

Recursive Transport Topology Detects Structural Regime Transitions Associated with Behavioral Failure in World Models

Jeffery Scott Allbright

scott.allbright@yahoo.com

Alla Technologies, Inc

Research Article

Keywords: world models, latent geometry, structural diagnostics, regime transition, representation stability, model-predictive control, Recursive Indexing

Posted Date: July 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10295344/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare potential competing interests as follows: Jeffery Scott Allbright reports pending intellectual property related to recursive indexing, transport topology, and representation architectures, and is a co-founder of Alla Technologies, Inc. and Forma Substrate, Inc., which may commercialize these ideas. The author declares no other competing interests.

Abstract

World models are typically evaluated through prediction accuracy or downstream task success, yet these measures do not directly test whether an internal representation has crossed into a structurally unstable operating regime. We evaluate Recursive Indexing (RIX), a fixed-reference analysis of latent transport topology, in a pretrained LeWM PushT world model exposed to controlled Gaussian observation stress. Progressive stress produced a sharp nonlinear reorganization of latent neighborhoods, fixed-reference basins, transition pathways, and pathway entropy while conventional pointwise latent displacement remained small. The dominant structural breakpoint localized to the 0.05–0.10 stress interval across all 320 analytical configurations. In an independent closed-loop calibration, task success declined from 80% at stress 0.00 to 20% at 0.05 and 0% at 0.10 and above. These results support a regime-level interpretation of world-model reliability: latent representations can remain pointwise close while their transport organization has already changed substantially, and the resulting structural transition can align with behavioral collapse.

1. Introduction

World models learn latent representations that support prediction, planning, and control [3–5]. Their reliability is usually judged through prediction error, reconstruction quality, uncertainty, or downstream task success. These measures are essential, but they do not directly test whether the organization of the latent state space remains suitable for navigation under perturbation.

This distinction matters because pointwise proximity does not imply structural preservation. A latent state can move only slightly while losing its original neighbors, changing basin membership, or entering transition pathways that were absent from the clean reference regime. For planning systems, these relational changes may matter before large vector displacement is visible.

We test whether a fixed-reference structural analysis can identify a reproducible operating-regime transition in a pretrained world model under controlled observation stress. Recursive Indexing (RIX), developed in prior work on transport organization and transfer structure [1, 2], is used here to measure latent transport topology relative to clean reference geometry. The structural transition is then compared with an independently collected closed-loop behavioral calibration.

The study addresses three questions:

- Does progressive observation stress produce a nonlinear transition in latent transport structure even when pointwise latent displacement remains small?
- Is the inferred structural transition robust to reasonable analytical choices?
- Does the independently identified structural transition align with the onset of closed-loop behavioral collapse?

The experimental sequence was deliberately ordered to separate structural discovery from behavioral interpretation. The offline stress analysis and robustness grid were completed before the closed-loop success curve was used to interpret the transition interval.

Contributions

- A fixed-reference world-model stress experiment showing large relational reorganization under very small pointwise latent displacement.
- A 320-configuration robustness analysis demonstrating stable breakpoint localization across projection and RIX resolution choices.
- A closed-loop calibration showing that task success collapses in the same stress region identified structurally.
- A regime-level formulation of world-model reliability that complements pointwise latent distance and downstream performance.

2. Methods

2.1 Model, environment, and observation stress

Experiments used a pretrained LeWorldModel (LeWM) PushT world model [3] and the associated stable-worldmodel model-predictive-control evaluation stack [4, 5]. PushT is a planar pushing benchmark used in visuomotor control research [6]. Observation stress was introduced as additive Gaussian noise on current observation pixels. Goal images remained clean, model parameters were frozen, and environment dynamics were unchanged.

The stress ladder was:

$$S = \{0.00, 0.05, 0.10, 0.15, 0.20, 0.30\}.$$

Matched offline observations were used for the structural stress experiment. For each matched frame, a deterministic noise realization was scaled by the six stress amplitudes so that the ladder changed perturbation magnitude rather than perturbation identity.

2.2 Latent extraction and compact representation

The pretrained encoder produced 49,344-dimensional latent vectors. Direct full-dimensional matched-frame displacement was retained as a baseline. For graph and RIX analyses, latents were mapped through a fixed deterministic sparse random projection. The primary analysis used 64 dimensions, followed by robustness tests at 32, 64, 128, and 256 projected dimensions.

2.3 Fixed clean-reference RIX geometry

The central methodological rule was that structural reference geometry was fit only on the clean condition. A scaler, PCA coordinate system, and quantile basin thresholds were learned from stress 0.00 and frozen. Each stressed latent was then assigned to the same clean coordinate system and basin partition.

This avoids a major interpretive ambiguity: if the geometry were re-fit at every stress level, each stressed condition could appear internally organized in its own coordinate system. The fixed-reference design instead asks whether a stressed representation preserves the organization of the healthy operating regime.

2.4 Structural metrics

The analysis tracked complementary levels of organization:

- Neighbor retention: overlap between clean and stressed k-nearest-neighbor sets for matched frames.
- Matched basin retention: fraction of matched frames retaining the same fixed-reference RIX basin assignment.
- Clean-path admissibility: fraction of stressed transitions belonging to the set of transitions observed in the clean reference.
- Transition JS divergence: Jensen-Shannon divergence between stressed and clean transition distributions.
- Path entropy: Shannon entropy of the transition-count distribution.
- Top-10 pathway fraction: concentration of transition mass in the ten most frequent pathways.
- Task locality: physical-state, proprioceptive, action, and temporal distances among latent neighbors.

2.5 Breakpoint analysis

For each stress-response curve, a simple two-segment linear breakpoint scan compared candidate split intervals while requiring at least two stress points on each side. Breakpoint interpretation was used as a descriptive localization tool rather than as proof of a universal phase transition.

2.6 Robustness grid

The robustness replication crossed:

- projection dimensions: 32, 64, 128, 256;
- projection seeds: five deterministic seeds;
- RIX PCA dimensions: 3, 4, 5, 6;
- RIX quantile bins per dimension: 3, 4, 5, 6.

$4 \times 5 \times 4 \times 4 = 320$ analytical configurations.

2.7 Closed-loop behavioral calibration

Closed-loop evaluation used the same six-level stress ladder and a fixed CPU-safe CEM planner budget across all conditions. Ten episodes were evaluated per stress level. The objective was comparative calibration of success probability under a constant planner budget, not reproduction of a maximum-compute benchmark.

3. Results

3.1 Small pointwise displacement accompanies large structural reorganization

Stress	Cosine displacement	Neighbor retention	Basin retention	Path admissibility	Transition JS
0.00	0.00000	100.0%	100.0%	100.0%	0.000
0.05	0.00087	19.9%	8.0%	6.1%	0.921
0.10	0.00333	5.5%	4.0%	2.0%	0.970
0.15	0.00711	3.4%	2.6%	1.0%	0.986
0.20	0.01204	2.7%	2.0%	0.4%	0.993
0.30	0.02470	2.4%	1.2%	0.6%	0.989

The strongest contrast appeared at the first nonzero stress level. At stress 0.05, mean cosine displacement from clean was only 0.00087, yet the representation retained 19.9% of its clean nearest neighbors, 8.0% of matched basin assignments, and 6.1% of clean-admissible transitions. Transition-distribution JS divergence had already increased to 0.921.

The response was therefore not well described as simple vector displacement. A low-amplitude perturbation produced large relational and transport reorganization while pointwise displacement remained numerically small.

3.2 Path organization disperses rapidly

The clean representation had path entropy of 6.85 bits and a top-10 pathway fraction of 24.65%. At stress 0.05, path entropy increased to 8.81 bits while the top-10 fraction fell to 5.45%. Thus the stressed representation did not merely move into different locations; its transitions became substantially more diffuse and less concentrated around dominant routes.

Task locality changed in the same direction. Mean physical-state locality distance increased from 0.575 in the clean condition to 2.224 at stress 0.05 and 3.146 at 0.10. Mean temporal neighbor gap increased from 1.27 steps in clean data to 18.26 at 0.05 and 31.77 at 0.10. Nearby latent states therefore ceased to correspond closely to nearby physical states and nearby moments in task progression.

3.3 The transition interval is robust across 320 analytical configurations

The full robustness grid reproduced the low-stress transition across projection dimensions, random projection seeds, RIX PCA dimensions, and basin resolutions.

Metric	Configurations selecting 0.05–0.10 breakpoint	Agreement
Neighbor retention	320 / 320	100.0%
Matched basin retention	320 / 320	100.0%
Clean-path admissibility	320 / 320	100.0%
Path entropy	320 / 320	100.0%
Top-10 pathway fraction	320 / 320	100.0%
Transition JS divergence	316 / 320	98.75%

Across the robustness grid, mean stress-0.05 values were 25.6% neighbor retention, 10.7% matched basin retention, 9.4% clean-path admissibility, and 0.875 transition JS divergence. The exact metric values varied with analytical configuration, but the qualitative transition and its localization were highly stable.

3.4 Behavioral collapse occurs in the same transition regime

Stress	Successful episodes	Success rate
0.00	8 / 10	80%
0.05	2 / 10	20%
0.10	0 / 10	0%
0.15	0 / 10	0%
0.20	0 / 10	0%
0.30	0 / 10	0%

Closed-loop success declined from 80% at stress 0.00 to 20% at 0.05 and 0% at 0.10 and above. The independently identified structural breakpoint interval was 0.05–0.10. Thus the dominant structural

transition and the onset of behavioral collapse localized to the same stress regime.

Figure 3 is the central synthesis of the study. The structural preservation index collapses across the same 0.05–0.10 stress interval in which closed-loop success falls from 20% to 0%, while the clean condition remains structurally intact and behaviorally successful. The alignment is regime-level rather than episode-level: the figure shows that the independently identified structural breakpoint and the behavioral collapse occupy the same operating region.

4. Discussion

The central finding is that pointwise stability and structural stability can separate. Under low observation stress, latent vectors remained close to their clean counterparts by cosine distance, yet neighborhood preservation, basin identity, admissible transport, and pathway concentration changed sharply. In this setting, the representation became structurally unfamiliar before it became far away in the ordinary pointwise sense.

The result is strengthened by the stability of the transition itself. Across 320 analytical configurations, the dominant breakpoint remained in the 0.05–0.10 interval for all configurations of neighborhood retention, basin retention, path admissibility, path entropy, and pathway concentration, and for 316 of 320 transition-divergence configurations. The persistence of a common breakpoint across projection dimensions, seeds, PCA dimensions, and basin resolutions argues that the transition reflects the stressed representation rather than a fragile analytical choice.

The independent closed-loop calibration gives this structural transition operational significance. Task success fell from 80% at stress 0.00 to 20% at 0.05 and 0% at 0.10 and above, placing behavioral collapse in the same regime identified structurally. The study does not establish topology as the cause of failure; it establishes a reproducible correspondence between a structural regime transition and loss of task performance.

4.1 Reliability as an operating-regime property

These findings support treating reliability as an operating-regime property. Pointwise diagnostics measure how far a representation moves; structural diagnostics test whether the relations that make prediction and planning navigable remain organized. The two views are complementary, but they answer different questions. Here, transport topology reorganized dramatically while pointwise displacement remained small, showing that latent distance alone can understate a meaningful change in operating condition.

4.2 Complementarity with conventional diagnostics

RIX is best interpreted as a complementary diagnostic layer. Prediction error, uncertainty estimation, density estimation, and downstream evaluation measure different aspects of reliability. Fixed-reference

transport topology adds a distinct question: does the latent system still occupy a structurally familiar regime? A practical reliability stack can combine these signals rather than treating any single metric as sufficient.

4.3 What this study does not establish

This study establishes regime detection and behavioral correspondence, not deterministic single-rollout prediction. The supported claim is that fixed-reference transport topology identifies a reproducible structural transition that aligns with behavioral collapse at the regime level. Prospective prediction of individual trajectory recovery or failure is a separate problem and requires separate evaluation.

5. Limitations

- The experiments use one pretrained world-model architecture and one control task family.
- The stress mechanism is Gaussian observation corruption; other perturbations may produce different structural responses.
- Closed-loop behavioral calibration used 10 episodes per stress level, sufficient for regime calibration but not precise probability estimation.
- The closed-loop planner used a fixed CPU-safe CEM budget rather than the highest-compute configuration associated with GPU execution.
- The breakpoint analysis is descriptive and should not be interpreted as proof of a universal thermodynamic phase transition.
- Cross-model, cross-task, and cross-stress replication are required before general claims about world-model reliability can be made.

Conclusion

Controlled observation stress produced a sharp, reproducible transition in the transport organization of a pretrained world model. The transition was visible as rapid loss of neighborhood preservation, fixed-reference basin identity, and clean-path admissibility, together with strong transition divergence and pathway dispersion, while pointwise latent displacement remained small.

The same breakpoint interval persisted across 320 analytical configurations and aligned with a sharp decline in closed-loop task success. Together, these results show that latent transport organization can change substantially while ordinary distance metrics still indicate only small displacement, and that fixed-reference transport topology can expose the operating-regime transition associated with behavioral collapse.

The central implication is simple: world-model reliability depends not only on where latent states are located, but on whether the structure that connects them remains organized.

Declarations

7. Reproducibility and Data Availability

All datasets and pretrained model resources used in this study are publicly available through the cited LeWorldModel, stable-worldmodel, and PushT resources [3-6]. The M17 program generated separate scripts and tabular outputs for latent stress extraction, compact projection, fixed-reference structural analysis, robustness replication, and closed-loop calibration. Processed datasets, analysis code, configuration manifests, summary CSV files, and figure-generation scripts will be made available on GitHub prior to formal publication. The repository link will be added in a revised version of this preprint.

8. Author Contributions

Jeffery Scott Allbright conceived the study, designed the analysis workflow, performed experiments, and wrote the manuscript draft.

9. Competing Interests

Jeffery Scott Allbright reports pending intellectual property related to recursive indexing, transport topology, and representation architectures, and is a co-founder of AIIA Technologies, Inc. and Forma Substrate, Inc., which may commercialize these ideas. The author declares no other competing interests.

References

1. Allbright JSA, Constraint-Weighted Transfer Model Predicts Cross-System Generalization in Protein Fitness Landscapes. Research Square preprint. 10.21203/rs.3.rs-9623030/v1
2. Allbright JS Recursive Transport Topology: Predictive and Causal Organization in Protein Fitness Landscapes. Research Square preprint. 10.21203/rs.3.rs-9889139/v1
3. Maes L, Le Lidec Q, Scieur D, LeCun Y, Balestrieri R (2026) LeWorldModel: Stable End-to-End Joint-Embedding Predictive Architecture from Pixels. arXiv:2603.19312
4. Maes L, Le Lidec Q, Haramati D, Massaudi N, Scieur D, LeCun Y, Balestrieri R (2026) stable-worldmodel-v1: Reproducible World Modeling Research and Evaluation. arXiv:2602.08968
5. Maes L, Le Lidec Q, Facury L, Massaudi N, Chaurasia A, Capuano F, Gao R, Gillin T, Haramati D, Scieur D, LeCun Y, Balestrieri R (2026) stable-worldmodel: A Platform for Reproducible World Modeling Research and Evaluation. arXiv:2605.21800
6. Chi C, Xu Z, Feng S, Cousineau E, Du Y, Burchfiel B, Tedrake R, Song S (2023) Diffus Policy: Visuomotor Policy Learn via Action Diffus arXiv :230304137

Figures

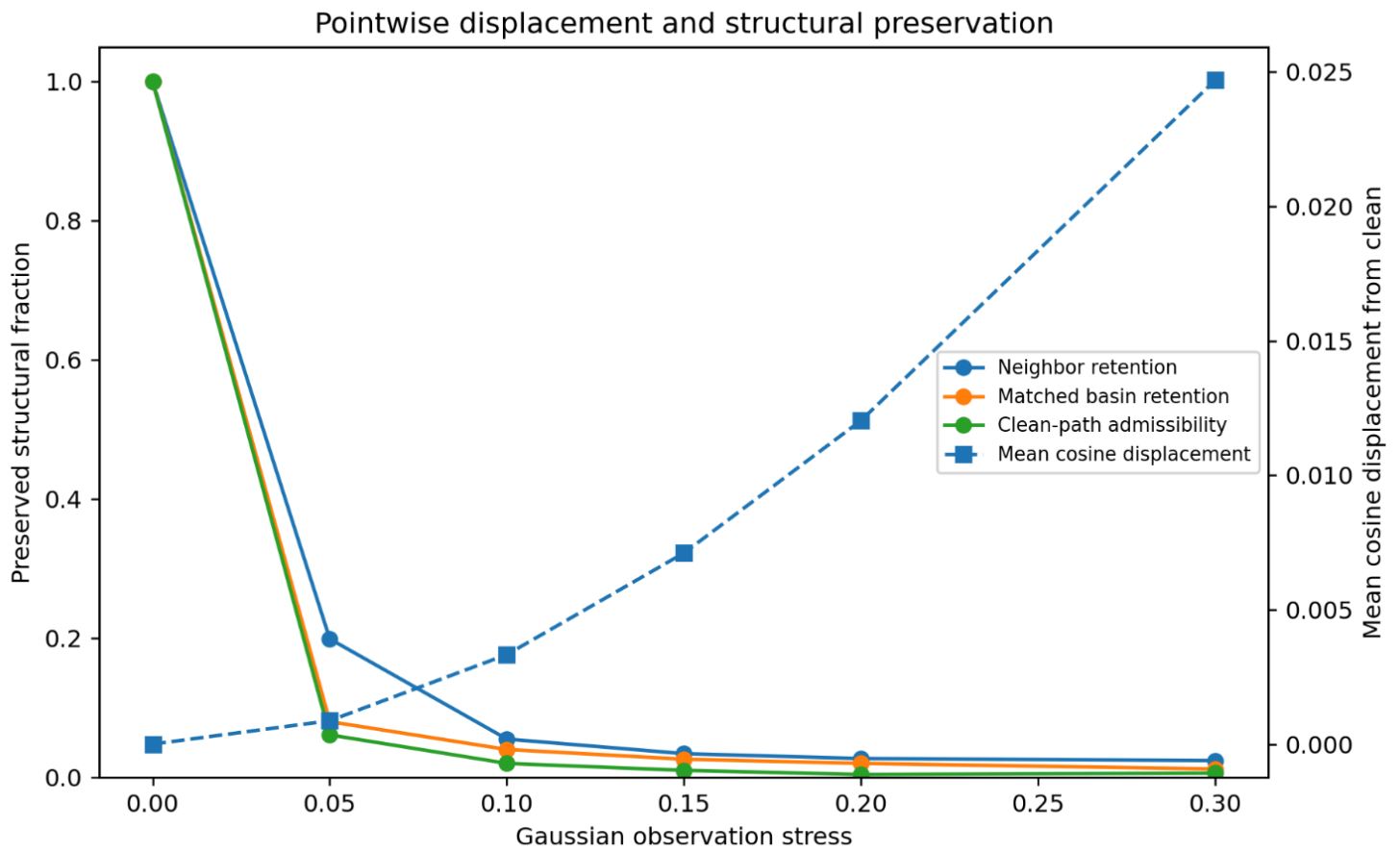


Figure 1

Pointwise latent displacement remains small while structural preservation collapses under low observation stress.

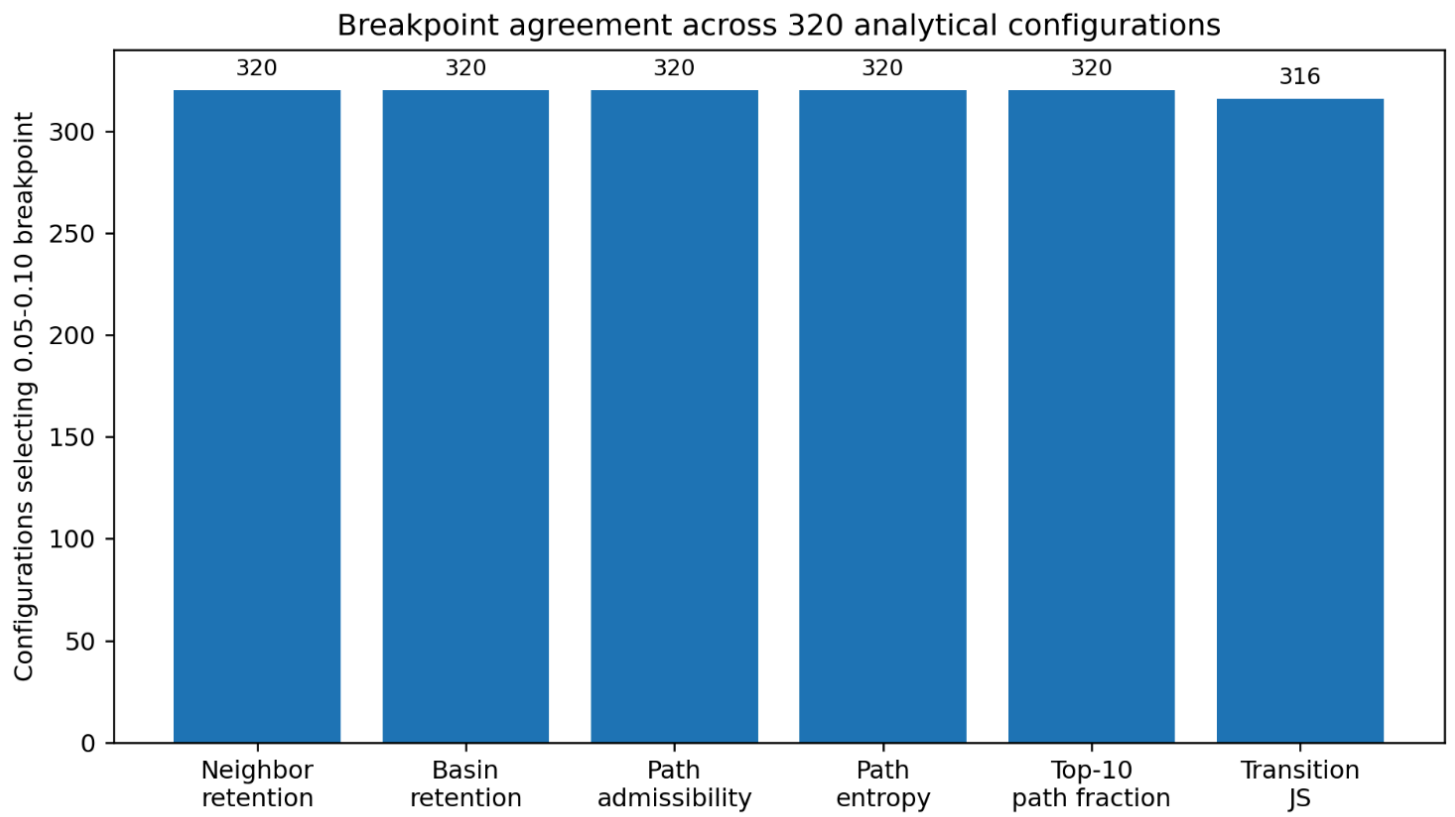


Figure 2

Breakpoint interval agreement across 320 analytical configurations.

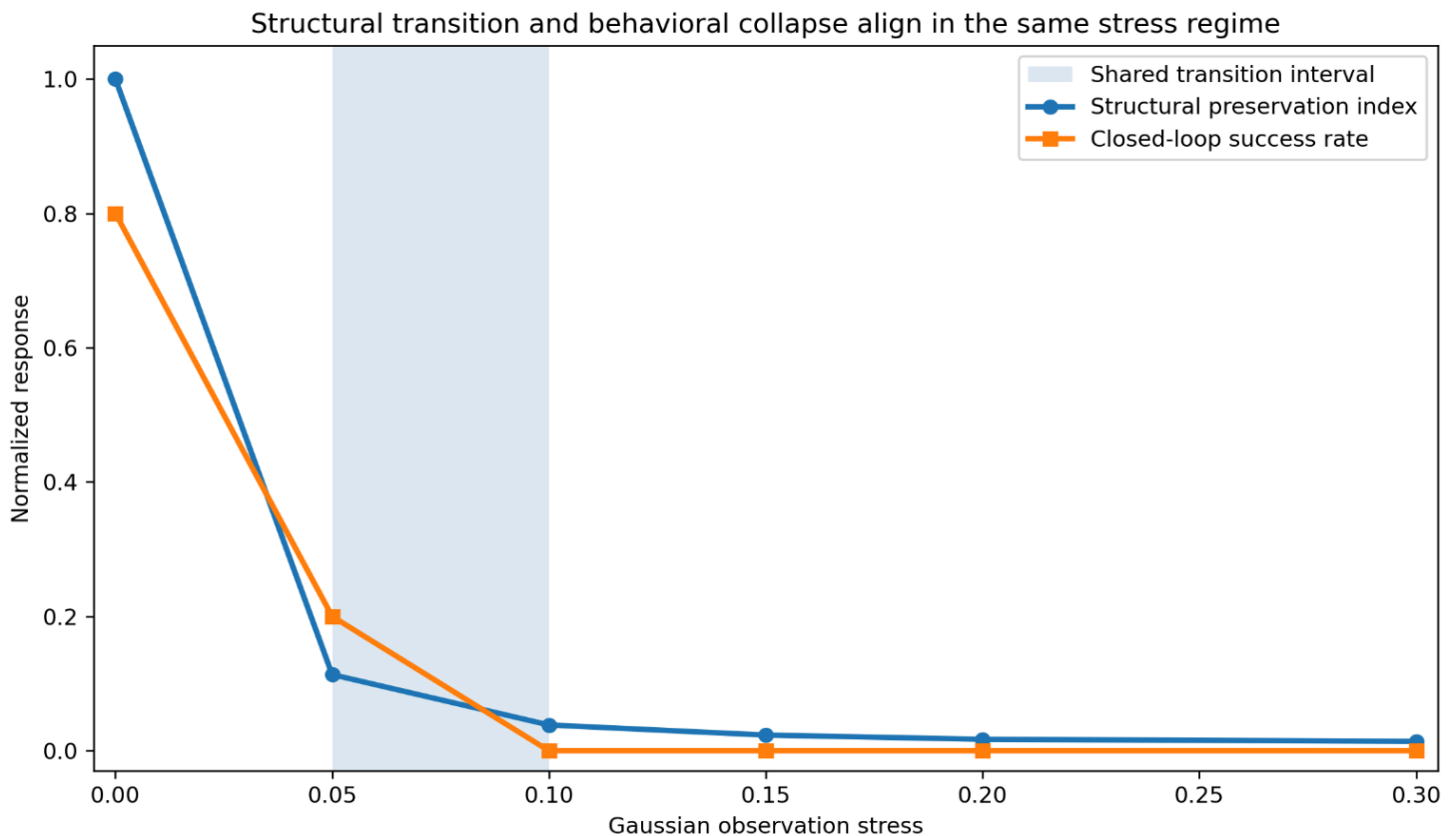


Figure 3

The central result: structural preservation and closed-loop success collapse across the same 0.05-0.10 stress interval. The structural preservation index is the mean of neighbor retention, matched basin retention, and clean-path admissibility.

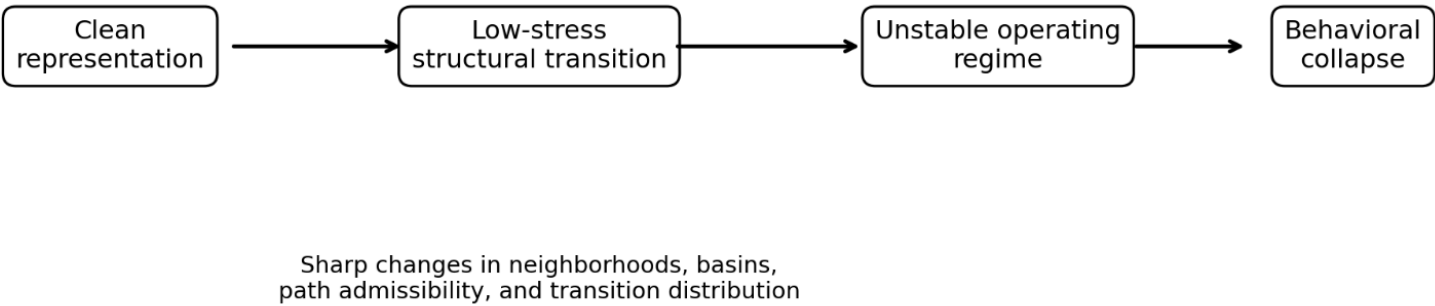


Figure 4

Conceptual interpretation of the observed sequence from clean representation to structural transition, unstable operating regime, and behavioral collapse.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)