Damping without tuning: the measurable stability threshold of the optimality-criteria iteration, and a controller that tracks it

Hyeongseok Kim
Independent Researcher, Uiwang, Republic of Korea

Preprint — August 17, 2026

Abstract

Every density-based topology optimization run using the optimality-criteria (OC) update makes one algorithmic choice: the damping exponent η, folklore-set to $\frac{1}{2}$. We measure what it controls. Exactly projected finite-difference Jacobians of the full update, across meshes, boundary conditions, volume fractions, filters and penalizations, give three empirical laws. *One:* oscillation is a period-doubling instability at $\eta^\star = 2/s_{\max}$, the top of the spectrum of the log–log stiffness of the filtered sensitivity field; the local multipliers $\mu = 1 - \eta s$ predict measured onset and decay rates to three decimals. *Two:* that spectrum saturates at $p{+}1$ on top while its lower edge sits at $s \approx 0$, and an edge at zero pins the min–max (Richardson) optimum onto the flip boundary itself: at $p = 3$ the folklore value is that optimum. The content of the coincidence is the spectral shape, not the optimum's worst-case rate, which the edge modes leave nearly vacuous; the damping's real work is done on the bulk at $s \approx 1$. *Three:* the observed $s_{\max}$ belongs to the fixed point the dynamics selects, and standard damping self-limits to $s_{\max} \leq 2/\eta$, separating from the naturally saturated family exactly at $p = 3$. The threshold is thus a moving target: under penalization continuation, fixed $\frac{1}{2}$ ends oscillating in 12 of 32 runs. It need not be. The trajectory estimators $\hat{\mu}, \hat{s}$ read the flip branch precisely when it is active and the $s \approx 0$ creep edge otherwise—a regime dependence that estimator-driven relaxation schemes do not account for—so a four-line additive-increase/multiplicative-decrease controller can track the threshold unaided: with one frozen parameter set it converges in all 64 continuation runs, matches the best hand tuning across an 11-configuration suite, and composes with Anderson acceleration so that each covers the other's failures.

1 Introduction

The optimality-criteria (OC) update is the workhorse of density-based topology optimization: rescale each element density by a power of the ratio between its strain-energy sensitivity and a volume multiplier,

$$x_e^+ = \operatorname{clip}(x_e B_e^\eta), \qquad B_e = \frac{-\partial c/\partial x_e}{\lambda\, \partial V/\partial x_e}, \tag{1}$$

with move limits, a bisection on λ for the volume constraint, and a damping exponent η [Bendsøe, 1989]. The educational codes that shaped the field hard-wire $\eta = \frac{1}{2}$ [Sigmund, 2001, Andreassen et al., 2011]; commercial documentation codifies the folklore that smaller exponents are more stable but slower [Abaqus]. When a run oscillates, practitioners lower η, enlarge the filter, or switch filters—by hand, per problem, without a quantity to aim for.

This paper supplies the quantity and then removes the hand. Viewed as a discrete dynamical system in logarithmic coordinates, the damped OC map has local multipliers obeying the affine law

$$\mu_m \;=\; 1 - \eta\, s_m, \tag{2}$$

where s_m are eigenvalues of the projected negative log–log sensitivity operator on the volume tangent space. Oscillation is a flip (period-doubling) instability of the modes with $s_m > 0$, with threshold $\eta^\star = 2/s_{\max}$; a statically determinate two-bar model gives $s = p{+}1$ exactly, placing the standard pair $(p, \eta) = (3, \frac{1}{2})$ exactly on the flip boundary. All of this is measured, not asserted: on instrumented variants of the 88-line code we build exactly projected finite-difference Jacobians of the full update and verify the law at the operator level, to three decimals, across meshes, boundary conditions, volume fractions, filters, and penalizations.

Three measured facts about that spectrum then decide how η should be chosen—and show that no fixed choice suffices.

First: the folklore is optimal, which explains both its success and its fragility. Every spectrum we measured has the same shape: bulk near $s \approx 1$, lower edge at $s \to 0^+$, upper tail at $s_{\max}$. For such spectra the min–max optimal damping—the classical Richardson step—coincides with the flip boundary (Section 5.3). At the canonical benchmark the optimum computed from the full measured spectrum is $\eta_{\text{opt}} = 0.499$: the folklore value is not a lucky guess but the optimum, sitting by necessity at the edge of instability. The coincidence is worth stating precisely, because its content is the spectral shape rather than the arithmetic: with the lower edge at zero the min–max optimum *must* land on the flip boundary, its worst-case rate is nearly 1 (the edge modes barely move at any η), and the damping's real work is done on the bulk at $s \approx 1$. Section 5.3 makes all three points quantitative, including how the optimum shifts when the near-zero modes are excluded. What the picture settles is the alternative: lowering η stabilizes the top of the spectrum and slows the bulk in equal measure, so "more stable but slower" is one trade-off with one constant.

Second: the target moves. The spectrum belongs to the design, and the design changes—within a run, as capacity adapts, and by construction under penalization continuation, where the natural family's threshold falls as $2/(p{+}1)$: from 1.0 at $p = 1$ through 0.5 exactly at $p = 3$ to 0.333 at $p = 5$. A fixed exponent cannot know this. In a continuation suite spanning four configurations, two ramp schedules and up to four starts each, fixed $\eta = \frac{1}{2}$ ends oscillating in 12 of its 32 runs—on one configuration because the ramp outruns settling, on another because the destination's own measured threshold (0.333) lies below $\frac{1}{2}$, which no schedule can repair (Section 6.3). Lowering η instead trades the failure for its dual: fixed $\eta = 0.3$ never oscillates there but carries the largest median compliance penalty of any converged method at a fixed budget.

Third: the iteration can feel its own threshold. The trajectory estimators $\hat{\mu}$ (dominant multiplier) and $\hat{s}$ (directional stiffness) cost nothing and satisfy the identity $\hat{\mu} = 1 - \eta\hat{s}$ in every run we measured. Which mode they read is regime-dependent, and usefully so: converging runs read the $s \approx 0$ creep edge—so a naive adaptive rule $\eta = c/\hat{s}$ diverges—while flip-active runs lock onto the flip branch, with $\eta\hat{s} = 2$ to three decimals in every saturated oscillation (Section 5.2). Gating $\hat{s}$ by the sign of $\hat{\mu}$ therefore yields a reliable threshold sensor, and a four-line controller built on it tracks the moving target: in the same continuation suite it converges in all 64 of its runs, following $2/(p(t){+}1)$ with a small lag that closes at the hold.

Our contributions:

1. **Three measured laws.** (a) The multiplier law (2) with the flip boundary $\eta^\star = 2/s_{\max}$ of the positive branch, verified at the operator level via exactly projected finite-difference Jacobians (η-invariant spectra; onset brackets containing $2/s_{\max}$ in every configuration; rates matching $|1 - \eta s_{\max}|$ with median deviation 0.001 over 42 probes), with exact two-bar models for both signs of s. (b) Spectral shape and its consequence: $s_{\max}$ saturates at $p{+}1$ while the lower edge sits at $s \approx 0$, so the Richardson optimum coincides with the flip boundary—the folklore $\eta = \frac{1}{2}$ is optimal at $p = 3$, and the three folk

remedies for oscillation (lower η, larger filter radius, density filtering) are one inequality, $\eta\, s_{\max} < 2$. (c) Fixed-point dependence: the observed $s_{\max}$ belongs to the fixed point the dynamics selects; standard damping self-limits to $s_{\max} \le 2/\eta$, separating from the naturally saturated family exactly at $p{+}1 = 2/\eta$, i.e. $p = 3$ for $\eta = \frac{1}{2}$.

2. **A free threshold sensor.** The trajectory estimators and their two regimes: the identity $\hat{\mu} = 1 - \eta\hat{s}$ holds in the wild (28 runs, three decimals), converging trajectories read the $s \approx 0$ edge, flip-active trajectories read the flip branch ($\eta\hat{s} = 2$, fourteen of fourteen saturated runs)—the dichotomy that makes the sensor usable only under a sign gate, and sufficient there.

3. **A tuning-free controller.** An additive-increase/multiplicative-decrease (AIMD) rule whose thresholds all live in multiplier space (no absolute amplitude scales), with one frozen constant set across every experiment in this paper. Benchmarks: single-problem rescues of over-driven runs; an 11-configuration sweep (mesh, boundary conditions, volume fraction, filter type and radius, $p \in \{3, 5\}$) with a tuning-failure-rate table against fixed damping, a practitioner ladder, and our MMA implementation; a penalization continuation where the fast ramp breaks fixed $\eta = \frac{1}{2}$ while the controller tracks the falling threshold point-for-point; and composition with Anderson acceleration, where either technique alone fails on some starts and the composition failed on none we tested.

Everything regenerates from the driver scripts in the public repository accompanying this paper; no claim rests on the toy models alone.

2 Related work

Convergence of the problem formulation. Rietz [2001], Stolpe and Svanberg [2001], and Martínez [2005] address the SIMP *problem*—penalization sufficiency, interpolation, convergence properties. These are orthogonal to our subject, the dynamics of the solver.

Resizing iterations as dynamical systems. The closest precedent is the fully-stressed-design literature: Mueller and Burns [2001] showed that some fully stressed frames are unstable fixed points of the stress-ratio iteration, with follow-ups on multiple load paths [Mueller et al., 2002, Liu and Burns, 2003], in a lineage reaching back to Razani [1965] and engineering variants [Patnaik et al., 1998]. They established "resizing iteration = dynamical system" for trusses and frames at $\eta = 1$, without a multiplier law, thresholds, or measured spectra; we import the viewpoint into density-based topology optimization and make it quantitative. Li et al. [2020] treated the OC iteration explicitly as a fixed-point map and accelerated it with periodic Anderson extrapolation; the oscillations they mitigate are the flip branch we analyze, and Section 6.5 shows the two approaches compose.

The statics of η. What η *is* in approximation terms is settled: Groenwold and Etman [2008] proved damped OC equivalent to dual sequential approximate optimization with exponential intervening variables, $\eta = 1/(1-a)$, identifying $\eta = \frac{1}{2}$ with the reciprocal approximation [see also Groenwold et al., 2007, Groenwold and Etman, 2009]. Svanberg's CCSA established that non-conservative separable approximations can cycle [Svanberg, 2002], with MMA the canonical realization [Svanberg, 1987]. None of these works analyzes the damped map's dynamics or measures its spectra; we attribute the η–a dictionary entirely to them and supply the dynamical complement. MMA's asymptote adaptation is itself a sign-based stabilization heuristic—per-element sign flips shrink the asymptote interval ($\times 0.7$), monotone progress widens it ($\times 1.2$)—which is dimensionless but uncalibrated: it does not know how far to retreat. Our controller differs in what it measures and what it moves: the measured multiplier calibrates the retreat ($\eta \leftarrow c/\hat{s}$ lands a known distance inside the boundary), and Section 6.4 shows that moving the step size instead of the damping bounds amplitude without removing the instability.

Estimating a relaxation parameter from successive iterates. Equation (5) belongs to a long lineage. Aitken's Δ^2 process, in the vector form of Irons and Tuck [1969], estimates a dominant multiplier from consecutive residuals and rescales the update accordingly; the same idea drives Aitken dynamic relaxation in partitioned fluid–structure interaction, where the relaxation factor is re-estimated at every step [Küttler and Wall, 2008], and, in optimization, the Barzilai–Borwein step sizes, which read a local curvature from two successive gradients [Barzilai and Borwein, 1988]. Anderson acceleration [Li et al., 2020] is the multi-secant member of the family. Two things distinguish what we do. First, the parameter: those methods adapt an *additive* relaxation weight or step length, whereas η is an *exponent*, and the object it must respect is the flip threshold $2/s_{\max}$ of the log-coordinate spectrum—a quantity none of these schemes estimates or targets. Second, and more practically, our measurements show that the estimate is only sometimes trustworthy: a trajectory that is converging reports the spectrum's $s \approx 0$ edge, not its top (Section 5.2), so a rule that consumes the estimate unconditionally—as Aitken relaxation and Barzilai–Borwein do by construction—diverges here. The regime gate of Section 6.1 is what the multiplier law buys over estimator-driven relaxation in general.

Modern convergent algorithms. Ananiev [2005] interpreted OC as projected gradient; SiMPL builds provably convergent mirror descent with Fermi–Dirac entropy [Kim et al., 2025, Keith et al., 2025]. We do not propose a replacement for these methods: we remove the one free parameter of the update everybody already runs, on the problem class where OC applies (compliance with a single volume constraint). Filtering enters through the standard constructions [Sigmund and Petersson, 1998, Bourdin, 2001].

3 The multiplier law

3.1 Log coordinates and the tangent-space spectrum

Write $z_e = \ln x_e$ and let $\tilde{D}$ be the filtered sensitivity. On the free set (elements pinned by neither box nor move limits) the pure OC step is

$$z^+ \;=\; z + \eta\left(\ln \tilde{D}(z) - \ln \lambda(z)\,\mathbf{1}\right), \tag{3}$$

an exponentiated-gradient-type update; the λ-bisection enforces the volume constraint to its tolerance (10^{-10} here), so the map's image lies on the volume isosurface. Linearizing at a fixed point gives, on the volume tangent space,

$$J_T = I - \eta\,\mathcal{S}, \qquad \mathcal{S} = -\,P\,\frac{\partial \ln \tilde{D}}{\partial z}, \tag{4}$$

so each eigenvalue s_m of $\mathcal{S}$ contributes $\mu_m = 1 - \eta s_m$: the law (2). Because the image lies on the isosurface, the constraint gradient w (in log coordinates $w_e = x_e\,\partial V/\partial x_e$ on the free set; $w_e = x_e$ for sensitivity filtering, $w_e = x_e \sum_i H_{ie}/\mathcal{H}_i$ for density filtering) is an exact left null vector of the unprojected Jacobian, $w^{\mathsf{T}} J = 0$, whence $\operatorname{eig}(J) = \{0\} \cup \operatorname{eig}(Q^{\mathsf{T}} J Q)$ for any orthonormal basis Q of $w^{\perp}$. The spurious $\mu = 0$ eigenvalue masquerades as $s = 1/\eta$—inside the physical range—so spectra must be exactly projected (Section 4). Move limits and clipping are inactive near interior fixed points: they leave thresholds unchanged and only bound post-onset amplitudes.

3.2 Two exact models, two branches

Proposition 1 (series model: flip branch)**.** *For $c(x) = \sum_e a_e x_e^{-p}$ (statically determinate; no force redistribution) with $\sum_e x_e = V$, the pure OC map is exactly affine in relative log coordinates with multiplier $\mu = 1 - \eta(p+1)$: geometric convergence for $0 < \eta < 2/(p+1)$, deadbeat at $\eta = 1/(p+1)$, flip boundary at $\eta = 2/(p+1)$.*

The standard pair $(3, \frac{1}{2})$ lies exactly on this boundary, and $(1, \frac{1}{2})$ recovers the one-step character of classical fully stressed resizing. In the toys these statements hold to machine precision, and adding clipping, move limits and the bisection changes nothing (Fig. 1).

Proposition 2 (parallel model: winner-take-all branch)**.** *For $c(x) = F^2/\sum_e a_e x_e^p$ (redundant, force-sharing), the symmetric point has $z^+ = (1 + \eta(p-1))z$ exactly: $s = -(p-1) < 0$, a monotone repulsion for every $\eta > 0$.*

Both propositions hold to machine precision in the toy: measured contraction rates match $|1 - \eta(p+1)|$ for the series model (Fig. 1b), and for the parallel model at $p = 3$, $\eta = \frac{1}{2}$ the measured z-multiplier is 2.000 against the predicted $1 + \eta(p-1) = 2$, with the pair reaching the bounds in six steps.

Remark 1 (branch-wise stability language)**.** *For complex s_m, $|1 - \eta s_m| < 1$ iff $\operatorname{Re} s_m > 0$ and $\eta < 2\operatorname{Re} s_m/|s_m|^2$. Real spectra contain both signs, so there is no global stability window: $2/s_{\max}$ bounds the* flip branch*; modes with $s_m < 0$ grow slowly at every η. In our measurements the negative branch appears as a benign, self-terminating symmetry-breaking channel; it changes no threshold statement and no controller decision, and we defer its analysis—growth at the predicted rate to four decimals, termination by resettling into an asymmetric fixed point—to Appendix A.*

3.3 The SAO dictionary

Groenwold and Etman [2008] proved damped OC equivalent to dual SAO with exponential intervening variables $y = x^a$ under $\eta = 1/(1-a)$; $a = -1$ (reciprocal) $\Leftrightarrow$ $\eta = \frac{1}{2}$. For the series model the exact fit $a = -p$ maps to the deadbeat $\eta = 1/(p+1)$: implicit in their framework, made explicit here as the midpoint of the stability window. At $p = 3$ the reciprocal approximation carries exactly half the true curvature, and half the curvature is, dynamically, the multiplier $\mu = -1$. The equivalence is theirs; the dynamical reading and the measurements are ours.

4 Measuring the threshold

All measurements run on instrumented variants of the 88-line code [Andreassen et al., 2011]: 2D compliance, SIMP with $E_{\min} = 10^{-9}$, move limit 0.2, box $[10^{-3}, 1]$, sensitivity or density filtering, bisection tolerance 10^{-10}. The canonical configuration R0 is the 60×20 MBB beam, volume fraction 0.5, $p = 3$, sensitivity filter $r_{\min} = 1.1$. A full step costs 9 ms at 60×20, 54 ms at 120×40.

Exactly projected Jacobians (the validation instrument). At a converged design we build the Jacobian of the entire update by log-coordinate finite differences over the free set ($\varepsilon = 10^{-5}$), project with a Householder basis of $w^{\perp}$, and diagonalize $Q^{\mathsf{T}}JQ$, reporting $s = (1-\mu)/\eta$. The block structure is verified directly: the row-null residual is $\|w^{\mathsf{T}}J\|/\|w\| = 2.5 \times 10^{-3}$ on V3 (1.2×10^{-3} relative to $\|J\|_2$; central differencing moves it only to 2.0×10^{-3} and the top-six spectrum by $\leq 1.1 \times 10^{-5}$, so the floor is the update's piecewise-smooth structure, not truncation), and the raw spectrum contains the predicted constraint artifact at

$s = 1/\eta$ to six digits (1.999990 at $\eta = \frac{1}{2}$). For large free sets an ARPACK operator variant with mean-centering deflation reproduces the dense $s_{\max}$ (fifty matrix-vector products versus a full build on the 120×40 problem); eigenpair authenticity is checked with independent-ε residuals and Rayleigh quotients. These spectra are the ground truth against which the cheap sensors below are judged.

Trajectory estimators (the sensor). On any trajectory, windowed regressions of successive centered log-increments give the dominant multiplier $\hat{\mu}$ and directional stiffness $\hat{s}$

$$\hat{\mu}_t = \frac{\langle \mathrm{c}(\mathrm{d}z_t), \mathrm{c}(\mathrm{d}z_{t-1})\rangle}{\|\mathrm{c}(\mathrm{d}z_{t-1})\|^2}, \qquad \hat{s}_t = -\frac{\langle \mathrm{c}(\Delta \ln \tilde{D}_t), \mathrm{c}(\mathrm{d}z_{t-1})\rangle}{\|\mathrm{c}(\mathrm{d}z_{t-1})\|^2}, \tag{5}$$

($\mathrm{c}(\cdot)$ = mean-centering on the interior mask). They cost nothing beyond quantities the update already computes, and they are the OC-specific member of the classical family of estimators read from successive iterates (Aitken/Irons–Tuck, Barzilai–Borwein, Anderson; Section 2): $\hat{\mu}$ is a power-iteration estimate of the dominant multiplier, $\hat{s}$ its translation into spectral coordinates through (2). We are explicit about a circularity: since both are read from the same trajectory, the identity $\hat{\mu} = 1 - \eta\hat{s}$ validates the multiplicative structure whichever mode the trajectory occupies—it does not certify that $\hat{s}$ reads $s_{\max}$. Which mode it reads is measured in Section 5.2, and is the design basis of the controller.

Behavioral thresholds. Mode-aligned probes kick a converged design along the measured leading eigenvector ($\delta = 0.005$) and track the signed mode-component ratio in x-space on mid-density elements; scanning η on a grid of width 0.08 brackets the onset. Two paid lessons: unkicked runs cannot see the onset (no seed of the unstable mode), and the probe must live in x-space, because the multiplicative drift of low-density gray elements ("log-creep") swamps log-norms. That creep is itself the theory: the spectrum's lower edge $s \to 0^+$ has $\mu \to 1$ at every η, so no damping removes it—persistence of the low-density tail is the $s \approx 0$ edge made visible.

Convergence protocols. Because measured thresholds turn out to be properties of the fixed point reached (Section 5.4), the protocol producing the design is part of the measurement: *reference* (converge at $\eta_0 = 0.3$) for saturation statements, *adapted* ($\eta_0 = 0.5$, stepping down only on oscillation) mimicking standard practice. Every number in this paper regenerates from the released driver scripts.

5 The threshold in practice

5.1 Operator-level validation

Across 28 instrumented runs ($\eta \in [0.3, 1.6]$, two filter radii) the free-set estimators satisfy $\hat{\mu} = 1 - \eta\hat{s}$ to three decimals in every run (Fig. 2). At the operator level, Jacobians built at the same design with $\eta = 0.5$ and 0.9 give s-spectra identical to three decimals (top five 3.974/3.932/3.893/3.784/3.763 at R0). This η-invariance is by construction—in (4) the operator $\mathcal{S}$ does not contain η—so we report it as an implementation-consistency check that the full code path (solver, filter, bisection, clipping) realizes the structure, not as a discovery. The R0 spectrum saturates just below the series-model value, $s_{\max} = 3.974 \approx p{+}1$, so $\eta^\star = 0.503$: the folklore damping operates 0.003 from the flip boundary. (This margin is an operator-level statement; the behavioral brackets, of width 0.08, cannot resolve margins of this size by construction.) The leading eigenmode coincides with the observed oscillation pattern ($|\cos| = 0.93$ and 0.96 at $r_{\min} = 1.1$ and 2.4), whose spatial character—localized joint

Table 1: Threshold law across configurations ($p = 3$, move 0.2; exactly projected FD Jacobians; onset brackets from mode-aligned probes, grid width 0.08).

case	configuration	$\lvert F\rvert$	$s_{\max}$	$\eta^\star$	onset bracket
R0	60×20 MBB, sens. $r_{\min}$1.1, vf .5	367	3.974	0.503	$(0.5, 0.6]$
V1	120×40 MBB, sens. $r_{\min}$1.1	902	3.625	0.552	$(0.45, 0.55]$
V2	120×40 MBB, sens. $r_{\min}$2.2	1485	3.850	0.519	$(0.479, 0.559]$
V3	60×20 cantilever, sens. $r_{\min}$1.1	352	3.486	0.574	$(0.534, 0.614]$
V4	60×20 MBB, sens. $r_{\min}$1.1, vf .4	281	2.888	0.693	$(0.653, 0.733]$
V5	60×20 MBB, density $r_{\min}$2.4	38	0.370	≈ 5.41	$(5.4, 6.0]$

patch versus global two-lobe—explains why sustained oscillation at large filter radius destroys compliance (312.7 versus 217.1 at $\eta = 1.0$, 400 iterations) while at small radius it barely dents it (196.4 versus 195.6) (Fig. 3).

Robustness: Table 1 sweeps mesh, boundary conditions, volume fraction and filter type. Every behavioral onset bracket contains $2/s_{\max}$; sub-threshold decay rates match $|1-\eta s_{\max}|$ to three decimals (V3: 0.583/0.583, 0.862/0.862, 1.140/1.141; V4: 0.886/0.886); mesh refinement at fixed physical radius moves $s_{\max}$ by -3%. Margins are configuration-dependent: 0.6% at R0, 4–13% across mesh/BC variants, 28% at volume fraction 0.4.

5.2 The two regimes of the trajectory sensor

Which mode does $\hat{s}$ read? Sweeping $\eta \in [0.3, 1.6]$ at both filter radii (Fig. 4a): *converging* runs read the creep edge, $\hat{s} \in [-0.04, 0.15]$—uninformative about $s_{\max}$, and fatal for any naive rule $\eta = c/\hat{s}$; *saturated* runs lock onto the flip branch with $\eta\hat{s} = 2.000$ to three decimals in all fourteen oscillating runs; in the *transition* band, $\hat{\mu}$ turns negative and $\hat{s}$ jumps toward $s_{\max}$. For a clean period-2 the saturated relation is the estimator identity evaluated at $\hat{\mu} = -1$, so its empirical content deserves stating plainly: the saturated attractor is a clean period-2—in none of our oscillating runs (the 28-run identity sweep, the 14-run regime sweep, and every post-onset probe) did we observe period-4 or higher, the tail multiplier sitting at -1.000 with a stationary two-phase pattern. We tested this directly rather than attributing it to clipping: lowering the box floor from 10^{-3} to 10^{-6} leaves every settled attractor a clean period-2 (period-2 residual below 10^{-6} of the amplitude for $\eta \in [0.9, 1.6]$), so the lower bound is *not* responsible. We therefore state only what we observe: no cascade is reached over the range we drive ($\eta \leq 1.6$, and the toy bifurcation diagram of Fig. 1(c) likewise shows two branches throughout); whether a bounded update would arrest one remains untested. The sensor reads the period-2 state exactly. The sign of $\hat{\mu}$ is therefore a gate: when it signals flip, $\hat{s}$ is trustworthy; otherwise it must be ignored.

5.3 Optimality of the folklore

The spectra do more than bound η; they select it. Every spectrum in this paper has bulk mass near $s \approx 1$ (median 0.95–1.0), lower edge $a = \min s_+ \in [0.004, 0.24]$, upper edge $b = s_{\max}$. Choosing η is then the classical Richardson problem: the minimizer of the worst-case rate $\max_m |1 - \eta s_m|$ over the positive branch is $\eta_{\mathrm{opt}} = 2/(a + b)$—numerically confirmed to four decimals against $\arg\min_\eta \max_m |1 - \eta s_m|$ on all eighteen measured spectra—and since $a/b \leq 0.02$ in seventeen of the eighteen, η_{opt} coincides with the flip boundary $2/s_{\max}$ to 1–2% (Fig. 4b). At R0, $\eta_{\mathrm{opt}} = 0.499$. *Marginality is the signature of optimality*: with the lower edge at zero, an optimally damped OC iteration must operate at the flip boundary. The same picture supplies the folklore's other half: at $s \approx 1$ the per-step rate is 0.5 at $\eta = \frac{1}{2}$ versus 0.75 at $\eta = \frac{1}{4}$—a factor 4×10^{17} in residual after 100 iterations. One trade-off, one constant.

Table 2: Penalization sweep (60×20 MBB, sens. $r_{\min}$1.1, vf 0.5). Onset brackets (grid 0.08) containing the corresponding $2/s_{\max}$ are marked ✓. Every adapted run converged at its initial $\eta_0 = 0.5$ (ladder never engaged); reference runs used $\eta_0 = 0.3$.

	reference ($\eta_0 = 0.3$)				adapted ($\eta_0 = 0.5$)	
p	$s_{\max}$	$\frac{s_{\max}}{p+1}$	$\eta^\star$	bracket	$s_{\max}$	bracket
1.5	2.403	0.961	0.832	$(0.792, 0.872]$ ✓	2.402	$(0.793, 0.873]$ ✓
2.0	2.927	0.976	0.683	$(0.643, 0.723]$ ✓	2.916	$(0.646, 0.726]$ ✓
2.5	3.394	0.970	0.589	$(0.549, 0.629]$ ✓	3.425	$(0.544, 0.624]$ ✓
3.0	3.976	0.994	0.503	$(0.463, 0.543]$ ✓	3.974	$(0.463, 0.543]$ ✓
3.5	4.465	0.992	0.448	$(0.408, 0.488]$ ✓	3.255	$(0.574, 0.654]$ ✓
4.0	4.977	0.995	0.402	$(0.362, 0.442]$ ✓	3.841	$(0.401, 0.481]$ †
4.5	5.464	0.993	0.366	$(0.326, 0.406]$ ✓	3.706	$(0.500, 0.580]$ ✓
5.0	5.969	0.995	0.335	$(0.300, 0.370]$ ✓	3.519	$(0.528, 0.608]$ ✓

† shallow convergence (step amplitude 4×10^{-5}) contaminated this probe; bracket edge at $|r| = 1.002$.

Three qualifications temper the result. First, with $a/b \leq 0.02$ the statements $\eta_{\text{opt}} = 2/(a+b)$ and $\eta^\star = 2/b$ are numerically the same claim; the content is the measured spectral shape (a lower edge at zero), not the coincidence as such. Second, the min–max value at the optimum, $(b-a)/(b+a) = 0.96$–1.0, is nearly vacuous—the edge modes barely converge at any η; the modes that do the work sit in the bulk at $s \approx 1$, where $\eta = \frac{1}{2}$ delivers rate 0.5. Third, the optimum depends on counting the creep modes:

lower cutoff on s	0	0.1	0.3	0.5
η_{opt}, R0	0.499	0.490	0.467	0.444
η_{opt}, median over 18 spectra	0.551	0.543	0.523	0.499

The folklore value is exactly optimal for the full measured spectrum and within 12% of optimal under every truncation we tried; the qualitative conclusion—the optimum sits at the flip boundary because the lower edge pins it there—survives every cutoff.

5.4 Saturation, self-limiting, and fixed points

Table 2 and Fig. 5 sweep $p \in [1.5, 5]$ under both protocols. Three findings. *(a)* Under the reference protocol $s_{\max}/(p+1) \in [0.961, 0.995]$ across the whole range: the series-model envelope is a near-identity, with no boundary in the range tested (an earlier preview of ours reporting a $p = 5$ breakdown traced to an undocumented protocol and is retracted). *(b)* Under the adapted protocol the measured $s_{\max}$ agrees for $p \leq 3$, then pins in $[3.26, 3.84]$, bounded by the capacity ceiling $2/\eta_0 = 4$; the step-down ladder never engaged. The protocols separate exactly at $p+1 = 2/\eta_0$: $(3, \frac{1}{2})$ *is the largest penalization whose naturally saturated fixed point standard damping can still reach*; beyond it the dynamics drifts to fixed points it can hold—a survivor bias that explains why $\eta = \frac{1}{2}$ "keeps working" at $p > 3$. *(c)* At $p = 5$ the two protocols deliver two fixed points of the same problem ($s_{\max} = 5.969$ and 3.519) whose measured onsets each contain their *own* $2/s_{\max}$ (0.335 and 0.568): the threshold is a property of the fixed point the dynamics selects. Pooling both protocols, all flip-active probes match the rate law with median deviation 0.001 (42 probes; maximum deviation 0.154, the single shallow-convergence probe flagged with †); the only complex leading pair ($p = 1.5$) shifts the threshold by under 10^{-3}.

Table 3: Filter-radius detuning (reference protocol; gray-element counts in parentheses).

$r_{\min}$	1.1	1.5	2.0	2.4
$s_{\max}$, $p = 3$	3.976 (276)	3.261 (284)	3.031 (336)	3.359 (443)
$s_{\max}$, $p = 5$	5.969 (237)	3.893 (225)	3.421 (263)	2.954 (362)

5.5 Filtering detunes the flip branch

Density filtering collapses the design-space spectrum ($s_{\max} = 0.370$, $\eta^\star \approx 5.41$; no oscillation for $\eta \le 4$ across three seeds; behavioral onset bracketed in $(5.4, 6.0]$)—strong detuning, not elimination (a tiny negative eigenvalue survives). Sensitivity filtering detunes with increasing radius as well—monotonically at $p = 5$, and up to $r_{\min} = 2.0$ at $p = 3$—more strongly at high penalization (Table 3): at $p = 5$, $s_{\max}$ falls from 5.97 ($r_{\min} = 1.1$) to 2.95 (2.4); at $p = 3$ the drop is 15–24%. Gray-element counts *rise* while $s_{\max}$ falls (monotonically at $p = 5$; at $p = 3$ up to $r_{\min} = 2.0$, beyond which the tail is non-monotonic), and the localization is directly measurable: the leading mode's participation ratio is 10.8 elements at $r_{\min} = 1.1$ (90% of $|v|^2$ in 10 of 360 free elements), delocalizing to 47.7 (93 elements) at $r_{\min} = 2.4$ as $s_{\max}$ falls. The mechanism is spatial averaging of a localized substructure—measured, not inferred. The non-monotonic $p = 3$ tail at $r_{\min} = 2.4$ re-verifies at deeper convergence (step amplitude 10^{-11}) and remains unexplained. Practically: the three folk remedies—lower η, larger radius, density filter—are one remedy, $\eta\, s_{\max} < 2$.

6 The controller

6.1 Design

Section 5.2 dictates the structure: the sensor is trustworthy exactly when it reports danger. This is the setting of additive-increase/multiplicative-decrease control (as in TCP congestion avoidance): probe gently, retreat decisively, and let the retreat be calibrated by the sensor.

$$\begin{array}{lll} \textbf{if } \hat{\mu}_w < -(1-\delta) \text{ and } \hat{s}_w > 0.1: & \eta \leftarrow c/\hat{s}_w & \text{(MD; cooldown 5)} \\ \textbf{elif } \hat{\mu}_w > 0.9 \text{ and } |\hat{s}_w| < 0.3: & \eta \leftarrow \eta & \text{(creep: hold)} \\ \textbf{else}: & \eta \leftarrow \eta + \alpha & \text{(AI)} \end{array} \tag{6}$$

with windowed medians ($w = 4$) of the per-step estimators, a warm-up of 25 iterations at the initial η, and $c = 1.7$, $\delta = 0.25$, $\alpha = 0.01$ *frozen across every experiment in this paper*. The design constants interlock: after MD the multiplier is $1 - c = -0.7$, outside the gate $-(1-\delta) = -0.75$, so a retreat does not re-trigger itself; near onset $\hat{s} \approx s_{\max}$ makes $c/\hat{s}$ the calibrated jump below threshold, while in saturation $\hat{s} = 2/\eta$ makes the same rule a geometric back-off ($\eta \leftarrow (c/2)\eta$)—the controller needs no knowledge of which regime it is in. The hold state uses the creep signature ($\hat{\mu} \to 1$, $\hat{s} \to 0$) rather than an amplitude tolerance: *every amplitude threshold in the controller lives in multiplier space*, which the law makes problem-independent; what remain are timescale constants (warm-up, window, cooldown, probe rate), measured in iterations—and Section 6.4 will show that the residual failures arise precisely there. (An earlier variant froze on an absolute step amplitude and failed on the 120×40 mesh, whose creep floor sits above any fixed tolerance; the regime test removes the scale. A slower probe rate $\alpha = 0.005$ was also tried: it polishes single-run behavior but degrades the Anderson composition of Section 6.5, whose interference cycling the gate must catch early; we freeze $\alpha = 0.01$. These three design iterations are the complete tuning history of the controller.)

6.2 Single-problem behavior

Fig. 7 shows the controller on the canonical problem ($r_{\min} = 1.1$) and the hard-filter problem ($r_{\min} = 2.4$), from $\eta_0 \in \{0.3, 0.5, 1.0\}$, against fixed damping, with a 400-iteration budget. From all six starts the controller converges, with landing points that are the problem's own despite the threefold spread of starts. The right comparison for those landings is not any nominal threshold but each endpoint's own: exactly projected spectra at the six endpoints give landing ratios $\eta_{\text{end}}/(2/s_{\max}^{\text{end}})$ of 0.77–0.95 in five runs and 1.01 in the sixth—the controller parks at, or within measurement resolution of, the boundary of the fixed point it actually reached, which is where the min–max optimum of Section 5.3 says an optimally damped iteration should sit. Fixed $\eta = 1.0$ oscillates throughout on both problems and destroys compliance on the hard filter (312.7 versus 217.1); the controller started at the same $\eta_0 = 1.0$ recovers the well-tuned compliance on both (195.3 and 217.1 at budget). Started over-damped ($\eta_0 = 0.3$), additive increase rides the design's capacity upward and slightly betters the fixed-0.5 compliance (194.6 versus 195.6). Started at the folklore value, it does no harm in compliance—here $\leq 0.1\%$, and at most $+0.7\%$ anywhere in the suite of Section 6.4—with a single stability-tag exception (boundary-riding at the 120×40 mesh, at $+0.0\%$ compliance), discussed there.

6.3 A moving threshold: penalization continuation

Continuation in p is the workflow where fixed damping meets a moving target: as p ramps, the threshold of the naturally saturated family falls as $2/(p{+}1)$—from 1.0 at $p = 1$ (where $\eta = \frac{1}{2}$ is twofold over-damped) through 0.5 exactly at $p = 3$ to 0.333 at $p = 5$ (where it is above threshold unless the design self-limits). We ramp $p : 1 \to 5$ in 0.25 increments (fast: every 20 iterations, budget 600; slow: every 40, budget 900) on three configurations, against the same frozen controller (Fig. 8).

Table 4 summarizes the suite; three outcomes. *(i) The ramp breaks the folklore, replicably.* Fixed $\eta = 0.5$ oscillates through the end of the budget on the fast cantilever ramp at the nominal start and at all three perturbed restarts, and on a fourth, leaner configuration (volume fraction 0.4) it fails on *both* schedules at every start (8/8). That group is worth dissecting, because there the failure is not a race but a destination. Every method that converges there ends on the naturally saturated family, $s_{\max}^{\text{end}} = 5.94$–$6.00 \approx p{+}1$, whose own threshold is $2/s_{\max} = 0.333$—below $\frac{1}{2}$, so the design the ramp is heading for is one that $\eta = \frac{1}{2}$ cannot hold, at any schedule. Fixed $\frac{1}{2}$ instead ends in a persistent oscillation whose (non-fixed-point, hence indicative only) spectrum sits at $s_{\max} = 3.7$–3.9, i.e. $\eta\, s_{\max} \approx 1.9$: marginal on the ceiling-limited side as well, so it neither settles there nor reaches the natural family. Fixed 0.3 survives the same group for the converse reason—$0.3 < 0.333$—and the controller, started from either 0.5 or 1.0, lands at $\eta \approx 0.33$–0.35 on endpoints with $s_{\max} \approx 6.0$: it finds the destination's threshold without being told it. In total fixed 0.5 ends oscillating in 12 of its 32 continuation runs and fixed 0.7 in 28 of 32, while the controller converges in all 64 of its runs; on the fast cantilever the controller started at $\eta_0 = 1.0$ gives the group best endpoint (182). On the slow cantilever schedule fixed 0.5 survives the same ramp: there the failure is a race between ramping and settling, which is precisely what a fixed parameter cannot know about. *(ii) The controller tracks the falling threshold it was never told about.* Started at $\eta_0 = 1.0$, its damping follows $2/(p(t){+}1)$ with a small positive lag that closes at the hold (0.81/0.80, 0.45/0.44, 0.39/0.35, 0.36/0.33 at successive ramp stages, and 0.331/0.333 once p stops moving)—the lag is the additive probe rate, one of the controller's timescale constants, working against a target that is still moving; started at 0.5 it first probes upward—correctly, since the threshold is still above it—until the falling boundary catches it near $p = 3$, then rides it down (Fig. 8a). On the cantilever it instead tracks its own endpoint's threshold ($\eta \approx 0.55$,

Table 4: Penalization continuation (p : $1 \to 5$): compliance at budget (endpoint $s_{\max}$ in parentheses; * oscillating at budget; bold = group best where unique). The natural-family endpoints show $s_{\max} \approx 6 = p{+}1$, the ceiling-limited ones $s_{\max} \approx 4 = 2/\eta$.

config	sched.	fixed 0.3	fixed 0.5	fixed 0.7	ladder	AIMD 0.5	AIMD 1.0
MBB $r1.1$	fast	199 (6.0)	**195 (4.0)**	197 (4.6)*	196 (4.3)	198 (5.3)	198 (6.0)
MBB $r2.4$	fast	**223 (3.2)**	224 (3.3)	248 (3.3)*	224 (2.9)	224 (3.1)	224 (3.3)
cant. $r1.1$	fast	188 (4.2)	188 (4.2)*	188 (3.8)*	183 (3.2)	187 (5.3)	**182 (3.7)**
MBB $r1.1$	slow	198 (5.3)	**196 (4.0)**	198 (3.6)*	199 (4.7)	198 (4.3)	196 (6.0)
MBB $r2.4$	slow	224 (3.0)	224 (3.3)	235 (2.8)*	224 (3.3)	224 (3.3)	224 (3.3)
cant. $r1.1$	slow	188 (5.4)	184 (3.4)	182 (3.3)	182 (3.5)	182 (3.5)	182 (3.7)
MBB vf0.4	fast	**243**	247*	252*	244	245	244
MBB vf0.4	slow	247	249*	253*	247	**246**	**246**
oscillating (6 groups)		0/6	1/6	5/6	0/6	0/6	0/6
median Δcompliance		+1.38%	+0.22%	+2.13%	+0.22%	+0.62%	**+0.02%**
worst Δcompliance		+3.2%	+2.9%	+11.0%	+1.4%	+2.7%	+1.5%
never reached 10^{-5}		1/6	3/6	6/6	4/6	**0/6**	**0/6**

Summary rows cover the six fully instrumented groups. The vf0.4 compliances come from the replication study; that group's endpoint spectra were measured separately (text). "Never reached 10^{-5}" counts groups whose step amplitude never settles below 10^{-5} after the ramp ends—the speed axis. Across three perturbed restarts per group, every oscillation tag replicates at 3/3 seeds, and no controller run fails at any start.

endpoint $s_{\max} = 3.7$): the controller follows the fixed point it is actually converging to, not a formula. *(iii) The two families of Section 5.4 emerge from a realistic workflow.* At $r_{\min} = 1.1$ the fixed-0.5 endpoint lands at $s_{\max} = 4.0$—the capacity ceiling $2/\eta$, on both schedules—while the controller from $\eta_0 = 1.0$ reaches the naturally saturated family at $s_{\max} = 6.0 = p{+}1$ (Fig. 8c). The basin caveat cuts both ways and we report it: on this problem the ceiling-limited design happens to have slightly *better* compliance (195 versus 198, +1.5%), while on the cantilever the controller's endpoint is best; threshold-tracking guarantees stability, not basin quality. On the hard filter ($r_{\min}2.4$) detuning keeps every method's endpoint near $s_{\max} \approx 3.3$ and 224 in compliance—continuation never stresses the folklore there, exactly as the filter results predict.

(iv) Why not simply lower η? Fixed $\eta = 0.3$ never oscillates in this suite, and a reader could stop there. The summary rows of Table 4 price that choice: over-damping converts stability into slowness, and at a fixed budget slowness is design quality. Fixed 0.3 carries the largest median compliance penalty of any non-oscillating method (+1.38%, worst +3.2%), against +0.02% (worst +1.5%) for the controller started at $\eta_0 = 1.0$; its advantage over the controller on the speed axis is likewise absent (both settle below 10^{-5} in five or six of six groups). This is the "more stable but slower" trade-off of Section 5.3 made operational: the fixed choices sit on one side of it or the other, while the controller moves along it.

One further closure comes free: at the $p = 5$ hold the controller's damping settles at a median 0.331 with endpoint $s_{\max} = 6.00$, i.e. $\eta\, s_{\max} = 1.99$ on both schedules—the iteration parks exactly on the boundary that Section 5.3 identifies as optimal and whose inequality the conclusion keeps.

6.4 Tuning-failure rates across the suite

Table 5 is the payload: eleven (configuration, p) groups—mesh, boundary conditions, volume fraction, filter type and radius, $p \in \{3, 5\}$—times nine methods, at a 400-iteration budget, one frozen constant set. Failure is reported two ways: still oscillating at budget (median tail amplitude $> 3 \times 10^{-3}$), or converged but worse than the group best by more than 2%.

Table 5: Tuning-failure rates over the 11-group suite (400 iterations; AIMD constants frozen; MMA is our single-constraint implementation with standard parameters, reported for scope reference).

method	osc. (ampl.)	masked ($\hat{\mu}\approx-1$)	quality ($> 2\%$)	median Δ	worst Δ
fixed $\eta = 0.5$	0/11	0/11	0/11	+0.36%	+0.8%
fixed $\eta = 0.7$	4/11	3/11	1/11	+0.81%	+17.1%
fixed $\eta = 1.0$	9/11	0/11	0/11	+0.74%	+44.5%
AIMD from 0.5	1/11	0/11	0/11	+0.36%	**+0.7%**
AIMD from 1.0	2/11	0/11	1/11	+0.17%	+18.4%
ladder from 1.0	1/11	2/11	2/11	+0.01%	+19.0%
MMA (reference)	0/11	—	1/11	+0.21%	+3.9%
adaptive move, $\eta = 0.5$	0/11	2/11	1/11	+0.41%	+2.3%
adaptive move, $\eta = 1.0$	0/11	8/11	1/11	+0.27%	+2.8%

Masked oscillation: quiet by the amplitude criterion (tail in $(10^{-5}, 3 \times 10^{-3})$) but flip-locked ($\hat{\mu}_{\text{tail}} \leq -0.9$; the move-$\eta$=1.0 cases sit at exactly -1.00 with the move limit at its floor). The three tags are exclusive; quality counts only genuinely converged runs; median and worst Δcompliance are over all 11 groups relative to the OC-family best. MMA's $\hat{\mu}$ is not instrumented (—). Across three perturbed restarts every amplitude count varies by at most ±1, except fixed 0.7 (3–6/11); see text.

Read honestly, the first row is a result in itself: fixed $\eta = \frac{1}{2}$ never failed at the nominal start—exactly what its Richardson optimality (Section 5.3) predicts, and the reason the folklore survived thirty-five years (though the perturbation study below shows even this row is not bulletproof). The tuning problem appears the moment one deviates (0.7: 4/11 oscillating; 1.0: 9/11) or composes with acceleration (Section 6.5). The controller's value is not beating $\frac{1}{2}$ at $p = 3$; it is making the choice irrelevant: from a hostile start it recovers the hand-tuned compliance to tenths of a percent, and it removes the one parameter a user could set wrongly. For a paper about worst cases the worst-case column is the one to read: started at the folklore value, the controller's worst compliance penalty across the entire suite is +0.7% at the nominal start, and +1.37%, +1.65%, +1.40% at the three perturbed starts—at each individual start no worse than fixed $\frac{1}{2}$ run from the same start (+0.8% nominal; +1.38%, +1.89%, +1.40% perturbed).

Repeating the entire suite from three perturbed starts (2% log-normal initial perturbation, seeds 5/7/11) leaves every count of Table 5 within ±1, with one exception and one lesson. The exception: fixed 0.7 fluctuates between 3/11 and 6/11 oscillating. The lesson: fixed 0.5 itself oscillates once at one seed—its clean row is not bulletproof—while the controller from 0.5 scores 0/11 at two seeds and 1/11 at the third. We claim no significance for 0-versus-1 distinctions; the seed-robust distinctions are fixed 1.0 (9/11 at every seed, worst case degrading to +61% under perturbation), fixed 0.7's instability, and the worst-case gap between the controller from the folklore start (+0.7–1.6% across seeds) and every deviating fixed choice (+15–61%). Its residual failures are diagnostic and we report them: at the 120×40 mesh the design is still structurally evolving within this budget, so additive increase keeps probing and the run rides the boundary in a small-amplitude sawtooth (tagged oscillating; compliance within 0.3% of best); the same boundary-riding tag appears for the over-driven start at $p = 5$ on the canonical mesh; and one genuinely treacherous group (volume fraction 0.4, $p = 5$)—where the practitioner ladder and fixed 0.7 also fail at +17–19%—defeats the controller from the over-driven start (from the folklore start it passes, +0.0%). More broadly, which fixed point the descent enters remains path-dependent: *the controller controls stability, not basin selection* (Section 7). The MMA reference column converges everywhere with compliance within a few percent either way of the OC family (better on two groups, 3.9% worse on one)—consistent with our scope claim that the controller addresses the OC update's one

Table 6: Anderson composition (60×20 MBB, sens. $r_{\min}2.4$, $p = 3$, 400 iterations). t_{conv}: first iteration with step amplitude below 10^{-3} sustained for 10.

method	compliance	tail amplitude	t_{conv}
fixed 0.5	217.10	2.7×10^{-5}	129
fixed 1.0	312.74	1.2×10^{-1}	—
AA + fixed 0.5	253.43	3.7×10^{-3}	—
AA + fixed 1.0	217.11	3.3×10^{-11}	116
damped AA + fixed 0.5	219.69	1.9×10^{-3}	—
AIMD(1.0)	217.10	9.0×10^{-5}	257
AA + AIMD(0.5)	217.11	2.0×10^{-6}	299
AA + AIMD(1.0)	217.11	4.2×10^{-11}	168

parameter rather than competing across method families.

The adaptive move-limit rows test the obvious alternative—stabilize by scheduling the move limit instead of the damping (an MMA-style rule: if the fraction of interior elements whose consecutive changes flip sign exceeds 0.35, multiply the move limit by 0.7, else by 1.2, bounds $[0.01, 0.2]$; η held fixed). The theory of Section 3 predicts the outcome: move limits bound amplitude but leave the multiplier untouched. That is exactly what happens. At $\eta = 1.0$, eight of eleven runs pass the amplitude criterion while their tail multiplier sits at $\hat{\mu} = -1.00$ with the move limit ground down to its floor: the oscillation has not been removed, it has been hidden below the tolerance, and the iteration can never converge—it grinds at the move floor indefinitely. Amplitude-based stopping criteria (including the default of the educational codes) are fooled by this state; the multiplier sensor is not. The move-limit rule is thus not a competitor but a cautionary exhibit: it also supplies the mechanism confirmation that the flip threshold belongs to η, not to the step size. And the masked column reveals that the hidden state is not peculiar to move scheduling: three of fixed 0.7's "converged" runs are quiet limit cycles at the standard move limit ($\hat{\mu} = -1.00$ at amplitudes 2×10^{-5}–9×10^{-4}—soft, supercritical onsets hiding below the tolerance), as are two of the ladder's. Counting both oscillation modes, fixed 0.7 fails 7 of 11 groups (8 including its quality failure). Masked oscillation is simply what near-threshold operation looks like under any amplitude criterion, and only the multiplier exposes it. Read down the masked column, the controller is the only method besides fixed $\frac{1}{2}$ with no hidden limit cycles from either start (0/11 both): it does not trade visible oscillation for invisible oscillation. That fixed $\frac{1}{2}$ also scores 0/11 while operating 0.6% from its threshold at R0 is explained by Section 5.4—self-limiting keeps the reachable fixed points at $s_{\max} \leq 2/\eta$, so the standard damping ends up inside its own boundary rather than on it. The third law predicts the failure-rate table.

6.5 Composition with Anderson acceleration

Anderson acceleration attacks OC oscillation by extrapolating through the cycle [Li et al., 2020]; damping control attacks it by stepping around it. Table 6 composes them on the hard-filter problem (periodic type-II AA, window 4, every fourth step, with box and volume projection; our variant, undamped unless noted).

The pattern is symmetric and, to our knowledge, unreported: AA rescues the over-driven run completely (extrapolation annihilates the period-2) but *breaks the well-tuned one*—at $\eta = 0.5$ its extrapolation interferes with a healthy contraction and lands 17% worse, and damping the mixing ($\beta = 0.5$) only mitigates. Fixed damping fails at 1.0; AA fails at 0.5; the controller alone fails at neither; and the composition converges deepest of all from both starts (tail amplitudes 10^{-6}–10^{-11}), with the controller intervening only a handful of times to break AA's interference and the acceleration doing the rest. Competition becomes complementarity:

AA supplies depth, the controller supplies safety, and neither requires the user to choose η.

7 Discussion

A landscape of fixed points. The same problem hosts multiple OC fixed points with widely different spectra ($s_{\max} = 5.97$, 4.33, 3.52 all observed at $p = 5$, $r_{\min} = 1.1$), and the iteration selects among them by its own stability. Reporting the convergence protocol is part of reporting a spectrum; our own retracted $p = 5$ preview is the cautionary tale. The controller inherits this: it stabilizes whatever basin the trajectory enters, and basin selection remains open. The cleanest exhibit is the canonical mesh at $p = 5$, where the converged methods of Table 5 land on fixed points spanning 2.7% in compliance (199.1–204.4) with no method reliably finding the best one.

Statics meets dynamics. The Groenwold–Etman dictionary makes η a curvature choice; the multiplier law makes it a spectral gain; the Richardson result closes the loop by making the folklore value optimal rather than lucky. Exact curvature $\Leftrightarrow$ deadbeat; half curvature $\Leftrightarrow$ flip boundary $\Leftrightarrow$ min–max optimum at $p = 3$.

Why $p + 1$? The saturation $s_{\max} \to p{+}1$ says the dangerous mode behaves like a statically determinate substructure; its measured localization (participation ratio ≈ 11 of 360 free elements at the canonical benchmark, ≈ 20 at $p = 5$), its suppression by spatial averaging, and its persistence across $p \in [1.5, 5]$ all fit, but we have no proof. Along with the critical behavior near onset (soft transitions, capacity adaptation), this is deferred to a separate, physics-oriented study.

8 Limitations

Local theory only; 2D compliance, single load case, single kick seed; onset brackets of width 0.08; the V5 density-filter bracket is semi-quantitative. The controller was tuned once (its three design iterations are documented in Section 6.1), then frozen; its residual failures are reported in Section 6.4, including one hard group at volume fraction 0.4, $p = 5$, and boundary-riding at 120×40 within the 400-iteration budget—the latter a timescale interaction (probe rate versus structural evolution within the budget), consistent with the fact that the controller's only non-dimensionless constants are timescales. The continuation evidence spans four configurations, two schedules, and up to four starts per group (fixed 0.5 failing 12 of 32 runs, the controller none of 64); 120×40 continuation was not run. The Anderson variant is ours (periodic, undamped type-II with projections); other AA flavors may not break at $\eta = 0.5$. The MMA column uses our compact single-constraint implementation with standard parameters, validated only by comparable compliance on the canonical problem. Compliance differences among converged methods are tenths of a percent and are not the claim; the claim is convergence without tuning. Basin selection is uncontrolled, and the controller addresses the flip branch only—the negative branch (Appendix A) is outside its authority, benign in all our measurements. Amplitude-based convergence criteria are fooled by masked oscillation—move-limited or near-threshold (Section 6.4); we recommend monitoring $\hat{\mu}$ alongside any amplitude tolerance. The $p{+}1$ saturation and filter-detuning mechanisms are measured regularities awaiting analysis.

9 Conclusion

The damping exponent of the OC iteration has been folklore for thirty-five years because it never had a measurable target. In logarithmic coordinates the target is one constant: keep $\eta\, s_{\max} < 2$, where $s_{\max}$ is the top of a spectrum the iteration itself can sense. The folklore $\eta = \frac{1}{2}$ turns out to sit at the min–max optimum for the spectra this problem class actually produces—at the edge of instability because a spectrum whose lower edge is zero puts the optimum there—and the three folk remedies for oscillation are one remedy. That optimality is a statement about one fixed point, though, and the threshold belongs to whichever fixed point the iteration is heading for: under continuation it moves, and no fixed exponent follows it. A four-line AIMD rule with all thresholds in multiplier space converges from starts that destroy fixed damping, lands at each problem's own threshold, tracks that threshold when penalization continuation moves it, composes with Anderson acceleration so that each covers the other's failures, and leaves the user nothing to tune.

Data and code availability. All code, checkpoints and figure scripts are openly available at `https://github.com/gitUserKHS/oc-flip`. Every table and figure in this paper is regenerated by the driver scripts listed in the repository README; the only dependencies are numpy, scipy and matplotlib. The conventions fixed across all experiments—finite-difference step 10^{-5}, OC bisection tolerance 10^{-10}, kick seed 5, onset grid width 0.08, and the frozen controller constants—are recorded there, as is the convergence protocol behind every reported spectrum, together with the full research record: the notes document a retracted preview result, a withdrawn mechanism attribution and a causal claim corrected by measurement. An archived snapshot with a persistent identifier will be referenced in a subsequent version of this preprint.

AI assistance. The experiments, figures and manuscript drafting were carried out with the assistance of a large language model under the author's direction. The research questions, the experimental design, the interpretation of the results, and the final text are the author's own, as is responsibility for their content; the commit history of the repository above records where that assistance was used.

A The negative branch

The parallel model's instability (Proposition 2) appears in real spectra as $s_{\min} < 0$. On the cantilever (V3, $s_{\min} = -0.04308$, exactly projected), kicking the converged design along the eigenvector and iterating: the mode grows by the predicted factor $1 - \eta s_{\min} = 1.0215$ per step to four decimals ($n = 76$ ratio samples); its support is organized in mirror pairs across the symmetry line (77% of $|v|^2$ below $x = 0.05$; the twenty largest components pair exactly), so the mode is an antisymmetric low-density competition— winner-take-all in symmetry coordinates. Termination is by resettling, not absorption: winners rise ($0.009 \to 0.032$), losers sink toward the floor, four elements leave the free set, and the iteration lands in a nearby *asymmetric* fixed point whose spectrum has shed the negative eigenvalue ($s_{\min} = +0.0375$), at compliance unchanged to 0.02% (Fig. 10). The symmetric design was a saddle; the escape is benign, registering in standard convergence checks only as a late small drift. Consistently, $s_{\min} < 0$ appears where low-density competition is unresolved and disappears as penalization removes gray coexistence ($s_{\min} = +0.119$ at $p = 5$, reference protocol).

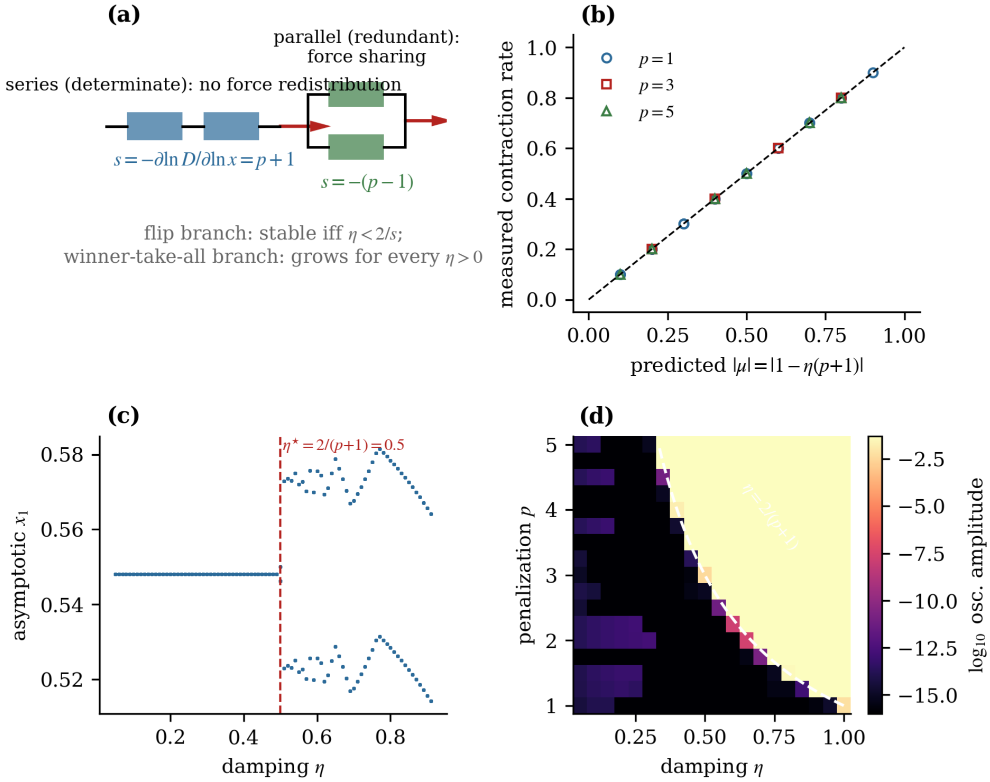

Figure 1: The two exact models (Section 3). (a) schematic: the series (statically determinate) model carries the flip branch, $s = p{+}1$; the parallel (redundant) model the winner-take-all branch, $s = -(p{-}1)$. (b) measured contraction rates versus the prediction $|1 - \eta(p{+}1)|$ for $p \in \{1, 3, 5\}$. (c) bifurcation diagram at $p = 3$ (move limit 0.05), splitting at $\eta = 0.500$. (d) oscillation amplitude in the (η, p) plane against the curve $\eta = 2/(p{+}1)$.

References

Abaqus. Stress ratio technique. Abaqus/CAE Documentation, Topology and Shape Optimization, 2017. States that smaller exponents give more stable but slower convergence—the folklore this paper quantifies.

S. Ananiev. On equivalence between optimality criteria and projected gradient methods with application to topology optimization problem. *Multibody System Dynamics*, 13(1):25–38, 2005. doi: 10.1007/s11044-005-2530-y.

E. Andreassen, A. Clausen, M. Schevenels, B. S. Lazarov, and O. Sigmund. Efficient topology optimization in MATLAB using 88 lines of code. *Structural and Multidisciplinary Optimization*, 43(1):1–16, 2011. doi: 10.1007/s00158-010-0594-7.

J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988. doi: 10.1093/imanum/8.1.141.

M. P. Bendsøe. Optimal shape design as a material distribution problem. *Structural Optimization*, 1(4):193–202, 1989. doi: 10.1007/BF01650949.

B. Bourdin. Filters in topology optimization. *International Journal for Numerical Methods in Engineering*, 50(9):2143–2158, 2001. doi: 10.1002/nme.116.

A. A. Groenwold and L. F. P. Etman. On the equivalence of optimality criterion and sequential approximate optimization methods in the classical topology layout problem. *International Journal for Numerical Methods in Engineering*, 73(3):297–316, 2008. doi: 10.1002/nme.2071.

A. A. Groenwold and L. F. P. Etman. A simple heuristic for gray-scale suppression in optimality criterion-based topology optimization. *Structural and Multidisciplinary Optimization*, 39(2):217–225, 2009. doi: 10.1007/s00158-008-0337-1.

A. A. Groenwold, L. F. P. Etman, J. A. Snyman, and J. E. Rooda. Incomplete series expansion for function approximation. *Structural and Multidisciplinary Optimization*, 34(1):21–40, 2007. doi: 10.1007/s00158-006-0070-6.

B. M. Irons and R. C. Tuck. A version of the Aitken accelerator for computer iteration. *International Journal for Numerical Methods in Engineering*, 1(3):275–277, 1969. doi: 10.1002/nme.1620010306.

B. Keith, D. Kim, B. S. Lazarov, and T. M. Surowiec. Analysis of the SiMPL method for density-based topology optimization. *SIAM Journal on Optimization*, 35(2):1134–1164, 2025. doi: 10.1137/24M1708863.

D. Kim, B. S. Lazarov, T. M. Surowiec, and B. Keith. A simple introduction to the SiMPL method for density-based topology optimization. *Structural and Multidisciplinary Optimization*, 68, 2025. doi: 10.1007/s00158-025-04008-9. Also arXiv:2411.19421.

U. Küttler and W. A. Wall. Fixed-point fluid–structure interaction solvers with dynamic relaxation. *Computational Mechanics*, 43(1):61–72, 2008. doi: 10.1007/s00466-008-0255-5.

W. Li, P. Suryanarayana, and G. H. Paulino. Accelerated fixed-point formulation of topology optimization: Application to compliance minimization problems. *Mechanics Research Communications*, 103:103469, 2020. doi: 10.1016/j.mechrescom.2019.103469.

M. Liu and S. A. Burns. Multiple fully stressed designs of steel frame structures with semi-rigid connections. *International Journal for Numerical Methods in Engineering*, 58(6):821–838, 2003. doi: 10.1002/nme.807.

J. M. Martínez. A note on the theoretical convergence properties of the SIMP method. *Structural and Multidisciplinary Optimization*, 29(4):319–323, 2005. doi: 10.1007/s00158-004-0479-8.

K. M. Mueller and S. A. Burns. Fully stressed frame structures unobtainable by conventional design methodology. *International Journal for Numerical Methods in Engineering*, 52(12):1397–1409, 2001. doi: 10.1002/nme.261.

K. M. Mueller, M. Liu, and S. A. Burns. Fully stressed design of frame structures and multiple load paths. *Journal of Structural Engineering*, 128(6):806–814, 2002. doi: 10.1061/(ASCE)0733-9445(2002)128:6(806).

S. N. Patnaik, A. S. Gendy, L. Berke, and D. A. Hopkins. Modified fully utilized design (MFUD) method for stress and displacement constraints. *International Journal for Numerical Methods in Engineering*, 41(7):1171–1194, 1998.

R. Razani. Behavior of fully stressed design of structures and its relationship to minimum-weight design. *AIAA Journal*, 3(12):2262–2268, 1965. doi: 10.2514/3.3355.

A. Rietz. Sufficiency of a finite exponent in SIMP (power law) methods. *Structural and Multidisciplinary Optimization*, 21(2):159–163, 2001. doi: 10.1007/s001580050180.

O. Sigmund. A 99 line topology optimization code written in Matlab. *Structural and Multidisciplinary Optimization*, 21(2):120–127, 2001. doi: 10.1007/s001580050176.

O. Sigmund and J. Petersson. Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Structural Optimization*, 16(1):68–75, 1998. doi: 10.1007/BF01214002.

M. Stolpe and K. Svanberg. An alternative interpolation scheme for minimum compliance topology optimization. *Structural and Multidisciplinary Optimization*, 22(2):116–124, 2001. doi: 10.1007/s001580100129.

K. Svanberg. The method of moving asymptotes—a new method for structural optimization. *International Journal for Numerical Methods in Engineering*, 24(2):359–373, 1987. doi: 10.1002/nme.1620240207.

K. Svanberg. A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM Journal on Optimization*, 12(2):555–573, 2002. doi: 10.1137/S1052623499362822.

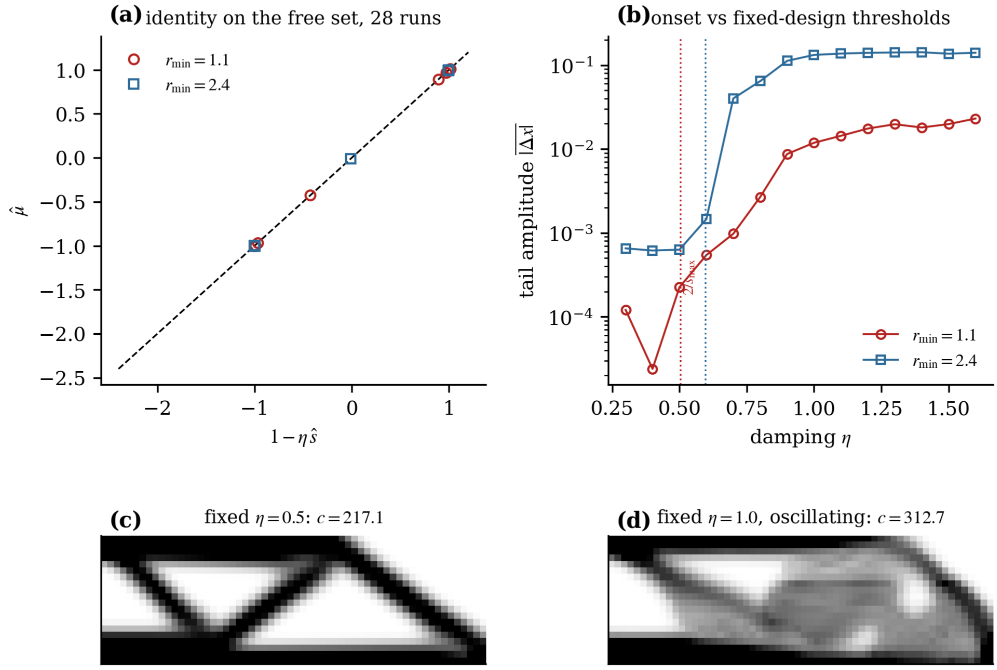

Figure 2: Trajectory estimators in the wild (Section 5.1). (a) the identity $\hat{\mu} = 1 - \eta\hat{s}$ across the 28-run sweep. (b) tail amplitude versus η at both filter radii; dotted lines mark the fixed-design thresholds $2/s_{\max}$ (0.503, 0.596). (c,d) converged versus persistently oscillating designs at $r_{\min} = 2.4$ (compliance 217.1 versus 312.7 at $\eta = 1.0$, 400 iterations): oscillation is not slow convergence but lock-in to a damaged design.

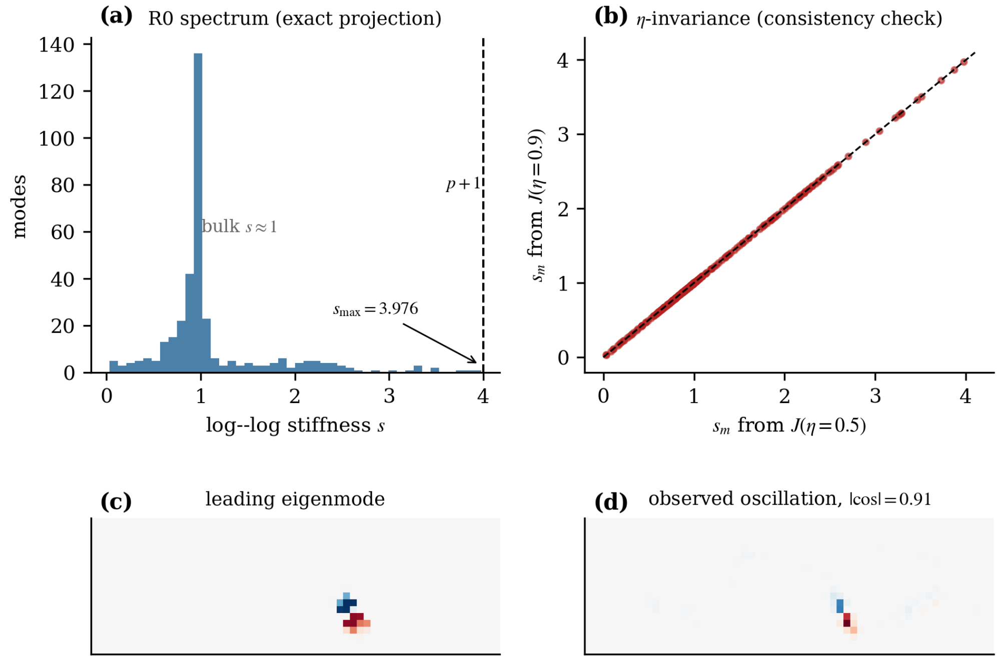

Figure 3: Operator-level spectra at R0 (Section 5.1). (a) the exactly projected s-spectrum: bulk at $s \approx 1$, lower edge near zero, tail saturating at $s_{\max} = 3.976$ just below $p+1$—the shape behind both the Richardson result and the creep phenomenology. (b) η-invariance of the spectrum (top-five difference 9×10^{-6} between $J(\eta{=}0.5)$ and $J(\eta{=}0.9)$). (c) leading eigenmode; (d) oscillation pattern measured just above threshold (independent reproduction, $|\cos| = 0.91$; the original measurements give 0.93 and 0.96).

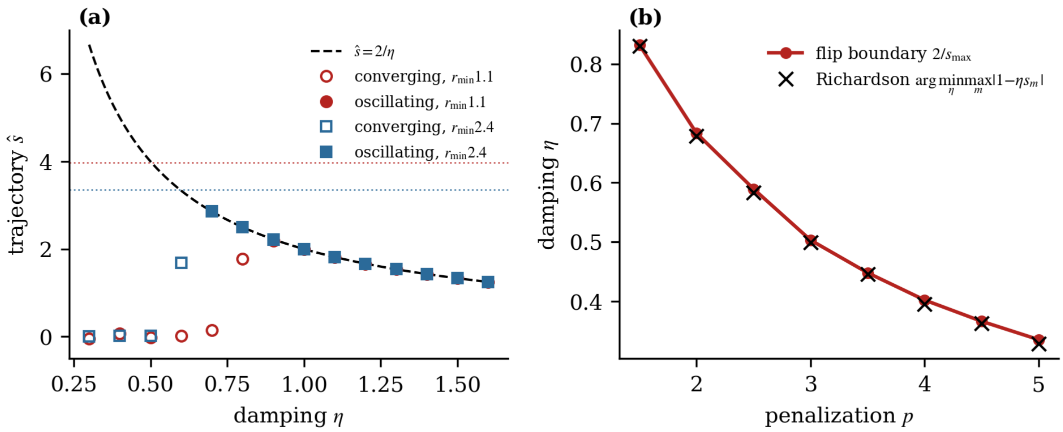

Figure 4: (a) The two regimes of the trajectory sensor (Section 5.2): converging runs read the $s \approx 0$ creep edge; saturated runs lock onto $\hat{s} = 2/\eta$; dotted lines mark operator $s_{\max}$. (b) The min–max optimal damping coincides with the flip boundary across the penalization sweep (Section 5.3).

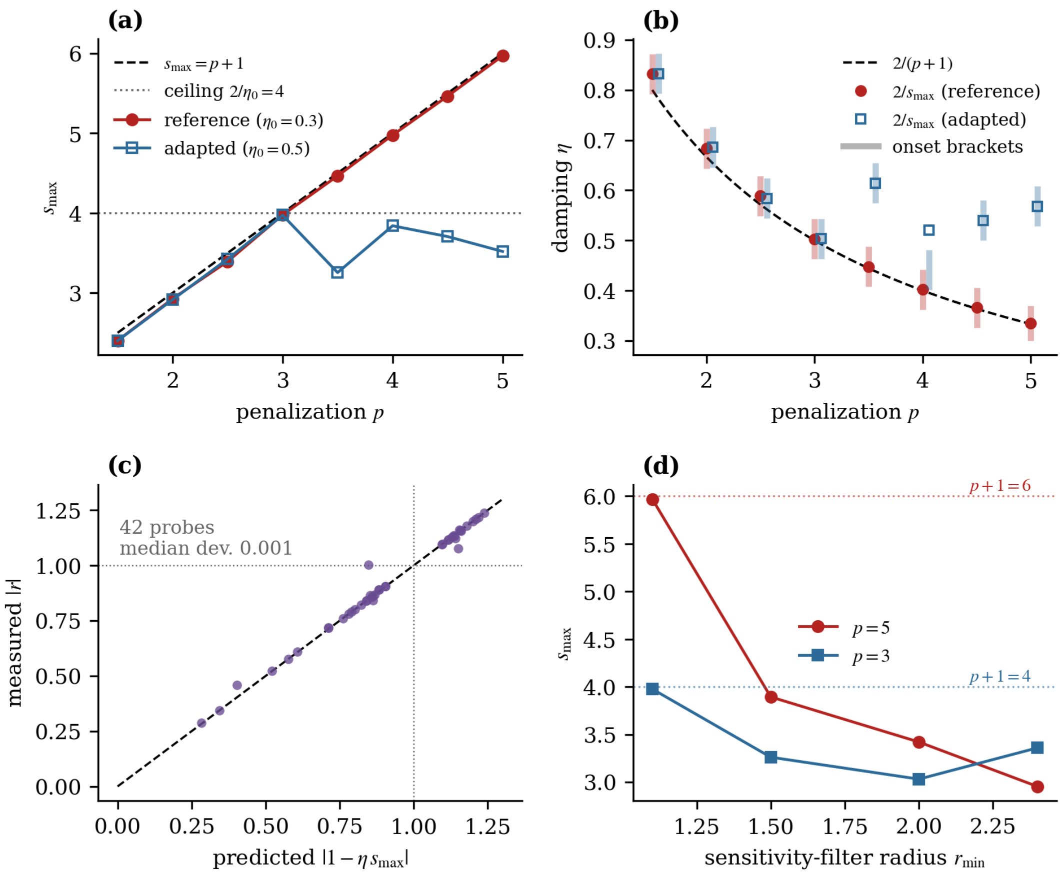

Figure 5: Penalization sweep (Section 5.4): saturation under the reference protocol versus self-limiting under standard damping, separating at $p = 3$; onset brackets following each fixed point's own $2/s_{\max}$; the rate law across both protocols; filter-radius detuning.

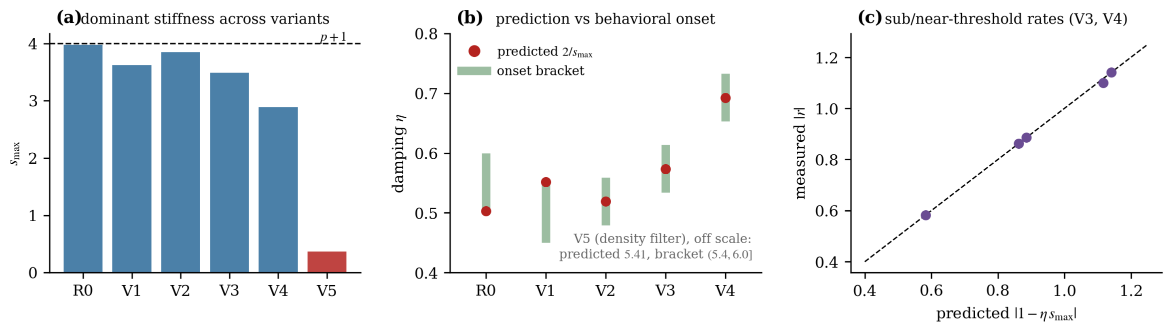

Figure 6: Robustness of the threshold law (Section 5.1, Table 1). (a) $s_{\max}$ across variants; the density filter (V5) collapses it to 0.37. (b) predicted $2/s_{\max}$ versus behavioral onset brackets for R0–V4 (V5 off scale: predicted 5.41, bracket $(5.4, 6.0]$). (c) flip-active probe rates versus $|1 - \eta\, s_{\max}|$ on V3/V4—three-decimal agreement below threshold.

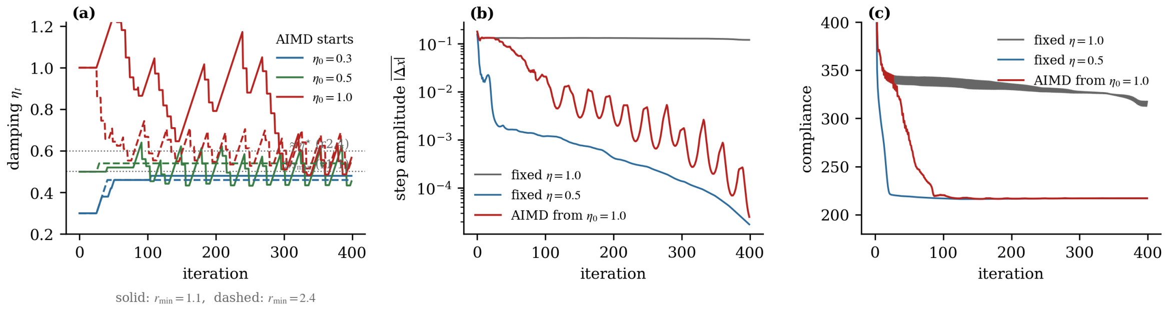

Figure 7: The controller on two problems from three starts (Section 6.2): damping trajectories landing in each problem's own threshold band; rescue of the over-driven hard-filter run; compliance recovery.

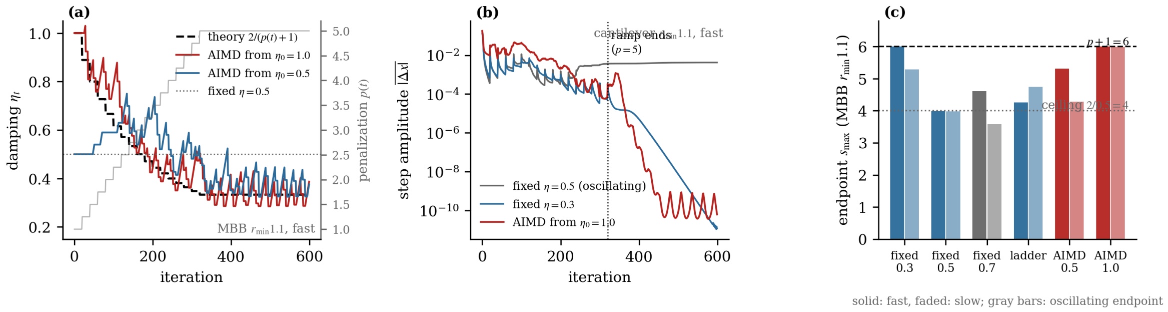

Figure 8: Penalization continuation ($p : 1 \to 5$; Section 6.3). (a) the controller's damping tracks the falling threshold $2/(p(t)+1)$ it was never told about (MBB $r_{\min}$1.1, fast schedule; right axis: $p(t)$). (b) the fast ramp breaks fixed $\eta = 0.5$ on the cantilever while the controller converges to 10^{-10}. (c) endpoint $s_{\max}$: fixed 0.5 lands on the capacity ceiling $2/\eta = 4$, the controller from $\eta_0 = 1.0$ reaches the naturally saturated family $p+1 = 6$, on both schedules.

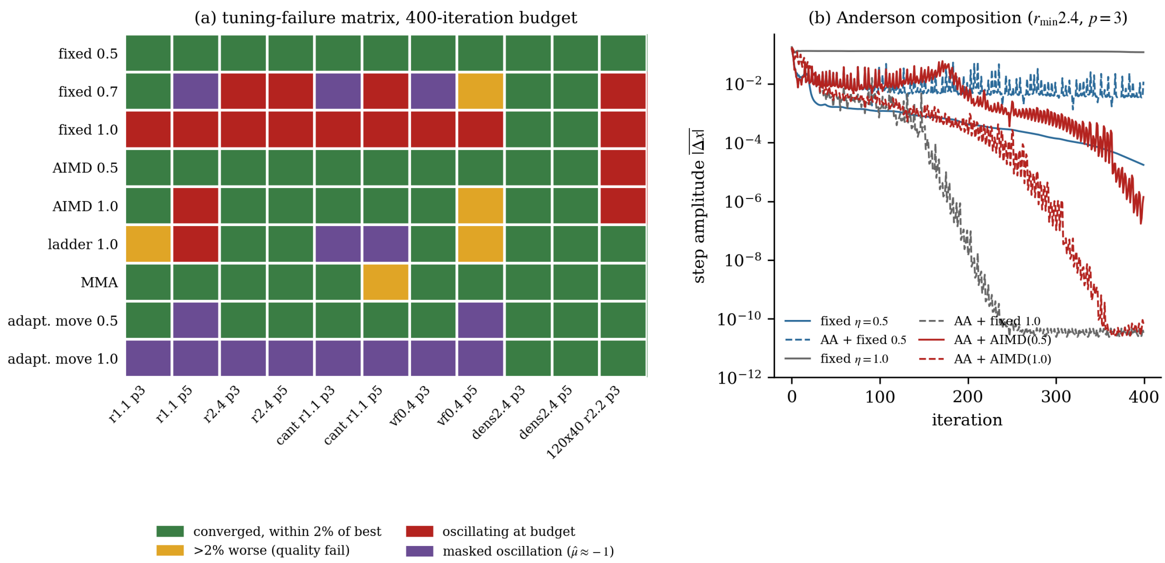

Figure 9: (a) Tuning-failure matrix across the 11-group suite (Section 6.4). (b) Anderson composition (Section 6.5): AA breaks the well-tuned run and fixes the over-driven one; the composition with the controller converges deepest from both starts.

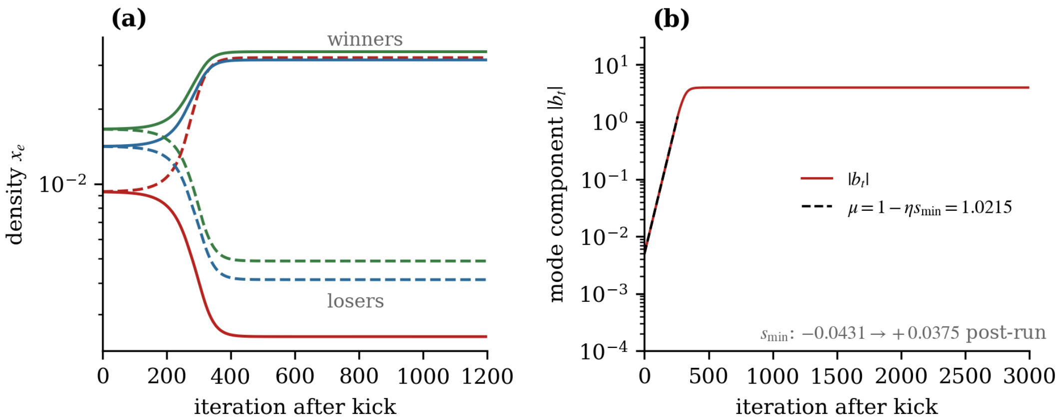

Figure 10: Appendix A: mirror-pair symmetry breaking of the negative branch and growth at the predicted rate, terminating in an asymmetric fixed point.