Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Structural and Spectral Criticality Observables for Topology-Free Power System Stability Assessment

Chi Hsing Wu

jcwu@taiscience.org

TaiScience https://orcid.org/0009-0001-3396-6835

Steven Anderson

TaiScience

Kai-Siang Chen

Chia Nan University of Pharmacy and Science https://orcid.org/0009-0004-3496-3371

Method Article

Keywords: critical slowing down, voltage stability, delayed voltage collapse, transient stability, GridSFM, power grid, PMU, ANDES

Posted Date: July 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10529105/v1

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

Additional Declarations: The authors declare no competing interests.

Structural and Spectral Criticality Observables for Topology-Free Power System Stability Assessment*

JC Wu[1], Steven Anderson[1], and Kai-Siang Chen[2]

[1]TaiScience Research Group, New Taipei City, Taiwan
[2]Chia Nan University of Pharmacy and Science, Tainan City, Taiwan
{jcwu, steve.anderson}@taiscience.org · cks888@mail.cnu.edu.tw

July 29, 2026

Abstract

Modern grid-AI systems can identify feasible operating points or classify disturbance trajectories without directly verifying whether the predicted state remains dynamically separated from an instability boundary. As foundation models for optimal power flow are deployed across heterogeneous, renewable-rich grids, the absence of an independent physical check on their outputs becomes a first-order reliability risk. We treat physical stability as a property to be verified rather than assumed, introducing a multi-observable Critical Slowing Down (CSD) framework that is topology-free, post-hoc, and model-agnostic. It combines a structural observable (a time-domain Jacobian proxy) with spectral observables (a low-frequency participation ratio and a spectral entropy drift). On IEEE 39-bus scenarios the structural indicator achieves strong boundary discrimination near the critical clearing-time limit (Area Under the Curve, AUC up to 0.88), and tracks inter-area stress monotonically on a Kundur sweep. On large-scale GridSFM scenarios spanning distinct renewable mixes, the observables remain directionally consistent across a different network scale and simulation platform. Under noisy, short-horizon delayed-collapse conditions they further retain significant associations with collapse risk where voltage-based indicators do not. These results position topology-free criticality observables as an independent, non-invasive layer that makes the physical stability of grid-AI outputs checkable rather than assumed.

Keywords— critical slowing down, voltage stability, delayed voltage collapse, transient stability, GridSFM, power grid, PMU, ANDES

1 Introduction

Voltage instability remains a major concern in modern power systems because a disturbed network may appear to recover immediately after fault clearing while still progressing toward delayed collapse. Conventional stability assessment therefore requires indicators that can distinguish true recovery from trajectories whose restorative dynamics are already weakening [1,2].

Critical slowing down provides a general early-warning principle for systems approaching a critical transition. As the dominant restoring rate decreases, state trajectories may exhibit slower recovery, stronger temporal dependence, and increased low-frequency energy, although the measurable form of these effects depends on the available data and the physical setting [3–5].

*This work has been submitted to ICT-PEP 2026 (IEEE) for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

In our previous cross-domain study [6], which transfers the IBM TokaMind plasma foundation model [7] to power-grid Phasor Measurement Unit (PMU) data, CSD indicators were used as a confidence gate for severe-event classification, improving selective prediction; this motivates their extension here as a post-hoc physical validation layer for voltage-stability assessment.

Machine-learning and LLM-driven approaches for transient stability assessment have demonstrated strong predictive performance [8–10]. This motivates the development of complementary physics-based validation layers that provide evidence under out-of-distribution operating conditions, measurement noise, and other data-quality uncertainties [11–13].

Foundation models have begun to address large-scale power grid modeling, including the IBM Foundation Model for the Electric Power Grid [14], which leverages graph-based architectures trained on diverse grid topologies. Microsoft GridSFM, trained across more than 150 grid topologies and approximately half a million scenarios, may generate a physically feasible operating point that is nevertheless close to a dynamic instability boundary. Recent work has leveraged Physics-Informed Neural Networks (PINNs) to accelerate Electromagnetic Transient (EMT) simulations [15] and Alternating Current Optimal Power Flow (AC-OPF) surrogate modeling [16], but these approaches do not address online instability monitoring under Out-of-Distribution (OOD) operating conditions—a gap that CSD addresses.

This study evaluates the proposed framework through IEEE 39-bus transient-stability simulations, large-scale GridSFM regional-grid scenarios [17], and noisy ANDES delayed-collapse cases [18]. The main contributions are threefold. First, the paper tests whether topology-free rolling indicators can distinguish stable and unstable trajectories near transient-stability boundaries using the IEEE 39-bus and Kundur benchmark systems. Second, it examines whether low-frequency critical-slowing-down observables remain directionally consistent when the validation environment is extended from benchmark systems to large-scale regional-grid simulations. Third, it tests whether these observables remain informative under noisy, short-horizon conditions where voltage-based indicators lose resolution.

2 Methodology

2.1 Multi-Observable CSD Framework

Physics-informed approaches have demonstrated utility in microgrid voltage stability analysis [19]. The present framework extends this direction through two observable families whose common basis is the weakening of restorative dynamics near a critical transition, although no equivalence among the variants is assumed [3–5]. Engine A is a *structural* observable derived from the leading eigenvalue of a locally estimated transition operator; Engine B is a *spectral* observable derived from post-fault low-frequency reorganization. The framework provides a standardized cross-validation interface enabling an AI system's predictions to be benchmarked against physical stability evidence [20, 21].

2.1.1 Engine A: Jacobian Observable

Engine A estimates local recovery dynamics from multichannel PMU-equivalent trajectories. The measurements are standardized and projected into a low-dimensional Principal Component Analysis (PCA) space, and within each sliding window a local transition operator $\widehat{\mathbf{J}}_t$ is estimated by ordinary least squares. The window-level score is

$$s_J(t) = \log\left(\max_i \left|\lambda_i\left(\widehat{\mathbf{J}}_t\right)\right| + \varepsilon\right) \tag{1}$$

where $\lambda_i(\widehat{\mathbf{J}}_t)$ are the eigenvalues of $\widehat{\mathbf{J}}_t$; larger values indicate the dominant recovery mode lies closer to marginal stability. Two aggregations of (1) are used: a *rolling* variant over all windows

along the trajectory, used for transient-boundary verification and stress-sweep monotonicity, and a *fault-locked* variant retaining only windows fully contained within the fault interval, aggregated by their 90th percentile, used for delayed-collapse ranking under noisy short-horizon observation.

2.1.2 Engine B: Spectral Observable

Engine B characterizes post-fault spectral reorganization from a sliding Fast Fourier transform (FFT) window. Let $P(f_i)$ denote the power spectrum and $p_i = P(f_i)/\sum_j P(f_j)$ the normalized spectral energy of the i-th component. Two variants are used. The low-frequency participation ratio (LFPR), the principal score for monotonic stress sweeps, is

$$\mathrm{LFPR} = \frac{\sum_{f_i \le f_{\mathrm{low}}} P(f_i)}{\sum_{f_i} P(f_i)}, \qquad f_{\mathrm{low}} = 2.0 \text{ Hz} \tag{2}$$

where a larger LFPR indicates spectral energy concentrating in the low-frequency band as restorative dynamics slow. The spectral entropy drift (ΔH), the principal score for near-boundary ranking under noise, uses the normalized spectral entropy

$$H = -\sum_{i=1}^{N_f} p_i \log p_i \tag{3}$$

with the Engine B score defined as

$$\Delta H = H_{\mathrm{post}} - H_{\mathrm{fault}} \tag{4}$$

where H_{fault} is computed during the fault interval and H_{post} from the causal post-clearing window; a larger ΔH indicates stronger post-fault spectral reorganization.

2.2 Experimental Settings

The IEEE 39-bus and delayed-collapse experiments use trajectories generated in ANDES, an open-source environment for dynamic power-system simulation. The simulator represents the network through coupled differential-algebraic equations, allowing controlled fault application, fault clearing, and post-fault trajectory extraction under reproducible operating conditions [1,2].

2.2.1 Experiment 1: IEEE 39-Bus Boundary and Kundur Stress Sweep

Experiment 1 comprises two complementary tests. First, 116 IEEE 39-bus scenarios (77 stable, 39 unstable) are sampled at a PMU-equivalent rate; a Bus-16 clearing-time sweep locates the stability boundary between 1.3 s (stable) and 1.4 s (unstable) [1,2]. Rolling CSD indicators are computed directly from the PMU-like trajectories without topology information or a trained classifier. Second, a Kundur K_A sweep evaluates rolling Engine A, Engine B LFPR, spectral entropy, and the combined A+B score against the inter-area oscillation-amplitude measure `da_std`, using $f_{\mathrm{low}} = 2.0$ Hz and a 120-sample FFT window at 60 Hz. Spectral entropy is retained only as a comparison indicator, with LFPR as the principal Engine B score.

2.2.2 Experiment 2: Large-Scale Spectral Sensitivity

Experiment 2 uses GridSFM regional-grid scenarios for Texas and California to test cross-platform stress sensitivity at a larger network scale. Wind- and solar-connected buses are selected and their Optimal Power Flow (OPF) angle changes are weighted by generator capacity and capacity factor, and rolling Engine A, Engine B LFPR, and A+B are evaluated against

a renewable-weighted perturbation scale α swept from 0 to 3.0. Because these PMU-like trajectories derive from static OPF operating points rather than dynamic fault simulations, this experiment is interpreted as a monotonicity test, not as evidence of dynamic voltage-collapse prediction or field-measured PMU validation. It complements the ANDES experiments by changing both the simulation platform and network scale.

2.2.3 Experiment 3: Spectral Criticality Ranking

Experiment 3 evaluates whether post-fault spectral reorganization provides information beyond conventional voltage-based observables in delayed-collapse scenarios. A near-boundary subset (S1) of 75 trajectories from the ANDES delayed-voltage-collapse benchmark is adopted. For each causal observation horizon, Engine B computes ΔH from a sliding FFT window; the rolling LFPR metric is retained only as a reference frequency-domain observable and is not used as the primary Engine B score here. Performance is evaluated by Spearman rank correlation between the observable score and collapse risk, represented by the negative survival time $-T_{\text{surv}}$, with voltage deviation (V_{dev}) as the baseline.

2.3 Baseline Measure and Evaluation Protocol

Voltage deviation is adopted as the conventional baseline observable:

$$V_{\text{dev}} = \frac{1}{N_t} \sum_{t=1}^{N_t} \left[\max_i |V_i(t) - V_i(0)| \right] \tag{5}$$

where $V_i(t)$ is the voltage magnitude of bus i and N_t the number of evaluated samples. Ranking performance uses the Spearman rank correlation between an observable score R and collapse risk,

$$\rho = \text{Spearman}(R, -T_{\text{surv}}) \tag{6}$$

where shorter survival times T_{surv} correspond to higher collapse risk. The primary evaluation horizon is fixed *a priori* at $k = 2$ (33 ms); additional horizons are reported only to illustrate temporal evolution and are not used for model selection.

3 Results

3.1 IEEE 39-Bus Boundary and Kundur Stress-Sweep Results

The three rolling time-domain indicators were evaluated near the Bus-16 clearing-time boundary, with their Area Under the Curve-Receiver Operating Characteristic (AUC-ROC) results summarized in Table 1.

Table 1: IEEE 39-Bus Boundary Verification

Indicator	Summary	AUC-ROC
Jacobian-proxy mean	Last-20 mean	0.8795
Rolling trend statistic	Local slope	0.8025
Weighted composite score	Mean + trend	0.8768

Figure 1 compares the CSD indicators against the inter-area oscillation-amplitude measure `da_std`. The upper panel shows the responses of rolling Engine A and Engine B LFPR across the tested stress levels, while the lower panel summarizes the Spearman correlations of all

indicators. Engine A, LFPR, and A+B achieve $\rho = 1.00$, whereas spectral entropy and V_{dev} yield approximately $\rho = -0.80$.

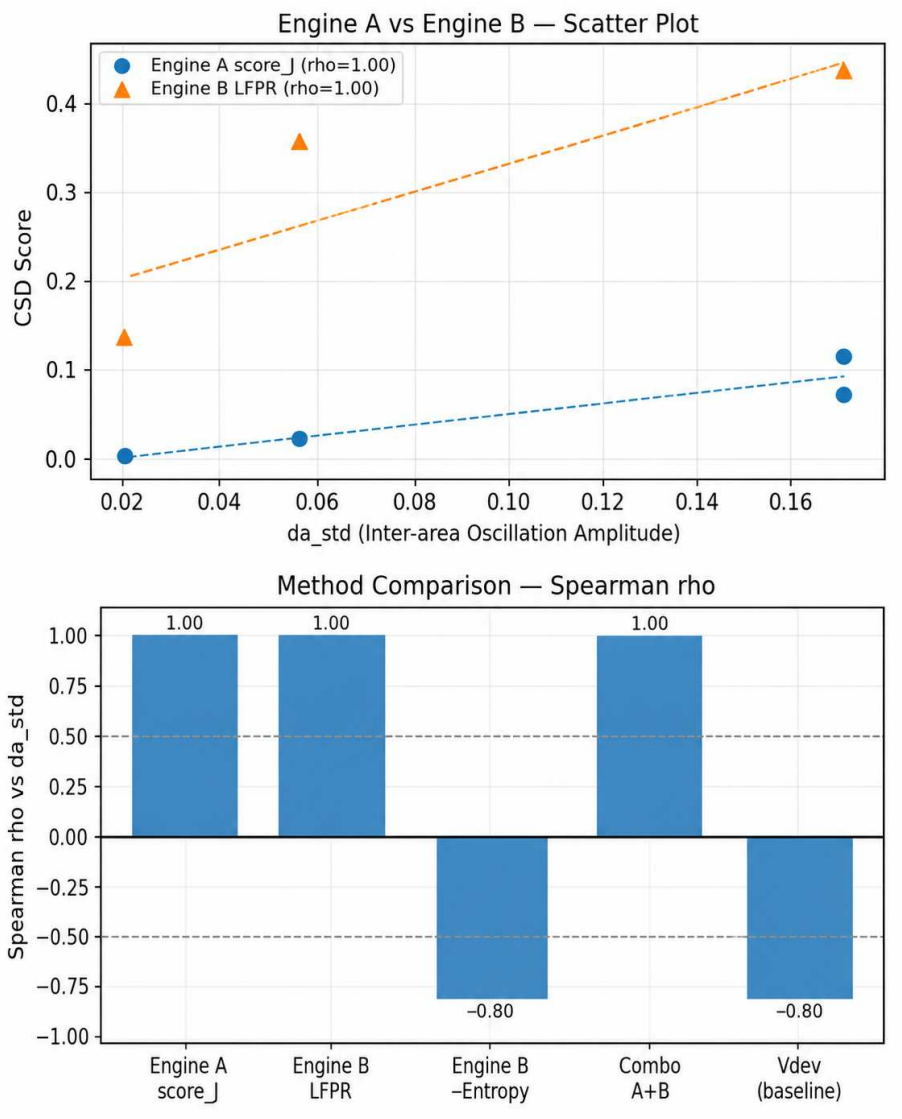


Figure 1: Comparison of CSD-based methods against the inter-area oscillation-amplitude indicator, denoted by `da_std`.

These results indicate monotonic sensitivity within the tested sweep, not field-level predictive accuracy. Figure 2 compares the two spectral indicators: the LFPR trajectories retain a clearer ordering across the K_A stress levels, particularly after the main disturbance, whereas the spectral-entropy responses overlap substantially. This supports LFPR as the principal Engine B indicator, with spectral entropy retained only as a comparison measure.

Table 2: Kundur K_A Stress-Sweep: Spearman Correlation of Each Observable Against the Inter-Area Oscillation Amplitude (`da_std`).

Observable	Spearman ρ
Rolling Engine A (s_J)	+1.00
Engine B LFPR	+1.00
Combined A+B	+1.00
Engine B spectral entropy	−0.80
V_{dev} (baseline)	−0.80

Table 2 reports the Kundur K_A stress-sweep results. Rolling Engine A, Engine B LFPR, and the combined A+B score each attain $\rho = +1.00$ against `da_std`, confirming monotonic sensitivity to the imposed inter-area stress. Spectral entropy and V_{dev} instead yield $\rho \approx -0.80$, i.e. a direction-inconsistent response, which is why spectral entropy is retained only as a comparison indicator and not as the principal Engine B score.

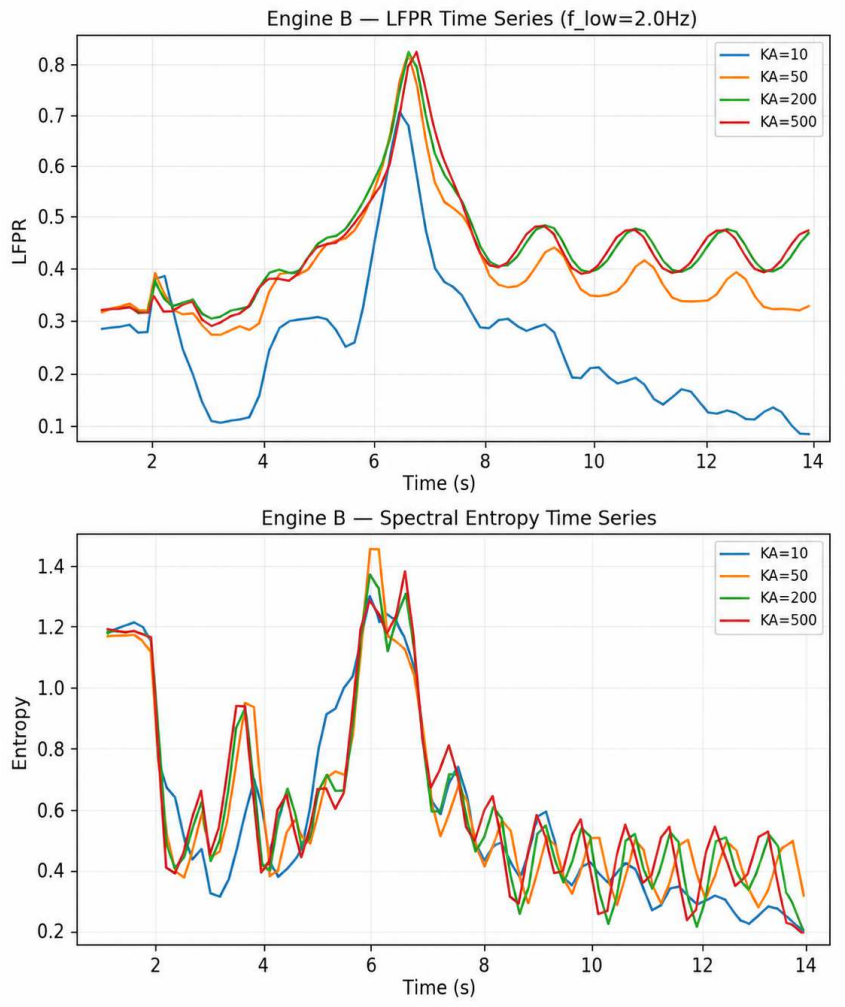


Figure 2: Engine B LFPR and spectral-entropy responses during the Kundur K_A sweep.

3.2 GridSFM Spectral Results

The GridSFM experiment evaluates whether the CSD observables remain monotonic under renewable-weighted perturbations in large regional-grid scenarios. Table 3 summarizes the

Spearman correlations with the renewable-weighted stress axis. In Texas, rolling Engine A shows a strong positive response ($\rho = +0.96$), while Engine B LFPR and A+B are perfectly monotonic ($\rho = +1.00$); in California, all three indicators achieve $\rho = +1.00$.

Table 3: GridSFM Renewable-Weighted Stress Validation

Region	VRE Share	Engine A	Engine B	A+B
Texas	29.3%	+0.96	+1.00	+1.00
California	12.4%	+1.00	+1.00	+1.00

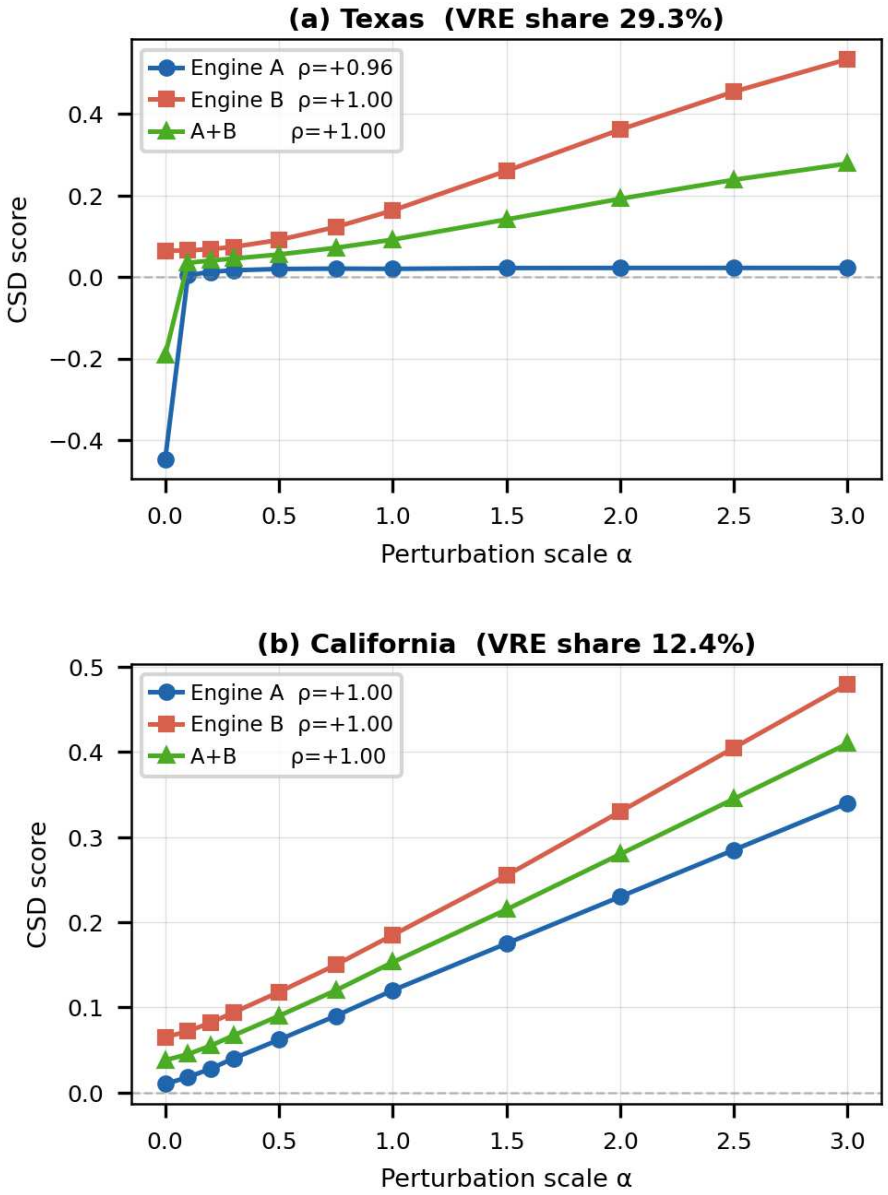


Figure 3: GridSFM renewable-weighted stress responses for Texas and California. Rolling Engine A (Jacobian proxy), Engine B (LFPR), and the combined A+B score are plotted against the perturbation scaling parameter α. Engine B and A+B preserve perfect monotonicity in both regions; Engine A shows a moderate positive response in Texas ($\rho = +0.96$) and perfect monotonicity in California ($\rho = +1.00$).

These correlations should be interpreted as monotonicity within the designed perturbation sweep rather than as independent evidence of dynamic voltage-collapse prediction. The GridSFM trajectories derive from static OPF conditions and synthetic PMU-like signals, whereas the ANDES experiments provide the dynamic fault-based evidence. To ensure CSD applicability across renewable scenarios, topology and control parameters were fixed and a progressive load perturbation was applied rather than inertia sweeps, thereby maintaining bifurcation-induced collapse dynamics and avoiding mechanism switching that would resemble rate-induced tipping.

3.3 Criticality Ranking Under Noisy Measurements

To evaluate robustness under measurement uncertainty, additive Gaussian noise (noise level $\sigma = 0.18$) was injected into the post-clearing voltage trajectories. Following the strict causal protocol adopted in this work, only the first three post-clearing PMU samples ($\approx 50\,\text{ms}$) were used for observable extraction. Under this short-horizon regime, binary stable–unstable discrimination remains relatively easy, indicating that short-horizon voltage trajectories already contain sufficient information for coarse stability assessment. Such coarse classification, however, provides limited resolution for distinguishing trajectories that are all near the stability boundary. The analysis therefore adopts criticality *ranking*—the monotonic association between an observable and collapse risk—rather than binary classification.

Table 4: Spearman correlation between the fault-locked Jacobian observable and voltage deviation (V_{dev}) on the near-boundary subset (S1, band = 0.03).

Horizon (ms)	ρ_J	p_J	ρ_{Vdev}	p_{Vdev}	Gain ($\Delta\rho$)
0–17	0.353	0.002	0.125	0.285	0.228
33	0.369	0.001	0.103	0.379	0.266
50	0.353	0.002	0.091	0.438	0.262
83	0.389	0.001	0.062	0.597	0.327
167	0.206	0.087	0.067	0.582	0.139

Table 4 summarizes the ranking performance of Engine A.

The evaluation was intentionally restricted to a challenging near-boundary subset ($n = 75$), where trajectories remain highly similar during the first 33 ms after fault clearing. Under this strictly causal protocol, only early post-clearing information is available for criticality assessment. At the primary horizon ($k = 2$, 33 ms), Engine A achieved $\rho = 0.369$ ($p = 0.001$), compared with $\rho = 0.103$ ($p = 0.380$) for the voltage-deviation baseline. The voltage baseline is statistically indistinguishable from chance ($p > 0.28$ at every horizon), whereas Engine A remains significant ($p < 0.01$) through 83 ms—precisely where an independent check is most needed.

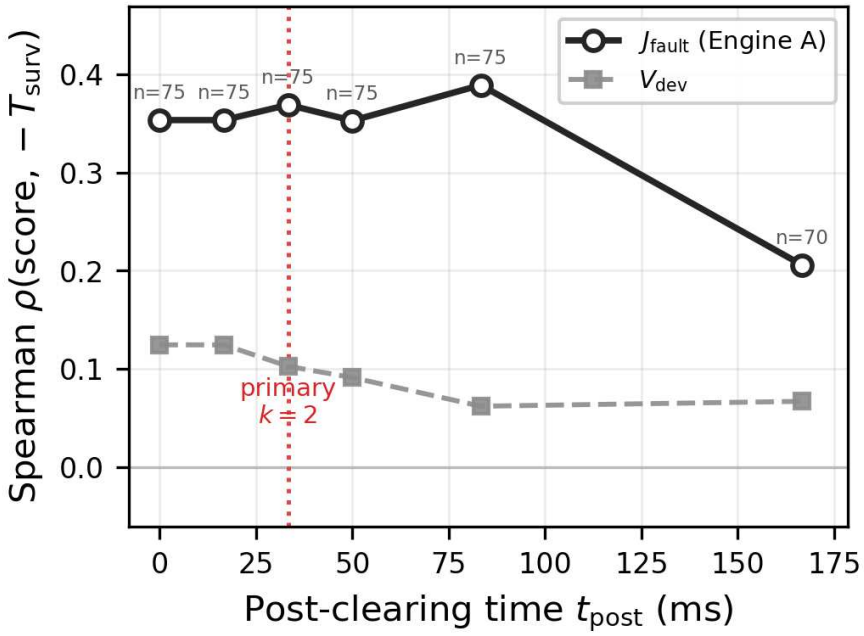


Figure 4: Engine A performance on the S1 delayed-collapse boundary subset. The fault-locked Jacobian observable (J_{fault}) consistently achieved higher rank correlation with collapse risk than the voltage-deviation baseline. The pre-specified horizon was fixed at $k = 2$ (33 ms).

Table 5 reports the corresponding results for Engine B. At the pre-specified primary horizon ($k = 2$, 33 ms), entropy drift achieved $\rho = 0.440$ ($p < 0.001$), exceeding both the voltage baseline and the Engine A ranking score, with positive and statistically significant correlations across all evaluated horizons. Although the strongest correlation occurred at $k = 20$ (333 ms, $\rho = 0.649$), this horizon was not selected as the primary evaluation point and is reported only as a secondary observation; the primary comparison remains the fixed 33 ms causal horizon adopted throughout.

Table 5: Spectral entropy drift (ΔH) as a spectral criticality observable on the S1 delayed-collapse boundary subset. Risk is encoded as $-T_{\text{surv}}$. The pre-specified primary horizon is fixed at $k = 2$ (33 ms) to match the causal protocol used throughout this study.

k	Horizon (ms)	$\rho(\Delta H, -T_{\text{surv}})$	p
2	33	+0.440	< 0.001
5	83	+0.484	< 0.001
10	167	+0.434	< 0.001
15	250	+0.431	< 0.001
20	333	+0.649	< 0.001

Overall, voltage trajectories alone are sufficient for binary stability classification but provide limited resolution for distinguishing trajectories near the stability boundary, whereas both Engine A and Engine B retain statistically significant ranking capability under noisy measurements, with Engine B providing the strongest association with collapse risk.

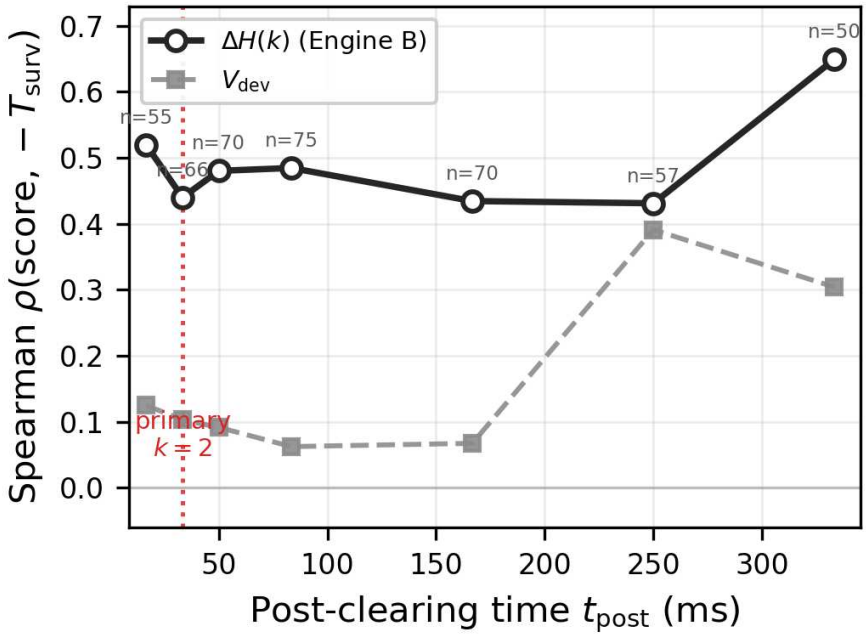


Figure 5: Engine B performance on the S1 delayed-collapse boundary subset. Spectral entropy drift (ΔH) achieved consistently higher rank correlation with collapse risk than the voltage-deviation baseline. The red dashed line marks the pre-specified primary evaluation horizon ($k = 2$, 33 ms).

4 Discussion

The proposed framework assumes bifurcation-driven critical transitions and does not address rate-induced tipping [3, 4, 22]. The GridSFM evaluation, moreover, rests on static OPF-derived operating conditions rather than field-measured PMU data, and the delayed-collapse results are obtained at a controlled delayed-collapse benchmark; both point to larger and field-validated event collections as the natural next step. What the three settings share is more telling than any single score: across a transient benchmark, a foundation-model-scale grid, and a noisy delayed-collapse regime, the same topology-free observables remain physically consistent without retraining or topology information—a property conventional voltage-based indicators alone do not consistently exhibit where it matters most, near the stability boundary.

5 Conclusion

This paper evaluates topology-free critical-slowing-down observables through transient-stability, regional-grid, and delayed-collapse scenarios. The results show that Jacobian-based and spectral observables can reveal stability-related information beyond conventional voltage measurements while remaining applicable across different simulation environments, supporting their use as complementary, non-invasive validation signals for grid-AI outputs. A key property of this post-hoc design is that it decouples a model's predictive speed from the physical assessment of its dynamic stability. Future work will focus on larger event collections, additional instability mechanisms, and validation using field-measured PMU data.

References

[1] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. V. Cutsem, and V. Vittal, "Definition and classification of power system stability," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[2] T. V. Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Boston, MA, USA: Springer, 2007.

[3] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, "Early-warning signals for critical transitions," *Nature*, vol. 461, no. 7260, pp. 53–59, 2009.

[4] V. Dakos, S. R. Carpenter, E. H. van Nes, and M. Scheffer, "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data," *PLoS ONE*, vol. 7, no. 7, p. e41010, 2012.

[5] C. Kuehn, "A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics," *Physica D: Nonlinear Phenomena*, vol. 240, no. 12, pp. 1020–1035, 2011.

[6] J. C. Wu, N. Lee, and K. Chen, "TokaMind for power grid: Cross-domain transfer from fusion plasma," 2026. [Online]. Available: https://arxiv.org/abs/2605.11033

[7] T. Boschi *et al.*, "TokaMind: A multimodal transformer foundation model for tokamak plasma dynamics," 2026. [Online]. Available: https://arxiv.org/abs/2602.15084

[8] J. Zhao, X. Han, C. Wang, J. Yang, and G. Zhang, "Contrastive-active transfer learning-based real-time adaptive assessment method for power system transient stability," *Sensors*, vol. 24, no. 15, p. 5052, Aug. 2024.

[9] M. Senyuk, M. Safaraliev, F. Kamalov, and H. Sulieman, "Power system transient stability assessment based on machine learning algorithms and grid topology," *Mathematics*, vol. 11, no. 3, p. 525, 2023.

[10] L. Hu, Y. Wang, and B. Pal, "LLM-driven transient stability assessment: From automated simulation to neural architecture design," Feb. 2026. [Online]. Available: https://arxiv.org/abs/2511.20276

[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[12] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, "Characterizing and quantifying noise in pmu data," in *2016 IEEE Power and Energy Society General Meeting*, 2016, pp. 1–5.

[13] Z. Li, H. Liu, J. Zhao, T. Bi, and Q. Yang, "A power system disturbance classification method robust to pmu data quality issues," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 130–142, 2021.

[14] H. F. Hamann *et al.*, "A perspective on foundation models for the electric power grid," Jul. 2024. [Online]. Available: https://arxiv.org/abs/2407.09434

[15] I. V. Nadal, M. K. Bakhshizadeh, P. Aristidou, N. Darii, R. Nellikkath, and S. Chatzivasileiadis, "Scalable physics-informed neural networks for accelerating electromagnetic transient stability assessment," Feb. 2026. [Online]. Available: https://arxiv.org/abs/2511.23046

[16] R. Nellikkath and S. Chatzivasileiadis, "Physics-informed neural networks for AC optimal power flow," *Electric Power Systems Research*, vol. 212, p. 108412, 2022.

[17] W. Yang, A. Britto, T. Spina, S. Fowers, B. Zhang, and C. M. White, "GridSFM: A foundation model for AC optimal power flow," Microsoft Research, Technical White Paper, May 2026. [Online]. Available: https://www.microsoft.com/en-us/research/wp-content/uploads/2026/05/GridFM_white_paper.pdf

[18] H. Cui, F. Li, and K. Tomsovic, "Hybrid symbolic–numeric framework for power system modeling and analysis," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1373–1384, Mar. 2021.

[19] R. Feng, F. E. Subhan, W. Khan, M. Faheem, J. Wang, and M. S. Bhutta, "Uniform physics informed neural network framework for microgrid and its application in voltage stability analysis," *IEEE Access*, vol. 13, pp. 4576–4589, 2025.

[20] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, "Slowing down as an early warning signal for abrupt climate change," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14 308–14 312, 2008.

[21] S. V. George, S. Kachhara, and G. Ambika, "Early warning signals for critical transitions in complex systems," *Physica Scripta*, vol. 98, p. 072002, 2023.

[22] Y. Huang, S. Bathiany, P. Ashwin, and N. Boers, "Deep learning for predicting rate-induced tipping," *Nature Machine Intelligence*, vol. 6, pp. 1556–1565, 2024.