

Compositional certificates for lossy power-grid stability, cross-validated by sum-of-squares programming

[TODO: Author Name(s)]¹

¹[TODO: Affiliation, Department, Institution, City, Country]

[TODO: corresponding author email]

Abstract

Certifying how much of a power grid’s true stable operating region a Lyapunov function can prove is exponentially costly for monolithic verification once transmission losses are present. On the IEEE New England 39-bus test system, we build a certificate that eliminates the lossy ten-generator Kron reduction’s speed variables in closed form, cutting an 18-D verification problem to 9-D, and show mechanistically why a scalar stability correction cannot compensate for lossy coupling – the required correction is $\approx 98\%$ rotational – fixing it instead with a learned vector correction that certifies 24.8% of the true region of attraction at 100% soundness, within half a point of that certificate’s own verification ceiling. An independently built, exact sum-of-squares certificate reproduces this result on every evaluation slice, and a convex program separates two previously entangled questions – how far stability can be proved, and how well the proved region’s shape matches reality – recovering an 86% relative coverage gain on the shape question alone once decoupled from the harder radius question, a gain six prior gradient-based searches missed entirely. Both the physics-only and learned-correction mechanisms transfer to a second, independently constructed benchmark topology, and the trained certificate is robust to a $\pm 15\%$ loading perturbation on two of three evaluation slices, collapsing on the third – reported exactly rather than smoothed into a single robustness claim.

Introduction

A power grid that suffers a fault either returns to synchronous operation or it does not, and operators need to know, in advance, how large a set of post-fault states this holds for – not just whether one simulated trajectory happens to recover. The classical answer, the transient energy function [1], gives an exact stability certificate but only for lossless networks, where the inter-generator coupling has a scalar potential. Every real transmission line has resistance, and once it does, no exact energy function exists; recovering a rigorous certificate for the lossy case is the problem this paper solves for one representative benchmark, together with an unusually thorough accounting of exactly where the resulting certificate’s limits come from.

Two modern toolkits attack pieces of this problem without solving it whole. Neural Lyapunov methods [2, 3] learn a certificate and verify it by sound bound propagation [4, 5, 6], but verification cost scales exponentially with state dimension, limiting practical use to a handful of generators. Sum-of-squares and occupation-measure methods [7, 8] compute provably convergent outer bounds on the stable region via semidefinite programming with no interval relaxation at all, but their cost grows combinatorially with dimension and polynomial degree. Neither has, to our knowledge, been applied to the IEEE New England 39-bus system [9] – a standard ten-generator lossy transient-stability benchmark – with a sound, quantified account of how much of the true stable region the result actually covers.

We report four findings. First, a compositional reduction that makes sound branch-and-bound verification tractable at this scale at all. Second, a mechanistic account of why lossy coupling defeats scalar stability corrections, and a vector-valued fix that reaches 24.8% certified coverage at full soundness – within half a

Table 1: Monolithic branch-and-bound cost vs. compositional cost, measured on this system.

Dim.	Active gens.	Monolithic leaves	Speedup
2	1	64	$1\times$
10	5	1.1×10^9	$2.4 \times 10^6\times$
20	10	1.15×10^{18}	$7.0 \times 10^{14}\times$

point of the best any verifier could ever certify for that same object. Third, an independently implemented, exact sum-of-squares certificate that reproduces the branch-and-bound result on every evaluation slice, the first exact cross-check in this line of work. Fourth, a structural decomposition showing that certified coverage depends on two logically separate questions – how far stability can be proved, and how well the proved region’s shape matches ground truth – and that a convex program solves the second question far better than six independent neural-network searches, once it is no longer entangled with the first.

Results

A compositional reduction makes ten-generator verification tractable

The plant is the swing-equation model of the IEEE New England 39-bus test system [9], Kron-reduced to its ten generator buses [10], with a closed-loop droop equilibrium and controller gains reused verbatim from Cui’s Lyapunov-regularized reinforcement learning benchmark (Methods, Eq. 1–2) [TODO: verify exact citation]. One generator is grounded as reference, leaving a 9-D relative angle vector ϕ and 9-D relative speed vector ω . For any Lyapunov candidate quadratic in speed – $V = \frac{1}{2}\omega^\top M\omega + U(\phi) + \Psi(\phi)^\top M\omega$, with Ψ a correction (“cross”) term – the speed variables eliminate in closed form (Methods, Eq. 3–6), collapsing the verification claim from 18-D to a 9-D condition $h(\phi) < 0$. This single structural choice is what makes sound interval branch-and-bound tractable on ten generators; verifying the same claim directly over the full 18-D state (see the negative-result row ‘c17’ below) is the exponential regime every prior neural-Lyapunov result on this system size runs into.

This is not asserted asymptotically – we measured it. A naive worst-case composition (bounding each coupling term by magnitude alone) reproduces the same verification claim at increasing generator count on this exact system: at 20-D, monolithic branch-and-bound needs on the order of 10^{18} leaf boxes to resolve what the compositional reduction resolves in $O(10^3)$ operations, a measured $7 \times 10^{14}\times$ gap (Table 1). This scaling alone is not sufficient for a good certificate – the same measurement shows worst-case composition itself degrades quickly, certifying only through 4-D before the bound turns positive – which is exactly why the sections below replace worst-case composition with a physics-structured, then learned, correction.

As a genuine, not proxy, head-to-head reference point: a companion verification effort in the accompanying repository independently trained and *monolithically* verified a neural Lyapunov function on this same network with CROWN/ α, β -CROWN linear-relaxation bound propagation [4, 6] – a materially different, generally tighter, bounding technology than the interval arithmetic used here. It certified cleanly at one generator and, without degradation, at two, with branch-and-bound subdomain count growing $\approx 5\times$ (roughly 10 to 35–61 across five seeds) and wall-clock per seed from ≈ 6 to ≈ 10 minutes – real growth, not yet an explosion. Three generators was explicitly deferred and never attempted, so this is the honest extent actually run, not a proven wall. The compositional reduction here reaches all nine.

Table 2: Held-out proved coverage, physics-only vs. trained vector correction, mean over three evaluation slices; zero-relaxation verifier ceiling for each.

Field	Coverage	Ceiling	Headroom
Physics only	23.2%	23.3%	0.1 pts
Trained vector correction	24.8%	25.3%	0.5 pts

Table 3: Interval branch-and-bound vs. independent sum-of-squares certificate, identical polynomial, all three evaluation slices.

Slice	BaB radius	BaB cov.	SOS radius	SOS cov.
(5, 6)	93.6°	27.8%	93.8°	28.0%
(4, 7)	88.9°	20.4%	88.5°	20.2%
(1, 9)	85.6°	17.7%	85.6°	17.7%
Mean		22.0%		22.0%

Lossy coupling defeats scalar corrections; a vector correction fixes it

With no correction term and the classical lossless energy function, the lossy residual enters h squared and positive, so it can only push the certified threshold down. A scalar residual added to the energy function can only contribute a gradient to the linear term of h , but the field required to cancel the lossy residual has curl $2G_{ij} \sin \delta_{ij}$ – measured at $\approx 98\%$ of the required correction’s magnitude on this system – and no gradient supplies a field that is almost entirely rotational. We confirmed this is not an optimization artefact by isolating the term: 700 steps of pure adversarial minimization with full network capacity converge to strictly *worse* than doing nothing (9.4% vs. 14.6% coverage; Table 5, row c8).

Moving the correction into a *vector* cross term removes the obstruction: cancelling the linear term becomes a pointwise algebraic condition on the field’s value, not a condition that it be a gradient. A shallow per-generator network, verified in-network by sound interval branch-and-bound and trained counterexample-guided from the physics baseline, reaches **24.8% certified coverage at 100% soundness** (Table 2), holding full structural margins – constraining the optimizer to retain them *improved* coverage rather than costing it. Measuring the largest radius any verifier could ever certify for this field, by dense evaluation of the exact condition, shows the proved result is within 0.5 percentage points of that ceiling: **the cap is the certificate’s own shape, not the tightness of the bound.**

An independent, exact certificate confirms the result

Every result above uses one verification technology and six architecturally distinct attempts to beat 24.8% (Table 5), all converging on the same band – suggestive, but every one of those attempts still used interval arithmetic somewhere. We built an independent, exact sum-of-squares (Positivstellensatz) certificate from the swing-equation physics directly: lifting a slice’s two varying angles to sine/cosine pairs makes the stability condition an exact low-degree polynomial (Methods, Eq. 7), and a degree-6 semidefinite program either finds an algebraic proof or does not, with no relaxation gap by construction. Run on all three evaluation slices against an identical, sum-of-squares-compatible baseline, it **ties interval branch-and-bound on every slice individually** (mean coverage 22.0% vs. 22.0%; Table 3) – to our knowledge the first exact, non-interval-relaxed confirmation in this line of work that a branch-and-bound-proved radius is not leaving verifiable region unclaimed.

Table 4: Coverage at each slice’s own fixed proved radius.

Slice	r^*	Classical shape	Naive disk	Convex fit
(5, 6)	92.5°	27.1%	41.6%	41.6%
(4, 7)	87.4°	19.8%	40.5%	40.5%
(1, 9)	84.5°	17.4%	37.3%	37.3%
Mean		21.4%	39.8%	39.8%

Table 5: Negative-result ledger. All coverage figures are 100% sound unless noted.

ID	Tested	Result
c8	Scalar potential residual	Strictly worse in isolation (9.4%)
c14	Exact closed-form cancellation, zero parameters	$b \equiv 0$ but a third term explodes; ceiling $\rightarrow 0.1\%$
c14b	Scalar blend with physics baseline	Flat, 23.6% peak vs. 23.3% baseline
c15/c15b	Potential-reshaping residual, frozen and jointly-trained cross term	Flat at 16.8%; best checkpoint = untrained start
c16	Full-state visibility, 2× capacity, no locality	17.4% adversarial, plateaued
c17	Free $V(\phi, \omega)$, no quadratic-in-speed restriction, full 18-D	Identical to baseline to one decimal (15.7%)

Coverage decomposes into an independent radius question and shape question

Coverage has depended on two logically separate quantities throughout: how far the stability condition can be proved (*radius*), and how well the proved region’s *shape* matches the true stable set at that radius. Crucially, coverage is defined entirely by the potential $U(\phi)$ ’s sublevel set – the cross term never enters that definition – yet six independent gradient-based searches to reshape U (Table 5, rows c15/c15b) found nothing distinguishable from noise, because gradient descent was solving both questions in one entangled, non-convex objective.

Decoupling them: fixing the radius at each slice’s already-proved boundary and posing the shape question alone as a single convex quadratic program (Methods, Eq. 8) – fit directly to Monte Carlo ground truth with a hard constraint against every sampled unstable point – recovers an **86% relative coverage gain** over the classical shape (mean 39.8% vs. 21.4%; Table 4), fully sound on all three slices, found by a single convex solve. Reported without rounding away the caveat: at every radius proved anywhere in this study, the sampled ground truth inside the disk is entirely stable (the nearest true unstable point on one slice is 114.9°, beyond any proved radius), so the fitted shape matches a plain disk almost exactly – a real, sound statement about how badly the classical shape under-uses a given proved radius, not yet evidence that shape optimization alone can extend the radius toward the true boundary.

Eight independently designed negative results bound the radius question

We report eight architecturally distinct attempts to extend the 24.8% radius, all negative, as findings rather than discarded drafts – their convergence is itself evidence about the certificate family’s limit (Table 5). The most structurally permissive of these (row c17) removes the very assumption – quadratic dependence on speed – that makes the closed-form reduction possible, verifying stability by direct automatic differentiation against the true nonlinear dynamics over the full, correctly-projected 18-D state. That even this maximally flexible certificate matches the fixed-structure baseline to one decimal place is the strongest single piece of evidence that the ceiling is a property of the verification problem on this system, not of any one certificate’s architecture.

Table 6: Second topology (IEEE/WSCC 9-bus), physics-only vs. learned field.

Field	Radius	Coverage	Sound
Physics only	88.5°	23.4%	100%
Learned cross term	104.7°	32.3%	100%

Generalization: operating-point robustness and a second topology

We reused this project’s own operating-point robustness method (scale equilibrium angle deviations from their mean, rebuild the linearized coupling, re-certify – avoiding the droop-only steady-state trap, since a raw loading change has no exact fixed point to linearize at) and applied it for the first time to the actual headline field, with no retraining. Result, reported exactly rather than smoothed into one claim: two of three slices are essentially fully robust across a $\pm 15\%$ loading band (under 5% change: $106.4^\circ \rightarrow 106.4^\circ \rightarrow 105.3^\circ$ and $105.3^\circ \rightarrow 103.1^\circ \rightarrow 100.8^\circ$); the third – slice (1,9), already the hardest throughout this study – collapses to 0° under either perturbation direction tested, revealing local overfitting to the exact nominal equilibrium there.

We also built a second, fully independent classical multimachine model – the IEEE/WSCC 9-bus 3-generator system [11] – from verified external sources rather than memory: topology and setpoints from MATPOWER’s official `case9.m`, generator inertia/reactance cross-confirmed from two independent sources, the AC power flow solved by `pandapower`, and the classical Kron reduction verified by two checks before anything downstream was trusted – an equilibrium-balance check (residual 10^{-16} after a units bug that conflated terminal and internal-node power was caught and fixed) and a Hurwitz stability check on the reference-grounded system (an earlier version showed exactly-zero eigenvalues at every damping level, traced to never grounding a reference generator – the same rotational-symmetry zero mode this project already discovered for the 39-bus system, not a new bug).

The physics-only baseline (23.4%) lands remarkably close to the 39-bus system’s own (23.2%) despite the networks sharing no topology or parameters, and the identical learned-cross-term procedure improves it to 32.3% – a 38% relative gain, larger than the primary system’s own 7%. Both mechanisms transfer to an independently constructed topology. Only the physics-baseline and learned-cross-term stages were re-run here, not the full SOS/shape-search arc or the coupling-LMI margin check (tied to the 39-bus data format) – evidence the approach generalizes, not a second complete experimental arc.

Discussion

The radius question is now unusually well bounded: eight independent methods, including one exact verification technology and the single most structurally permissive change available inside this certificate paradigm, converge on a $93\text{--}106^\circ$ band. The shape question, by contrast, is genuinely open – the convex program here is a first, minimal instance of a much larger family, and has not yet been tested at a radius where real boundary information exists to fit against. Extending the proved radius even modestly and re-running the same convex shape fit there is the most direct next step this study’s own evidence points toward. More broadly, this report now spans two topologies and a loading-condition band on the primary one, short of a systematic sweep; broader generalization and comparison against occupation-measure or reachability-based alternatives at this scale remain open.

Data availability

[TODO: repository URL and license]. All experiment scripts and raw results (JSON) referenced in this paper are version-controlled in the accompanying repository and available on request pending publication.

Code availability

[TODO: repository URL]. Code is released under [TODO: license].

Author contributions

[TODO: CRediT-style contribution statement].

Competing interests

[TODO: The authors declare no competing interests. / disclose as applicable.]

Acknowledgements

[TODO: funding / acknowledgements].

References

- [1] Chiang, H.-D., Chu, C.-C. & Cauley, G. Direct stability analysis of electric power systems using energy functions: theory, applications, and perspective. *Proc. IEEE* **83**, 1497–1529 (1995).
- [2] Chang, Y.-C., Roohi, N. & Gao, S. Neural Lyapunov control. In *Advances in Neural Information Processing Systems* (2019).
- [3] Dawson, C., Gao, S. & Fan, C. Safe control with learned certificates: a survey of neural Lyapunov, barrier, and contraction methods for robotics and control. *IEEE Trans. Robotics* (2023).
- [4] Zhang, H., Weng, T.-W., Chen, P.-Y., Hsieh, C.-J. & Daniel, L. Efficient neural network robustness certification with general activation functions. In *Advances in Neural Information Processing Systems* (2018).
- [5] Katz, G., Barrett, C., Dill, D. L., Julian, K. & Kochenderfer, M. J. Reluplex: an efficient SMT solver for verifying deep neural networks. In *Computer Aided Verification* (2017).
- [6] Xu, K. *et al.* Automatic perturbation analysis for scalable certified robustness and beyond. In *Advances in Neural Information Processing Systems* (2020).
- [7] Papachristodoulou, A. & Prajna, S. On the construction of Lyapunov functions using the sum of squares decomposition. In *IEEE Conference on Decision and Control* (2002).
- [8] Henrion, D. & Korda, M. Convex computation of the region of attraction of polynomial control systems. *IEEE Trans. Autom. Control* **59**, 297–312 (2014).
- [9] Athay, T., Podmore, R. & Virmani, S. A practical method for the direct analysis of transient stability. *IEEE Trans. Power Appar. Syst.* **PAS-98**, 573–584 (1979). [TODO: verify page range] – origin of the IEEE New England 39-bus test system.
- [10] Chow, J. H. *Power System Toolbox*. [TODO: verify citation form] – source of the NE39 network data used in the Kron reduction to ten generator buses.

- [11] Anderson, P. M. & Fouad, A. A. *Power System Control and Stability* 2nd edn (IEEE Press, 2003). – origin of the IEEE/WSCC 9-bus 3-machine test system; network data cross-checked against MATPOWER’s `case9.m`.

Methods

System model

The plant is

$$M_i \ddot{\delta}_i = P_i(\boldsymbol{\delta}) - d_i \dot{\delta}_i, \quad (1)$$

$$P_i(\boldsymbol{\delta}) = P_{m,i} - \sum_j [F_{ij} \sin \delta_{ij} + G_{ij} \cos \delta_{ij}] \quad (2)$$

with $\delta_{ij} := \delta_i - \delta_j$, F_{ij} the conservative susceptance coupling and G_{ij} the transfer conductance. The region of attraction is estimated by dense Monte Carlo integration (12s horizon, fourth-order Runge–Kutta, $\Delta t = 0.02$ s) from angle-only initial conditions on three held-out 2-D angle slices (generator pairs (5, 6), (4, 7), (1, 9)), with the remaining active angles pinned at equilibrium. Coverage is the fraction of Monte-Carlo-confirmed-stable grid points a certified set soundly claims, averaged over the three slices; soundness is 100% minus the percentage of certified points later found unstable.

Closed-form speed elimination

For $V = \frac{1}{2} \boldsymbol{\omega}^\top M \boldsymbol{\omega} + U(\phi) + \boldsymbol{\Psi}(\phi)^\top M \boldsymbol{\omega}$, differentiating along the dynamics gives $\dot{V} = \boldsymbol{\omega}^\top Q \boldsymbol{\omega} + b(\phi)^\top \boldsymbol{\omega} + c_0(\phi)$ with

$$Q = -\text{diag}(d) + \Omega \text{sym}(MJ), \quad J := \partial \boldsymbol{\Psi} / \partial \phi, \quad (3)$$

$$b_i = -L_i(\phi) - d_i \Psi_i(\phi), \quad L_i(\phi) := \sum_j G_{ij} (\cos \delta_{ij} - \cos \delta_{ij}^*), \quad (4)$$

$$c_0 = \boldsymbol{\Psi}(\phi)^\top P(\phi). \quad (5)$$

For $Q \prec 0$ this is concave in $\boldsymbol{\omega}$, so

$$h(\phi) := \max_{\boldsymbol{\omega}} \dot{V} = \frac{1}{4} b(\phi)^\top (-Q(\phi))^{-1} b(\phi) + c_0(\phi) \quad (6)$$

is closed form; proving $h(\phi) < 0$ over an angle-only box proves $\dot{V} < 0$ for every speed in that box.

Interval verification

$h(\phi)$ is verified by interval branch-and-bound with $-Q = D - E$ split into diagonal and off-diagonal parts, E dominated by a per-row Gershgorin bound in Loewner order for a sound closed-form upper bound per box; a small central cube around equilibrium is excluded (covered by a local Hessian argument plus dense sampling) since $h \rightarrow 0$ there by construction.

Sum-of-squares certificate

Lifting a slice’s two varying angles to (s_1, c_1, s_2, c_2) via product-to-sum identities and using $\Psi_i = \epsilon \sin \phi_i$ (bounded and polynomial, unlike the affine baseline), $H := 4q_1 q_2 h$ is an exact degree-5 polynomial. A degree-6 program searches for SOS multipliers σ_k , free multipliers λ_k , and a maximized slack γ such that

$$-H - \gamma = \sigma_0 + \sigma_1 g_1 + \sigma_2 g_2 + \lambda_1 h_1 + \lambda_2 h_2 \quad (7)$$

over the box/circle constraints g_k, h_k , implemented via Gram-matrix semidefinite programming (`cvxpy`, Clarabel/SCS solvers), with the central neighbourhood excluded by two region certificates whose union covers the annulus BaB itself proves over.

Convex shape search

$p(\phi) = z(\phi)^\top Q z(\phi)$, $Q \succeq 0$, z the degree- ≤ 2 monomial basis in (s_1, c_1, s_2, c_2) :

$$\min_{Q \succeq 0} \sum_{\text{stable}} \text{hinge}(p(x) - 1) \text{ s.t. } p(x) \geq 1 + m \ \forall \text{ unstable } x, \ p(x) \geq 1 + m \ \forall x \notin \text{disk}(r^*), \quad (8)$$

solved once per slice at r^* = that slice’s proved radius minus a 0.02 rad safety margin. Containment of $\{p \leq 1\}$ inside r^* is checked on a dense post-hoc grid, not by a Positivstellensatz containment certificate – the one non-rigorous ingredient in this study, on the same footing as the sampling-based central-cube argument used throughout.