A domain-specific modal-growth detector clears a matched false-alarm bar on power-grid instability, and an eigenvalue regime map shows when its form transfers

Miroslav Šotek
Anulum Institute
ORCID 0009-0009-3560-0851 protoscience@anulum.li

August 2026

Abstract

Under a matched-false-alarm, permutation-tested evaluation protocol, generic early-warning detectors are statistically at chance on power-grid growing oscillations, as they are on brain, heart and climate data (companion manuscript). This paper shows the complementary positive result and bounds it. A domain-specific detector that reads the canonical wide-area-monitoring instability quantity — the exponential growth rate σ of the most unstable bus's cross-bus voltage-deviation envelope — leads 36 of 90 generator-trip transitions of the PSML 23-bus corpus at a matched 10 % false alarm (label-permutation $p = 0.0001$; held-out half 24 of 45, $p = 0.0002$), on a non-circular disturbance-type split where the best generic detector is at chance (13 of 90, $p = 0.19$). The operating point was selected on a development half through a disclosed variant search and validated on the held-out half. Stress tests then bound the result. Run as a causal stream, the certified per-window operating point is not deployable, and an honest recalibration with a persistence debounce and an exponential-fit-quality gate recovers 11 of 45 held-out transitions at a matched 10 % stream false alarm. On two cross-dataset corpora (three real ISO-NE PMU captures; the 13 synthetic WECC 240-bus cases of the 2021 IEEE-NASPI contest), the certified numeric threshold does not transfer in either direction, while the frozen shape with per-system matched-false-alarm calibration remains usable. The form does not transfer to scalp EEG or to few-timepoint molecular indices. Finally, against ANDES small-signal eigenvalues of the IEEE-39 and Kundur systems, the analytic Jacobians of the fold, pitchfork and Hopf normal forms, and the mean-field eigenvalues of unimodal and bimodal Kuramoto models, the estimated quantity is confirmed to be the eigenvalue's real part, yielding a regime map: the envelope-growth family sizes the eigenvalue on oscillatory modes, the autocorrelation family on non-oscillatory ones. All evidence records are content-addressed and byte-reproducible.

1 Background

When a disturbance leaves a power-system electromechanical mode under-damped, the amplitude of the cross-bus voltage deviation grows exponentially: the real part σ of the mode's eigenvalue is positive, the canonical small-signal-stability quantity of wide-area monitoring (Kundur 1994). A detector that reads this quantity directly is domain-specific in the strict sense: it presupposes the physics of a linear instability rather than a generic statistical signature of critical slowing down.

A companion manuscript establishes the baseline this paper builds on: under a matched-false-alarm protocol with label-permutation significance — thresholds calibrated to a 10 % false alarm on no-transition nulls, detection counted only when an alarm falls in the pre-onset segment, 10 000-draw permutation p-values, hash-sealed evidence — generic early-warning detectors (critical slowing down, rising synchronisation, ordinal-transition entropy, their fusion, and the Dakos AR(1)-Kendall-τ trend) are statistically at chance on grid data, as on brain, heart and palaeoclimate corpora. The questions here are: does a detector that reads the physical instability quantity clear that same bar; under what deployment conditions does the certified result survive; and is the quantity it estimates externally checkable?

2 Methods

2.1 The modal envelope-growth detector

The detector estimates σ directly as the slope of the log cross-bus voltage-deviation envelope over a scoring window. Two physically motivated refinements are selected on a development half only: a *focal* aggregation takes the growth rate of the most unstable bus (instability is localised to a mode/bus cluster, so the per-bus maximum, un-diluted, is the signal; the matched-false-alarm threshold is calibrated on that same per-bus maximum over the nulls, absorbing the multiplicity), and a *recency weighting* up-weights the later samples of the window (a real instability accelerates toward the disturbance). Every variant tried is disclosed in Section 3.2.

2.2 The evaluation protocol and the non-circular split

The protocol is the companion manuscript's, unchanged: identical two-second pre-onset segments, thresholds calibrated continuously to a 10 % false-alarm target on the same nulls (achieved 9.1 %), a transition counted as led only on a pre-onset alarm, label-permutation significance with 10 000 draws and a fixed seed. The PSML corpus (Zheng et al. 2021; 23-bus co-simulation at 238 Hz) is split by disturbance *type*: a transition is any generator-trip scenario (90 usable), a null any damped bus-fault or branch-trip scenario (88) — a physical annotation independent of the growth statistic the detector measures, so a growing-oscillation detector is never scored against a growth-defined label. Three scenarios with a non-monotonic time column are dropped and the count is recorded in the sealed payload. The operating point is chosen on the even-index development half and validated on the odd-index held-out half.

2.3 A counterpoint detector on a domain without the signature

To test whether domain specificity alone wins, the same protocol runs a domain-specific detector on a domain without a deterministic signature: a scalp-EEG spectral detector (beta 13–30 Hz / delta 0.5–4 Hz band-power ratio, a-priori bands from the seizure-prediction literature, Kendall-τ rising trend, whole-head and focal aggregations) on the six chb01 preictal segments.

3 Results on the certification corpus

3.1 Head-to-head against the generic suite

The modal detector leads 36 of 90 transitions (40 %) at $p = 0.0001$; every generic member is at chance, the best (critical slowing down) reaching 13 of 90 at $p = 0.19$. The held-out half confirms the operating point: 24 of 45 (53 %) at $p = 0.0002$, the unbiased estimate.

Detector	Transitions led / 90	Permutation p
generic critical slowing down	13 / 90	0.185
generic weighted ensemble	11 / 90	0.324
generic ordinal-transition entropy	8 / 90	0.629
generic synchronisation	3 / 90	0.976
modal envelope-growth (focal, recency)	**36 / 90**	**0.0001**

Table 1: Identical segments, nulls, budget and test; only the detector differs.

3.2 The disclosed variant search

Variant	Development led / 45, p	Held-out led / 45, p
whole-network log-slope σ	11, 0.051	14, 0.009
per-bus maximum σ (focal)	14, 0.008	22, 0.0001
matrix-pencil, PCA-dominant mode	10, 0.072	3, 0.80
matrix-pencil, per-bus maximum	8, 0.189	6, 0.39
Hilbert-envelope σ, per-bus	2, 0.90	3, 0.80
recency-weighted focal σ	**16, 0.002**	**24, 0.0002**

Table 2: Choice made on development, reported on held-out; no held-out data touched during selection.

Notably, the gold-standard wide-area-monitoring estimator — matrix-pencil modal damping — recovers a planted σ exactly on synthetic data, yet on the short two-second, 238 Hz windows it is at chance (held-out $p = 0.39$–0.80); the robust direct log-envelope slope wins on real noisy segments. A recency-ratio sensitivity sweep (2–8) beats the unweighted per-bus detector across the whole range, so the choice is robust rather than knife-edge.

3.3 The counterpoint

The scalp-EEG spectral detector is itself at chance (1 of 6 led, $p \approx 0.56$, in both aggregations): the preictal state is murky and the corpus small. Domain specificity does not win by itself; the protocol certifies the difference between a detector reading a genuine physical signature and a plausible but empty one.

3.4 The streaming operating point

The head-to-head is a per-window result. Run as a live causal stream (every sliding window scored as a PMU feed would deliver it), the certified per-window threshold leads 82 of 90 transitions but false-alarms on 73 % of damped scenarios — a damped bus-fault or branch-trip produces a short transient growth window the continuous monitor alarms on, so the per-window operating point is operationally unusable as-is. Recalibrating at a matched stream-level false alarm, with a persistence debounce and an exponential-fit-quality gate that rejects a fault's step-like transient (a fault fits an exponential poorly, a genuine instability well), the gated detector leads 11 of 45 held-out transitions (24 %) at a matched 10 % stream false alarm — well below the offline per-window 40 %. The gap is physical: separating a sustained instability from a damped fault online requires observing the damping sign over several cycles. The streaming operating point

— sealed in `examples/real_data/psml_modal_growth/grid_modal_stream_operating_point.json` — is the deployable figure, and the downstream advisory record is passive and review-only: it surfaces σ, the most-unstable bus and the certified operating point, carries the ∼24 % recall as a first-class field, never actuates, and states explicitly that the absence of an advisory is not evidence of stability.

4 Cross-dataset transfer

Two pre-registered cross-dataset evaluations — protocols fixed before any detector run, every amendment disclosed in the sealed records — test what survives a dataset change. Both are small, and the conclusions are scoped to these corpora.

The certified numeric threshold did not transfer on either corpus. On three real ISO-NE PMU captures of documented sustained oscillations (UTK test-cases library), the frozen PSML threshold (1.3203 σ/s, two-second windows) fires on ∼32 % of pre-onset ambient windows against the certified 9.09 %. On the 13 synthetic WECC 240-bus forced-oscillation cases of the 2021 IEEE-NASPI contest (bus-voltage observable, the certified observable family), the same threshold crosses nothing: 0 of 611 pooled ambient windows and 0 transition windows. On both corpora tested, the certified threshold behaved as a corpus-specific quantity; deployment therefore calibrates per system on the target system's own ambient data and re-seals.

The frozen shape with local calibration, stated conservatively. On ISO-NE (window scaled to five cycles of the documented mode, threshold at a matched 10 % false alarm on pooled pre-onset ambient windows), the detector catches 1 of 3 events, the 1.13 Hz regional mode, 57.1 s before the estimated onset ($p = 0.332$ at $n = 3$): a case study, no significance claim. On WECC the corpus structure dominates: 9 of 13 cases are already past three times the ambient in-band envelope at the exact forcing start ($t = 30$ s, known by simulation design) — an effectively instantaneous onset with no early-warning window for any detector at these 90 s records — and of the 4 resolvable cases the locally calibrated shape leads 2 (+5.4 s, +4.1 s; $p = 0.41$, not significant). A secondary branch reports post-onset detection descriptively (12 of 13, median latency ∼1 s) with an explicit per-case chance bound showing the count is chance-compatible at this budget on 60 s regions; no significance is claimed. Records: `examples/real_data/iso_ne_forced_oscillation/` and `examples/real_data/wecc_240_osl/`.

5 Where the form does not transfer

Scalp EEG: a model mismatch. Scoring the same a-priori beta/delta band-power trajectory by the exponential growth rate of its envelope — identical focal aggregation, recency weighting and R^2 gate — leads 0 of 6 preictal segments ($p = 1.00$) in both aggregations, below the rank trend's 1 of 6; the gate, built to reject non-exponential fault transients, rejects the preictal signal too. A robustness grid over window length, recency and gate stays at 0 of 6 at every setting: a model mismatch, not a parameter artefact — the statistic presupposes a linear instability's eigenvalue, and the preictal rise, if present, is a slow non-exponential drift. (A 600 s horizon surfaces a rank-trend lead of 3 of 6, $p = 0.04$; one subject, one forking path — disclosed as exploratory, not claimed.) Record: `examples/real_data/chb01_seizures/seizure_modal_transfer.json`.

Few-timepoint molecular indices: a resolution limit. On the dynamical-network-biomarker rising limbs (three points single-cell, four points bulk), every monotone rise fits an exponential well enough ($R^2 = 0.67$–0.86) to pass the gate, and the growth rate re-orders the lineages exactly as the existing linear slope does: the transfer collapses to the detector DNB already uses. Record: `examples/real_data/mojtahedi_fate/dnb_modal_transfer.json`.

Together these bound the transfer: the form needs both a genuine exponential instability and a trajectory long enough to fit and gate it.

6 External validation against eigenvalue ground truth

6.1 Real systems: ANDES small-signal analysis

On the IEEE 39-bus New England system and the Kundur two-area system (open-source ANDES simulator; Cui et al. 2021), we sweep the operating point, take the dominant electromechanical mode's growth rate σ from small-signal eigenvalue analysis — an independent method — and score a ringdown from a small disturbance at each point. The *coherent* aggregation (cross-bus mean or dominant spatial mode) recovers the true σ trend on both systems (Spearman $\rho = 0.87$ on IEEE-39, 0.60–0.66 on Kundur); the *focal* aggregation — the PSML winner — does not ($\rho = -0.59$, -0.27): on a slow coherent inter-area mode the per-bus maximum locks onto local excursions. The quantity generalises; the best aggregation is regime-dependent. Limits: simulated ground truth, stable operating points only (a damping-ranking test), fourteen points per system. Record: `examples/real_data/grid_external_validation/grid_eigenvalue_external_validation.json`.

6.2 Analytic normal forms: fold, pitchfork, Hopf

For non-oscillatory codimension-one bifurcations the ground truth is closed-form: sweeping the control parameter quasi-statically through the stochastic normal form, the autocorrelation channel recovers the Jacobian eigenvalue λ ($-2\sqrt{\mu}$ fold, μ pitchfork) in rank and magnitude (Spearman and Pearson $\rho = 0.98/0.96$), because the lag-one autocorrelation of a linear-response process is $\exp(\lambda\,\Delta t)$; the variance channel rises in step ($\rho = 0.98$). On the Hopf normal form (oscillatory, eigenvalue $\alpha \pm i\omega$) the roles reverse: the envelope-growth family recovers α in rank and magnitude ($\rho = 0.97$, mean $|\sigma - \alpha| = 0.04$, frequency-invariant across 0.2–1.2 Hz), while the autocorrelation family tracks α only in rank (its lag-one autocorrelation is pinned near $\cos(\omega\,\Delta t)$). The envelope recovery degrades with ringdown SNR (ρ 0.97 $\rightarrow$ 0.62), a physical floor. Records: `examples/real_data/csd_bifurcation/csd_bifurcation_external_validation.json`, `examples/real_data/hopf_bridge/hopf_bridge_external_validation.json`.

6.3 The collective coordinate: unimodal and bimodal Kuramoto

The same two regimes appear one level up, on the emergent collective coordinate of a high-dimensional system. For the noisy Kuramoto model ($N = 512$, Lorentzian frequencies, phase diffusion D), the incoherent state loses stability at $K_c = 2(\gamma + D)$ with a real mean-field eigenvalue $\lambda(K) = (K - K_c)/2$ (Sakaguchi 1988): the autocorrelation family recovers λ — but only through the signed mean-field component $\mathrm{Re}(Z)$ (magnitude slope $1.19 \approx 1$); the folded amplitude $|Z|$ ranks it ($\rho = 0.98$) but cannot size it (slope 2.66). For a symmetric bimodal frequency law (Martens et al. 2009) the sub-critical eigenvalue is complex and the roles reverse exactly as the Hopf bridge predicts: the envelope-growth family on the sub-population order parameter $|Z_+|$ recovers $\mathrm{Re}(\lambda)$ in magnitude ($\rho = 0.99$, slope 0.93; $N = 4000$ ringdown), the autocorrelation family is confounded by the oscillation Ω (slope 15.8), the global $|Z|$ is a standing wave that cannot be fit, and the measured Ω matches the analytic value to $\sim$7 %. Limits: finite-N ($O(1/\sqrt{N})$), quasi-static or noiseless ringdown sweeps, coupling below K_c. Records: `examples/real_data/kuramoto_` `synchronization/kuramoto_synchronization_external_validation.json`, `examples/real_data/kuramoto_bimodal/kuramoto_bimodal_hopf_external_validation.json`.

6.4 The regime map

Mode	Systems	Eigenvalue truth	Estimator	Recovery
Oscillatory, real (§6.1)	IEEE-39, Kundur	ANDES σ	envelope, coherent	rank (ρ 0.87 / 0.60+)
Non-oscillatory (§6.2)	fold, pitchfork	normal-form λ	AR1	magnitude (ρ 0.98 / 0.96)
Oscillatory (§6.2)	Hopf normal form	normal-form α	envelope σ	magnitude (ρ 0.97)
Collective, real λ (§6.3)	Kuramoto unimodal	mean-field λ	AR1 on $\mathrm{Re}(Z)$	magnitude (slope 1.19)
Collective, complex λ (§6.3)	Kuramoto bimodal	mean-field $\mathrm{Re}\lambda$	envelope on $\lvert Z_+ \rvert$	magnitude (slope 0.93)

Table 3: The eigenvalue's real part is the common quantity; the magnitude-correct estimator depends on whether the mode oscillates.

The split is mechanical, not empirical accident: on an oscillatory mode the lag-one autocorrelation is confounded by the oscillation while the envelope carries the decay cleanly; on a non-oscillatory mode there is no envelope to fit and the linear-response autocorrelation is exactly $\exp(\lambda\,\Delta t)$. Pick the estimator by whether the mode oscillates; the estimated quantity is the same eigenvalue throughout.

7 Discussion

The regime map of Table 3 is the paper's main theoretical output. Across five independent ground truths — a power system simulator's small-signal analysis, three analytic normal forms, and two mean-field Kuramoto reductions — the quantity the detector estimates is the eigenvalue's real part, and which estimator recovers its magnitude is set mechanically by whether the mode oscillates: the envelope-growth family on oscillatory modes, the autocorrelation family on non-oscillatory ones, and on collective coordinates the signed mean field rather than the folded order-parameter amplitude. This turns the empirical grid result into a claim about a named first-principles quantity with a stated domain of applicability.

The empirical results bound that claim in deployment-relevant ways. A detector reading the instability quantity directly clears an operational bar that generic detectors do not, while the counterpoint shows domain specificity alone is not sufficient — the same idea on a domain without the signature stays at chance. The streaming recalibration replaces the per-window figure with an honest $\sim$24 % held-out recall at a held false alarm; the two cross-dataset corpora show the certified numeric threshold behaving as corpus-specific in both directions, leaving the frozen shape plus per-system calibration as the transferable unit; and the transfer boundaries on EEG and molecular data delimit where the exponential form applies at all.

8 Limitations

- The certification corpus is one dataset at one sampling rate with two-second windows; the held-out count (24/45) is the unbiased estimate and the full-corpus count (36/90) is the fitted-model deployment figure; both are sealed.
- The cross-dataset corpora are small (three real events; 13 synthetic cases, mostly with effectively instantaneous onsets), so the transfer conclusions are scoped to these corpora; the

WECC early-warning null partly reflects corpus structure rather than detector skill.

- The eigenvalue validations use simulated or analytic ground truths (no field-PMU eigenvalue reference), stable operating points or sub-critical sweeps, and finite-N Kuramoto runs.
- The EEG counterpoint and transfer test are one subject, six seizures.

9 Reproduction and data availability

Every result regenerates deterministically (fixed permutation seed; no other randomness). The pipelines are Python (NumPy/SciPy; pinned lock files in the repository); the PSML head-to-head and the external validations each complete in minutes on a workstation, the ANDES sweeps in tens of minutes. Entry points:

```
python -m bench.grid_modal_head_to_head        PSML OUT
python -m bench.early_warning_leadtime_isone   --data-dir ISONE --output OUT
python -m bench.early_warning_leadtime_wecc    --data-dir WECC  --output OUT
python -m bench.csd_bifurcation_external_validation   OUT
python -m bench.dnb_modal_transfer                    OUT
```

Sealed records and integrity tests are committed under `examples/real_data/` in the source repository (https://github.com/anulum/scpn-phase-orchestrator); a fresh run reproduces every `content_hash` bit-for-bit. Raw corpora are public and cited, not redistributed: PSML (Zheng et al. 2021), the ISO-NE and WECC 240-bus cases of the UTK oscillation test-cases library (https://web.eecs.utk.edu/~kaisun/Oscillation/; Maslennikov et al. 2016; NREL reduced 240-bus model, credit DOE/NREL/Alliance), and CHB-MIT via PhysioNet.

References

[1] Kundur P. *Power System Stability and Control.* McGraw-Hill (1994).

[2] Zheng X, et al. A multi-scale time-series dataset with benchmark for machine learning in electricity (PSML) (2021).

[3] Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS* 105(38):14308 (2008).

[4] Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature* 461:53 (2009).

[5] Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: the long and winding road. *Brain* 130(2):314 (2007).

[6] Cui H, Li F, Tomsovic K. Hybrid symbolic-numeric framework for power system modeling and analysis. *IEEE Transactions on Power Systems* 36(2):1373 (2021); the ANDES simulator.

[7] Maslennikov S, Wang B, Zhang Q, Ma F, Luo X, Sun K, Litvinov E. A test cases library for methods locating the sources of sustained oscillations. IEEE PES General Meeting (2016).

[8] Yuan H, Biswas RS, Tan J, Zhang Y. Developing a reduced 240-bus WECC dynamic model for frequency response study of high renewable integration. 2020 IEEE/PES T&D (2020).

[9] Sakaguchi H. Cooperative phenomena in coupled oscillator systems under external fields. *Progress of Theoretical Physics* 79(1):39 (1988).

[10] Martens EA, Barreto E, Strogatz SH, Ott E, So P, Antonsen TM. Exact results for the Kuramoto model with a bimodal frequency distribution. *Physical Review E* 79:026204 (2009).

[11] Shoeb A. Application of machine learning to epileptic seizure onset detection and treatment. PhD thesis, MIT (2009); CHB-MIT via PhysioNet.