AMORTIZED EXPLOITABILITY CERTIFICATES FOR LEARNED RESPONSE EQUILIBRIA

Anonymous authors
Paper under double-blind review

ABSTRACT

A learned response map can reproduce observed best responses without certifying low regret in the game that generated them. We derive a playerwise exploitability certificate for games with compact convex product action sets and playerwise-convex costs. A regional audit bounds each player's true-gradient magnitude and the mismatch between the learned response displacement and the true pseudo-gradient. These bounds yield $\varepsilon(x) \le \max_i D_i \min\{\bar{M}_i, \rho_i(x) + \bar{\delta}_i\}$ throughout the audited region. Prediction loss does not control $\bar{\delta}_i$: curvature, cost normalization, and active constraints remain relevant even for a zero-error best-response predictor. A finite covering constructs the regional bounds only when valid regularity and noise bounds are supplied, and its cost grows exponentially with the region's intrinsic dimension. A fresh pointwise audit always favors the model-free candidate. Regional auditing avoids later game queries, but this reuse is not unique to the learned map: retaining the cover gradients gives a profile-dependent model-free certificate that dominates the learned candidate in our regional experiment. The remaining experiments show that contraction does not characterize low exploitability and that playerwise convexity is essential.

1 INTRODUCTION

Decision-making in multi-agent settings can be characterized as *non-cooperative games*, where heterogeneous and selfish entities, i.e. *agents or players*, act with independent decisions to optimize their own objective, irrespective of other players' goals. With applications that traverse socio-technical systems (Ekpo et al., 2026), regional science and transportation planning (Facchinei & Pang, 2003), multi-agent/robot path-finding (Zhang et al., 2024; Wang et al., 2024; Ren et al., 2021), and multi-agent reinforcement learning (Zhang et al., 2021), these games are often encumbered with incomplete or conflicting information and multiple conflicting goals. (Generalized) Nash equilibrium (Facchinei & Kanzow, 2010; Pavel & Hong, 2022; Li et al., 2022) is a solution concept used to characterize the strategic interactions and possible outcomes of such games. Modeling the players' interactions and analyzing the resulting equilibria beyond a small number of players proves intractable: the normal-form representation grows exponentially in the number of players, and even computing an approximate equilibrium is PPAD-hard (Rubinstein, 2015). Beyond tractability, for systems where cost functions are unavailable or players are boundedly rational, the Nash solution concept is practically ineffective for characterizing best responses.

This has motivated learning equilibria from players' interaction data using structured learning methods (Scarabaggio et al., 2024). These data-driven formulations extend fixed-point theorems that analyze solution existence under convexity (Kakutani, 1941) to response analysis from data. Relaxations such as the first-order Nash equilibrium notion of Scarabaggio et al. (2026) retain the players' objectives while weakening the equilibrium concept, so that each player's action is a fixed point of its own projected gradient, with existence following from algebraic topology rather than from convex analysis. Scarabaggio et al. (2024) replace the objectives, characterizing best responses as a composition of differentiable neural network layers fitted to players' observed actions. Feasibility is enforced by a projection of the network output onto each player's strategy set. Related formulations learn a parametric equilibrium map trained against the Nikaido–Isoda gap instead of regressing on observed play (Bemporad & Tatarenko, 2026a), or take the reacting object to be the data distribution rather than the response (Narang et al., 2023). This paper works in the second setting and certifies

against the first: we measure the residual of the learned response map, which costs one forward pass, and bound through it the projected-gradient residual whose zeros define the first-order notion.

Suppose that $\mathcal{N} := \{1, \dots, N\}$ indexes the players and player i chooses an action x_i from a nonempty compact convex set $X_i \subseteq \mathbb{R}^{n_i}$. Learning-based Nash equilibrium (LBNE) (Scarabaggio et al., 2024) fits a map $f_i(x_{-i})$ predicting player i's response to the other players' actions, x_{-i}. The *behavioral equilibrium* is a profile $x^\star$ for which $x_i^\star = f_i(x_{-i}^\star)$ for every $i \in \mathcal{N}$. Each f_i consists of an optional feature map, a response head, and a Euclidean projection onto X_i; when no feature transformation is needed, the feature map is the identity on x_{-i}. The projection enforces feasibility. Given the continuity of each f_i and the convexity and compactness of the product X, Scarabaggio et al. (2024) establish that the joint map $F = (f_1, \dots, f_N)$ admits a fixed point by Brouwer's theorem (Brouwer, 1911). However, convergence of the prescribed Picard–Banach iteration $x^{k+1} = F(x^k)$ is secured only under an aggregate contraction condition $\left(\sum_{i \in \mathcal{N}} \gamma_i^2\right)^{1/2} < 1$ on the playerwise Lipschitz constants γ_i, which excludes the nonexpansive case at which the iteration can fail to converge at all. Contraction hypotheses of this kind are replete across response-learning and decision-dependent formulations, where they inform uniqueness, a linear rate, and a perturbation bound $\varsigma/(1-q)$ under a uniform map error ς (Mignoni, 2025; Narang et al., 2023; Bauschke & Combettes, 2017).

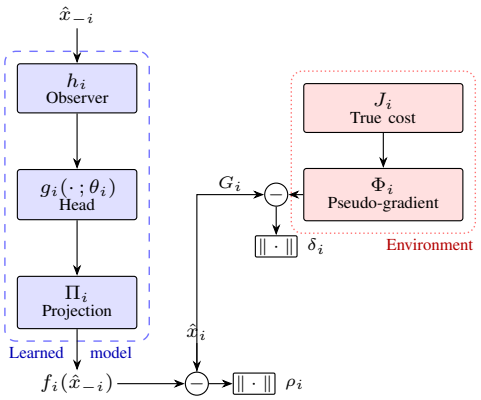


Fig. 1: **Response prediction and operator alignment.** The learned displacement $G_i(\hat{x}) = \hat{x}_i - f_i(\hat{x}_{-i})$ has residual $\rho_i(\hat{x}) = \|G_i(\hat{x})\|$. The operator-alignment error $\delta_i(\hat{x}) = \|G_i(\hat{x}) - \nabla_{x_i} J_i(\hat{x})\|$ compares that displacement with the true own-action gradient; it is not a response-prediction error.

These guarantees constrain the learned map rather than the game. A contractive F has a unique fixed point of the *learned* game, which need not be an equilibrium of the game that generated the data: the certificate bounds the sensitivity of a fitted network, whereas the quantity of interest is the regret a player would incur at the predicted profile. For arbitrary strategies $\hat{x} \in X$, a constant map $F \equiv \hat{x}$ is nonexpansive and has zero fixed-point residual at $\hat{x}$, while its exploitability is determined by the environment alone and is unconstrained by that residual. Enforcing contraction restricts the hypothesis class where the restriction is least affordable i.e. an interaction whose true best-response map is not contractive admits no contractive model since multi-agent retraining dynamics admit chaotic orbits (Piliouras & Yu, 2023).

While Scarabaggio et al. (2024) introduced the projected observer–head response map and showed that its fixed point solves a variational inequality of the learned game, Narang et al. (2023) proved uniqueness and linear convergence of repeated retraining to their decision-dependent equilibrium under an analogous ratio condition. In a similar vein, Bemporad & Tatarenko (2026a) trained a network against the Nikaido–Isoda gap (Nikaidô & Isoda, 1955) of a surrogate game, while Bemporad & Tatarenko (2026b) instead enforce monotonicity of the learned game by construction. In contrast, Gürkan & Pang (2009) characterized when equilibria of approximating games converge to equilibria of the true game; however, the results are asymptotic with no explicit constant, and Bergerault et al. (2026) proved that the Nash gap, the exploitability of a profile, is *not* determined uniformly over games by a measure-matching error, establishing a divergence between prediction loss and any equilibrium guarantee. Neither approach yields a number that can be evaluated at a given profile *in-silico*: the first certifies the model under a hypothesis that may fail on the interaction of interest, and the second delimits what fitting can achieve without quantifying the shortfall anywhere. A certificate should therefore bound regret in the true game of the prediction model's profile, and should (ideally) be finite when the learned map does not contract.

Write ρ_i for player i's fixed-point residual and δ_i for the operator-alignment error between the learned response operator and the true own-action cost gradient. For a region Ω, let $\bar{M}_i$ and $\bar{\delta}_i$ upper-bound the true-gradient magnitude and this discrepancy throughout Ω. We call the pair $(\bar{M}, \bar{\delta})$ a *regional envelope*. Figure 1 illustrates the profile-specific quantities. Our **contributions** are as follows:

(1) We certify the exploitability of every profile $\hat{x} \in \Omega$ with a playerwise envelope, without requiring contraction, monotonicity, or equilibrium uniqueness. (2) We derive a quadratic exploitability bound in the distance to an interior best response, while the general envelope certificate remains linear in $\rho_i + \bar{\delta}_i$. (3) We separate regional reuse from the learned map. A regional audit makes a model-free bound reusable without later evaluations. Retaining the cover gradients also gives a profile-dependent model-free certificate that dominates the learned candidate under the regularity construction used in our regional experiment. (4) Across two product-action games, seven response families, and 20 random seeds, the certificate upper-bounds computed exploitability in all 280 in-scope configurations. In a separate sweep of 1,000 configurations, a small-gain contraction test is neither necessary nor sufficient for low exploitability, showing that contraction and equilibrium quality measure different properties.

The rest of this paper is organized as follows: Section 2 formulates the game and its learned response map, Section 3 states the certificate, Section 4 gives the envelope and its cost, and Section 5 reports the measurements. Proofs of our main results, discussions and limitations, as well as additional supporting empirical results are available in the appendices.

2 Problem Formulation

This section states the assumptions, fixes notation, and introduces the objects used in our results.

2.1 Notations and Preliminaries

Let $\mathcal{N} = \{1, \dots, N\}$ denote players' indices. For each player i, let x_i denote its action in a nonempty compact convex set $X_i \subseteq \mathbb{R}^{n_i}$. Player i observes other players' actions $x_{-i} = (x_j)_{j \neq i}$ and incurs a local cost $J_i : X \to \mathbb{R}$ with $J = (J_1, \dots, J_N)$ as a game's cost profile. The joint action set is the product $X = \prod_{i \in \mathcal{N}} X_i \subseteq \mathbb{R}^d$, $d = \sum_{i \in \mathcal{N}} n_i$, and $X_{-i} = \prod_{j \neq i} X_j$. The action profile $x = (x_1, \dots, x_N) \in X$ is also written as $x = (x_i, x_{-i})$ when player i's action is singled out. *We reserve $x^\star$ for an equilibrium and $\hat{x}$ for an arbitrary strategy profile at which a certificate is evaluated.* For vectors $v_1, \dots, v_N$ we write $\mathrm{col}(v_1, \dots, v_N)$ for the column vector obtained by stacking them. For a closed convex set S, the Euclidean projection $\Pi_S(x) := \arg\min_{y \in S} \|x - y\|$ is the unique element of S nearest x; since X is a product, $\Pi_X = (\Pi_{X_1}, \dots, \Pi_{X_N})$ acts blockwise. For differentiable J_i we write $\nabla_{x_i} J_i(x)$ for its gradient in player i's own action.

We now state the assumptions that underlie our theoretical machinery.

Assumption 2.1 (Compact convex product actions)**.** Each action set X_i is nonempty, compact, and convex, and the joint action set is the product $X = \prod_{i \in \mathcal{N}} X_i$.

Assumption 2.2 (Playerwise convexity)**.** For each player i and each $x_{-i} \in X_{-i}$, the cost $J_i(\cdot, x_{-i})$ is convex and continuously differentiable on X_i. The *pseudo-gradient* stacks the players' own-action gradients,

$$\Phi(x) \triangleq \mathrm{col}\big(\Phi_1(x), \dots, \Phi_N(x)\big) \in \mathbb{R}^d, \qquad \Phi_i(x) \triangleq \nabla_{x_i} J_i(x), \tag{1}$$

so that $\Phi_i(x) \in \mathbb{R}^{n_i}$ is player i's block of the pseudo-gradient.

We use dimensionless action and cost coordinates and fix the projected-gradient step to one. Thus x_i, $G_i(x)$, and $\Phi_i(x)$ share the numerical scale used below. This normalization is part of the certificate: multiplying a player's cost by a positive constant leaves its best responses unchanged but rescales its gradient, exploitability, and operator-alignment error. With physical units, the same development replaces Φ_i by $\gamma_i \Phi_i$ for a stated positive step size γ_i; this paper fixes $\gamma_i = 1$.

Players' regret measures a profile's quality. The *exploitability* of $x \in X$, equivalently the max-over-players variant of the (Nikaidô & Isoda, 1955) gap, whose original form sums over players is

$$\varepsilon(x) \triangleq \max_{i \in \mathcal{N}} \Big[J_i(x) - \min_{y_i \in X_i} J_i(y_i, x_{-i}) \Big], \tag{2}$$

where each bracket is nonnegative, since $x_i \in X_i$ is feasible for the inner minimization. Hence $x \in X$ is a Nash equilibrium when $\varepsilon(x) = 0$, and $\varepsilon(x)$ is the smallest tolerance for which x is an ε-Nash equilibrium (Radner, 1980).

Definition 2.1 (Variational inequality)**.** Let $\mathcal{K} \subseteq \mathbb{R}^d$ be nonempty and let T map $\mathbb{R}^d$ into itself. The *variational inequality problem*, $\mathrm{VI}(\mathcal{K}, T)$, finds a vector $x^\star \in \mathcal{K}$ such that

$$T(x^\star)^\top (y - x^\star) \geq 0 \quad \text{for all } y \in \mathcal{K}. \tag{3}$$

Under Assumption 2.1 and 2.2 the equilibria coincide with the solutions of $\mathrm{VI}(X, \Phi)$ (Rosen, 1965; Facchinei & Pang, 2010). The *natural map* of $\mathrm{VI}(\mathcal{K}, T)$ is $x \mapsto x - \Pi_{\mathcal{K}}\big(x - T(x)\big)$, which vanishes at the solutions of (3). This is the *natural residual* (Facchinei & Pang, 2003, §1.5).

Definition 2.2 (Learned Response Map)**.** Given a dataset of players' observed actions, player i's response map may first apply an *observer* h_i that summarizes the other players' actions in a feature vector $h_i(x_{-i})$. When no feature transformation is needed, h_i is the identity on the concatenated rival actions. A *response head* $g_i(\cdot\,; \theta_i)$, with parameters θ_i, maps these features to an action, and projection onto X_i enforces feasibility:

$$f_i(x_{-i}; \theta_i) = \Pi_{X_i}\big(g_i(h_i(x_{-i}); \theta_i)\big). \tag{4}$$

Writing $F(x) = \mathrm{col}\,\big(f_1(x_{-1}), \ldots, f_N(x_{-N})\big)$ for the joint response of all players, a *behavioral equilibrium* is a profile $x^\star \in X$ for which every player's prediction reproduces its own input,

$$x_i^\star = f_i(x_{-i}^\star; \theta_i) \quad \text{for all } i \in \mathcal{N}, \tag{5}$$

equivalently $x^\star = F(x^\star)$ (Scarabaggio et al., 2024).

Because Π_{X_i} maps into X_i, we have $F(\mathbb{R}^d) \subseteq X$ and in particular $F : X \to X$, and because the projection onto a closed convex set is nonexpansive (Facchinei & Pang, 2003, §1.5), continuity of f_i reduces to continuity of the learned part $g_i \circ h_i$. Under Assum. 2.1 the joint action set is nonempty, compact and convex, so that Brouwer's theorem (Brouwer, 1911) supplies a profile satisfying (5) whenever each $g_i \circ h_i$ is continuous. The *induced operator* of the projected response map is

$$G(x) \triangleq x - F(x), \tag{6}$$

so that $x - G(x) = F(x) \in X$. Thus, $F(x) = \Pi_X\big(x - G(x)\big)$. The fixed-point residual $x - F(x) = G(x)$ is thus the natural map of $\mathrm{VI}(X, G)$ (Facchinei & Pang, 2003, §1.5). The learned map is connected to the game via (i) the fixed-point residual of F; (ii) the operator-consistency error between (6) and (1); and (iii) the natural residual of the variational inequality $\mathrm{VI}(X, \Phi)$. Only (i) can be computed without access to the game; any certificate that uses (ii) or (iii) must query the environment.

Small prediction loss can make the operator-consistency error uncontrollable: a response map fitted to observed actions constrains G where data was observed, while Φ enters the exploitability (2) through its values on the region being certified. Bergerault et al. (2026) make an analogous point for multi-agent imitation learning and recover imitation-to-Nash bounds only when the demonstrated expert has additional structure, such as the dominant-strategy behavior or continuity of best responses around that expert.

3 An Amortized Exploitability Certificate

This section constructs the amortized exploitability certificate. We first bound each player's regret using the true projected-gradient residual and local, player-specific constants. We then relate that residual to the mismatch of the learned response and its discrepancy from the true game. Finally, we replace the profile-specific game quantities by upper bounds that hold throughout a region and can be reused at later profiles. All proofs are in Appendix C;

For a profile $x \in X$ write the gradient magnitude, blockwise residual, and the diameter of each player's block as

$$M_i(x) \triangleq \|\Phi_i(x)\|, \;\; D_i \triangleq \operatorname{diam} X_i, \;\; r_i(x) \triangleq \big\|x_i - \Pi_{X_i}\big(x_i - \Phi_i(x)\big)\big\|, \tag{7}$$

where $\operatorname{diam} X_i = \sup_{u,v \in X_i} \|u - v\|$ is finite by Assum. 2.1. Block i's $r_i(x) := \|\dot{x}\|$, where $\dot{x}$ is projected pseudo-gradient flow $\dot{x} = \Pi_X\big(x - \Phi(x)\big) - x$. Every $r_i(x)$ vanishes at the flow's equilibrium points.

Lemma 3.1 (Blockwise gap bound). *Suppose that Assumption 2.1 and 2.2 hold. Then for every $x \in X$,*

$$\varepsilon(x) \;\le\; \max_{i\in\mathcal{N}} \min\big\{\big(M_i(x)+D_i\big)\, r_i(x),\;\; M_i(x)\, D_i\big\}. \tag{8}$$

Lemma 3.1 follows from the projection inequality (Facchinei & Pang, 2003, §1.5). $M_i(x)$ and D_i are local and blockwise constants; we evaluate M_i at x. Error bounds of this kind control the distance from x to the solution set, for which the constant exists only under strong monotonicity (Yamashita & Fukushima, 1997). Equation (8) instead controls exploitability. Its model-free candidate is $M_i(x)D_i$.

3.1 Transfer to a learned map

Write $G_i(x) = x_i - f_i(x_{-i})$ for the i-th block of the induced operator (6), and define the blockwise residual ρ_i and the operator-alignment error δ_i by

$$\rho_i(x) \triangleq \big\|G_i(x)\big\|, \qquad \delta_i(x) \triangleq \big\|G_i(x) - \Phi_i(x)\big\|, \tag{9}$$

respectively. The quantity δ_i compares a response displacement with the true gradient under the fixed unit-step normalization; it is not prediction error relative to a best response.

Lemma 3.2 (Blockwise operator transfer). *Let Assum. 2.1 hold and let Φ_i be defined on X. Then, for every $x \in X$ and every player i,*

$$r_i(x) \le \rho_i(x) + \delta_i(x), \qquad\qquad M_i(x) \le \rho_i(x) + \delta_i(x). \tag{10}$$

3.2 Amortizing game queries over a region

A regional envelope immediately supplies the model-free candidate $\bar{M}_i D_i$ throughout Ω; applying it later requires neither a game query nor a model evaluation. Evaluating the learned map additionally supplies $\rho_i(x)$ and hence a profile-dependent candidate. Both candidates reuse the same audit; the learned map does not create the reuse.

Definition 3.1 (Regional envelope). Let $\Omega \subseteq X$ be closed. A *regional envelope* of a trained model on Ω is a pair of vectors $(\bar{M}, \bar{\delta})$ that uniformly upper-bound, for every player i, the true-gradient magnitude and the model–game discrepancy:

$$\bar{M}_i \ge \sup_{x\in\Omega} M_i(x), \qquad \bar{\delta}_i \ge \sup_{x\in\Omega} \delta_i(x). \tag{11}$$

Call $D_i\big(\rho_i(\hat{x}) + \bar{\delta}_i\big)$ the *learned-alignment candidate* and $\bar{M}_i D_i$ the *regional model-free candidate*.

Theorem 3.1 (Amortized exploitability certificate). *Let Assumption 2.1 and 2.2 hold, let $(\bar{M}, \bar{\delta})$ be a regional envelope of a trained model on $\Omega \subseteq X$, and let F have the form (4). Every profile $\hat{x} \in \Omega$ satisfies*

$$\varepsilon(\hat{x}) \le \max_{i\in\mathcal{N}} D_i \min\big\{\bar{M}_i,\; \rho_i(\hat{x}) + \bar{\delta}_i\big\} \tag{12}$$

$$= \max_{i\in\mathcal{N}} \min\big\{\bar{M}_i D_i,\; D_i\big(\rho_i(\hat{x}) + \bar{\delta}_i\big)\big\}\,. \tag{13}$$

The regional model-free candidate requires no evaluation at $\hat{x}$; the learned-alignment candidate requires only the residuals $\rho_i(\hat{x})$.

Proposition 3.1 (The learned map cannot improve a pointwise certificate). *Specialize Theorem 3.1 to the singleton region $\Omega = \{\hat{x}\}$, with $\bar{M}_i = M_i(\hat{x})$ and $\bar{\delta}_i = \delta_i(\hat{x})$. Then, for every player,*

$$D_i \min\{M_i(\hat{x}),\; \rho_i(\hat{x}) + \delta_i(\hat{x})\} = M_i(\hat{x}) D_i. \tag{14}$$

Thus the pointwise certificate is $\max_i M_i(\hat{x}) D_i$.

The reverse triangle inequality gives $M_i(\hat{x}) \le \rho_i(\hat{x}) + \delta_i(\hat{x})$, so that a freshly evaluated learned candidate cannot improve the model-free candidate. Regional reuse is a separate regime.

Corollary 3.1 (Non-vacuity threshold). *Fix a player i. If $D_i > 0$, then the learned-alignment candidate is strictly smaller than the regional model-free candidate if and only if*

$$\rho_i(\hat{x}) + \bar{\delta}_i < \bar{M}_i. \tag{15}$$

If $D_i = 0$, both candidates vanish.

Let $B_{\mathrm{MF}} = \max_i \bar{M}_i D_i$. If $B_{\mathrm{MF}} > 0$, the full certificate is strictly smaller than B_{MF} if and only if the threshold (15) holds for every player attaining B_{MF}. Training can reduce ρ_i or improve operator alignment, but it cannot change the audited threshold $\bar{M}_i$.

Proposition 3.2 (A spurious stationary profile). *Let $X_i = [-c, c]$ with $c > (1-\kappa)^{-1/2}$ and let*

$$J_i(x_i, x_{-i}) = \tfrac{1}{4}x_i^4 - \tfrac{1}{2}w_i(x_{-i})\,x_i^2, \qquad w_i(x_{-i}) = 1 + \tfrac{\kappa}{N-1}\sum_{j\neq i} x_j^2, \tag{16}$$

with $\kappa \in [0, 1)$. Then $\Phi(0) = 0$, hence $r_i(0) = 0$ for every player, while $\varepsilon(0) = \frac{1}{4}$.

The origin is a strict local maximum of each $J_i(\cdot, x_{-i})$ — the conditional mean of play whenever the data-generating process is symmetric across the two minimizers $\pm\sqrt{w_i(x_{-i})}$.

4 Envelope Costs

This section states the cost and assumptions of constructing the regional envelope in Definition 3.1. A *gradient oracle* returns $\Phi_i(x)$, while a *cost oracle* returns $J_i(x) + Z$, where the query errors Z are independent, mean-zero, and σ-sub-Gaussian. Let $e_{i,k}$ be the unit vector for coordinate k of player i's action block, so that a radius $\tau > 0$ implies

$$\big[\widehat{\Phi}_\tau(x)\big]_{i,k} = \frac{J_i(x + \tau e_{i,k}) - J_i(x - \tau e_{i,k})}{2\tau}, \tag{17}$$

Noisy oracle returns give this estimate $O(\tau^2)$ bias and $O(\sigma/\tau)$ noise.

4.1 Regional envelopes

Finite probes control a continuum only through supplied moduli of continuity. For every player, let $L_{\delta,i}$ and $L_{M,i}$ be valid Lipschitz upper bounds for δ_i and M_i on Ω. Because $G_i(x) = x_i - f_i(x_{-i})$, a valid bound for $L_{\delta,i}$ combines the identity block, a certified Lipschitz bound for the projected response f_i, and a valid bound for Φ_i.

Proposition 4.1 (Envelope from a finite covering). *Let $\Omega \subseteq X$ be nonempty and let $\mathcal{Z} \subseteq \Omega$ be a supplied finite Euclidean η-cover: every $x \in \Omega$ lies within distance η of some $z \in \mathcal{Z}$. For every player i, suppose supplied constants $L_{\delta,i}$ and $L_{M,i}$ are valid Lipschitz bounds for δ_i and M_i on Ω. Let $\tau > 0$, suppose every required probe segment lies in a domain on which J_i has a C^3 extension satisfying $|\partial^3 J_i / \partial x_{i,k}^3| \le B_3$, and let cost-query errors be independent, mean-zero, and σ-sub-Gaussian. Define*

$$E_i(\tau, \alpha) \triangleq \tfrac{1}{6}B_3\tau^2\sqrt{n_i} + \tfrac{\sigma}{\tau}\sqrt{n_i \log(2d|\mathcal{Z}|/\alpha)}. \tag{18}$$

After the $2d|\mathcal{Z}|$ scalar cost evaluations in (17), the event of probability at least $1-\alpha$ simultaneously yields, for every player,

$$\bar{M}_i = \max_{z\in\mathcal{Z}} \big\|\widehat{\Phi}_{\tau,i}(z)\big\| + L_{M,i}\eta + E_i(\tau,\alpha), \tag{19}$$

$$\bar{\delta}_i = \max_{z\in\mathcal{Z}} \big\|G_i(z) - \widehat{\Phi}_{\tau,i}(z)\big\| + L_{\delta,i}\eta + E_i(\tau,\alpha). \tag{20}$$

These quantities satisfy both inequalities in (11) for all players.

One complete instance. Appendix D gives an end-to-end two-firm Cournot example. A four-point cover and analytic regularity constants satisfy every assumption above; 16 scalar cost queries give, at the reported profile,

$$0.00125 = \varepsilon(q) \le 0.0361 < 0.0624 < 0.0924, \tag{21}$$

where the last three values are the theorem certificate, the retained-cover model-free bound, and the constant regional model-free bound.

Remark 4.1 (Probes at the boundary). Since every f_i ends in a projection onto X_i, certified profiles frequently lie on ∂X, and then $x \pm \tau e_{i,k}$ leaves X, where the costs are neither defined nor assumed smooth. Proposition 4.1 uses a C^3 extension to a τ-neighborhood of X, so that the central Taylor

expansion remains valid along the complete probe segment. When such an extension is unavailable, one may instead use a one-sided difference on coordinates within τ of the boundary. Its bias is $O(B_2\tau)$ when B_2 bounds the second derivative. The balanced radius then scales as $\tau \asymp (\sigma/B_2)^{1/2}$, and the corresponding penalty is $O(\sigma^{1/2}B_2^{1/2})$, up to the $\sqrt{n_i}$ and logarithmic factors in (19).

Balancing the bias and noise terms in (19) at $\tau \asymp \left(\sigma\sqrt{\log(2d|\mathcal{Z}|/\alpha)}/B_3\right)^{1/3}$ leaves a penalty of order $\sqrt{n_i}\left(\log(2d|\mathcal{Z}|/\alpha)\right)^{1/3}\sigma^{2/3}B_3^{1/3}$, which vanishes with the oracle noise. The covering term $L_{\delta,i}\eta$ is a modeling cost, and at fixed relative resolution the covering number of a box-like m-dimensional region is $O((\operatorname{diam}\Omega/\eta)^m)$. A rigorous envelope is therefore affordable only for regions with small covering numbers, such as a line segment between two candidate profiles, the trajectory of a response iteration, or a neighborhood of a single deployment point.

Remark 4.2 (The finite-query construction is conditional on supplied regularity). Proposition 4.1 requires valid upper bounds $L_{\delta,i}$, B_3, and $L_{M,i}$, together with a justified independent mean-zero sub-Gaussian noise model. Finite cost values cannot establish these uniform regional quantities because they do not control behavior between queried points. Weight-derived bounds can control the learned response f_i through the composite $g_i \circ h_i$ and the nonexpansive projection. If these inputs are unavailable, the finite-cover construction is a diagnostic rather than a uniform certificate.

5 Experiments

The experiments test certificate validity, playerwise scaling, regional reuse, and the separation between contraction and exploitability. Reference best responses are analytic or globally verified. We fit seven response families: an affine baseline; two-layer 32–32 tanh, ReLU, spectral-normalized, and Almost-Orthogonal-Layer networks (Prach & Lampert, 2022); and 64–64 conditional diffusion and flow models.

Bound validity and tightness. A tightness study fits seven response families over 20 seeds on three games: a coupled quadratic game whose coupling parameter fixes the monotonicity modulus, a capacitated Cournot game whose capacities bind at equilibrium, and a Rosen game carrying one constraint shared across players. The quadratic and Cournot games have product action sets: $2 \times 7 \times 20 = 280$ cells. We also use a Rosen game as a stress diagnostic apart from theorem coverage. The shared game has an analytically specified best-response oracle; however, its variational equilibrium is computed to the numerical tolerance.

For each profile we compare the following scalar upper bounds: (i) *model-free candidate*, $B_{\mathrm{MF}} = \max_i \bar{M}_i D_i$; (ii) the *learned-alignment candidate*, $B_{\mathrm{LA}} = \max_i D_i(\rho_i + \bar{\delta}_i)$; and (iii) a *direct-probe candidate* queries the true gradient at the evaluated profile, computes the projected-gradient residual r_i, and reports $B_{\mathrm{DP}} = \max_i(M_i + D_i)r_i$. Note that the blockwise certificate in (12) takes the smaller candidate separately for each player before maximizing.

The analysis contains 420 game–model–seed configurations. The 280 product-set configurations satisfy the theorem assumptions; the 140 shared-constraint configurations are stress diagnostics. The model-free candidate determines every pointwise blockwise certificate, as Proposition 3.1 predicts. Its median tightness is 108.7, compared with 134.7 for the direct-probe candidate. Their median paired difference is -0.001165 in $\log_{10}$ tightness ($p_{\mathrm{Holm}} = 0.1519$), so that this test detects no difference. The model-free candidate is the standard convexity/Cauchy–Schwarz bound.

Playerwise scaling. The scaling study varies the player count N while holding the observational-data fixed. At a fixed displacement from the reference equilibrium, the aggregate ratio of certificate to true exploitability increases from 38.15 at $N = 2$ to 1157 at $N = 256$. The corresponding blockwise ratio changes from 13.92 to 6.628 (see Fig. 2) at both action dimensions. In Fig. 2, the dash-dot rule marks where the response fit are under-determined. Table 1 lists scalar actions. The aggregated form grows with N because exploitability is a maximum over players whereas an aggregated residual is a joint norm. The blockwise reports tightness of the model-free candidate.

The aggregate form uses one global constant and a joint residual, whereas the playerwise form uses (7); here the blockwise series equals the model-free candidate. The fixed 240-transition budget makes

Table 1: **Certificate-exploitability ratio as players are added.**

Players N	2	8	64	128	256
Aggregated: $(M+D)(\rho+\delta)$	38.15	84.85	431.1	696.1	1157
Blockwise: $\max_i \bar{M}_i D_i$	13.92	10.78	7.862	7.176	6.628

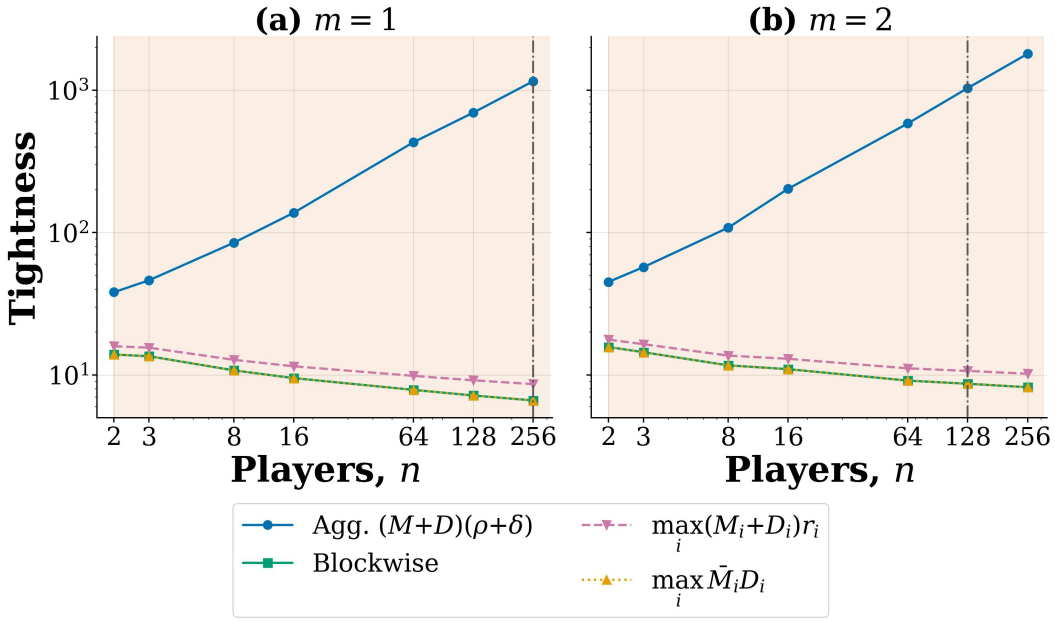


Fig. 2: **Constants degrade with the player count.** Coupled quadratic game, (a) scalar and (b) two-dimensional actions, medians over 20 seeds.

the largest fits underdetermined, which can inflate learned-map residuals but not the model-free candidate.

The bound-and-tightness and coupling-sweep experiments supply 240 noisy observational transitions; neural responders use Adam with batch size 128 for 200 epochs at learning rate 3×10^{-3} (AOL: 5×10^{-3}), with least-squares, denoising, and flow-matching losses as appropriate. The player-count scaling experiment uses its stated reduced training budget. A cell is one game, response family, and seed; comparisons are paired within cells, with 10,000-resample bootstrap intervals (Efron, 1979), Wilcoxon signed-rank tests (Wilcoxon, 1945) and Holm's step-down correction (Holm, 1979). The full protocol appears in Appendix G.

Regional audits trade up-front queries for later reuse. A regional audit queries each player's gradient block on an η-cover $\mathcal{Z}$ of Ω once and then certifies later profiles without another game query. A fresh pointwise audit costs N gradient-block queries at each profile, whereas the regional audit costs $N|\mathcal{Z}|$ queries once. The costs are equal after $|\mathcal{Z}|$ profiles, and regional reuse saves queries beyond that point. By Proposition 3.1, the fresh pointwise certificate reduces here to the model-free bound $\max_i M_i(x) D_i$.

The experiment uses regions embedded in the six-dimensional joint action space. Let D_{box} denote the diagonal length of the joint bounding box. The segment connects the learned fixed point $\hat{x}$ to

$$\Pi_X(x^\star + 0.2 D_{\mathrm{box}} u)\,, \tag{22}$$

for a seeded random unit vector u. A coordinate box of intrinsic dimension $k \in \{1, 2, 3\}$ varies its first k coordinates within $\pm 5\%$ of their coordinate ranges around $\hat{x}$, clipped to the action bounds. Each seed supplies 256 profiles drawn uniformly from each region.

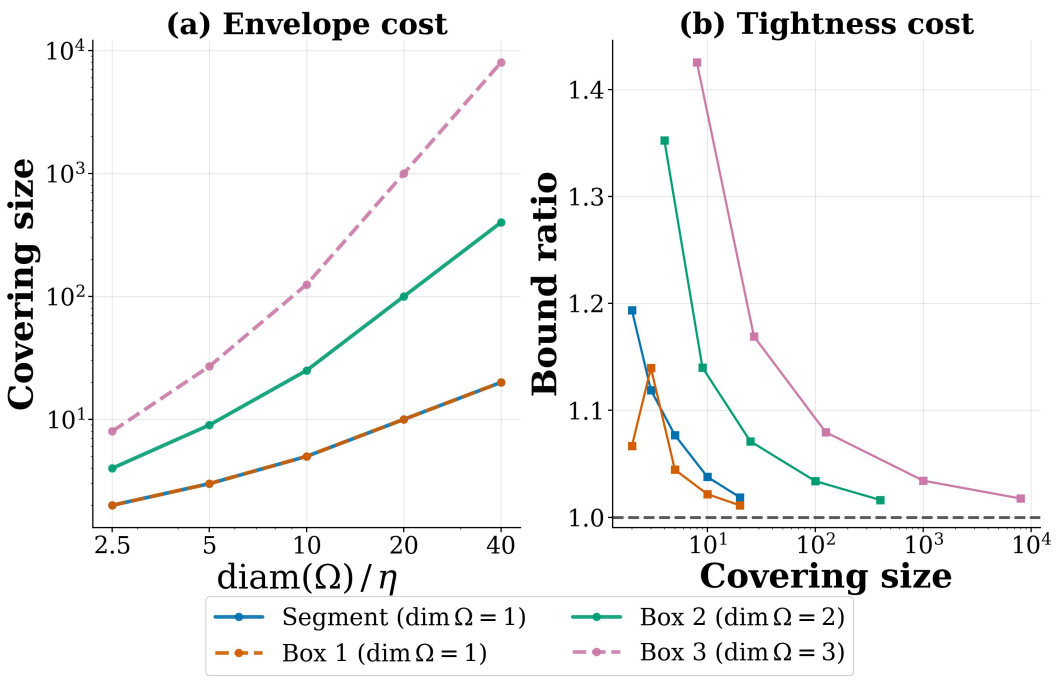


Fig. 3: **Cost and tightness of regional reuse.** Three-player quadratic product-set game with two-dimensional actions and a two-hidden-layer 32–32 tanh response model. Curves show a segment and coordinate boxes of intrinsic dimension $k \in \{1, 2, 3\}$. Values are medians over ten seeds, each with 256 held-out profiles per region. **(a)** Cover size versus $\operatorname{diam}(\Omega)/\eta$. **(b)** Regional-to-pointwise model-free certificate ratio versus cover size; the dashed line marks equal bounds.

At $\eta/\operatorname{diam}\Omega = 0.025$, the median cover size is 20 for the segment and 8,000 for the three-dimensional box. The cover and Lipschitz inflation establish the regional envelope; the held-out evaluation found no violations among the 2,560 profiles in either region. The growth in cover size limits this method to regions with small intrinsic dimension that will be reused often.

Retaining the cover gradients gives a stronger reusable model-free candidate. At a later profile x, define

$$U_i(x) = \min_{z \in \mathcal{Z}} \{M_i(z) + L_{M,i}\|x - z\|\}. \tag{23}$$

Then $M_i(x) \le U_i(x)$, so that $D_i U_i(x)$ is valid without another game query. In this experiment $L_{\delta,i} = L_{G,i} + L_{M,i}$; the triangle inequality therefore makes $D_i U_i(x)$ no larger than the learned-alignment candidate. At relative resolution 0.025, the median same-cover model-free bound is 1.019 times the fresh pointwise bound on the segment and 1.018 times it on the three-dimensional box. The theorem certificate is respectively 1.393 and 1.604 times the same-cover bound. The same-cover bound is valid and no larger than the learned candidate at every held-out profile. The regional experiment consequently demonstrates query reuse, not an empirical tightening attributable to the learned map. Its costs also exclude storage and nearest-cover-point search.

The audit converts cover-point maxima into regional bounds using valid Lipschitz moduli. The affine quadratic pseudo-gradient has modulus $L_{M,i} = \|A_{i,:}\|_2$. If L_{f_i} is the responder's weight-derived Lipschitz bound, then $G_i(x) = x_i - f_i(x_{-i})$ admits $L_{\delta,i} = \sqrt{1 + L_{f_i}^2} + \|A_{i,:}\|_2$. For a k-dimensional coordinate box with edge lengths s_j, a midpoint grid with $n_j = \left\lceil s_j\sqrt{k}/(2\eta) \right\rceil$ has cell half-diagonal at most η and is an η-cover. This gradient-oracle experiment tests the covering, Lipschitz-inflation, and reuse components of Proposition 4.1, but not finite-difference bias or oracle noise. The responder trains on actions; the independent audit uses game gradients and externally justified Lipschitz bounds.

USE OF AI TOOLS

In this work, generative AI assistance was used to augment our literature review through search and retrieval of related work, edit prose, and edit and audit the experiment code. We did not use generative AI tools to produce the theoretical contributions or empirical claims. We reviewed all AI-assisted work, including literature-survey and prose edits. The authors take full responsibility for the final content of the paper, including text, claims, and artifacts produced with the aid of generative AI.

REFERENCES

Yakov Babichenko. Query Complexity of Approximate Nash Equilibria. *Journal of the ACM*, 63(4): 1–24, 2016. doi: 10.1145/2908734.

Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, Cham, 2nd edition, 2017. doi: 10.1007/978-3-319-48311-5.

Giuseppe Belgioioso, Peng Yi, Sergio Grammatico, and Lacra Pavel. Distributed Generalized Nash Equilibrium Seeking: An Operator-Theoretic Perspective. *IEEE Control Systems Magazine*, 42(4): 87–102, 2022. doi: 10.1109/MCS.2022.3171480.

Alberto Bemporad and Tatiana Tatarenko. Learning Approximate Solutions to Multiparametric Generalized Nash Equilibrium Problems, 2026a. arXiv preprint arXiv:2605.28757.

Alberto Bemporad and Tatiana Tatarenko. Learning Parametric Monotone Games, 2026b. arXiv preprint arXiv:2609.02494.

Antoine Bergerault, Volkan Cevher, and Negar Mehr. Matching Multiple Experts: On the Exploitability of Multi-Agent Imitation Learning, 2026. arXiv preprint arXiv:2602.21020.

Jérôme Bolte and Edouard Pauwels. Conservative set valued fields, automatic differentiation, stochastic gradient methods and deep learning. *Mathematical Programming*, 188(1):19–51, 2021. doi: 10.1007/s10107-020-01501-5.

L. E. J. Brouwer. Über Abbildung von Mannigfaltigkeiten. *Mathematische Annalen*, 71(1):97–115, 1911. doi: 10.1007/BF01456931.

Frank H. Clarke. *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, New York, 1983. Reprinted as SIAM Classics in Applied Mathematics 5, 1990, doi:10.1137/1.9781611971309.

Stella Dafermos. Sensitivity analysis in variational inequalities. *Mathematics of Operations Research*, 13(3):421–434, 1988. doi: 10.1287/moor.13.3.421.

Alexander Davydov, Saber Jafarpour, and Francesco Bullo. Non-Euclidean Contraction Theory for Robust Nonlinear Stability. *IEEE Transactions on Automatic Control*, 67(12):6667–6681, 2022. doi: 10.1109/TAC.2022.3183966.

Roy Dong, Heling Zhang, and Lillian J. Ratliff. Approximate Regions of Attraction in Learning with Decision-Dependent Distributions. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 206 of *Proceedings of Machine Learning Research*, pp. 11172–11184. PMLR, 2023.

Asen L. Dontchev and R. Tyrrell Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization*, 6(4):1087–1105, 1996. doi: 10.1137/S1052623495284029.

B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.1214/aos/1176344552.

Promise Ekpo, Angelique Taylor, and Lekan Molu. A Generalized Nash Equilibrium-Seeking Scheme for Trauma Resuscitation. In *Proceedings of the IFAC World Congress*, Busan, Republic of Korea, 2026.

Francisco Facchinei and Christian Kanzow. Generalized Nash Equilibrium Problems. *Annals of Operations Research*, 175(1):177–211, 2010. doi: 10.1007/s10479-009-0653-x.

Francisco Facchinei and Jong-Shi Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2003. doi: 10.1007/b97544. Two volumes, continuously paginated pp. 1–1234.

Francisco Facchinei and Jong-Shi Pang. Nash Equilibria: The Variational Approach. In Daniel P. Palomar and Yonina C. Eldar (eds.), *Convex Optimization in Signal Processing and Communications*, pp. 443–493. Cambridge University Press, Cambridge, 2010. doi: 10.1017/CBO9780511804458.013. Published online 3 December 2009.

Mahyar Fazlyab, Alexander Robey, Hamed Hassani, Manfred Morari, and George J. Pappas. Efficient and Accurate Estimation of Lipschitz Constants for Deep Neural Networks. In *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, 2019.

Masao Fukushima. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming*, 53(1–3):99–110, 1992. doi: 10.1007/BF01585696.

Gül Gürkan and Jong-Shi Pang. Approximations of Nash Equilibria. *Mathematical Programming*, 117(1–2):223–253, 2009. doi: 10.1007/s10107-007-0156-y. Published online 21 August 2007.

John C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23, 1973. doi: 10.1007/BF01737554.

Sergiu Hart and Andreu Mas-Colell. Uncoupled Dynamics Do Not Lead to Nash Equilibrium. *American Economic Review*, 93(5):1830–1836, 2003. doi: 10.1257/000282803322655581.

Aaron J. Havens, Alexandre Araujo, Siddharth Garg, Farshad Khorrami, and Bin Hu. Exploiting Connections between Lipschitz Structures for Certifiably Robust Deep Equilibrium Models. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, pp. 21658–21674, 2023.

Michael D. Hirsch, Christos H. Papadimitriou, and Stephen A. Vavasis. Exponential lower bounds for finding Brouwer fix points. *Journal of Complexity*, 5(4):379–416, 1989. doi: 10.1016/0885-064X(89)90017-4.

Josef Hofbauer and Ed Hopkins. Learning in perturbed asymmetric games. *Games and Economic Behavior*, 52(1):133–152, 2005. doi: 10.1016/j.geb.2004.06.006.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979. doi: 10.2307/4615733.

Guoqiang Hu, Yipeng Pang, Chao Sun, and Yiguang Hong. Distributed Nash Equilibrium Seeking: Continuous-Time Control-Theoretic Approaches. *IEEE Control Systems Magazine*, 42(4):68–86, 2022. doi: 10.1109/MCS.2022.3171479.

Saber Jafarpour, Alexander Davydov, Anton V. Proskurnikov, and Francesco Bullo. Robust Implicit Networks via Non-Euclidean Contractions. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, pp. 9857–9868, 2021.

Shizuo Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Mathematical Journal*, 8 (3):457–459, 1941. doi: 10.1215/S0012-7094-41-00838-4.

Torbjörn Larsson and Michael Patriksson. A class of gap functions for variational inequalities. *Mathematical Programming*, 64(1–3):53–79, 1994. doi: 10.1007/BF01582565.

Tao Li, Guanze Peng, Quanyan Zhu, and Tamer Başar. The confluence of networks, games, and learning a game-theoretic framework for multiagent decision making over networks. *IEEE Control Systems Magazine*, 42(4):35–67, 2022. doi: 10.1109/MCS.2022.3171478.

Eric Mazumdar, Lillian J. Ratliff, and S. Shankar Sastry. On Gradient-Based Learning in Continuous Games. *SIAM Journal on Mathematics of Data Science*, 2(1):103–131, 2020. doi: 10.1137/18M1231298.

Richard D. McKelvey and Thomas R. Palfrey. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, 10(1):6–38, 1995. doi: 10.1006/game.1995.1023.

Panayotis Mertikopoulos and Zhengyuan Zhou. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173(1–2):465–507, 2018. doi: 10.1007/s10107-018-1254-8.

Nicola Mignoni. *Non-Cooperative Game Theoretical Control for Green and Efficient Energy Communities*. PhD thesis, Politecnico di Bari, Bari, Italy, 2025. URL https://hdl.handle.net/11589/284440. Chapter 6.

Adhyyan Narang, Evan Faulkner, Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J. Ratliff. Multiplayer Performative Prediction: Learning in Decision-Dependent Games. *Journal of Machine Learning Research*, 24(202):1–56, 2023. URL https://jmlr.org/papers/v24/22-0131.html.

Hukukane Nikaidô and Kazuo Isoda. Note on non-cooperative convex game. *Pacific Journal of Mathematics*, 5(5):807–815, 1955. doi: 10.2140/pjm.1955.5.807.

Jong-Shi Pang. Error bounds in mathematical programming. *Mathematical Programming*, 79(1–3): 299–332, 1997. doi: 10.1007/BF02614322.

Lacra Pavel and Yiguang Hong. Distributed games and nash equilibrium seeking in multiagent systems over networks: An introduction to the special issue. *IEEE Control Systems Magazine*, 42 (4):32–34, 2022. doi: 10.1109/MCS.2022.3171477.

Ji-Ming Peng. Equivalence of variational inequality problems to unconstrained minimization. *Mathematical Programming*, 78(3):347–355, 1997. doi: 10.1007/BF02614360.

Juan C. Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative Prediction. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7599–7609. PMLR, 2020. URL https://proceedings.mlr.press/v119/perdomo20a.html.

Georgios Piliouras and Fang-Yi Yu. Multi-Agent Performative Prediction: From Global Stability and Optimality to Chaos. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*, pp. 1047–1074. ACM, 2023. doi: 10.1145/3580507.3597759.

Bernd Prach and Christoph H. Lampert. Almost-orthogonal layers for efficient general-purpose lipschitz networks. In *Computer Vision – ECCV 2022*, volume 13681 of *Lecture Notes in Computer Science*, pp. 350–365. Springer, 2022. doi: 10.1007/978-3-031-19803-8_21.

Roy Radner. Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of Economic Theory*, 22(2):136–154, 1980. doi: 10.1016/0022-0531(80)90037-X.

Zhongqiang Ren, Sivakumar Rathinam, and Howie Choset. Multi-objective conflict-based search for multi-agent path finding. In *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 8786–8791, 2021. doi: 10.1109/ICRA48506.2021.9560985.

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

J. B. Rosen. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3):520–534, 1965. doi: 10.2307/1911749.

Aviad Rubinstein. Inapproximability of Nash Equilibrium. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 409–418, 2015. doi: 10.1145/2746539.2746578.

Paolo Scarabaggio, Nicola Mignoni, Raffaele Carli, and Mariagrazia Dotoli. Equilibrium Seeking in Learning-Based Noncooperative Nash Games. In *IEEE 63rd Conference on Decision and Control (CDC)*, pp. 210–215, Milan, Italy, December 2024. IEEE. doi: 10.1109/CDC56724.2024.10886614.

Paolo Scarabaggio, Nicola Mignoni, Raffaele Carli, and Mariagrazia Dotoli. Equilibria Existence in Games: A Perspective from the Lefschetz–Nielsen Theory. In *European Control Conference (ECC)*, pp. 2938–2943, Reykjavík, Iceland, July 2026. EUCA. ISBN 978-3-907144-13-8.

Tatiana Tatarenko and Maryam Kamgarpour. Payoff-based Learning of Nash Equilibria in Merely Monotone Games. *IEEE Transactions on Control of Network Systems*, 11(4):1790–1799, 2024. doi: 10.1109/TCNS.2024.3355035.

Aladin Virmaux and Kevin Scaman. Lipschitz Regularity of Deep Neural Networks: Analysis and Efficient Estimation. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, 2018.

Anna von Heusinger and Christian Kanzow. Optimization reformulations of the generalized Nash equilibrium problem using Nikaido–Isoda-type functions. *Computational Optimization and Applications*, 43(3):353–377, 2009. doi: 10.1007/s10589-007-9145-6.

Fangji Wang, Han Zhang, Sven Koenig, and Jiaoyang Li. Efficient approximate search for multi-objective multi-agent path finding. In *International Conference on Automated Planning and Scheduling (ICAPS)*, volume 34, pp. 613–622, 2024. doi: 10.1609/icaps.v34i1.31524.

Ruigang Wang and Ian R. Manchester. Direct Parameterization of Lipschitz-Bounded Deep Networks. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, volume 202 of *Proceedings of Machine Learning Research*, pp. 36093–36110. PMLR, 2023.

Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945. doi: 10.2307/3001968.

Ezra Winston and J. Zico Kolter. Monotone Operator Equilibrium Networks. In *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, pp. 10718–10728, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/798d1c2813cbdf8bcdb388db0e32d496-Abstract.html.

Nobuo Yamashita and Masao Fukushima. Equivalent unconstrained minimization and global error bounds for variational inequality problems. *SIAM Journal on Control and Optimization*, 35(1): 273–284, 1997. doi: 10.1137/S0363012994277645.

Han Zhang, Oren Salzman, Ariel Felner, Carlos Hernández Ulloa, and Sven Koenig. A-a*pex: Efficient anytime approximate multi-objective search. In *Symposium on Combinatorial Search (SoCS)*, volume 17, pp. 179–187, 2024. doi: 10.1609/socs.v17i1.31556.

Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. In *Handbook of Reinforcement Learning and Control*, pp. 321–384. Springer, 2021. doi: 10.1007/978-3-030-60990-0_12.

APPENDICES

CONTENTS

A Related Work and Distinctions **16**

B Discussion and Limitations **18**

C A Catalog of (Deferred) Proofs **20**

D A Complete Finite-Query Example **22**

E Constants and Rates of the Blockwise Bound **24**

E.1 Average prediction error does not control the consistency error 24

E.2 Sharpness of the blockwise constants 24

E.3 The quadratic regime 24

F Response Models and Multiple-Equilibrium Recovery **25**

G Experimental Setup **25**

H Robustness and Scope **28**

H.1 Calibration of the central-difference estimator 28

H.2 Convexity failure 28

A Related Work and Distinctions

A Need for Certification. No uncoupled dynamic converges to a Nash equilibrium across every game in a neighborhood of a given game (Hart & Mas-Colell, 2003). The counterexample of Hart & Mas-Colell (2003) already satisfies our main structural assumptions: two players act on unit disks, each cost is smooth and strongly convex in the player's own action, and the equilibrium is unique and strict. Convergence is therefore a stronger demand than evaluating whether a given profile is near equilibrium. Our certificate addresses the latter question.

Learned response maps. Response models replace each player's optimization with a map fitted to observed actions. Scarabaggio et al. (2024) introduced the projected observer–head architecture, proved fixed-point existence through Brouwer's theorem, and linked its fixed points to a variational inequality. Their uniqueness and convergence guarantees require the aggregate contraction condition $\left(\sum_i \gamma_i^2\right)^{1/2} < 1$, with Krasnosel'skiĭ iteration at the nonexpansive boundary. Those guarantees concern the learned game. In contrast, Theorem 3.1 bounds exploitability in the game that generated the data, permits a nonzero fixed-point residual, and imposes no condition on the γ_i.

Mignoni (2025, Chapter 6) combine projected maps, existence, and aggregate contraction in the same manner. Their variational inequality restates the learned fixed-point condition and does not constrain the data-generating game. Without a learned component, Scarabaggio et al. (2026) characterize stationary profiles of the true costs as fixed points of the blockwise projected gradient and prove existence and homotopy persistence through Lefschetz–Nielsen theory. Our direct-probe candidate uses the corresponding blockwise residual r_i.

Perturbed equilibria use a different ε. In the purification of Harsanyi (1973), ε scales a payoff perturbation and the equilibrium solves the perturbed game. In quantal response (McKelvey & Palfrey, 1995) and perturbed best-response learning (Hofbauer & Hopkins, 2005), it scales smoothing and the equilibrium is a fixed point of a modified operator. Neither construction bounds exploitability in the original game. Our ε instead measures regret in the unmodified game through (2).

Decision-dependent and performative games. When deployment changes the data distribution, repeated retraining defines the relevant fixed point. Perdomo et al. (2020) introduced performative prediction, and Narang et al. (2023) extended the framework to multiple players. They prove existence, uniqueness, and linear convergence when the ratio of sensitivity to strong convexity is below one. That ratio plays the role of a contraction modulus and produces the same $1/(1-q)$ degradation discussed in Remark B.3. Outside this regime, multi-agent retraining can be chaotic (Piliouras & Yu, 2023), while Dong et al. (2023) obtain only local regions of attraction. These results motivate a certificate that remains finite at the contraction boundary.

Certified operators and Lipschitz networks. Another literature constrains the learned operator by construction. Monotone operator equilibrium networks produce well-posed implicit layers (Winston & Kolter, 2020); non-Euclidean one-sided Lipschitz conditions provide weaker alternatives to Euclidean contraction (Jafarpour et al., 2021; Davydov et al., 2022). Lipschitz bounds can be imposed by direct parameterization (Wang & Manchester, 2023; Havens et al., 2023) or bounded after training by semidefinite relaxation (Fazlyab et al., 2019); computing the smallest global bound is NP-hard (Virmaux & Scaman, 2018).

Two limitations affect certification. Products of layerwise spectral norms can be loose, and the Clarke chain rule for a composition is an inclusion rather than an equality (Clarke, 1983). Automatic differentiation can also return an element of a conservative field rather than the Clarke Jacobian at activation or projection boundaries (Bolte & Pauwels, 2021). We therefore use weight-derived upper bounds for certification and report sampled Jacobian norms only as lower-bound diagnostics.

Error bounds and gap functions. The inequality behind Lemma 3.1 belongs to the error-bound theory of finite-dimensional variational inequalities (Facchinei & Pang, 2003, Ch. 6); see also Pang (1997). Regularized gaps, D-gaps, and related merit functions convert a variational inequality into an optimization problem (Fukushima, 1992; Larsson & Patriksson, 1994; Peng, 1997; von Heusinger & Kanzow, 2009). Most associated error bounds control distance to the solution set and require strong monotonicity (Yamashita & Fukushima, 1997). Our bound controls exploitability instead, so that compactness and playerwise convexity suffice. Classical sensitivity results similarly rely on strong regularity or strong monotonicity of the environment (Robinson, 1980; Dontchev & Rockafellar, 1996; Dafermos, 1988).

A profile certificate differs from a convergence guarantee. Mertikopoulos & Zhou (2018) study compact convex product action sets with costs that are C^1 in each player's own action. Their global guarantees concern limits of dual averaging and require pseudo-concavity, variational stability, monotonicity, or a concave potential. Their $\mathcal{O}(1/\sqrt{n})$ estimate applies to an ergodic average that may differ from the played sequence. Variational stability also depends on an unknown equilibrium and therefore does not yield a number that can be evaluated at a candidate profile. For C^2 costs it is equivalent to local asymptotic stability (Mazumdar et al., 2020), excluding unstable equilibria that may still require external evaluation.

Continuous-time equilibrium-seeking methods likewise drive a trajectory toward an equilibrium rather than certify an arbitrary profile. Existing consensus, passivity, gradient-free, and operator-splitting methods typically assume monotonicity and use Lyapunov functions involving the unknown equilibrium (Hu et al., 2022; Belgioioso et al., 2022). The quantitative result of Belgioioso et al. (2022) is an ergodic $\mathcal{O}(1/k)$ rate for a fixed-point residual. Their operator-to-game link applies when that residual vanishes; it does not convert a positive residual into an exploitability bound.

Observed play does not identify equilibrium quality. Gürkan & Pang (2009) give asymptotic conditions under which equilibria of approximating games converge to equilibria of the true game, but provide no finite constant. Bemporad & Tatarenko (2026a) train parametric equilibrium maps with a Nikaido–Isoda objective, while Bemporad & Tatarenko (2026b) impose monotonicity on the learned game. Bergerault et al. (2026) show that measure-matching error does not determine the Nash gap. Their positive results require additional structure on the demonstrated expert, such as dominant-strategy behavior or continuity of best responses near the expert equilibrium.

Our assumptions constrain different objects. Playerwise convexity constrains the true one-shot game, while the regional envelope bounds the discrepancy between the learned operator and the true pseudo-gradient. Neither assumption set implies the other. This is why the consistency error in (9) measures operator agreement rather than distributional fit, and why Section 4 states environment access as a cost. Payoff-based equilibrium learning supplies the same access model: players evaluate costs but do not differentiate them (Tatarenko & Kamgarpour, 2024).

Generative responders. Conditional diffusion and flow models can represent multimodal response distributions rather than only conditional means. Proposition 3.2 shows why this distinction matters when playerwise convexity fails. Under playerwise convexity, each best-response set is convex; multimodality must then arise from multiple equilibria rather than a multimodal best response. Our experiments evaluate the generative families through their conditional means and therefore do not establish a benefit from sampling or generativity.

A certificate for a genuinely set-valued response remains open because the residual in (9) requires a single-valued evaluation. The same limitation applies to a generalist policy deployed as several players when those playerwise responses share parameters.

B Discussion and Limitations

Regional auditing requires supplied regularity and a manageable cover. The finite-cover construction in Proposition 4.1 assumes valid regularity and noise bounds; cost queries do not infer them. Without those inputs, Theorem 3.1 remains valid for an envelope established by another method, but Proposition 4.1 does not construct one. The audit also requires a finite cover. For an intrinsic-dimension-k box, the cover size grows as $O((\operatorname{diam}\Omega/\eta)^k)$, limiting the construction to regions with small covering numbers.

The regional-audit experiment illustrates this tradeoff. At $\eta/\operatorname{diam}\Omega = 0.025$, the median cover contains 20 points on the segment and 8,000 points on the three-dimensional box. A gradient-block oracle makes a pointwise audit cost N queries and a regional audit cost $N|\mathcal{Z}|$ queries. Under the scalar cost-oracle model of Section 4, central differences instead cost $2d$ evaluations per point and $2d|\mathcal{Z}|$ per cover. Both accounting rules break even after $|\mathcal{Z}|$ later profiles. This comparison excludes cover construction, storage, training, model evaluation, and rollout replication. It also assumes the required regularity constants are already available.

Regional auditing buys reuse rather than a tighter certificate. Retaining the cover gradients gives the profile-dependent model-free upper bound

$$D_i \min_{z\in\mathcal{Z}}\{M_i(z) + L_{M,i}\|x - z\|\}. \tag{24}$$

Under the Lipschitz construction used in our regional experiment, this bound dominates the learned candidate at every profile. The experiment therefore does not establish a computational or tightness advantage for the learned map. This ordering is not universal: the worked Cournot example in Appendix D uses a sharp modulus for δ_i that captures cancellation between G_i and Φ_i, and its learned candidate is smaller than the retained-cover model-free bound. Replacing the cover with random samples removes the uniform guarantee unless a separate distributional result is proved.

Response accuracy does not imply operator alignment. The discrepancy δ_i compares the displacement $x_i - f_i(x_{-i})$ with the true gradient; it is not prediction error relative to a best response. For example, consider

$$J_i(y_i, x_{-i}) = \frac{a_i}{2}\|y_i - b_i(x_{-i})\|^2, \qquad a_i > 0. \tag{25}$$

A zero-error response predictor $f_i = b_i$ has

$$\delta_i(x) = |1 - a_i|\,\|x_i - b_i(x_{-i})\|. \tag{26}$$

Thus curvature and cost normalization affect the certificate even when response prediction is perfect. Constraints introduce another mismatch: at a boundary best response, $0 \in \Phi_i(b_i, x_{-i}) + N_{X_i}(b_i)$, so that the raw gradient need not vanish even though the response displacement does. The audit measures this alignment directly; it does not derive it from training loss.

Coupled feasibility sets are outside the theory. Assum. 2.1 requires a product action set. This requirement makes projection blockwise and allows one player's action to vary while the others remain fixed. Games with shared feasibility constraints therefore fall outside every theorem in the paper. The shared-constraint experiment is a stress diagnostic, not evidence of theorem coverage.

Playerwise convexity is necessary. The witness in Proposition 3.2 has zero variational residual and positive exploitability once Assum. 2.2 fails. Under playerwise convexity, each best-response set is convex, excluding the witness's bimodal response. Multimodality can still arise from multiple equilibria of a non-monotone game.

Our experiments evaluate diffusion and flow models through their conditional means because the residual ρ_i in (9) requires a single-valued map. The comparison therefore does not evaluate sampling quality or isolate generativity from architecture. A linear responder fitted across attraction basins lands between equilibria on a fraction 0.724 of instances, while the nonlinear responders avoid that failure more often. A certificate for a genuinely set-valued response remains open. Fixed-point existence for correspondences with nonempty closed convex values is classical (Kakutani, 1941; Facchinei & Kanzow, 2010); the missing object is a computable residual that measures approximate fixed-point agreement.

The quadratic rate is not a computable certificate. Proposition E.2 gives a quadratic rate when best responses lie in the relative interior of their action sets. Evaluating its discrepancy term requires

distances to an $\arg\min$ and hence N constrained solves in the unknown costs. We therefore report it as a rate and measure it only when ground truth is available. A computable version would also require probe-based evidence of interiority and a curvature modulus.

Sampled Jacobians are diagnostics. Weight-derived bounds provide the rigorous operator-norm quantities used for certification. Sampled Jacobians are lower-bound diagnostics and do not enter a certificate. The distinction matters because the Clarke chain rule for a composition is an inclusion (Clarke, 1983), and automatic differentiation may return an element of a conservative field rather than the Clarke Jacobian at activation or projection boundaries (Bolte & Pauwels, 2021).

Remark B.1 (Why certify rather than solve). The oracle that constructs an envelope could also drive projected-gradient play. Certification still serves three different purposes. (i) One regional envelope certifies every later profile in Ω without another game query, whereas solving the game pays along each trajectory and returns one profile. This reuse also applies to the regional model-free candidate. (ii) Projected-gradient play has no convergence guarantee under our assumptions. Even monotonicity can leave a nontrivial invariant set (Hu et al., 2022), and multi-player retraining outside the contraction regime can be chaotic (Piliouras & Yu, 2023). (iii) A learned response map may be needed for counterfactual simulation or prediction regardless of whether it is used to compute an equilibrium. The certificate evaluates that existing model.

Remark B.2 (What the certificate does not claim). The bound in (12) concerns profiles in Ω. It does not imply that $\mathrm{Fix}(F)$ is a singleton or that a certified profile is close to a particular equilibrium. For a nonexpansive map, the fixed-point set is nonempty, closed, and convex but may contain many points, as $F = \mathrm{Id}$ shows. We therefore refer to *a* behavioral equilibrium.

Remark B.3 (Behavior at the contraction boundary). For a contraction with modulus $q < 1$, a uniform map error ς yields the classical displacement factor $\varsigma/(1-q)$ (Bauschke & Combettes, 2017). The factor diverges as $q \to 1$. The certificate in (12) contains no q and therefore remains finite for nonexpansive or expansive learned maps. This separates it from the aggregate-contraction condition of Scarabaggio et al. (2024) and the sensitivity-to-curvature condition of Narang et al. (2023).

Remark B.4 (Regularity of the learned map does not certify the game). Let $F \equiv \hat{x}$ for an arbitrary $\hat{x} \in X$. The map is nonexpansive and satisfies $\rho_i(\hat{x}) = 0$, while $\varepsilon(\hat{x})$ can be as large as the game permits. Its induced operator has $G_i(\hat{x}) = 0$, so that $\delta_i(\hat{x}) = M_i(\hat{x})$. The learned-alignment candidate is then $D_i \bar{M}_i$ on the singleton region, equal to the model-free candidate. Regularity of the learned map alone contains no information about the true game.

Remark B.5 (Fixed-point computation near the contraction boundary). A contraction F with modulus $q < 1$ gives geometric convergence of Picard iteration. Reducing the fixed-point residual below ϵ requires

$$\mathcal{O}\left(\frac{\log(1/\epsilon)}{1-q}\right) \tag{27}$$

value queries when q is near one. If $q = 1 - \varepsilon$ with $0 < \varepsilon \ll 1$, the cost is $\mathcal{O}(\varepsilon^{-1}\log(1/\epsilon))$. This contraction argument disappears at $q = 1$. General value-oracle lower bounds for Brouwer fixed-point computation are exponential in dimension and requested accuracy under the corresponding Lipschitz assumptions (Hirsch et al., 1989); related reductions give query lower bounds for equilibrium computation (Babichenko, 2016). These results concern finding a fixed point. Our result instead certifies a supplied profile without assuming contraction.

Remark B.6 (Sampling changes the guarantee). Replacing the cover with independent draws from a distribution ν on Ω removes the covering cost and the deterministic uniform guarantee. Under exchangeability, an order-statistic or tolerance result could provide a fresh-draw or quantile guarantee, but not a bound on the regional supremum. We prove no such distributional result and treat sample maxima only as empirical diagnostics.

The boundary-stencil count is fixed by the implementation. The reported fraction of coordinates using the one-sided stencil of Remark 4.1 is zero by construction: the implementation always uses an unconstrained central difference. The measurable companion quantity is the fraction of probe steps that leave the action set, which is zero at every tested radius. This result describes the tested arenas and does not validate the one-sided remedy.

C A CATALOG OF (DEFERRED) PROOFS

This appendix catalogs the proofs of the results of Section 3 and Section 4, in order.

Proof of Lemma 3.1. Fix a player i and let $y_i^\star \in \arg\min_{y_i \in X_i} J_i(y_i, x_{-i})$, which exists because X_i is compact by Assum. 2.1 and $J_i(\cdot, x_{-i})$ is continuous by Assum. 2.2. Convexity of $J_i(\cdot, x_{-i})$ implies that $J_i(y_i^\star, x_{-i}) \geq J_i(x) + \langle \Phi_i(x), y_i^\star - x_i \rangle$. Hence, player i's regret obeys

$$J_i(x) - J_i(y_i^\star, x_{-i}) \;\leq\; \big\langle \Phi_i(x),\, x_i - y_i^\star \big\rangle. \tag{28}$$

Set $p_i = \Pi_{X_i}\big(x_i - \Phi_i(x)\big)$ and $w_i = x_i - p_i$, so that $\|w_i\| = r_i(x)$ by (7). Projecting $x_i - \Phi_i(x)$ onto X_i implies $\langle x_i - \Phi_i(x) - p_i,\, y_i - p_i \rangle \leq 0$ for every $y_i \in X_i$, the variational characterization of the projection (Facchinei & Pang, 2003, §1.5); equivalently, p_i solves $\mathrm{VI}\big(X_i,\, u \mapsto u - (x_i - \Phi_i(x))\big)$ in the sense of Definition 2.1. It is easy to verify that $\langle \Phi_i(x),\, p_i - y_i \rangle \leq \langle w_i,\, p_i - y_i \rangle$; applying the Cauchy–Schwarz inequality with $\|p_i - y_i\| \leq D_i := \operatorname{diam} X_i$, since $p_i, y_i \in X_i$, gives

$$\langle \Phi_i(x),\, p_i - y_i \rangle \;\leq\; r_i(x)\, D_i \qquad \text{for every } y_i \in X_i. \tag{29}$$

Splitting (28) at p_i as $\langle \Phi_i(x), x_i - y_i^\star \rangle = \langle \Phi_i(x), w_i \rangle + \langle \Phi_i(x), p_i - y_i^\star \rangle$, then bounding the first term by $M_i(x) r_i(x)$ with the Cauchy–Schwarz inequality and the second by (29), we find that the bound on the player i's regret is $\big(M_i(x) + D_i\big) r_i(x)$, which is the first entry of (8). Separately, (28) and Cauchy–Schwarz give $J_i(x) - J_i(y_i^\star, x_{-i}) \leq M_i(x) \|x_i - y_i^\star\| \leq M_i(x) D_i$. Both bounds apply to the same player, hence player i's regret is at most their minimum, and (8) follows by taking the maximum over players. Writing $a_i = \big(M_i(x) + D_i\big) r_i(x)$ and $b_i = M_i(x) D_i$, the inequality $\max_i \min\{a_i, b_i\} \leq \min\{\max_i a_i, \max_i b_i\}$ shows that taking the minimum after the maximum is valid but never tighter, and strict when no single player has both $a_i \geq m$ and $b_i \geq m$ for $m = \min\{\max_j a_j, \max_j b_j\}$. □

Proof of Lemma 3.2. Since $x_i - G_i(x) = f_i(x_{-i}) \in X_i$, projection fixes this point. Inserting it into $r_i(x) = \|x_i - \Pi_{X_i}(x_i - \Phi_i(x))\|$ and using the triangle inequality gives $r_i(x) \leq \rho_i(x) + \|\Pi_{X_i}(x_i - G_i(x)) - \Pi_{X_i}(x_i - \Phi_i(x))\|$. Nonexpansiveness of Π_{X_i} bounds the second term by $\|G_i(x) - \Phi_i(x)\| = \delta_i(x)$. The triangle inequality also gives $M_i(x) = \|\Phi_i(x)\| \leq \|G_i(x)\| + \|G_i(x) - \Phi_i(x)\| = \rho_i(x) + \delta_i(x)$, proving both inequalities. □

Proof of Proposition 3.1. The reverse triangle inequality gives $M_i(\hat{x}) \leq \rho_i(\hat{x}) + \delta_i(\hat{x})$. Therefore the minimum in (14) selects $M_i(\hat{x})$, including when either $M_i(\hat{x})$ or D_i is zero. □

Proof of Theorem 3.1. Fix $\hat{x} \in \Omega$. By (11), $M_i(\hat{x}) \leq \bar{M}_i$ and $\delta_i(\hat{x}) \leq \bar{\delta}_i$. The operator-transfer inequality (10) therefore gives

$$M_i(\hat{x}) \leq \min\big\{\bar{M}_i,\; \rho_i(\hat{x}) + \bar{\delta}_i\big\}. \tag{30}$$

The model-free part of Lemma 3.1 bounds player i's regret by $M_i(\hat{x}) D_i$. Substitution gives the playerwise minimum in (12); taking the maximum over players proves the claim. The envelope is fixed before $\hat{x}$ is selected, so that the conclusion holds simultaneously throughout Ω. □

Proof of Corollary 3.1. When $D_i > 0$, dividing $D_i\big(\rho_i(\hat{x}) + \bar{\delta}_i\big) < D_i \bar{M}_i$ by D_i gives (15). If $D_i = 0$, both candidates equal zero. □

Proof of Proposition 3.2. Differentiating (16) gives $\Phi_i = x_i\big(x_i^2 - w_i(x_{-i})\big)$, which vanishes at $x = 0$ for every x_{-i}. Since $0 \in X_i = [-c, c]$, the projection fixes it, so that $r_i(0) = \|\Pi_{X_i}(0)\| = 0$ for every player. Each $J_i(\cdot, x_{-i})$ is even in x_i with minimizers $\pm\sqrt{w_i(x_{-i})}$ and minimum value $-\frac{1}{4} w_i(x_{-i})^2$, while $J_i(0, x_{-i}) = 0$. At the origin $w_i(0) = 1$, hence every player's regret is $\frac{1}{4}$. The condition on c places $\pm\sqrt{w_i}$ inside X_i for every $x_{-i} \in X_{-i}$. □

Proof of Proposition 4.1. Fix a player i, a covering point $z \in \mathcal{Z}$ and a coordinate k of player i's block, and write $\varphi(t) = J_i(z + t e_{i,k})$, so that $\varphi'(0) = [\Phi_i(z)]_k$ is the quantity (17) estimates. The assumed extension makes φ three times continuously differentiable on $[-\tau, \tau]$, hence Taylor's

theorem with Lagrange remainder gives $\varphi(\pm\tau) = \varphi(0) \pm \tau\varphi'(0) + \frac{1}{2}\tau^2\varphi''(0) \pm \frac{1}{6}\tau^3\varphi'''(\xi_\pm)$ for some $|\xi_\pm| \le \tau$. Subtracting the two expansions cancels the even orders, so that

$$\frac{\varphi(\tau) - \varphi(-\tau)}{2\tau} \; - \; \varphi'(0) \; = \; \frac{\tau^2}{12}\big(\varphi'''(\xi_+) + \varphi'''(\xi_-)\big). \tag{31}$$

Since $\varphi''' = \partial^3 J_i / \partial x_{i,k}^3$ is bounded by B_3 on the probe segment, the right side of (31) is at most $\frac{1}{6}B_3\tau^2$ in magnitude. Stacking the n_i coordinates, player i's bias vector has norm at most $\frac{1}{6}B_3\tau^2\sqrt{n_i}$.

For the noise, a difference of two independent σ-sub-Gaussian variables is $\sqrt{2}\sigma$-sub-Gaussian, so that dividing by 2τ makes each difference quotient s-sub-Gaussian about its mean with $s = \sigma/(\sqrt{2}\tau)$. There are $d|\mathcal{Z}|$ scalar estimates in all, one per coordinate and covering point, and each has two-sided tail $2\exp(-t^2/2s^2)$. Equating $2d|\mathcal{Z}|\exp(-t^2/2s^2)$ with α and solving for t, a union bound gives that every coordinatewise deviation is at most $s\sqrt{2\log(2d|\mathcal{Z}|/\alpha)} = \frac{\sigma}{\tau}\sqrt{\log(2d|\mathcal{Z}|/\alpha)}$ simultaneously, with probability at least $1-\alpha$; the $\sqrt{2}$ in s cancels the one from the tail. On that event player i's noise vector has norm at most $\frac{\sigma}{\tau}\sqrt{n_i\log(2d|\mathcal{Z}|/\alpha)}$.

Hence, on the same event and for every $z \in \mathcal{Z}$ at once, the triangle inequality bounds $\delta_i(z) = \|G_i(z) - \Phi_i(z)\|$ by $\|G_i(z) - \widehat{\Phi}_{\tau,i}(z)\| + E_i(\tau,\alpha)$ and bounds $M_i(z)$ by $\|\widehat{\Phi}_{\tau,i}(z)\| + E_i(\tau,\alpha)$. Finally, given $x \in \Omega$, choose $z \in \mathcal{Z}$ with $\|x - z\| \le \eta$. Lipschitzness gives $\delta_i(x) \le \delta_i(z) + L_{\delta,i}\eta$ and $M_i(x) \le M_i(z) + L_{M,i}\eta$. Taking the two maxima over the covering yields (19). The coordinatewise union bound already ranges over all d coordinates and all covering points, so that both envelope inequalities hold for every player on this one event. □

D A COMPLETE FINITE-QUERY EXAMPLE

The certificate and its finite-query construction impose assumptions on the game, learned response, oracle, and audited region. We use one constructed instance to show that these assumptions can hold simultaneously and to trace every quantity from the cost queries to the final bound. The example is deliberately low-dimensional: it establishes satisfiability and non-vacuity only.

Consider a two-firm Cournot game in which firm i chooses output $q_i \in X_i = [0, 0.6]$, and let $Q = q_1 + q_2$ denote total output. Start from the standard linear inverse demand $P(Q) = a - bQ$, where a is the price intercept and $b > 0$ is the rate at which price falls with total output. Choose $b = 0.5$, unit marginal cost $c = 1$, and net intercept $a - c = 0.6$, which gives

$$P(Q) = 1.6 - 0.5Q. \tag{32}$$

Only $a - c$ affects the game, so that the unit marginal cost is a normalization. Firm i earns profit $\pi_i(q_i, q_j) = q_i\big(P(q_i + q_j) - c\big)$. Under our cost-minimization convention, define $J_i = -\pi_i$ i.e.

$$J_i(q_i, q_j) = -q_i\big(P(q_i + q_j) - c\big) = \frac{1}{2}q_i^2 + \frac{1}{2}q_iq_j - 0.6q_i. \tag{33}$$

The cost has $\partial^2 J_i/\partial q_i^2 = 1$, so that its first-order condition gives the unique unconstrained minimizer:

$$0 = \frac{\partial J_i}{\partial q_i} = q_i + 0.5q_j - 0.6 \quad \Longrightarrow \quad \mathrm{BR}_i(q_j) = 0.6 - 0.5q_j. \tag{34}$$

This minimizer lies in $[0.3, 0.6]$ for every $q_j \in [0, 0.6]$; thus, it is also the constrained best response on $X_i = [0, 0.6]$. To represent a learned responder with a small slope error, take

$$f_i(q_j) = 0.6 - 0.49q_j. \tag{35}$$

This map is feasible, and its slope error reveals the discrepancy from the true gradient i.e. , $|G_i(q) - \Phi_i(q)| = 0.01q_j$.

Suppose that we audit the region $\Omega = [0.3, 0.5]^2$ using the following

$$\mathcal{Z} = \{0.35, 0.45\}^2, \qquad \eta = \frac{\sqrt{2}}{20}, \qquad \tau = 0.01, \qquad \alpha = 0.05. \tag{36}$$

Use a noiseless cost oracle, so that the sub-Gaussian scale σ of the query error is zero. The Cartesian product $\mathcal{Z}$ contains $(0.35, 0.35)$, $(0.35, 0.45)$, $(0.45, 0.35)$, and $(0.45, 0.45)$. These points center the four 0.1×0.1 squares obtained by splitting each coordinate interval of Ω at 0.4. Every point in a square lies within its half-diagonal $\sqrt{0.05^2 + 0.05^2} = \sqrt{2}/20 =: \eta$ of the center, so that $\mathcal{Z}$ is a Euclidean η-cover. Since the costs are quadratic on $\mathbb{R}^2$, every probe segment is valid, $B_3 = 0$, and central differences recover the gradients without bias.

Table 2: Cournot Game Example: Verification of Assumptions.

Requirement	Verification
Product geometry	$X = [0, 0.6]^2$ is compact and convex; $D_1 = D_2 = 0.6$.
Playerwise convexity	$\partial^2 J_i/\partial q_i^2 = 1$.
Feasible learned response	$f_i([0, 0.6]) = [0.306, 0.6] \subseteq X_i$.
Gradient modulus	$M_i(q) = \lvert q_i + 0.5q_j - 0.6\rvert$ is $L_{M,i} = \sqrt{5}/2$-Lipschitz.
Alignment modulus	$\delta_i(q) = \lvert G_i(q) - \Phi_i(q)\rvert = 0.01q_j$ on Ω, so that $L_{\delta,i} = 0.01$.
Probe smoothness	The global quadratic extension gives $B_3 = 0$; $\tau = 0.01$ remains feasible.
Noise model	Zero noise is independent, mean-zero, and 0-sub-Gaussian.
Finite cover	$\lvert\mathcal{Z}\rvert = 4$ and Proposition 4.1 uses $2d\lvert\mathcal{Z}\rvert = 16$ scalar cost queries.

On the cover, $\max_z M_i(z) = 0.075$ and $\max_z \delta_i(z) = 0.0045$. Proposition 4.1 therefore gives

$$\bar{M}_i = 0.075 + \frac{\sqrt{10}}{40} = 0.15406, \qquad \bar{\delta}_i = 0.0045 + \frac{\sqrt{2}}{2000} = 0.005207. \tag{37}$$

At $q = (0.5, 0.3)$, the learned residuals are $(\rho_1, \rho_2) = (0.047, 0.055)$. Both best responses are interior, and each cost has unit own-action curvature, so that direct minimization gives

$$\varepsilon(q) = \frac{1}{2}(0.05)^2 = 0.00125. \tag{38}$$

The certificate in (12) is

$$\max_i D_i \min\{\bar{M}_i, \rho_i + \bar{\delta}_i\} = 0.03612. \tag{39}$$

For comparison, retaining the cover gradients gives

$$\max_i D_i \min_{z \in \mathcal{Z}}\{M_i(z) + L_{M,i}\|q - z\|\} = 0.06243, \tag{40}$$

while the constant regional model-free value is $\max_i D_i \bar{M}_i = 0.09243$. Thus

$$0.00125 = \varepsilon(q) \leq 0.03612 < 0.06243 < 0.09243. \tag{41}$$

The learned candidate improves here because the sharp modulus $L_{\delta,i} = 0.01$ captures cancellation between G_i and Φ_i. Replacing it with the generic triangle-inequality modulus $L_{G,i} + L_{M,i}$ remains valid but removes this advantage. The construction demonstrates simultaneous satisfiability and non-vacuity; its four-point cover does not address high-dimensional scalability.

E CONSTANTS AND RATES OF THE BLOCKWISE BOUND

This appendix establishes three limits of Lemma 3.1 and Theorem 3.1. Average prediction error does not control uniform model–game consistency, neither coefficient in the blockwise bound can be reduced uniformly, and the rate becomes quadratic only under an interior best-response condition.

E.1 AVERAGE PREDICTION ERROR DOES NOT CONTROL THE CONSISTENCY ERROR

A small training loss does not control the operator-consistency error in Lemma 3.2. The following one-player construction isolates the gap between an average error and a uniform bound; no multi-player coupling is needed.

Proposition E.1 (Small mean-squared error, displaced fixed point)**.** *Let $X = [0,1]$, let the reference response be $F_0 \equiv 0$, and for $0 < h < \frac{1}{4}$ let*

$$T_h(x) = \tfrac{1}{2} \max\left\{1 - |x - \tfrac{1}{2}|/h,\ 0\right\}. \tag{42}$$

Then under observations uniform on X, $\mathbb{E}\left|T_h(x) - F_0(x)\right|^2 = h/6 \to 0$ as $h \to 0$, while $\sup_{x\in X}\left|T_h(x) - F_0(x)\right| = \frac{1}{2}$ for every h. The map T_h retains the reference fixed point at 0 and acquires a second fixed point at $\frac{1}{2}$, whose fixed-point residual is zero and whose distance from the reference fixed point is $\frac{1}{2}$.

Proof. The triangle T_h is supported on $[\frac{1}{2} - h, \frac{1}{2} + h]$ with peak $\frac{1}{2}$ at $x = \frac{1}{2}$, so that its supremum is $\frac{1}{2}$ and $\int_0^1 T_h^2 = h/6$ by direct integration of the squared triangular profile. Evaluating (42) gives $T_h(0) = 0$ and $T_h(\frac{1}{2}) = \frac{1}{2}$, hence both points are fixed. The residual at $\frac{1}{2}$ vanishes because $\frac{1}{2}$ is a fixed point. □

To read this as a statement about the consistency error of Lemma 3.2, use a two-player decoupled game on $X_1 = X_2 = [0,1]$ with $J_i(x) = \frac{1}{2}x_i^2$, hence $\Phi_i(x) = x_i$. Define the admissible response map $F(x_1, x_2) = (T_h(x_2), T_h(x_1))$. Then $G_i(x) = x_i - T_h(x_{-i})$ and $\delta_i(x) = T_h(x_{-i})$, so that $\sup_X \delta_i = \frac{1}{2}$ while its squared expectation under the uniform distribution on X is $h/6$. Both $(0,0)$ and $(\frac{1}{2}, \frac{1}{2})$ are fixed points. On the full region X, $\bar{M}_i = D_i = 1$ and $\bar{\delta}_i = \frac{1}{2}$, so that at the displaced fixed point the learned-alignment candidate is $\frac{1}{2}$, compared with the model-free candidate 1. The bound still holds because each player's regret there is $\frac{1}{8}$. Thus vanishing average prediction error remains compatible with a displaced fixed point and a nonzero uniform alignment error.

The construction has three consequences. First, an average prediction error of $h/6$ coexists with a uniform map error of $\frac{1}{2}$, so that average loss cannot control the regional discrepancy without a uniform regularity condition. Second, the displaced fixed point has zero residual, echoing Remark B.4. Third, no contraction result applies because T_h has slope magnitude $1/(2h) > 1$ on its support. The regional envelope in Proposition 4.1 supplies the missing uniform control by probing a finite cover rather than relying on training loss.

E.2 SHARPNESS OF THE BLOCKWISE CONSTANTS

Neither coefficient in (8) can be reduced uniformly. The construction below satisfies $\varepsilon = MD$ for every parameter value, so that the model-free candidate always binds. It establishes sharpness of the coefficients only in that branch of the certificate.

Take $N = 1$, $X = [0,1]$, and $J(x) = ax$, and evaluate at $x = 1$. Then $M = a$, $D = 1$, and $r = \min\{1, a\}$. For $a \le 1$, the attainment ratio $\varepsilon/[(M + D)r] = 1/(1 + a)$ tends to one as $a \to 0$, ruling out replacement of D by cD for $c < 1$. For $a \ge 1$, the ratio $a/(1 + a)$ tends to one as $a \to \infty$, ruling out replacement of M by cM. This argument rules out uniform rescaling of either coefficient; it does not rule out a bound with a different form.

E.3 THE QUADRATIC REGIME

Whether the rate in (12) is linear or quadratic turns on whether best responses lie in the relative interior of the action sets. Write $\mathrm{BR}_i(x_{-i}) = \arg\min_{y_i \in X_i} J_i(y_i, x_{-i})$.

Proposition E.2 (Quadratic certificate under interior best responses)**.** *Let Assum. 2.1 hold, let $J_i(\cdot, x_{-i})$ be differentiable with L_i-Lipschitz gradient $\Phi_i(\cdot, x_{-i})$ on X_i, and let some minimizer $b_i \in \mathrm{BR}_i(\hat{x}_{-i})$ lie in the relative interior of X_i for every player i. With $\rho_i = \rho_i(\hat{x})$ and $\delta_i^{\mathrm{BR}} \triangleq \|g_i(h_i(\hat{x}_{-i})) - b_i\|$,*

$$\varepsilon(\hat{x}) \;\le\; \max_{i \in \mathcal{N}} \tfrac{L_i}{2} \left(\rho_i + \delta_i^{\mathrm{BR}}\right)^2. \tag{43}$$

Proof. Fix a player i. Since $b_i \in X_i$ we have $\Pi_{X_i} b_i = b_i$, hence nonexpansiveness of Π_{X_i} gives $\|\hat{x}_i - b_i\| \le \rho_i + \delta_i^{\mathrm{BR}}$ as in Lemma 3.2. Minimality of b_i gives $\langle \Phi_i(b_i, \hat{x}_{-i}), y_i - b_i \rangle \ge 0$ for every $y_i \in X_i$. Because b_i lies in the relative interior of X_i, for each such y_i the point $b_i - t(y_i - b_i)$ also lies in X_i for all small $t > 0$, which supplies the reverse inequality; hence $\langle \Phi_i(b_i, \hat{x}_{-i}), y_i - b_i \rangle = 0$ for every $y_i \in X_i$, including at $y_i = \hat{x}_i$. The descent lemma along the segment $[b_i, \hat{x}_i] \subseteq X_i$ then bounds player i's regret by $\frac{L_i}{2}\|\hat{x}_i - b_i\|^2$. Taking the maximum over players gives (43). □

Relative interiority is the required geometric condition. At a relative-interior point of a lower-dimensional X_i, the gradient need not vanish, but its inner product with every feasible direction does. Playerwise convexity is not needed, so that (43) also applies outside Assum. 2.2.

The refinement requires more information than the regional envelope. The quantity δ_i^{BR} is a distance to an $\arg\min$, so that evaluating it requires N constrained optimization problems in the unknown costs. We therefore report (43) as the rate of the interior regime, not as a computable certificate, and measure it only when ground truth is available.

F Response Models and Multiple-Equilibrium Recovery

Multiple equilibria and what recovers them. Non-monotone coupling produces multiple equilibria even under Assum. 2.2. Active-set enumeration finds at least two Nash equilibria in 113 of 120 playerwise-convex instances, with a worst residual of 1.42×10^{-14}. A linear regressor trained on play from both attraction basins often predicts an action between the equilibria and recovers both in 43.1% of instances. The nonlinear deterministic network recovers both in 86.2%, compared with 22.4% for flow matching and 5.17% for diffusion.

The experiment evaluates every family through a single-valued map. For diffusion and flow matching, that map is the conditional mean rather than a sample. The result therefore compares fitted mean maps and does not measure sampling quality. It also does not isolate generativity from architecture: the generative heads have roughly four times as many parameters per player as the deterministic network. The supported conclusion is narrower. A linear map can average across attraction basins, while the nonlinear deterministic map avoids that failure more often in this experiment.

Set-valuedness where convexity fails. The nonconvex witness of Proposition 3.2 has a conditional-mean profile with zero variational residual and exploitability 0.25. Mode-selected samples reduce median exploitability from 0.256 to 0.0286 for diffusion and from 0.252 to 0.0175 for flow matching over 20 paired seeds ($p_{\mathrm{Holm}} = 1.91 \times 10^{-6}$ for both comparisons). Because playerwise convexity fails, this experiment diagnoses the scope of the theorem and does not test the certificate's validity.

G Experimental Setup

Each game isolates one hypothesis or limitation of Theorem 3.1. Analytic solutions or documented numerical procedures provide the reference values, allowing errors in the certificate to be distinguished from errors in an inner solver.

Coupled quadratic: monotonicity on a knob. Set $Q_i = I$ and $A = I + \kappa(W + \frac{1}{2}Z)$ with W symmetric and Z antisymmetric, both with zero diagonal blocks, W rescaled so that $\lambda_{\min}(W) = -1$ and Z so that $\|Z\|_2 = 1$. Then

$$\tfrac{1}{2}\left(A + A^\top\right) = I + \kappa W, \qquad \text{hence} \qquad \lambda_{\min}\left(\tfrac{1}{2}(A + A^\top)\right) = 1 - \kappa, \tag{44}$$

so the sweep crosses the monotonicity boundary at $\kappa = 1$. The skew component keeps $A \neq A^\top$, making the pseudo-gradient non-potential. Varying κ moves the same game family across the monotonicity boundary while controlling the modulus, so that the sweep tests the claim that the certificate does not require monotonicity.

Capacitated Cournot: an active projection. Firms choose scalar quantities on $[0, \mathrm{cap}_i]$ with capacities alternating between 0.6 and 5.0, so that the tight capacities bind at equilibrium. This tests the projection-dependent part of the blockwise gap bound. A fresh true-gradient query can make the direct projected-gradient candidate $(M_i + D_i)r_i$ sharper than the simpler model-free candidate $M_i D_i$. Proposition 3.1 still rules out an improvement by the learned-alignment candidate $D_i(\rho_i + \delta_i)$ when the model–game discrepancy is evaluated at that same profile.

Rosen shared constraint: outside the product assumption. The game intersects box constraints with one coupled row $w^\top x \leq h$, so that player i optimizes over $\mathcal{X}_i(x_{-i}) = \{x_i \in [\mathrm{lo}_i, \mathrm{hi}_i] : w_i^\top x_i \leq h - \sum_{j \neq i} w_j^\top x_j\}$. This violates the product-set assumption in Assum. 2.1. We therefore report the game only as a stress diagnostic. A globally convergent bisection on the single multiplier computes the best response. The reported D_i is conservative because it is the box diameter rather than the diameter of the coupled feasible set.

Bimodal coordination: convexity removed. With J_i and w_i as in (16), $\partial^2 J_i / \partial x_i^2(0) = -w_i < 0$, so that playerwise convexity fails. All 2^N sign patterns of $s = (1 - \kappa)^{-1/2}$ are equilibria in closed form. At the origin, the variational residual vanishes while exploitability equals $\frac{1}{4}$. This is the witness in Proposition 3.2; it tests the necessity of Assum. 2.2, not the validity of the certificate.

Nonexpansive boundary: monotone yet not contractive. Let $J_i(x) = \frac{1}{2}(x_i - \sigma_i x_{i+1})^2$ cyclically with $\sigma_i = +1$ for $i < N - 1$ and $\sigma_{N-1} = -1$. Then $F(x) = Px$ with P a signed cyclic permutation, so that $P^\top P = I$ and the best-response modulus is one, making the map nonexpansive but not contractive. The equilibrium is unique and interior at the origin because $P^N = (\prod_i \sigma_i) I = -I$ makes $I - P$ invertible. Picard iteration is $2N$-periodic from every $x^0 \neq 0$, and its residual remains constant. The game is strongly monotone with a unique equilibrium, yet naive fixed-point iteration fails. It therefore separates equilibrium quality from contraction-based convergence and provides the solver-comparison test case.

Ground truth. The product-set games have closed-form best responses. The Rosen shared-constraint diagnostic computes each best response by globally convergent bisection on its single multiplier. Exploitability is therefore evaluated against a global best response rather than inferred from the learned model or a local optimizer. Reference equilibria have residual below 10^{-10} when an equilibrium is required. Active-set enumeration yields the capacitated Cournot equilibrium, projected extragradient computes the shared-constraint variational equilibrium to its stated tolerance, and the bimodal game has 2^N closed-form equilibria.

Evaluation protocol. The experiments address three groups of questions. The bound-and-tightness study tests certificate validity across 420 game–model–seed configurations: 280 product-set configurations satisfy the theorem assumptions, while 140 shared-constraint configurations serve only as stress diagnostics. The coupling study evaluates 1,120 configurations across eight coupling values to test validity on both sides of the monotonicity boundary. Both studies use seeds 0–19 and the seven response families described below.

The remaining studies isolate distinct properties of the method. The contraction comparison uses the five response families with weight-derived Lipschitz bounds. Separate experiments test the necessity of playerwise convexity, recovery of multiple equilibria, convergence of Picard, Krasnosel'skiĭ–Mann, and Halpern iterations, finite-difference bias and noise, scaling with the player count, and the cost of a regional audit.

Comparisons use the same seed, game, response family, and coupling whenever those factors are shared. Confidence intervals use $10{,}000$ paired bootstrap resamples. Wilcoxon signed-rank tests use Holm correction within each declared family of comparisons.

Response families. Seven response models receive the concatenated rival profile x_{-i}: a least-squares affine map, 32–32 tanh and ReLU networks, a spectral-normalized network, an Almost-Orthogonal-Layer (AOL) Lipschitz-certified network (Prach & Lampert, 2022), a conditional diffusion sampler, and a conditional flow-matching model. The bound-and-tightness and coupling studies train every family on the same 240 noisy observational transitions per seed. The multiple-equilibrium study instead uses its game-specific multimodal response process.

The first five families provide rigorous weight-derived Lipschitz bounds. The two samplers provide no such bound. Tests that require a global learned-map Lipschitz bound therefore exclude the samplers or record the result as unavailable; lack of a bound is not evidence that the corresponding condition fails.

The deterministic neural heads have two hidden layers of width 32. The generative heads have two hidden layers of width 64 and a time embedding, giving them roughly four times as many parameters per player. Their larger parameter budget rules out smaller capacity as an explanation for poorer equilibrium recovery, but the comparison does not isolate generativity from other architectural differences.

All neural responders use Adam with batch size 128 for 200 epochs. The learning rate is 3×10^{-3}, except for AOL, which uses 5×10^{-3}. Affine and feedforward responders minimize mean-squared prediction error, diffusion uses denoising-score matching, and flow matching uses its conditional-flow objective.

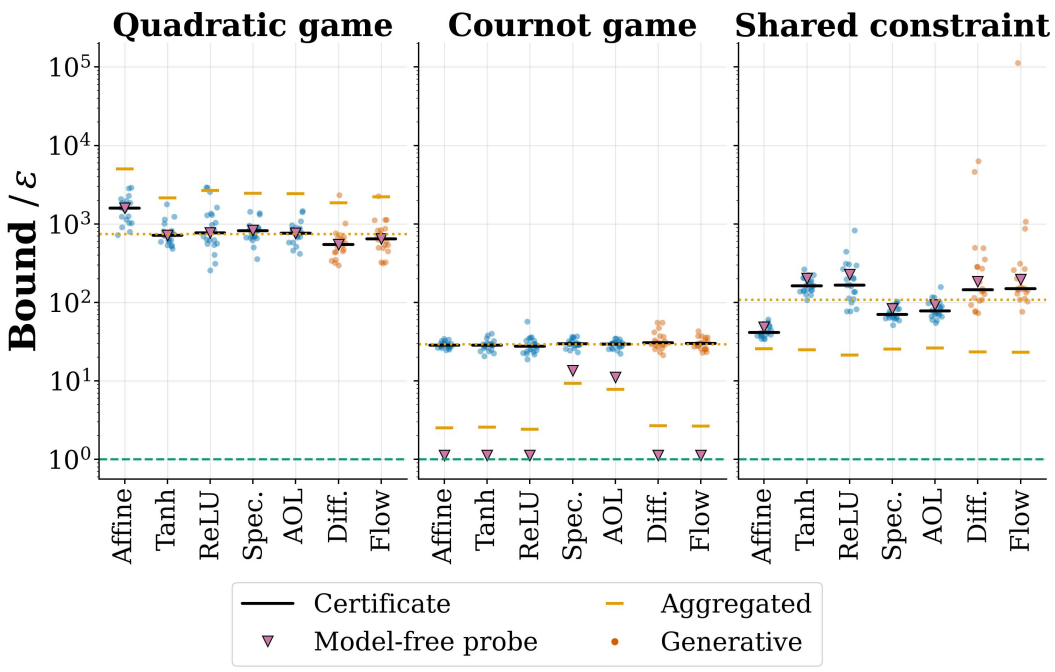


Fig. 4: **Tightness by response family.** Certified bound over true exploitability. Blue marks families with a rigorous Lipschitz bound, orange those without.

H Robustness and Scope

This appendix reports the supporting robustness and scope diagnostics.

Fig. 4 illustrates the certified bound of (12) over true exploitability, testing the pointwise result against architecture changes and game parameters. Here, the dots denote seeds and the bars denote medians. We compare seven response families across three games. Within each game, tightness varies little across response families.

Fig. 5 varies the coupling parameter, which sets the smallest eigenvalue of the symmetrized pseudo-gradient Jacobian and crosses the monotonicity boundary at one. The certificate remains valid in every cell. Median tightness decreases from 3.01×10^3 to 9.86 across the grid; this change concerns looseness, not validity.

H.1 Calibration of the central-difference estimator

Fig. 6 reports the central-difference error in estimating Φ_i against the probe radius τ at five noise scales, with medians over twelve probe points at each radius. The first panel has bias $B_3 > 0$ and exhibits the interior minimum of the bias–noise prediction (19); the second panel sets $B_3 = 0$, where no interior minimum can arise and none does. Dotted verticals mark the balance radius of the prediction of Proposition 4.1. The minimizer over the radius is 2.14 against a multiplicative grid step 2.15. The first panel is the multimodal arena, where $B_3 > 0$. A bias order $B_3\tau^2$ dominates at large τ while noise order σ/τ dominated at small τ. The second panel is a control with $B_3 = 0$, wherein no interior minimum can occur.

H.2 Convexity failure

Fig. 7 compares the exploitability reached from a conditional-mean response with that reached from mode-selected samples, over 20 paired seeds. At the origin, the variational residual vanishes analytically for an exploitability that is 0.25. Mode selection lowers exploitability to 0.0286 and

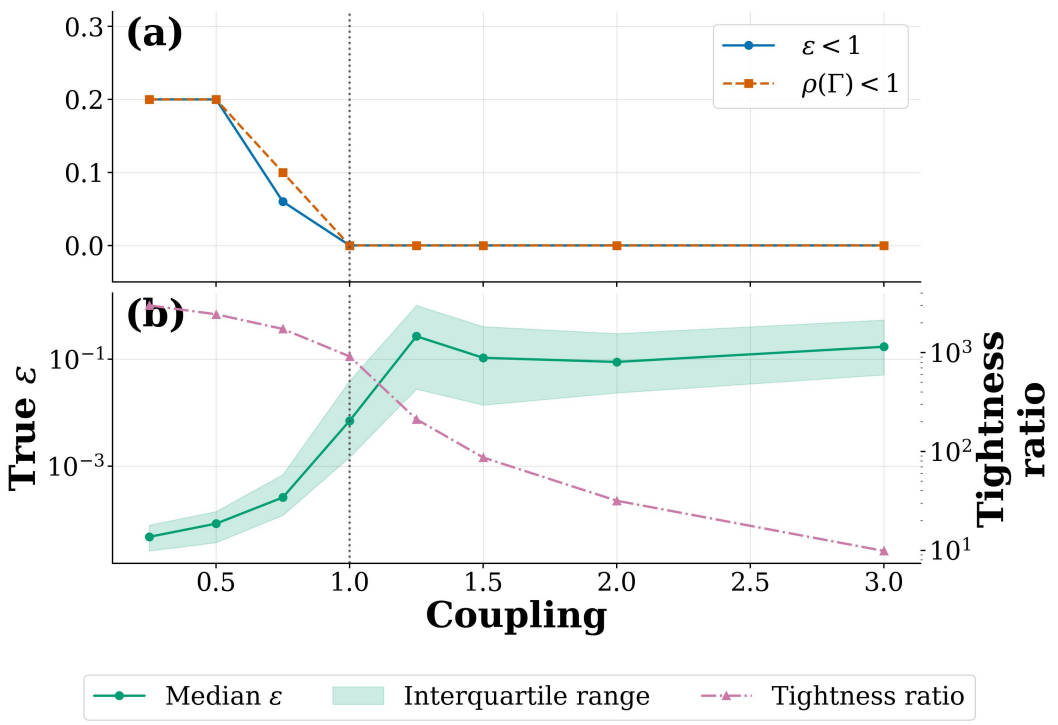


Fig. 5: **Coupling sweep across the monotone boundary.** Coupled quadratic game; the coupling parameter sets $\lambda_{\min}$ of the symmetrized pseudo-gradient Jacobian, so that the grid crosses monotonicity at one.

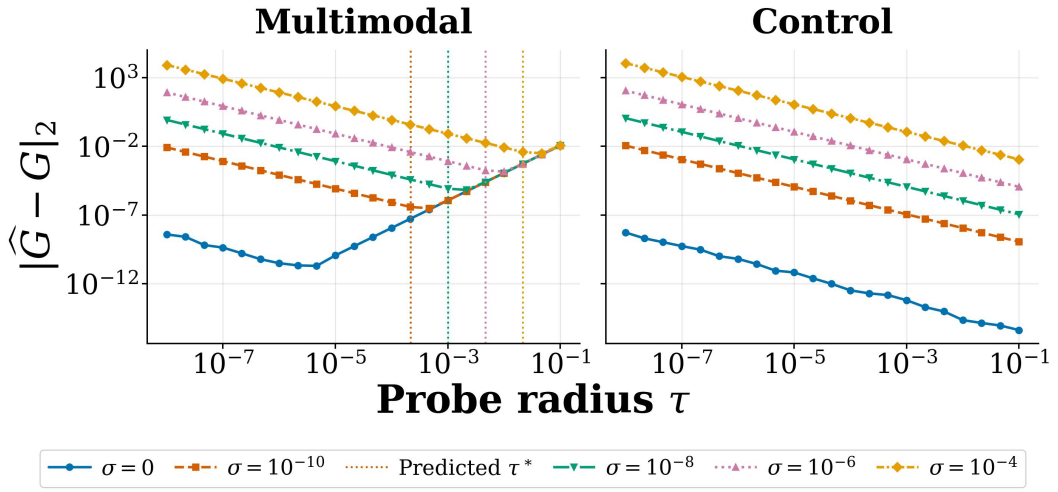


Fig. 6: **Probe error vs. probe radius.** Central-difference error in estimating Φ_i versus τ, at several noise scales σ; medians over twelve probe points and twenty-four noise replicates each.

0.0175. The comparison is paired and between groups: a fraction 0.825 of mode-selected profiles carry both a smaller exploitability and a larger residual than the conditional-mean profile.

Fig. 8 cross-classifies the small-gain condition $\rho(\Gamma) < 1$ and the joint Jacobian spectral-norm condition. The empty cell illustrates the fits, since $\rho(\Gamma) \le \|J_F\|_2$ held in all 1100 of them with

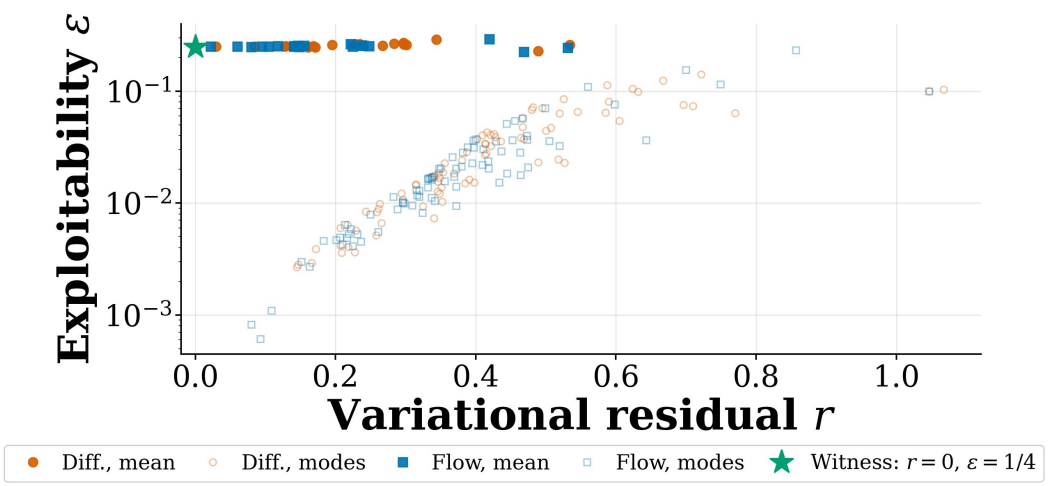


Fig. 7: **A vanishing residual is no evidence of a small gap.** Exploitability against variational residual. Filled markers: conditional-mean fixed points; hollow markers: four profiles per seed; star: analytic result at zero residual and $1/4$ exploitability.

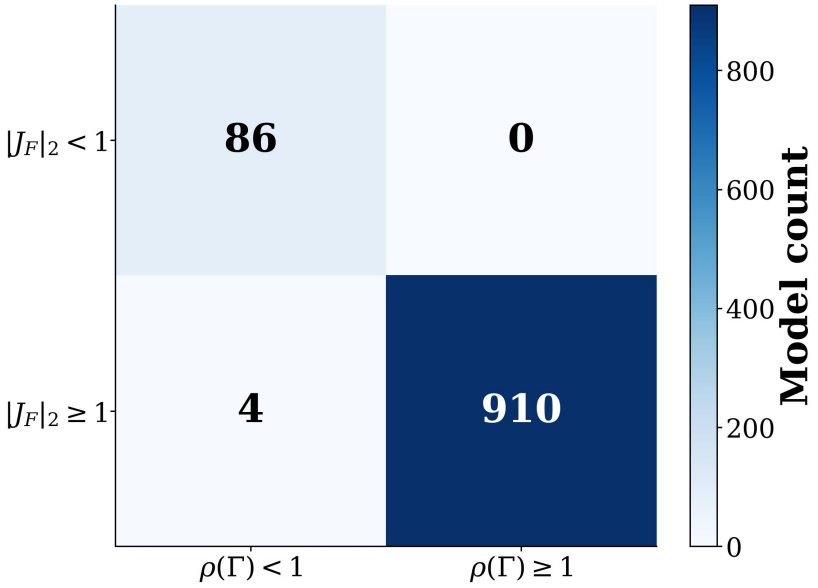


Fig. 8: **Joint Jacobian contraction versus small-gain contraction.**

median slack 0.03969. The direction separating the two conditions analytically is realized nowhere in this population, with all 106 certified cells are affine.

Pooling the theorem-supported product-set cells from the tightness and coupling studies gives 1,000 model cells. The small-gain test accepts 0.09, while 0.524 have exploitability at most 0.01. The contingency counts $(70, 20, 454, 456)$ show that contraction is neither necessary nor sufficient in this population. Shared-constraint diagnostics and generative families without a rigorous global Lipschitz bound are excluded.