Neural–Spectral and Source-Frozen Picard Control Reconstruction for Semilinear Wave Equations

Tan-Phuc Nguyen Minh-Binh Tran Son Tu

Presented by: **Tan-Phuc Nguyen**

Texas A&M University

August 16, 2026

Workshop on Pure and Applied Analysis 2026

Controllability theory: mature; computing controls: still hard.

Established Theory

- Exact / null / approximate controllability: decades of results for wave, heat, dispersive equations.
- Linear wave equation: the Geometric Control Condition (GCC) is the sharp controllability and observability criterion (every generalized optical ray must meet ω before time T).

Controllability theory: mature; computing controls: still hard.

Established Theory

- Exact / null / approximate controllability: decades of results for wave, heat, dispersive equations.
- Linear wave equation: the Geometric Control Condition (GCC) is the sharp controllability and observability criterion (every generalized optical ray must meet ω before time T).

Computational gap

- Existence proofs are *not* algorithms.
- Nonlinearity brings loss of compactness, nonconvexity, stiffness, instability.
- Difficulty grows with dimension and with the number of modes needed to resolve the wave–actuator interaction.

What should we prove about computing controls?

The algorithm *Will it converge?*	**The computed control** *Is it accurate?*
Is the iteration provably *convergent*? Is the constant small enough for the criterion to actually be *met* on the benchmark we run?	Does *this* candidate certify an exact control nearby?
Thm 9, 11; Prop 41; validated $\chi_{8,2.5} < 0.161$	Thm 13, 14 (solver-independent); Thm 28

The control problem

Let $\Omega = (0,1)^N$, $\omega \subset \Omega$ open and nonempty, $T > 0$:

$$\begin{cases} \partial_{tt} u - \Delta u + g(u) + \alpha u + f\,\mathbf{1}_\omega = 0, \\ u|_{\partial\Omega} = 0, \\ u(0,\cdot) = u^0, \quad \partial_t u(0,\cdot) = u^1. \end{cases}$$

Goal. Construct an *interior* control f, supported in ω, steering the state to a prescribed (u_T^0, u_T^1) at time T.

The control problem

Let $\Omega = (0,1)^N$, $\omega \subset \Omega$ open and nonempty, $T > 0$:

$$\begin{cases} \partial_{tt} u - \Delta u + g(u) + \alpha u + f\,\mathbf{1}_\omega = 0, \\ u|_{\partial\Omega} = 0, \\ u(0,\cdot) = u^0, \quad \partial_t u(0,\cdot) = u^1. \end{cases}$$

Goal. Construct an *interior* control f, supported in ω, steering the state to a prescribed (u_T^0, u_T^1) at time T.

First-order form. With $\Phi = (u, \partial_t u)$, $G(\Phi) = (0, g(u))$ and $\mathbf{F} = (0, f)$,

$$\partial_t \Phi + A\Phi + \mathbf{F}\mathbf{1}_\omega + G(\Phi) = 0, \qquad \Phi(0) = \Phi_0, \quad \Phi(T) = \Phi_T,$$

on the energy space $\mathcal{E} := H_0^1(\Omega) \times L^2(\Omega)$, where A is skew-adjoint, so $S(t) = e^{-tA}$ is unitary.

Two standing hypotheses and one error source

(H1) Nonlinearity. g induces a globally Lipschitz Nemytskii map

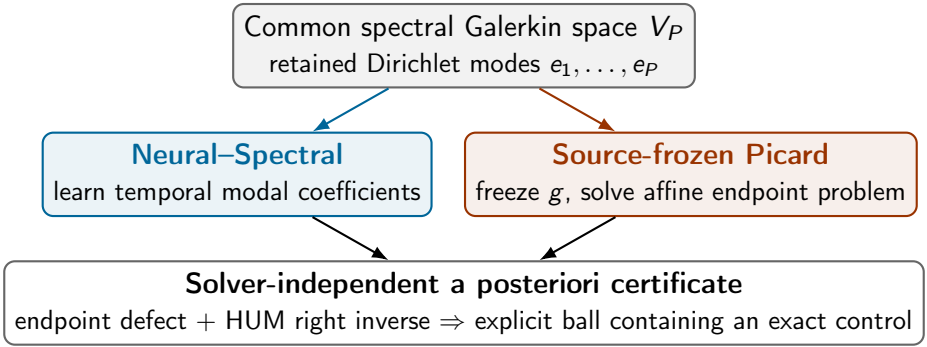
$$g : H_0^1(\Omega) \to L^2(\Omega), \qquad \|g(v) - g(w)\|_{L^2(\Omega)} \le L_g \|v - w\|_{H_0^1(\Omega)}.$$

$g(u) = \beta \tanh u$ is inside; $g(u) = \beta u^3$ is outside (kept as a stress test).

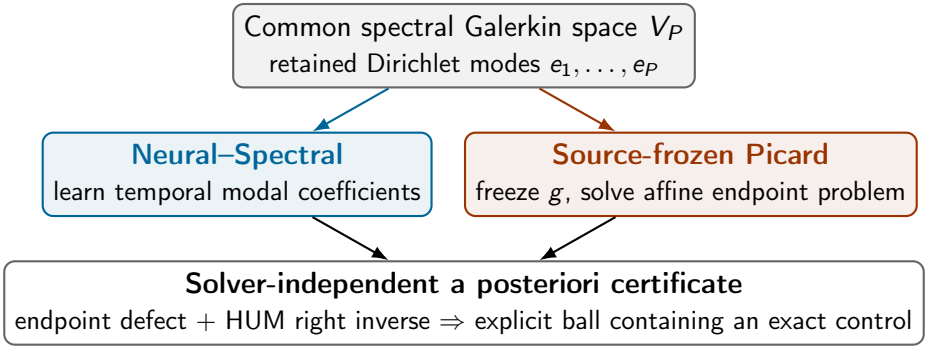
(H2) Geometry. The linear part needs the GCC on (ω, T). Every benchmark picks a collar and a horizon that satisfy it (e.g., $\delta = 0.3$, $T = 2.5$).

(E) Truncation (neither assumable nor avoidable). Everything is computed on V_P. But the control is localized by multiplication with a *discontinuous* indicator, which generates spatial frequencies above P. They are invisible to the retained system, but physically present.

Two paradigms on one common Galerkin space



Two paradigms on one common Galerkin space



- **Neural–Spectral**: the spatial eigenstructure is built in *exactly*; only the temporal coefficients $a_j(t), c_j(t)$ are learned. No fixed-point iteration.
- **Source-frozen Picard**: two *equivalent* realizations: HUM, and a reusable controllability Gramian.

The common spectral Galerkin setting

Spectrum of A (Lemma 4). For $k \in (\mathbb{N}^*)^N$, $\lambda_k^{\pm} = \pm \mathbf{i}\sqrt{\mu_k + \alpha}$, $\mu_k = \sum_{j=1}^{N} k_j^2 \pi^2$, and the corresponding eigenfunctions form an orthonormal basis of $\mathcal{E}$.

Retained space (Lemma 5). After a 1D re-indexing, V_P corresponds to the first P scalar Dirichlet modes $\{e_j\}_{j=1}^{P}$.

The common spectral Galerkin setting

Spectrum of A (Lemma 4). For $k \in (\mathbb{N}^*)^N$, $\lambda_k^{\pm} = \pm \mathbf{i}\sqrt{\mu_k + \alpha}$, $\mu_k = \sum_{j=1}^{N} k_j^2 \pi^2$, and the corresponding eigenfunctions form an orthonormal basis of $\mathcal{E}$.

Retained space (Lemma 5). After a 1D re-indexing, V_P corresponds to the first P scalar Dirichlet modes $\{e_j\}_{j=1}^{P}$.

Projected dynamics: where the control enters. Testing against e_j:

$$a_j'' + (\mu_j + \alpha) a_j + \underbrace{\langle g(\textstyle\sum_\ell a_\ell e_\ell), e_j \rangle}_{\text{nonlinear source}} + \sum_{\ell=1}^{P} \underbrace{M_{j\ell}}_{\int_\omega e_j e_\ell} c_\ell = 0.$$

The common spectral Galerkin setting

Spectrum of A (Lemma 4). For $k \in (\mathbb{N}^*)^N$, $\lambda_k^{\pm} = \pm \mathbf{i}\sqrt{\mu_k + \alpha}$, $\mu_k = \sum_{j=1}^{N} k_j^2 \pi^2$, and the corresponding eigenfunctions form an orthonormal basis of $\mathcal{E}$.

Retained space (Lemma 5). After a 1D re-indexing, V_P corresponds to the first P scalar Dirichlet modes $\{e_j\}_{j=1}^{P}$.

Projected dynamics: where the control enters. Testing against e_j:

$$a_j'' + (\mu_j + \alpha) a_j + \underbrace{\left\langle g\left(\sum_\ell a_\ell e_\ell\right), e_j \right\rangle}_{\text{nonlinear source}} + \sum_{\ell=1}^{P} \underbrace{M_{j\ell}}_{\int_\omega e_j e_\ell} c_\ell = 0.$$

M is the actuator-overlap matrix. The lifted control is localized only *after* reconstruction, so multiplying by $\mathbf{1}_\omega$ couples the modes through M. Being symmetric positive definite, M makes (A_P, B_P) controllable (Lemma 31), which is what makes every affine subproblem solvable, and M reappears in both branches.

From GCC to a bounded "source $\mapsto$ trajectory" operator

Set $U = L^2(\omega)$, $Bv = (0, \tilde{v})$, and define the linear endpoint operator

$$\mathcal{C}_T v := -\int_0^T S(T-s)\, Bv(s)\, ds, \qquad \mathcal{C}_T : L^2(0,T;U) \to \mathcal{E}.$$

From GCC to a bounded "source $\mapsto$ trajectory" operator

Set $U = L^2(\omega)$, $Bv = (0, \tilde{v})$, and define the linear endpoint operator

$$\mathcal{C}_T v := -\int_0^T S(T-s)\, Bv(s)\, ds, \qquad \mathcal{C}_T : L^2(0,T;U) \to \mathcal{E}.$$

Under the GCC, observability $\Rightarrow \mathcal{C}_T$ is onto $\Rightarrow$ the Gramian $\mathcal{W}_T = \mathcal{C}_T \mathcal{C}_T^*$ is boundedly invertible. Put

$$\Gamma_T := \|\mathcal{W}_T^{-1}\|^{1/2}, \qquad C_T^{\mathrm{aff}} := \sqrt{T}\big(1 + \sqrt{T}\,\|B\|\,\Gamma_T\big).$$

From a frozen source to HUM, and the estimate it obeys

Now *freeze* a source $h \in L^2(0,T;\mathcal{E})$ and solve the **affine** endpoint problem $\Psi' + A\Psi + Bv + h = 0$, $\Psi(0) = \Phi_0$, $\Psi(T) = \Phi_T$:

$$d_h = \Phi_T - S(T)\Phi_0 + \int_0^T S(T-s)h(s)\,ds,$$

$$v_h = \boxed{\mathcal{C}_T^* \mathcal{W}_T^{-1} d_h} \qquad \text{(minimum-norm HUM control).}$$

From a frozen source to HUM, and the estimate it obeys

Now *freeze* a source $h \in L^2(0,T;\mathcal{E})$ and solve the **affine** endpoint problem $\Psi' + A\Psi + Bv + h = 0$, $\Psi(0) = \Phi_0$, $\Psi(T) = \Phi_T$:

$$d_h = \Phi_T - S(T)\Phi_0 + \int_0^T S(T-s)h(s)\,ds,$$

$$v_h = \boxed{\mathcal{C}_T^* \mathcal{W}_T^{-1} d_h} \qquad \text{(minimum-norm HUM control).}$$

Lemma 8 (Energy-space affine source-to-trajectory estimate)

For $h_1, h_2 \in L^2(0,T;\mathcal{E}))$,

$$\|v_{h_1} - v_{h_2}\|_{L^2} \le \Gamma_T \sqrt{T}\, \|h_1 - h_2\|_{L^2},$$

$$\|\mathcal{H}_T(h_1) - \mathcal{H}_T(h_2)\|_{C([0,T];\mathcal{E})} \le C_T^{\mathrm{aff}} \|h_1 - h_2\|_{L^2}.$$

Constants depend on T and the observability constant, not on any truncation.

Theorem 9: global energy-space convergence

Define the full source-frozen HUM map on $\mathcal{X} = C([0,T];\mathcal{E})$:

$$\mathcal{T}_\infty(\Phi) := \mathcal{H}_T\big(G(\Phi)\big), \qquad G(\Phi) = (0, g(u)).$$

Theorem 9: global energy-space convergence

Define the full source-frozen HUM map on $\mathcal{X} = C([0,T];\mathcal{E})$:

$$\mathcal{T}_\infty(\Phi) := \mathcal{H}_T\big(G(\Phi)\big), \qquad G(\Phi) = (0, g(u)).$$

Theorem 9 (Global convergence of the source-frozen HUM map)

Set

$$q_\infty := C_T^{\mathrm{aff}} \sqrt{T}\, L_g.$$

If $q_\infty < 1$ then $\mathcal{T}_\infty$ is a contraction on $\mathcal{X}$; its unique fixed trajectory Φ^* and control $v^* = v_{G(\Phi^*)}$ steer the full semilinear equation from Φ_0 to Φ_T in the mild sense.

Theorem 9: consequences and proof

Geometric convergence from *any* start:

$$\|\Phi^{(n)} - \Phi^*\|_{\mathcal{X}} \le q_\infty^{\,n} \|\Phi^{(0)} - \Phi^*\|_{\mathcal{X}}$$

$$\|v^{(n+1)} - v^*\|_{L^2(0,T;\mathcal{U})} \le \Gamma_T T L_g \, q_\infty^{\,n} \|\Phi^{(0)} - \Phi^*\|_{\mathcal{X}}$$

Computable a posteriori bound:

$$\|\Phi^{(n)} - \Phi^*\|_{\mathcal{X}} \le \frac{\|\Phi^{(n)} - \mathcal{T}_\infty(\Phi^{(n)})\|_{\mathcal{X}}}{1 - q_\infty}$$

Relaxation is free: $(1-\eta)I + \eta\mathcal{T}_\infty$ has the same fixed point and factor $\le 1 - \eta + \eta q_\infty$.

Proof in one line: Assumption 1 gives $\|G(\Phi) - G(\Psi)\| \le \sqrt{T} L_g \|\Phi - \Psi\|_{\mathcal{X}}$; compose with Lemma 8 and apply Banach's fixed point theorem.

The coarse criterion is too lossy and what replaces it

Proposition 36 (Quantitative lower bounds for the coarse criteria)

For $\Omega = (0,1)$, $T = 2.5$, $g(u) = \beta \tanh u$ (so $L_g^* = \beta/\pi$):

$$q_\infty \;\geq\; \frac{2T\beta}{\pi} = \frac{5\beta}{\pi} > 1, \qquad q_8^{\mathrm{glob}} \;\geq\; TL_g^*\sqrt{\mu_{\max}(8)} = 20\beta > 1.$$

Both exceed one already at $\beta = 1$. The sup-norm criteria certify nothing on the actual benchmark.

Why: C_T^{aff} bounds the source response and the HUM endpoint correction *separately*. But the endpoint correction is *computed from* the source response, so the two are strongly correlated, and the discarded cancellation is lossy by a factor that does not improve with resolution.

The exact-response criterion

Fix (Theorem 11). Let $\mathcal{H}_{\infty,T}$ be the *exact endpoint-corrected* source response and $\chi_{\infty,T} = \|\mathcal{H}_{\infty,T}\|$. Then

$$\widehat{q}_\infty := \ell_g \, \chi_{\infty,T} < 1 \implies \text{contraction, unique fixed displacement, mild steering pair.}$$

The exact-response criterion

Fix (Theorem 11). Let $\mathcal{H}_{\infty,T}$ be the *exact endpoint-corrected* source response and $\chi_{\infty,T} = \|\mathcal{H}_{\infty,T}\|$. Then

$$\widehat{q}_\infty := \ell_g \, \chi_{\infty,T} < 1 \implies \text{contraction, unique fixed displacement, mild steering pair.}$$

Proposition 37 (*Validated* retained-model certificate)

For $P = 8$, $T = 2.5$, collar $\delta = 0.3$, computer-assisted interval arithmetic (600 cells, rank-500 trace–Ritz, outward rounding) gives

$$\chi_{8,2.5} < 0.160561 < \tfrac{1}{6} \implies \text{global contraction for } \beta \le 6.$$

Remark: the certificate is for the *retained* $P = 8$ map; Theorem 11 provides the full-PDE statement.

Two realizations, one map

At iteration n, freeze $s^{(n)} = G_P(a^{(n)})$, $b^{(n)} = \begin{pmatrix} 0 \\ -s^{(n)} \end{pmatrix}$: what remains is the **affine** system $Y' = A_P Y + B_P c + b^{(n)}$ with $Y(0) = Y_0$, $Y(T) = Y_T$.

SF–Picard–HUM

- adjoint solve *every* iteration
- returns the min-modal-norm $c^{(n+1)}$

Picard–Gramian

$$c^{(n+1)}(t) = B_P^\top e^{A_P^\top (T-t)} \mathcal{W}_T^{-1} d^{(n)}$$

- A_P, B_P, T are *independent* of the frozen source

$\Rightarrow$ assemble and factorize $\mathcal{W}_T$ once, reuse it

Two realizations, one map

At iteration n, freeze $s^{(n)} = G_P(a^{(n)})$, $b^{(n)} = \begin{pmatrix} 0 \\ -s^{(n)} \end{pmatrix}$: what remains is the **affine** system $Y' = A_P Y + B_P c + b^{(n)}$ with $Y(0) = Y_0$, $Y(T) = Y_T$.

SF–Picard–HUM

- adjoint solve *every* iteration
- returns the min-modal-norm $c^{(n+1)}$

Picard–Gramian

$$c^{(n+1)}(t) = B_P^\top e^{A_P^\top (T-t)} \mathcal{W}_T^{-1} d^{(n)}$$

- A_P, B_P, T are *independent* of the frozen source

$\Rightarrow$ assemble and factorize $\mathcal{W}_T$ once, reuse it

Proposition 42 (equivalence)

Same space, endpoints, frozen source, actuator matrix and cost $\Rightarrow$ both return the *same* minimizer: $\mathcal{T}_P^{\mathrm{PG}} = \mathcal{T}_P^{\mathrm{HUM}}$. So the difference is computational, not mathematical.

When contraction is not available: damping, exactly characterized

Local criterion (Proposition 38). $D\mathcal{T}_P(a^*) = \mathcal{H}_{P,T}\, DG_P(a^*)$; spectral radius $< 1 \Rightarrow$ local convergence.

When contraction is not available: damping, exactly characterized

Local criterion (Proposition 38). $D\mathcal{T}_P(a^*) = \mathcal{H}_{P,T}\, DG_P(a^*)$; spectral radius $< 1 \Rightarrow$ local convergence.

Proposition 41 (exact scalar-damping stabilizability criterion)

Let $J = D\mathcal{T}_P(a^*)$ in a finite-dimensional time discretization of the trajectory map. Then

$$\exists\, \eta > 0 : \; r\big((1-\eta)I + \eta J\big) < 1 \quad \Longleftrightarrow \quad \operatorname{Re}\lambda < 1 \;\; \forall\, \lambda \in \sigma(J).$$

More precisely, every $0 < \eta < \min_{\lambda \in \sigma(J)} \frac{2(1-\operatorname{Re}\lambda)}{|\lambda-1|^2}$ is locally stabilizing. Hence $\operatorname{Re}\lambda \geq 1$ cannot be stabilized by any positive scalar under-relaxation.

When contraction is not available: diagnostics

β	certified upper for $\widehat{q}_8^{\mathrm{glob}}$ (exact-response, *not* Prop. 36's q_8^{glob})	$r(D\mathcal{T}_8)$	selected η	$r\big((1-\eta)I+\eta D\mathcal{T}_8\big)$
1	0.1606	–	–	–
2	0.3212	–	–	–
4	0.6423	–	–	–
6	0.9634	0.589	1	0.589
8	1.2845	0.776	1	0.776
9	1.4451	0.868	1	0.868
10	1.6057	0.958	0.75	0.588
11	1.7662	1.046	0.75	0.620
12	1.9268	1.133	0.75	0.651

When contraction is not available: diagnostics

β	certified upper for $\widehat{q}_8^{\mathrm{glob}}$ (exact-response, *not* Prop. 36's q_8^{glob})	$r(D\mathcal{T}_8)$	selected η	$r\big((1-\eta)I+\eta D\mathcal{T}_8\big)$
1	0.1606	–	–	–
2	0.3212	–	–	–
4	0.6423	–	–	–
6	0.9634	0.589	1	0.589
8	1.2845	0.776	1	0.776
9	1.4451	0.868	1	0.868
10	1.6057	0.958	0.75	0.588
11	1.7662	1.046	0.75	0.620
12	1.9268	1.133	0.75	0.651

Rigorous certificate stops at $\beta = 6$. Rows 8–12 are *numerical* local diagnostics (100 time subintervals; refinement $50 \to 100$ moves the radii by $\leq 1.7 \times 10^{-3}$).

$\beta = 10$ is slow but locally *stable*; $\beta = 11, 12$ are locally *unstable* undamped; and $\mathrm{Re}\,\lambda < 1$ still holds, which is exactly why η exists.

Neural–Spectral: learn the time, keep the space exact

Ansatz. Expand in the retained Dirichlet basis; only the *temporal* coefficients are approximated by neural networks:

$$\widetilde{u}_{\theta_a}(t,x) = \sum_{j=1}^{P} a_j(t;\theta_a)\, e_j(x), \qquad \widetilde{f}_{\theta_c}(t,x) = \sum_{j=1}^{P} c_j(t;\theta_c)\, e_j(x),$$

$$a_j(t;\theta_a) = A_j^a \cdot \sigma(W_j^a t + B_j^a)$$
$$c_j(t;\theta_c) = A_j^c \cdot \sigma(W_j^c t + B_j^c)$$
(shallow, per mode).

Neural–Spectral: residual and training objective

Projected residual (the actuator overlap M sits inside it):

$$R_j(t;\theta) = a_j'' + (\mu_j + \alpha)a_j + \underbrace{\big\langle g\big(\textstyle\sum_\ell a_\ell e_\ell\big), e_j\big\rangle}_{Q_j} + \sum_{\ell=1}^{P} M_{j\ell}\, c_\ell.$$

Training objective (unsupervised).

$$\min_\theta \ \sum_{j=1}^{P} \|R_j(\cdot;\theta)\|^2_{L^2(0,T)} \ + \ \big\|\widetilde{\Phi}_\theta(0) - \Phi_0\big\|^2_{\mathcal{E}} \ + \ \big\|\widetilde{\Phi}_\theta(T) - \Phi_T\big\|^2_{\mathcal{E}}.$$

No fixed point, no adjoint solve, no target data. Since steering controls are nonunique, an optional penalty $\gamma \int_0^T c^\top M c\, dt = \gamma \|\widetilde{f}_{\theta_c} \mathbf{1}_\omega\|^2_{L^2}$ selects a *cheap* one.

Can the ansatz represent a steering pair at all?

Assumption 16. $\sigma \in C^2(\mathbb{R})$, not a polynomial. **Lemma 17.** Shallow σ-networks are dense in $W^{2,2}(0,T)$ (*two* derivatives, which is what a wave residual needs).

Theorem 24 (density relative to an admissible pair)

Assume Assumptions 1 and 16, and fix an admissible exact pair $(\Phi, \mathbf{F1}_\omega)$. For every $\varepsilon > 0$ there exist P, a common width m, and $\theta \in (\mathbb{R}^m)^{6P}$ with

$$\sup_{[0,T]} \left\| \Phi - \widetilde{\Phi}_\theta \right\|_{\mathcal{E}} < \varepsilon \quad \text{and} \quad \left\| \mathbf{F1}_\omega - \widetilde{\mathbf{F}}_{\theta_c} \mathbf{1}_\omega \right\|_{L^2} < \varepsilon.$$

Can the ansatz represent a steering pair at all?

Assumption 16. $\sigma \in C^2(\mathbb{R})$, not a polynomial. **Lemma 17.** Shallow σ-networks are dense in $W^{2,2}(0,T)$ (*two* derivatives, which is what a wave residual needs).

Theorem 24 (density relative to an admissible pair)

Assume Assumptions 1 and 16, and fix an admissible exact pair $(\Phi, \mathbf{F1}_\omega)$. For every $\varepsilon > 0$ there exist P, a common width m, and $\theta \in (\mathbb{R}^m)^{6P}$ with

$$\sup_{[0,T]} \left\| \Phi - \widetilde{\Phi}_\theta \right\|_{\mathcal{E}} < \varepsilon \quad \text{and} \quad \left\| \mathbf{F1}_\omega - \widetilde{\mathbf{F}}_{\theta_c} \mathbf{1}_\omega \right\|_{L^2} < \varepsilon.$$

- **Cor. 25:** if $q_\infty < 1$, Theorem 9 *supplies* the pair. But Prop. 36 gives $q_\infty > 1$ here, so the pair must come from an independent controllability result.
- **Rmk. 26:** Thm 24 $\neq$ Thm 22, since $\Pi_P g(u) \neq \Pi_P g(\Pi_P u)$.
- Rmk. 20: qualitative: no width rates. And Assumption 1 excludes the cubic: none of this covers the stress test.

From a small residual to a certified control

Let $\rho_P(\theta) = \left(\sum_j \|R_j\|_{L^2}^2\right)^{1/2}$, $e_{\mathrm{IC}}, e_{\mathrm{TC}}$ the endpoint mismatches, and $\widehat{\Phi}_\theta^P$ be the trajectory obtained by independently integrating the retained system under the learned control.

Theorem 28 (continuous-residual forward validation)

$$\left\|\widehat{\Phi}_\theta^P(T) - \Phi_T\right\|_{\mathcal{E}} \le \underbrace{E_{\mathrm{tar}}(P)}_{\text{target truncation}} + e_{\mathrm{TC}}(\theta) + e^{L_g T}\Big(e_{\mathrm{IC}}(\theta) + \sqrt{T}\,\rho_P(\theta)\Big).$$

The constant is independent of P, the network width, and the temporal sampling rule. No optimizer hypothesis.

From a small residual to a certified control

Let $\rho_P(\theta) = \left(\sum_j \|R_j\|_{L^2}^2\right)^{1/2}$, $e_{\mathrm{IC}}, e_{\mathrm{TC}}$ the endpoint mismatches, and $\widehat{\Phi}_\theta^P$ be the trajectory obtained by independently integrating the retained system under the learned control.

Theorem 28 (continuous-residual forward validation)

$$\left\|\widehat{\Phi}_\theta^P(T) - \Phi_T\right\|_{\mathcal{E}} \le \underbrace{E_{\mathrm{tar}}(P)}_{\text{target truncation}} + e_{\mathrm{TC}}(\theta) + e^{L_g T}\Big(e_{\mathrm{IC}}(\theta) + \sqrt{T}\,\rho_P(\theta)\Big).$$

The constant is independent of P, the network width, and the temporal sampling rule. No optimizer hypothesis.

- **Corollary 29** descends to the *implemented* discrete loss, splitting the error into $E_{\mathrm{net}} + \eta_Q + E_{\mathrm{opt}}$ (approximation, certified quadrature, optimization gap).

The bridge: certify *any* candidate, then correct it

Nonlinear endpoint map $F(v) := \Phi_v(T) - \Phi_T \in \mathcal{E}$. Let $\bar{v}$ be any computed control.

Theorem 13 (a posteriori exact-control certificate)

Suppose on $B_\rho(\bar{v})$: $\|DF(v) - DF(w)\| \le L\|v - w\|$, and $DF(\bar{v})$ has a right inverse $R_{\bar{v}}$ satisfying $\|R_{\bar{v}}\| \le \gamma$. Put $\varepsilon = \|F(\bar{v})\|_{\mathcal{E}}$ and

$$h := 2L\gamma^2\varepsilon.$$

If $h < 1$ and $\gamma r_* \le \rho$ with $r_* = \frac{1-\sqrt{1-h}}{L\gamma^2}$, then an exact control v^* exists with

$$\|v^* - \bar{v}\|_{\mathcal{V}} \le \gamma r_*.$$

The bridge: Newton–HUM and computable constants

Theorem 14 (Newton–HUM).
$v_{n+1} = v_n - R_{v_n} F(v_n)$ gives

$$\varepsilon_{n+1} \le \frac{L\gamma^2}{2}\,\varepsilon_n^2,$$

quadratic convergence.

At the retained level, all three are computable:

$$L = L_P^{\mathrm{end}}, \quad \gamma = \gamma_P(c), \quad \varepsilon = \|F_P(c)\|.$$

- L_P^{end}: closed form (Prop. 35), only for $\Omega = (0,1)$, $\alpha = 0$
- $\gamma_P(c) = \lambda_{\min}(W_{P,c}^E)^{-1/2}$, provided the energy-scaled *linearized* Gramian is positive definite
- ε is the retained terminal energy error

The bridge: Newton–HUM and computable constants

Theorem 14 (Newton–HUM). $v_{n+1} = v_n - R_{v_n} F(v_n)$ gives

$$\varepsilon_{n+1} \le \frac{L\gamma^2}{2}\,\varepsilon_n^2,$$

quadratic convergence.

At the retained level, all three are computable:

$$L = L_P^{\mathrm{end}}, \quad \gamma = \gamma_P(c), \quad \varepsilon = \|F_P(c)\|.$$

- L_P^{end}: closed form (Prop. 35), only for $\Omega = (0,1)$, $\alpha = 0$
- $\gamma_P(c) = \lambda_{\min}(W_{P,c}^E)^{-1/2}$, provided the energy-scaled *linearized* Gramian is positive definite
- ε is the retained terminal energy error

Solver-independent: neural, Picard, reduced-order or hand-built; the theorem never asks where $\bar{v}$ came from. It separates *certifying the control* from *optimizing the parameterization*.

Benchmark 1: setup and how everything is scored

Configuration. $\Omega = (0, 1)$, collar $\delta = 0.3$ (GCC holds), $T = 2.5$, $\alpha = 0$, $P = 8$, $\beta \in [1, 30]$; three-mode initial displacement $(a_1^0, a_2^0, a_3^0) = \frac{1}{\sqrt{2}}(1.0, 0.5, -0.3)$, $u^1 = 0$, steered to rest.

One metric, method-independent

The recovered spectral control is fed into a high-accuracy DOP853 integration of the retained nonlinear system; we report $E_{\text{same}} = \|\Phi(T) - \Phi_T\|_{\mathcal{E}}$. Neither method is scored by its own loss.

Benchmark 1: setup and how everything is scored

Configuration. $\Omega = (0, 1)$, collar $\delta = 0.3$ (GCC holds), $T = 2.5$, $\alpha = 0$, $P = 8$, $\beta \in [1, 30]$; three-mode initial displacement $(a_1^0, a_2^0, a_3^0) = \frac{1}{\sqrt{2}}(1.0, 0.5, -0.3)$, $u^1 = 0$, steered to rest.

One metric, method-independent

The recovered spectral control is fed into a high-accuracy DOP853 integration of the retained nonlinear system; we report $E_{\text{same}} = \|\Phi(T) - \Phi_T\|_{\mathcal{E}}$. Neither method is scored by its own loss.

A second, harder metric

The *physical* control field is embedded into a richer Galerkin system, $P_{\text{val}} \in \{32, 64\}$, and re-integrated: E_{out}. The 32- and 64-mode values agree to better than 1%, so E_{out} is a property of the control, not of the validator.

Benchmark 1: setup and how everything is scored

Configuration. $\Omega = (0,1)$, collar $\delta = 0.3$ (GCC holds), $T = 2.5$, $\alpha = 0$, $P = 8$, $\beta \in [1, 30]$; three-mode initial displacement $(a_1^0, a_2^0, a_3^0) = \frac{1}{\sqrt{2}}(1.0,\ 0.5,\ -0.3)$, $u^1 = 0$, steered to rest.

One metric, method-independent

The recovered spectral control is fed into a high-accuracy DOP853 integration of the retained nonlinear system; we report $E_{\text{same}} = \|\Phi(T) - \Phi_T\|_{\mathcal{E}}$. Neither method is scored by its own loss.

A second, harder metric

The *physical* control field is embedded into a richer Galerkin system, $P_{\text{val}} \in \{32, 64\}$, and re-integrated: E_{out}. The 32- and 64-mode values agree to better than 1%, so E_{out} is a property of the control, not of the validator.

Picard convergence is declared on the *unrelaxed* proposal residual $r_{\text{fp}}^{(n)} = \|\hat{a}^{(n+1)} - a^{(n)}\|_{L^2}$, tolerance 10^{-7}, so "converged" and "steers well" are two separate questions.

Benchmark 1: the deterministic sweep

Undamped iteration converges

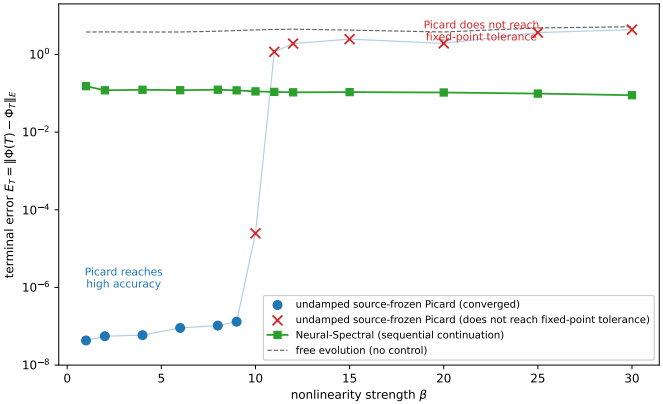
β	Picard E_T	E_T^{free}
1	4.31×10^{-8}	3.78
2	5.55×10^{-8}	3.80
4	5.90×10^{-8}	3.78
6	9.06×10^{-8}	3.78
8	1.04×10^{-7}	3.97
9	1.31×10^{-7}	4.12

It does not

β	Picard E_T	E_T^{free}
10	2.47×10^{-5}	4.28
11	1.17	4.40
12	1.90	4.47
15	2.48	4.23
20	1.92	3.78
25	3.65	4.81
30	4.34	5.16

- $\beta \leq 9$: six to seven orders below free evolution.
- $\beta = 10$ is transitional: the iteration stalls ($r_{\text{fp}} = 2.31 \times 10^{-5}$) yet the returned iterate still steers ($2.47 \times 10^{-5} \ll 4.28$). *Failing the tolerance $\neq$ failing the task.*
- $\beta \geq 11$: error is $\mathcal{O}(1)$ (no longer a steering result).

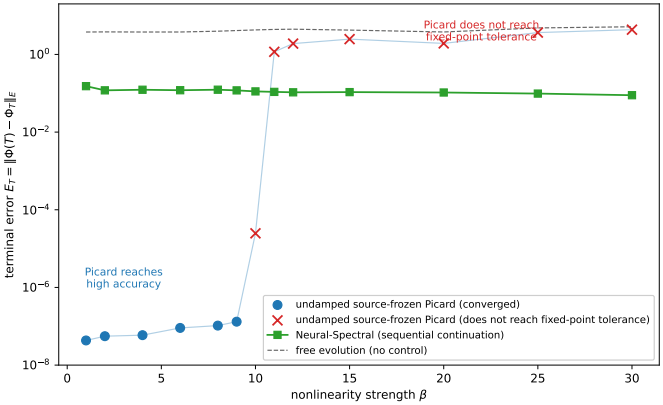
Benchmark 1: both methods on one axis



The Neural–Spectral curve.

- trained *once* at $\beta = 1$, then warm-started along the grid;
- a finite control at every sampled β, no per-β tuning;
- $E_{\mathrm{same}} \approx 0.9\text{–}1.5 \times 10^{-1}$ throughout.

Benchmark 1: both methods on one axis



The Neural–Spectral curve.

- trained *once* at $\beta = 1$, then warm-started along the grid;
- a finite control at every sampled β, no per-β tuning;
- $E_{\text{same}} \approx 0.9\text{–}1.5 \times 10^{-1}$ throughout.

Two different claims: converged or stabilized Picard gives the smallest validated errors; Neural–Spectral supplies a finite candidate across the whole transition.

Damping removes the transition

Predeclared schedule $\eta : 1 \to 0.75 \to 0.5 \to 0.25$, first value reaching $r_{\mathrm{fp}} < 10^{-7}$ within 200 iterations. *Fixed in advance, not chosen after seeing the errors.*

β	selected η	N_{iter}	E_T	$\|f\mathbf{1}_\omega\|_{L^2}$
10	0.75	26	4.7×10^{-7}	8.87
11	0.75	29	5.2×10^{-7}	9.56
12	0.75	32	6.6×10^{-7}	10.29
20	0.50	117	1.3×10^{-6}	17.42

At $\beta = 20$ both $\eta = 1$ *and* $\eta = 0.75$ fail, so the schedule steps to 0.5; it is genuinely adaptive, not post hoc. And the selected control there has norm 17.42 against 7.94 for the failed undamped iterate: amplitude is one contributor to the improved steering, not just better convergence.

Benchmark 1: richer-system validation

Validate the control in a *richer* system ($P = 8 \to P_{\mathrm{val}} = 32$)

		Source-frozen Picard		Neural–Spectral (median)	
β	η	E_{same}	$E_{\mathrm{out}}^{(32)}$	E_{same}	$E_{\mathrm{out}}^{(32)}$
1	1.00	4.31×10^{-8}	1.77×10^{-2}	0.153	0.227
10	0.75	4.70×10^{-7}	2.56×10^{-2}	0.191	0.217
20	0.50	1.34×10^{-6}	5.28×10^{-2}	0.266	0.311

The localized control excites modes *above* P. In-truncation validation alone understates the true error by more than five orders of magnitude (4.31×10^{-8} vs. 1.77×10^{-2} at $\beta = 1$).

At the common level, deterministic beats neural by $12.8\times$, $8.5\times$, $5.9\times$ at $\beta = 1, 10, 20$.

Theorems 13 and 14: certifying the neural controls

Three median-error **Neural–Spectral** controls. No retraining. Independently integrate, assemble the linearized Gramian, use the closed-form L_P^{end}, evaluate the indicator

$$h_P = 2L_P^{\mathrm{end}}\bar{\gamma}_P^2\,\bar{\varepsilon}_P.$$

β	seed	L_P^{end}	$\bar{\gamma}_P$	$\bar{\varepsilon}_0$	$h_P(c_0)$
1	0	6.43×10^{-1}	2.0	1.53×10^{-1}	0.787
10	0	7.40×10^{2}	2.6	1.92×10^{-1}	1.92×10^{3}
20	20	1.10×10^{6}	3.2	2.66×10^{-1}	5.98×10^{6}

$\bar{\varepsilon}_0$ barely moves, but L_P^{end} grows six orders. Prop. 35 has $\kappa = \beta/2\pi$ and a factor $e^{\kappa T}(\cosh \kappa T - 1)/\kappa^2$, so L is exponential in β. The indicator degrades because the *problem* stiffens, not because the candidate worsens.

Theorems 13 and 14: after Newton–HUM correction

β	Newton steps	$\bar{\varepsilon}_{\mathrm{corr}}$	$h_P(c_{\mathrm{corr}})$
1	2	1.00×10^{-8}	5.14×10^{-8}
10	2	1.00×10^{-8}	1.00×10^{-4}
20	2	1.25×10^{-8}	2.81×10^{-1}

Theorems 13 and 14: after Newton–HUM correction

β	Newton steps	$\bar{\varepsilon}_{\mathrm{corr}}$	$h_P(c_{\mathrm{corr}})$
1	2	1.00×10^{-8}	5.14×10^{-8}
10	2	1.00×10^{-8}	1.00×10^{-4}
20	2	1.25×10^{-8}	2.81×10^{-1}

- At $\beta = 1$ the indicator was already below one before any correction: Theorem 13 then gives a nominal radius 0.419 at *solver-level* accuracy, not a validated-interval proof like Prop. 37, though the same formula accepts enclosures.
- Two steps bring *all three* indicators below one and the retained defect to solver tolerance, exactly the quadratic rate of Thm 14.
- A deterministic post-processor bolted onto a stochastic procedure: no retraining, no optimizer hypothesis.

Newton–HUM correction: retained and richer errors

	before correction		after 2 steps	
β	E_{same}	$E_{\text{out}}^{(32)}$	E_{same}	$E_{\text{out}}^{(32)}$
1	0.152	0.227	1.9×10^{-10}	0.168
10	0.191	0.217	6.5×10^{-10}	0.102
20	0.266	0.286	1.3×10^{-9}	0.106

- The richer error does *not* vanish ($0.227 \to 0.168$): the correction lives in the retained space, while the defect it cannot see lives above P.
- The correction norm equals the theorem radius in all three rows (0.188, 0.235, 0.358): the Newton step lands exactly on the boundary Theorem 13 predicted.

Benchmark 2: a focusing cubic the ladder cannot reach

Same geometry, same T, same P. Only the nonlinearity changes: $g(u) = -\beta u^3$, $\beta = 11$ (non-saturating and anti-restoring).

Benchmark 2: a focusing cubic the ladder cannot reach

Same geometry, same T, same P. Only the nonlinearity changes: $g(u) = -\beta u^3$, $\beta = 11$ (non-saturating and anti-restoring).

The traced branch.

Newton–Krylov then pseudo-arclength continuation from $\beta = 0$:

- the branch folds near $\beta^* \simeq 10.68$;
- on the *upper sheet* the dominant eigenvalue is real, positive, > 1: by Prop. 41 no $\eta > 0$ stabilizes *those* fixed points;
- at $\beta = 11$ the continued branch carries no fixed point at all, so Prop. 41 has *nothing to govern* there: it rules out the upper-sheet fixed points, and is not claimed to cause the divergence.

The predeclared ladder, run anyway

$\eta \in \{1, 0.75, 0.5, 0.25, 0.125, 0.0625, 0.03125\}$;
$N_{\max} = 1000$.

- All 28 runs diverge (4 configurations $\times$ 7 values of η: $P = 8$ at $\beta = 11, 12$; $P = 12$ and $P = 16$ at $\beta = 11$).
- By overflow, not by budget: residual passes 10^{12}.

Below the fold the branch is two-sheeted, and plain iteration finds the *stable* point. That is why the benchmark sits at $\beta = 11$, past β^*, rather than at $\beta = 10$.

Focusing cubic: what the direct method still delivers

Neural-Spectral on the same problem: three cold seeds

All three complete the budget with finite applied-control norm:

$$E_{\text{same}} = 0.174,\ 0.200,\ 0.179; \quad E_{\text{out}}^{(32)} \approx 0.21$$

the two richer validators agree to 0.14%; every forward solve is checked to reach T.

What is *not* claimed: that no fixed point exists; that no other η/initialization/globalization would work; that the neural optimizer converges.

Only: *this* map, *this* initialization, *these* seven parameters do not produce a control here; and the direct formulation does.

Benchmark 3: the classical reference

$\Omega = (0,1)^2$, $T = 7$, collar $\delta = 0.3$ (84% of the domain), *linear* dynamics, $4 \times 4 = 16$ retained modes, validated at 8×8.

control	E_{same}	$E_{\text{out}}^{(8\times8)}$	$\|c\|_{L^2}^{\text{modal}}$	$\|\mathbf{1}_{\omega_\delta} f\|_{L^2}$	wall (s)
modal-coeff. cost ($R = I$)	1.5×10^{-10}	$\mathbf{4.0 \times 10^{-3}}$	**2.10**	2.04	1.2 [1.2, 1.3]
physical energy ($R = M$)	1.4×10^{-10}	7.1×10^{-3}	2.62	**1.97**	1.3 [1.2, 1.4]

Benchmark 3: the classical reference

$\Omega = (0,1)^2$, $T = 7$, collar $\delta = 0.3$ (84% of the domain), *linear* dynamics, $4 \times 4 = 16$ retained modes, validated at 8×8.

control	E_{same}	$E_{\text{out}}^{(8\times8)}$	$\|c\|_{L^2}^{\text{modal}}$	$\|\mathbf{1}_{\omega_\delta} f\|_{L^2}$	wall (s)
modal-coeff. cost ($R = I$)	1.5×10^{-10}	$\mathbf{4.0 \times 10^{-3}}$	**2.10**	2.04	1.2 [1.2, 1.3]
physical energy ($R = M$)	1.4×10^{-10}	7.1×10^{-3}	2.62	**1.97**	1.3 [1.2, 1.4]

Each weighting wins only its own objective

$R = I$ minimizes the modal cost ($2.10 < 2.62$); $R = M$ minimizes the physical field ($1.97 < 2.04$). But $R = M$ pays for it with a larger richer-system error (7.1 vs 4.0×10^{-3}): a genuine trade-off between effort and out-of-truncation accuracy.

Wall time is the median [min, max] over 11 repetitions, one core, excluding the one-time actuator-overlap construction. Quadrature refinement ($2 \times 10^3 \to 8 \times 10^3$ Simpson subintervals)

Benchmark 3: the neural formulations

method	median E_T	range	E_{rel}	$\|\mathbf{1}_{\omega_\delta} f\|_{L^2}$
Neural–Spectral (3 seeds)	0.520	[0.517, 0.664]	**0.18**	11.4
Vanilla PINN (single run)	2.69	—	0.95	0.64
Hard-BC PINN, $\lambda_{ep}=500$ (3 seeds)	2.83	[2.82, 2.83]	1.00	0.067

$E_{\mathrm{free}} = 2.83$. Norms are physical applied-control norms; the Neural–Spectral value is the median-error seed (3-seed median 11.32).

Benchmark 3: the neural formulations

method	median E_T	range	E_{rel}	$\|\mathbf{1}_{\omega_\delta} f\|_{L^2}$
Neural–Spectral (3 seeds)	0.520	[0.517, 0.664]	**0.18**	11.4
Vanilla PINN (single run)	2.69	—	0.95	0.64
Hard-BC PINN, $\lambda_{ep}=500$ (3 seeds)	2.83	[2.82, 2.83]	1.00	0.067

$E_{\mathrm{free}} = 2.83$. Norms are physical applied-control norms; the Neural–Spectral value is the median-error seed (3-seed median 11.32).

The comparison hides an amplitude gap

Neural–Spectral is $\approx 5\times$ better than either PINN, while applying $\approx 18\times$ more control. Penalize the effort ($\lambda_f = 3$, seed 0): norm drops to 0.74 and $E_T = 2.531$, right next to the vanilla PINN's 0.64/2.697.

On the hard-BC row: it learned essentially no control (norm 0.067, unchanged across $\lambda_{ep} \in \{100, 300, 500\}$); it bounds *this* budget and parameterization on *this* geometry, not hard-constraint formulations.

Terminal accuracy is not comparable across methods minimizing different objectives at different control amplitudes.

Conclusions

1. **Operator level.** Under GCC and a smallness condition the source-frozen HUM map contracts. The coarse criterion is too lossy on the actual benchmark; the *exact-response* constant $\chi_{P,T}$ replaces it, with a validated enclosure $\chi_{8,2.5} < 0.160561$ certifying $\beta \in \{1, 2, 4, 6\}$.

2. **Algorithmic level.** SF–Picard–HUM and Picard–Gramian are the *same* map, so one factorization is assembled and reused. Scalar damping is characterized exactly: it works iff $\operatorname{Re}\lambda < 1$.

Conclusions

1. **Operator level.** Under GCC and a smallness condition the source-frozen HUM map contracts. The coarse criterion is too lossy on the actual benchmark; the *exact-response* constant $\chi_{P,T}$ replaces it, with a validated enclosure $\chi_{8,2.5} < 0.160561$ certifying $\beta \in \{1, 2, 4, 6\}$.

2. **Algorithmic level.** SF–Picard–HUM and Picard–Gramian are the *same* map, so one factorization is assembled and reused. Scalar damping is characterized exactly: it works iff $\mathrm{Re}\,\lambda < 1$.

3. **Certification level.** An endpoint defect, a HUM right inverse and a Lipschitz bound yield an explicit ball containing an exact control, for any candidate, from any solver, with a quadratically convergent corrector.

4. **Experiments.** Stabilized Picard is the accurate reference wherever it converges; **Neural–Spectral** stays operational where the entire predeclared ladder diverges. Out-of-truncation validation is not optional.

Selected references

- C. Bardos, G. Lebeau, J. Rauch. *Sharp sufficient conditions for the observation, control and stabilization of waves from the boundary.* SIAM J. Control Optim., 1992.
- A. Münch, E. Trélat. *Constructive exact control of semilinear 1D wave equations by a least-squares approach.*
- S. Claret, J. Lemoine, A. Münch. *Fixed-point approaches for semilinear boundary control.*
- Y. Du, N. Chalapathi, A. Krishnapriyan. *Neural spectral methods.* ICLR.
- M. Feng et al. *Optimal control operator perspective and a neural adaptive spectral method.* AAAI 39, 2025.
- M. Raissi, P. Perdikaris, G. E. Karniadakis. *Physics-informed neural networks.* J. Comput. Phys., 2019.

Thank you!