# CERTIFYING PHYSICS-INFORMED SURROGATES FOR PDE-CONSTRAINED OPTIMIZATION

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Physics-informed neural surrogates can reduce the cost of PDE-constrained optimization, but an optimizer itself can drive iterates into regions where the surrogate is least reliable. We introduce SA-CPITR, a solver-assisted certified trust-region method that treats a learned PDE surrogate as an inexact optimization oracle whose errors are bounded online. A posteriori majorants convert PDE residual information into computable error bounds for the state and sensitivities, which we assemble into certificates for the reduced objective and its gradient. Since the majorants hold for any conforming candidate, the same formulas certify both the neural surrogate and the high-fidelity solver outputs. These certificates control trust-region acceptance and termination, guaranteeing objective descent for every accepted step without evaluating the true objective. When a certificate becomes insufficient, the method spends state or gradient high-fidelity solves to generate a reusable anchor that tightens subsequent certificates. If the target tolerance cannot be certified within the anchor event budget, the algorithm returns its best certified bound rather than an uncertified solution. We prove certificate validity, guaranteed descent, and convergence to first-order criticality under explicit assumptions. On five PDE control problems, from linear elliptic to nonlinear and hyperbolic classes, SA-CPITR certifies terminal criticality on all five with a total of only three anchor events, while the corresponding solver-free variant certifies two. A sixth problem extends the method to state constraints, with certified feasibility and descent. The experiments use a fixed quadrature grid to evaluate integrals inside the optimization loop, and then independently re-verify certificates at every returned point by exact integration or continuum-valid enclosures.

## 1 INTRODUCTION

PDE-constrained optimization is computationally expensive due to repeated PDE evaluations (Tröltzsch, 2010; Hinze et al., 2009). Neural surrogates, such as physics-informed neural networks (PINNs) (Raissi et al., 2019) and neural operators (Lu et al., 2021a; Kovachki et al., 2023), can reduce this cost significantly. They can solve optimality systems directly, with the control as a trainable input (Mowlavi & Nabi, 2023; Lu et al., 2021b), or approximate the control-to-state map, i.e. train a surrogate offline and optimize over its inputs (Hwang et al., 2022). However, none of these methods provides a guarantee at the returned optimal control/design, i.e. a certified solution. The difficulty is structural: an optimizer can exploit regions where the surrogate is most inaccurate, returning a spurious solution, and PINN failure modes are already well documented (Krishnapriyan et al., 2021; Wang et al., 2022).

The two systematic approaches to solving this trust issue cost high-fidelity solves: trust-region model management restores first-order consistency using high-fidelity solves (Alexandrov et al., 1998; Conn et al., 2000), and certified reduced-basis trust-region methods enrich a projection-based model with full-order solves when its online error bound degrades (Qian et al., 2017; Keil et al., 2021; Kartmann et al., 2024). Inexact-gradient trust-region theory instead assumes sufficiently accurate gradients (Carter, 1991; Heinkenschloss & Vicente, 2002; Kouri et al., 2013; 2014; Ziems & Ulbrich, 2011). However, for learned surrogates, such as PINNs, this accuracy must be checked online during optimization. Even though PINNs reach beyond the setting in which reduced-basis certification is most mature (meshfree, differentiable in the control, applicable to nonlinear state equations), they lack exactly what made the reduced-basis approach rigorous: an online, computable, guaranteed



Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

error bound. Existing generalization bounds (Mishra & Molinaro, 2023) involve unknown constants and are not computable certificates, and a posteriori bounds for trained networks (Muzalevskiy & Repin, 2023; Berrone et al., 2022; Hillebrecht & Unger, 2022; Fanaskov et al., 2024), including continuum-valid residual bounds obtained by bound propagation through the network (Eiras et al., 2024), have not been connected to an optimization loop with convergence guarantees. We connect computable state and sensitivity error bounds to trust-region acceptance and termination, and establish certified neural surrogates for PDE-constrained optimization. To our knowledge, this is the first framework that certifies a neural surrogate through its own PDE residuals inside an optimization method with convergence guarantees. Methods that use trust regions to train PINNs (Cheng & Na, 2024) address a different optimization problem. Appendix I discusses related work in full.

**Parametrized controls.** We work in the setting where the control function is an affine combination of $p$ fixed basis functions, $z(x) = z_0(x) + \sum_{j=1}^{p} c_j \varphi_j(x)$, and the optimization variable is the coefficient vector $c \in \mathbb{R}^p$. This is how engineering control and design problems usually arrive (actuator amplitudes, heater powers, spline-parametrized time profiles of a flux, shape parameters (Tröltzsch, 2010; Hinze et al., 2009)), and it is exactly the finite-element, spline or modal ansatz for the control, which approximates any smooth control to high order with moderate $p$. This also makes a single parametric surrogate possible: one network takes $(x, c)$ as input, represents the whole control-to-state map, and delivers sensitivities $\partial_{c_j} u_\theta$ by automatic differentiation.

In this paper, we introduce the SA-CPITR (solver-assisted certified physics-informed trust region) method, which closes the certification gap with a strong design goal: the solver is called only when a computable certificate proves the surrogate's information content is insufficient, and each call comes with a theorem stating what is achieved by it (Figure 1). Throughout, the surrogate is an inexact optimization oracle whose errors are bounded online. The framework is based on a family of functional majorants (Repin, 2000; 2008): computable functionals $\mathcal{M}(v; c)$ that bound $\|u(c) - v\|$ for any conforming input $v$. When evaluated at the surrogate, they provide certified error bounds on the surrogate objective and the reduced gradient, which gate every trust-region step so that accepted steps reduce the true objective $\widehat{J}$, even though $\widehat{J}$ is never evaluated. When evaluated at a solver output, the same functional certifies the solver: the output becomes an anchor, a reusable certified reference point that sharpens the surrogate's certificates nearby and supplies a certified gradient when the surrogate's own gradient certificate stalls. The proposed SA-CPITR reproduces its zero-solve variant, CPITR, exactly on problems the surrogate can certify alone. As the surrogate degrades, SA-CPITR spends its given budget of at most $N_{\max}$ anchor events and returns its best certified bound if the target tolerance cannot be reached.

**Contributions.** The contributions are: (i) a model-agnostic certificate framework (Section 3.1): objective and gradient certificates assembled from PDE state and sensitivity majorants, including reusable anchored (solver) bounds; (ii) the SA-CPITR algorithm (Section 3.2): a trust-region method with separate solve events for the objective value and the gradient; (iii) convergence theory (Section 4): guaranteed descent and conditional convergence to a certified $\mathrm{tol}$-critical point; (iv) five PDE instantiations with full proofs (Appendix D): linear elliptic (C1), parabolic (C2), semilinear (C3), wave (C4) and steady Burgers (C5, with a verified fixed-point closure), and a state-constrained extension with a certified feasibility margin (C6); and (v) a controlled study (Section 5) that benchmarks the method against frozen surrogates and a solver-based optimization reference. Full proofs are in the appendix.

## 2 Preliminaries

### 2.1 Problem setup

Let $\Omega \subset \mathbb{R}^d$ ($d \leq 2$) be a bounded Lipschitz domain and let $\mathcal{C} \subseteq \mathbb{R}^p$ be a nonempty, closed and convex set of admissible coefficient vectors; $P_{\mathcal{C}}$ denotes the Euclidean projection onto $\mathcal{C}$. A coefficient vector $c \in \mathcal{C}$ defines the control function $z(c) = z_0 + \sum_{j=1}^{p} c_j \varphi_j$ in a fixed basis (Figure 1A), and $u(c)$ denotes the exact state of the governing PDE under the control $z(c)$; $w_j(c) := \partial_{c_j} u(c)$ are its sensitivities. We minimize the reduced objective

$$\min_{c \in \mathcal{C}} \ \widehat{J}(c) := J\big(u(c), c\big), \tag{1}$$

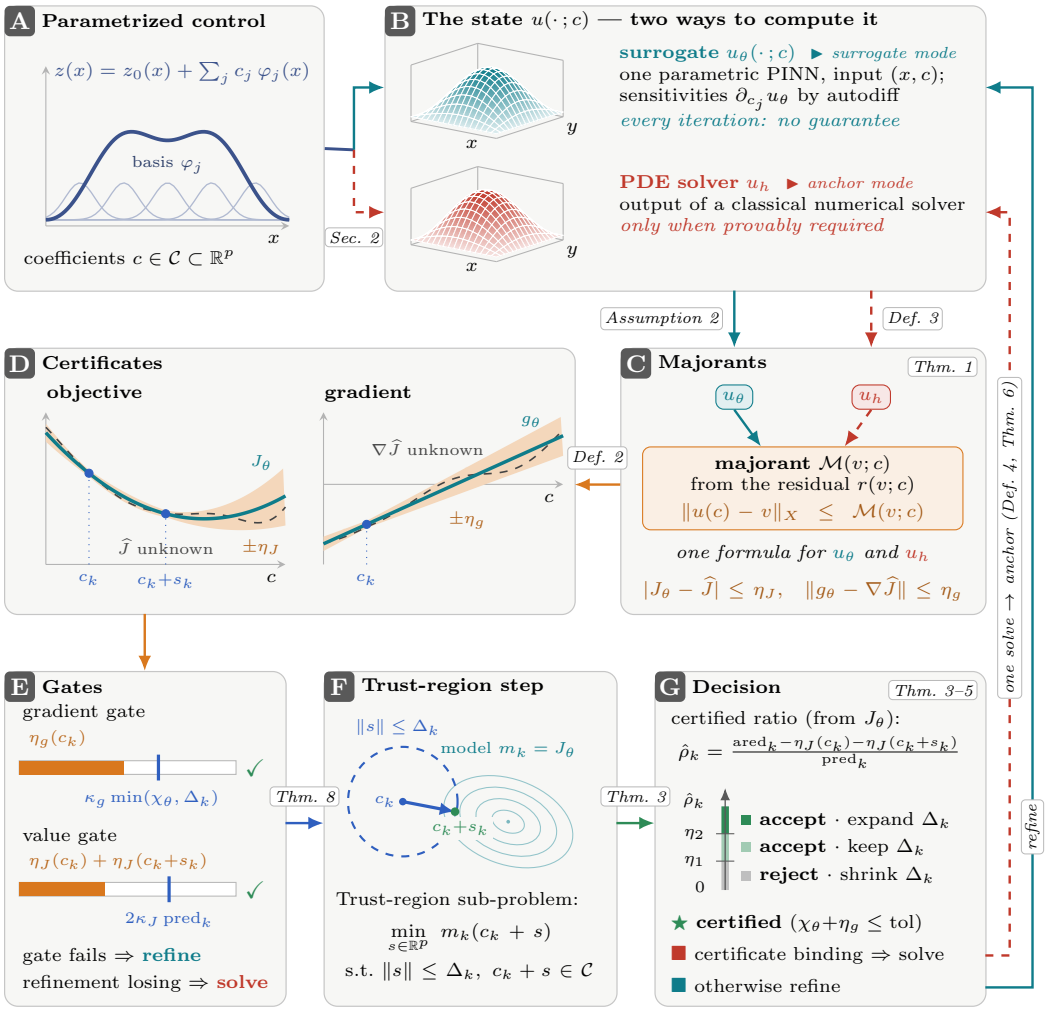



Figure 1: **Overview of the SA-CPITR method.**

with $J$ of tracking type: $\frac{1}{2}\|u(c) - u_d\|_H^2 + \frac{\alpha}{2}|c|^2$ on a separable Hilbert space $H$ (stationary tracking, $H = L^2(\Omega)$; space–time tracking, $H = L^2(0, T; L^2(\Omega))$), or the terminal "steer-and-stop" form $\frac{1}{2}\|u(T) - u_d\|^2 + \frac{1}{2}\|u_t(T)\|^2 + \frac{\alpha}{2}|c|^2$ for second-order-in-time problems. The surrogate (Figure 1B) is a hard-constraint parametric PINN $u_\theta(\cdot; c)$ (boundary and initial conditions are enforced exactly by a smooth mask, so only the strong residual is nonzero), with sensitivities $w_{\theta,j} := \partial_{c_j} u_\theta$ by automatic differentiation, and

$$J_\theta(c) := J\big(u_\theta(c), c\big), \qquad g_\theta := \nabla J_\theta.$$

Criticality is measured by the projected-gradient residuals $\chi(c) := \left\|P_{\mathcal{C}}(c - \nabla\widehat{J}(c)) - c\right\|$ (truth) and $\chi_\theta(c) := \|P_{\mathcal{C}}(c - g_\theta(c)) - c\|$ (surrogate), and, on convex $\mathcal{C}$, $\chi(c) = 0$ characterizes first-order criticality. Thus, a certificate $\chi \le \mathrm{tol}$ bounds the distance to the global optimum for convex reduced objectives (C1, C2, C4; Theorem 7) and establishes stationarity on the nonlinear C3 and C5. On C6 the run projects onto a certified inner feasible set $\mathcal{C}_\theta$ that need not be convex, so no criticality claim is made there, and the only claims are certified feasibility and descent. The optimality gap of the returned point on the true feasible set $\mathcal{C}$ is verified offline with the exact solution map and is not certified by the method (Appendix D.8). The norm $\|\cdot\|$ is the $L^2$ norm on the relevant domain for functions and the Euclidean norm on $\mathbb{R}^p$, written $|c|$ in the Tikhonov term.

### 2.2 STANDING ASSUMPTIONS

The convergence theory rests on four standing hypotheses, collected as Assumption 1 in Appendix C.1: (A1) a training oracle (finitely many surrogate refinement rounds drive the majorants of Definition 1 below any tolerance); (A2) bounded curvature of the trust-region models; (A3) smoothness and lower boundedness of the true reduced objective $\widehat{J}$; and (A4) solver consistency. We prove (A3) for C1–C6 in Appendix D, for C5 under the explicit bound $\|z(c)\| < 2\nu^2/C_{\mathrm{F}}^2$ on the feasible set, which is satisfied by the experimental data. Assumption (A4) ensures that triggered solves can succeed, and the curvature condition (A2), the one hypothesis about the trained network, cannot be verified in advance: Proposition 1 restates it as a bound on quantities every run computes.

## 3 METHOD

### 3.1 CERTIFICATES: ONE FUNCTIONAL FOR THE NETWORK AND THE SOLVER

The basis of the certificate framework is a bound-producing functional (majorant) that certifies, i.e. bounds errors for, any conforming candidate of the certifiable PDE class, network or solver output (Figure 1C).

**Assumption 2** (Certifiable problem class)**.** For the problem class there are an *observation* $\Lambda = (\Lambda_1, \dots, \Lambda_\ell)$ with values in Hilbert spaces $H_1, \dots, H_\ell$ (Template A: $\ell = 1$, $\Lambda_1 v = v \in H$; Template B: $\ell = 2$, $\Lambda_1 v = v(T)$, $\Lambda_2 v = \partial_t v(T)$, $H_1 = H_2 = L^2(\Omega)$), an admissible class $\mathcal{V}$ of conforming functions (satisfying the essential boundary/initial conditions exactly and regular enough to evaluate the strong residual; $\Lambda$ is defined on $\mathcal{V}$, on $u(c)$ and on every $w_j(c)$), and computable functionals such that, with $\|v\|_X := (\|\Lambda_i v\|_{H_i})_{i \le \ell} \in \mathbb{R}^\ell_{\ge 0}$ and all inequalities, sums, minima and monotonicity statements between such vectors read componentwise, for every $v, \omega \in \mathcal{V}$ and every $c \in \mathcal{C}$:

(i) (*majorants*) $\|u(c) - v\|_X \le \mathcal{M}(v; c) \in [0, \infty]^\ell$, and $\|w_j(c) - \omega\|_X \le \mathcal{M}_j(\omega; c, v, \mu) \in [0, \infty]^\ell$ whenever $\|u(c) - v\|_X \le \mu$, with $\mathcal{M}_j$ nondecreasing in $\mu$. The state majorant $\mathcal{M}$ takes no $\mu$ argument; the tangent majorant may depend on a bound $\mu$ for the state error (it does not in the linear classes).

(ii) (*validity flag*) a computable predicate $\mathrm{val}(v; c) \in \{0, 1\}$ is evaluated with every majorant and $\mathrm{val} = 0$ sets every component to $+\infty$ (so (i) holds for all $c$): an invalid certificate is an automatic gate failure, never a silent guarantee.

(iii) (*control shift*) there is a computable nondecreasing $\varsigma : [0, \infty) \to [0, \infty)^\ell$ with $\varsigma(0) = 0$ such that $\mathcal{M}(v; c') \le \mathcal{M}(v; c) + \varsigma(\|c' - c\|)$ for every fixed $v$ and all $c, c'$ with $\mathrm{val}(v; c') = 1$.

The norms in (i) are exactly the $H_i$-norms consumed by Theorem 1; for $\ell = 1$ everything is scalar.

Assumption 2 is a theorem for each of C1–C6 (Template A: C1–C3, C5, C6; Template B: C4), proved with all constants in Appendix D.

**Definition 1** (Surrogate certificates)**.** We define the surrogate certificates as

$$\mu(c) := \mathcal{M}(u_\theta(c); c), \qquad \mu_j(c) := \mathcal{M}_j\big(w_{\theta,j}(c); c, u_\theta(c), \mu(c)\big), \quad j = 1, \dots, p.$$

**Theorem 1** (Assembly)**.** *Let $H$ be a separable Hilbert space with norm $\|\cdot\|$ and let $c \in \mathcal{C}$.*

(A: tracking.) *Let $\widehat{J}(c) = \frac{1}{2}\|u(c) - u_d\|^2 + \frac{\alpha}{2}|c|^2$ with $u(c), u_d \in H$, and for a candidate pair $(v, \{\omega_j\}) \subset H$ set $J^v(c) := \frac{1}{2}\|v - u_d\|^2 + \frac{\alpha}{2}|c|^2$ and $g_j^v := \langle v - u_d, \omega_j \rangle + \alpha c_j$. If $\|u(c) - v\| \le \mu$ and $\|w_j(c) - \omega_j\| \le \mu_j$, then*

$$|J^v(c) - \widehat{J}(c)| \le \tfrac{1}{2}\big(2\|v - u_d\| + \mu\big)\mu =: \mathcal{E}_J(v; \mu), \tag{2}$$

$$\big|g_j^v - \partial_{c_j}\widehat{J}(c)\big| \le \mu\|\omega_j\| + \big(\|v - u_d\| + \mu\big)\mu_j =: \mathcal{E}_{g,j}(v, \omega_j; \mu, \mu_j), \tag{3}$$

*and $(\sum_j \mathcal{E}_{g,j}^2)^{1/2} \ge \Big\|g^v - \nabla\widehat{J}(c)\Big\|_2$. All right-hand sides are computable and nondecreasing in $\mu, \mu_j$. The statement covers stationary tracking ($H = L^2(\Omega)$) and space–time tracking ($H = L^2(0, T; L^2(\Omega))$, the Bochner space of square-integrable $L^2(\Omega)$-valued functions).*

(B: terminal tracking with velocity.) *Let $V \subset L^2(\Omega) \subset V^*$ be the energy space of a second-order-in-time problem, let the state and the sensitivities lie in $C([0,T];V) \cap C^1([0,T];L^2(\Omega))$, with $\partial_{c_j}(u(T), u_t(T)) = (w_j(T), \partial_t w_j(T))$, let $\widehat{J}(c) = \frac{1}{2}\|u(T) - u_d\|^2 + \frac{1}{2}\|u_t(T)\|^2 + \frac{\alpha}{2}|c|^2$, and let the candidate $v$ (a network or a conforming reconstruction) define $J^v$ in the same way. If $\mu^{(0)} \geq \|u(T) - v(T)\|$ and $\mu^{(1)} \geq \|u_t(T) - v_t(T)\|$, then*

$$|J^v(c) - \widehat{J}(c)| \leq \mathcal{E}_J\big(v(T); \mu^{(0)}\big) + \mathcal{E}_J^0\big(v_t(T); \mu^{(1)}\big),$$

*where $\mathcal{E}_J^0$ is (2) with $u_d = 0$, and the gradient error is bounded by the sum of the two corresponding terms (3), formed with majorants for $w_j(T)$ and $\partial_t w_j(T)$.*

**Definition 2** (Assembled surrogate certificates)**.** We define the assembled surrogate certificates as

$$\eta_J(c) := \mathcal{E}_J\big(u_\theta(c); \mu(c)\big), \qquad \eta_g(c) := \Big\| \big(\mathcal{E}_{g,j}(u_\theta(c), w_{\theta,j}(c); \mu(c), \mu_j(c))\big)_{j=1}^p \Big\|_2$$

(Template B analogously); by Theorem 1, $|J_\theta - \widehat{J}| \leq \eta_J$ and $\Big\|g_\theta - \nabla\widehat{J}\Big\|_2 \leq \eta_g$ at every $c$ (the tubes of Figure 1D).

In (3) the sensitivity majorant $\mu_j$ is multiplied by the tracking residual $\|u - u_d\|$, which is $O(1)$ at a typical iterate, meaning that state accuracy alone is insufficient. We therefore train the surrogate on both state and tangent residuals (Appendix F.2).

The theory uses exact integrals or validated enclosures, such as a quadrature with certified remainder bounds, interval enclosures of the network's derivatives on each cell (Appendix F.6) or network bound propagation (Eiras et al., 2024). The experiments evaluate every integral on a fixed-grid quadrature without such a bound (the midpoint rule for the network, Gauss rules aligned with the solver's breakpoints for anchors), which leads to two tiers. Returned points: the terminal certificate of Theorem 4 does not depend on the history of the run, and it is re-verified at every returned point of C1–C5 by exact integration or by continuum-valid enclosures (Section 5.2). Inside the loop: the gates, the acceptance ratio and hence the per-step descent guarantee of the network branch are grid-evaluated; as in PINN generalization bounds (Mishra & Molinaro, 2023), quadrature error is a separate term that is not controlled here, so they hold up to quadrature and are audited against ground truth (Section 5.2).

**Definition 3** (Anchor)**.** An *anchor* at $c_a \in \mathcal{C}$ is a tuple $a = (c_a, u_h, \{w_{h,j}\}, \delta_a, \{\delta_{a,j}\})$ where $(u_h, w_{h,j}) \in \mathcal{V}$ is a solver output ((A4)), $\delta_a := \mathcal{M}(u_h; c_a)$ and $\delta_{a,j} := \mathcal{M}_j(w_{h,j}; c_a, u_h, \delta_a)$; a *state-only* anchor omits the $w_{h,j}$. With $\mathcal{A}_k$ the anchors collected up to iteration $k$, the *anchored* majorants at any $c \in \mathcal{C}$ are

$$\widehat{\mu}(c) := \min\Big\{\mathcal{M}(u_\theta(c); c),\ \min_{a \in \mathcal{A}_k}\big[\mathcal{M}(u_h; c) + \|u_h - u_\theta(c)\|_X\big]\Big\}, \tag{4}$$

and analogously $\widehat{\mu}_j(c)$ with $\mathcal{M}_j(w_{h,j}; c, u_h, \mathcal{M}(u_h; c)) + \|w_{h,j} - w_{\theta,j}(c)\|_X$ in the anchor branch (11); $\widehat{\eta}_J(c), \widehat{\eta}_g(c)$ are the assemblies of Theorem 1 with the surrogate pair and $(\widehat{\mu}, \widehat{\mu}_j)$. The *anchored gradient* at the anchor point and its certificate are

$$(g_a)_j := \langle u_h - u_d, w_{h,j}\rangle + \alpha(c_a)_j, \qquad \eta_g^a(c_a) := \Big(\textstyle\sum_j \mathcal{E}_{g,j}(u_h, w_{h,j}; \delta_a, \delta_{a,j})^2\Big)^{1/2}, \tag{5}$$

and $\eta_J^a(c_a) := \mathcal{E}_J(u_h; \delta_a)$.

$\mathcal{M}(u_h; c)$ in (4) is evaluated at the query control $c$: the anchor function is fixed and its residual at control $c$ is computable, so an anchor remains usable at every $c$, with degradation quantified by Assumption 2(iii); with $\mathcal{A}_k = \emptyset$ all hatted quantities reduce to Definition 2. For $\ell = 2$ (see Assumption 2) the anchored majorants (4), (11) are formed componentwise, $\|u_h - u_\theta(c)\|_X$ being the vector of component norms.

**Theorem 2** (Anchored certificates)**.** *Let Assumption 2 hold and let $\mathcal{A}_k$ be any set of anchors. At every $c \in \mathcal{C}$ with* val $= 1$*: (i) (*validity*) $|J_\theta(c) - \widehat{J}(c)| \leq \widehat{\eta}_J(c)$ and $\Big\|g_\theta(c) - \nabla\widehat{J}(c)\Big\|_2 \leq \widehat{\eta}_g(c)$; (ii) (*dominance*) $\widehat{\eta}_J \leq \eta_J$ and $\widehat{\eta}_g \leq \eta_g$: anchoring never loosens a certificate, and every gate that passes without anchors passes with them; (iii) (*persistence*) for every anchor $\mathcal{M}(u_h; c) \leq \delta_a + \varsigma(\|c - c_a\|)$; (iv) (*anchored value and gradient*) $|J(u_h, c_a) - \widehat{J}(c_a)| \leq \eta_J^a(c_a)$ and $\Big\|g_a - \nabla\widehat{J}(c_a)\Big\|_2 \leq \eta_g^a(c_a)$.*

Surrogate and anchor branches may live on different quadrature grids (Lemma 1), and a state-only anchor sharpens $\widehat{\eta}_J$ while leaving $\widehat{\eta}_g$ to the surrogate (Appendix B.3), which is exactly what a value-limited iteration needs.

### 3.2 THE CERTIFIED TRUST-REGION ITERATION

Given $c_k \in \mathcal{C}$ and a radius $\Delta_k > 0$, with acceptance thresholds $0 < \eta_1 \le \eta_2 < 1$, radius factors $0 < \gamma_s < 1 < \gamma_e$, Cauchy fraction $\beta \in (0, 1]$, radius cap $\Delta_{\max} \in [\Delta_0, \infty]$, and a tolerance $\mathrm{tol} > 0$, one iteration in the default surrogate mode follows the loop D–E–F–G of Figure 1 (the complete pseudo-code is Algorithm 1 in Appendix E.2): it refines the surrogate until the gradient gate at the center (Figure 1E)

$$\widehat{\eta}_g(c_k) \;\le\; \max\{\kappa_g \min(\chi_\theta(c_k), \Delta_k),\; \kappa_g \mathrm{tol}\} \tag{6}$$

holds; minimizes the model $m_k = J_\theta$ over $\{c_k + s \in \mathcal{C} : \|s\| \le \Delta_k\}$ (Figure 1F) with a general non-linear programming (NLP) solver, keeping the better of its point and the best point on the projected Cauchy arc $t \mapsto P_{\mathcal{C}}(c_k - t g_\theta(c_k))$, so that $\mathrm{pred}_k := m_k(c_k) - m_k(c_k + s_k) \ge \beta$ [Cauchy decrease] (Theorem 6) whatever the NLP returned; refines until the value gate at both endpoints $\widehat{\eta}_J(c_k), \widehat{\eta}_J(c_k + s_k) \le \kappa_J \mathrm{pred}_k$ holds; and accepts iff the certified ratio (Figure 1G)

$$\hat{\rho}_k := \frac{\big(J_\theta(c_k) - J_\theta(c_k + s_k)\big) - \widehat{\eta}_J(c_k) - \widehat{\eta}_J(c_k + s_k)}{\mathrm{pred}_k} \;\ge\; \eta_1, \tag{7}$$

with $\Delta_{k+1} = \min\{\gamma_e \Delta_k, \Delta_{\max}\}$ if $\hat{\rho}_k \ge \eta_2$, $\Delta_{k+1} = \Delta_k$ if $\eta_1 \le \hat{\rho}_k < \eta_2$, and $\Delta_{k+1} = \gamma_s \Delta_k$ otherwise. The run terminates with a certificate when the terminal test $\chi_\theta(c_k) + \widehat{\eta}_g(c_k) \le \mathrm{tol}$ holds (Theorem 4); every certified iterate is stored and the run returns the one with the best certified terminal bound. The gate constants satisfy (Lemma 4)

$$\kappa_g \le \beta(1 - \eta_2)/16, \qquad \kappa_J \le (1 - \eta_2)/8. \tag{8}$$

### 3.3 ANCHORED MODE: A CERTIFIED FIRST-ORDER CORRECTION

When a gradient anchor is created at $c_a = c_k$ (event $\mathrm{E}_g$ below) and the anchored terminal test $\chi_a + \eta_g^a(c_k) \le \mathrm{tol}$ fails, the algorithm enters mode A at that center: the model becomes

$$m_k^A(c) := J_\theta(c) + \langle g_a - g_\theta(c_k),\, c - c_k \rangle, \tag{9}$$

whose gradient at $c_k$ is the anchored gradient $g_a$ and whose Hessian equals that of $J_\theta$. The model criticality is $\chi_a := \|P_{\mathcal{C}}(c_k - g_a) - c_k\|$, the gradient gate tests the computable $\eta_g^a(c_k)$, and the value gate and (7) are unchanged (they use the surrogate values $J_\theta$ with the anchored certificates $\widehat{\eta}_J$). The anchor persists through rejections (no re-solve) and mode A ends with the first accepted step. Model (9) is the additive first-order correction of Alexandrov et al. (1998), with consistency certified against the PDE within $\eta_g^a(c_k)$ rather than enforced against a high-fidelity model treated as truth.

### 3.4 CERTIFICATE-BASED SOLVE POLICY

A step can be blocked by the objective certificate (the ratio (7) fails although the raw decrease was fine) or by the gradient certificate (the gate (6) or the terminal test fails although $\chi_\theta$ is small), and each has a matching solve (Figure 1G). The value event $\mathrm{E}_J$ creates a state anchor after $n_{\mathrm{ns}}$ consecutive rejections caused by objective certificate widths. The gradient event $\mathrm{E}_g$ creates a gradient anchor when refinement has not reduced the gradient-certificate deficit for $n_g$ iterations, then tests anchored termination and otherwise starts mode A. These two events, within a budget of $N_{\max}$ events, are the full solve policy (Definition 5, Algorithms 1 and 2). Once the budget is spent, the run continues without solves and returns its best certified bound.

## 4 THEORETICAL GUARANTEES

The theoretical results apply to both surrogate mode and mode A, since both modes share the convergence requirements: bounded model curvature, attainable gates, and a fraction of the projected Cauchy decrease (Definition 4). Theorems 2 to 4 require only Assumption 2 and exactly evaluated integrals (Theorem 4 also requires a convex feasible set), whereas Theorem 5 is the only one that also needs the gates, their attainability, and a bounded observed curvature (Proposition 1).

**Theorem 3** (Guaranteed descent)**.** *If step $k$ is accepted by (7) (in either mode, with any valid certificates $\widehat{\eta}_J$, and even if the network used at $c_k + s_k$ differs from the one used at $c_k$), then*

$$\widehat{J}(c_k) - \widehat{J}(c_k + s_k) \;\ge\; \eta_1 \mathrm{pred}_k \;>\; 0.$$

The true decrease is thus guaranteed but never observed: the algorithm never evaluates $\widehat{J}$.

**Theorem 4** (Certified criticality)**.** *Let $c \in \mathcal{C}$ and let $g \in \mathbb{R}^p$ be* any *vector with a certified bound $\eta \geq \left\| g - \nabla \widehat{J}(c) \right\|_2$. Then $\left| \|P_{\mathcal{C}}(c-g) - c\| - \chi(c) \right| \leq \eta$; in particular $\|P_{\mathcal{C}}(c-g) - c\| + \eta \leq \mathrm{tol}$ implies $\chi(c) \leq \mathrm{tol}$. Instances: $g = g_\theta(c)$, $\eta = \widehat{\eta}_g(c)$ (surrogate mode) and $g = g_a$, $\eta = \eta_g^a(c)$ (anchored terminal test).*

**Theorem 5** (Convergence)**.** *Let Assumption 2 and (A1)–(A3) of Assumption 1 hold, let $\mathcal{C}$ be nonempty closed convex, let the gates be enforced with constants (8), and let every step satisfy the projected Cauchy decrease (Theorem 6) for its mode's model. Then, whatever events fire and in whatever order, provided each iteration's gates are eventually satisfied ((A1) in surrogate mode; (A4) via Lemma 5 in mode A):*

*(i) with the floored gate (6) the run reaches, or returns a point with, certified criticality $\chi \leq (1+\kappa_g)\mathrm{tol}$;*

*(ii) with the unfloored gate ($\kappa_g \mathrm{tol}$ removed) $\liminf_k \chi(c_k) = 0$;*

*(iii) every accepted step decreases $\widehat{J}$ by at least $\eta_1 \mathrm{pred}_k$; after the event budget is exhausted, (i)–(ii) apply to the remaining zero-solve loop.*

The claim actually made by a run is (i): certified $\mathrm{tol}$-criticality at the returned point, verified a posteriori by Theorem 4, and no run of finite tolerance claims more. The convergence proof follows classical inexact-gradient trust-region theory (Conn et al., 2000; Heinkenschloss & Vicente, 2002; Kouri et al., 2013), with the accuracy conditions supplied by certificates instead of being postulated. Assumption (A1) is exactly the hypothesis that refinement can meet them. The curvature bound (A2) cannot be verified in advance, and therefore Proposition 1 restates Theorem 5 conditionally on the observed curvature $\widehat{\kappa}$ staying bounded along the whole run. A finite run reports its running $\widehat{\kappa}$ maximum, which shows that the hypothesis has not failed so far but says nothing about later iterations. Our experiments are limited to at most 10 iterations, so they exercise the certificate layer (Theorems 2 to 4), not the asymptotic statements of Theorem 5. For a convex reduced objective the terminal certificate also yields a certified distance to the optimizer (Theorem 7), the "certified gap" of Table 9.

## 5 Experiments

### 5.1 Setup

We run SA-CPITR on five cases with explicit constants and reference solutions (Appendix D, Table 5): linear elliptic (C1), parabolic (C2), semilinear (C3), wave (C4), steady Burgers (C5) and, as an extension, elliptic with a pointwise state constraint (C6), with 1–16 control coefficients. A $\tanh$ MLP with four hidden layers of width 64–128 takes $(x, c)$ as input, is trained offline on the weighted state and tangent residual loss (31) and refined when a gate fails. The trust-region subproblem is solved by an interior-point NLP with the projected Cauchy safeguard, and the parameters used are $\mathrm{tol} = 10^{-4}$, $\eta_1 = 0.1$, $\eta_2 = 0.75$, $\beta = 1$, $N_{\max} = 5$, $\vartheta = \frac{1}{2}$, $n_{\mathrm{ns}} = 2$, $n_g = 2$, iteration cap 10. Anchors come from classical discretizations (modal, Galerkin or spline solvers) returning conforming reconstructions, so the same majorant certifies them. Solver cost is counted in anchor events and in PDE solves: one for a state anchor, $1 + p$ for a gradient anchor (two with the unimplemented adjoint variant; Remark 2 and Appendix F.4). Both budgets run the same code: $N_{\max} = 0$ is CPITR, the certified solver-free variant of the method and itself a contribution of this paper. External comparisons use a frozen PINN surrogate (trained once and minimized as is) and the solver-in-the-loop optimizer that produced the reference optima.

**Implementation and verification.** The implementation follows the theory with three simplifications made for practicality: (i) every integral (the majorants, the objective and its gradient) is evaluated on a fixed quadrature grid without a bound on the quadrature error (Section 3.1); (ii) refinement budgets are finite, and a step may be computed after a budget is exhausted even when a gate fails; and (iii) C5's mean-zero constraint also uses quadrature-evaluated coefficients (Appendix F.6). A run is reported as certified when the terminal test of Theorem 4 holds with the certificates evaluated

Table 1: **Summarized results** (seed 0, $\mathrm{tol} = 10^{-4}$, cap 10). crit.: terminal criticality (anch.: gradient anchor); feas.: C6 feasibility; other statuses: stopping rules. bound: $\chi_\theta + \widehat{\eta}_g$ or $\chi_a + \eta_g^a$; C6: optimality gap verified offline, not certified by the method. diag.: diagnostic at the frozen surrogate's point. PDE: solves during optimization, excluding later verification anchors; wall clock: Appendix F.4.

| | frozen | CPITR (solver-free, $N_{\max} = 0$) | | SA-CPITR (budget 5) | | | |
|---|---|---|---|---|---|---|---|
| | diag. | status | bound | status | bound | event | PDE |
| C1 | $2.78 \cdot 10^{-5}$ | crit. | $4.16 \cdot 10^{-5}$ | crit. | $4.16 \cdot 10^{-5}$ | — | 0 |
| C2 | $2.72 \cdot 10^{-2}$ | gap target | $1.77 \cdot 10^{-4}$ | crit. (anch.) | $1.39 \cdot 10^{-6}$ | grad. | 3 |
| C3 | $9.49 \cdot 10^{-4}$ | crit. | $3.44 \cdot 10^{-5}$ | crit. | $3.44 \cdot 10^{-5}$ | — | 0 |
| C4 | $8.72 \cdot 10^{-3}$ | max iters | $1.42 \cdot 10^{-4}$ | crit. (anch.) | $3.62 \cdot 10^{-5}$ | grad. | 13 |
| C5 | $1.54 \cdot 10^{-2}$ | futile | $2.27 \cdot 10^{-4}$ | crit. (anch.) | $5.04 \cdot 10^{-5}$ | grad. | 17 |
| C6 (ext.) | $3.24 \cdot 10^{-4}$ | feas. | $5.43 \cdot 10^{-6}$ | feas. | $5.73 \cdot 10^{-6}$ | state | 1 |
| diag. $\leq$ tol on C1–C5: $1/5$ | | crit. $2/5$, 0 solves | | crit. $5/5$, 3 events (33 PDE solves) | | | |

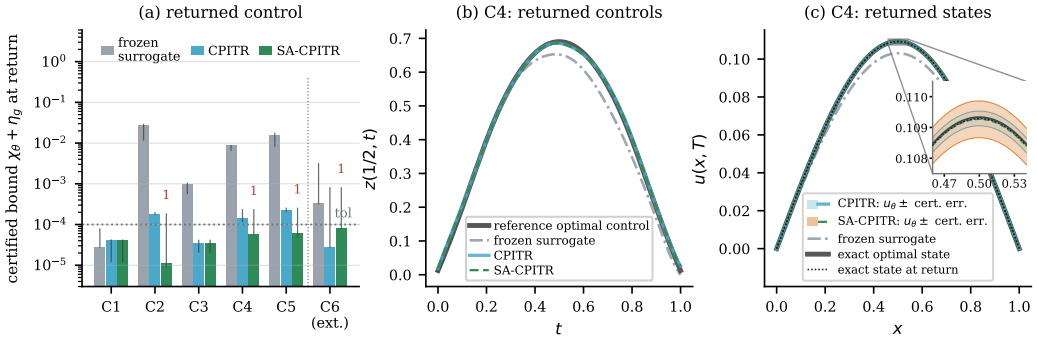



Figure 2: **Results comparison.** (a) Certified bound at the returned control (bars seed 0, whiskers five seeds, red digits anchor events); C6 (ext.): diagnostic on its inner set. (b), (c) C4: returned controls and states against the exact optimum, with the certified error bands of the network states (Appendix F.5).

in this way inside the loop. The certificate at every returned point is then re-verified with exact integration or continuum-valid enclosures (Section 5.2), whereas the in-loop layer (gates, acceptance ratio, descent) is grid-evaluated.

### 5.2 Results

SA-CPITR certifies terminal criticality on all of C1–C5 with three anchor events and 33 PDE solves in total (Table 1). CPITR certifies C1 and C3 at zero solves and coincides with SA-CPITR where the latter spends nothing. On each remaining case, C2, C4 and C5, where it does not certify, a single gradient anchor establishes the SA-CPITR certificate. On C6, both methods return certified feasible points (Theorem 15, sharp margin; Appendix D.8), and every accepted step decreases the true objective, but no criticality claim is made. The single state event SA-CPITR spends on C6 (1 PDE solve) does not change this outcome. Against the exact solution map, the returned objectives are within $5.7 \cdot 10^{-6}$ (SA-CPITR) and $5.4 \cdot 10^{-6}$ (CPITR) of the optimum, whereas the feasible control of the frozen surrogate is $3.7 \cdot 10^{-4}$ above it; this optimality gap is verified offline and is not a certificate of the method (Appendix D.8).

**Baselines and ablation.** A surrogate trained once, frozen and minimized by a plain gradient method stops where its own criticality is between $4 \cdot 10^{-13}$ and $3 \cdot 10^{-9}$ on C1–C5, but its certified criticality bound exceeds tol by a factor of about 10 to 270 on four of the five cases (Figure 2a). On C4, its control lies 0.11 from the reference optimum and its maximum state error is $6 \cdot 10^{-3}$,

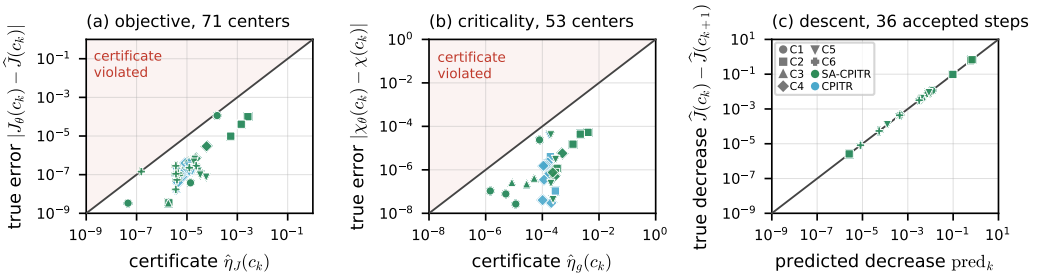



Figure 3: **Certificates against the truth** at every logged iterate.

compared with $3 \cdot 10^{-2}$ and $4 \cdot 10^{-5}$, respectively, for both CPITR and SA-CPITR (Figure 2b,c). The solver-in-the-loop optimizer that computed the reference optima uses 16–4641 PDE solves per case (Appendix F.4). In a 240-run ablation over seeds, widths, training budgets and surrogate quality (Appendix G), adding solves never cost a certificate: across 74 paired configurations on C1–C5 (64 unique), SA-CPITR certifies 53 and CPITR 26. Overall, 27 pairs are decided in favour of SA-CPITR and none against, at 0.72 of the solver-free wall clock and with at most two anchor events per run. The observed curvature of Proposition 1 stayed bounded in all runs (Appendix F.7).

**Certificate validity against ground truth.** The certificates are evaluated on a quadrature grid inside the loop, so we separate the returned points, whose certificates are re-verified exactly, from the in-loop layer, which is audited against ground truth. Exact integration reproduces the anchored certificates of C2, C4 and C5 to five digits, and additional anchors placed after the run at the returned controls of C1 and C3 give exact certificates of $4.00 \cdot 10^{-5}$ and $2.63 \cdot 10^{-5}$ (Appendix F.6); these post-run solves are excluded from Table 1. On C3 a continuum-valid evaluation of the network certificate by cellwise interval enclosures gives $3.46 \cdot 10^{-5}$ versus the grid value $3.44 \cdot 10^{-5}$. These validation tests support the five SA-CPITR terminal claims independently of the in-loop quadrature. The grid-based audit (Figure 3 and Appendix F.5) found valid objective bounds at all 71 logged centers, with at least 7% slack, and valid criticality transfer at all 53 C1–C5 centers. Also, every accepted step decreased the true objective by at least $\eta_1 \mathrm{pred}_k$, and the returned points of all ten C1–C5 arms were tol-critical. Hence the gates, the acceptance ratio and the per-step descent guarantee of the network branch are grid-evaluated and audited, but not proved. In addition, Appendix H shows out-of-distribution tests of the method.

## 6 Conclusion

SA-CPITR uses a posteriori majorants to determine when a learned PDE surrogate supports an optimization decision and when a solver anchor is needed. The same bounds certify network and solver outputs. The experiments show that a small number of anchor events can establish terminal certificates that the solver-free method does not obtain.

**Limitations and future work.** The experiments evaluate integrals by quadrature without bounds on its error and use finite refinement budgets, so they validate the theory empirically. Only C3 has a continuum-verified network terminal certificate; the other terminal claims are verified with solver anchors (Section 5.2). A fully verified implementation requires continuum-valid evaluation of both the network and the solver branch inside the optimization loop, and a non-resetting allowance for gate-failing steps, after which refinement continues until the gates hold. Gradient anchors currently cost $1+p$ PDE solves; an adjoint-based certificate (Remark 2) would reduce this to one state and one adjoint solve. Finally, the test problems are small by design, but the theory never uses the dimension of $c$, so the next step is infinite-dimensional controls with neural-operator surrogates.

## Reproducibility Statement

The anonymized code is provided as supplementary material. It contains the implementation of SA-CPITR and of its solver-free variant CPITR, the six test problems, the scripts that define every

run of the paper, the logged records of all runs, and scripts that regenerate all result figures and tables from these records on a CPU. The problem data, constants and reference solutions are given in Appendix D, the full algorithm and its implementation in Appendix E, and the experimental protocol and hyperparameters in Appendix F; all proofs are in Appendices B and C.

## AI USE STATEMENT

In this work, we used generative AI tools to formulate mathematical claims, to provide ingredients for their proofs and to assist in writing the proofs, to implement the method, to provide feedback on the design of the experiments, and to help interpret the results. We have not used generative AI tools to develop the conceptual framework, to propose the research hypotheses or for translation, and the generation of synthetic data sets, the cleaning and reformatting of data sets and qualitative or thematic data analysis are not applicable to this work. Additionally, we used generative AI tools to create and edit software code and figures, to draft parts of the paper and edit it for readability, to search and summarize the literature, and to format references. We have reviewed all AI-assisted work: the authors checked all theorems and proofs, tested the implementation against classical reference solvers and exact solutions, regenerated every reported number from the logged runs, and checked every reference against its published version. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

## REFERENCES

Mark Ainsworth and Justin Dong. Galerkin neural networks: A framework for approximating variational equations with error control. *SIAM Journal on Scientific Computing*, 43(4):A2474–A2501, 2021. doi: 10.1137/20M1366587.

Natalia M. Alexandrov, John E. Dennis, Jr., Robert Michael Lewis, and Virginia Torczon. A trust-region framework for managing the use of approximation models in optimization. *Structural Optimization*, 15(1):16–23, 1998. doi: 10.1007/BF01197433.

Ross Anderson, Joey Huchette, Will Ma, Christian Tjandraatmadja, and Juan Pablo Vielma. Strong mixed-integer programming formulations for trained neural networks. *Mathematical Programming*, 183(1–2):3–39, 2020. doi: 10.1007/s10107-020-01474-5.

Eyal Arian, Marco Fahl, and Ekkehard W. Sachs. Trust-region proper orthogonal decomposition for flow control. ICASE Report 2000-25, Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA, 2000. URL `https://apps.dtic.mil/sti/pdfs/ADA377382.pdf`. NASA/CR-2000-210124.

Eduard Bader, Mark Kärcher, Martin A. Grepl, and Karen Veroy. Certified reduced basis methods for parametrized distributed elliptic optimal control problems with control constraints. *SIAM Journal on Scientific Computing*, 38(6):A3921–A3946, 2016. doi: 10.1137/16M1059898.

Afonso S. Bandeira, Katya Scheinberg, and Luís N. Vicente. Convergence of trust-region methods based on probabilistic models. *SIAM Journal on Optimization*, 24(3):1238–1264, 2014. doi: 10.1137/130915984.

Stefan Banholzer, Tim Keil, Luca Mechelli, Mario Ohlberger, Felix Schindler, and Stefan Volkwein. An adaptive projected Newton non-conforming dual approach for trust-region reduced basis approximation of PDE-constrained parameter optimization. *Pure and Applied Functional Analysis*, 7(5):1561–1596, 2022. URL `https://arxiv.org/abs/2012.11653`.

Robert J. Baraldi and Drew P. Kouri. A proximal trust-region method for nonsmooth optimization with inexact function and gradient evaluations. *Mathematical Programming*, 201(1–2):559–598, 2023. doi: 10.1007/s10107-022-01915-3.

Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T. Patera. An "empirical interpolation" method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathématique*, 339(9):667–672, 2004. doi: 10.1016/j.crma.2004.08.006.

Jostein Barry-Straume, Arash Sarshar, Andrey A. Popov, and Adrian Sandu. Physics-informed neural networks for PDE-constrained optimization and control. *Communications on Applied Mathematics and Computation*, 8(4):1283–1306, 2026. doi: 10.1007/s42967-025-00499-x. Published online 2025.

Roland Becker, Hartmut Kapp, and Rolf Rannacher. Adaptive finite element methods for optimal control of partial differential equations: Basic concept. *SIAM Journal on Control and Optimization*, 39(1):113–132, 2000. doi: 10.1137/S0363012999351097.

Stefano Berrone, Claudio Canuto, and Moreno Pintore. Solving PDEs by variational physics-informed neural networks: an a posteriori error analysis. *Annali dell'Università di Ferrara*, 68 (2):575–595, 2022. doi: 10.1007/s11565-022-00441-6.

Jose Blanchet, Coralia Cartis, Matt Menickelly, and Katya Scheinberg. Convergence rate analysis of a stochastic trust-region method via supermartingales. *INFORMS Journal on Optimization*, 1 (2):92–119, 2019. doi: 10.1287/ijoo.2019.0016.

David Brookes, Hahnbeom Park, and Jennifer Listgarten. Conditioning by adaptive sampling for robust design. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 773–782. PMLR, 2019. URL `https://proceedings.mlr.press/v97/brookes19a.html`.

Paul H. Calamai and Jorge J. Moré. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39(1):93–116, 1987. doi: 10.1007/BF02592073.

Claudio Canuto, Timo Tonn, and Karsten Urban. A posteriori error analysis of the reduced basis method for nonaffine parametrized nonlinear PDEs. *SIAM Journal on Numerical Analysis*, 47(3): 2001–2022, 2009. doi: 10.1137/080724812.

Lianghao Cao, Thomas O'Leary-Roseberry, Prashant K. Jha, J. Tinsley Oden, and Omar Ghattas. Residual-based error correction for neural operator accelerated infinite-dimensional Bayesian inverse problems. *Journal of Computational Physics*, 486:112104, 2023. doi: 10.1016/j.jcp.2023.112104.

Liyuan Cao, Albert S. Berahas, and Katya Scheinberg. First- and second-order high probability complexity bounds for trust-region methods with noisy oracles. *Mathematical Programming*, 207 (1–2):55–106, 2024. doi: 10.1007/s10107-023-01999-5.

Richard G. Carter. On the global convergence of trust region algorithms using inexact gradient information. *SIAM Journal on Numerical Analysis*, 28(1):251–265, 1991. doi: 10.1137/0728014.

Eduardo Casas. Control of an elliptic problem with pointwise state constraints. *SIAM Journal on Control and Optimization*, 24(6):1309–1318, 1986. doi: 10.1137/0324078.

Eduardo Casas and Fredi Tröltzsch. Second order optimality conditions and their role in PDE control. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 117(1):3–44, 2015. doi: 10.1365/s13291-014-0109-3.

Thierry Cazenave and Alain Haraux. *An Introduction to Semilinear Evolution Equations*, volume 13 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 1998. doi: 10.1093/oso/9780198502777.001.0001. Translated by Yvan Martel.

Francesco Ceccon, Jordan Jalving, Joshua Haddad, Alexander Thebelt, Calvin Tsay, Carl D. Laird, and Ruth Misener. OMLT: Optimization & machine learning toolkit. *Journal of Machine Learning Research*, 23(349):1–8, 2022. URL `https://jmlr.org/papers/v23/22-0277.html`.

Ruobing Chen, Matt Menickelly, and Katya Scheinberg. Stochastic optimization using a trust-region method and random models. *Mathematical Programming*, 169(2):447–487, 2018. doi: 10.1007/s10107-017-1141-8.

Xiaoran Cheng and Sen Na. Physics-informed neural networks with trust-region sequential quadratic programming. *arXiv preprint arXiv:2409.10777*, 2024.

Ze Cheng, Zhuoyu Li, Xiaoqiang Wang, Jianing Huang, Zhizhou Zhang, Zhongkai Hao, and Hang Su. Accelerating PDE-constrained optimization by the derivative of neural operators. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 10090–10105. PMLR, 2025. URL https://proceedings.mlr.press/v267/cheng25f.html.

Andrew R. Conn, Nicholas I. M. Gould, and Philippe L. Toint. *Trust-Region Methods*, volume 1 of *MPS-SIAM Series on Optimization*. SIAM, Philadelphia, 2000. doi: 10.1137/1.9780898719857.

Wojciech M. Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Swirszcz, and Razvan Pascanu. Sobolev training for neural networks. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/758a06618c69880a6cee5314ee42d52f-Abstract.html.

Tim De Ryck and Siddhartha Mishra. Numerical analysis of physics-informed neural networks and related models in physics-informed machine learning. *Acta Numerica*, 33:633–713, 2024. doi: 10.1017/S0962492923000089.

Nicola Demo, Maria Strazzullo, and Gianluigi Rozza. An extended physics informed neural network for preliminary analysis of parametric optimal control problems. *Computers & Mathematics with Applications*, 143:383–396, 2023. doi: 10.1016/j.camwa.2023.05.004.

Asen L. Dontchev and R. Tyrrell Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2nd edition, 2014. doi: 10.1007/978-1-4939-1037-3.

Francisco Eiras, Adel Bibi, Rudy R. Bunel, Krishnamurthy Dj Dvijotham, Philip Torr, and M. Pawan Kumar. Efficient error certification for physics-informed neural networks. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12318–12347. PMLR, 2024. URL https://proceedings.mlr.press/v235/eiras24a.html.

Lewin Ernst, Nikolaos Rekatsinas, and Karsten Urban. A posteriori certification for physics-informed neural networks. *arXiv preprint arXiv:2502.20336v1*, 2025.

Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2nd edition, 2010. doi: 10.1090/gsm/019.

Marco Fahl and Ekkehard W. Sachs. Reduced order modelling approaches to PDE-constrained optimization based on proper orthogonal decomposition. In Lorenz T. Biegler, Matthias Heinkenschloss, Omar Ghattas, and Bart van Bloemen Waanders (eds.), *Large-Scale PDE-Constrained Optimization*, volume 30 of *Lecture Notes in Computational Science and Engineering*, pp. 268–280. Springer, Berlin, Heidelberg, 2003. doi: 10.1007/978-3-642-55508-4_16.

Vladimir Fanaskov, Tianchi Yu, Alexander Rudikov, and Ivan Oseledets. Astral: training physics-informed neural networks with error majorants. *arXiv preprint arXiv:2406.02645*, 2024. ICLR 2026 Workshop on AI and Partial Differential Equations.

Clara Fannjiang and Jennifer Listgarten. Autofocused oracles for model-based design. In *Advances in Neural Information Processing Systems*, volume 33, pp. 12945–12956, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/972cda1e62b72640cb7ac702714a115f-Abstract.html.

Matteo Fischetti and Jason Jo. Deep neural networks and mixed integer linear optimization. *Constraints*, 23(3):296–309, 2018. doi: 10.1007/s10601-018-9285-6.

Alexander I. J. Forrester and Andy J. Keane. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, 45(1-3):50–79, 2009. doi: 10.1016/j.paerosci.2008.11.001.

Serge Gratton, Annick Sartenaer, and Philippe L. Toint. Recursive trust-region methods for multiscale nonlinear optimization. *SIAM Journal on Optimization*, 19(1):414–444, 2008. doi: 10.1137/050623012.

Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains*, volume 69 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 2011. doi: 10.1137/1.9781611972030. Reprint of the 1985 original.

Johannes Gründler, Moritz Maibaum, and Philipp Petersen. Certified and accurate computation of function space norms of deep neural networks. *arXiv preprint arXiv:2603.06431*, 2026.

Bernard Haasdonk, Hendrik Kleikamp, Mario Ohlberger, Felix Schindler, and Tizian Wenzel. A new certified hierarchical and adaptive RB-ML-ROM surrogate model for parametrized PDEs. *SIAM Journal on Scientific Computing*, 45(3):A1039–A1065, 2023. doi: 10.1137/22M1493318.

Zhongkai Hao, Chengyang Ying, Hang Su, Jun Zhu, Jian Song, and Ze Cheng. Bi-level physics-informed neural networks for PDE constrained optimization using Broyden's hypergradients. In *International Conference on Learning Representations (ICLR)*, 2023. URL https://openreview.net/forum?id=kkpL4zUXtiw.

Emil Haugen, Alexei Stepanenko, and Anders C. Hansen. Trustworthy AI in numerics: On verification algorithms for neural network-based PDE solvers. *arXiv preprint arXiv:2509.26122*, 2025.

Matthias Heinkenschloss and Luís N. Vicente. Analysis of inexact trust-region SQP algorithms. *SIAM Journal on Optimization*, 12(2):283–302, 2002. doi: 10.1137/S1052623499361543.

Birgit Hillebrecht and Benjamin Unger. Certified machine learning: A posteriori error estimation for physics-informed neural networks. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2022. doi: 10.1109/IJCNN55064.2022.9892569.

Birgit Hillebrecht and Benjamin Unger. Rigorous a posteriori error bounds for PDE-defined PINNs. *IEEE Transactions on Neural Networks and Learning Systems*, 36(1):1583–1593, 2025. doi: 10.1109/TNNLS.2023.3335837.

M. Hintermüller and R. H. W. Hoppe. Goal-oriented adaptivity in control constrained optimal control of partial differential equations. *SIAM Journal on Control and Optimization*, 47(4):1721–1743, 2008. doi: 10.1137/070683891.

Michael Hintermüller and Karl Kunisch. Feasible and noninterior path-following in constrained minimization with low multiplier regularity. *SIAM Journal on Control and Optimization*, 45(4):1198–1221, 2006. doi: 10.1137/050637480.

Michael Hinze, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, Dordrecht, 2009. doi: 10.1007/978-1-4020-8839-1.

Philipp Holl, Vladlen Koltun, and Nils Thuerey. Learning to control PDEs with differentiable physics. In *International Conference on Learning Representations (ICLR)*, 2020. URL https://openreview.net/forum?id=HyeSin4FPB.

Jun-Ting Hsieh, Shengjia Zhao, Stephan Eismann, Lucia Mirabella, and Stefano Ermon. Learning neural PDE solvers with convergence guarantees. In *International Conference on Learning Representations (ICLR)*, 2019. URL https://openreview.net/forum?id=rklaWn0qK7.

Jianguo Huang, Haoqin Wang, and Haizhao Yang. Int-Deep: A deep learning initialized iterative method for nonlinear problems. *Journal of Computational Physics*, 419:109675, 2020. doi: 10.1016/j.jcp.2020.109675.

Rakhoon Hwang, Jae Yong Lee, Jin Young Shin, and Hyung Ju Hwang. Solving PDE-constrained control problems using operator learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):4504–4512, 2022. doi: 10.1609/aaai.v36i4.20373.

Prashant K. Jha. Residual-based error corrector operator to enhance accuracy and reliability of neural operator surrogates of nonlinear variational boundary-value problems. *Computer Methods in Applied Mechanics and Engineering*, 419:116595, 2024. doi: 10.1016/j.cma.2023.116595.

Mark Kärcher and Martin A. Grepl. A certified reduced basis method for parametrized elliptic optimal control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 20(2):416–441, 2014. doi: 10.1051/cocv/2013069.

Mark Kärcher, Zoi Tokoutsi, Martin A. Grepl, and Karen Veroy. Certified reduced basis methods for parametrized elliptic optimal control problems with distributed controls. *Journal of Scientific Computing*, 75(1):276–307, 2018. doi: 10.1007/s10915-017-0539-z.

Michael Kartmann, Tim Keil, Mario Ohlberger, Stefan Volkwein, and Barbara Kaltenbacher. Adaptive reduced basis trust region methods for parameter identification problems. *Computational Science and Engineering*, 1(1):3, 2024. doi: 10.1007/s44207-024-00002-z.

Tim Keil and Mario Ohlberger. A relaxed localized trust-region reduced basis approach for optimization of multiscale problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 58 (1):79–105, 2024. doi: 10.1051/m2an/2023089.

Tim Keil, Luca Mechelli, Mario Ohlberger, Felix Schindler, and Stefan Volkwein. A non-conforming dual approach for adaptive trust-region reduced basis approximation of PDE-constrained parameter optimization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 55(3):1239–1269, 2021. doi: 10.1051/m2an/2021019.

Tim Keil, Hendrik Kleikamp, Rolf J. Lorentzen, Micheal B. Oguntola, and Mario Ohlberger. Adaptive machine learning-based surrogate modeling to accelerate PDE-constrained optimization in enhanced oil recovery. *Advances in Computational Mathematics*, 48(6):73, 2022. doi: 10.1007/s10444-022-09981-z.

Dmitrii Kochkov, Jamie A. Smith, Ayya Alieva, Qing Wang, Michael P. Brenner, and Stephan Hoyer. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences*, 118(21):e2101784118, 2021. doi: 10.1073/pnas.2101784118.

Drew P. Kouri and Denis Ridzal. Inexact trust-region methods for PDE-constrained optimization. In Harbir Antil, Drew P. Kouri, Martin-D. Lacasse, and Denis Ridzal (eds.), *Frontiers in PDE-Constrained Optimization*, volume 163 of *The IMA Volumes in Mathematics and its Applications*, pp. 83–121. Springer, New York, NY, 2018. doi: 10.1007/978-1-4939-8636-1_3.

Drew P. Kouri, Matthias Heinkenschloss, Denis Ridzal, and Bart G. van Bloemen Waanders. A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. *SIAM Journal on Scientific Computing*, 35(4):A1847–A1879, 2013. doi: 10.1137/120892362.

Drew P. Kouri, Matthias Heinkenschloss, Denis Ridzal, and Bart G. van Bloemen Waanders. Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty. *SIAM Journal on Scientific Computing*, 36(6):A3011–A3029, 2014. doi: 10.1137/140955665.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89):1–97, 2023. URL https://jmlr.org/papers/v24/21-1524.html.

Aditi S. Krishnapriyan, Amir Gholami, Shandian Zhe, Robert M. Kirby, and Michael W. Mahoney. Characterizing possible failure modes in physics-informed neural networks. In *Advances in Neural Information Processing Systems*, volume 34, pp. 26548–26560, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/df438e5206f31600e6ae4af72f2725f1-Abstract.html.

Aviral Kumar and Sergey Levine. Model inversion networks for model-based optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 5126–5137, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/373e4c5d8edfa8b74fd4b6791d0cf6dc-Abstract.html.

Samuel Lanthaler, Siddhartha Mishra, and George E. Karniadakis. Error estimates for DeepONets: a deep learning framework in infinite dimensions. *Transactions of Mathematics and Its Applications*, 6(1):tnac001, 2022. doi: 10.1093/imatrm/tnac001.

Jacques-Louis Lions and Enrico Magenes. *Non-Homogeneous Boundary Value Problems and Applications, Vol. I*, volume 181 of *Die Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, 1972. doi: 10.1007/978-3-642-65161-8.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021a. doi: 10.1038/s42256-021-00302-5.

Lu Lu, Raphaël Pestourie, Wenjie Yao, Zhicheng Wang, Francesc Verdugo, and Steven G. Johnson. Physics-informed neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing*, 43(6):B1105–B1132, 2021b. doi: 10.1137/21M1397908.

Dingcheng Luo, Thomas O'Leary-Roseberry, Peng Chen, and Omar Ghattas. Efficient PDE-constrained optimization under high-dimensional uncertainty using derivative-informed neural operators. *SIAM Journal on Scientific Computing*, 47(4):C899–C931, 2025. doi: 10.1137/23M157956X.

Andrew March and Karen Willcox. Provably convergent multifidelity optimization algorithm not requiring high-fidelity derivatives. *AIAA Journal*, 50(5):1079–1089, 2012. doi: 10.2514/1.J051125.

Stefano Markidis. The old and the new: Can physics-informed deep-learning replace traditional linear solvers? *Frontiers in Big Data*, 4:669097, 2021. doi: 10.3389/fdata.2021.669097.

Dominik Meidner and Boris Vexler. Adaptive space-time finite element methods for parabolic optimization problems. *SIAM Journal on Control and Optimization*, 46(1):116–142, 2007. doi: 10.1137/060648994.

P. Minakowski and T. Richter. A priori and a posteriori error estimates for the Deep Ritz method applied to the Laplace and Stokes problem. *Journal of Computational and Applied Mathematics*, 421:114845, 2023. doi: 10.1016/j.cam.2022.114845.

Siddhartha Mishra and Roberto Molinaro. Estimates on the generalization error of physics-informed neural networks for approximating PDEs. *IMA Journal of Numerical Analysis*, 43(1):1–43, 2023. doi: 10.1093/imanum/drab093.

Saviz Mowlavi and Saleh Nabi. Optimal control of PDEs using physics-informed neural networks. *Journal of Computational Physics*, 473:111731, 2023. doi: 10.1016/j.jcp.2022.111731.

A. V. Muzalevskiy and S. I. Repin. A posteriori error control of approximate solutions to boundary value problems found by neural networks. *Journal of Mathematical Sciences*, 273(4):492–510, 2023. doi: 10.1007/s10958-023-06516-9.

Thomas O'Leary-Roseberry, Peng Chen, Umberto Villa, and Omar Ghattas. Derivative-informed neural operator: An efficient framework for high-dimensional parametric derivative learning. *Journal of Computational Physics*, 496:112555, 2024. doi: 10.1016/j.jcp.2023.112555.

A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*, volume 44 of *Applied Mathematical Sciences*. Springer, New York, 1983. doi: 10.1007/978-1-4612-5561-1.

Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018. doi: 10.1137/16M1082469.

Elizabeth Qian, Martin Grepl, Karen Veroy, and Karen Willcox. A certified trust region reduced basis approach to PDE-constrained optimization. *SIAM Journal on Scientific Computing*, 39(5):S434–S460, 2017. doi: 10.1137/16M1081981.

Maziar Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

Sergey Repin. A posteriori error estimation for variational problems with uniformly convex functionals. *Mathematics of Computation*, 69(230):481–500, 2000. doi: 10.1090/S0025-5718-99-01190-4.

Sergey Repin. *A Posteriori Estimates for Partial Differential Equations*, volume 4 of *Radon Series on Computational and Applied Mathematics*. De Gruyter, Berlin, 2008. doi: 10.1515/9783110203042.

Stephen M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5(1):43–62, 1980. doi: 10.1287/moor.5.1.43.

T. D. Robinson, M. S. Eldred, K. E. Willcox, and R. Haimes. Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping. *AIAA Journal*, 46(11):2814–2822, 2008. doi: 10.2514/1.36043.

Julian Roth, Max Schröder, and Thomas Wick. Neural network guided adjoint computations in dual weighted residual error estimation. *SN Applied Sciences*, 4(2):62, 2022. doi: 10.1007/s42452-022-04938-9.

Artur M. Schweidtmann and Alexander Mitsos. Deterministic global optimization with artificial neural networks embedded. *Journal of Optimization Theory and Applications*, 180(3):925–948, 2019. doi: 10.1007/s10957-018-1396-0.

Yeonjong Shin, Jérôme Darbon, and George Em Karniadakis. On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type PDEs. *Communications in Computational Physics*, 28(5):2042–2074, 2020. doi: 10.4208/cicp.OA-2020-0193.

Khemraj Shukla, Vivek Oommen, Ahmad Peyvan, Michael Penwarden, Nicholas Plewacki, Luis Bravo, Anindya Ghoshal, Robert M. Kirby, and George Em Karniadakis. Deep neural operators as accurate surrogates for shape optimization. *Engineering Applications of Artificial Intelligence*, 129:107615, 2024. doi: 10.1016/j.engappai.2023.107615.

Hwijae Son, Jin Woo Jang, Woo Jin Han, and Hyung Ju Hwang. Sobolev training for physics-informed neural networks. *Communications in Mathematical Sciences*, 21(6):1679–1705, 2023. doi: 10.4310/CMS.2023.v21.n6.a11.

Yongcun Song, Xiaoming Yuan, and Hangrui Yue. The ADMM-PINNs algorithmic framework for nonsmooth PDE-constrained optimization: A deep learning approach. *SIAM Journal on Scientific Computing*, 46(6):C659–C687, 2024. doi: 10.1137/23M1566935.

Shigeng Sun and Jorge Nocedal. A trust region method for noisy unconstrained optimization. *Mathematical Programming*, 202(1–2):445–472, 2023. doi: 10.1007/s10107-023-01941-9.

Roger Temam. *Navier–Stokes Equations: Theory and Numerical Analysis*, volume 2 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1977.

Brandon Trabucco, Aviral Kumar, Xinyang Geng, and Sergey Levine. Conservative objective models for effective offline model-based optimization. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 10358–10368. PMLR, 2021. URL https://proceedings.mlr.press/v139/trabucco21a.html.

Brandon Trabucco, Xinyang Geng, Aviral Kumar, and Sergey Levine. Design-bench: Benchmarks for data-driven offline model-based optimization. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 21658–21676. PMLR, 2022. URL https://proceedings.mlr.press/v162/trabucco22a.html.

Fredi Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. doi: 10.1090/gsm/112.

Kiwon Um, Robert Brand, Yun (Raymond) Fei, Philipp Holl, and Nils Thuerey. Solver-in-the-loop: Learning from differentiable physics to interact with iterative PDE-solvers. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6111–6122, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/43e4e6a6f341e00671e123714de019a8-Abstract.html.

Karen Veroy and Anthony T. Patera. Certified real-time solution of the parametrized steady incompressible Navier–Stokes equations: rigorous reduced-basis a posteriori error bounds. *International Journal for Numerical Methods in Fluids*, 47(8–9):773–788, 2005. doi: 10.1002/fld.867.

Sifan Wang, Mohamed Aziz Bhouri, and Paris Perdikaris. Fast PDE-constrained optimization via self-supervised operator learning. *arXiv preprint arXiv:2110.13297*, 2021a. doi: 10.48550/arXiv.2110.13297.

Sifan Wang, Hanwen Wang, and Paris Perdikaris. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Science Advances*, 7(40):eabi8605, 2021b. doi: 10.1126/sciadv.abi8605.

Sifan Wang, Xinling Yu, and Paris Perdikaris. When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, 449:110768, 2022. doi: 10.1016/j.jcp.2021.110768.

Kaidi Xu, Zhouxing Shi, Huan Zhang, Yihan Wang, Kai-Wei Chang, Minlie Huang, Bhavya Kailkhura, Xue Lin, and Cho-Jui Hsieh. Automatic perturbation analysis for scalable certified robustness and beyond. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1129–1141, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/0cbc5671ae26f67871cb914d81ef8fc1-Abstract.html.

Kailai Xu and Eric Darve. Trust region method for coupled systems of PDE solvers and deep neural networks. *arXiv preprint arXiv:2105.07552*, 2021.

Pengfei Yin, Guangqiang Xiao, Kejun Tang, and Chao Yang. AONN: An adjoint-oriented neural network method for all-at-once solutions of parametric optimal control problems. *SIAM Journal on Scientific Computing*, 46(1):C127–C153, 2024. doi: 10.1137/22M154209X.

Jeremy Yu, Lu Lu, Xuhui Meng, and George Em Karniadakis. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems. *Computer Methods in Applied Mechanics and Engineering*, 393:114823, 2022. doi: 10.1016/j.cma.2022.114823.

Yao Yue and Karl Meerbergen. Accelerating optimization of parametric linear systems by model order reduction. *SIAM Journal on Optimization*, 23(2):1344–1370, 2013. doi: 10.1137/120869171.

Matthew J. Zahr and Charbel Farhat. Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *International Journal for Numerical Methods in Engineering*, 102(5):1111–1135, 2015. doi: 10.1002/nme.4770.

Marius Zeinhofer, Rami Masri, and Kent-André Mardal. A unified framework for the error analysis of physics-informed neural networks. *IMA Journal of Numerical Analysis*, 45(5):2988–3025, 2025. doi: 10.1093/imanum/drae081.

Enrui Zhang, Adar Kahana, Alena Kopaničáková, Eli Turkel, Rishikesh Ranade, Jay Pathak, and George Em Karniadakis. Blending neural operators and relaxation methods in PDE numerical solvers. *Nature Machine Intelligence*, 6(11):1303–1313, 2024. doi: 10.1038/s42256-024-00910-x.

J. Carsten Ziems and Stefan Ulbrich. Adaptive multilevel inexact SQP methods for PDE-constrained optimization. *SIAM Journal on Optimization*, 21(1):1–40, 2011. doi: 10.1137/080743160.

Enrique Zuazua. Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Review*, 47(2):197–243, 2005. doi: 10.1137/S0036144503432862.

## APPENDIX OVERVIEW

The appendix contains the standing assumptions, full statements and proofs of every result, the six instantiations, the implementation layer, and all experimental details. Restated results carry their original numbers.

- Appendix A: nomenclature.
- Appendix B: proofs of the certificate results of Section 3.1 (assembly, anchoring, grid transfer).
- Appendix C: standing assumptions; proofs of the convergence theory of Section 4.
- Appendix D: the five instantiations C1–C5 and the state-constrained extension C6 with explicit constants.
- Appendix E: the two solve events (Definition 5); complete pseudo-code (Algorithms 1 and 2); mapping of the two events onto the reference implementation.
- Appendix F: experimental details, the audit of the theory against ground truth, the ground-truth comparison of every case, and additional figures.
- Appendix G: the 240-run ablation study and the frozen-surrogate baseline.
- Appendix H: surrogates out of distribution, where mode A executes.
- Appendix I: related work.

## A NOMENCLATURE

Table 2: Symbols used throughout the paper.

| symbol | meaning | where |
|---|---|---|
| Problem | | |
| $\Omega$, $Q = \Omega \times (0,T)$ | spatial domain; space–time cylinder | Section 2 |
| $c \in \mathbb{R}^p$, $\mathcal{C}$ | control coefficients; nonempty closed convex feasible set | Section 2 |
| $z(c) = z_0 + \sum_j c_j \varphi_j$ | parametrized control function | Section 2 |
| $u(c)$, $w_j(c) = \partial_{c_j} u(c)$ | exact state and sensitivities | Section 2 |
| $\widehat{J}(c) = J(u(c), c)$, $\chi(c)$ | true reduced objective; true projected-gradient criticality | Section 2 |
| $P_{\mathcal{C}}$ | Euclidean projection onto $\mathcal{C}$ (1-Lipschitz) | Section 2 |
| $H$, $V \subset L^2(\Omega) \subset V^*$ | separable Hilbert space of the tracking term; Gelfand triple | Theorem 1 |
| $\Omega_{\mathrm{sc}}$ | region of the pointwise state constraint (C6) | Appendix D.8 |
| Surrogate | | |
| $u_\theta(\cdot; c)$, $w_{\theta,j}$ | parametric PINN state and its sensitivities (autodiff) | Section 2 |
| $J_\theta$, $g_\theta = \nabla J_\theta$, $\chi_\theta$ | surrogate objective, gradient, criticality | Section 2 |
| $r(v; c)$, $r_{V,j}(\omega; \cdot)$ | strong state residual; tangent residual of a candidate | Appendix D |
| Certificates | | |
| $\mathcal{M}(v; c)$, $\mathcal{M}_j(\omega; c, v, \mu)$ | state and tangent majorant functionals (Assumption 2) | Section 3.1 |
| $\mu(c)$, $\mu_j(c)$ | majorants evaluated at the surrogate | Definition 1 |
| $\mathcal{E}_J$, $\mathcal{E}_{g,j}$ | assembly maps from majorants to objective/gradient bounds | Theorem 1 |
| $\eta_J$, $\eta_g$ | assembled surrogate certificates, $\lvert J_\theta - \widehat{J}\rvert \le \eta_J$, $\left\lVert g_\theta - \nabla\widehat{J}\right\rVert \le \eta_g$ | Definition 2 |
| val | validity flag ($0 \Rightarrow$ certificate $= +\infty$) | Assumption 2 |
| $\varsigma$, $C_B$ | control-shift function and constant | Assumption 2, Table 4 |
| Anchors | | |
| $a = (c_a, u_h, \{w_{h,j}\}, \delta_a, \{\delta_{a,j}\})$ | anchor: solver output with its own majorants | Definition 3 |
| $\mathcal{A}_k$ | anchors collected up to iteration $k$ | Definition 3 |
| $\widehat{\mu}$, $\widehat{\mu}_j$, $\widehat{\eta}_J$, $\widehat{\eta}_g$ | anchored majorants and certificates (minima over branches) | Definition 3 |

continued on the next page

Table 2 (continued)

| symbol | meaning | where |
|---|---|---|
| $g_a$, $\eta_g^a$, $\eta_J^a$, $\chi_a$ | anchored gradient, its certificate, anchored value certificate, anchored criticality | Definition 3 |
| Trust region | | |
| $\Delta_k$, $s_k$, $m_k$ | radius, step, model ($J_\theta$ in surrogate mode, $m_k^A$ in mode A) | Section 3.2 |
| $\mathrm{pred}_k$, $\mathrm{ared}_k$, $\mathrm{ared}_k^-$ | predicted, true and certified decrease | Section 3.2, Theorem 3 |
| $\hat{\rho}_k$, $\rho_k^{\mathrm{raw}}$ | certified and raw acceptance ratios | (7), Definition 5 |
| $\eta_1, \eta_2, \gamma_s, \gamma_e$ | acceptance thresholds and radius factors | Section 3.2 |
| $\Delta_{\max}$ | radius cap | Section 3.2 |
| $\kappa_g$, $\kappa_J$, $\beta$ | gate constants and Cauchy fraction, (8) | Section 3.2 |
| tol | target criticality tolerance | Section 3.2 |
| $\chi_k^m$, $g_k^m$ | model criticality and gradient at $c_k$ (either mode) | Definition 4 |
| $\kappa_H$, $\widehat{\kappa}_k^{\mathrm{C}}$, $\widehat{\kappa}_k^{\mathrm{s}}$, $\widehat{\kappa}$ | model curvature bound; observed arc/step curvature; their supremum | (A2), (17) |
| $L$, $J_{\inf}$ | Lipschitz constant of $\nabla\widehat{J}$; lower bound of $\widehat{J}$ | (A3) |
| $\zeta$, $\widehat{\zeta}$ | strong-monotonicity modulus of $\nabla\widehat{J}$ (C1, C2, C4, C6) and its estimate | Theorem 7 |
| Events | | |
| $d_k$, $\tau_k$ | deficit of the gradient certificate and current demand | Definition 5 |
| $\mathrm{E}_J$, $\mathrm{E}_g$ | value-limited / gradient-limited events | Definition 5 |
| $N_{\max}$, $\vartheta$, $n_{\mathrm{ns}}$, $n_g$ | solve budget, gain factor, patience parameters | Definition 5 |
| Constants of the instantiations | | |
| $C_{\mathrm{F}}$ | sharp Friedrichs constant ($1/\pi$ on $(0,1)$, $1/(\pi\sqrt{2})$ on $(0,1)^2$) | Lemma 6 |
| $D$, $a$, $\nu$ | diffusivity (C2), wave speed (C4), viscosity (C5) | Appendix D |
| $\kappa(d)$, $d_-$, $\kappa_{\mathrm{lin}}$ | Burgers coercivity margin, verified error radius, linearized margin | Appendix D.7 |
| $\varepsilon_\infty$ | certified $L^\infty$ margin of the state constraint | Appendix D.8 |

# B Proofs and supporting results for Section 3.1

## B.1 Proof of the assembly theorem

**Theorem 1** (Assembly). *Let $H$ be a separable Hilbert space with norm $\|\cdot\|$ and let $c \in \mathcal{C}$.*

(A: tracking.) *Let $\widehat{J}(c) = \frac{1}{2}\|u(c) - u_d\|^2 + \frac{\alpha}{2}|c|^2$ with $u(c), u_d \in H$, and for a candidate pair $(v, \{\omega_j\}) \subset H$ set $J^v(c) := \frac{1}{2}\|v - u_d\|^2 + \frac{\alpha}{2}|c|^2$ and $g_j^v := \langle v - u_d, \omega_j\rangle + \alpha c_j$. If $\|u(c) - v\| \le \mu$ and $\|w_j(c) - \omega_j\| \le \mu_j$, then*

$$|J^v(c) - \widehat{J}(c)| \le \tfrac{1}{2}\big(2\|v - u_d\| + \mu\big)\mu =: \mathcal{E}_J(v;\mu), \tag{2}$$

$$\big|g_j^v - \partial_{c_j}\widehat{J}(c)\big| \le \mu\|\omega_j\| + \big(\|v - u_d\| + \mu\big)\mu_j =: \mathcal{E}_{g,j}(v, \omega_j; \mu, \mu_j), \tag{3}$$

*and $(\sum_j \mathcal{E}_{g,j}^2)^{1/2} \ge \Big\|g^v - \nabla\widehat{J}(c)\Big\|_2$. All right-hand sides are computable and nondecreasing in $\mu, \mu_j$. The statement covers stationary tracking ($H = L^2(\Omega)$) and space–time tracking ($H = L^2(0,T;L^2(\Omega))$, the Bochner space of square-integrable $L^2(\Omega)$-valued functions).*

(B: terminal tracking with velocity.) *Let $V \subset L^2(\Omega) \subset V^*$ be the energy space of a second-order-in-time problem, let the state and the sensitivities lie in $C([0,T];V) \cap C^1([0,T];L^2(\Omega))$, with $\partial_{c_j}(u(T), u_t(T)) = (w_j(T), \partial_t w_j(T))$, let $\widehat{J}(c) = \frac{1}{2}\|u(T) - u_d\|^2 + \frac{1}{2}\|u_t(T)\|^2 + \frac{\alpha}{2}|c|^2$, and let the candidate $v$ (a network or a conforming reconstruction) define $J^v$ in the same way. If $\mu^{(0)} \ge \|u(T) - v(T)\|$ and $\mu^{(1)} \ge \|u_t(T) - v_t(T)\|$, then*

$$|J^v(c) - \widehat{J}(c)| \le \mathcal{E}_J\big(v(T);\mu^{(0)}\big) + \mathcal{E}_J^0\big(v_t(T);\mu^{(1)}\big),$$

*where $\mathcal{E}_J^0$ is (2) with $u_d = 0$, and the gradient error is bounded by the sum of the two corresponding terms (3), formed with majorants for $w_j(T)$ and $\partial_t w_j(T)$.*

*Proof. (A), objective.* Let $H$ be a separable Hilbert space; all functions below are elements of $H$ (the concrete spaces of the instantiations are recalled in Remark 1 below). With $a := v - u_d$, $b := u - u_d$ and $a - b = v - u$, the polarization identity $\|a\|^2 - \|b\|^2 = \langle a + b, a - b\rangle$ gives $J^v - \widehat{J} = \frac{1}{2}\langle a + b, a - b\rangle$. By Cauchy–Schwarz and $\|a + b\| \le 2\|a\| + \|a - b\|$,

$$|J^v - \widehat{J}| \le \tfrac{1}{2}\big(2\|a\| + \|a - b\|\big)\|a - b\| \le \tfrac{1}{2}\big(2\|a\| + \mu\big)\mu,$$

using $\|a - b\| = \|v - u\| \le \mu$ and monotonicity of $t \mapsto (2\|a\| + t)t$ on $t \ge 0$. This is (2).

*(A), gradient.* The reduced objective is differentiable with $\partial_{c_j}\widehat{J} = \langle u - u_d, w_j\rangle + \alpha c_j$ (chain rule in $H$, justified for every instantiation by Proposition 2). The $\alpha c_j$ terms cancel against those of $g_j^v$; adding and subtracting $\langle u - u_d, \omega_j\rangle$,

$$g_j^v - \partial_{c_j}\widehat{J} = \langle v - u, \omega_j\rangle + \langle u - u_d, \omega_j - w_j\rangle.$$

Cauchy–Schwarz bounds this by $\mu\|\omega_j\| + \|u - u_d\|\mu_j$, and $\|u - u_d\| \le \|v - u_d\| + \mu$ yields (3), whose right-hand side involves only $v, \omega_j$ and the majorants. Summing squares gives the Euclidean bound.

*(B).* By hypothesis $u$ and the $w_j$ lie in $C([0,T]; V) \cap C^1([0,T]; L^2(\Omega))$, so $u(T), w_j(T) \in V \subset L^2(\Omega)$ and $u_t(T), \partial_t w_j(T) \in L^2(\Omega)$ are well defined (for the wave equation this is the energy regularity of Lemma 8, and the derivative identity is Proposition 2); the candidate $v$ is smooth. Apply the polarization step of (A) in $L^2(\Omega)$ separately to $\frac{1}{2}\|u(T) - u_d\|^2$ (with the bound $\mu^{(0)}$) and to $\frac{1}{2}\|u_t(T)\|^2$ (with $u_d = 0$ and the bound $\mu^{(1)}$) and sum:

$$|J^v - \widehat{J}| \le \big(\|v(T) - u_d\| + \tfrac{1}{2}\mu^{(0)}\big)\mu^{(0)} + \big(\|v_t(T)\| + \tfrac{1}{2}\mu^{(1)}\big)\mu^{(1)}; \tag{10}$$

the gradient statement is identical term by term, with $\partial_{c_j}\widehat{J} = \langle u(T) - u_d, w_j(T)\rangle + \langle u_t(T), \partial_t w_j(T)\rangle + \alpha c_j$. Monotonicity in $\mu, \mu_j$ is inspected directly in (2)–(10). Nothing in the argument uses the dimension of $H$ or of $c$. □

*Remark* 1 (The function spaces behind the templates). For space–time tracking, $H = L^2(0,T; L^2(\Omega))$ is the Bochner space of strongly measurable $L^2(\Omega)$-valued functions with $\int_0^T \|u(t)\|^2\,dt < \infty$, isometric to $L^2(Q)$; for the second-order-in-time template, $V = H_0^1(\Omega)$and $V \subset L^2(\Omega) \cong L^2(\Omega)^* \subset V^*$ with dense continuous embeddings. Separability of $H$ guarantees that the $H$-valued functions appearing in the space–time and terminal templates are strongly measurable and that the Bochner spaces are the classical ones. Time continuity is derived, not assumed: the parabolic solutions of Appendix D.4 lie in the Hilbert space $W(0,T) := \{u \in L^2(0,T;V) : u_t \in L^2(0,T;V^*)\}$, which embeds continuously into $C([0,T]; L^2(\Omega))$ and carries the integration-by-parts formula $\frac{d}{dt}\|u(t)\|^2_{L^2(\Omega)} = 2\langle u_t(t), u(t)\rangle_{V^*,V}$ (Lions & Magenes, 1972, Ch. 1, Thm. 3.1); (Evans, 2010, Thm. 5.9.3); the hyperbolic solutions are energy solutions (Lemma 8) and lie in $C([0,T]; V) \cap C^1([0,T]; L^2(\Omega))$.

### B.2 Anchored certificates: proof of Theorem 2

For completeness, the anchored tangent majorant of Definition 3 is

$$\begin{aligned}\widehat{\mu}_j(c) := \min\Big\{&\mathcal{M}_j(w_{\theta,j}(c); c, u_\theta(c), \widehat{\mu}(c)),\\ &\min_{a \in \mathcal{A}_k}\big[\mathcal{M}_j(w_{h,j}; c, u_h, \mathcal{M}(u_h; c)) + \|w_{h,j} - w_{\theta,j}(c)\|_X\big]\Big\}.\end{aligned} \tag{11}$$

**Theorem 2** (Anchored certificates)**.** *Let Assumption 2 hold and let $\mathcal{A}_k$ be any set of anchors. At every $c \in \mathcal{C}$ with* val $= 1$*: (i) (*validity*) $|J_\theta(c) - \widehat{J}(c)| \le \widehat{\eta}_J(c)$ and $\Big\|g_\theta(c) - \nabla\widehat{J}(c)\Big\|_2 \le \widehat{\eta}_g(c)$; (ii) (*dominance*) $\widehat{\eta}_J \le \eta_J$ and $\widehat{\eta}_g \le \eta_g$: anchoring never loosens a certificate, and every gate that passes without anchors passes with them; (iii) (*persistence*) for every anchor $\mathcal{M}(u_h; c) \le \delta_a + \varsigma(\|c - c_a\|)$; (iv) (*anchored value and gradient*) $|J(u_h, c_a) - \widehat{J}(c_a)| \le \eta_J^a(c_a)$ and $\Big\|g_a - \nabla\widehat{J}(c_a)\Big\|_2 \le \eta_g^a(c_a)$.*

*Proof.* (i) The first branch of (4) bounds $\|u(c) - u_\theta(c)\|_X$ by Assumption 2(i) with $v = u_\theta(c)$. For the anchor branch, the triangle inequality in $X$ and Assumption 2(i) with $v = u_h$ give

$$\|u(c) - u_\theta(c)\|_X \le \|u(c) - u_h\|_X + \|u_h - u_\theta(c)\|_X \le \mathcal{M}(u_h; c) + \|u_h - u_\theta(c)\|_X ;$$

a minimum of valid upper bounds is a valid upper bound. For the sensitivities, the first branch is Assumption 2(i) with $\omega = w_{\theta,j}(c)$, $v = u_\theta(c)$, $\mu = \widehat{\mu}(c)$, admissible because $\widehat{\mu}(c) \ge \|u(c) - u_\theta(c)\|_X$ by the state claim just proved and $\mathcal{M}_j$ is nondecreasing in $\mu$; the anchor branch is the triangle inequality $\|w_j(c) - w_{\theta,j}(c)\|_X \le \|w_j(c) - w_{h,j}\|_X + \|w_{h,j} - w_{\theta,j}(c)\|_X$ followed by Assumption 2(i) with $\omega = w_{h,j}$, $v = u_h$, $\mu = \mathcal{M}(u_h; c)$. The assembled claims are Theorem 1 with the surrogate pair and the just-validated majorants.

(ii) The first branches with $\mu = \widehat{\mu}(c) \le \mu(c)$ are respectively equal to, and by monotonicity of $\mathcal{M}_j$ no larger than, the unanchored certificates; a minimum over a superset cannot be larger; assembled dominance follows from the monotonicity in Theorem 1. Gates: left-hand sides shrink, right-hand sides are unchanged.

(iii) is Assumption 2(iii) applied to the fixed $v = u_h$ between $c_a$ and $c$.

(iv) is Theorem 1 applied to the pair $(u_h, w_{h,j})$ with the majorants $(\delta_a, \delta_{a,j})$, valid for that pair by Assumption 2(i) and Definition 3; (5) is exactly the assembled gradient expression of (3) for that pair. □

*Remark* 2 (Adjoint option). For large $p$, one adjoint solve replaces the $p$ tangent solves in (5): with $p_h$ approximating the adjoint state, the assembled gradient is $\langle p_h, \partial_{c_j} f\rangle + \alpha c_j$, certified by the same assembly argument with an adjoint majorant. The statements of this paper, and the implementation, use the tangent route.

### B.3 Certificate transfer between evaluation grids; state-only anchors

Definition 3 presumes that the surrogate branch and the anchor branch of each minimum are evaluated with the same quadrature. An implementation will in general not do this: the surrogate carries its own certification grid, the reference solver its own, and forcing either onto the other would be wasteful and, for the solver, impossible (its majorants are exact only on the grid the reconstruction was built for).

**Lemma 1** (Certificate transfer between evaluation grids)**.** *Let $(J_1, g_1)$ and $(J_2, g_2)$ be two computed pairs approximating the reduced objective and gradient at the same control $c$, with certified bounds $|J_2 - \widehat{J}(c)| \le E_J$ and $\left\|g_2 - \nabla\widehat{J}(c)\right\|_2 \le E_g$. Then $|J_1 - \widehat{J}(c)| \le E_J + |J_1 - J_2|$ and $\left\|g_1 - \nabla\widehat{J}(c)\right\|_2 \le E_g + \|g_1 - g_2\|_2$, with computable correction terms. Consequently, if $(J_1, g_1)$ carries its own certificates $(E_J^{(1)}, E_g^{(1)})$, then $\min\{E_J^{(1)}, E_J + |J_1 - J_2|\}$ and $\min\{E_g^{(1)}, E_g + \|g_1 - g_2\|_2\}$ are valid certificates for $(J_1, g_1)$, never larger than $(E_J^{(1)}, E_g^{(1)})$.*

*Proof.* Triangle inequality; a minimum of valid upper bounds is a valid upper bound. □

The reference implementation applies the lemma with $(J_1, g_1)$ the surrogate pair on its own grid and $(J_2, g_2)$ the anchored assembly on the solver's grid, which gives the anchored certificates $\widehat{\eta}_J, \widehat{\eta}_g$; the statements of Section 4 use only their validity and dominance. A state-only anchor (event $\mathrm{E}_J$) has no $w_{h,j}$, so its tangent branch is $+\infty$: it sharpens $\widehat{\eta}_J$, whose assembly does not involve $\mu_j$, and leaves $\widehat{\eta}_g$ to the surrogate, which is what a value-limited iteration needs.

## C Standing assumptions and proofs for Section 4

### C.1 Standing assumptions

**Assumption 1** (Standing hypotheses)**.** (A1) (*training oracle*) For every $c$ in the visited region and every $\tau > 0$, finitely many surrogate refinement rounds (training steps, width increases) produce a surrogate whose majorants satisfy $\mu(c) \le \tau$ and $\mu_j(c) \le \tau$ (Definition 1).

(A2) (*bounded model curvature*) The trust-region models $m_k$ of Section 3.2 satisfy $\sup_k \sup_{\|c-c_k\|\le\Delta_k} \left\|\nabla^2 m_k(c)\right\|_2 \le \kappa_H < \infty$.

(A3) (*smoothness of the truth*) $\widehat{J}$ is bounded below on $\mathcal{C}$ by $J_{\inf}$ and $\left\|\nabla\widehat{J}(y) - \nabla\widehat{J}(c)\right\| \le L\,\|y-c\|$ for all $c \in S_0 := \{c \in \mathcal{C} : \widehat{J}(c) \le \widehat{J}(c_0)\}$ and all $y \in \mathcal{C}$ with $\|y-c\| \le \Delta_{\max}$.

(A4) (*solver consistency*) For every visited $c$ and every $\tau > 0$ there is a computable solver output, a conforming pair $(u_h, \{w_{h,j}\}_{j\le p})$ in the admissible class of Assumption 2, with $\mathcal{M}(u_h; c) \le \tau$ and $\mathcal{M}_j(w_{h,j}; c, u_h, \mathcal{M}(u_h; c)) \le \tau$ for all $j$.

(A3) is proved for C1–C6 in Appendix D (for C5 under the explicit bound $\|z(c)\| < 2\nu^2/C_{\mathrm{F}}^2$ on the feasible set, which holds for the data of the experiment). (A1) is an honest hypothesis about training, not guaranteed by it; its violation is detectable at run time (Theorem 16). (A4) is the only place where discretization convergence enters; it is never used inside a certificate, only to prove that triggered solves can succeed. (A2) is different in kind: $m_k$ is (a linear correction of) the network $J_\theta$, so $\kappa_H$ is a property of the trained network on the current trust region, not of the PDE, and it changes every time the network is refined. It cannot be proved per case, which is why it is retained only as the name of the classical hypothesis: Proposition 1 restates it as a condition on quantities the algorithm computes anyway, and every statement below that invokes (A2) may equivalently invoke Proposition 1(ii), and is then conditional on $\widehat{\kappa} < \infty$.

### C.2 Guaranteed descent

**Theorem 3** (Guaranteed descent)**.** *If step $k$ is accepted by (7) (in either mode, with any valid certificates $\widehat{\eta}_J$, and even if the network used at $c_k + s_k$ differs from the one used at $c_k$), then*

$$\widehat{J}(c_k) - \widehat{J}(c_k + s_k) \;\ge\; \eta_1 \mathrm{pred}_k \;>\; 0.$$

*Proof.* Let $\theta$ and $\theta'$ be the networks used at $c_k$ and at $c_k + s_k$ (possibly equal). Validity of the certificates at the respective (point, network) pairs (Theorem 2(i); for $\mathcal{A}_k = \emptyset$, Definition 2) gives $J_\theta(c_k) \ge \widehat{J}(c_k) - \widehat{\eta}_J(c_k)$ and $J_{\theta'}(c_k + s_k) \le \widehat{J}(c_k + s_k) + \widehat{\eta}_J(c_k + s_k)$; subtracting,

$$\mathrm{ared}_k - 2\widehat{\eta}_J(c_k) - 2\widehat{\eta}_J(c_k + s_k) \;\le\; \mathrm{ared}_k^- \;\le\; \mathrm{ared}_k, \qquad \mathrm{ared}_k := \widehat{J}(c_k) - \widehat{J}(c_k + s_k), \tag{12}$$

where $\mathrm{ared}_k^- := \big(J_\theta(c_k) - J_{\theta'}(c_k + s_k)\big) - \widehat{\eta}_J(c_k) - \widehat{\eta}_J(c_k + s_k)$ is the certified numerator of (7): the right inequality is the two validity bounds, the left one their reverses. Acceptance means $\mathrm{ared}_k^- \ge \eta_1 \mathrm{pred}_k > 0$, so $\mathrm{ared}_k \ge \mathrm{ared}_k^- \ge \eta_1 \mathrm{pred}_k$. □

### C.3 Projected-path facts and Cauchy decrease

**Lemma 2** (Projected path)**.** *Let $\mathcal{C} \subseteq \mathbb{R}^p$ be nonempty, closed and convex, $c \in \mathcal{C}$, $g \in \mathbb{R}^p$, and $s(t) := P_{\mathcal{C}}(c - tg) - c$ for $t \ge 0$. Then (i) $\langle g, s(t)\rangle \le -\|s(t)\|^2 / t$ for $t > 0$; (ii) $t \mapsto \|s(t)\|$ is nondecreasing; (iii) $t \mapsto \|s(t)\| / t$ is nonincreasing on $t > 0$; (iv) $\|s(t)\| \le t\,\|g\|$ and $s(\cdot)$ is continuous.*

*Proof.* Write $p_t := P_{\mathcal{C}}(c - tg)$. The variational characterization of the projection onto a closed convex set is

$$\langle c - tg - p_t,\, y - p_t\rangle \le 0 \qquad \forall y \in \mathcal{C}. \tag{13}$$

(i) Take $y = c$: since $c - p_t = -s(t)$, $\langle -tg - s(t), -s(t)\rangle \le 0$, i.e. $t\langle g, s(t)\rangle + \|s(t)\|^2 \le 0$. (ii)–(iii) are the classical monotonicity properties of the projected gradient path (Calamai & Moré, 1987, Lemma 3.1); see also (Conn et al., 2000, §12.1). (iv) $P_{\mathcal{C}}$ is 1-Lipschitz and $P_{\mathcal{C}}(c) = c$. □

**Theorem 6** (Generalized Cauchy decrease)**.** *Let $\mathcal{C}$ be closed convex, $c_k \in \mathcal{C}$, and let $m(s) = m(0) + \langle g, s\rangle + \frac{1}{2}\langle s, Hs\rangle$ with $\|H\|_2 \le \kappa_H$. Put $\chi^m := \|P_{\mathcal{C}}(c_k - g) - c_k\|$. Then there is $t \in (0, 1]$ such that $s^{\mathrm{C}} := s(t)$ satisfies $c_k + s^{\mathrm{C}} \in \mathcal{C}$, $\left\|s^{\mathrm{C}}\right\| \le \Delta$ and*

$$m(0) - m(s^{\mathrm{C}}) \;\ge\; \tfrac{1}{2}\chi^m \min\big(\Delta, \chi^m/\kappa_H, \chi^m\big). \tag{14}$$

*Consequently any step with* $\mathrm{pred}_k \geq \beta[m(0) - m(s^{\mathrm{C}})]$, $\beta \in (0, 1]$, *satisfies* $\mathrm{pred}_k \geq \frac{\beta}{2}\chi^m \min(\Delta, \chi^m/\kappa_H, \chi^m)$. *The same conclusion holds for a* $C^{1,1}$ *model* $m$ *whose gradient is* $\kappa_H$*-Lipschitz on the ball, by Taylor's theorem with integral remainder.*

*Proof.* Set $\sigma := \min(\Delta, \chi^m/\kappa_H, \chi^m)$ ($\chi^m/\kappa_H = +\infty$ if $\kappa_H = 0$). If $\chi^m = 0$ the claim is trivial; else $\sigma > 0$. Since $\|s(0)\| = 0$, $\|s(1)\| = \chi^m \geq \sigma$ and $t \mapsto \|s(t)\|$ is continuous, there is $t \in (0, 1]$ with $\|s(t)\| = \sigma$; feasibility is automatic and $\|s(t)\| \leq \Delta$. By Lemma 2(iii), $\|s(t)\|/t \geq \chi^m$, hence $1/t \geq \chi^m/\sigma$. Using Lemma 2(i) and $|\langle s, Hs\rangle| \leq \kappa_H \|s\|^2$,

$$m(0) - m(s(t)) \;\geq\; \frac{\|s(t)\|^2}{t} - \frac{\kappa_H}{2}\|s(t)\|^2 \;\geq\; \chi^m\sigma - \frac{\kappa_H}{2}\sigma^2 \;\geq\; \tfrac{1}{2}\chi^m\sigma,$$

the last step since $\sigma \leq \chi^m/\kappa_H$. □

### C.4 Mode-uniform acceptance theory

**Definition 4** (Mode-uniform properties)**.** At iteration $k$, with model $m_k$, model gradient $g_k^m := \nabla m_k(c_k)$, model criticality $\chi_k^m := \|P_{\mathcal{C}}(c_k - g_k^m) - c_k\|$ and certified bounds $\eta_k^g, \eta_k^J(\cdot)$:

(P1) $\left\|\nabla^2 m_k\right\|_2 \leq \kappa_H$ on the trust region;

(P2) $\left\|g_k^m - \nabla\widehat{J}(c_k)\right\| \leq \eta_k^g$ and the gate $\eta_k^g \leq \kappa_g \min(\chi_k^m, \Delta_k)$;

(P3) $|J_\theta - \widehat{J}| \leq \eta_k^J \leq \kappa_J \mathrm{pred}_k$ at $c_k$ and at $c_k + s_k$ (each with the network used there), and acceptance uses (7) with these certificates;

(P4) $\mathrm{pred}_k \geq \frac{\beta}{2}\chi_k^m \min(\Delta_k, \chi_k^m/\kappa_H, \chi_k^m)$.

Surrogate mode realizes these with $m_k = J_\theta$, $g_k^m = g_\theta(c_k)$, $\eta_k^g = \widehat{\eta}_g(c_k)$, $\eta_k^J = \widehat{\eta}_J$; mode A with $m_k = m_k^A$, $g_k^m = g_a$, $\eta_k^g = \eta_g^a(c_k)$, $\eta_k^J = \widehat{\eta}_J$ (validity of all these bounds: Theorems 2 and 4).

**Lemma 3** (Model–truth discrepancy)**.** *Under (A2)–(A3) and (P2), for any feasible step with* $\|s_k\| \leq \Delta_k$*,*

$$|\mathrm{ared}_k - \mathrm{pred}_k| \leq \eta_k^g\Delta_k + \tfrac{1}{2}(L + \kappa_H)\Delta_k^2. \tag{15}$$

*Proof.* By (A3) ($c_k \in S_0$ by Theorem 3, and $\|s_k\| \leq \Delta_k \leq \Delta_{\max}$) and Taylor's theorem, $\widehat{J}(c_k + s_k) = \widehat{J}(c_k) + \langle\nabla\widehat{J}(c_k), s_k\rangle + R$ with $|R| \leq \frac{L}{2}\|s_k\|^2$. Also $\mathrm{pred}_k = -\langle g_k^m, s_k\rangle - R_k^m$ with the model remainder $R_k^m := m_k(c_k + s_k) - m_k(c_k) - \langle g_k^m, s_k\rangle$, $|R_k^m| \leq \frac{\kappa_H}{2}\|s_k\|^2$ by (P1). Therefore $\mathrm{ared}_k - \mathrm{pred}_k = \langle g_k^m - \nabla\widehat{J}(c_k), s_k\rangle - R + R_k^m$, and (15) follows from Cauchy–Schwarz, (P2) and $\|s_k\| \leq \Delta_k$. (In mode A the correction term of (9) lives inside $m_k$ and never appears explicitly: only the certified anchored gradient error enters.) □

**Lemma 4** (Acceptance at small radius)**.** *Assume (P1)–(P4) with* $\chi_k^m \geq \varepsilon > 0$ *and the gate constants (8). If*

$$\Delta_k \leq \Delta_\star(\varepsilon) := \min\Big(\varepsilon,\ \frac{\varepsilon}{\kappa_H},\ \frac{\beta(1 - \eta_2)\varepsilon}{8(L + \kappa_H)}\Big), \tag{16}$$

*then* $\hat{\rho}_k \geq \eta_2$*; in particular the step is accepted and* $\Delta_{k+1} \geq \Delta_k$*.*

*Proof.* By (16) and $\chi_k^m \geq \varepsilon$ the minimum in (P4) equals $\Delta_k$, so $\mathrm{pred}_k \geq \frac{\beta}{2}\varepsilon\Delta_k$. From (12), $|\mathrm{ared}_k^- - \mathrm{pred}_k| \leq |\mathrm{ared}_k - \mathrm{pred}_k| + 2\eta_k^J(c_k) + 2\eta_k^J(c_k + s_k)$, so by Lemma 3 and (P3)

$$|\hat{\rho}_k - 1| \leq \frac{\eta_k^g\Delta_k + \frac{1}{2}(L + \kappa_H)\Delta_k^2}{\mathrm{pred}_k} + 4\kappa_J \leq \frac{2\Delta_k}{\beta\varepsilon}\Big(\kappa_g + \frac{L + \kappa_H}{2}\Big) + 4\kappa_J,$$

using the gate $\eta_k^g \leq \kappa_g\Delta_k$. By (16), $\frac{2\Delta_k}{\beta\varepsilon}\frac{L+\kappa_H}{2} \leq \frac{1-\eta_2}{8}$; by $\Delta_k \leq \varepsilon$ and (8), $\frac{2\Delta_k}{\beta\varepsilon}\kappa_g \leq \frac{2\kappa_g}{\beta} \leq \frac{1-\eta_2}{8}$; and $4\kappa_J \leq \frac{1-\eta_2}{2}$. Summing, $|\hat{\rho}_k - 1| \leq \frac{3}{4}(1 - \eta_2) < 1 - \eta_2$, hence $\hat{\rho}_k > \eta_2$. □

**Observed model adequacy.** The proof of Theorem 5 uses the curvature bound $\kappa_H$ of (A2) (Assumption 1) in exactly two places: the projected Cauchy floor $\mathrm{pred}_k \geq \frac{\beta}{2}\chi_k^m \min(\Delta_k, \chi_k^m/\kappa_H, \chi_k^m)$ and the model's Taylor remainder along the accepted step. Both are quantities the algorithm computes. Let $g_k^m, \chi_k^m$ be the model gradient and criticality at $c_k$ and define, for every iteration with $\chi_k^m > 0$,

$$\widehat{\kappa}_k^{\mathrm{C}} := \begin{cases} 0, & \mathrm{pred}_k \geq \frac{\beta}{2}\chi_k^m \min(\Delta_k, \chi_k^m), \\ \beta(\chi_k^m)^2/(2\mathrm{pred}_k), & \text{otherwise}, \end{cases} \qquad \widehat{\kappa}_k^{\mathrm{s}} := \frac{2\,|R_k^m|}{\|s_k\|^2}, \tag{17}$$

with $R_k^m := m_k(c_k + s_k) - m_k(c_k) - \langle g_k^m, s_k\rangle$ the model's Taylor remainder along the step: the implied arc curvature (the smallest $\kappa$ for which the observed $\mathrm{pred}_k$ satisfies the Cauchy floor with $\kappa_H = \kappa$) and the observed step curvature.

**Proposition 1** (Observable Cauchy condition)**.** *Let* $\widehat{\kappa} := \sup_k \max(\widehat{\kappa}_k^{\mathrm{C}}, \widehat{\kappa}_k^{\mathrm{s}})$.

*(i) The properties (P1) and (P4) of Definition 4 hold at every iteration with $\kappa_H$ replaced by $\widehat{\kappa}_k^{\mathrm{s}}$ and $\widehat{\kappa}_k^{\mathrm{C}}$ respectively, by construction and without any hypothesis on the network.*

*(ii) If $\widehat{\kappa} < \infty$, then Lemmas 3 and 4 and Theorem 5 hold verbatim with $\kappa_H := \widehat{\kappa}$; in particular the radius floor $\Delta_k \geq \Delta_{\min}$ of the proof holds with the observed constant.*

*(iii) If $\widehat{\kappa} = \infty$ along a subsequence with $\chi_k^m \geq \varepsilon > 0$, the progress guarantees of Theorem 5 are void for that run. Along such a subsequence either the guaranteed decrease $\eta_1 \mathrm{pred}_k$ of accepted steps tends to zero, or all but finitely many steps are rejected and the radius collapses at a center whose certified criticality stays above $\varepsilon - \eta_k^g$; both regimes are observable and reported (Appendix E). Every certificate-level claim is unaffected: validity, Theorem 3 at accepted steps, and the terminal certificate of Theorem 4 are proved without $\kappa_H$.*

*Proof.* (i) By definition of $\widehat{\kappa}_k^{\mathrm{C}}$, if $\mathrm{pred}_k \geq \frac{\beta}{2}\chi_k^m \min(\Delta_k, \chi_k^m)$ then (P4) holds with $\kappa_H = 0$ (the middle argument of the minimum is $+\infty$); otherwise $\chi_k^m/\widehat{\kappa}_k^{\mathrm{C}} = 2\mathrm{pred}_k/(\beta\chi_k^m) < \min(\Delta_k, \chi_k^m)$, so the minimum equals $\chi_k^m/\widehat{\kappa}_k^{\mathrm{C}}$ and $\frac{\beta}{2}\chi_k^m \cdot \chi_k^m/\widehat{\kappa}_k^{\mathrm{C}} = \mathrm{pred}_k$: (P4) holds with equality. For (P1), the only use of the Hessian bound in the proofs is the model remainder $|R_k^m| \leq \frac{\kappa_H}{2}\|s_k\|^2$ in Lemma 3, and $\widehat{\kappa}_k^{\mathrm{s}}$ is defined so that this holds with equality. Both are identities between computed numbers.

(ii) Lemma 3 holds with $\kappa_H$ replaced by $\widehat{\kappa}_k^{\mathrm{s}}$, Lemma 4 with $\kappa_H$ replaced by $\max(\widehat{\kappa}_k^{\mathrm{C}}, \widehat{\kappa}_k^{\mathrm{s}})$ (the first argument of the minimum in (P4) and the remainder bound are the only two entries), and since $\Delta_\star(\varepsilon)$ is nonincreasing in $\kappa_H$, the uniform constant $\widehat{\kappa}$ gives a uniform threshold $\Delta_\star(\varepsilon)$ valid at every iteration. The proof of Theorem 5 uses (P1) and (P4) only through Lemmas 3 and 4 and through the decrease bound $\eta_1 \frac{\beta}{2}\varepsilon \min(\Delta_{\min}, \varepsilon/\kappa_H, \varepsilon)$, which is monotone in $\kappa_H$; it therefore goes through verbatim with $\kappa_H := \widehat{\kappa}$.

(iii) $\widehat{\kappa}_k^{\mathrm{C}} \to \infty$ with $\chi_k^m \geq \varepsilon$ forces $\mathrm{pred}_k \to 0$ by (17), and $\widehat{\kappa}_k^{\mathrm{s}} \to \infty$ forces $|R_k^m|/\|s_k\|^2 \to \infty$. Neither forces rejection: acceptance tests the certified true decrease $\mathrm{ared}_k^- \geq \eta_1 \mathrm{pred}_k$, whose sign the model remainder does not control, so a step with vanishing $\mathrm{pred}_k$ can still be accepted. If infinitely many steps of the subsequence are accepted, their guaranteed decrease $\eta_1 \mathrm{pred}_k$ tends to zero and the summable-decrease argument of Theorem 5 no longer yields a contradiction: no progress conclusion is available, which is the claim. If all but finitely many are rejected, $\Delta_k \to 0$ geometrically while the certified criticality stays above $\varepsilon - \eta_k^g$: a stalled radius collapse (trigger E3 of Appendix E.4). The certificate layer is proved without $\kappa_H$: validity is Theorem 2(i), descent is Theorem 3, and Theorem 4 uses only the 1-Lipschitz property of $P_{\mathcal{C}}$. □

For a run that stops after $K$ iterations the same quantity $\widehat{\kappa}_K := \max_{k \leq K} \max(\widehat{\kappa}_k^{\mathrm{C}}, \widehat{\kappa}_k^{\mathrm{s}})$ is computed, and every per-iteration statement holds up to iteration $K$ with $\kappa_H := \widehat{\kappa}_K$; nothing is claimed about later iterations. The network hypothesis of Theorem 5 is thus monitored: in the campaign $\mathrm{pred}_k$ exceeded the curvature-free floor $\frac{1}{2}\chi_k^m \min(\Delta_k, \chi_k^m)$ at every stepping iteration by factors between 2.6 and 780, i.e. $\widehat{\kappa}_k^{\mathrm{C}} = 0$ throughout (Appendix F.7).

### C.5 CERTIFIED CRITICALITY, THE INEXACT-GRADIENT BRIDGE, AND THE OPTIMALITY GAP

**Theorem 4** (Certified criticality). *Let $c \in \mathcal{C}$ and let $g \in \mathbb{R}^p$ be* any *vector with a certified bound $\eta \geq \left\| g - \nabla \widehat{J}(c) \right\|_2$. Then $\left| \|P_{\mathcal{C}}(c-g) - c\| - \chi(c) \right| \leq \eta$; in particular $\|P_{\mathcal{C}}(c-g) - c\| + \eta \leq \mathrm{tol}$ implies $\chi(c) \leq \mathrm{tol}$. Instances: $g = g_\theta(c)$, $\eta = \widehat{\eta}_g(c)$ (surrogate mode) and $g = g_a$, $\eta = \eta_g^a(c)$ (anchored terminal test).*

*Proof.* $P_{\mathcal{C}}$ is 1-Lipschitz, hence by the reverse triangle inequality

$$\begin{aligned} \left| \|P_{\mathcal{C}}(c-g) - c\| - \left\| P_{\mathcal{C}}(c - \nabla \widehat{J}(c)) - c \right\| \right| &\leq \left\| P_{\mathcal{C}}(c-g) - P_{\mathcal{C}}(c - \nabla \widehat{J}(c)) \right\| \\ &\leq \left\| g - \nabla \widehat{J}(c) \right\| \;\leq\; \eta. \end{aligned}$$

The proof is indifferent to the provenance of $g$, which is what makes the terminal test candidate-agnostic. □

The anchored instance turns the classical inexact-gradient hypothesis into a finite test: by Theorem 2(iv) the mode-A gate compares the computable $\eta_g^a(c_k)$ against its demand, and (P2) of Definition 4 holds whenever it passes. The next lemma shows the test can always be passed at a noncritical center; it is where Theorem 5 consumes (A4).

**Lemma 5** (Anchored gate passability). *Let (A4) hold, let $c_k$ be a center with $\chi(c_k) > 0$ and let $\underline{\Delta} > 0$ be any lower bound for the radius. Then, with $\delta := \max(\delta_a, \max_j \delta_{a,j})$ the solver accuracy, $\eta_g^a(c_k) \leq P(\delta)$ for a polynomial $P$ with $P(0) = 0$, and finitely many mesh refinements produce an anchor at $c_k$ whose certificate passes the mode-A gate of Section 3.3; each $\vartheta$-reduction round of Definition 5 can be met the same way.*

*Proof.* By (3), $\eta_g^a(c_k)^2 = \sum_j \big(\delta_a \|w_{h,j}\| + (\|u_h - u_d\| + \delta_a)\delta_{a,j}\big)^2$ (for C4, the sum of this expression over position and velocity, Theorem 1(B)), and along solver outputs of (A4) with $\delta \to 0$, Assumption 2(i) keeps the observed norms of $w_{h,j}$ and of $u_h - u_d$ bounded ($\|w_{h,j}\|$ and $\|u_h - u_d\|$ for Template A, their values and time derivatives at $T$ for C4), because it bounds the errors in exactly these norms, so $\eta_g^a(c_k) \leq P(\delta)$ for a polynomial $P$ with $P(0) = 0$. By Theorem 4, $\chi_a \geq \chi(c_k) - \eta_g^a(c_k)$, so once $\eta_g^a(c_k) \leq \frac{1}{2}\chi(c_k)$ we have $\chi_a \geq \frac{1}{2}\chi(c_k) > 0$ and the gate threshold $\kappa_g \min(\chi_a, \Delta_k) \geq \kappa_g \min(\frac{1}{2}\chi(c_k), \underline{\Delta})$ is independent of $\delta$, while $\eta_g^a(c_k) \to 0$; (A4) supplies such an anchor after finitely many refinements. □

**Theorem 7** (Optimality gap for a convex reduced objective). *Let $\mathcal{C}$ be closed convex and let $\nabla \widehat{J}$ be $L$-Lipschitz and $\zeta$-strongly monotone on $\mathcal{C}$, i.e. $\langle \nabla \widehat{J}(c) - \nabla \widehat{J}(c'), c - c' \rangle \geq \zeta \|c - c'\|^2$ for all $c, c' \in \mathcal{C}$ (the affine-state cases C1, C2, C4 and C6, Proposition 3), and let $c^\star$ be the minimizer of $\widehat{J}$ on $\mathcal{C}$. Set $\tau := \zeta/L^2$ and $\chi_\tau(c) := \frac{1}{\tau}\left\| P_{\mathcal{C}}(c - \tau \nabla \widehat{J}(c)) - c \right\|$. Then for every $c \in \mathcal{C}$*

$$\|c - c^\star\| \leq \frac{2}{\zeta}\chi_\tau(c), \qquad \widehat{J}(c) - \widehat{J}(c^\star) \leq \begin{cases} \left\| \nabla \widehat{J}(c) \right\| \cdot 2\chi_\tau(c)/\zeta, & \\ \frac{L}{2}(2\chi_\tau(c)/\zeta)^2 & \text{if } \nabla \widehat{J}(c^\star) = 0, \end{cases} \tag{18}$$

*and, with Theorem 4 applied at step $\tau$, the fully computable anchored bound $\|c - c^\star\| \leq \frac{2}{\zeta}\big(\frac{1}{\tau}\|P_{\mathcal{C}}(c - \tau g_a) - c\| + \eta_g^a(c)\big)$ holds.*

*Proof.* Write $T(c) := P_{\mathcal{C}}(c - \tau \nabla \widehat{J}(c))$, so $\|c - T(c)\| = \tau \chi_\tau(c)$. *Contraction, existence and uniqueness.* For $c, c' \in \mathcal{C}$ put $\Delta c := c - c'$ and $\delta := \nabla \widehat{J}(c) - \nabla \widehat{J}(c')$. Nonexpansiveness of $P_{\mathcal{C}}$ gives $\|T(c) - T(c')\| \leq \|\Delta c - \tau \delta\|$, and expanding the square, $\|\Delta c - \tau\delta\|^2 = \|\Delta c\|^2 - 2\tau\langle \Delta c, \delta \rangle + \tau^2 \|\delta\|^2 \leq (1 - 2\tau\zeta + \tau^2 L^2)\|\Delta c\|^2$, using $\langle \Delta c, \delta \rangle \geq \zeta \|\Delta c\|^2$ (strong monotonicity) and $\|\delta\| \leq L\|\Delta c\|$ (Lipschitz continuity). With $\tau = \zeta/L^2$ the bracket is $q^2 := 1 - \zeta^2/L^2 \in [0, 1)$, so $T$ is a $q$-contraction of the complete set $\mathcal{C}$ and has a unique fixed point (Banach). By (13) the fixed points of $T$ are exactly the solutions of the variational inequality $\langle \nabla \widehat{J}(c^\star), y - c^\star \rangle \geq 0$ ($y \in \mathcal{C}$), which for the convex $\widehat{J}$ are exactly the minimizers; hence the minimizer $c^\star$ exists, is unique and is the fixed point

of $T$. With $\Delta c := c - c^\star$ and $T(c^\star) = c^\star$, $\|\Delta c\| \le \|c - T(c)\| + \|T(c) - c^\star\| \le \tau\chi_\tau(c) + q\,\|\Delta c\|$ and, using $1 - \sqrt{1-x} \ge x/2$ for $x \in [0,1]$, $(1-q) \ge \zeta^2/(2L^2)$, whence the first bound. The objective bounds are the convexity inequality with Cauchy–Schwarz, respectively $L$-smoothness at a stationary $c^\star$. The anchored bound is Theorem 4 at step $\tau$ (the same argument at step $\tau$ gives $\chi_\tau(c) \le \frac{1}{\tau}\,\|P_{\mathcal{C}}(c - \tau g) - c\| + \eta$). □

*Remark* 3 (Nonconvex reduced objectives). On the nonlinear cases C3 and C5 the reduced objective is not convex on $\mathcal{C}$. The proof above goes through verbatim on $\mathcal{C} \cap N$ if $\nabla\widehat{J}$ is $L$-Lipschitz and $\zeta$-strongly monotone on a convex neighbourhood $N$ of a local minimizer $c^\star$ that satisfies the second-order sufficient condition (Casas & Tröltzsch, 2015; Robinson, 1980; Dontchev & Rockafellar, 2014) and if $c \in N$; the conclusions then hold with $c^\star$ the unique minimizer on $\mathcal{C} \cap N$. These hypotheses are strong in general: they need a certified lower bound on the Hessian of $\widehat{J}$ over $N$, hence second-order sensitivities with their own majorants, and a proof that $c^\star \in N$. We do not pursue either. On C3 and C5 the terminal certificate therefore certifies stationarity and nothing about the gap (Appendices D.5 and D.7); on C5 even the implementation's surrogate-Hessian estimate of $\zeta$ is not positive at the returned point.

**Corollary 1** (From the unit-step to the step-normalized measure)**.** *Let $c \in \mathcal{C}$, $g \in \mathbb{R}^p$, $\eta \ge \left\|g - \nabla\widehat{J}(c)\right\|_2$ and $\chi_1^{\mathrm{cert}} := \|P_{\mathcal{C}}(c - g) - c\| + \eta$. Then $\chi_\tau(c) \le \chi_1^{\mathrm{cert}}$ for $\tau \ge 1$ and $\chi_\tau(c) \le \chi_1^{\mathrm{cert}}/\tau$ for $\tau < 1$.*

*Proof.* Apply Lemma 2 to $\nabla\widehat{J}(c)$: (iii) gives $\|s(\tau)\|/\tau \le \|s(1)\|$ for $\tau \ge 1$ and (ii) gives $\|s(\tau)\| \le \|s(1)\|$ for $\tau < 1$; finally $\chi_1(c) \le \chi_1^{\mathrm{cert}}$ by Theorem 4. □

This is what lets the implementation report the gap bound without evaluating a projection at the step $\tau = \widehat{\zeta}/\widehat{L}^2$: the unit-step residual it already computes is converted with a certified inflation factor $\max(1, 1/\tau)$; certifying computable $(\zeta, L)$ by second-order tangent majorants is possible but not pursued, and $(\zeta, L)$ are taken per case from Appendix D.

### C.6 Convergence: proof of Theorem 5

**Theorem 5** (Convergence)**.** *Let Assumption 2 and (A1)–(A3) of Assumption 1 hold, let $\mathcal{C}$ be nonempty closed convex, let the gates be enforced with constants (8), and let every step satisfy the projected Cauchy decrease (Theorem 6) for its mode's model. Then, whatever events fire and in whatever order, provided each iteration's gates are eventually satisfied ((A1) in surrogate mode; (A4) via Lemma 5 in mode A):*

- *(i) with the floored gate (6) the run reaches, or returns a point with, certified criticality $\chi \le (1 + \kappa_g)\mathrm{tol}$;*
- *(ii) with the unfloored gate ($\kappa_g\mathrm{tol}$ removed) $\liminf_k \chi(c_k) = 0$;*
- *(iii) every accepted step decreases $\widehat{J}$ by at least $\eta_1\mathrm{pred}_k$; after the event budget is exhausted, (i)–(ii) apply to the remaining zero-solve loop.*

*Proof.* Both modes satisfy (P1)–(P4) with the same constants: (P1) by (A2) and the construction (9); (P2) by Theorem 2(i)/Theorem 4 with the respective gates; (P3) by Theorem 2(i) and the value gate (the acceptance numerator uses surrogate values and anchored certificates in both modes); (P4) by Theorem 6 and the safeguard. Hence Theorem 3 and Lemmas 3 and 4 hold at every iteration with mode-independent constants. We first prove (ii) (unfloored gate) and then derive (i).

Suppose, for contradiction, that $\chi(c_k) \ge \varepsilon_0 > 0$ for all $k \ge k_0$. By Theorem 4 applied to the mode's gradient and certificate, $|\chi_k^m - \chi(c_k)| \le \eta_k^g$, and the gate gives $\eta_k^g \le \kappa_g\chi_k^m$; hence

$$\chi_k^m \;\ge\; \frac{\chi(c_k)}{1 + \kappa_g} \;\ge\; \varepsilon := \frac{\varepsilon_0}{1 + \kappa_g} > 0 \qquad (k \ge k_0) \tag{19}$$

in both modes. (Well-definedness of each iteration under this hypothesis: in surrogate mode (A1) makes the gates attainable in finitely many refinements whenever $\chi_k^m > 0$; in mode A, Lemma 5 with $\underline{\Delta} := \Delta_{\min}$ below.)

*Step 1: the radius is bounded below.* By Lemma 4, whenever $\Delta_k \le \Delta_\star(\varepsilon)$ the step is accepted with $\hat{\rho}_k \ge \eta_2$ and $\Delta_{k+1} \ge \Delta_k$. Hence the radius can be reduced only from values above $\Delta_\star(\varepsilon)$, each time by the factor $\gamma_s$ (radius holds never decrease it), so $\Delta_k \ge \Delta_{\min} := \min\{\Delta_{k_0}, \gamma_s \Delta_\star(\varepsilon)\} > 0$ for all $k \ge k_0$ (by induction: if $\Delta_k \le \Delta_\star(\varepsilon)$ the radius does not shrink, otherwise it drops to at least $\gamma_s \Delta_k > \gamma_s \Delta_\star(\varepsilon)$).

*Step 2: infinitely many accepted steps.* Otherwise every step is eventually rejected and, since the number of consecutive radius holds is bounded by $n_{\mathrm{hold}}$, $\Delta_k \to 0$ geometrically, contradicting Step 1.

*Step 3: summable decrease.* For every accepted step, by Theorem 3, (P4) and (19), $\widehat{J}(c_k) - \widehat{J}(c_{k+1}) \ge \eta_1 \frac{\beta}{2} \varepsilon \min(\Delta_{\min}, \varepsilon/\kappa_H, \varepsilon) =: \delta_\star > 0$, and on rejected steps $c_{k+1} = c_k$, so $\widehat{J}$ is non-increasing along the whole sequence. Summing over infinitely many accepted steps contradicts $\widehat{J} \ge J_{\inf}$. Hence no such $\varepsilon_0$ exists and $\liminf_k \chi(c_k) = 0$, which is (ii).

(i) The argument runs on the computable bound $\chi_k^m + \eta_k^g$, not on the unknown $\chi(c_k)$. Suppose $\chi_k^m + \eta_k^g > (1+\kappa_g)\mathrm{tol}$ for all $k \ge k_1$. At each such $k$ the floored gate (6) leaves two cases. Unfloored branch: $\eta_k^g \le \kappa_g \min(\chi_k^m, \Delta_k)$, so $(1+\kappa_g)\chi_k^m \ge \chi_k^m + \eta_k^g > (1+\kappa_g)\mathrm{tol}$ and also $\eta_k^g \le \kappa_g \Delta_k$. Floored branch: $\eta_k^g \le \kappa_g \mathrm{tol}$, so $\chi_k^m > (1+\kappa_g)\mathrm{tol} - \kappa_g \mathrm{tol} = \mathrm{tol}$. In both branches $\chi_k^m > \mathrm{tol} =: \varepsilon$ for all $k \ge k_1$. Steps 1–3 above run verbatim on this direct lower bound (they used the true criticality only through (19), which is bypassed): Lemma 4 applies with $\chi_k^m \ge \varepsilon$, through $\eta_k^g \le \kappa_g \Delta_k$ in the unfloored branch and, in the floored branch, through $\eta_k^g \le \kappa_g \varepsilon$, which bounds the only gate term of its proof, $\eta_k^g \Delta_k / \mathrm{pred}_k$, by the same $2\kappa_g/\beta$, and the summable-decrease contradiction follows as before. Hence the run reaches an iterate with $\chi_k^m + \eta_k^g \le (1+\kappa_g)\mathrm{tol}$, a computable certified bound; the best-certified return, which tracks the smallest computable bound over the run (Appendix E.4), returns a point carrying it, and Theorem 4 converts it into certified $\chi \le (1+\kappa_g)\mathrm{tol}$.

(iii) Descent is Theorem 3. After the event budget is exhausted no event fires; every subsequent iteration is surrogate mode with certificates $\widehat{\mu}, \widehat{\mu}_j$, which satisfy validity (Theorem 2(i)) and are dominated by the unanchored ones (Theorem 2(ii)), so (A1) still makes the gates attainable and the tail is the gate-enforced zero-solve loop, to which Steps 1–3 and the argument for (i) apply. The best-certified return selects the stored iterate with the smallest certified terminal bound over a set that only grows. □

*Remark* 4 (Why flooring is the right design). As $\chi \downarrow 0$ the unfloored demand $\kappa_g \chi$ requires unbounded *relative* accuracy, whereas the terminal certificate requires only absolute accuracy $\eta^g \lesssim \mathrm{tol}$. Near convergence the unfloored gate is strictly harder to satisfy than the goal it serves and can block a certificate otherwise within reach; since a run with finite tolerance never claims more than tol-criticality, the floored gate costs nothing that is actually claimed.

# D THE SIX INSTANTIATIONS

This appendix proves Assumption 2, and (A3), for the six classes of Table 5, with every constant explicit. Throughout, candidates $v, \omega \in \mathcal{V}$ are conforming functions satisfying the essential boundary/initial conditions exactly and regular enough to evaluate the strong residual ($\Delta v \in L^2$ where required); solver outputs enter as such candidates after conforming reconstruction ((A4)); nothing in the proofs uses the provenance of the candidate. The notation of the state equations is: $D > 0$ the diffusivity (C2), $a > 0$ the wave speed (C4), $\nu > 0$ the viscosity (C5); $\varphi_j$ the control basis functions, so that the load is the control function $z(c) = z_0 + \sum_j c_j \varphi_j$ of Section 2 (the fixed part $z_0$ vanishes in every experiment); in the time-dependent problems C2 and C4 the basis functions, and hence the load, may depend on $t$ (the experiments use time-independent loads for C2 and time profiles for C4). $\sigma_B := (\sum_j \|\varphi_j\|^2)^{1/2}$ in the case norm.

## D.1 FUNCTION-SPACE SETTING

Table 3 fixes, for every case, the space in which the objective is tracked and the spaces in which the majorants of Assumption 2 measure errors. Template A (Theorem 1(A)) covers C1–C3, C5 and C6; Template B covers C4, where the observation $\Lambda$ of Assumption 2 is the pair (position,

velocity) at $t = T$. The pivot space of a Gelfand triple is written $L^2(\Omega)$, not $H$, to keep it apart from the tracking space $H$. The $L^\infty$ arguments (Lemma 7, the tangent bound of C3, C5 and C6) use $H_0^1(0,1) \hookrightarrow L^\infty(0,1)$ and are intrinsically one-dimensional; the C1 certificate uses only the Friedrichs inequality and integration by parts.

Table 3: Function-space setting of the six instantiations ($Q = (0,1) \times (0,T)$). $V$ is the energy space of the state equation, $C_{\mathrm{F}}$ the Friedrichs constant of Lemma 6 on the relevant domain, "state map" refers to $c \mapsto u(c)$.

| | state equation | $d$ | tracking space $H$ | error norm(s) $X$ | $V$ | $C_{\mathrm{F}}$ | state map |
|---|---|---|---|---|---|---|---|
| C1 | $-\Delta u = z(c)$ | 2 | $L^2(\Omega), \Omega = (0,1)^2$ | $L^2(\Omega)$ | $H_0^1(\Omega)$ | $1/(\pi\sqrt{2})$ | affine |
| C2 | $u_t - Du_{xx} = z(c)$ | 1 | $L^2(Q)$ | $L^2(Q)$ | $H_0^1(0,1)$ | $1/\pi$ | affine |
| C3 | $-u'' + u^3 = z(c)$ | 1 | $L^2(0,1)$ | $L^2(0,1)$ | $H_0^1(0,1)$ | $1/\pi$ | nonlinear, monotone |
| C4 | $u_{tt} - a^2 u_{xx} = z(c)$ | 1 | $L^2(0,1)$ at $t = T$ (pos., vel.) | $L^2(0,1)^2$ | $H_0^1(0,1) \times L^2(0,1)$ | $1/\pi$ | affine |
| C5 | $-\nu u'' + (u^2/2)' = z(c)$ | 1 | $L^2(0,1)$ | $L^2(0,1)$ | $H_0^1(0,1)$ | $1/\pi$ | nonlinear, unique for $\|z\| < z_{\mathrm{crit}}$ |
| C6 | $-u'' = z(c)$ | 1 | $L^2(0,1)$ | $L^2(0,1)$, plus $L^\infty$ margin | $H_0^1(0,1)$ | $1/\pi$ | affine |

**Proposition 2** (Control-to-state maps and sensitivities)**.** *In each of C1–C6 there is an open set $\mathcal{O} \supseteq \mathcal{C}$ ($\mathcal{O} = \mathbb{R}^p$ in C1–C4 and C6; $\mathcal{O} = \{c : \|z(c)\| < z_{\mathrm{crit}}\}$ in C5, which contains $\mathcal{C}$ for the data of the experiment, Appendix F.1) on which the exact state $u(c)$ is uniquely defined with values in the state space $\mathcal{S} := H_0^1(\Omega)$ (C1, C3, C5, C6), $W(0,T)$ (C2) or $C([0,T];V) \cap C^1([0,T];L^2(0,1))$ (C4). The map $c \mapsto u(c) \in \mathcal{S}$ is $C^\infty$ (affine in C1, C2, C4, C6), and $\partial_{c_j} u(c) = w_j(c)$, where $w_j(c) \in \mathcal{S}$ is the unique solution of the tangent equation: the state equation (linear part) with load $\varphi_j$ in C1, C2, C4, C6; $-w'' + 3u^2 w = \varphi_j$ in C3; $-\nu w'' + (uw)' = \varphi_j$ in C5. Consequently, for the observation $\Lambda$ of Assumption 2 (a bounded linear map $\mathcal{S} \to H$, resp. $\mathcal{S} \to L^2(0,1)^2$), $\widehat{J}$ is $C^\infty$ on $\mathcal{O}$ and the chain rule used in Theorem 1 holds: $\partial_{c_j} \widehat{J} = \langle \Lambda u - u_d, \Lambda w_j \rangle + \alpha c_j$ for Template A and its Template B analogue.*

*Proof. C1, C2, C4, C6.* The solution operator $z \mapsto u$ is bounded and linear into $\mathcal{S}$: Lax–Milgram in $H_0^1$ (C1, C6, with Lemma 6), the energy estimate of (Evans, 2010, §7.1) for $z \in L^2(Q)$ (C2), and Lemma 8 for $z \in L^1(0,T;L^2)$ (C4). As $z(c) = z_0 + \sum_j c_j \varphi_j$, $u(c) = u_0 + \sum_j c_j w_j$.

*C3, C5.* $F(u,z) = -u'' + u^3 - z$ (C3), $F(u,z) = -\nu u'' + (u^2/2)' - z$ (C5) is $C^\infty$ from $H_0^1(0,1) \times L^2(0,1)$ to $H^{-1}(0,1)$, since $u \mapsto u^3$ and $u \mapsto (u^2/2)'$ are bounded polynomial maps $H_0^1 \to L^2 \subset H^{-1}$ ($H^1(0,1)$ is an algebra). $\partial_u F(u,z)w = -w'' + 3u^2 w$, resp. $-\nu w'' + (uw)'$, satisfies $\langle \partial_u F\, w, w \rangle \geq \|w'\|^2$ (C3, as $3u^2 \geq 0$), resp. $\geq \kappa_{\min}(z) \|w'\|^2$ (C5, Lemma 9(iii), on $\|z\| < z_{\mathrm{crit}}$), hence is an isomorphism $H_0^1 \to H^{-1}$ by Lax–Milgram. The state is unique (C3: strict monotonicity, Theorem 10; C5: Lemma 9(iii)), so the implicit function theorem gives that $z \mapsto u$ is $C^\infty$ and, differentiating $F(u(z), z) = 0$ in the direction $\varphi_j$, that $\partial_{c_j} u = w_j$; $c \mapsto z(c)$ is affine, and $\mathcal{O}$ is open in C5 by continuity of $z(\cdot)$. The statement on $\widehat{J}$ is the chain rule, $\Lambda$ being bounded and linear. □

### D.2 Two sharp embedding lemmas

**Lemma 6** (Sharp Friedrichs constants)**.** *For $v \in H_0^1(\Omega)$, $\|v\| \leq C_{\mathrm{F}} \|\nabla v\|$ with $C_{\mathrm{F}} = \lambda_1(\Omega)^{-1/2}$, $\lambda_1$ the first Dirichlet eigenvalue of $-\Delta$. In particular $C_{\mathrm{F}} = 1/\pi$ for $\Omega = (0,1)$ and $C_{\mathrm{F}} = 1/(\pi\sqrt{2})$ for $\Omega = (0,1)^2$.*

*Proof.* This is the variational characterization $\lambda_1 = \min_{v \in H_0^1 \setminus \{0\}} \|\nabla v\|^2 / \|v\|^2$, attained at the first eigenfunction (Evans, 2010, §6.5.1, Thm. 2); the eigenfunctions are $\sin \pi x$ on $(0,1)$ and $\sin \pi x \sin \pi y$ on $(0,1)^2$. □

**Lemma 7** (Sharp 1D embedding)**.** *For $v \in H_0^1(0,1)$, $\|v\|_{L^\infty} \leq \frac{1}{2} \|v'\|_{L^2}$, and $\frac{1}{2}$ is attained.*

*Proof.* $2v(x) = \int_0^x v' - \int_x^1 v'$, so $2|v(x)| \leq \int_0^1 |v'| \leq \|v'\|$; equality for the hat function $\min(x, 1-x)$. □

### D.3 C1: LINEAR ELLIPTIC

**Theorem 8** (C1 certificate)**.** *Let $-\Delta u = z(c)$ in $\Omega = (0,1)^2$, $u \in H_0^1(\Omega)$, $z(c)$ affine in $c$, and let $v$ be conforming with $v|_{\partial\Omega} = 0$, $\Delta v \in L^2$, residual $r(v;c) := \Delta v + z(c)$. Then*

$$\|\nabla(u-v)\| \le C_{\mathrm{F}} \|r(v;c)\|, \qquad \|u-v\| \le C_{\mathrm{F}}^2 \|r(v;c)\|, \tag{20}$$

*and the same bounds hold for each sensitivity $w_j$ with $r$ replaced by the tangent residual $r_{V,j}(\omega) := \Delta\omega + \varphi_j$. Assumption 2 holds with $\mathcal{M}(v;c) = C_{\mathrm{F}}^2 \|r(v;c)\|$, $\mathcal{M}_j(\omega;c) = C_{\mathrm{F}}^2 \|r_{V,j}(\omega)\|$ ($\mu$-independent),* val $\equiv 1$*, and $\varsigma(t) = C_B t$ with $C_B = C_{\mathrm{F}}^2 \sigma_B$; the tangent majorants are independent of $c$, since in the linear classes C1, C2, C4 and C6 the sensitivity $w_j$ itself does not depend on $c$.*

*Proof.* Let $e := u - v \in H_0^1$. For all $\psi \in H_0^1$, $\langle \nabla u, \nabla\psi\rangle = \langle z(c), \psi\rangle$ and, integrating by parts (legitimate since $\Delta v \in L^2$), $\langle \nabla v, \nabla\psi\rangle = -\langle \Delta v, \psi\rangle$; subtracting, $\langle \nabla e, \nabla\psi\rangle = \langle r, \psi\rangle$. Taking $\psi = e$, Cauchy–Schwarz and Lemma 6 give $\|\nabla e\|^2 \le \|r\| \, \|e\| \le C_{\mathrm{F}} \|r\| \, \|\nabla e\|$, the first bound; the second follows from Lemma 6 again. The tangent problem is $-\Delta w_j = \varphi_j$ with homogeneous data and no dependence on $c$ or $u$, so the identical argument applies. For (iii): with $v$ fixed, $r(v;c') - r(v;c) = \sum_j (c'_j - c_j)\varphi_j$, so $\|r(v;c')\| \le \|r(v;c)\| + \sigma_B \|c' - c\|_2$; multiply by $C_{\mathrm{F}}^2$. □

**Proposition 3** (Smoothness of the truth in the linear classes)**.** *In C1, C2, C4 and C6 the control-to-state map $c \mapsto u(c)$ is affine, so $\widehat{J}$ is a convex quadratic, $\nabla\widehat{J}$ is Lipschitz with $L = \|\mathcal{G}\|^2 + \alpha$ where $\mathcal{G} : \mathbb{R}^p \to H$, $c \mapsto \sum_j c_j w_j$ (with values in $H \times H$ for C4; for C1 the columns are $(-\Delta)^{-1}\varphi_j$), so that $\nabla^2\widehat{J} = \mathcal{G}^*\mathcal{G} + \alpha I$ and $\nabla\widehat{J}$ is $\zeta$-strongly monotone with $\zeta = \sigma_{\min}(\mathcal{G})^2 + \alpha$, and $\widehat{J} \ge 0$ is bounded below: (A3) holds. Here $\zeta > 0$: in C1 ($p = 1$, $\alpha = 0$) $\zeta = \|w_1\|^2 > 0$; in C2 ($\alpha = 0$) the sensitivities are proportional to $(1 - e^{-D\lambda_k t}) \sin k\pi x$, $k = 1, 2$, hence linearly independent, so $\zeta > 0$; in C4 $\zeta = \alpha$ since $\mathcal{G}$ has rank at most $2 < p$; in C6 $\zeta \ge \alpha$. In C3 and C5 the control-to-state map is $C^2$ on the bounded sets visited (Proposition 2; implicit function theorem: the linearized operators $-\partial_{xx} + 3u^2$ and $-\nu\partial_{xx} + \partial_x(u\,\cdot)$ are boundedly invertible on $H_0^1$ under the monotonicity (C3), respectively under the uniform coercivity $\kappa_{\min} > 0$ of Lemma 9(iii), which holds on all of $\mathcal{C}$ when $\|z(c)\| < z_{\mathrm{crit}}$ there (C5)), so $\widehat{J}$ is $C^2$ with bounded Hessian on $\mathcal{C} \cap (S_0 + \bar{B}_{\Delta_{\max}})$ (bounded because $\alpha > 0$ in C3 and because $\mathcal{C}$ is compact in C5) and (A3) holds there. (A2), by contrast, concerns the network $J_\theta$ and is covered by Proposition 1, not by any property of the PDE.*

### D.4 C2: PARABOLIC

Let $V := H_0^1(0,1)$, $V^* = H^{-1}(0,1)$, the Gelfand triple $V \subset L^2(0,1) \subset V^*$, and $W(0,T) := \{u \in L^2(0,T;V) : u_t \in L^2(0,T;V^*)\} \hookrightarrow C([0,T];L^2(0,1))$ (Evans, 2010, §5.9.2). The state equation is

$$\partial_t u - D\,\partial_{xx} u = z(c) \ \text{ in } Q = (0,1)\times(0,T), \qquad u(0,\cdot) = u(1,\cdot) = 0, \qquad u(\cdot,0) = 0, \tag{21}$$

with $D > 0$ and $z(c) \in L^2(Q)$ affine in $c$; the problem is ill posed for $D \le 0$ (for $D < 0$ it is a backward heat equation; for $D = 0$ there is no smoothing and no Friedrichs absorption). For $z \in L^2(0,T;L^2(0,1))$ there is a unique weak solution $u \in W(0,T)$ (Evans, 2010, Thm. 7.1.3–4). The objective tracks the state over the whole cylinder $Q$, $\widehat{J}(c) = \frac{1}{2}\|u(c) - u_d\|_{L^2(Q)}^2$, i.e. Template A with $H = L^2(0,T;L^2(0,1))$.

**Theorem 9** (C2 certificate)**.** *Let $v \in W(0,T)$ satisfy the boundary and initial conditions of (21) exactly, with residual $r(v;c) := z(c) - \partial_t v + D\partial_{xx} v \in L^2(Q)$. Then, with $C_{\mathrm{F}} = 1/\pi$ and $e := u - v$,*

$$\|\partial_x e\|_{L^2(Q)} \le \frac{C_{\mathrm{F}}}{D}\|r\|_{L^2(Q)}, \quad \|e\|_{L^2(Q)} \le \frac{C_{\mathrm{F}}^2}{D}\|r\|_{L^2(Q)}, \quad \|e(T)\|_{L^2(0,1)} \le \frac{C_{\mathrm{F}}}{\sqrt{2D}}\|r\|_{L^2(Q)}. \tag{22}$$

*The tangent equations $\partial_t w_j - D\partial_{xx} w_j = \varphi_j$ carry the same homogeneous data and the same bounds. Assumption 2 holds with these functionals ($\mu$-independent tangents,* val $\equiv 1$*) and $C_B = (C_{\mathrm{F}}^2/D)\sigma_B$ with $\sigma_B$ in $L^2(Q)$.*

*Proof. Energy route.* The error $e = u - v \in W(0,T)$ satisfies $\langle e_t(t), \psi\rangle_{V^*,V} + D\langle e_x(t), \psi_x\rangle = \langle r(t), \psi\rangle$ for all $\psi \in V$ and a.e. $t$, with $e(0) = 0$. Since $e \in W(0,T)$, the map $t \mapsto \|e(t)\|_{L^2(0,1)}^2$

is absolutely continuous with $\frac{d}{dt}\|e(t)\|^2_{L^2(0,1)} = 2\langle e_t(t), e(t)\rangle_{V^*,V}$ a.e. (Evans, 2010, Thm. 5.9.3); this is the step that requires the Gelfand triple, since $e_t(t) \notin L^2(0,1)$ in general. Testing with $\psi = e(t)$ and integrating over $(0,T)$,

$$\tfrac{1}{2}\|e(T)\|^2_{L^2(0,1)} + D\|e_x\|^2_{L^2(Q)} = \int_0^T \langle r(t), e(t)\rangle dt \\ \le \|r\|_{L^2(Q)}\|e\|_{L^2(Q)} \le C_{\mathrm{F}}\|r\|_{L^2(Q)}\|e_x\|_{L^2(Q)},$$

by Cauchy–Schwarz in $L^2(Q)$ and Lemma 6 applied for a.e. $t$ followed by integration in $t$. If $\|e_x\|_{L^2(Q)} = 0$ then $e = 0$ a.e. by Lemma 6 and all bounds are trivial; otherwise dividing by $D\|e_x\|_{L^2(Q)}$ gives the first bound, the second follows from Lemma 6 once more, and discarding the gradient term instead gives $\frac{1}{2}\|e(T)\|^2 \le \|r\|\|e\| \le (C_{\mathrm{F}}^2/D)\|r\|^2$, i.e. $\|e(T)\| \le \sqrt{2}\,C_{\mathrm{F}}D^{-1/2}\|r\|$.

*Modal route (same constants, sharper terminal bound, and the validation solution).* Expand $e(\cdot,t) = \sum_{k\ge1} e_k(t)\sqrt{2}\sin(k\pi x)$ and $r$ likewise; each mode solves $e_k' + D\lambda_k e_k = r_k$, $e_k(0) = 0$, $\lambda_k = k^2\pi^2$, whence $e_k(t) = \int_0^t e^{-D\lambda_k(t-s)}r_k(s)\,ds$. Young's convolution inequality gives $\|e_k\|_{L^2(0,T)} \le \|e^{-D\lambda_k\cdot}\|_{L^1(0,\infty)}\|r_k\|_{L^2(0,T)} = (D\lambda_k)^{-1}\|r_k\|_{L^2(0,T)}$, so $\|e\|^2_{L^2(Q)} = \sum_k \|e_k\|^2 \le (D\lambda_1)^{-2}\|r\|^2_{L^2(Q)}$ and $\|e_x\|^2_{L^2(Q)} = \sum_k \lambda_k\|e_k\|^2 \le (D^2\lambda_1)^{-1}\|r\|^2$: exactly the two constants of the energy route, now seen to be attained in the limit of a slowly varying first-mode residual on a long horizon. For the terminal value, Cauchy–Schwarz in $s$ gives $|e_k(T)| \le \|r_k\|_{L^2(0,T)}\big(\int_0^T e^{-2D\lambda_k s}ds\big)^{1/2} \le (2D\lambda_k)^{-1/2}\|r_k\|$, hence $\|e(T)\| \le (2D\lambda_1)^{-1/2}\|r\|_{L^2(Q)} = C_{\mathrm{F}}(2D)^{-1/2}\|r\|_{L^2(Q)}$, which is the third bound in (22) and improves the energy route by a factor 2. The same expansion gives the exact solution by superposition: $u(c) = u_0 + \sum_j c_j U_j$ with $U_j(x,t) = \sum_k\big(\int_0^t e^{-D\lambda_k(t-s)}\hat{\varphi}_{j,k}(s)ds\big)\sqrt{2}\sin k\pi x$; for the time-independent loads of the experiment ($\varphi_k = \sin k\pi x$, $k = 1,2$) this is $u = \sum_{k=1}^2 c_k\frac{1-e^{-D\lambda_k t}}{D\lambda_k}\sin k\pi x$, the "Fourier solution" used for validation in Appendix F.1. The shift bound is as in Theorem 8, with the $L^2(Q)$ norms of the loads. □

### D.5 C3: SEMILINEAR MONOTONE

The reduced objective of this case is not convex on $\mathcal{C}$, so Theorem 7 does not apply (Remark 3); the terminal certificate certifies stationarity.

**Theorem 10** (C3 certificates)**.** *Let $-u'' + u^3 = z(c)$ on $(0,1)$, $u \in H^1_0$, $z(c)$ affine in $c$, and let $v \in H^1_0$ with $v'' \in L^2$ have residual $r(v;c) := v'' - v^3 + z(c)$. Then with $C_{\mathrm{F}} = 1/\pi$,*

$$\|(u-v)'\| \le C_{\mathrm{F}}\|r\|, \qquad \|u-v\| \le C_{\mathrm{F}}^2\|r\|, \qquad \|u-v\|_\infty \le \tfrac{1}{2}C_{\mathrm{F}}\|r\|. \tag{23}$$

*Let $w_j$ solve the tangent equation $-w_j'' + 3u^2 w_j = \varphi_j$ and let $\omega \in H^1_0$ have tangent residual $r_{V,j}(\omega;v) := \omega'' - 3v^2\omega + \varphi_j$. Then, whenever $\|u-v\| \le \mu$,*

$$\|w_j - \omega\| \le C_{\mathrm{F}}^2\big(\|r_{V,j}(\omega;v)\| + 3M_v\|\omega\|_\infty\mu\big) =: \mathcal{M}_j(\omega;c,v,\mu), \tag{24}$$

*where $M_v := 2\|v\|_\infty + \frac{1}{2}C_{\mathrm{F}}\|r(v;c)\| \ge \|u+v\|_\infty$ is computable. Assumption 2 holds with these functionals ($\mathcal{M}_j$ affine, hence nondecreasing, in $\mu$;* val $\equiv 1$*), and the state shift bound is the C1 formula.*

*Proof. Existence and uniqueness.* $A(u) := -u'' + u^3$ is strictly monotone: $\langle A(u_1) - A(u_2), u_1 - u_2\rangle = \|(u_1-u_2)'\|^2 + \int(u_1^3 - u_2^3)(u_1-u_2) \ge \|(u_1-u_2)'\|^2$, coercive and hemicontinuous, so by Browder–Minty it is a bijection $H^1_0 \to H^{-1}$ (Tröltzsch, 2010; Hinze et al., 2009). *State bound.* With $e = u - v$, for all $\psi \in H^1_0$, $\langle e', \psi'\rangle + \langle u^3 - v^3, \psi\rangle = \langle r, \psi\rangle$; taking $\psi = e$ and discarding the nonnegative monotone term, $\|e'\|^2 \le \langle r, e\rangle \le C_{\mathrm{F}}\|r\|\|e'\|$, which yields (23) with Lemmas 6 and 7. *Tangent bound.* Let $\epsilon := w_j - \omega$; for $\psi \in H^1_0$, $\langle \epsilon', \psi'\rangle + \langle 3u^2w_j - 3v^2\omega, \psi\rangle = \langle r_{V,j}, \psi\rangle$. Using $3u^2w_j - 3v^2\omega = 3u^2\epsilon + 3(u^2 - v^2)\omega$ and $\psi = \epsilon$, $\|\epsilon'\|^2 + 3\int u^2\epsilon^2 = \langle r_{V,j}, \epsilon\rangle - 3\int(u^2 - v^2)\omega\epsilon$. The second term on the left is discarded; pointwise $|u^2 - v^2| \le M_v|e|$ with $\|u+v\|_\infty \le 2\|v\|_\infty + \|e\|_\infty \le M_v$ by (23); hence $\|\epsilon'\|^2 \le (\|r_{V,j}\| + 3M_v\|\omega\|_\infty\|e\|)C_{\mathrm{F}}\|\epsilon'\|$ and (24) follows with $\|e\| \le \mu$ and Lemma 6. □

### D.6 C4: HYPERBOLIC (WAVE)

Let $u_{tt} - a^2 u_{xx} = z(c)$ on $(0,1)\times(0,T)$, $a > 0$, with homogeneous Dirichlet lateral conditions and $u(\cdot,0) = u_t(\cdot,0) = 0$, $z(c) \in L^1(0,T;L^2(0,1))$ affine in $c$ (time dependence of the load through the basis is allowed and used). With $V = H_0^1(0,1)$ the state $u(c)$ is the energy solution of Lemma 8 with load $z(c)$, so $(u,u_t) \in C([0,T];V\times L^2(0,1))$ and $u(T) \in V$, $u_t(T) \in L^2(0,1)$ are well defined; $u(c) = u_0 + \sum_j c_j w_j$ is affine in $c$, $w_j$ being the energy solution for the load $\varphi_j$ (this settles differentiability of $c \mapsto (u(T), u_t(T))$ for C4), so Template B of Theorem 1 applies with $\widehat{J} = \frac{1}{2}\|u(T) - u_d\|^2 + \frac{1}{2}\|u_t(T)\|^2 + \frac{\alpha}{2}|c|^2$.

**Lemma 8** (Energy solutions of the wave equation)**.** *Let $V = H_0^1(0,1)$, $a > 0$, and let $\mathfrak{X} := V \times L^2(0,1)$ carry the energy norm $\|(y,\eta)\|_{\mathfrak{X}}^2 := a^2\|y_x\|^2 + \|\eta\|^2$. The operator $\mathcal{A}(y,\eta) := (\eta,\, a^2 y_{xx})$, $D(\mathcal{A}) = (H^2 \cap V)\times V$, generates a unitary $C_0$-group $S(t)$ on $\mathfrak{X}$. For $g \in L^1(0,T;L^2(0,1))$ the mild solution $Y(t) = \int_0^t S(t-s)(0,g(s))\,ds$ lies in $C([0,T];\mathfrak{X})$, satisfies $\|Y(T)\|_{\mathfrak{X}} \le \int_0^T \|g(s)\|\,ds$, and $Y = (y,y_t)$ with $y \in C([0,T];V)\cap C^1([0,T];L^2(0,1))$. We call $y$ the* energy solution *of $y_{tt} - a^2 y_{xx} = g$, $y(0) = y_t(0) = 0$ (Cazenave & Haraux, 1998).*

*Proof.* $\mathcal{A}$ is skew-symmetric by integration by parts, and $I \pm \mathcal{A}$ is onto: $(y \mp \eta, \eta \mp a^2 y_{xx}) = (h_1,h_2) \in V\times L^2$ reduces to $-a^2 y_{xx} + y = h_1 \pm h_2 \in L^2$, solved in $H^2\cap V$ (Lax–Milgram), with $\eta = \pm(y - h_1) \in V$. Hence $\mathcal{A}$ is skew-adjoint and generates a unitary group by Stone's theorem (Pazy, 1983, Ch. 1, §10). For the mild solution, continuity and $\|Y(T)\| \le \int_0^T \|S(T-s)(0,g(s))\|\,ds = \int_0^T \|g(s)\|\,ds$ are (Pazy, 1983, Ch. 4, §2) together with unitarity. Finally the first row of $\mathcal{A}$ is the bounded map $(y,\eta)\mapsto\eta$, so the integrated equation $Y = \mathcal{A}\int_0^{\cdot} Y + \int_0^{\cdot}(0,g)$ gives $y = \int_0^{\cdot} y_2$ for $Y = (y,y_2)$, i.e. $y \in C^1([0,T];L^2)$ with $y_t = y_2$. □

**Theorem 11** (Wave energy certificate)**.** *Let $v \in C^2([0,T];L^2(0,1)) \cap C^1([0,T];V) \cap C([0,T];H^2(0,1)\cap V)$ with $v(0) = v_t(0) = 0$ (tanh networks with a smooth mask and spline–modal reconstructions are of this form), and have residual $r(v;c) := v_{tt} - a^2 v_{xx} - z(c) \in L^1(0,T;L^2)$. Set $e := u - v$ and $R(v;c) := \int_0^T \|r(v;c)(\cdot,t)\|\,dt$. Then*

$$\|e_t(T)\| \le R, \qquad a\,\|e_x(T)\| \le R, \qquad \|e(T)\| \le \frac{C_{\mathrm{F}} R}{a}, \qquad C_{\mathrm{F}} = \tfrac{1}{\pi}, \tag{25}$$

*and the same bounds hold for each tangent variable with $r$ replaced by the tangent residual. Assumption 2 holds with these functionals (Template B assembly,* val $\equiv 1$*) and with $C_B^{(1)} = \big(\sum_j (\int_0^T \|\varphi_j(t)\|\,dt)^2\big)^{1/2}$ for the velocity/gradient components and $C_B^{(0)} = \frac{C_{\mathrm{F}}}{a} C_B^{(1)}$ for the position component; tangent majorants are $c$-independent.*

*Proof.* Let $u$ be the energy solution of Lemma 8 for the load $z(c)$. $Y_v := (v, v_t) \in C^1([0,T];\mathfrak{X}) \cap C([0,T];D(\mathcal{A}))$ is a classical solution of $Y_v' = \mathcal{A}Y_v + (0, g_v)$ with $g_v := v_{tt} - a^2 v_{xx} = z(c) + r$, hence the mild solution for that forcing. The mild formula is linear in the forcing, so $Y_e := (u,u_t) - Y_v = (e,e_t)$ is the mild solution with forcing $(0,-r)$, and unitarity gives $\|Y_e(T)\|_{\mathfrak{X}} \le \int_0^T \|r(s)\|\,ds = R$, i.e. $a^2\|e_x(T)\|^2 + \|e_t(T)\|^2 \le R^2$, i.e. $E(T) \le R^2/2$ for the energy $E := \frac{1}{2}\|(e,e_t)\|_{\mathfrak{X}}^2$. The first two bounds of (25) follow, the third from Lemma 6. The tangent $w_j$ is the energy solution for the load $\varphi_j$ ($u(c) = u_0 + \sum_j c_j w_j$), and the argument is identical with $r_{V,j}(\omega) = \omega_{tt} - a^2\omega_{xx} - \varphi_j$. The shift bound follows from $R(v;c') \le R(v;c) + \int_0^T \Big\|\sum_j \delta c_j \varphi_j(t)\Big\|\,dt$ with Minkowski's inequality and Cauchy–Schwarz in $\mathbb{R}^p$. □

*Remark* 5 (Closed form, dependence on the horizon, and the limits of the certificate). In the instantiation the load is a single spatial mode, $z(x,t;c) = g(t;c)\sin\pi x$ with $g(t;c) = \sum_k c_k B_k(t)$ in a B-spline basis. Then $u(x,t) = y(t)\sin\pi x$ with the modal ODE $y'' + (a\pi)^2 y = g$, $y(0) = y'(0) = 0$, solved by Duhamel's formula

$$y(t) = \frac{1}{a\pi}\int_0^t \sin\big(a\pi(t-s)\big) g(s)\,ds, \qquad y'(t) = \int_0^t \cos\big(a\pi(t-s)\big) g(s)\,ds.$$

Both terminal values are linear in $c$, $y(T) = \langle q_0, c\rangle$, $y'(T) = \langle q_1, c\rangle$ with $(q_0)_k = \frac{1}{a\pi}\int_0^T \sin(a\pi(T-s))B_k(s)ds$ and $(q_1)_k = \int_0^T \cos(a\pi(T-s))B_k(s)ds$, so (with $u_d = \gamma \sin \pi x$ and $\|\sin \pi x\|^2 = \frac{1}{2}$) the reduced objective is the explicit convex quadratic $\widehat{J}(c) = \frac{1}{4}\big[(\langle q_0, c\rangle - \gamma)^2 + \langle q_1, c\rangle^2\big] + \frac{\alpha}{2}|c|^2$, with Hessian $\frac{1}{2}(q_0 q_0^\top + q_1 q_1^\top) + \alpha I$, hence $L = \frac{1}{2}(\|q_0\|^2 + \|q_1\|^2) + \alpha$ and strong convexity modulus $\zeta = \alpha$. This is the "modal Duhamel" reference solver of Appendix F.4, and it makes the dependence on the horizon explicit: the reachable set of $(y(T), y'(T))$ is all of $\mathbb{R}^2$ for every $T > 0$ (the windows $\sin(a\pi(T-\cdot))$ and $\cos(a\pi(T-\cdot))$ are linearly independent on any interval), so this one-mode problem is exactly controllable for every horizon, but $\|q_0\| = O(T^{3/2})$ and $\|q_1\| = O(T^{1/2})$ as $T \to 0$, so the control needed to reach a given target grows like $T^{-3/2}$ and the energy budget becomes the binding constraint. For the full wave equation with a control acting on a strict subregion the reachable set itself depends on $T$ through the geometric control condition (on $(0,1)$, a time of order twice the distance the waves must travel), and the corresponding observability constants degenerate as $T$ approaches the critical time (Zuazua, 2005); the certificate (25) is unaffected by this (it is a statement about $\|u - v\|$ for the given $c$), but its right-hand side $R$ grows linearly in $T$ for a residual of fixed size: hyperbolic problems provide no contraction, so certified accuracy degrades with the horizon, a property of the PDE rather than of the certification technique; long horizons require restarts or sharper flux-based majorants.

### D.7 C5: NONLINEAR ADVECTIVE (STEADY BURGERS)

As for C3, Theorem 7 does not apply (Remark 3).

**Theorem 12** (Conditional state certificate)**.** *Let $-\nu u'' + (u^2/2)' = z(c)$ on $(0,1)$, $\nu > 0$, $u \in H_0^1$, and let $v \in H_0^1$ have residual $r(v;c) := \nu v'' - vv' + z(c)$. Put $M := \|u + v\|_{L^\infty}$ and $\kappa := \nu - \frac{1}{2}MC_{\mathrm{F}}$, $C_{\mathrm{F}} = 1/\pi$. If $\kappa > 0$ then, with $e = u - v$,*

$$\|e'\| \le \frac{C_{\mathrm{F}}\,\|r\|}{\kappa}, \qquad \|e\| \le \frac{C_{\mathrm{F}}^2\,\|r\|}{\kappa}, \qquad \|e\|_\infty \le \frac{C_{\mathrm{F}}\,\|r\|}{2\kappa}. \tag{26}$$

*Proof.* For $\psi \in H_0^1$: $\nu\langle e', \psi'\rangle + \langle \frac{1}{2}(u^2 - v^2)', \psi\rangle = \langle r, \psi\rangle$. Taking $\psi = e$ and integrating the advection term by parts, $\nu\|e'\|^2 - \int_0^1 \frac{u+v}{2}\, e\, e'\, dx = \langle r, e\rangle$. With $\|(u+v)/2\|_\infty = M/2$ and Lemma 6, $|\int \frac{u+v}{2} ee'| \le \frac{M}{2}C_{\mathrm{F}}\|e'\|^2$; hence $\kappa\|e'\|^2 \le \langle r, e\rangle \le C_{\mathrm{F}}\|r\|\,\|e'\|$, which is (26) after division and Lemmas 6 and 7. □

**Lemma 9** (Burgers: a-priori bound, uniqueness, coercivity)**.** *Let $\nu > 0$, $z \in L^2(0,1)$, $C_{\mathrm{F}} = 1/\pi$ and $z_{\mathrm{crit}} := 2\nu^2/C_{\mathrm{F}}^2$. (i) $-\nu u'' + (u^2/2)' = z$ has a weak solution $u \in H_0^1(0,1)$. (ii) Every weak solution satisfies $\nu\|u'\| \le C_{\mathrm{F}}\|z\|$ and $\|u\|_{L^\infty} \le U_a(z) := C_{\mathrm{F}}\|z\|/(2\nu)$. (iii) If $\|z\| < z_{\mathrm{crit}}$ the solution is unique, and with $\kappa_{\min}(z) := \nu(1 - \|z\|/z_{\mathrm{crit}}) > 0$*

$$\nu\|w'\|^2 + \langle (uw)', w\rangle \;\ge\; \kappa_{\min}(z)\,\|w'\|^2 \qquad \forall w \in H_0^1.$$

*Hence $w \mapsto -\nu w'' + (uw)'$ is an isomorphism $H_0^1 \to H^{-1}$ and $z \mapsto u$ is $C^\infty$ from $\{\|z\| < z_{\mathrm{crit}}\} \subset L^2$ to $H_0^1$.*

*Proof.* (ii) Testing with $u$, $\int_0^1 (u^2/2)' u\, dx = [u^3/3]_0^1 = 0$, so $\nu\|u'\|^2 = \langle z, u\rangle \le C_{\mathrm{F}}\|z\|\,\|u'\|$; then Lemma 7. (i) Galerkin approximation in the sine basis and Brouwer's theorem, as for the stationary Navier–Stokes equations (Temam, 1977, Ch. II, Lemma 1.4): by (ii), $\langle F_n(u), u\rangle = \nu\|u'\|^2 - \langle z, u\rangle > 0$ on a large sphere; the bound is uniform in $n$, and the quadratic term passes to the limit by the compact embedding $H_0^1(0,1) \hookrightarrow C([0,1])$. (iii) For two solutions and $e = u_1 - u_2$, testing the difference with $e$ gives $\nu\|e'\|^2 = \int \frac{u_1+u_2}{2} ee' \le U_a C_{\mathrm{F}}\|e'\|^2$ by (ii), and $\nu - C_{\mathrm{F}}U_a = \kappa_{\min} > 0$ forces $e = 0$. The same bound with $\|u\|_\infty \le U_a$ gives the coercivity; the isomorphism is Lax–Milgram and the smoothness of $z \mapsto u$ the implicit function theorem, as in the proof of Proposition 2. □

The margin $\kappa$ involves the unknown $u$ through $M$; the next theorem replaces it by a computable quantity (using the a-priori bound of Lemma 9), the validity flag of Assumption 2(ii) for C5.

**Theorem 13** (Verified fixed-point closure). *Let $U := \|v\|_{L^\infty}$, $\rho := \|r(v;c)\|$, $\kappa(d) := \nu - (U + \frac{d}{2})C_{\mathrm{F}}$, and consider $d\,\kappa(d) = \frac{1}{2}C_{\mathrm{F}}\rho$, i.e.*

$$\tfrac{C_{\mathrm{F}}}{2}d^2 - (\nu - UC_{\mathrm{F}})d + \tfrac{1}{2}C_{\mathrm{F}}\rho = 0, \qquad \mathcal{D} := (\nu - UC_{\mathrm{F}})^2 - C_{\mathrm{F}}^2\rho. \tag{27}$$

*Assume (V1) $\nu - UC_{\mathrm{F}} > 0$ and $\mathcal{D} \geq 0$, with $d_- := (\nu - UC_{\mathrm{F}} - \sqrt{\mathcal{D}})/C_{\mathrm{F}}$ the smaller root (so $\kappa(d_-) > 0$); (V2′) $U + U_a(z(c)) < d_+$, where $d_+ := (\nu - UC_{\mathrm{F}} + \sqrt{\mathcal{D}})/C_{\mathrm{F}}$ is the larger root and $U_a$ is the bound of Lemma 9. Then every weak solution $u$ satisfies $\|e\|_\infty \leq d_-$, and (26) holds with $\kappa$ replaced by the computable $\kappa(d_-) > 0$.*

*Proof.* Let $u$ be any weak solution and $d^* := \|u - v\|_\infty$. By Lemma 9(ii), $d^* \leq U + U_a < d_+ \leq d_- + d_+$, hence $\kappa(d^*) > 0$ (the sum of the roots is $2(\nu - UC_{\mathrm{F}})/C_{\mathrm{F}}$, the zero of $\kappa$). Since $M = \|u + v\|_\infty \leq 2U + d^*$, the margin of Theorem 12 is $\geq \kappa(d^*) > 0$, and (26) with Lemma 7 gives $d^*\kappa(d^*) \leq \frac{1}{2}C_{\mathrm{F}}\rho$, i.e. $\frac{C_{\mathrm{F}}}{2}d^{*2} - (\nu - UC_{\mathrm{F}})d^* + \frac{1}{2}C_{\mathrm{F}}\rho \geq 0$, so $d^* \leq d_-$ or $d^* \geq d_+$; the latter is excluded. As $\kappa$ is nonincreasing, the margin is $\geq \kappa(d^*) \geq \kappa(d_-) = \frac{1}{2}\big((\nu - UC_{\mathrm{F}}) + \sqrt{\mathcal{D}}\big) > 0$. □

(V2′) is a computable inequality; it is not binding for small $\rho$, since $d_+ \to 2(\nu - UC_{\mathrm{F}})/C_{\mathrm{F}}$ as $\rho \to 0$.

**Theorem 14** (Linearized (tangent) certificate). *Let $w$ solve $-\nu w'' + (uw)' = \varphi_j$ in $H_0^1$ and let $\omega \in H_0^1$ have tangent residual $r_V(\omega; v) := \nu\omega'' - (v\omega)' + \varphi_j$. Put $\kappa_{\mathrm{lin}} := \nu - \|u\|_{L^\infty} C_{\mathrm{F}}$ and $\underline{\kappa}_{\mathrm{lin}} := \max\{\kappa_{\min}(z(c)),\ \nu - (\|v\|_\infty + d_-)C_{\mathrm{F}}\}$; the first term counts only if $\|z(c)\| < z_{\mathrm{crit}}$ (Lemma 9), the second is computable via Theorem 13, and both are lower bounds for $\kappa_{\mathrm{lin}}$. If $\underline{\kappa}_{\mathrm{lin}} > 0$ then, whenever $\|u - v\| \leq \mu$,*

$$\begin{aligned} \|(w - \omega)'\| &\leq \frac{1}{\underline{\kappa}_{\mathrm{lin}}}\big(C_{\mathrm{F}}\,\|r_V(\omega; v)\| + \|\omega\|_{L^\infty}\,\mu\big), \\ \|w - \omega\| &\leq \frac{C_{\mathrm{F}}}{\underline{\kappa}_{\mathrm{lin}}}\big(C_{\mathrm{F}}\,\|r_V(\omega; v)\| + \|\omega\|_{L^\infty}\,\mu\big). \end{aligned} \tag{28}$$

*The validity flag for C5 tests $\nu - UC_{\mathrm{F}} > 0$, $\mathcal{D} \geq 0$, (V2′) and $\underline{\kappa}_{\mathrm{lin}} > 0$.*

*Proof.* With $\epsilon := w - \omega$ and $(uw)' - (v\omega)' = (u\epsilon)' + ((u - v)\omega)'$, for $\psi \in H_0^1$: $\nu\langle\epsilon', \psi'\rangle + \langle(u\epsilon)', \psi\rangle = \langle r_V, \psi\rangle - \langle((u - v)\omega)', \psi\rangle$. Take $\psi = \epsilon$: integrating by parts, $|\langle(u\epsilon)', \epsilon\rangle| = |\int u\epsilon\epsilon'| \leq \|u\|_\infty C_{\mathrm{F}}\|\epsilon'\|^2$, $|\langle((u - v)\omega)', \epsilon\rangle| = |\int(u - v)\omega\epsilon'| \leq \|\omega\|_\infty \|u - v\|\,\|\epsilon'\|$ and $|\langle r_V, \epsilon\rangle| \leq C_{\mathrm{F}}\|r_V\|\,\|\epsilon'\|$. Collecting gives the claim; the two lower bounds for $\kappa_{\mathrm{lin}}$ use $\|u\|_\infty \leq U_a$ (Lemma 9) and $\|u\|_\infty \leq \|v\|_\infty + d_-$ (Theorem 13), and the $L^2$ bound follows from Lemma 6. □

For fixed $v$, $\rho(c') \leq \rho(c) + \sigma_B\,\|c' - c\|$, and $\rho \mapsto C_{\mathrm{F}}^2\rho/\kappa(d_-(\rho))$ is nondecreasing, so Assumption 2(iii) holds for C5 with $\varsigma$ given by this map evaluated at $\rho(c) + \sigma_B\,\|c' - c\|$.

### D.8 C6: POINTWISE STATE CONSTRAINTS

C6 is an *extension* of the framework, not an instance of its full theory: the state constraint makes the feasible set depend on the state, whereas the theory of Section 4 needs one fixed convex feasible set (Lemma 2, Theorem 4). What is proved and certified for C6 is feasibility of every returned point (Theorem 15) and descent of every accepted step (Theorem 3); no criticality claim is made. Let $\Omega_{\mathrm{sc}} \subseteq [0, 1]$ be a closed set (in the experiment $\Omega_{\mathrm{sc}} = [0.4, 0.6]$) on which the state is required to satisfy $u(x) \leq u_{\max}$; the feasible set of the true problem is $\mathcal{C} = \{c \in B : |c|^2 \leq P_{\max},\ u(x; c) \leq u_{\max}\ \forall x \in \Omega_{\mathrm{sc}}\}$, where $B = \prod_j [c_j^-, c_j^+]$ is the box of coefficient bounds of Table 5 (Appendix F.1) which is convex because $c \mapsto u(x; c)$ is affine.

**Theorem 15** (Certified feasibility). *Let $-u'' = z(c)$ on $(0, 1)$, $u \in H_0^1$, $z(c)$ affine in $c$, and let $v \in H_0^1$ have residual $r(v; c) = v'' + z(c)$. Define the computable margin $\varepsilon_\infty(v; c) := \frac{1}{2}C_{\mathrm{F}}\,\|r(v; c)\| \geq \|u - v\|_{L^\infty}$, $C_{\mathrm{F}} = 1/\pi$. If the tightened surrogate constraint*

$$v(x) + \varepsilon_\infty(v; c) \leq u_{\max} \qquad \text{for all } x \in \Omega_{\mathrm{sc}} \tag{29}$$

*holds, then the true state satisfies $u(x) \leq u_{\max}$ on $\Omega_{\mathrm{sc}}$. Assumption 2 holds with the one-dimensional C1 functionals, and $\varepsilon_\infty(v; c') \leq \varepsilon_\infty(v; c) + \frac{1}{2}C_{\mathrm{F}}\sigma_B\,\|c' - c\|$.*

*Proof.* By the one-dimensional case of Theorem 8, $\|(u-v)'\| \le C_{\mathrm{F}}\|r\|$, and by Lemma 7 $\|u-v\|_\infty \le \frac{1}{2}C_{\mathrm{F}}\|r\| = \varepsilon_\infty$. Hence for $x \in \Omega_{\mathrm{sc}}$, $u(x) \le v(x) + \varepsilon_\infty \le u_{\max}$ by (29). □

The implementation solves the inner subproblem with $\varepsilon_\infty$ frozen at the value certified at the center and re-certifies the returned point with a fresh margin; only a point that passes the fresh test (29) is claimed feasible, and by Theorem 15 it then is. Anchored margins $\min\{\varepsilon_\infty(u_\theta(c);c),\ \frac{1}{2}C_{\mathrm{F}}\|r(u_h;c)\| + \|u_h - u_\theta(c)\|_\infty\}$ are valid by the triangle inequality and Lemma 7.

**Scope of the C6 claim.** The algorithm projects onto the certified inner approximation $\mathcal{C}_\theta := \{c : u_\theta(x;c) + \varepsilon_\infty(c) \le u_{\max} \text{ on } \Omega_{\mathrm{sc}}\} \cap (\text{box} \cap \text{power ball}) \subseteq \mathcal{C}$, which is convex only if $u_\theta$ is affine in $c$ there, while convexity of the feasible set enters the theory through Lemma 2 and Theorem 4. No criticality claim is therefore made for C6; the run's $\chi_\theta + \widehat{\eta}_g$ relative to $\mathcal{C}_\theta$ ($8.0 \cdot 10^{-5}$) is a diagnostic. What is certified is feasibility of every returned point and descent of every accepted step. The optimality gap is *not* certified by the method: it is verified offline with the exact solution map of the one-dimensional problem (piecewise polynomial in $x$, affine in $c$), which the majorants of Theorem 15 do not enter. With this map the exact quadratic objective is minimized over the node relaxation of $\mathcal{C}$ (state constraint at 2001 nodes of $\Omega_{\mathrm{sc}}$, box, power ball; a superset of $\mathcal{C}$, so its minimum is a lower bound on $\min_{\mathcal{C}} \widehat{J}$), and weak duality at computed multipliers, evaluated in exact arithmetic, certifies that lower bound. The same computation applies to any feasible control: it gives $\widehat{J}(c_{\mathrm{ret}}) - \min_{\mathcal{C}} \widehat{J} \le 5.73 \cdot 10^{-6}$ for the SA-CPITR point, $5.43 \cdot 10^{-6}$ for the CPITR point and $3.70 \cdot 10^{-4}$ for the (feasible but uncertified) frozen surrogate, with dual and primal values agreeing to $8 \cdot 10^{-16}$. On feasibility: the runs use a margin larger by $\sqrt{2}$ than Theorem 15; on the validation grid the solver-free point of the campaign misses that margin by $2 \cdot 10^{-7}$ (surrogate maximum $0.059863$ plus $1.373 \cdot 10^{-4}$) but passes the sharp margin with slack $4 \cdot 10^{-5}$. Over the 26 non-frozen C6 runs of the ablation, 24 pass the sharp test and all 26 are truly feasible (smallest slack $1.2 \cdot 10^{-4}$). The exact criticality of the SA-CPITR point relative to $\mathcal{C}$ is $2.0 \cdot 10^{-3}$, twenty times $\mathrm{tol}$: the conservatism of the inner set shows in the criticality and is negligible in the objective.

Lemma 7 has no analogue for $d \ge 2$ ($H^1 \not\hookrightarrow L^\infty$); for $d \le 3$ one may use $H^2 \hookrightarrow C^0$, which requires a computable $H^2$ bound available only under additional structure (Grisvard, 2011); classical theory for state-constrained elliptic control is (Casas, 1986; Hintermüller & Kunisch, 2006).

### D.9 Anchor-growth constants and existence of certifying surrogates

Table 4: Anchor-growth constants of Assumption 2(iii). They are diagnostic: the algorithm re-evaluates $\mathcal{M}(u_h;c)$ at the query control, so no growth constant enters any certificate.

| case | $C_{\mathrm{F}}$ | state $C_B$ | tangent growth | conditional? |
|---|---|---|---|---|
| C1 | $1/(\pi\sqrt{2})$ | $C_{\mathrm{F}}^2\sigma_B$ | none | no |
| C2 | $1/\pi$ | $(C_{\mathrm{F}}^2/D)\sigma_B$ | none | no |
| C3 | $1/\pi$ | $C_{\mathrm{F}}^2\sigma_B$ | via $\mu$ only | no |
| C4 | $1/\pi$ | $C_B^{(m)}$ per component | none | no |
| C5 | $1/\pi$ | monotone formula (Appendix D.7) | via $\kappa_{\mathrm{lin}}, \mu$ | yes |
| C6 | $1/\pi$ | $C_{\mathrm{F}}^2\sigma_B$ (+ margin) | none | no |

No guarantee is claimed that training produces a network satisfying (A1), and PINN training is known to fail on stiff or advection-dominated regimes (Krishnapriyan et al., 2021; Wang et al., 2022). (A1) is therefore a hypothesis; its violation is detectable at run time by Theorem 16, and there the solve events of Definition 5 take over.

## E The full algorithm, the events, and the implementation layer

### E.1 The two solve events

**Definition 5** (Events, roles, discipline, budget)**.** Let $d_k := \widehat{\eta}_g(c_k)/\tau_k$ be the *deficit* of the gradient certificate, where $\tau_k$ is the current demand: the right-hand side of (6) outside tolerance, and $\mathrm{tol} -$

$\chi_\theta(c_k)$ once $\chi_\theta(c_k) \le \frac{1}{2}\mathrm{tol}$. A rejection is *certificate-width dominated* if $\rho_k^{\mathrm{raw}} := (J_\theta(c_k) - J_\theta(c_k + s_k))/\mathrm{pred}_k \ge \eta_1 > \hat{\rho}_k$. Solves are spent only on two events, subject to an event budget $N_{\max}$ (counted in anchor events; the PDE solves one event costs are stated in Section 5.1) and a gain factor $\vartheta \in (0, 1)$: *$E_J$ (value-limited → state anchor)*: $n_{\mathrm{ns}}$ consecutive certificate-width-dominated rejections; the solve is committed only if it sharpens the anchored state majorant by the factor $\vartheta$, else the role is disabled at that center. *$E_g$ (gradient-limited → gradient anchor)*: $d_k > 1$ after the refinement budget for $n_g$ consecutive iterations without a decrease of $d_k$ (refinement is losing the race against the demand, Theorem 16); the anchor is evaluated first in the anchored terminal test $\chi_a + \eta_g^a(c_k) \le \mathrm{tol}$ (Theorem 4), with mesh refinements each required to reduce $\eta_g^a$ by $\vartheta$, and if the test fails outside tolerance the algorithm enters mode A. Events never fire once $N_{\max}$ events are spent; the run then continues in surrogate mode with the (dominant) anchored certificates and reports its best-certified iterate.

### E.2 The full algorithm

Algorithm 1 states SA-CPITR in full: the offline training, the outer trust-region loop of Section 3.2, the mode switch of Section 3.3 and the two events of Definition 5; the panel letters in the margin comments refer to Figure 1. To state both modes in one loop we write, at the center $c_k$,

$$(m_k,\ g_k^m,\ \chi_k^m,\ \eta_{g,k}^m) := \begin{cases} \big(J_\theta,\ g_\theta(c_k),\ \chi_\theta(c_k),\ \widehat{\eta}_g(c_k)\big) & \text{in surrogate mode,} \\ \big(m_k^A,\ g_a,\ \chi_a,\ \eta_g^a(c_k)\big) & \text{in mode A, model (9),} \end{cases} \tag{30}$$

for the current model, its gradient, its criticality $\chi_k^m = \|P_{\mathcal{C}}(c_k - g_k^m) - c_k\|$ and its gradient certificate, so that the terminal test, the gradient gate, the deficit and the subproblem read the same in both modes. The value gate and the certified ratio always use the surrogate values $J_\theta$ with the anchored certificates $\widehat{\eta}_J$. Every quantity in the algorithm is computable; the only calls to a PDE solver are the two lines marked SOLVE.

**Reading the algorithm.** Lines 4–6 of Algorithm 1 are the certificate layer (Section 3.1); the terminal test of line 6 is the only claim a run makes, an inequality between computed numbers at the returned point (Theorem 4). Lines 7–9 are the gradient gate with its refinement loop and the deficit that drives $\mathrm{E}_g$; line 10 and the first procedure of Algorithm 2 are $\mathrm{E}_g$ itself. Line 11 computes the step with the projected Cauchy safeguard, so Theorem 6 holds whatever the NLP solver returned; lines 12–14 are the value gate and the certified ratio (7), and lines 15–21 the acceptance test with the radius update and $\mathrm{E}_J$ (line 19 and the second procedure of Algorithm 2). Theorem 3 is the statement about line 16, Theorem 5 about the whole loop, and Theorem 4 about lines 6, 6 and 22. The reference implementation adds the devices (H1)–(H6) of Appendix E.4, none of which touches a certificate (Proposition 4).

**Algorithm 1** SA-CPITR: the outer loop. Trust-region constants $0 < \eta_1 \le \eta_2 < 1$, $0 < \gamma_s < 1 < \gamma_e$, Cauchy fraction $\beta \in (0, 1]$, radius cap $\Delta_{\max} \in [\Delta_0, \infty]$, gate constants $\kappa_g, \kappa_J$ as in (8); event parameters $n_g, n_{\rm ns}, n_{\rm hold} \in \mathbb{N}$, gain factor $\vartheta \in (0, 1)$, event budget $N_{\max}$ (anchor events), refinement budget $n_{\rm ref}$ per gate. The two events are the procedures of Algorithm 2, the only lines that call a PDE solver.

**Input:** initial control $c_0 \in \mathcal{C}$, radius $\Delta_0 > 0$, tolerance $\mathrm{tol} > 0$
**Output:** $c_{\rm ret}$ with certified $\chi(c_{\rm ret}) \le \mathrm{tol}$ (Theorem 4), or the best certified iterate and an abstention

1: **offline:** train the parametric surrogate $(u_\theta, \{w_{\theta,j}\})$ on state and tangent residuals (Appendix F.2)
2: $\mathcal{A} \leftarrow \emptyset$ (anchors); $N \leftarrow 0$ (events spent); mode $\leftarrow$ surrogate; $\mathcal{S} \leftarrow \emptyset$ (snapshots); $n_{\rm race}, n_{\rm wd} \leftarrow 0$
3: **for** $k = 0, 1, 2, \ldots$ **do**
*— certificates at the center (C–D) —*
4: anchored majorants (4)–(11) at $c_k$ over $\mathcal{A}$; assemble $\widehat{\eta}_J(c_k), \widehat{\eta}_g(c_k)$ (Theorems 1 and 2); form the current model (30)
5: $\mathcal{S} \leftarrow \mathcal{S} \cup \{(c_k,\ \chi_k^m + \eta_{g,k}^m)\}$ ▷ store every iterate with its bound
6: **if** $\chi_k^m + \eta_{g,k}^m \le \mathrm{tol}$ **then return** $c_k$: certified $\chi(c_k) \le \chi_k^m + \eta_{g,k}^m \le \mathrm{tol}$ ▷ terminal test, Theorem 4
*— gradient gate at the center and refinement (E) —*
7: **while** $\eta_{g,k}^m > \max\{\kappa_g \min(\chi_k^m, \Delta_k),\ \kappa_g \mathrm{tol}\}$ **for at most** $n_{\rm ref}$ rounds **do** ▷ gate (6) fails
8: refine the surrogate at $c_k$ (training rounds); re-evaluate line 4
9: $d_k \leftarrow \eta_{g,k}^m / \tau_k$ with the demand $\tau_k$ of Definition 5; $n_{\rm race} \leftarrow n_{\rm race} + 1$ **if** $d_k > 1$ **and** $d_k \ge d_{k-1}$, **else** 0 ▷ Theorem 16
10: **if** $n_{\rm race} \ge n_g$ **and** $N < N_{\max}$ **and** mode = surrogate **then** GRADIENTANCHOR
▷ event $\mathrm{E}_g$, Algorithm 2: may return certified or enter mode A
*— trust-region step on the current model (F) —*
11: $s_k \leftarrow$ the better of an approximate minimizer of $m_k(c_k + s)$ over $\{c_k + s \in \mathcal{C} : \|s\| \le \Delta_k\}$ (general NLP solver) and the best point of $m_k$ on the projected Cauchy arc within the radius;
$\mathrm{pred}_k \leftarrow m_k(c_k) - m_k(c_k + s_k)$ ▷ $\ge \beta$ [Cauchy decrease], Theorem 6
*— value gate at both endpoints and certified ratio (E, G) —*
12: **while** $\max\{\widehat{\eta}_J(c_k), \widehat{\eta}_J(c_k + s_k)\} > \kappa_J \mathrm{pred}_k$ **for at most** $n_{\rm ref}$ rounds **do** ▷ value gate fails
13: refine the surrogate at the failing endpoint(s); re-evaluate $\widehat{\eta}_J$
14: $\hat{\rho}_k \leftarrow \big[J_\theta(c_k) - J_\theta(c_k + s_k) - \widehat{\eta}_J(c_k) - \widehat{\eta}_J(c_k + s_k)\big]/\mathrm{pred}_k$;
$\rho_k^{\rm raw} \leftarrow \big[J_\theta(c_k) - J_\theta(c_k + s_k)\big]/\mathrm{pred}_k$ ▷ (7), Definition 5
15: **if** $\hat{\rho}_k \ge \eta_1$ **then** ▷ accept: $\widehat{J}(c_{k+1}) \le \widehat{J}(c_k) - \eta_1 \mathrm{pred}_k$, Theorem 3
16: $c_{k+1} \leftarrow c_k + s_k$; $\Delta_{k+1} \leftarrow \min\{\gamma_e \Delta_k, \Delta_{\max}\}$ **if** $\hat{\rho}_k \ge \eta_2$, **else** $\Delta_k$; mode $\leftarrow$ surrogate;
$n_{\rm race}, n_{\rm wd} \leftarrow 0$ ▷ mode A ends with the first accepted step
17: **else** ▷ reject: $c_{k+1} \leftarrow c_k$; anchors persist, no re-solve
18: $n_{\rm wd} \leftarrow n_{\rm wd} + 1$ **if** $\rho_k^{\rm raw} \ge \eta_1$ (certificate-width dominated), **else** 0
19: **if** $n_{\rm wd} \ge n_{\rm ns}$ **and** $N < N_{\max}$ **and** the state-anchor role is enabled at $c_k$ **then** STATEANCHOR
▷ event $\mathrm{E}_J$, Algorithm 2
20: **else if** $0 < n_{\rm wd} \le n_{\rm hold}$ **then** $\Delta_{k+1} \leftarrow \Delta_k$; refine the surrogate at $c_k$ ▷ radius hold (H5)
21: **else** $\Delta_{k+1} \leftarrow \gamma_s \Delta_k$
22: **return** the best entry of $\mathcal{S}$ (an abstention if its bound exceeds tol) ▷ budget exhausted, or (H2)–(H3)

**Algorithm 2 The two anchor events.** Entered at the current center $c_k$ of Algorithm 1; the lines marked SOLVE are the only PDE-solver calls (one state solve, or a state and $p$ tangent solves, per event). Solver mesh levels $\iota = 0, 1, \ldots$

**procedure** GRADIENTANCHOR ▷ $\mathrm{E}_g$: the gradient certificate is binding (Definition 5)
SOLVE at $c_k$, level $\iota$: $(u_h, \{w_{h,j}\})$; $\mathcal{A} \leftarrow \mathcal{A} \cup \{a\}$, $a$ the anchor at $c_k$ (Definition 3); $N \leftarrow N + 1$
$(g_a, \eta_g^a(c_k))$ by (5); $\chi_a \leftarrow \|P_{\mathcal{C}}(c_k - g_a) - c_k\|$
**while** $\chi_a + \eta_g^a(c_k) > \mathrm{tol}$ **and** $\chi_\theta(c_k) \le \frac{1}{2}\mathrm{tol}$ **and** a level $\iota + 1$ exists **do** ▷ inside tolerance
re-solve on level $\iota + 1$; **commit** only if $\eta_g^a(c_k)$ drops by the factor $\vartheta$, **else break** ▷ $\vartheta$-discipline
**if** $\chi_a + \eta_g^a(c_k) \le \mathrm{tol}$ **then return** $c_k$ from Algorithm 1: certified $\chi(c_k) \le \chi_a + \eta_g^a(c_k) \le \mathrm{tol}$
**else if** $\chi_\theta(c_k) > \frac{1}{2}\mathrm{tol}$ **then** mode $\leftarrow$ A with the model (9) ▷ outside tolerance, Section 3.3
$n_{\rm race} \leftarrow 0$ ▷ no switch inside tolerance: the anchored certificates stay in use

**procedure** STATEANCHOR ▷ $\mathrm{E}_J$: the objective certificate is binding (Definition 5)
SOLVE at $c_k$: $u_h$; $a \leftarrow$ state-only anchor at $c_k$ (Definition 3); $N \leftarrow N + 1$
**commit** $\mathcal{A} \leftarrow \mathcal{A} \cup \{a\}$ if it reduces $\widehat{\mu}(c_k)$ by the factor $\vartheta$, **else** disable the role at $c_k$ ▷ gain discipline
$\Delta_{k+1} \leftarrow \Delta_k$; $n_{\rm wd} \leftarrow 0$

### E.3 THE RACE TEST AND ITS REGIME GUARD

The event $\mathrm{E}_g$ of Definition 5 fires when the gradient certificate is losing against its demand. Outside tolerance the demand is the gate threshold $\kappa_g \min(\chi_\theta, \Delta_k)$, which shrinks on every rejection; the following result makes "the gate is losing against the shrinking radius" precise and shows that the test cannot fire while refinement is winning.

**Theorem 16** (Race test)**.** *Consider the regime $\mathcal{R}_k := \{\Delta_k < \chi_\theta(c_k)$, gate (6) violated after the refinement budget, step rejected$\}$, and on $\mathcal{R}_k$ let $d_k = \widehat{\eta}_g(c_k)/(\kappa_g \Delta_k)$. Then: (i) on $\mathcal{R}_k$ the gate threshold decays geometrically, $\kappa_g \Delta_{k+1} = \gamma_s \kappa_g \Delta_k$; (ii) if refinement productivity satisfies $\widehat{\eta}_g(c_{k+1}) \geq (1-\delta)\widehat{\eta}_g(c_k)$ with $\delta < 1 - \gamma_s$, then $d_{k+1} > d_k$; conversely, if $d_{k+1} \leq d_k$ then $\widehat{\eta}_g(c_{k+1}) \leq \gamma_s \widehat{\eta}_g(c_k)$, i.e. refinement reduced the certificate at least as fast as the demand; (iii) consequently the test "flag after $n_g$ consecutive increases of $d_k$ within $\mathcal{R}$" never fires while refinement is winning in the sense of (ii).*

*Proof.* (i) In $\mathcal{R}_k$ the step is rejected, so $\Delta_{k+1} = \gamma_s \Delta_k$, and the binding threshold is $\kappa_g \Delta$ because $\Delta_k < \chi_\theta(c_k)$. (ii) $d_{k+1}/d_k = (\widehat{\eta}_g(c_{k+1})/\widehat{\eta}_g(c_k))/\gamma_s$; if the ratio of certificates is $\geq 1 - \delta > \gamma_s$ this exceeds 1; the converse is the contrapositive. (iii) is immediate. □

Outside $\mathcal{R}_k$ the binding threshold is $\kappa_g \chi_\theta(c_k)$, and along a convergent subsequence $d_k \to \infty$ even when $\widehat{\eta}_g(c_k) \downarrow 0$ (take $\widehat{\eta}_g = 1/k$, $\chi_\theta = 1/k^2$); a test without the restriction to $\mathcal{R}_k$ would fire when the optimizer succeeds.

### E.4 THE IMPLEMENTATION LAYER

The implementation has four triggers and three anchor roles; they are the two events of Definition 5 split by regime:

| implementation | regime | this paper | anchor |
|---|---|---|---|
| E2 noise streak → R1 recentering | any | $\mathrm{E}_J$ | state anchor |
| E4 terminal shortfall → R3 terminal certification | $\chi_\theta \leq \frac{1}{2}\mathrm{tol}$ | $\mathrm{E}_g$ inside tolerance | gradient anchor, terminal test |
| E1 lost race → R2 gradient anchoring (mode A) | $\Delta_k < \chi_\theta$, Theorem 16 | $\mathrm{E}_g$ outside tolerance | gradient anchor, mode A |
| E3 stalled radius collapse → R2 + radius restart | $\Delta_k$ below a floor, $\chi_\theta + \widehat{\eta}_g > \mathrm{tol}$ | $\mathrm{E}_g$ outside tolerance | gradient anchor, mode A |

The reference implementation contains six devices the theory does not model. None of them can invalidate a certificate (Proposition 4), because the guarantee attached to a returned point is the a posteriori terminal test of Theorem 4, which is indifferent to the history that produced the point.

- (H1) Terminal-refinement mode. When $\chi_\theta(c_k) \leq \frac{1}{2}\mathrm{tol}$ but $\chi_\theta(c_k) + \widehat{\eta}_g(c_k) > \mathrm{tol}$, the iteration performs refinement only and computes no step (a step cannot reduce $\widehat{\eta}_g$); released as soon as $\chi_\theta$ rises above the threshold.
- (H2) Plateau termination. Inside (H1), if $\widehat{\eta}_g$ improves by less than a fixed relative amount over a fixed number of refinement iterations, the run stops and reports a capacity limit.
- (H3) Futility termination. Outside (H1), the deficit is also used to stop: if the projected number of iterations needed to close the gate exceeds a budget for several consecutive iterations, the run stops. Theorem 16 licenses $d_k$ only as a trigger; its use as a stopping rule is a heuristic.
- (H4) Symmetric scheduler. When the gate is met but $\widehat{\eta}_g \geq \varrho\, \chi_\theta$ for a fixed $\varrho = O(1)$, a bounded number of additional refinement rounds is performed before the step.
- (H5) Noise-retry radius hold. On a certificate-width-dominated rejection the radius is held rather than shrunk, for at most $n_{\mathrm{hold}}$ consecutive iterations, and refinement is performed instead.
- (H6) Scale-relative stopping. An optional second stopping rule $\chi_\theta + \widehat{\eta}_g \leq \mathrm{tol}_{\mathrm{rel}} S$ with a certified curvature scale $S$, disabled by default.

**Proposition 4** (The implementation layer does not affect certificates)**.** *Let $c_{\mathrm{ret}}$ be returned by a run in which any subset of (H1)–(H6) is active, with reported guarantee $\chi_\theta(c_{\mathrm{ret}}) + \widehat{\eta}_g(c_{\mathrm{ret}}) \leq \mathrm{tol}$ (or its anchored form). Then $\chi(c_{\mathrm{ret}}) \leq \mathrm{tol}$ holds unchanged, and every accepted step still*

*satisfies Theorem 3. Moreover (H4) preserves every hypothesis of Theorem 5 (additional refinement only tightens certificates); (H5) preserves its conclusion provided the number of consecutive holds is bounded (the radius is then multiplied by $\gamma_s$ at least once every $n_{\mathrm{hold}} + 1$ rejections, so Step 2 of the proof applies); (H1)–(H3) and (H6) are stopping rules that terminate the sequence, so Theorem 5 neither applies nor is contradicted: a run that invokes them reports a certificate or an honest abstention, but no asymptotic claim.*

*Proof.* Theorem 4 is an inequality between quantities evaluated at $c_{\mathrm{ret}}$ and quantifies over no history; (H1)–(H4), (H6) alter only which iterations refine and when the loop exits; (H5) alters the radius update, which Theorem 3 does not involve. The statements about Theorem 5 are read off its proof in Appendix C.6. □

# F EXPERIMENTAL DETAILS

This appendix fixes the discrete data of the six cases (Appendix F.1), the surrogate and its training (Appendix F.2) and the certificate evaluation (Appendix F.3); it then gives the reference solvers, costs, protocol and solve events (Appendix F.4), the comparison with ground truth (Appendix F.5), the exact evaluation of the solver-side certificates (Appendix F.6), the observed curvature (Appendix F.7) and the fields (Appendix F.8).

## F.1 DISCRETE PROBLEM DATA

Throughout, $\Omega = (0,1)$ or $(0,1)^2$, $Q = \Omega \times (0,T)$, and $\|\cdot\|$ is the $L^2$ norm computed by the midpoint rule on the certification grid of Appendix F.3. All problems use $\mathrm{tol} = 10^{-4}$, $\eta_1 = 0.1$, $\eta_2 = 0.75$, $\gamma_s = 0.5$, $\gamma_e = 2$, hence $\kappa_g = \beta(1-\eta_2)/16 = 1.5625 \cdot 10^{-2}$ and $\kappa_J = (1-\eta_2)/8 = 3.125 \cdot 10^{-2}$ with $\beta = 1$.

No radius cap is used ($\Delta_{\max} = \infty$); the largest radius in the logged runs is 8 (C1, C2), 4 (C4, C5) and 2 (C3). C6 also reaches 8, so any $\Delta_{\max} \geq 8$ reproduces every campaign run; with such a finite cap the pair form of (A3) also holds for C3, whose set $S_0 + \bar{B}_{\Delta_{\max}}$ is then bounded (Proposition 3).

Table 5: **The six instantiations and their discrete data.** $p$ control coefficients, Tikhonov weight $\alpha$, initial radius $\Delta_0$, certification-grid parameter $n_{\mathrm{cert}}$. Only C3's basis is $L^2$-orthonormal; elsewhere certified radii are coefficient statements up to a fixed Gram factor.

| | PDE | $p$ | basis | $\alpha$ | constraints | $\Delta_0$ | $n_{\mathrm{cert}}$ |
|---|---|---|---|---|---|---|---|
| C1 | Poisson, $(0,1)^2$ | 1 | single mode | 0 | box | 1.0 | 64 |
| C2 | heat, $Q$ | 2 | $\sin k\pi x$ | 0 | box | 1.0 | 64 |
| C3 | semilinear, $(0,1)$ | 4 | $\sqrt{2}\sin k\pi x$ | $10^{-2}$ | none | 0.5 | 256 |
| C4 | wave, $Q$ | 12 | cubic B-spline in $t$ | $10^{-3}$ | box, energy budget | 1.0 | 48 |
| C5 | Burgers, $(0,1)$ | 16 | cubic B-spline in $x$ | $10^{-3}$ | box, mean-zero, power | 1.0 | 192 |
| C6 | Poisson, $(0,1)$ | 10 | cubic B-spline in $x$ | $10^{-4}$ | box, power, state | 1.0 | 128 |

**C1 (linear elliptic, scalar control).** $-\Delta u = c\varphi$ on $(0,1)^2$, $\varphi = \sin \pi x \sin \pi y$, $u = 0$ on $\partial\Omega$; $J(c) = \frac{1}{2}\|u - u_d\|^2$ with $u_d$ the exact state at $c^\star = 7.5$; $c \in [1,10]$, $c_0 = 1$, $\alpha = 0$. The analytic solution $u = c\varphi/(2\pi^2)$ is used for validation only.

**C2 (parabolic, two-parameter inverse).** $\partial_t u - D\partial_{xx} u = c_1 \sin \pi x + c_2 \sin 2\pi x$ on $Q = (0,1)^2$, $D = 0.1$, homogeneous data; $J(c) = \frac{1}{2}\|u - u_d\|^2_{L^2(Q)}$ with the target generated at $c^\star = (5,2)$; $c \in [-1,7] \times [-1,4]$, $c_0 = (1,1)$, $\alpha = 0$. The Fourier solution of Theorem 9 is used for validation only. C1 and C2 were run with the objective $\|u - u_d\|^2$, twice the $J$ stated here; the objective values, gradients and certificates reported for C1 and C2 refer to this scaling (the bounds of Theorem 1 scale with the factor).

**C3 (semilinear).** $-u'' + u^3 = z$, $u(0) = u(1) = 0$, $z = \sum_{k=1}^{4} c_k \sqrt{2} \sin k\pi x$; $J(c) = \frac{1}{2}\|u - u_d\|^2 + \frac{\alpha}{2}|c|^2$, $u_d = 0.2 \sin \pi x$, $\alpha = 10^{-2}$, $c_0 = 0$, no constraints; offline controls from $[-2, 2]^4$.

**C4 (wave, energy budget).** $u_{tt} - a^2 u_{xx} = z(x,t) = g(t) \sin \pi x$ on $(0,1) \times (0,1)$, $a = 1$, $T = 1$, homogeneous data, $g(t) = \sum_{k=1}^{12} c_k B_k(t)$ in a cubic B-spline basis; $J(c) = \frac{1}{2}\|u(\cdot, T) - u_d\|^2 + \frac{1}{2}\|u_t(\cdot, T)\|^2 + \frac{\alpha}{2}|c|^2$, $u_d = 0.15 \sin \pi x$, $|c|^2 \le E_{\max} = 9$, $c_k \in [-4, 4]$, $\alpha = 10^{-3}$, $c_0 = 0$; offline controls from $[-2, 2]^{12}$. This is the classical "steer to a target and bring to rest" problem; its closed form is Remark 5.

**C5 (steady Burgers, active power budget).** $-\nu u'' + uu' = z$, $u(0) = u(1) = 0$, $\nu = 0.2$, $z = \sum_{k=1}^{16} c_k B_k$; $J(c) = \frac{1}{2}\|u - u_d\|^2 + \frac{\alpha}{2}|c|^2$, $u_d = 0.30 \sin \pi x$, $\alpha = 10^{-3}$, subject to $\int_0^1 z = 0$, $|c|^2 \le P_{\max} = 3$ (active at the optimum) and $c_k \in [-6, 6]$; $c_0 = 0$, offline controls from $[-3, 3]^{16}$. The certificate is valid whenever the computable flag of Theorems 13 and 14 holds; the run reports $\kappa_{\text{lin}}$. On the feasible set $\|z(c)\| \le 0.47 < z_{\text{crit}} = 2\nu^2 / C_{\text{F}}^2 = 0.79$ (Lemma 9), so $u(c)$ exists, is unique and uniformly coercive, with $\kappa_{\text{lin}} \ge 0.08$ on all of $\mathcal{C}$. The runs did not evaluate (V2′) of Theorem 13; evaluated after the run at every logged C5 center of both arms (with $U$ within $10^{-3}$ of the exact $\|u(c)\|_\infty$, since $d_- < 5 \cdot 10^{-4}$ there), $U + U_a(z(c)) \le 0.47$ against $d_+ \ge 0.96$.

**C6 (pointwise state constraint with certified margin).** $-u'' = z$, $u(0) = u(1) = 0$, $z = \sum_{k=1}^{10} c_k B_k$; $J(c) = \frac{1}{2}\|u - u_d\|^2 + \frac{\alpha}{2}|c|^2$, $u_d = 0.15 \sin \pi x$, $\alpha = 10^{-4}$, subject to $u(x) \le u_{\max} = 0.06$ for $x \in \Omega_{\text{sc}} = [0.4, 0.6]$, $|c|^2 \le 64$ and $c_k \in [-8, 8]$; $c_0 = 0$, offline controls from $[-4, 4]^{10}$. The state constraint is handed to the inner solver in the tightened form $u_\theta(x_j) + \varepsilon_\infty \le u_{\max}$ at the certification nodes $x_j \in \Omega_{\text{sc}}$ (Theorem 15). The runs use the constant $1/\sqrt{2}$ in place of the sharp $\frac{1}{2}$ of Lemma 7, so their margin $\varepsilon_\infty = C_{\text{F}}\|r_\theta\|/\sqrt{2}$ ($1.36 \cdot 10^{-4}$ at the SA-CPITR return) is $\sqrt{2}$ times that of Theorem 15, and the feasibility guarantee holds a fortiori. Theorem 15 asks for the tightened inequality on all of $\Omega_{\text{sc}}$, the campaign checks it at the $n_{\text{cert}}$ nodes. The exact solution map of C6 is piecewise polynomial, so the supremum over $\Omega_{\text{sc}}$ is computable by root finding: at the SA-CPITR control the true state's maximum is $0.0598538$ (at $x = 0.508$; slack $1.46 \cdot 10^{-4}$ to $u_{\max}$), $4 \cdot 10^{-9}$ above the 41-node maximum, and all three arms are feasible on the continuum.

### F.2 The parametric surrogate

**Architecture.** A single fully connected $\tanh$ MLP with 4 hidden layers and width 64 (C1–C3) or 128 (C4–C6) takes the coordinate and the control as input and returns the state; sensitivities are obtained by differentiating with respect to $c$. Weights use PyTorch's default initialization with seed 0. Essential conditions are imposed exactly by a smooth mask ($\mathcal{N}_\theta$ the raw MLP):

C1: $u_\theta(x, y; A) = x(1-x)y(1-y)\,\mathcal{N}_\theta(x, y, A/10)$,
C2: $u_\theta(x, t; A) = x(1-x)t\,\mathcal{N}_\theta(x, t, A/6)$,
C3, C6: $u_\theta(x; c) = x(1-x)\,\mathcal{N}_\theta(x, c/\lambda)$, $\lambda = 2$ (C3), $4$ (C6),
C4: $u_\theta(x, t; c) = x(1-x)t^2\,\mathcal{N}_\theta(x, t, c)$, C5: $u_\theta(x; c) = x(1-x)\,\mathcal{N}_\theta(x, c)$.

The mask $t^2$ in C4 enforces $u = u_t = 0$ at $t = 0$; the control rescalings keep the inputs $O(1)$.

**Training loss.** With $r(\cdot; c)$ the strong state residual and $r_{V,k} = \partial_{c_k} r$ the tangent residuals (both by automatic differentiation of the masked network; the identity $\partial_{c_k} r =$ tangent residual is exact), training minimizes the certificate-weighted residual loss

$$\mathcal{L}_\theta = p\,\frac{\overline{r^2}}{(\overline{r^2})_{\text{det}}^{1/2}} + \sum_{k=1}^{p} \frac{\overline{r_{V,k}^2}}{(\overline{r_{V,k}^2})_{\text{det}}^{1/2}}, \tag{31}$$

where $\overline{\cdot}$ is the empirical mean over the collocation batch and "det" marks a detached normalizer. The certificate $\eta_g$ is linear in the residual norms, and the gradient of $n^2/n_{\text{det}}$ equals that of $2n$, so (31) descends a weighted sum of norms rather than of squared norms (a flat sum of squares underweights small-but-binding terms near the certificate floor); the state norm feeds all $p$ tangent majorants through the coupling terms of (3), hence the weight $p$.

**Collocation, offline and refinement stages.** Collocation points ($n_{\mathrm{col}} = 1024$ per epoch) are redrawn uniformly at every epoch; in the offline stage the control is redrawn with the same batch size from the case's sampling distribution (the design box for C1, C2; the cube of radius $c_{\mathrm{rad}}$ otherwise). Offline: Adam, learning rate $10^{-3}$, 6000 epochs (C4: 3000), no validation split and no early stopping. Refinement (called when a gate fails) is per instance: the control is clamped at the requested $c$ and (31) is minimized over collocation points only, in two phases: (a) annealed Adam with persistent state, learning rate $\max(3 \cdot 10^{-4} \cdot 0.97^{\text{rounds}}, 10^{-5})$, 800 epochs per round (C4: 300); (b) an L-BFGS polish (strong Wolfe, history 40, 120 iterations; C4: 60) on a fixed collocation batch, kept only if it decreases the certified $\eta_g$ at the clamped control. The polish is therefore monotone in the certified metric; the round is not (the stochastic phase can end above its start), which is why the plateau and futility diagnostics are per polish.

Table 6: Surrogate and training configuration. All cases: $\tanh$ MLP, 4 hidden layers, Adam, $n_{\mathrm{col}} = 1024$, offline learning rate $10^{-3}$, refinement learning rate $3 \cdot 10^{-4}$ annealed by 0.97 per round with floor $10^{-5}$, seed 0.

| | width | params | offline epochs | refine epochs | L-BFGS steps | refine rounds | $c_{\mathrm{rad}}$ |
|---|---|---|---|---|---|---|---|
| C1 | 64 | 12 801 | 6000 | 800 | 120 | 10 | box |
| C2 | 64 | 12 865 | 6000 | 800 | 120 | 10 | box |
| C3 | 64 | 12 929 | 6000 | 800 | 120 | 10 | 2 |
| C4 | 128 | 51 585 | 3000 | 300 | 60 | 5 | 2 |
| C5 | 128 | 51 969 | 6000 | 800 | 120 | 10 | 3 |
| C6 | 128 | 51 201 | 6000 | 800 | 120 | 10 | 4 |

### F.3 Certificate evaluation and the inner subproblem

The network certificates are evaluated on a fixed midpoint tensor grid of parameter $n_{\mathrm{cert}}$ (nodes $(i - \frac{1}{2})/n_{\mathrm{cert}}$): residuals and tangent residuals by automatic differentiation at those nodes, the case formulas of Appendix D applied to the resulting norms, and the surrogate objective and gradient from the same midpoint sums. The anchored branch integrates the solver reconstruction by composite Gauss–Legendre quadrature of the same parameter (Appendix F.6). No bound on the quadrature remainder is added to any of these integrals (Section 5.1). The certificates at the returned points are re-verified exactly in Appendix F.6. The subproblem is solved by SciPy's `trust-constr` (IPOPT when available) with at most 200 iterations, followed by a 33-point projected Cauchy-arc safeguard, so the step is never worse than the best sampled arc point; the finite sampling of the arc is covered by the observable Cauchy condition, since $\widehat{\kappa}_k^{\mathrm{C}} = 0$ at every stepping iterate of all 240 runs (Appendix F.7).

### F.4 Reference solvers, costs, protocol and solve events

Anchors come from classical discretizations that return conforming reconstructions, so the same majorant certifies them: the exact single-mode solution (C1), the modal solution of Theorem 9 (C2), a sine-series Galerkin solution with Newton's method (C3), the modal Duhamel solution of Remark 5 (C4), a quintic B-spline Galerkin space (C5) and exact piecewise-polynomial integration of the B-spline load (C6). The reference optima $c^\star$ are computed on the same discretizations by SLSQP at tolerance $10^{-14}$, with the objective and constants of the run, from the case's own $c_0$; multi-start checks on the nonconvex cases found no better basin. This optimizer spends $16/30/55/91/4641/242$ PDE solves on C1–C6 (one state and $p$ tangent solves per objective–gradient evaluation), the campaign's anchors $0/3/0/13/17/1$. A complete run, offline training included (GPU surrogate, CPU solver), took $3/13/7/6/38/56$ minutes for SA-CPITR, $3/30/8/33/108/69$ for CPITR and 0.30 h in total for the frozen surrogate; one reference query costs 0.1–79 s.

The main campaign (budget 5, cap 10), the paired CPITR arm (budget 0, same seed and training budgets) and a C4 CPITR arm at cap 20 ran in one session on one machine. The runs are deterministic: where SA-CPITR spent no solve, its log equals CPITR's row for row. The code and the commands that regenerate every table and figure are in the supplementary material. Each of the three gradient anchors replaced the plateaued surrogate certificate $\eta_g \approx 2$–$3 \cdot 10^{-4}$ by an anchored

Table 7: **The theory against ground truth, per arm.** $n$ logged centers; (A) and (B): worst ratio of true error to certificate for the objective and for the criticality transfer; steps: accepted steps, all with true descent; $\rho$: range of the true ratio; gate: gradient-gate failures / those at which the transfer error met the demand; last three columns: exact $\chi$ at the returned point, the reported terminal bound, and its effectivity. SF = solver-free arm. C1, C3: both arms identical. C6: criticality not comparable; its claim is feasibility and descent (the gap of Table 9 is verified offline).

| | arm | $n$ | (A) | (B) | steps | $\rho$ | gate | $\chi(c_{\mathrm{ret}})$ | bound | eff. |
|---|---|---|---|---|---|---|---|---|---|---|
| C1 | both | 4 | 0.72 | 0.30 | 3 | $[0.997, 1.003]$ | 0/0 | $4.00 \cdot 10^{-5}$ | $4.16 \cdot 10^{-5}$ | 1.04 |
| C2 | SA | 5 | 0.04 | 0.02 | 4 | $[0.998, 1.004]$ | 2/2 | $1.39 \cdot 10^{-6}$ | $1.39 \cdot 10^{-6}$ | 1.00 |
| C2 | SF | 8 | 0.04 | 0.02 | 4 | $[0.998, 1.004]$ | 5/4 | $1.39 \cdot 10^{-6}$ | $1.77 \cdot 10^{-4}$ | 127 |
| C3 | both | 3 | 0.02 | 0.03 | 2 | $[1.000, 1.001]$ | 1/1 | $2.63 \cdot 10^{-5}$ | $3.44 \cdot 10^{-5}$ | 1.31 |
| C4 | SA | 3 | 0.05 | 0.01 | 2 | $[0.993, 1.000]$ | 3/3 | $3.62 \cdot 10^{-5}$ | $3.62 \cdot 10^{-5}$ | 1.00 |
| C4 | SF | 10 | 0.05 | 0.01 | 2 | $[0.993, 1.000]$ | 10/8 | $3.62 \cdot 10^{-5}$ | $1.38 \cdot 10^{-4}$ | 3.82 |
| C5 | SA | 4 | 0.03 | 0.01 | 3 | $[1.003, 1.018]$ | 4/3 | $5.00 \cdot 10^{-5}$ | $5.04 \cdot 10^{-5}$ | 1.01 |
| C5 | SF | 9 | 0.03 | 0.01 | 3 | $[1.003, 1.018]$ | 9/9 | $5.00 \cdot 10^{-5}$ | $2.27 \cdot 10^{-4}$ | 4.54 |
| C6 | SA | 9 | 0.93 | — | 4 | $[0.993, 1.000]$ | — | — | — | — |
| C6 | SF | 9 | 0.05 | — | 4 | $[0.993, 1.000]$ | — | — | — | — |

$\eta_g^a$ of $2.9 \cdot 10^{-13}$ (C2), $6.3 \cdot 10^{-9}$ (C4) and $3.5 \cdot 10^{-7}$ (C5) (Figure 4); the state anchor of C6 cut $\eta_J$ from $3.9 \cdot 10^{-6}$ to $1.5 \cdot 10^{-7}$.

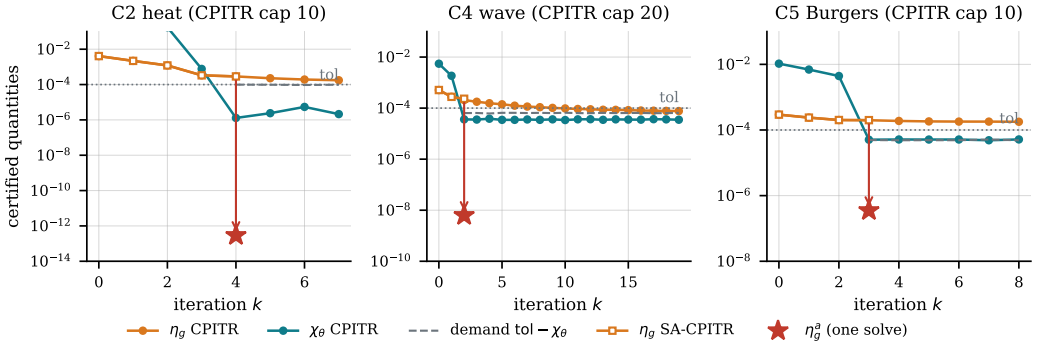



Figure 4: **One solve versus the refinement plateau** on the three decisive cases. $\eta_g$ (orange) along CPITR (circles; C4: the cap-20 re-run) and SA-CPITR (squares), $\chi_\theta$ (teal), the demand $\mathrm{tol} - \chi_\theta$ (dashed), and the anchored $\eta_g^a$ (red star).

### F.5 Ground truth

**Along the runs.** The exact objective $\widehat{J}$ and criticality $\chi$ of the reference discretization are evaluated at every logged control of both seed-0 arms and compared with what the loop saw (Figure 3, Table 7). Every conclusion the theory draws from the certificates held: objective validity at all 71 centers, the criticality transfer of Theorem 4 at all 53 centers of C1–C5 (C6's $\chi_\theta$ refers to the inner set), true descent of at least $\eta_1 \mathrm{pred}_k$ at all 36 accepted steps, and $\chi \le \mathrm{tol}$ at the returned points of all ten C1–C5 arms, including the three returned uncertified. The gradient gate never passed on the certificate while failing on the exact transfer error; at 31 of its 35 failures the truth would have passed. C5 is audited on the run's feasible set (Appendix F.6).

**At the returned controls.** Table 9 compares the certified and the true state errors. The returned controls lie $3.12 \cdot 10^{-2}$ (C1), $2.88 \cdot 10^{-5}$ (C2), $2.44 \cdot 10^{-3}$ (C3), $3.06 \cdot 10^{-2}$ (C4), $1.39 \cdot 10^{-2}$ (C5) and $1.41 \cdot 10^{-2}$ (C6) from the reference optima. Constraint activity is as designed: C5 terminates with $|c|^2 = 3.000 = P_{\max}$ and coercivity margin $\kappa_{\mathrm{lin}} = 0.154 > 0$, so its conditional certificate is valid; C4 uses 2.22 of its energy budget 9.

Table 8: **What each certificate rests on.** Evaluation of the quantities behind the SA-CPITR entries of Table 1, and the check applied.

| entry | branch | evaluation | checked by |
|---|---|---|---|
| C1, C3 terminal criticality | network | midpoint grid | exact $\chi$ at the returned point $(4.00,\ 2.63 \cdot 10^{-5})$; post-hoc anchor, exact quadrature $(4.00,\ 2.63 \cdot 10^{-5})$; C3: continuum-valid network certificate $3.46 \cdot 10^{-5}$ |
| C2, C4, C5 terminal criticality | solver | Gauss grid | exact quadrature reproduces the bounds to five digits; exact $\chi$ at the returned point |
| C6 feasibility, optimality gap | solver | exact | exact solution map, root finding; dual bound |
| descent at accepted steps | network | midpoint grid | true decrease $\geq \eta_1 \mathrm{pred}_k$ at all 36 steps |
| gate decisions along the runs | network | midpoint grid | no false gate pass in 53 centers; gates failed on 35, at 31 of them the truth would have passed |

Table 9: **Ground-truth validation of the campaign.** "cert."/"true": certified and independently computed state errors, effectivity their ratio. Last column: the optimality gap of Theorem 7 on the convex cases; (Appendix D.8) verified offline with the exact solution map, not a certificate of the method; —: not claimed (C3, C5: Remark 3) or attempted and declined (C4).

| | $J_{\mathrm{true}}$ | cert. error | true error | effectivity | certified gap |
|---|---|---|---|---|---|
| C1 | $6.242 \cdot 10^{-7}$ | $2.832 \cdot 10^{-5}$ | $2.434 \cdot 10^{-6}$ | 11.6 | $6.74 \cdot 10^{-7}$ (global) |
| C2 | $1.815 \cdot 10^{-11}$ | $4.448 \cdot 10^{-4}$ | $1.193 \cdot 10^{-5}$ | 37.3 | $1.61 \cdot 10^{-9}$ (global) |
| C3 | $4.938 \cdot 10^{-3}$ | $2.608 \cdot 10^{-5}$ | $2.314 \cdot 10^{-7}$ | 112.7 | — |
| C4 | $1.525 \cdot 10^{-3}$ | $3.757 \cdot 10^{-4}$ | $2.261 \cdot 10^{-5}$ | 16.6 | — |
| C5 | $9.663 \cdot 10^{-3}$ | $1.500 \cdot 10^{-4}$ | $3.527 \cdot 10^{-6}$ | 42.5 | — |
| C6 | $1.950 \cdot 10^{-3}$ | $6.139 \cdot 10^{-5}$ | $4.940 \cdot 10^{-6}$ | 12.4 | $5.73 \cdot 10^{-6}$ (offline, exact map) |

**Per case.** Figures 5 to 10 compare the three arms with the reference optimum of each case: (a) the distance $\|c_k - c^\star\|$; (b) the certified interval $J_\theta \pm \eta_J$ against the exact objective, which no run evaluates; (c) the returned controls against the optimal control; (d) the network states $u_\theta$ at the returned controls with a certified pointwise error band $|u - u_\theta|(x) \leq \|G(x,\cdot)\|\ \|r_\theta\|$, where $r_\theta$ is the PDE residual of the network and $G$ the Green function of the state operator: the Laplace kernel on C1 and, by the comparison principle, C3; the heat kernel on C2; the energy estimate on C4, which gives $\sqrt{x(1-x)}\, c^{-1} \int_0^T \|r_\theta(t)\|\ dt$; and on C5 $\sqrt{x(1-x)}\, C_F \|r_\theta\| / \kappa$ with the coercivity margin $\kappa$. On C6 the band is the margin the runs impose, $u_\theta \pm \varepsilon_\infty$; its upper edge meets $u_{\max}$ on $\Omega_{\mathrm{sc}}$, so the tightened constraint is active. The residual norms are quadratures, as for all certificates here. The exact objective lies in the certified interval at every iterate, and every band contains the exact state at the returned control, at 11–106 times the largest error. The band bounds the PDE error at the returned control, not its distance to the optimum, so the exact optimal state can lie outside it.

### F.6 Exact evaluation of the solver-side certificates

The anchored certificates depend only on the solver reconstruction $(u_h, w_{h,j})$, the load and $u_d$. In the campaign these are quintic splines in $t$ (C2, C4) or $x$ (C5), the exact single-mode solution (C1) and a 48-mode sine Galerkin solution (C3), so every integrand behind $\delta_a$, $\delta_{a,j}$, $\chi_a$ and $\eta_g^a$ is a piecewise polynomial or a finite sine series. Gauss–Legendre rules of exact degree on the union of breakpoints and knots, root splitting for $\int |\rho|$ and supremum norms, and Parseval for sine series integrate them exactly (for the sine-times-spline integrals of C5 the remainder is below $10^{-72}$).

**The three anchored terminal certificates.** Rebuilt from the logged controls, the three gradient anchors reproduce the logged $\chi_a$ and $\eta_g^a$ (to $10^{-9}$ relative; C2's $\eta_g^a \approx 3 \cdot 10^{-13}$ is at round-off, within 3%). Exact integration gives $\chi_a + \eta_g^a = 1.393995 \cdot 10^{-6}$ (C2), $3.622670 \cdot 10^{-5}$ (C4) and $5.036749 \cdot 10^{-5}$ (C5, on the run's feasible set), against the reported $1.393995 \cdot 10^{-6}$, $3.622670 \cdot 10^{-5}$ and $5.036744 \cdot 10^{-5}$.

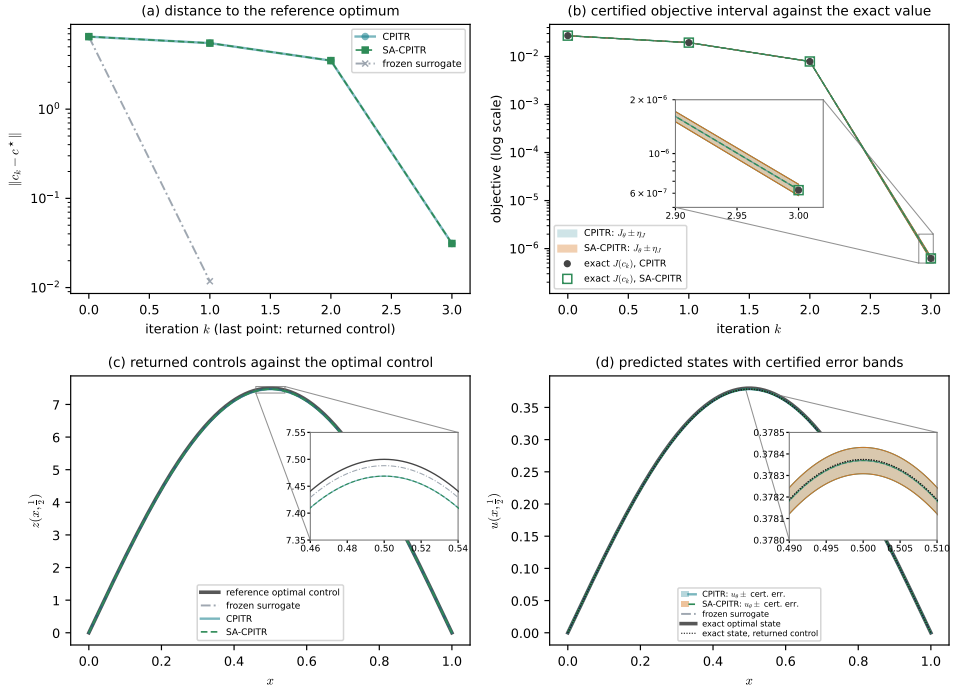



Figure 5: **C1, linear elliptic, against the exact optimum.** Panels as in Appendix F.5.

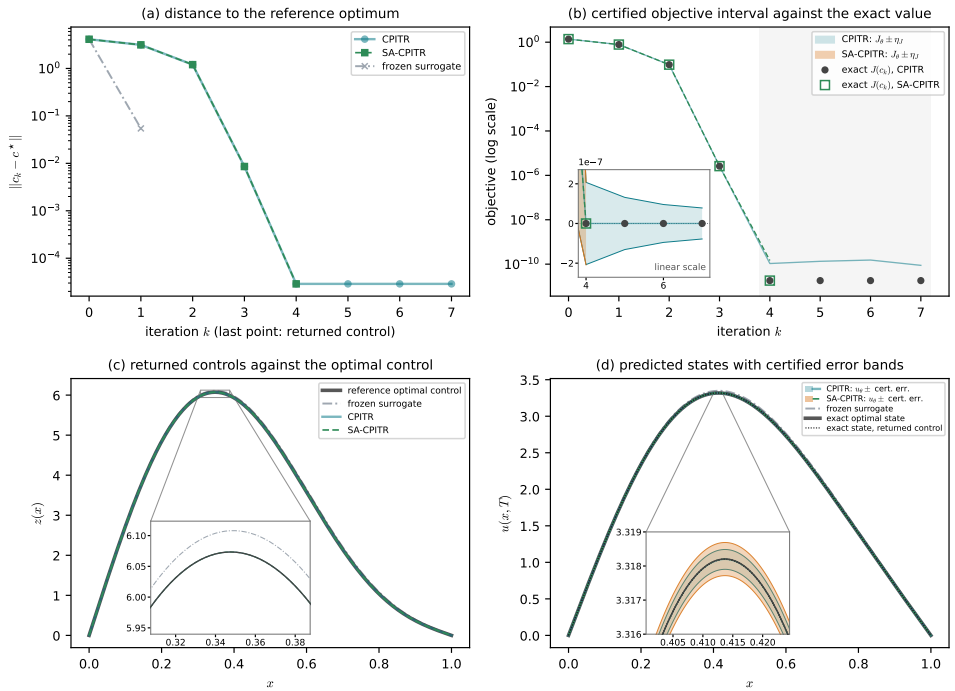



Figure 6: **C2, parabolic, against the exact optimum.** Panels as in Appendix F.5.

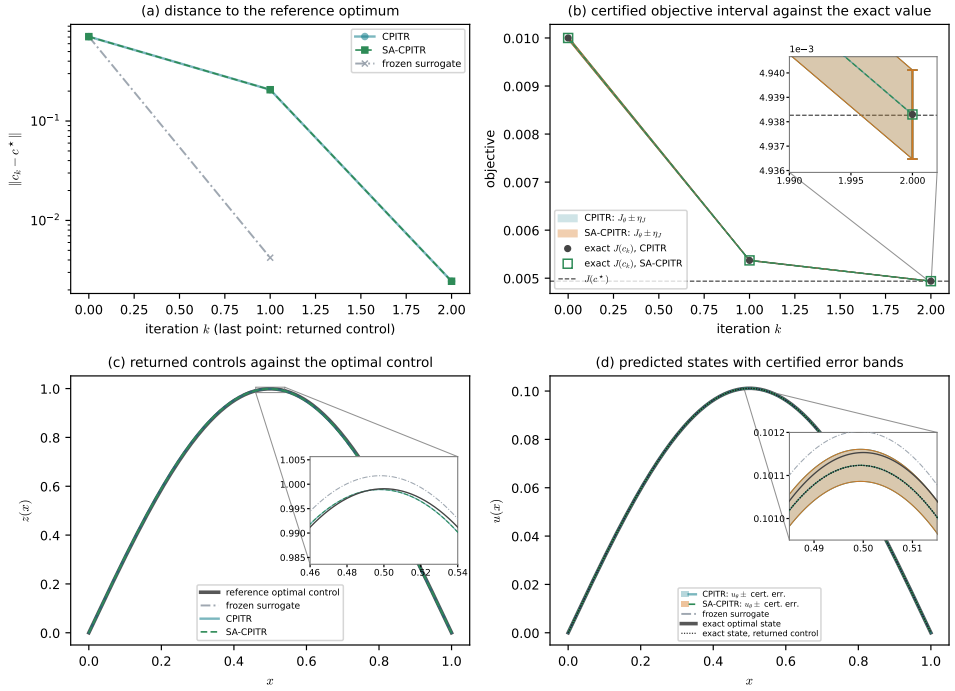



Figure 7: **C3, semilinear, against the exact optimum.** Panels as in Appendix F.5.

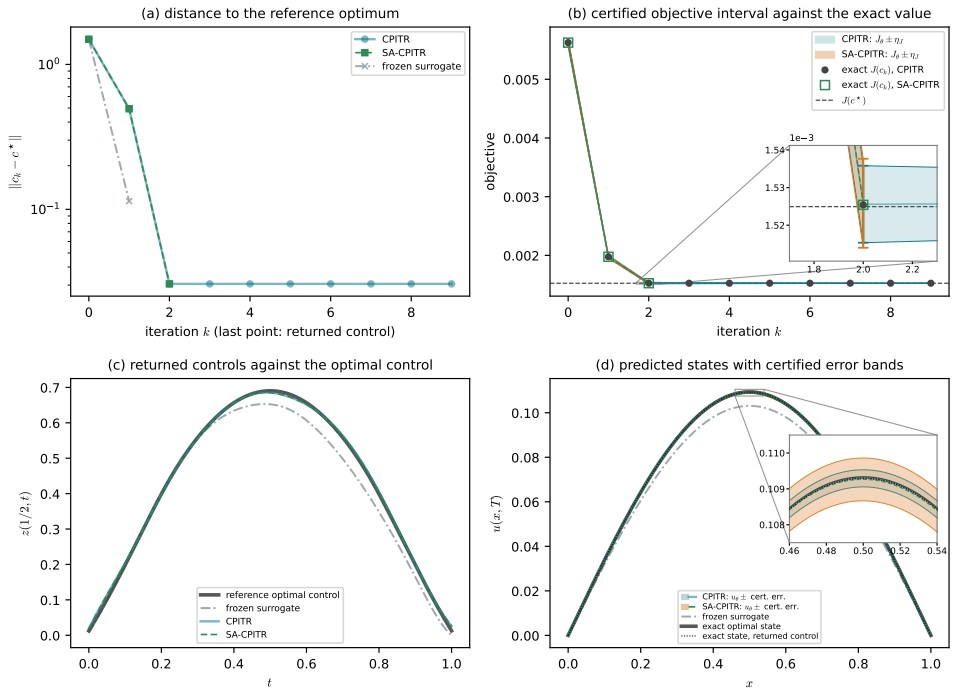



Figure 8: **C4, wave, against the exact optimum.** Panels as in Appendix F.5.

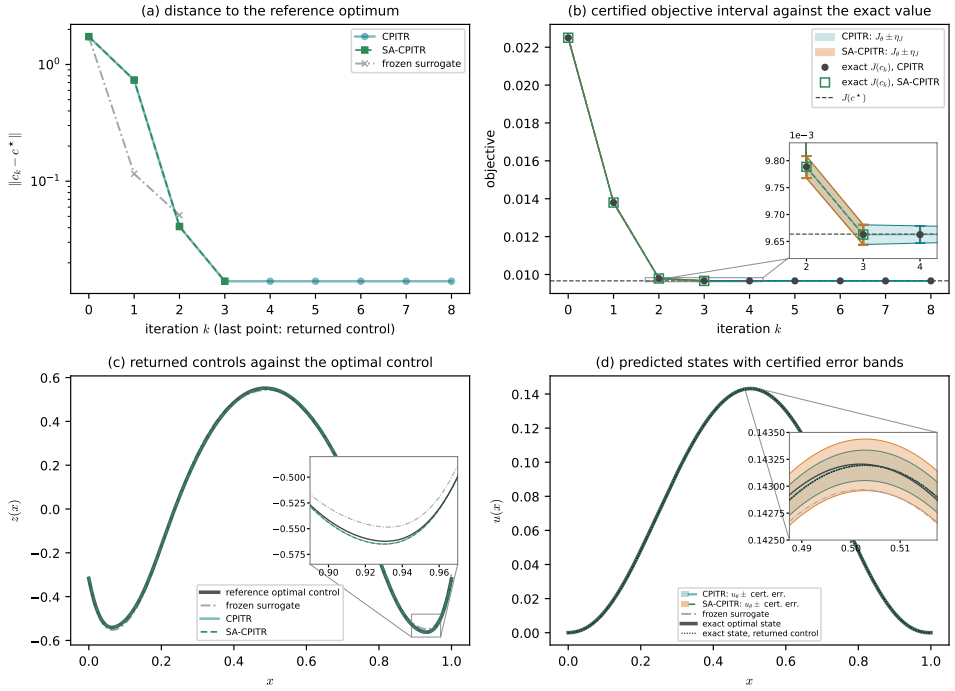



Figure 9: **C5, steady Burgers, against the exact optimum.** Panels as in Appendix F.5.

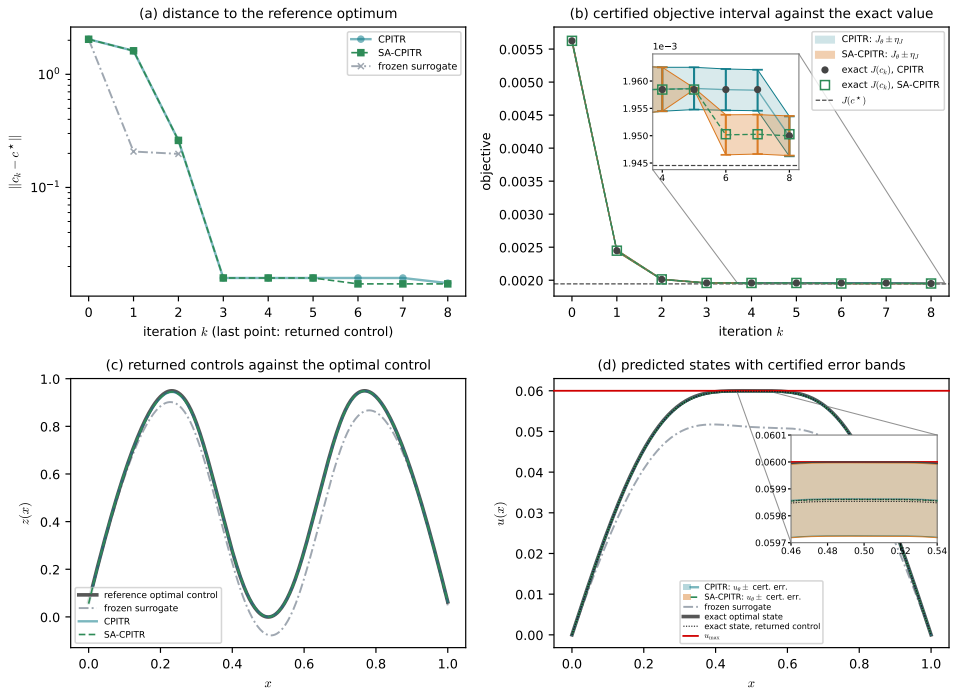



Figure 10: **C6, state constraint, against the exact optimum.** Panels as in Appendix F.5.

**C5 and the exactly feasible set.** The run imposes the mean-zero constraint of C5 with the midpoint means $w_k = \frac{1}{n_{\mathrm{cert}}} \sum_i B_k(x_i)$ instead of $\int_0^1 B_k$, so the returned control violates the exact constraint by $1.9 \cdot 10^{-5}$ and lies $7.4 \cdot 10^{-5}$ from the exactly feasible set. Its exact-set criticality is $\chi = 8.927 \cdot 10^{-5}$ ($\chi + \eta_g^a = 8.962 \cdot 10^{-5} \leq \mathrm{tol}$), and at its projection onto that set a fresh anchor (17 PDE solves) certifies $\chi_a + \eta_g^a = 5.082 \cdot 10^{-5}$.

**Verification anchors for C1 and C3.** Anchors built after the run at the returned controls of C1 (two PDE solves) and C3 (five) give exact terminal certificates of $4.002501 \cdot 10^{-5}$ and $2.628698 \cdot 10^{-5}$, below the surrogate's grid certificates ($4.16 \cdot 10^{-5}$, $3.44 \cdot 10^{-5}$) and equal to the exact $\chi$ of Table 7 to six digits; they are not part of the campaign.

**A continuum evaluation of the network certificate on C3.** At the returned control of C3 the saved network is evaluated in interval arithmetic: on each of $8 \cdot 10^4$ cells, interval Taylor jets in $x$ (third order) and $c$ (first order) enclose the residuals, the misfit, the tangents and their $x$-derivatives, and the cellwise bound $|\int_{\mathrm{cell}} f - hf(x_{\mathrm{mid}})| \leq \frac{h^2}{4} \sup_{\mathrm{cell}} |f'|$ turns the midpoint sums into rigorous enclosures of every integral in the certificate. After reproducing the logged grid certificate to ten digits the computation gives $\chi_\theta + \widehat{\eta}_g \leq 3.457 \cdot 10^{-5}$ against the grid value $3.444 \cdot 10^{-5}$ (quadrature remainder $3.1\%$ of $\|\rho\|^2$; $10^4$ cells give $4.33 \cdot 10^{-5}$); the global midpoint remainder, with $\left\|\partial_x^2(\rho^2)\right\|_\infty \leq 2 \cdot 10^{18}$ bounded from the weights, would add $1.3 \cdot 10^{12}$ to $\|\rho\|^2 \approx 6.7 \cdot 10^{-8}$. Only the terminal networks were saved, so the certificates along the runs cannot be evaluated this way.

### F.7 The observable Cauchy condition along the campaign

Proposition 1 makes the curvature hypothesis of Theorem 5 checkable from the logged $(\chi_\theta, \Delta_k, \mathrm{pred}_k)$. At every iteration of every run in which a step was taken, the predicted decrease exceeded the curvature-free Cauchy floor $\frac{1}{2}\chi_\theta \min(\Delta_k, \chi_\theta)$, so the implied arc curvature $\widehat{\kappa}_k^{\mathrm{C}}$ of (17) is zero throughout; the smallest margins are $2.6\times$ (C2, $k = 0$), $2.9\times$ (C6, $k = 3$) and $12.9\times$ (C5, $k = 2$), the largest $779\times$ (C1). Table 10 lists the per-case minima. The step remainder $\widehat{\kappa}_k^{\mathrm{s}}$ is logged by the reference implementation and reported for the 240 runs of Appendix G (Figure 16b): it stays below $0.179$ (median $0.0038$) and $\widehat{\kappa}_k^{\mathrm{C}}$ is zero in every run, so $\widehat{\kappa} < \infty$ and Proposition 1(ii) applies throughout.

Table 10: Minimum over stepping iterations of the observed ratio $\mathrm{pred}_k / [\frac{1}{2}\chi_\theta \min(\Delta_k, \chi_\theta)]$ per case and campaign; a value $\geq 1$ means $\widehat{\kappa}_k^{\mathrm{C}} = 0$ at every step.

| | C1 | C2 | C3 | C4 | C5 | C6 |
|---|---|---|---|---|---|---|
| SA-CPITR (budget 5) | 222 | 2.6 | 45 | 244 | 12.9 | 2.9 |
| solver-free (budget 0) | 222 | 2.6 | 45 | 244 | 12.9 | 2.9 |

### F.8 Control, state and sensitivity fields

A run saves only the best and the final network, so the surrogate panels below show the terminal network evaluated at the trajectory's controls, while the reference fields are solved at $c_k$ exactly.

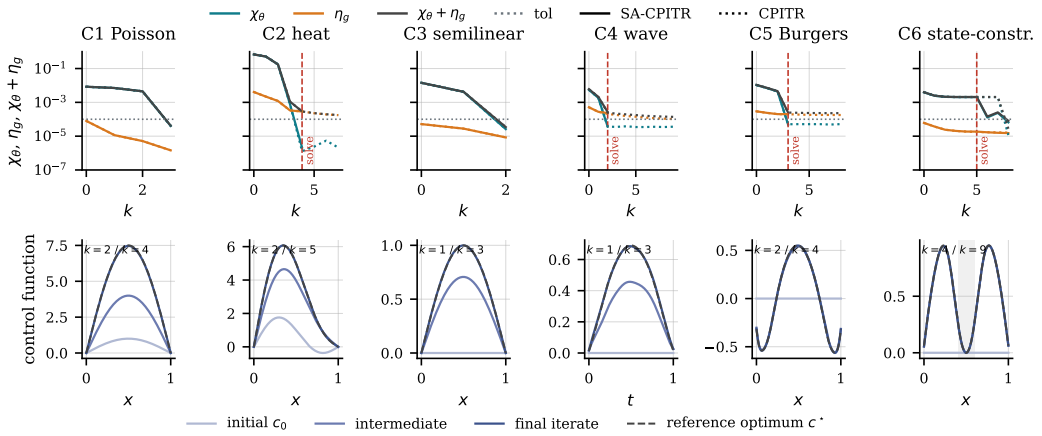



Figure 11: **Per-case trajectories.** SA-CPITR (solid) and CPITR (dotted): $\chi_\theta$, $\eta_g$ and $\chi_\theta + \eta_g$ against tol with anchor events marked (top); the control at the initial, an intermediate and the final iterate against the reference optimum (bottom). Fields at the same iterates: Figures 12 to 14.

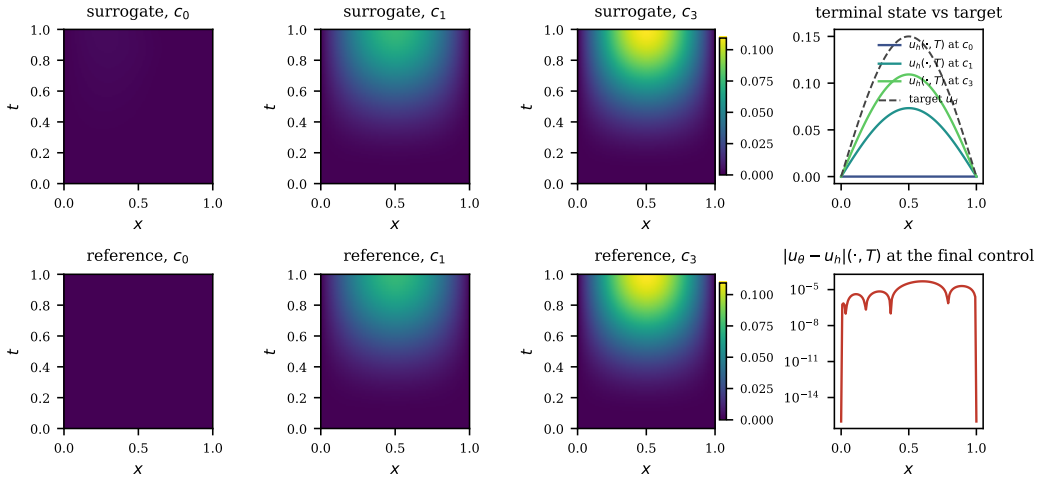



Figure 12: **C4 state along the optimization.** $u(x,t)$ at the initial, intermediate and final control (surrogate, top; reference, bottom); terminal profiles $u(\cdot,T)$ against $u_d = 0.15 \sin \pi x$ (the velocity and Tikhonov terms keep the optimum off the target); surrogate–reference difference at $t = T$.

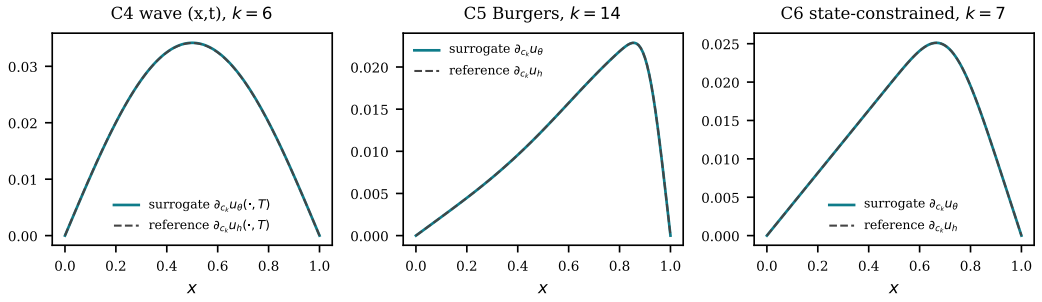



Figure 13: One sensitivity $\partial u / \partial c_k$ at the final control (the component with the largest $|c_k|$), surrogate against reference, for C4 (terminal profile), C5 and C6. The tangent residuals in the training loss (31) are what make these agree; the assembly (3) multiplies their error by the tracking residual.

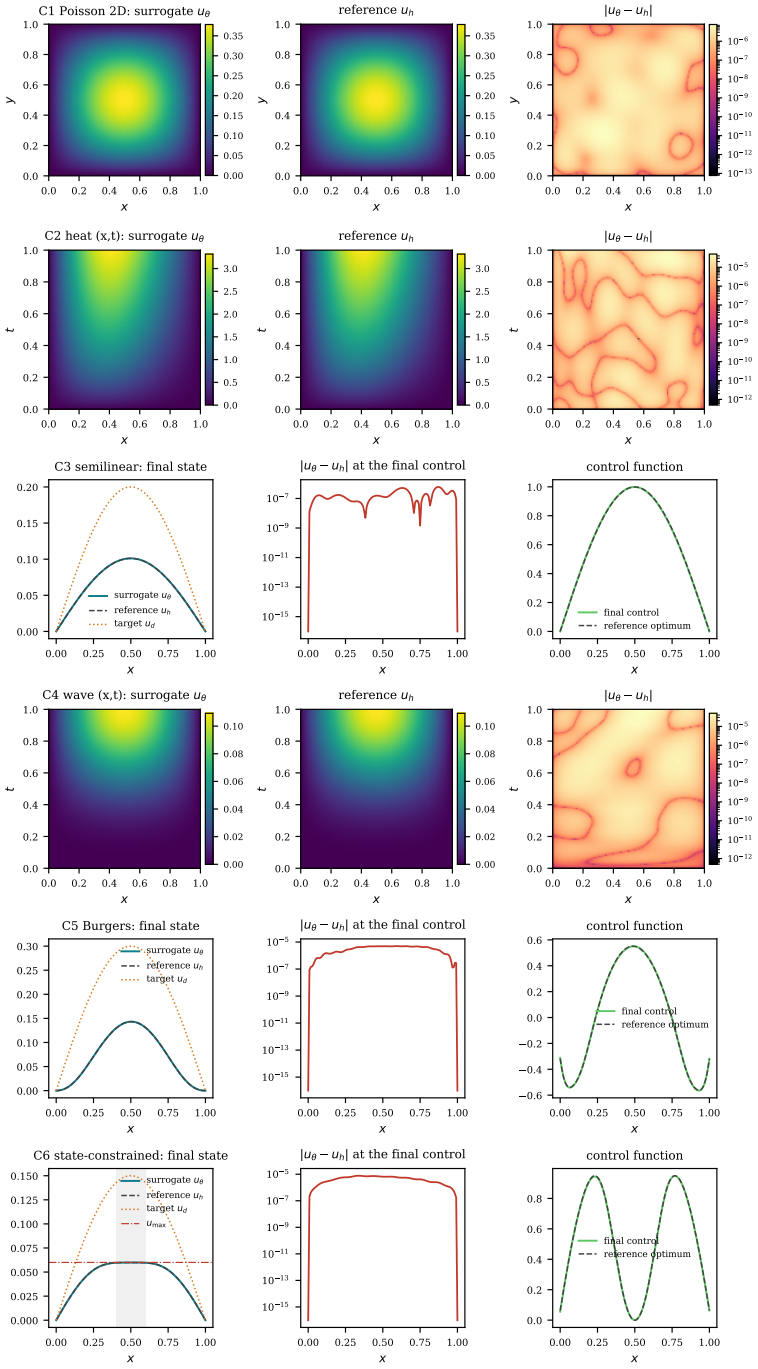



Figure 14: **Final-iterate fields.** Space–time and 2D cases: $u_\theta$, $u_h$, $|u_\theta - u_h|$ (log scale). Steady 1D cases: state against reference and target, pointwise difference, control against the reference optimum; C6: shaded band $\Omega_{\text{sc}} = [0.4, 0.6]$, dash-dotted line $u_{\max} = 0.06$.

# G Ablations and the frozen-surrogate baseline

This appendix tests whether the campaign's outcome survives a different seed, a smaller network, a shorter training run, or a deliberately degraded surrogate, in a 240-run study; the campaign is its seed-0, default-architecture slice and is reproduced there iterate for iterate.

## G.1 Design and protocol

Table 11 lists the six studies. Every study that varies the surrogate is run as a pair, the same configuration with $N_{\max} = 5$ and with $N_{\max} = 0$, so a pair differs only in the solve policy. Depth, activations, optimizers and the tolerance are not swept: they probe the PINN literature, not the claim that certification is bought with few solves. All other settings are those of Appendix F.4.

Table 11: The ablation study. "Pairs" counts configurations run at both $N_{\max} = 5$ and $N_{\max} = 0$. Seeds are 0–4; the decisive cases are C2, C4 and C5, on which the campaign's gradient anchors fired.

| | what varies | runs | question |
|---|---|---|---|
| S1 | seed $\in \{0, \dots, 4\}$, all six cases | 60 | is the outcome a seed artifact? |
| S2 | width $\in \{16, 32, 64, 128\}$, seed 0 | 48 | how does the method degrade with capacity? |
| S3 | training budget $\in \{\frac{1}{4}, \frac{1}{2}, 1, 2\}$, seed 0 | 48 | the same along the training axis |
| S4 | frozen offline surrogate, 2 widths, 5 seeds | 60 | the practice baseline |
| S5 | bad surrogate (width 16, $\frac{1}{4}$ training), 3 seeds | 18 | can anchors rescue it? |
| S6 | two regime guards of $E_g$, seed 0 | 6 | guard consistency (identical runs) |

## G.2 Results

Across the 74 pairs of S1–S3 and S5 on C1–C5, SA-CPITR certifies 53 and the solver-free arm 26; no pair loses a certificate by adding solves, 27 are decided in favour of the method and none against (Figure 15a, Table 12). The budget-5 runs spend 50 solves, never more than two per run, and take 34.0 h against 47.0 h. The 13 C6 pairs (extension) are not counted: there the 17 solves (21.6 h against 23.4 h) do not change the outcome, a control passing the feasibility test of Theorem 15 is returned in 12 of 13 runs of each arm, and all 26 returned controls are truly feasible.

Table 12: Paired outcomes by study. "SA only" counts pairs in which SA-CPITR certifies and the solver-free arm does not; the reverse column is empty by construction of the data, not by convention. Counts are over C1–C5; the C6 extension is reported in the text.

| | certified | | decided by the solve | | | wall (h) | |
|---|---|---|---|---|---|---|---|
| study | SA | solver-free | SA only | SF only | solves | SA | SF |
| S1 seeds | 20/25 | 10/25 | 10 | 0 | 16 | 9.8 | 15.1 |
| S2 width | 14/20 | 8/20 | 6 | 0 | 13 | 12.1 | 16.2 |
| S3 training budget | 16/20 | 8/20 | 8 | 0 | 15 | 8.3 | 11.5 |
| S5 bad surrogate | 3/9 | 0/9 | 3 | 0 | 6 | 3.8 | 4.2 |
| total | 53/74 | 26/74 | 27 | 0 | 50 | 34.0 | 47.0 |

**Seeds.** S1 repeats the campaign for five seeds (Table 13): the method certifies 20/25 (C1–C5) against 10/25, and on C2, C4 and C5, where the solver-free arm never certifies, it certifies 10/15, each with a single gradient anchor. Under the budget-before-abstention rule (R1) of Appendix H, SA-CPITR certifies C5 for all five seeds instead of three.

**Capacity and training budget.** At every width in S2 and every training budget in S3 (Figure 15b,c) the method certifies more cases than the solver-free arm (3–5 of the five cases C1–C5 against 2); neither arm is monotone in the training budget.

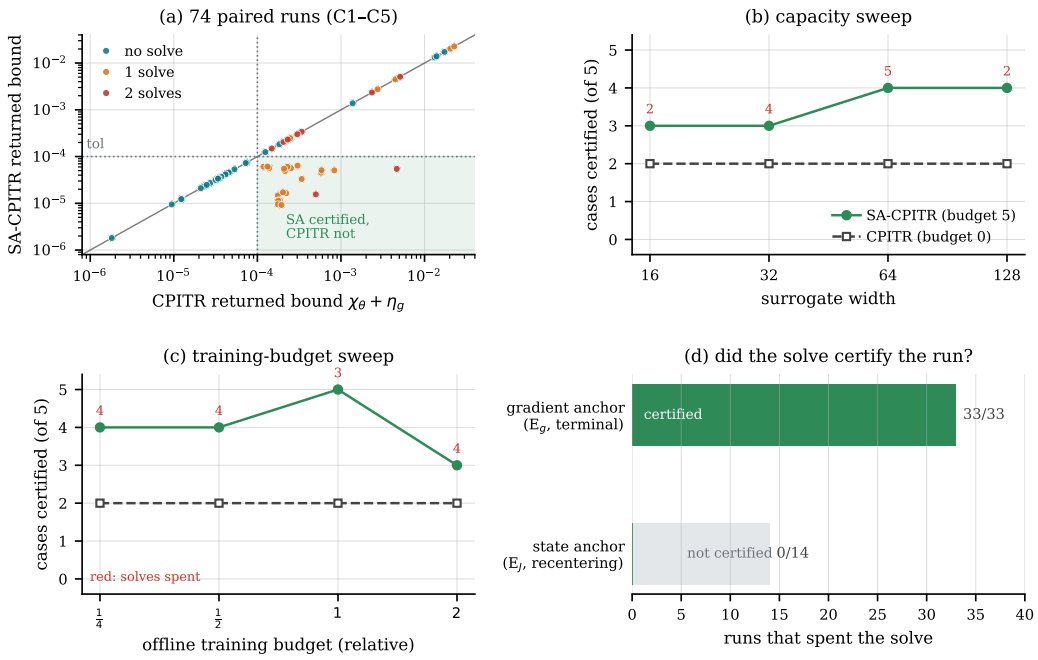



Figure 15: **The ablation study.** (a) The 74 pairs on C1–C5, CPITR against SA-CPITR bound, coloured by events spent; diagonal points spent none, the shaded quadrant holds the 27 pairs the solve decides. (b), (c) Cases certified out of five against surrogate capacity and training; red digits the events of the budget-5 arm. (d) Certification rate of the runs that spent each event role.

Table 13: S1: certified runs out of five seeds, and mean wall clock per run. Solves are the total the budget-5 arm spent over the five seeds. C6 (extension): runs returning a control that passes the feasibility test of Theorem 15.

| | certified (of 5) | | | mean wall (min) | |
|---|---|---|---|---|---|
| case | SA-CPITR | solver-free | solves | SA-CPITR | solver-free |
| C1 Poisson | 5 | 5 | 0 | 3.8 | 3.7 |
| C2 heat | 4 | 0 | 4 | 14.0 | 22.3 |
| C3 semilinear | 5 | 5 | 0 | 7.8 | 7.7 |
| C4 wave | 3 | 0 | 7 | 26.9 | 38.8 |
| C5 Burgers | 3 | 0 | 5 | 65.3 | 109.2 |
| total C1–C5 | 20 | 10 | 16 | 9.8 h vs 15.1 h | |
| C6 (ext.), feasible | 5 | 5 | 5 | 104.5 | 123.2 |

**The frozen-surrogate baseline.** S4 minimizes the frozen offline surrogate with a plain gradient method; none of its 60 runs returns a certificate. At the stopping point the surrogate's own median criticality is $2 \cdot 10^{-12}$ to $2 \cdot 10^{-8}$ on C1–C5 ($8 \cdot 10^{-6}$ on C6; default width), while the certified bound on the true criticality is four to ten orders larger on C1–C5 and exceeds tol in 40 of the 50 C1–C5 runs (Figure 16a). On C2 the frozen runs land $4.6 \cdot 10^{-2}$ from the exact optimum against $2.4 \cdot 10^{-5}$ for SA-CPITR, in 1.6–5.0 minutes per case against 3.8–105.

**Rescuing a bad surrogate.** S5 combines both degradations (width 16, a quarter of the training) on the decisive cases. On C2 all three seeds certify with a single gradient anchor ($4.9 \cdot 10^{-5}$ against the solver-free arm's $6.7 \cdot 10^{-4}$); on C4 and C5 no seed certifies in either arm, and the method abstains after one solve across the three C4 seeds and two across the three C5 seeds instead of spending its budget. Mode A executes in none of the 240 runs; it does out of distribution (Appendix H).

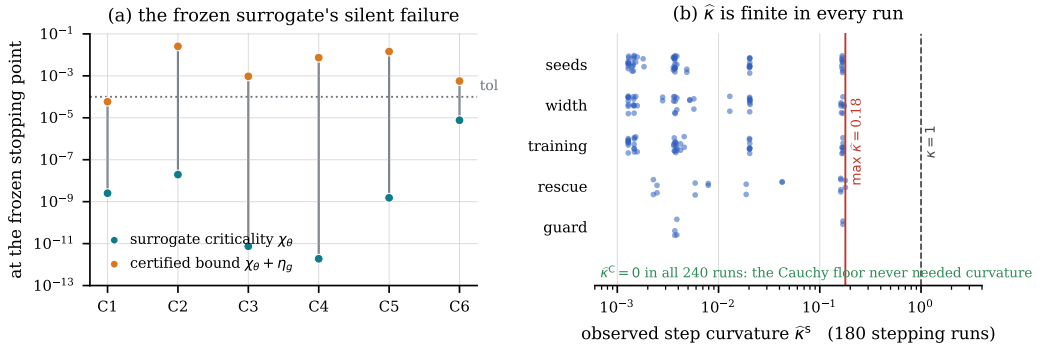



Figure 16: **The frozen surrogate, and the observed curvature.** (a) Own criticality $\chi_\theta$ (teal) against the certified bound $\chi_\theta + \widehat{\eta}_g$ (orange) at the frozen surrogate's returned point; medians over five seeds, default width. (b) Observed step curvature $\widehat{\kappa}^s$ of (17) in every stepping run, by study.

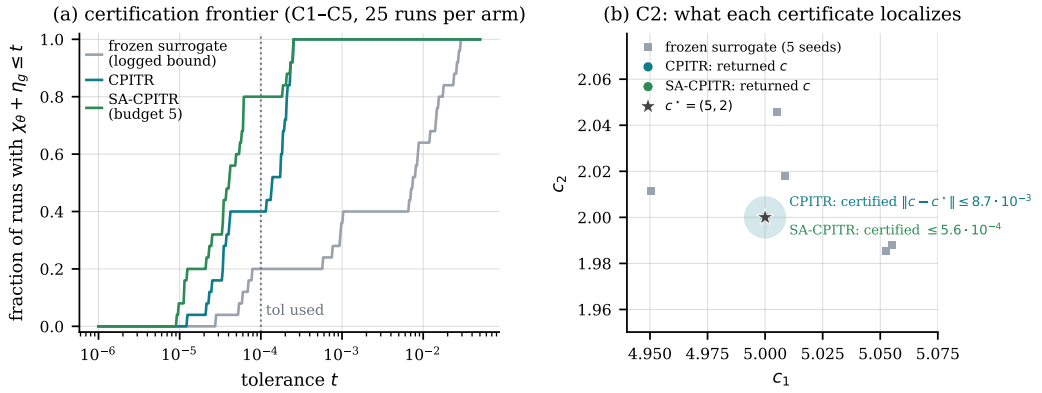



Figure 17: **Certification frontier and certified balls.** (a) Fraction of the 25 seed-study runs per arm on C1–C5 with returned bound $\chi_\theta + \widehat{\eta}_g \le t$, read a posteriori ($\mathrm{tol} = 10^{-4}$ dotted). (b) C2 in coefficient space: certified balls around the solver-free control ($d = 8.7 \cdot 10^{-3}$, teal) and SA-CPITR's ($d = 5.6 \cdot 10^{-4}$, green), the exact optimum $c^\star = (5, 2)$, and the five uncertified frozen-surrogate controls.

**Where the tolerance sits.** On C2, C4 and C5 the solver-free arm ends with a gradient certificate of $0.66$–$1.97 \cdot 10^{-4}$ (five seeds each), so $\mathrm{tol} = 10^{-4}$ sits at or just below the surrogate's own certificate floor. Read a posteriori at other tolerances (Figure 17a), the two arms certify 20 and 10 of the 25 C1–C5 seed-study runs at tol, both reach 25 at $3 \cdot 10^{-4}$, and at $3 \cdot 10^{-5}$ the counts drop to 8 and 4. In design space the certified gap of Theorem 7 localizes the C2 optimum to $8.6$–$9.6 \cdot 10^{-3}$ for the solver-free arm and to $4.6$–$8.5 \cdot 10^{-4}$ for SA-CPITR (Figure 17b), which suggests choosing tol from a distance target as $\mathrm{tol} \approx \widehat{\zeta}\, d_{\mathrm{target}}/2$.

**The two solve roles.** All 33 runs that spent a gradient anchor ended with a certificate; of the 14 C1–C5 runs that spent only state anchors, none did (Figure 15d).

## H SURROGATES OUT OF DISTRIBUTION

The surrogate is trained on a box that excludes the optimum, and the run starts inside it. On C2 the box is $[-1, h_1] \times [-1, h_2]$ with $(h_1, h_2) = (3, 3)$ (moderate) or $(2, 1)$ (hard), so that $c^\star = (5, 2)$ lies 2.0 or 3.2 outside, in the design space $[-1, 20]^2$; on C5 it is $[-r, r]^{16}$ with $r = 0.10$ or $0.03$ (1.4 and 1.6 outside). The frozen surrogate, CPITR and SA-CPITR share each surrogate (three seeds, 36 runs). The event layer is the campaign's with two rules changed: (R1) before an abstention rule

may stop a run, the run spends the anchor its budget still allows (the anchored terminal test inside tolerance, a gradient anchor and mode A outside it); (R2) $\mathrm{E}_J$ fires only if a state anchor with the observed gain could flip the rejection that triggered it.

Table 14: C2 out of distribution: six runs per arm (two levels $\times$ three seeds). "certified" is the a-posteriori criticality certificate $\chi$-bound $\leq$ tol; the solver-free arm holds the design-space gap certificate of Theorem 7 instead in all six runs.

| | certified | $\|c_T - c^\star\|$ | bound at return | wall [min] |
|---|---|---|---|---|
| frozen surrogate | 0/6 | $2.6 \cdot 10^{-2}$–1.04 | $6.7 \cdot 10^{-2}$–3.7 | 2.4–3.1 |
| CPITR | 0/6 | $6.0 \cdot 10^{-6}$–$1.3 \cdot 10^{-4}$ | $1.77$–$1.96 \cdot 10^{-4}$ | 28–51 |
| SA-CPITR (one anchor) | 6/6 | $6.0 \cdot 10^{-6}$–$1.3 \cdot 10^{-4}$ | $1.3$–$4.7 \cdot 10^{-5}$ | 15–19 |

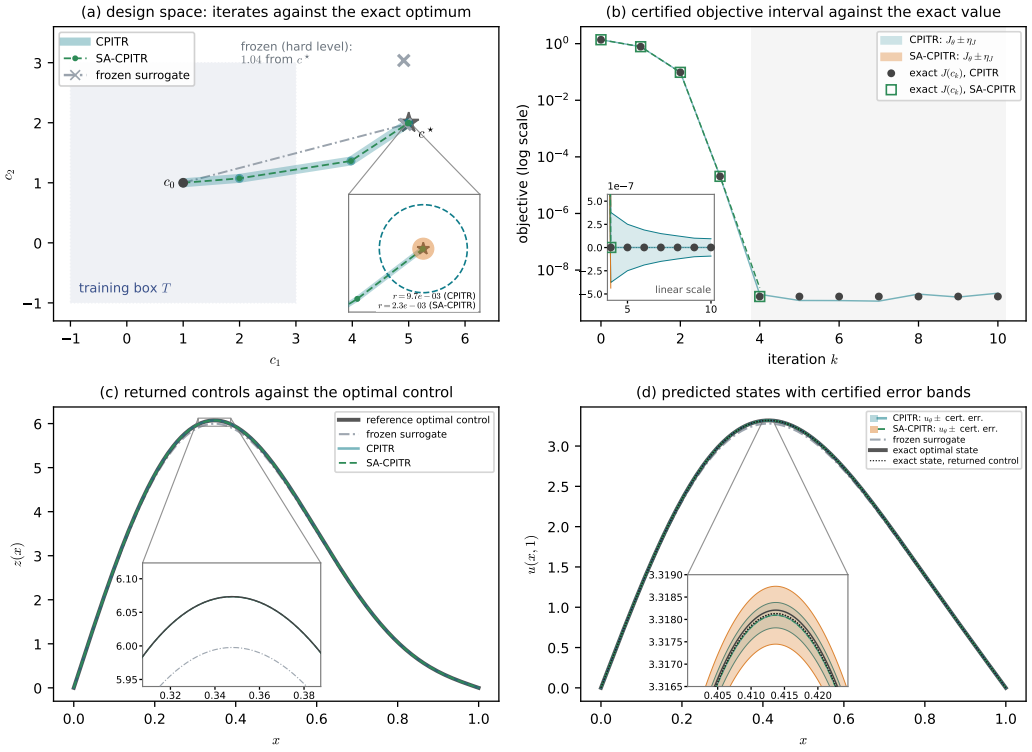



Figure 18: **C2 out of distribution against the exact optimum** (seed 0, moderate level). (a) Iterates in design space; inset: the certified balls around $c^\star$. (b)–(d) As in Appendix F.5.

**C2.** Every stepping run leaves the training box and lands on $c^\star$ to $6 \cdot 10^{-6}$–$1.3 \cdot 10^{-4}$ (Table 14). Without solves the certificate stalls at 1.8–2.0 $\times$ tol; one gradient anchor certifies all six SA runs at $1.3$–$4.7 \cdot 10^{-5}$. The frozen surrogate misses $c^\star$ by up to 1.04 while its own criticality reads $\chi_\theta \approx 10^{-11}$; its certified bound, $6.7 \cdot 10^{-2}$–3.7, flags the failure. C2 has the closed-form solution map $u(x,t;c) = \sum_{k\leq 2} c_k U_k(x,t)$, $U_k = \lambda_k^{-1}(1 - e^{-\lambda_k t}) \sin k\pi x$, $\lambda_k = Dk^2\pi^2$, so Figure 18 compares the seed-0 runs with the exact optimum. The exact objective lies inside $J_\theta \pm \eta_J$ at every logged iterate; both stepping arms return the same control, inside the certified balls $\|c - c^\star\| \leq \frac{2}{\widehat{\zeta}}(\chi + \eta_g)$ of Theorem 7 ($\widehat{\zeta} = 0.0404$) of radius $9.7 \cdot 10^{-3}$ without and $2.3 \cdot 10^{-3}$ with the anchor; and the state bands of both networks contain the exact state.

**C5.** Mode A executes here. At the hard level the frozen surrogate's coercivity fixed point fails at its returned control, so its bound is $+\infty$. On two hard seeds the solver-free arm abstains at $k = 3$; SA-

CPITR instead spends a gradient anchor under (R1), which sharpens $\eta_g$ from $6.0 \cdot 10^{-3}$ to $8.1 \cdot 10^{-8}$ and $4.8 \cdot 10^{-7}$, and accepts one mode-A step in each run ($\rho = 0.55$, $0.44$) before the time budget ends it uncertified. With eight hours instead of three (moderate level, seed 0) the solver-free run stops by futility after $5.1\,\mathrm{h}$ at bound $2.09 \cdot 10^{-3}$, while SA-CPITR spends two state anchors ($\eta_J$ sharpened $49\times$ and $67\times$), keeps accepting and ends at $1.05 \cdot 10^{-3}$, with true objective within $1.0\%$ of the reference optimum; neither certifies at $\mathrm{tol}$. No run of this appendix abstains with budget left or claims a certificate its returned bound does not support.

# I Related work

**Learning-based PDE-constrained optimization.** Neural methods for PDE-constrained optimization fall into two groups. The first trains the control together with a network for the state and sometimes the adjoint: penalty formulations with PINNs (Mowlavi & Nabi, 2023; Barry-Straume et al., 2026), hard-constraint inverse design (Lu et al., 2021b), all-at-once solvers for parametric optimal control (Demo et al., 2023; Yin et al., 2024), operator splitting for nonsmooth objectives (Song et al., 2024), and bi-level schemes in which an inner PINN solves the constraint and the outer loop follows implicit hypergradients (Hao et al., 2023). The second trains a surrogate of the control-to-state map offline and optimizes over its inputs: operator learning for control (Hwang et al., 2022), physics-informed DeepONets (Wang et al., 2021b) as differentiable surrogates for design and control (Wang et al., 2021a), derivative-informed neural operators (O'Leary-Roseberry et al., 2024) for optimization under uncertainty (Luo et al., 2025), and shape optimization (Shukla et al., 2024); differentiable solvers are also used to learn controllers (Holl et al., 2020). In all of these the returned control is assessed against a reference solver after the optimization or on test instances, and no computable bound on the error of the objective or the gradient enters the iteration. Closest in intent is Cheng et al. (2025), which counters the derailment of gradient-based optimization by inaccurate neural-operator predictions with training on optimization trajectories, better derivatives and interleaved numerical solves. No error bound governs that combination; here the certificates of Section 3.1 decide when to rely on the surrogate, and they give guaranteed descent and a certified bound at the returned point.

**When the optimizer exploits the surrogate.** Offline model-based optimization, where the objective is a black box known only through data, documents how an optimizer is drawn to inputs where a learned model is wrong. Remedies there keep the search near the training distribution (Brookes et al., 2019), learn an inverse map from scores to inputs (Kumar & Levine, 2020), retrain the model as the search moves (Fannjiang & Listgarten, 2020), or learn an objective model that is conservative away from the data (Trabucco et al., 2021; 2022). Mixed-integer and deterministic global optimization over trained networks (Fischetti & Jo, 2018; Anderson et al., 2020; Schweidtmann & Mitsos, 2019; Ceccon et al., 2022) solves the surrogate problem to global optimality, which says nothing about the system the network models. In our setting the PDE is known, so the surrogate error at a queried control is bounded from its residual instead of estimated from data, also outside the training distribution (Appendix H).

**A posteriori error estimates for neural PDE solutions.** Functional a posteriori estimates (Repin, 2000; 2008) bound the energy error of any conforming approximation by computable residual terms without Galerkin orthogonality, so they apply to network outputs. They have been used to control the error of network solutions of boundary value problems (Muzalevskiy & Repin, 2023) and as a training loss (Fanaskov et al., 2024). Related estimates exist for variational PINNs (Berrone et al., 2022), the Deep Ritz method (Minakowski & Richter, 2023), Galerkin neural networks (Ainsworth & Dong, 2021) and PINNs of ODE and PDE systems (Hillebrecht & Unger, 2022; 2025), together with two-sided bounds from Riesz representers of the residual (Ernst et al., 2025) and networks that supply the adjoint weights of dual-weighted residual estimates (Roth et al., 2022). These works bound the error of a computed solution or of a goal functional. We assemble the state and tangent majorants into certificates for the reduced objective and its gradient (Theorem 1), evaluate them at every iterate and use them in the acceptance, gate and termination tests; because the majorant holds for any conforming candidate, it also certifies the solver reconstructions.

**A priori analysis and continuum-valid certification.** Error analyses of PINNs (Shin et al., 2020; Mishra & Molinaro, 2023; Zeinhofer et al., 2025), reviewed by De Ryck & Mishra (2024), and of

operator networks (Lanthaler et al., 2022) explain when small residuals imply small errors; they give rates and stability estimates, not computed bounds for a given trained network. Computed bounds need the integrals of the residuals, which our runs evaluate on fixed quadrature grids without a remainder bound (Section 5.1). Rigorous alternatives exist: linear bound propagation through computational graphs (Xu et al., 2020), applied to PINN residuals by Eiras et al. (2024), and interval or quadrature enclosures of network norms and their derivatives (Haugen et al., 2025; Gründler et al., 2026). They are the missing step to continuum-valid certificates along the run; the re-verification of the C3 certificate (Appendix F.6) applies this idea at the returned point.

**Derivative accuracy and residual-based correction.** Sobolev training fits derivatives together with values (Czarnecki et al., 2017), with PINN analogues in Sobolev-norm losses (Son et al., 2023) and gradient-enhanced residuals (Yu et al., 2022); derivative-informed neural operators learn parametric Jacobians from derivative data (O'Leary-Roseberry et al., 2024). Our tangent residuals serve the same end without derivative data, and the tangent majorants bound the remaining sensitivity error at query time. Residual-based error correction uses the PDE at query time in another way: one linear variational solve built from the residual reduces the error of a neural-operator prediction quadratically (Cao et al., 2023; Jha, 2024). An anchor also spends a solve at the queried control, but its output is certified by the same majorant and it is spent only when a certificate is insufficient.

**Trust-region methods with inexact information.** Global convergence of trust-region methods with inexact gradients goes back to Carter (1991) and Conn et al. (2000). In PDE-constrained optimization, inexact SQP and adaptive-accuracy methods (Heinkenschloss & Vicente, 2002; Ziems & Ulbrich, 2011; Kouri et al., 2013; 2014; Kouri & Ridzal, 2018), proximal methods with inexact evaluations (Baraldi & Kouri, 2023) and multilevel methods (Gratton et al., 2008) rest on an accuracy condition that ties the model error to the gradient norm and the radius and is enforced by refining a discretization or a sample; our majorants make it checkable for a neural oracle (Theorem 4). Trust regions for noisy evaluations assume a known noise level (Sun & Nocedal, 2023) or oracles that are accurate with some probability (Bandeira et al., 2014; Chen et al., 2018; Blanchet et al., 2019; Cao et al., 2024); our bounds are pointwise and deterministic, and an anchor can make them smaller.

**Model management and reduced-basis trust regions.** The first-order corrected surrogate inside a trust region (Alexandrov et al., 1998) is the basis of our anchored mode and of trust-region POD (Arian et al., 2000; Fahl & Sachs, 2003) and multifidelity methods (Robinson et al., 2008; March & Willcox, 2012), surveyed in (Forrester & Keane, 2009; Peherstorfer et al., 2018); in the taxonomy of Peherstorfer et al. (2018) our method is of adaptation type with certified switching. Certified reduced-basis trust regions (Qian et al., 2017; Keil et al., 2021; Banholzer et al., 2022; Keil & Ohlberger, 2024; Kartmann et al., 2024; Yue & Meerbergen, 2013; Zahr & Farhat, 2015) and certified reduced bases for optimal control (Kärcher & Grepl, 2014; Bader et al., 2016; Kärcher et al., 2018) are the closest precedent for certificate-driven solver economy. Their rigorous bounds are most mature under affine parameter dependence and extend to nonaffine and nonlinear problems through empirical interpolation and problem-specific stability analysis (Barrault et al., 2004; Veroy & Patera, 2005; Canuto et al., 2009); they measure the distance to the full-order model, while our majorants bound the distance to the PDE solution; and their full-order solves are triggered by an error-bound threshold, ours by optimization-side events with soundness proofs. Networks enter this line as a further level of a certified hierarchy whose reduced-basis estimator certifies the network output (Haasdonk et al., 2023), or as a surrogate trained along the optimization path and checked with full-order evaluations (Keil et al., 2022); we certify the network directly, without a reduced space.

**Error-controlled discretization and hybrid solvers.** Adaptive finite elements for optimal control refine the mesh where dual-weighted residual estimates of the cost are large (Becker et al., 2000), also for parabolic problems (Meidner & Vexler, 2007) and control constraints (Hintermüller & Hoppe, 2008), and adaptive multilevel SQP ties the discretization accuracy to the progress of the optimizer (Ziems & Ulbrich, 2011); network refinement and the mesh ladder of the anchors play this role here, driven by a majorant instead of an estimate. Hybrid schemes combine learned and numerical solvers the other way round: networks correct a coarse solver (Um et al., 2020; Kochkov et al., 2021), modify the iterations of a convergent solver (Hsieh et al., 2019), supply initial guesses for Newton's method (Huang et al., 2020), or alternate with relaxation sweeps (Zhang et al., 2024;

Markidis, 2021). In SA-CPITR the network does the work, and the solver is called when a certificate shows that the network is not accurate enough, as it can be after the known failure modes of PINN training (Krishnapriyan et al., 2021; Wang et al., 2022). Trust-region methods have also been used to train networks (Xu & Darve, 2021; Cheng & Na, 2024), which concerns the training problem. To our knowledge, no prior work combines a neural surrogate, PDE-level error certificates for the objective and its gradient, and an event-triggered solver policy with convergence theory.