed H makes this a finite sum, not necessarily a small one. If the reference radius has density at most $f_{\max}$ near the success threshold, then $\beta(\epsilon) \le 2 \cdot f_{\max} \cdot \epsilon$. When the optimizing ϵ lies in that neighborhood,

$$\delta_{M,H} \le \min\{1, 2 \cdot \sqrt{2 \cdot f_{\max} \cdot \epsilon^{\rho}_{M,H}}\}. \tag{122}$$

For finite second moments, Cauchy–Schwarz gives

$$\epsilon_{\mathrm{plan},H} = \frac{\sqrt{\mathrm{Var}(\ell_{H,c}) \cdot \mathrm{Var}(d_c)}}{s_c \cdot \bar{\ell}_{H,c}}. \tag{123}$$

If instead $\ell \in [\ell_-, \ell_+]$ and $d \in [d_-, d_+]$, a valid bound is $(\ell_+ - \ell_-) \cdot (d_+ - d_-)/(4 \cdot s_c \cdot \bar{\ell}_{H,c})$. Bounded state norms do not bound the sensitivity away from zero or make its variation small.

Optional stability justification. If the execution system obeys a state/action Lipschitz bound with constants ℓ_s, ℓ_a, and the common planner satisfies the local bound

$$\|\operatorname{Plan}_u(f_M, \mathbf{x}, \mathbf{g}, \zeta) - \operatorname{Plan}_u(f_0, \mathbf{x}', \mathbf{g}, \zeta)\| \le \ell_z \cdot \|\mathbf{x} - \mathbf{x}'\| + \ell_f \cdot \delta_M,$$

where δ_M is a uniform model discrepancy on the relevant domain, paired rollouts obey $d_{t+1} \le (\ell_s + \ell_a \cdot \ell_z) \cdot d_t + \ell_a \cdot \ell_f \cdot \delta_M$, $d_0 = 0$. Hence $d_H \le \ell_a \cdot \ell_f \cdot \delta_M \cdot \sum_{j=0}^{H-1} (\ell_s + \ell_a \cdot \ell_z)^j$. Here $\mathbf{x}$ includes the information needed to close the state dynamics. This is a sufficient local stability argument, not an automatic property of CEM or MPC. Average one-step MSE cannot replace δ_M without an additional bound, and trajectory proximity alone does not bound derivative differences.

Optional Gaussian shape and scale control. If the stacked innovations satisfy $\boldsymbol{\eta}_M \sim \mathcal{N}(\boldsymbol{\mu}_M, \boldsymbol{\Gamma}_M)$, choose a common centered Gaussian reference $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_0)$. The common-normal coupling gives

$$\mathbb{E}\|\boldsymbol{\eta}_M - \sqrt{v_M} \cdot \boldsymbol{\xi}\| \le \|\boldsymbol{\mu}_M\| + \|\boldsymbol{\Gamma}_M^{1/2} - \sqrt{v_M} \cdot \boldsymbol{\Gamma}_0^{1/2}\|_F. \tag{124}$$

Writing $\boldsymbol{\Gamma}_M = e^{\eta}_M \cdot \mathbf{Q}_M$ and $\boldsymbol{\Gamma}_0 = \gamma \cdot \mathbf{Q}_0$, with unit-trace shapes and positive trace energies, bounds the right side by

$$\chi_M = \|\boldsymbol{\mu}_M\| + \sqrt{e^{\eta}_M} \cdot \|\mathbf{Q}_M^{1/2} - \mathbf{Q}_0^{1/2}\|_F + |\sqrt{e^{\eta}_M} - \sqrt{\gamma \cdot v_M}|.$$

With $\mathbf{T}_H = [\boldsymbol{\Phi}^0_{H,1} \ \cdots \ \boldsymbol{\Phi}^0_{H,H}]$, the innovation contribution to the endpoint budget is at most $\|\mathbf{T}_H\|_{\mathrm{op}} \cdot \chi_M$. The covariance includes temporal correlations. A coupling of innovations can be extended to trajectories by their conditional distributions, preserving the actual innovation–trajectory dependence. Gaussianity is not needed elsewhere in the argument.

Scope for general task predicates. For a general finite-horizon success event, augment the environment state with the goal, necessary history, and any success flag. Let K_t^M be the actual closed-loop kernels, let v_t^{G} be the `Global` continuation success probability, and let d_t^{R} be the `Regional` occupancy. Telescoping the two kernel products gives

$$\Delta p_H = \sum_{t=0}^{H-1} \mathbb{E}_{\mathbf{x}_t \sim d_t^{\mathrm{R}}}[(K_t^{\mathrm{R}} - K_t^{\mathrm{G}}) \cdot v_{t+1}^{\mathrm{G}}(\mathbf{x}_t)]. \tag{125}$$

This identity shows why action selection, occupancy, and task sensitivity remain relevant under a shared planner. For non-ball success predicates, one may replace the reference curve by any independently justified nonincreasing family and prove a corresponding event-boundary bound; Equation (117) and the selection results then apply unchanged. A terminal latent-ball predicate should not be assumed equivalent to the actual benchmark predicate without justification.

G.6 Robust Selection and the Exact Special Case

Theorem 4 (Robust selection with explicit remainder budgets)**.** *Assume the nested risk comparison, its actual-model bridge, and Assumptions 3–5 in their two-sided form. Let* $\epsilon_H = \epsilon_{\mathrm{risk}} + \epsilon_{\mathrm{plan},H} + \delta_H/\alpha_H$. *Then*

$$\left| \frac{\Delta p_H}{\alpha_H} - (q - \kappa) \right| \le \epsilon_H. \tag{126}$$

If $|q - \kappa| > \epsilon_H$, *then* $\operatorname{sign}(\Delta p_H) = \operatorname{sign}(q - \kappa)$. *If an estimated score satisfies* $|\widehat{q} - q| \le \epsilon_q$, *then*

$$|\widehat{q} - \kappa| > \epsilon_q + \epsilon_H \quad \Longrightarrow \quad \operatorname{sign}(\Delta p_H) = \operatorname{sign}(\widehat{q} - \kappa). \tag{127}$$

Proof. Equation (119) and the triangle inequality prove the bound. A perturbation smaller than the nonzero reference margin cannot change its sign, and $\alpha_H > 0$. For the estimated score add $q - \widehat{q}$ to the remainder. □

Proposition 3 (One-sided `Regional` success certificate)**.** *Assume Proposition 2, the actual-model risk bridge with budget ϵ_{extra}, and Assumption 5. If $|\widehat{q} - q| \le \epsilon_q$, let $q_- = [\widehat{q} - \epsilon_q]_+$. Then*

$$\frac{\Delta p_H}{\alpha_H} \ge [\sqrt{q_-} - \bar{\delta}]_+^2 - \bar{\kappa} - \epsilon_{\text{extra}} - \epsilon_{\text{plan},H} - \delta_H/\alpha_H. \tag{128}$$

A strictly positive right-hand side implies $p_{\text{R},H} > p_{\text{G},H}$.

Proof. Use the lower risk bound in Equation (103) in the first identity of Equation (119). Bound the covariance and coupling terms from below, and use monotonicity of $\underline{g}$ and $q \ge q_-$. □

When budgets are uniform and $\alpha_H \ge \alpha_{\min} > 0$, a convenient sufficient condition is

$$q_- > \left(\bar{\delta} + \sqrt{\bar{\kappa} + \epsilon_{\text{extra}} + \epsilon_{\text{plan},H} + \delta_H/\alpha_{\min}}\right)^2. \tag{129}$$

If its right side is at least $1/2$, the sufficient condition is empty. Failure to certify `Regional` never establishes that `Global` is better.

Notation for the main-text selection theorem. The theoretical threshold and total deviation bound in the original Worst-Case Partition Gain score units are

$$\tau_\star = 2 \cdot \kappa, \qquad \epsilon(c) = 2 \cdot \epsilon_H(c). \tag{130}$$

Thus $|b(\Pi) - \tau_\star| > \epsilon(c)$ is exactly the condition $|q - \kappa| > \epsilon_H(c)$.

Proof of Theorem 3. Apply Theorem 4 and Equation (130), using $b(\Pi) = 2 \cdot q$. □

Corollary 2 (Exact threshold on a restricted comparison class)**.** *If the common reference cost is $\kappa \in (0, 1/2)$, $\alpha_H(c) > 0$, and the total remainder $r_{\text{risk}} + r_{\text{plan},H} + r_{\text{link},H}$ vanishes for every configuration in a specified class, then*

$$p_{\text{R},H} > p_{\text{G},H} \iff b(\Pi) > \tau_\star, \qquad \tau_\star = 2 \cdot \kappa. \tag{131}$$

Proof. The decomposition reduces to $\Delta p_H = \alpha_H \cdot (q - \kappa)$. □

Zero total remainder is an additional structural restriction, not a consequence of matched protocols. The controlled-margin theorem above does not require it. In particular, replacing the original common-zero argument by this explicit special case prevents matched control from being used to infer $H_u(0) = 0$ without a task connection.

Corollary 3 (Transfer under common budgets)**.** *The same theoretical threshold $2 \cdot \kappa$ applies to any partitioning method, region count, or matched planner for which the two-sided budgets hold and the configuration satisfies $|q - \kappa| > \epsilon_H(c)$.*

Proof. Apply Theorem 4 separately to each such configuration. □

These are conditional transfer statements. Both the excess-risk cost and the success connection can vary with K, the representation, and the planner. Uniform budgets require independent justification. All formulas so far condition on a trained candidate pair; averaging over training seeds requires averaging the full identity, including α_H.

G.7 Finite-Sample Dynamics Selection

Let the available latent transitions be

$$\mathcal{D}_N := \{(\mathbf{z}_n, \mathbf{a}_n, \mathbf{z}'_n)\}_{n=1}^N, \qquad \Delta\mathbf{z}_n := \mathbf{z}'_n - \mathbf{z}_n. \tag{132}$$

Let $\mathcal{S}$ be the prespecified set of partition draws or seeds. For each $\ell \in \mathcal{S}$, assign

$$r_n^{(\ell)} := \Pi^{(\ell)}(\mathbf{z}_n). \tag{133}$$

Let

$$\widehat{\boldsymbol{\mu}}_a := \frac{1}{N} \cdot \sum_{n=1}^N \mathbf{a}_n, \qquad \widehat{\boldsymbol{\Sigma}}_a := \frac{1}{N} \cdot \sum_{n=1}^N (\mathbf{a}_n - \widehat{\boldsymbol{\mu}}_a) \cdot (\mathbf{a}_n - \widehat{\boldsymbol{\mu}}_a)^\top. \tag{134}$$

Within each region, the selection procedure estimates the affine action response by ridge regression:

$$(\widehat{\mathbf{b}}_k^{(\ell)}, \widehat{\mathbf{J}}_k^{(\ell)}) := \arg\min_{\mathbf{b}_0, \mathbf{J}} \sum_{n: r_n^{(\ell)} = k} \|\Delta\mathbf{z}_n - \mathbf{b}_0 - \mathbf{J} \cdot (\mathbf{a}_n - \widehat{\boldsymbol{\mu}}_a)\|_2^2 + \lambda \cdot \|\mathbf{J}\|_F^2. \tag{135}$$

It then forms

$$\widehat{\mathbf{C}}_k^{(\ell)} := \widehat{\mathbf{J}}_k^{(\ell)} \cdot \widehat{\boldsymbol{\Sigma}}_a \cdot (\widehat{\mathbf{J}}_k^{(\ell)})^\top, \qquad \widehat{e}_k^{(\ell)} := \mathrm{tr}(\widehat{\mathbf{C}}_k^{(\ell)}). \tag{136}$$

The corresponding empirical centered Gaussian response distribution is

$$\widehat{\mathcal{P}}_k^{(\ell)} := \mathcal{N}\Big(\mathbf{0}, \widehat{\mathbf{C}}_k^{(\ell)}\Big). \tag{137}$$

The procedure then evaluates

$$\widehat{b}_{ij}^{(\ell)} := \frac{W_2^2\Big(\widehat{\mathcal{P}}_i^{(\ell)}, \widehat{\mathcal{P}}_j^{(\ell)}\Big)}{\widehat{e}_i^{(\ell)} + \widehat{e}_j^{(\ell)}}. \tag{138}$$

The empirical Worst-Case Partition Gain score uses a fixed order of aggregation: first take the weakest pair within each draw, then average across the prespecified draw set $\mathcal{S}$:

$$\widehat{b}_K := \frac{1}{|\mathcal{S}|} \cdot \sum_{\ell \in \mathcal{S}} \min_{1 \le i < j \le K} \widehat{b}_{ij}^{(\ell)}. \tag{139}$$

The seed set and region count are specified before observing the results.

The operational selection is

$$\widehat{g}_{\mathrm{sel}} := \begin{cases} \texttt{Regional}, & \widehat{b}_K > \tau_{\mathrm{cal}}, \\ \texttt{Global}, & \widehat{b}_K \le \tau_{\mathrm{cal}}, \end{cases} \qquad \tau_{\mathrm{cal}} = 0.508947854338762. \tag{140}$$

The corresponding operational threshold in the q-scale is

$$\kappa_{\mathrm{cal}} := \frac{\tau_{\mathrm{cal}}}{2} = 0.254473927169381. \tag{141}$$

The threshold is calibrated on the LeWM $K = 4$ calibration configurations and then frozen for every cross-K and cross-backbone evaluation.

G.7.1 From Estimation Error to Correct Selection

We now show when response-estimation errors leave the selection outcome unchanged. We first fix one matched configuration and its candidate partition Π, suppressing the draw index ℓ. The population and estimated responses below refer to this same partition.

Recall the population response factor $\mathbf{A}_k = \mathbf{J}_k \cdot \boldsymbol{\Sigma}_a^{1/2}$, and define its estimate by

$$\widehat{\mathbf{A}}_k := \widehat{\mathbf{J}}_k \cdot \widehat{\boldsymbol{\Sigma}}_a^{1/2}, \qquad \widehat{\mathbf{C}}_k = \widehat{\mathbf{A}}_k \cdot \widehat{\mathbf{A}}_k^\top, \qquad \widehat{e}_k := \mathrm{tr}(\widehat{\mathbf{C}}_k). \tag{142}$$

For this fixed partition, write

$$\widehat{\mathcal{P}}_k := \mathcal{N}(\mathbf{0}, \widehat{\mathbf{C}}_k), \qquad \widehat{d}_{ij} := W_2(\widehat{\mathcal{P}}_i, \widehat{\mathcal{P}}_j), \qquad \widehat{b}_{ij} := \frac{\widehat{d}_{ij}^2}{\widehat{e}_i + \widehat{e}_j}, \qquad \widehat{b}(\Pi) := \min_{i<j} \widehat{b}_{ij}. \tag{143}$$

Using factor error accounts for uncertainty in both the `regional` Jacobian and the common action covariance. In particular,

$$\|\widehat{\mathbf{A}}_k - \mathbf{A}_k\|_F \le \|(\widehat{\mathbf{J}}_k - \mathbf{J}_k)\cdot \boldsymbol{\Sigma}_a^{1/2}\|_F + \|\widehat{\mathbf{J}}_k\|_F \cdot \|\widehat{\boldsymbol{\Sigma}}_a^{1/2} - \boldsymbol{\Sigma}_a^{1/2}\|_{\mathrm{op}}, \tag{144}$$

where $\|\cdot\|_{\mathrm{op}}$ is the operator norm, or largest singular value. The first term is the Jacobian error weighted by the reference action distribution; the second is the additional error from estimating that distribution through its covariance. The second term vanishes when the reference covariance is known and used directly.

Theorem 5 (Worst-Case Partition Gain score stability and selection with controlled errors)**.** *For a fixed partition, suppose*

$$\max_k \|\widehat{\mathbf{A}}_k - \mathbf{A}_k\|_F \le \eta, \quad \min_{i<j}(e_i + e_j) \ge e_{\min} > 0, \quad \sqrt{2}\cdot\eta < \sqrt{e_{\min}}. \tag{145}$$

Then estimated pair energies are positive, and with $\epsilon_{\mathrm{score}} = \min\{1, 8\cdot\eta/\sqrt{e_{\min}}\}$,

$$\max_{i<j} |\widehat{b}_{ij} - b_{ij}| \le \epsilon_{\mathrm{score}}, \qquad |\widehat{b}(\Pi) - b(\Pi)| \le \epsilon_{\mathrm{score}}. \tag{146}$$

If the two-sided risk and planning budgets also hold, let $\epsilon_q = \epsilon_{\mathrm{score}}/2$ *and suppose* $|\tau_{\mathrm{cal}}/2 - \kappa| \le \epsilon_{\mathrm{cal}}$. *The operational decision is correct whenever*

$$|\widehat{q} - \tau_{\mathrm{cal}}/2| > \epsilon_q + \epsilon_{\mathrm{cal}} + \epsilon_H, \qquad \widehat{q} = \widehat{b}(\Pi)/2. \tag{147}$$

Calibration on a small set of configurations does not itself establish ϵ_{cal} *or the structural budgets.*

Proof. Write $e = e_i + e_j$, $\widehat{e} = \widehat{e}_i + \widehat{e}_j$, $d = W_2(\mathcal{P}_i, \mathcal{P}_j)$, and $\widehat{d} = W_2(\widehat{\mathcal{P}}_i, \widehat{\mathcal{P}}_j)$. The reverse triangle inequality for the stacked factors gives $|\sqrt{\widehat{e}} - \sqrt{e}| \le \sqrt{2}\cdot\eta$, so $\widehat{e} > 0$. The Procrustes form gives $|\widehat{d} - d| \le 2\cdot\eta$. Since $\widehat{d}/\sqrt{\widehat{e}} \le 1$,

$$|\sqrt{\widehat{b}_{ij}} - \sqrt{b_{ij}}| \le \frac{|\widehat{d} - d|}{\sqrt{e}} + \frac{\widehat{d}}{\sqrt{\widehat{e}}}\cdot\frac{|\sqrt{\widehat{e}} - \sqrt{e}|}{\sqrt{e}} \le \frac{(2+\sqrt{2})\cdot\eta}{\sqrt{e_{\min}}}.$$

Multiply by $\sqrt{\widehat{b}_{ij}} + \sqrt{b_{ij}} \le 2$ and use the range $[0,1]$. Taking minima cannot increase a uniform pairwise error, proving the score bound. Finally,

$$\left|\frac{\Delta p_H}{\alpha_H} - (\widehat{q} - \tau_{\mathrm{cal}}/2)\right| \le \epsilon_H + \epsilon_q + \epsilon_{\mathrm{cal}}.$$

The stipulated nonzero margin determines its sign. □

A data-dependent alternative for score error. If $\|\widehat{\mathbf{A}}_k - \mathbf{A}_k\|_F \le \eta_k$, define $t_{ij} = \eta_i + \eta_j$ and

$$u_{ij} = (2\cdot\widehat{d}_{ij} + t_{ij})\cdot t_{ij}, \qquad v_{ij} = \sum_{k\in\{i,j\}} \eta_k\cdot(2\cdot\|\widehat{\mathbf{A}}_k\|_F + \eta_k). \tag{148}$$

These bound the squared-distance and pair-energy errors. If true and estimated pair energies are both at least $e_{\min}$, then

$$|\widehat{q} - q| \le \epsilon_q := \min\left\{\tfrac{1}{2}, \frac{\max_{i<j}(u_{ij} + v_{ij})}{2\cdot e_{\min}}\right\}. \tag{149}$$

Indeed, expand the difference of $\widehat{d}^2/(2\cdot\widehat{e})$ and $d^2/(2\cdot e)$, and use $d^2 \le e$; the minimum over pairs is one-Lipschitz in the uniform norm. For the one-sided success certificate, use $q_- = [\widehat{q} - \epsilon_q]_+$.

Identifying the reference response from data. The risk derivation uses a product start/action reference. In logged data, actions may depend on the state or region, and ridge regression includes regularization bias. Factor error bounds should include the consequences of this sampling difference, non-affine responses, covariance estimation, and regularization. A pooled action covariance alone does not make the logged conditional action distributions identical. Conditional mean-zero noise, excitation, suitable reference weighting, or another explicit identification argument is needed for the chosen bound. For a partition estimated from the same data, the bounds are assumed to hold for that data-dependent partition; one may use independent estimation data or a valid uniform analysis. We do not infer such guarantees from the algebraic perturbation bound alone.

Averaging across partition draws. For prespecified draws $\ell \in \mathcal{S}$, define $q_\ell = b(\Pi^{(\ell)})/2$, $\bar{q} = |\mathcal{S}|^{-1} \cdot \sum_\ell q_\ell$, and $\widehat{\bar{q}} = \widehat{b}_K/2$. If the score bounds hold simultaneously, then

$$|\widehat{\bar{q}} - \bar{q}| \le \bar{\epsilon}_q := |\mathcal{S}|^{-1} \cdot \sum_\ell \epsilon_{q,\ell}. \tag{150}$$

This certifies score estimation, not an average success ordering by itself. For an evaluation that averages the corresponding draw-specific branch success probabilities, write $\Delta p_\ell = \alpha_\ell \cdot (q_\ell - \kappa + r_\ell^{\mathrm{tot}})$, with $|r_\ell^{\mathrm{tot}}| \le \epsilon_{H,\ell}$. Uniform averaging over draws gives

$$\overline{\Delta p} = \bar{\alpha} \cdot (\bar{q} - \kappa) + \mathrm{Cov}_\ell(\alpha_\ell, q_\ell) + \overline{\alpha_\ell \cdot r_\ell^{\mathrm{tot}}}. \tag{151}$$

Consequently a sufficient condition for the implemented average-score decision to match the average success ordering is

$$|\widehat{\bar{q}} - \tau_{\mathrm{cal}}/2| > \bar{\epsilon}_q + \epsilon_{\mathrm{cal}} + \frac{|\mathrm{Cov}_\ell(\alpha_\ell, q_\ell)| + \overline{\alpha_\ell \cdot \epsilon_{H,\ell}}}{\bar{\alpha}}. \tag{152}$$

Alternatively, certifying each draw on the same side with the single-draw operational margin suffices. For the one-sided route, let L_ℓ be the right side of Equation (128) for draw ℓ; then $\overline{\Delta p} \ge \overline{\alpha_\ell \cdot L_\ell}$. This expectation should correspond to the partitions actually used in the evaluated models. Merely averaging auxiliary score draws does not certify a different deployed partition.

Probability and training-seed guarantees. If score, cost, model, propagation, boundary, and sensitivity budgets are probabilistic, the conclusions hold on their joint event. Apply a union bound to valid simultaneous component guarantees; independence is not required for that step. Averaging over training seeds also requires averaging the complete success identity, since α_H and its remainders can depend on the trained candidates. No numerical structural budgets are claimed to have been measured in the reported experiments.