When Low-Rank Payoff Compression Changes Equilibrium Structure

Anonymous authors
Paper under double-blind review

Abstract

Low-rank compression is a common way to shrink the payoff matrix of a large zero-sum game, and the best rank-k spectral approximation is the default compressor. Spectral accuracy measures how faithfully payoffs are reproduced, so it leaves open whether the equilibrium set survives. We show that it can fail at any accuracy: the equilibrium sets of a game and of its best rank-k approximation can stay a constant distance apart while the spectral error tends to zero. The same failure appears in the 337×337 sequence form of Leduc poker and in three structurally different matrix-game families, at spectral errors down to 2.9×10^{-3}. The mechanism is selection. When a perturbation of scale ϵ breaks a tie among equilibria, a compression residual η is amplified by ϵ^{-1} on the equilibrium face once global support containment holds, and an explicit family attains the displacement $\eta/(2(4\epsilon + \eta))$. Two guarantees separate the two questions. Set fidelity is local and conditional: face exactness, a uniform separation of outside actions, and a local gap-to-distance bound jointly identify the equilibrium set near a selected face. Output fidelity is pointwise: best responses on the original matrix certify the exploitability of the returned strategy whatever face generated it, and the certificate lifts from actions to strategy populations. The experiments reproduce the failure at finite spectral error on four testbeds, measure the cost of discovering the protected face without an oracle, and run the certificate inside double-oracle loops on three OpenSpiel games. The certified meta-solver returns the exact meta-solution's support in all 396 rounds, with numerically evaluated exploitability below 10^{-13} on the perturbed current meta-game, while a blind meta-game truncation distorts the support and leaves exploitability at the payoff scale.

1 Introduction

Computing equilibria in large zero-sum games requires reasoning over many competing strategies. Double-oracle (DO) methods and policy-space response oracles (PSRO) grow a strategy population through best-response expansion (McMahan et al., 2003; Lanctot et al., 2017; Bighashdel et al., 2024). Each round solves a *meta-game*, the payoff matrix over the strategies found so far (Wellman, 2006). The matrix grows with the population, and its entries become expensive to evaluate.

Prior work shrinks such games in three ways, and each leaves the equilibrium set open. Sparsification rewrites the payoff matrix with fewer nonzero entries while keeping the game exact (Zhang & Sandholm, 2020a; Farina & Sandholm, 2022), so it reduces storage without reducing the strategic problem. Abstraction solves a smaller game over coarser actions or information states and bounds the lifted solution (Sandholm & Singh, 2012; Kroer & Sandholm, 2014; 2018). The lifted bound scales with the abstraction error, and refinement can be non-monotone (Waugh et al., 2009). Low-rank spectral compression is the natural candidate for a growing meta-game: the truncated singular value decomposition (SVD) is the best rank-k approximation in spectral norm (Eckart & Young, 1936; Mirsky, 1960). Spectral optimality measures payoff approximation, and a direction with a small singular value can still determine which actions are optimal. Removing it can create or delete equilibria even when the matrix approximation is accurate, and in population-based training the resulting mixture also determines which best responses enter the next round.

We study what compression must preserve to retain the equilibrium set, and how to certify a returned strategy when that structure is unknown. We focus on the *face* attached to a strategy pair, the product

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

of the simplices spanned by the actions that can carry equilibrium probability. Our face-anchored compression (FACE) augments the spectral subspace with the corresponding coordinate directions, so payoffs are exact on the selected face (Figure 1b, Appendix B). We pair this construction with double-oracle best-response checks (McMahan et al., 2003) on the original matrix, which expand an insufficient support until the returned strategy is certified (Figure 1c).

The analysis distinguishes *set fidelity* from *output fidelity*. Set fidelity combines face exactness, a uniform outside-action margin that survives compression, and a local gap-to-distance bound. Output fidelity certifies one strategy's exploitability directly on the original matrix, without requiring a correct support estimate. The underlying failure mechanism is equilibrium selection: compression error competes with a tie-breaking perturbation and, under support containment and local regularity, is amplified by the inverse perturbation scale.

The experiments trace this mechanism from matrix games to population dynamics (Section 6). In the 337×337 Leduc sequence form (Koller et al., 1996; Southey et al., 2005), blind truncation misses at all five ranks, and the point-to-set distance stays above 10^2 at rank 128. Certification returns a strategy with gap below 10^{-8} from every non-oracle seed support (Figure 1d). Three structurally different matrix-game families fail down to spectral error 2.9×10^{-3}. In DO runs on three OpenSpiel games (Lanctot et al., 2019), the certified meta-solver returns the exact meta-solution's support in all 396 rounds under fixed tie-breaking, while blind truncation distorts the support.

Our contributions are:

- **What spectral compression cannot preserve, and when it can.** A 2×2 family rules out any uniform spectral-error modulus for equilibrium-set distance, and an explicit selection family shows compression error amplified by the inverse tie-breaking scale (Theorems 4.1 and 4.3). Perturbation theory bounds the distance from a point to a fixed program's solution set under data error (Hoffman, 2003; Facchinei & Pang, 2003). These results answer a different question: whether the equilibrium set is preserved. They locate the error in the direction truncation discards and give a sufficient set-stability condition (Theorem 4.5).
- **A diagnostic and a certificate that need no correct support.** FACE makes payoffs exact on a candidate face, so remaining distortion is due to outside actions. Algorithm 1 pairs support-restricted solving with an original-matrix certificate that holds for any seed and lifts to populations (Theorem 5.3); the seed sets its cost, not its correctness.
- **A negative result on compressed meta-solving.** On Leduc, three matrix-game families, and double-oracle pools, blind truncation distorts the equilibrium set to spectral error 2.9×10^{-3}, the certificate returns a certified point from every seed, and in the tested SVD-seeded pipeline a truncated seed costs 12 to 27 times the certified exact full linear program (LP) at every tested rank. The certificate audits correctness but provides no speedup (Section 6).

2 Related Work

Double-oracle methods and PSRO alternate a restricted solve with best-response expansion (Bosansky et al., 2014; McAleer et al., 2021; Tang et al., 2023; Lanctot et al., 2017; Bighashdel et al., 2024; McAleer et al., 2024), and empirical game-theoretic analysis builds the finite game from simulated strategies (Wellman, 2006; Konicki et al., 2022). The closest matrix-level work sparsifies payoffs while keeping enough structure for equilibrium computation (Zhang & Sandholm, 2020a; Farina & Sandholm, 2022). Sparsification re-expresses the same game exactly; spectral truncation discards directions. Working-set expansion alone is constraint generation on the equilibrium LP (Bosansky et al., 2014; McAleer et al., 2021), and our diagnostic separates when such a scheme is sound under a spectral candidate. Three known lines meet here. Hoffman-type bounds control the distance from a point to the solution set of a fixed program under data error (Hoffman, 2003; Facchinei & Pang, 2003). We ask when the equilibrium set of the compressed game coincides with the original one, identify the constant as the inverse tie-breaking scale in the direction truncation discards, and show that no spectral modulus exists without it. Coordinate anchoring then becomes a diagnostic that isolates outside-action error. Best-response certification is standard in double oracle (McMahan et al., 2003; Bosansky et al., 2014). We separate its correctness, which holds for any seed, from its cost, which is what a compressed seed changes. Appendix D reviews abstraction, strategic stability (Kohlberg & Mertens, 1986; Van Damme, 1991), and regularized targets.

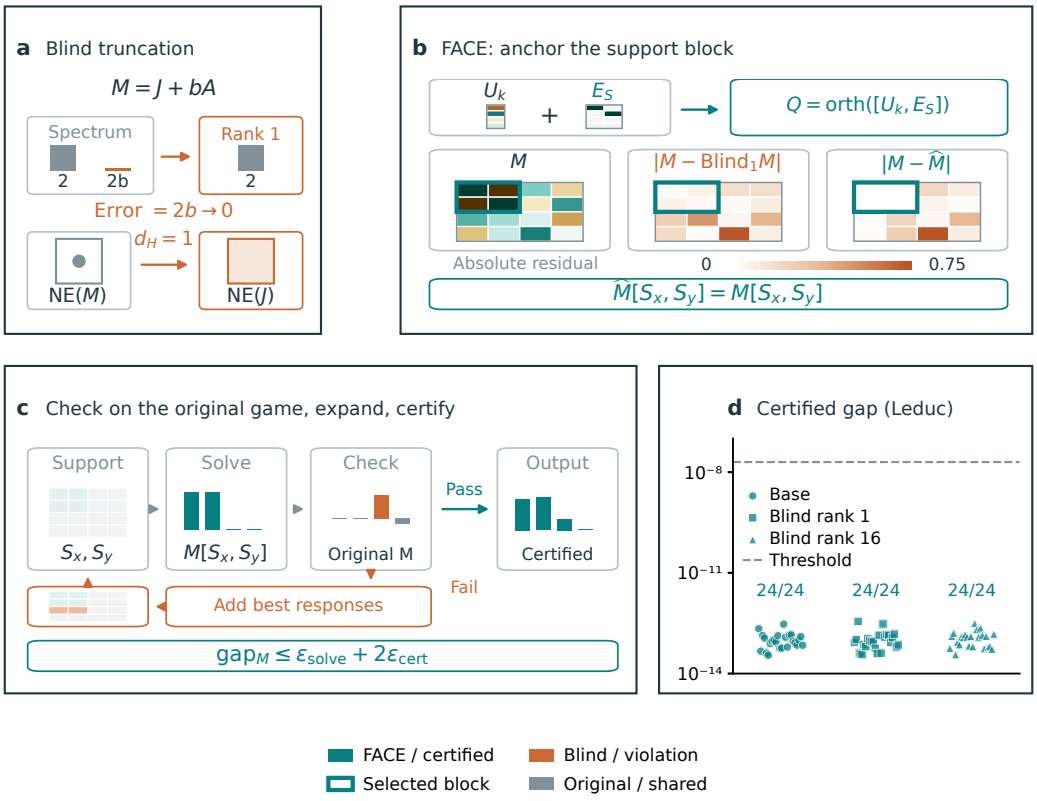


Figure 1: Face-anchored compression with original-game certification. (a) In $M = J + bA$, with A matching pennies, blind rank-one truncation keeps the constant direction and discards the strategic one: the spectral error $2b$ vanishes as $b \to 0$, yet the equilibrium set jumps from the single point $(1/2, 1/2)$ to the whole strategy square, at Hausdorff distance 1 (Theorem 4.1). (b) FACE augments the leading singular directions with the coordinate anchors of a candidate support $S_x \times S_y$. On the 4×4 example of Appendix A the blind rank-one residual corrupts the selected block, while the anchored residual is zero on it (Proposition 5.1). (c) Algorithm 1 solves the restricted game, computes best responses on the original M, and adds violating actions; on return, $\mathrm{Gap}_M \le \epsilon_{\mathrm{solve}} + 2\epsilon_{\mathrm{cert}}$ whatever the seed support (Theorem 5.3). (d) Certified gap on the 337×337 Leduc sequence form for the same 24 perturbed instances from three non-oracle seed supports. All 72 runs terminate below $2\epsilon_{\mathrm{cert}} = 2 \times 10^{-8}$ (Table 5).

3 Problem Setup

We use the standard zero-sum formulation (Nisan et al., 2007). A matrix $M \in \mathbb{R}^{m \times n}$ holds the row player's payoffs: the row player chooses $x \in \Delta_m$ to maximize $x^\top M y$, and the column player chooses $y \in \Delta_n$ to minimize it. For a set of coordinates S, $\Delta(S)$ is the simplex supported on S, so $\Delta_m = \Delta(\{1, \dots, m\})$; we call S the *carrier* of that simplex. The game value is $v(M) = \max_x \min_y x^\top M y$. For $\epsilon \ge 0$ the player-wise approximate solution sets are

$$X_\epsilon(M) = \{x \in \Delta_m : M^\top x \ge (v(M) - \epsilon)\mathbf{1}\},$$
$$Y_\epsilon(M) = \{y \in \Delta_n : My \le (v(M) + \epsilon)\mathbf{1}\},$$

with $\mathrm{NE}_\epsilon(M) = X_\epsilon(M) \times Y_\epsilon(M)$ and $\mathrm{NE}(M) = \mathrm{NE}_0(M)$. The exploitability, or gap, of a pair $z = (x, y)$ is

$$\mathrm{Gap}_M(z) = \max_i (My)_i - \min_j (x^\top M)_j,$$

which is nonnegative and vanishes exactly at equilibrium. For compact sets A, B we use the two-sided Euclidean Hausdorff distance $d_H(A, B)$. For a matrix A we write $\|A\|_{\max} = \max_{ij} |A_{ij}|$ for its largest entry in absolute value.

Compression operators. Blind rank truncation keeps the k leading singular directions. For $1 \le k \le \min(m,n)$, with $\sigma_{\min(m,n)+1} = 0$, we set $\mathrm{Blind}_k(M) = U_k \Sigma_k V_k^\top$, so that $\|M - \mathrm{Blind}_k(M)\|_2 = \sigma_{k+1}(M)$, a best rank-$k$ approximation in spectral norm (Eckart & Young, 1936; Mirsky, 1960). It is single-valued whenever $\sigma_k \neq \sigma_{k+1}$, and when $\sigma_k = \sigma_{k+1}$ we fix a deterministic pairing of leading singular subspaces, which keeps the operator single-valued globally. The selection reads only the spectral data of M: it is blind to which entries determine equilibrium indifference.

Face-anchored compression takes candidate supports S_x, S_y. With E_S the coordinate columns of a set S and $\mathrm{orth}(\cdot)$ an orthonormal basis of the column span, it forms $Q_x = \mathrm{orth}([U_k, E_{S_x}])$ and $Q_y = \mathrm{orth}([V_k, E_{S_y}])$ and returns $\widehat{M} = Q_x Q_x^\top M Q_y Q_y^\top$. The residual $E = M - \widehat{M}$ carries two measures. The *full payoff perturbation* $\eta = \|E\|_{\max}$ is the largest change in any pure action's payoff against a mixed opponent. The *face-direction perturbation* $\eta_F = \max\{\max_{j \in S_y} \|(M - \widehat{M})e_j\|_\infty, \max_{i \in S_x} \|e_i^\top (M - \widehat{M})\|_\infty\}$ restricts the opponent to $\Delta(S_y)$ or $\Delta(S_x)$ and satisfies $\eta_F \le \eta$. A linear function attains its supremum over a simplex at a vertex, so this quantity has the closed form above, and it is the only residual that a protected equilibrium is exposed to.

Face geometry and the margin. Fix an equilibrium $(x^\star, y^\star)$ with support $S_x \times S_y$ and let $F = \Delta(S_x) \times \Delta(S_y)$ be its *face*. Write v_F for the value of the restricted game $M[S_x, S_y]$, and let Z_F be its equilibrium set, embedded in the full simplices. The *uniform outside-action margin* is

$$\gamma = \min\Big\{ \inf_{y \in Y_0(M[S_x,S_y])} \big[v_F - \max_{i \notin S_x} (My)_i\big], \quad \inf_{x \in X_0(M[S_x,S_y])} \big[\min_{j \notin S_y} (x^\top M)_j - v_F\big] \Big\}.$$

It measures how far outside actions are from becoming best responses. The infima run over the projections of Z_F, because a basic LP solution can overstate the margin near indifference, and a side without outside actions contributes $+\infty$. In the double-oracle experiments of Section 6, M is the empirical meta-game over the current strategy populations (Wellman, 2006; Lanctot et al., 2017) and the support is the equilibrium mixture over those members.

Selection and local regularity. A face is *flat* with margin $\gamma > 0$ when $M_{ij} = v$ on $S_x \times S_y$, when $M_{ij} \le v - \gamma$ for $i \notin S_x$ and $j \in S_y$, and when $M_{ij} \ge v + \gamma$ for $i \in S_x$ and $j \notin S_y$. Flatness forces $\mathrm{NE}(M) = \Delta(S_x) \times \Delta(S_y)$ (Appendix A), the geometry of Theorem 4.1 and of the explicit family below. A *perturbation direction* Δ with $\|\Delta\|_{\max} \le 1$ defines the *selection game* $\Delta_F = \Delta[S_x, S_y]$ on the carrier simplices, and we write $\mathrm{Gap}^{\mathrm{sel}}_\Delta(z)$ for the gap of a face-supported pair z in that game. We call the selection game *regular* when it has a unique saddle point z_Δ and satisfies the local error bound below; we then write κ_Δ and r_Δ for its constants.

Definition 3.1 (Local gap-to-distance error bound)**.** A game G has local constant κ and radius $r > 0$ around a set Z if $\mathrm{dist}(z, Z) \le \kappa\, \mathrm{Gap}_G(z)$ whenever $\mathrm{Gap}_G(z) \le r$, with dist the product ℓ_1 distance. When Z is the full equilibrium set of a finite matrix game, this is a Hoffman-type error bound whose constant depends on the active inequalities and the payoff scale (Hoffman, 2003); we use the definition only for such sets.

Certificates and sequence form. Algorithm 1 solves a restricted game to tolerance $\epsilon_{\mathrm{solve}}$ and stops when the two outside best-response residuals fall below a per-side tolerance ϵ_{cert}, so the returned pair has gap at most $\epsilon_{\mathrm{solve}} + 2\epsilon_{\mathrm{cert}}$ on the original matrix (Theorem 5.3). Sequence-form games replace the simplices by realization-plan polytopes $P_p = \{x \ge 0 : F_p x = f_p\}$ over n_p sequences, and a *behavioral support* is a set of (information set, action) pairs. It is *closed* when it holds the parent of every sequence in it and a continuation at every information set it reaches; closure makes the restricted polytope a face of P_p (Koller et al., 1996). Write C_p for the number of rows of F_p.

4 Theory

Spectral Error Does Not Control the Equilibrium Set. A low-energy direction can be strategically decisive. The cleanest example adds a small matching-pennies component to a large constant matrix. The constant dominates the spectrum and leaves every incentive unchanged, while the small component determines the equilibrium.

Theorem 4.1 (No uniform spectral-error modulus)**.** *There is no function $\omega : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ with $\omega(t) \to 0$ as $t \to 0$ such that*

$$d_H\big(\mathrm{NE}(M), \mathrm{NE}(\mathrm{Blind}_1(M))\big) \le \omega\big(\|M - \mathrm{Blind}_1(M)\|_2\big)$$

holds for every two-player zero-sum matrix game M.

The construction is a one-parameter family $M_b = J + bA$ with $0 < b < 1$, where J denotes the all-ones matrix and A the matching-pennies payoff. The row and column spaces of J and bA are orthogonal, with nonzero singular values 2 and $2b$, so $\mathrm{Blind}_1(M_b) = J$ and the spectral error is $2b \to 0$. The gap $2 > 2b$ is strict, so the rank-one truncated SVD is unambiguous on this witness. Adding the constant matrix J preserves best responses (von Stengel, 2007), so M_b has the unique equilibrium $(1/2, 1/2)$ of matching pennies, while every pair of mixed strategies is an equilibrium of J. In the Euclidean product metric of Section 3 the farthest pure corner from $(1/2, 1/2)$ lies at distance exactly 1. The Hausdorff distance between the two equilibrium sets is therefore 1 for every $b \in (0, 1)$, and the discarded direction has genuinely nonzero singular value $2b$.

A positive result therefore needs geometry that keeps discarded directions from changing best responses, as in Theorem 4.5, or certification of the returned strategy on the original game, as in Theorem 5.3 and Algorithm 1. The proof is in Appendix A.

Selection Amplifies Compression Error. Theorem 4.1 rules out a uniform modulus. When the base game has a non-singleton equilibrium set, a sharper statement holds under the flat-carrier and unique-selection hypotheses of Section 3. A small perturbation that breaks the tie fixes the scale on which compression error acts, and the error is amplified by the inverse perturbation size.

The selection identity makes this precise. For a face-supported pair z and $0 \le \epsilon < \gamma/2$,

$$\mathrm{Gap}_{M+\epsilon\Delta}(z) = \epsilon\, \mathrm{Gap}^{\mathrm{sel}}_{\Delta}(z) \qquad \text{(Appendix A)},$$

because on the face the base game contributes only the constant v, so an approximation error of size η perturbs a signal of size ϵ. The margin pins equilibria to the face with zero outside mass:

Lemma 4.2 (Support containment)**.** *Let M have a flat face $S_x \times S_y$ with margin $\gamma > 0$. If $\|W' - M\|_{\max} \le \rho$ and $4\rho(\gamma + \|M\|_{\max}) < \gamma^2$, then every $z' \in \mathrm{NE}(W')$ is supported on $S_x \times S_y$ and $|v(W') - v| \le \rho$.*

Hence:

Theorem 4.3 (Selection stability under compression)**.** *Let M have a flat carrier face with margin $\gamma > 0$, let Δ have $\|\Delta\|_{\max} \le 1$, and let the selection game $\Delta[S_x, S_y]$ be regular. Let $\widehat{W}$ be any matrix with $\|M + \epsilon\Delta - \widehat{W}\|_{\max} \le \eta$. If $\epsilon > 0$, $4(\epsilon + \eta)(\gamma + \|M\|_{\max}) < \gamma^2$, and $2\eta/\epsilon \le r_\Delta$, then z_Δ is the unique equilibrium of $M + \epsilon\Delta$, and every $\hat{z} \in \mathrm{NE}(\widehat{W})$ is face-supported with*

$$\mathrm{dist}(\hat{z}, z_\Delta) \;\le\; \min\Big\{\, 2\kappa_\Delta\, \frac{\eta}{\epsilon},\; 4 \Big\}.$$

The displacement is at most $2\kappa_\Delta\eta/\epsilon$, with κ_Δ set by the selection game. Compression error acts relative to the tie-breaking scale, so an error far below the payoff scale dominates the selection once it is comparable to ϵ. The next result shows that the rate is tight.

Proposition 4.4 (Explicit family: the rate is attained)**.** *For every $\gamma > 0$, $\epsilon \in (0, \gamma/4)$, and $\eta \in [0, \gamma/2]$ there is a 3×3 game M with flat carrier face $\{1, 2\}^2$ and margin γ, a perturbation Δ with $\|\Delta\|_{\max} = 1$, and an approximation $\widehat{W}$ of $M + \epsilon\Delta$ with $\|M + \epsilon\Delta - \widehat{W}\|_{\max} = \eta$, such that both perturbed games have unique equilibria. Write $x^\star(\epsilon)$ for the row strategy of the equilibrium of $M + \epsilon\Delta$ and $\hat{x}$ for that of $\widehat{W}$. The two differ by*

$$\|\hat{x} - x^\star(\epsilon)\|_\infty = \frac{\eta}{2(4\epsilon + \eta)}.$$

Both players are displaced by this amount, so the product-ℓ_1 distance is $2\eta/(4\epsilon+\eta)$, at least $\eta/(4\epsilon)$ for $\eta \le 4\epsilon$, and it approaches 2 as $\eta/\epsilon \to \infty$.

The bounds meet in the family itself. Its selection game is matching pennies, whose duality gap equals the product-ℓ_1 distance to its saddle point, so $\kappa_\Delta = 1$. Both bounds apply on the intersection of the explicit-family regime with the hypotheses of Theorem 4.3, in particular $\eta \le \epsilon$, $4(\epsilon + \eta)(\gamma + \|M\|_{\max}) < \gamma^2$, and $2\eta/\epsilon \le r_\Delta$. There the theorem's $2\eta/\epsilon$ and the attained $2\eta/(4\epsilon+\eta) \ge 2\eta/(5\epsilon)$ differ by a universal constant. The proposition gives the worst-case rate over entrywise errors of size η; a compressor attains it when its residual aligns with the selection-sensitive face direction, which we measure for rank-k truncation in Section 6. This reframes Theorem 4.1: the face residual of blind truncation takes over once it reaches the perturbation scale. Proofs are in Appendix A.

Conditional Set Stability. In the counterexample of Theorem 4.1, truncation fails because it changes payoffs on the equilibrium face, destroying the indifference that defines it. Truncation can equally erase the margin that keeps outside actions out. A positive result must rule out both. Fix a face $F = \Delta(S_x) \times \Delta(S_y)$ and its equilibrium set Z_F as in Section 3. Compression is *face exact* when $(M - \widehat{M})[S_x, S_y] = 0$.

Theorem 4.5 (Conditional two-sided stability)**.** *Suppose that (i) $\widehat{M}$ is exact on F; (ii) every action outside S_x or S_y is worse than the restricted value by the uniform margin $\gamma > 0$, the setup margin of Section 3 infimized over the projections of Z_F; (iii) the face-direction perturbation satisfies $\eta_F < \gamma$; and (iv) M satisfies Definition 3.1 around Z_F, with constant κ and radius r. Then* $\mathrm{NE}(M) = \mathrm{NE}(\widehat{M}) = Z_F$*, and for every $\epsilon \geq 0$ with $2(\epsilon + \eta) \leq r$,*

$$d_H\big(\mathrm{NE}_\epsilon(M), \mathrm{NE}_\epsilon(\widehat{M})\big) \leq 2\kappa(\epsilon + \eta).$$

At $\epsilon = 0$ the first conclusion gives exact set equality; for $\epsilon > 0$ the displayed bound is the operative statement.

Conditions (i) to (iii) identify the same exact equilibrium set in both games. Face exactness gives a common restricted game. The uniform margin keeps every Z_F strategy a full-game equilibrium and confines any full-game equilibrium to F, and $\eta_F < \gamma$ preserves that separation. Condition (iv) controls how nearby gap sublevels contract to the common set. Payoff perturbation moves exploitability by at most 2η:

Lemma 4.6 (Gap propagation)**.** *If $\|M - \widehat{M}\|_{\max} \leq \eta$, then $|\operatorname{Gap}_M(z) - \operatorname{Gap}_{\widehat{M}}(z)| \leq 2\eta$ for all z.*

The Nash set is convex (Nisan et al., 2007). Proofs are in Appendix A.

5 METHOD

Face-Anchored Compression. Theorem 4.5 assumes the compressed game is exact on a chosen face. FACE enforces this condition by construction.

Proposition 5.1 (Face exactness)**.** *For every matrix M, rank budget k, and candidate support $S_x \times S_y$, the* FACE *projection satisfies*

$$\widehat{M}[S_x, S_y] = M[S_x, S_y].$$

Face exactness is unconditional: it holds for every singular-value decay, every margin, and every support estimate. What compression changes is how the outside gradient coordinates move.

Proposition 5.2 (Gradient separation)**.** *Let $\widehat{M} = Q_x Q_x^\top M Q_y Q_y^\top$ with $\mathrm{range}(Q_x) \supseteq \mathrm{span}\{e_i : i \in S_x\}$ and $\mathrm{range}(Q_y) \supseteq \mathrm{span}\{e_j : j \in S_y\}$. For any pair (x, y) with $\mathrm{supp}(x) \subseteq S_x$ and $\mathrm{supp}(y) \subseteq S_y$: (i) $(\widehat{M}y)_i = (My)_i$ for every $i \in S_x$, and $(x^\top \widehat{M})_j = (x^\top M)_j$ for every $j \in S_y$; (ii) $(\widehat{M}y)_i = (Q_x Q_x^\top My)_i$ for $i \notin S_x$, and $(x^\top \widehat{M})_j = (x^\top M Q_y Q_y^\top)_j$ for $j \notin S_y$.*

Restricted solves therefore coincide with or without compression, and face-confined solvers see identical gradients. Full-space solvers can differ even at a face-supported iterate through the outside gradient coordinates (Appendix A.11). The residual an outside action sees against a face strategy is $(M - \widehat{M})y = (I - Q_x Q_x^\top)My$, which gives the computable envelope

$$\eta_{\mathrm{face}} = \|(I - Q_x Q_x^\top)M\|_{2\to\infty} + \|(I - Q_y Q_y^\top)M^\top\|_{2\to\infty},$$

where $\|A\|_{2\to\infty} = \sup_{\|u\|_2 \leq 1} \|Au\|_\infty$ and $\eta_F \leq \eta_{\mathrm{face}}$ (Appendix A). The diagnostic $\eta_{\mathrm{face}} < \hat{\gamma}/2$ screens for condition (iii) without an oracle. Here $\hat{\gamma}$ is the margin of Section 3 evaluated at the solved equilibrium; the setup margin infimizes the same quantity over Z_F. Since $\hat{\gamma} \geq \gamma$, the diagnostic certifies condition (iii) when Z_F is a point and serves as a heuristic otherwise. The evaluation measures a one-sided sample, d_{diag} (Appendix B.10); η_{face} itself is a two-sided operator quantity.

The decisive step is the original-game best-response certificate used in double-oracle methods (McMahan et al., 2003; Bosansky et al., 2014), which Algorithm 1 decomposes as follows.

Algorithm 1 Certified support-restricted solving

Require: payoff matrix M, nonnegative tolerances $\epsilon_{\text{solve}}, \epsilon_{\text{cert}} \geq 0$, nonempty supports S_0, S_1

1: **repeat**
2: solve $M[S_0, S_1]$ to restricted gap $g_S \leq \epsilon_{\text{solve}}$; embed (x, y) in the full action spaces
3: compute on M the best responses $i^\star \in \arg\max_i (My)_i$ and $j^\star \in \arg\min_j (x^\top M)_j$
4: $r_x \leftarrow (My)_{i^\star} - \max_{i \in S_0} (My)_i$; $r_y \leftarrow \min_{j \in S_1} (x^\top M)_j - (x^\top M)_{j^\star}$
5: **if** $r_x \leq \epsilon_{\text{cert}}$ and $r_y \leq \epsilon_{\text{cert}}$ **then**
6: **return** (x, y), with $\text{Gap}_M(x, y) \leq \epsilon_{\text{solve}} + 2\epsilon_{\text{cert}}$ by Theorem 5.3
7: **end if**
8: $S_0 \leftarrow S_0 \cup \{i^\star\}$; $S_1 \leftarrow S_1 \cup \{j^\star\}$
9: **until** return

Theorem 5.3 (Certified output fidelity)**.** *Let a restricted candidate (x, y) satisfy*

$$g_S = \max_{i \in S_x} (My)_i - \min_{j \in S_y} (x^\top M)_j \leq \epsilon_{\text{solve}}.$$

Define the two outside residuals

$$r_x = \max_i (My)_i - \max_{i \in S_x} (My)_i,$$
$$r_y = \min_{j \in S_y} (x^\top M)_j - \min_j (x^\top M)_j.$$

If the loop stops only when $r_x, r_y \leq \epsilon_{\text{cert}}$, then

$$\text{Gap}_M(x, y) = g_S + r_x + r_y \leq \epsilon_{\text{solve}} + 2\epsilon_{\text{cert}}.$$

This guarantee is independent of η.

Theorem 4.5 supplies the set-level guarantee when its conditions hold. The certificate supplies the pointwise guarantee in every case, independent of how the candidate face was generated, and costs one matrix-vector product per side with no support oracle, margin, or non-degeneracy constant (Section A.3). Proofs and a 4×4 numerical illustration are in Appendix A.

Certified Support-Restricted Solving. Algorithm 1 uses double-oracle restricted solving and best-response expansion (McMahan et al., 2003; Bosansky et al., 2014). The supports $S_0 \subseteq [m]$ and $S_1 \subseteq [n]$ are index sets, $M[S_0, S_1]$ is the corresponding submatrix, and the restricted game is the zero-sum game $M[S_0, S_1]$ on $\Delta(S_0) \times \Delta(S_1)$. Sequence-form games replace the restricted simplices by restricted realization-plan polytopes whose supports are closed (Section 3); Proposition B.1 states that version.

Lemma 5.4 (Termination)**.** *With nonempty initial supports and nonnegative tolerances, each round of Algorithm 1 that does not return adds at least one action to S_0 or to S_1. The loop therefore returns after at most $m + n$ rounds.*

The worst case is expansion to the full game, and the favorable case is a single round. The two together make the algorithm a prediction and certification scheme: a cheap support estimate is a prediction, the true-matrix best responses certify it, and a wrong prediction costs extra solves. Proposition B.1 in Appendix B transfers the certificate and the termination count verbatim to sequence-form (Koller et al., 1996) realization-plan polytopes, which is the form the Leduc experiments run.

Proposition 5.5 (One-shot certification)**.** *Suppose the restricted game $M[S_0, S_1]$ has a unique equilibrium and its embedding is an equilibrium of M. With an exact restricted solve and the nonnegative tolerances of Algorithm 1, it returns in the first round.*

By hypothesis the seed is *support-stable*: its restricted equilibrium is already an equilibrium of the full game. The Leduc seeds of Section 6 are not support-stable and pay 8 to 22 restricted solves.

The certificate also lifts to policy populations.

Proposition 5.6 (Population lift)**.** *With a fixed payoff u over finite classes $\mathcal{P}_0, \mathcal{P}_1$, Algorithm 1 given exact best responses in place of line 3 returns a mixture with $\text{Gap}_u \leq \epsilon_{\text{solve}} + 2\epsilon_{\text{cert}}$ over $\Delta(\mathcal{P}_0) \times \Delta(\mathcal{P}_1)$.*

Table 1: Leduc oracle-face diagnostic, interquartile means over 24 instances per rank. $\|M - M_k\|_2$ is the rank-k spectral error; d_{set} is the point-to-set distance to the 10^{-6}-relaxed equilibrium polytopes of the original game; P/R is support precision/recall against the oracle support; $|S_0|/|S_1|$ are median oracle support sizes. The certified columns are identical across ranks: the pair comes from one oracle-seeded run of Algorithm 1 on the original matrix, which no rank or compressor enters.

		Blind-k				Oracle-seeded certified solve				
k	$\|M - M_k\|_2$	hits	d_{set}	gap	P/R	hits	d_{set}	gap	P/R	$\|S_0\|/\|S_1\|$
1	1.49	0/24	135.93	9.04	.70/.17	24/24	0	$\approx 9 \times 10^{-14}$	.994/.978	181.5/191.5
4	1.49	0/24	134.00	8.71	.68/.18	24/24	0	$\approx 9 \times 10^{-14}$	.994/.978	181.5/191.5
16	1.21	0/24	128.32	9.01	.68/.16	24/24	0	$\approx 9 \times 10^{-14}$	.994/.978	181.5/191.5
64	0.92	0/24	124.73	2.51	.67/.13	24/24	0	$\approx 9 \times 10^{-14}$	.994/.978	181.5/191.5
128	0.41	0/24	115.33	1.89	.76/.23	24/24	0	$\approx 9 \times 10^{-14}$	.994/.978	181.5/191.5

The oracle and the certificate share one payoff function, and the guarantee is per round rather than per pool: it bounds the gap of the returned mixture on the pool it was solved against. When the two differ, as in the perturbed meta-game of Section 6, the certified gap transfers to the original payoff only up to the perturbation, and a pool-level statement needs the outer loop's own convergence.

6 EXPERIMENTS

All numerical claims are recomputed from saved per-instance records. Leduc uses full-game equilibrium supports as oracles to isolate face exactness. A *hit* requires $d_{\text{set}} \leq 10^{-5}$ and gap $\leq 10^{-6}$. A 3×3 probe separates point fidelity from set fidelity: blind rank two still loses a member of the true support family at spectral error 5×10^{-4}, so preserving every equilibrium is strictly stronger than keeping one solution accurate (Section B.1). The displacement law of Proposition 4.4 is checked in Section B.5.

Leduc sequence form. The 337×337 sequence form (Koller et al., 1996) of Leduc poker (Southey et al., 2005) carries the failure to a real game. Each instance is perturbed by a masked Rademacher matrix of unit Frobenius norm scaled by $\epsilon \in \{10^{-3}, 10^{-2}, 10^{-1}, 3 \times 10^{-1}\}$, and six seeds, four levels, and five ranks give 24 instances per rank (Appendix B). Table 1 reports the outcome. Blind truncation never meets the hit criterion. At $k = 128$ the spectral error has fallen to 0.41, while the gap is 1.89 and $d_{\text{set}} = 115.33$. Feasible realization plans carry an ℓ_1 mass between 25 and 91 per player, so this displacement rearranges a constant fraction of the plan. Seeded with the oracle face, the certified arm of Algorithm 1 terminates in all 24 instances at every rank, with a numerically checked certified gap (Appendix B) whose interquartile mean (IQM) (Agarwal et al., 2021) is 9.2×10^{-14}. The corner alone does not certify, and the loop takes a median of 8 rounds with a range of 6 to 10, so the successes belong to seeded certification. The anchored projector spans 56 to 85% of the ambient dimension, so face exactness costs more storage than the original sparse matrix. A resource-matched run shows the compressed path costs more time: seeding Algorithm 1 from the rank-k truncated solve takes a median of 20 to 22 restricted solves, and 12 to 27 times the wall-clock of the exact full LP with its certificate, at every $k \in \{1, 16, 64, 128\}$ (Appendix B).

Non-oracle support discovery and repair. A deterministic 6×6 game with a 3×3 rock-paper-scissors core and three outside actions per player, all worse by a margin γ, measures discovery without an oracle. Support is estimated from an anchored quantal response equilibrium (QRE) temperature path (McKelvey & Palfrey, 1995) that retains probabilities above $\rho = 10^{-3}$ of the maximum. Sweeping γ from 10^{-1} to 10^{-5} (Table 4) leaves the estimate exact down to $\gamma = 10^{-3}$, where the path already misses its residual tolerance. Non-convergence and support error are therefore distinct failure modes. Below the softmax resolution boundary the estimate expands to all six actions per side with recall one and precision halved, and the full-matrix gap stays certified: the cost is lost dimension reduction and correctness is preserved. The same repair mechanism transfers to Leduc from three non-oracle seeds, the equilibrium support of the unperturbed base game and those of the rank-1 and rank-16 truncations. All three seed types terminate certified in all 24 instances, with a median certified gap of 9.5×10^{-14}. Expansion raises median recall to 0.94 to 0.98 from a seed range of 0.07 to 0.85; the smallest final recall across the 72 instances is 0.88 (Table 5). The solve count reflects how far a seed sits from a certified equilibrium: 8 restricted solves

Table 2: Matrix games of different structure, protocol as in Table 1, with 24 instances per rank. $|S_0|/|S_1|$: median oracle best-response face sizes (Appendix B.3). "miss": pairs failing the hit criterion; "min error": smallest rank-wise IQM spectral error among all-miss pairs; "oracle-seeded hit": a single restricted solve on that face in the original matrix. Expansion covers 72 instances from base and rank-1/16 seeds; "recall": IQM of $\min(\mathrm{rec}_0, \mathrm{rec}_1)$ against the oracle best-response face.

		Blind-k		oracle seed	non-oracle expansion		
game	$\lvert S_0\rvert/\lvert S_1\rvert$	miss	min error	hit	cert.	solves	recall
random sparse	117/117	120/120	4.01	120/120	72/72	127	1.00
column-noise	134.5/136.5	120/120	2.9×10^{-3}	120/120	72/72	136	0.98
wide face	247.5/250	36/120	1.73	115/120	72/72	3	0.18

from the base seed against 22 and 20 from the truncation seeds. None of these seeds meets the support-stability hypothesis of Proposition 5.5, so the counts measure the cost of the other regime. The local constant of Theorem 4.5 is instance-dependent: on the 6×6 game the product-ℓ_1 ratio $\mathrm{dist}(z, Z_F)/\operatorname{Gap}_M(z) = 2/\gamma$ diverges as $\gamma \to 0$, so κ is certified per game (Appendix B).

Structurally different matrix games. We repeat the Leduc protocol on three 250×250 zero-sum games that differ from poker (Appendix B.3). Each family perturbs one fixed base matrix with its six instance seeds. *Sparse random* draws 7500 coordinates with replacement and independent standard-normal values, leaving a spectrum no rank budget $k \leq 128$ can approximate. *Column noise* hides a rank-4 backbone under 10^{-3}-scale normal noise in the first half of the columns, placing the noise near the tail of the spectrum. *Wide face* places a rock–paper–scissors core of value 0 on the first three rows and columns, adds 10^{-4}-scale noise to the 4×4 block on rows and columns 3–6, and leaves the rest exactly zero, so the last 243 rows and columns stay zero and $\mathrm{NE}(M)$ contains $\Delta(\{7, \ldots, 249\})^2$ in every instance. The sparse random family fails at every rank, with spectral error 4.0 to 10.5. The column-noise family also fails at every rank, down to spectral error 2.9×10^{-3} at $k = 128$ with distance 1.23. The wide-face family is low-rank by construction and forms the contrast: blind truncation succeeds once the rank budget covers the core and fails only at ranks 1 and 4. The oracle face is the *best-response face* of the LP pair (x°, y°): the rows i with $(My^\circ)_i \geq v - 10^{-6}$ and the columns j with $(M^\top x^\circ)_j \leq v + 10^{-6}$, which on the wide-face family run to 247–250 rows and columns. With the oracle best-response face as seed, a single restricted solve returns a certified point of the original equilibrium set in all 120 instance-rank pairs of the sparse and column-noise families and in 115 of 120 on the wide-face family. The five misses are one instance at all five ranks: its restricted LP selects a restricted equilibrium that an excluded column exploits by 1.1×10^{-4}. The expansion loop of Algorithm 1 then certifies a point after adding that column (Appendix B.3). Non-oracle expansion certifies all 72 instances of every game, with restricted solves near 127 and 136 on the first two families and 3 on the wide-face family, where any corner of the large face is already a valid point certificate. Table 2 collects the results, and Figure 3 shows the same failures in the spectral-error and distance plane (Appendix B.3).

Certified meta-solving in double oracle. A double-oracle (McMahan et al., 2003) loop on three OpenSpiel games (Lanctot et al., 2019), Kuhn (Kuhn, 1950), Leduc, and one-die liar's dice with six seeds each, grows policy pools by exact best responses against the aggregate mixture. Each round the exact meta-game receives a fixed seed-indexed entry perturbation of scale $\epsilon = 10^{-2}$ that is invariant under pool growth. It then goes to one of three meta-solvers: exact LP, blind truncation at rank $\lceil T/2 \rceil$ of the pool size, or Algorithm 1 seeded from the previous round's final working set. Table 7 reports the outcome on the uncompressed perturbed meta-game. The blind meta-solution is uncertified, and its exploitability reaches the payoff scale, with a median of 0.33 on Leduc that is 33 times the perturbation, while its support differs from the reference solution's in 68 to 88% of rounds. The population therefore grows in directions chosen by truncation error. Under the fixed tie-breaking of the runs, the certified meta-solver's mixture has the same support as the exact LP solution of the same pool in all 396 logged rounds. A gap g certified on the perturbed pool bounds the true-game gap over that pool by $g + 2\epsilon$ (Appendix B.6). On one shared pool trajectory at rank $\lceil T/4 \rceil$, the truncated candidate's support differs from the exact LP solution in 361 of 396 rounds. Algorithm 1 seeded from that candidate returns the exact LP support in all 396 rounds, at a median of 3 to 8 restricted solves against 2 from the previous round's working set.

AI USE STATEMENT

AI assistance took two forms in this work: creating and editing the experimental code, and editing the manuscript for readability. Such assistance was not used to produce synthetic data, build theoretical models, state or prove mathematical claims, choose the experimental design, or interpret the results. Every piece of AI-assisted code was executed, and its output checked against the recorded experimental logs. All numbers, proofs, and conclusions in this manuscript were written and verified by the authors, who are responsible for the final content, including any text, claims, or artifacts that AI tools helped produce.

REPRODUCIBILITY STATEMENT

Every matrix perturbation instance is fixed by its integer seed; the double-oracle entry perturbations are fixed within a run but keyed by a process-dependent hash, so they are not reproduced across processes (Appendix C). Solver outputs are reproducible up to LP tie-breaking, which is fixed within a run of the released code. All experimental details, the complete hyperparameter table, and the reproducibility protocol are given in Appendix C (Reproducibility Statement). An anonymous supplementary package provides compiled implementations, minimal runners, and configurations for the main experiments; its README specifies the supported environment and experiment coverage.

REFERENCES

Rishabh Agarwal, Max Schwarzer, Pablo Samuel Castro, Aaron C Courville, and Marc Bellemare. Deep reinforcement learning at the edge of the statistical precipice. *Advances in neural information processing systems*, 34:29304–29320, 2021.

David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. In *Soda*, volume 7, pp. 1027–1035, 2007.

Ariyan Bighashdel, Yongzhao Wang, Stephen McAleer, Rahul Savani, and Frans A Oliehoek. Policy space response oracles: a survey. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 7951–7961, 2024.

Branislav Bosansky, Christopher Kiekintveld, Viliam Lisy, and Michal Pechoucek. An exact double-oracle algorithm for zero-sum extensive-form games with imperfect information. *Journal of Artificial Intelligence Research*, 51:829–866, 2014.

Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Weiqiang Zheng. Last-iterate convergence properties of regret-matching algorithms in games. In *The Thirteenth International Conference on Learning Representations*, 2025.

Thomas M Cover, Joy A Thomas, and John Kieffer. *Elements of information theory*, volume 2. wiley New York, 1991.

Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2003.

Gabriele Farina and Tuomas Sandholm. Fast payoff matrix sparsification techniques for structured extensive-form games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 4999–5007, 2022.

Andrew Gilpin and Tuomas Sandholm. Better automated abstraction techniques for imperfect information games, with application to texas hold'em poker. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pp. 1–8, 2007.

Samid Hoda, Andrew Gilpin, Javier Pena, and Tuomas Sandholm. Smoothing techniques for computing nash equilibria of sequential games. *Mathematics of Operations Research*, 35(2):494–512, 2010.

Alan J Hoffman. On approximate solutions of systems of linear inequalities. In *Selected Papers Of Alan J Hoffman: With Commentary*, pp. 174–176. World Scientific, 2003.

QHJAJ Huangfu and JA Julian Hall. Parallelizing the dual revised simplex method. *Mathematical programming computation*, 10(1):119–142, 2018.

Elon Kohlberg and Jean-Francois Mertens. On the strategic stability of equilibria. *Econometrica: Journal of the Econometric Society*, pp. 1003–1037, 1986.

Daphne Koller, Nimrod Megiddo, and Bernhard Von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and economic behavior*, 14(2):247–259, 1996.

Christine Konicki, Mithun Chakraborty, and Michael P Wellman. Exploiting extensive-form structure in empirical game-theoretic analysis. In *International Conference on Web and Internet Economics*, pp. 132–149. Springer, 2022.

Christian Kroer and Tuomas Sandholm. Extensive-form game abstraction with bounds. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 621–638, 2014.

Christian Kroer and Tuomas Sandholm. Imperfect-recall abstractions with bounds in games. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 459–476, 2016.

Christian Kroer and Tuomas Sandholm. A unified framework for extensive-form game abstraction with bounds. *Advances in Neural Information Processing Systems*, 31, 2018.

Harold W. Kuhn. A simplified two-person poker. In Harold W. Kuhn and Albert W. Tucker (eds.), *Contributions to the Theory of Games, Volume I*, volume 24 of *Annals of Mathematics Studies*, pp. 97–103. Princeton University Press, 1950.

Marc Lanctot, Vinícius Flores Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. A unified game-theoretic approach to multiagent reinforcement learning. In *NIPS*, pp. 4190–4203, 2017.

Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, Vinicius Zambaldi, Satyaki Upadhyay, Julien Pérolat, Sriram Srinivasan, Finbarr Timbers, Karl Tuyls, Shayegan Omidshafiei, Daniel Hennes, Dustin Morrill, Paul Muller, Timo Ewalds, Ryan Faulkner, János Kramár, Bart De Vylder, Brennan Saeta, James Bradbury, David Ding, Sebastian Borgeaud, Matthew Lai, Julian Schrittwieser, Thomas Anthony, Edward Hughes, Ivo Danihelka, and Jonah Ryan-Davis. Openspiel: A framework for reinforcement learning in games. *arXiv preprint arXiv:1908.09453*, 2019.

Boning Li and Longbo Huang. Efficient online pruning and abstraction for imperfect information extensive-form games. In *The Thirteenth International Conference on Learning Representations*, 2025.

Boning Li and Longbo Huang. Agents that certify their own exploits: Confidence-scheduled restricted responses for safe opponent exploitation. *arXiv preprint arXiv:2607.28520*, 2026a.

Boning Li and Longbo Huang. Effective, efficient, and general information abstraction for imperfect-information extensive-form games. *arXiv preprint arXiv:2605.10900*, 2026b.

Boning Li, Zhixuan Fang, and Longbo Huang. Rl-cfr: improving action abstraction for imperfect information extensive-form games with reinforcement learning. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 27752–27770, 2024.

Richard J Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pp. 36–41, 2003.

Stephen McAleer, John B Lanier, Kevin A Wang, Pierre Baldi, and Roy Fox. Xdo: A double oracle algorithm for extensive-form games. *Advances in Neural Information Processing Systems*, 34: 23128–23139, 2021.

Stephen McAleer, Kevin Wang, John Lanier, Marc Lanctot, Pierre Baldi, Tuomas Sandholm, and Roy Fox. Anytime psro for two-player zero-sum games. *arXiv preprint arXiv:2201.07700*, 2022.

Stephen Marcus McAleer, JB Lanier, Kevin A Wang, Pierre Baldi, Tuomas Sandholm, and Roy Fox. Toward optimal policy population growth in two-player zero-sum games. In *The Twelfth International Conference on Learning Representations*, 2024.

Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

H Brendan McMahan, Geoffrey J Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pp. 536–543, 2003.

Linjian Meng, Tianpei Yang, Youzhi Zhang, Zhenxing Ge, Shangdong Yang, Tianyu Ding, Wenbin Li, Bo An, and Yang Gao. Efficient last-iterate convergence in solving extensive-form games. In *Advances in Neural Information Processing Systems*, 2025a. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/918c67cd7910578f1637beb56717f455-Abstract-Conference.html.

Linjian Meng, Youzhi Zhang, Zhenxing Ge, Tianyu Ding, Shangdong Yang, Zheng Xu, Wenbin Li, and Yang Gao. Last-iterate convergence of smooth regret matching$^+$ variants in learning nash equilibria. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025b. URL https://openreview.net/forum?id=v771lscnlS.

Leon Mirsky. Symmetric gauge functions and unitarily invariant norms. *The quarterly journal of mathematics*, 11(1):50–59, 1960.

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge university press, 2007.

Tuomas Sandholm and Satinder Singh. Lossy stochastic game abstraction with bounds. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 880–897, 2012.

Martin Schmid, Matej Moravcik, and Milan Hladik. Bounding the support size in extensive form games with imperfect information. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

Finnegan Southey, Michael Bowling, Bryce Larson, Carmelo Piccione, Neil Burch, Darse Billings, and Chris Rayner. Bayes' bluff: opponent modelling in poker. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, pp. 550–558, 2005.

Xiaohang Tang, Le Cong Dinh, Stephen Marcus McAleer, and Yaodong Yang. Regret-minimizing double oracle for extensive-form games. In *International Conference on Machine Learning*, pp. 33599–33615. PMLR, 2023.

Eric Van Damme. *Stability and perfection of Nash equilibria*, volume 339. Springer, 1991.

Bernhard von Stengel. Equilibrium computation for two-player games in strategic and extensive form. In *Algorithmic Game Theory*, pp. 53–78. Cambridge University Press, 2007.

Kevin Waugh, David Schnizlein, Michael H Bowling, and Duane Szafron. Abstraction pathologies in extensive games. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, volume 2, pp. 781–788, 2009.

Michael P Wellman. Methods for empirical game-theoretic analysis. In *AAAI*, pp. 1552–1556, 2006.

Brian Zhang and Tuomas Sandholm. Sparsified linear programming for zero-sum equilibrium finding. In *International Conference on Machine Learning*, pp. 11256–11267. PMLR, 2020a.

Brian Hu Zhang and Tuomas Sandholm. Small nash equilibrium certificates in very large games. In *NeurIPS*, 2020b. URL https://proceedings.neurips.cc/paper/2020/hash/4fbe073f17f161810fdf3dab1307b30f-Abstract.html.

Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. *Advances in neural information processing systems*, 20, 2007.

A PROOFS

A.1 IMPOSSIBILITY AND THE SELECTION LOWER BOUND

Let

$$J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad M_b = J + bA, \quad 0 < b < 1.$$

The row and column spaces of J and A are orthogonal. The nonzero singular values of J and bA are 2 and $2b$, respectively, so $\mathrm{Blind}_1(M_b) = J$ and $\|M_b - J\|_2 = 2b \to 0$.

Adding the constant matrix J shifts every payoff by one and leaves best responses unchanged, a standard payoff-invariance property (von Stengel, 2007). Thus M_b has the same unique equilibrium as matching pennies, $x^\star = y^\star = (1/2, 1/2)$. By contrast, every pair of mixed strategies is an equilibrium of J, because every entry is equal. Under the Euclidean norm on the concatenated pair (x, y), the pure pair $(e_1, e_1) \in \mathrm{NE}(J)$ is at distance

$$\sqrt{\|e_1 - x^\star\|_2^2 + \|e_1 - y^\star\|_2^2} = 1$$

from $\mathrm{NE}(M_b)$. Hence $d_H(\mathrm{NE}(M_b), \mathrm{NE}(J)) = 1$ for every $b \in (0, 1)$, while the spectral error tends to zero. The discarded direction has positive singular value $2b$, so the failure holds with a genuinely nonzero discarded direction and rules out any modulus ω.

A.2 PROOFS OF LEMMA 4.2, THEOREM 4.3, AND THE SELECTION IDENTITY

Throughout, dist is with respect to the product norm $\|z - z'\| = \|x - x'\|_1 + \|y - y'\|_1$, the local error bound constant $(\kappa_\Delta, r_\Delta)$ is taken in this norm, and $B = \|M\|_{\max}$.

The flat carrier face carries the whole equilibrium set. Every pure pair in $S_x \times S_y$ has payoff v; a row deviation $i \notin S_x$ earns at most $v - \gamma$ against any face column strategy and a column deviation $j \notin S_y$ concedes at least $v + \gamma$ against any face row strategy, so every pair in $\Delta(S_x) \times \Delta(S_y)$ is an equilibrium with value v. Conversely the pairing argument of Theorem 4.5 (with the entrywise margin in place of condition (ii)) shows every equilibrium is face-supported, so $\mathrm{NE}(M) = \Delta(S_x) \times \Delta(S_y)$.

Proof of Lemma 4.2. Let $(\hat{x}, \hat{y}) \in \mathrm{NE}(W')$ with value $\hat{v}$, and let $\alpha = \sum_{i \notin S_x} \hat{x}_i$, $\beta = \sum_{j \notin S_y} \hat{y}_j$. *(i) Value.* Any $x_F \in \Delta(S_x)$ earns $x_F^\top M y \geq (1-\beta')v + \beta'(v+\gamma) \geq v$ against any y with outside mass β', so $x_F^\top W' y \geq v - \rho$ and $\hat{v} \geq v - \rho$; symmetrically $\hat{v} \leq v + \rho$. *(ii) Outside mass is small.* Playing x_F against $\hat{y}$ earns at least $v + \beta\gamma - \rho$ in W', and $\hat{y}$ is a minimax strategy, so $v + \beta\gamma - \rho \leq \hat{v} \leq v + \rho$, giving $\beta \leq 2\rho/\gamma$; symmetrically $\alpha \leq 2\rho/\gamma$. *(iii) Outside mass is zero.* Suppose $\hat{x}_i > 0$ for some $i \notin S_x$. Complementarity gives $(W'\hat{y})_i = \hat{v} \geq v - \rho$, while $(M\hat{y})_i \leq (1-\beta)(v-\gamma) + \beta B$, so

$$v - \rho \leq (1-\beta)(v-\gamma) + \beta B + \rho \quad \Longrightarrow \quad \gamma \leq 2\rho + \beta(B - v + \gamma) \leq \frac{2\rho}{\gamma}\big(2\gamma + B - v\big) \leq \frac{4\rho(\gamma + B)}{\gamma},$$

using $\beta \leq 2\rho/\gamma$ and $|v| \leq B$; this contradicts $4\rho(\gamma + B) < \gamma^2$. Hence $\alpha = 0$, and the symmetric computation with $(\hat{x}^\top M)_j \geq (1-\alpha)(v+\gamma) - \alpha B$ gives $\beta = 0$. □

The selection identity. Let $z = (x, y)$ be face-supported and $0 \leq \epsilon < \gamma/2$. For $i \in S_x$, $((M + \epsilon\Delta)y)_i = v + \epsilon(\Delta_F y_F)_i \geq v - \epsilon$; for $i \notin S_x$, $((M + \epsilon\Delta)y)_i \leq v - \gamma + \epsilon < v - \epsilon$. So the row maximum is attained on S_x and equals $v + \epsilon \max_{i \in S_x}(\Delta_F y_F)_i$; the column side is symmetric, and subtracting cancels v:

$$\mathrm{Gap}_{M+\epsilon\Delta}(z) = \epsilon \ \mathrm{Gap}_\Delta^{\mathrm{sel}}(z). \tag{1}$$

Because z ranges over the full product of carrier simplices, $\mathrm{Gap}_\Delta^{\mathrm{sel}}$ is exactly the duality gap of the selection game Δ_F, and its pure-action form $\max_{i \in S_x}(\Delta_F y_F)_i - \min_{j \in S_y}(x_F^\top \Delta_F)_j$ coincides with the mixed-deviation form by linearity.

Proof of Theorem 4.3. The condition $4(\epsilon + \eta)(\gamma + B) < \gamma^2$ implies $\epsilon < \gamma/4$, hence $2\epsilon < \gamma$. Since $\|M + \epsilon\Delta - M\|_{\max} \leq \epsilon$, Lemma 4.2 with $\rho = \epsilon$ confines $\mathrm{NE}(M + \epsilon\Delta)$ to the face, where by (1) equilibria are exactly the saddle points of Δ_F; uniqueness gives $\mathrm{NE}(M + \epsilon\Delta) = \{z_\Delta\}$. Now take $\hat{z} \in \mathrm{NE}(\widehat{W})$. Since $\|\widehat{W} - M\|_{\max} \leq \epsilon + \eta$, the lemma with $\rho = \epsilon + \eta$ makes $\hat{z}$ face-supported.

Lemma 4.6 gives $\mathrm{Gap}_{M+\epsilon\Delta}(\hat{z}) \le \mathrm{Gap}_{\widehat{W}}(\hat{z}) + 2\eta = 2\eta$, so (1) yields $\mathrm{Gap}^{\mathrm{sel}}_{\Delta}(\hat{z}) \le 2\eta/\epsilon \le r_\Delta$, and the error bound of the selection game gives $\mathrm{dist}(\hat{z}, z_\Delta) \le 2\kappa_\Delta\eta/\epsilon$. The alternative bound 4 is the diameter of the product of two simplices in the product ℓ_1 norm. Both bounds use $\|\Delta\|_{\max} \le 1$: the confinement steps apply Lemma 4.2 with residuals $\rho = \epsilon$ and $\rho = \epsilon + \eta$, which require $\|M + \epsilon\Delta - M\|_{\max} \le \epsilon$. □

Proof of Proposition 4.4. Take

$$M = \begin{bmatrix} 0 & 0 & \gamma \\ 0 & 0 & \gamma \\ -\gamma & -\gamma & 0 \end{bmatrix}, \qquad \Delta = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \widehat{W} = M + \epsilon\Delta + \eta B_0, \quad B_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$\mathrm{NE}(M)$ is the face $\Delta(\{1,2\})^2$: the third row loses γ against every face column strategy and the third column loses γ symmetrically, so the margin is γ and the carrier is $\{1,2\}^2$.

In $M + \epsilon\Delta$, $x_F = y_F = (1/2, 1/2, 0)$ is a saddle pair with value 0: the first two rows each earn $\pm(\epsilon/2 - \epsilon/2) = 0$ and the third row earns $-\gamma$ against y_F, while the third column costs γ against x_F and the first two columns cost 0. No equilibrium places mass on the third actions (a row strategy with outside mass a earns below the value against y_F, and the column side is symmetric), so the equilibria of $M + \epsilon\Delta$ are exactly those of the face game $\epsilon\Delta_F$ with $\Delta_F = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$: matching pennies, with unique equilibrium $x^\star(\epsilon) = y^\star(\epsilon) = (1/2, 1/2, 0)$.

In $\widehat{W}$, the same saddle-pair argument (using $\epsilon + \eta \le 3\gamma/4 < \gamma$) confines equilibria to the face game

$$\epsilon\Delta_F + \eta B_F = \begin{bmatrix} \epsilon + \eta & -\epsilon \\ -\epsilon & \epsilon \end{bmatrix},$$

whose value is $\widehat{v} = \epsilon\eta/(4\epsilon + \eta) < \epsilon < \gamma$. Its unique equilibrium $\hat{x} \propto (2\epsilon,\, 2\epsilon + \eta)$ and $\hat{y} \propto (2\epsilon,\, 2\epsilon + \eta)$ earn $\widehat{v}$ on both face rows and columns, while the third row earns $-\gamma < \widehat{v}$ against any face column strategy and the third column costs $\gamma > \widehat{v}$ against any face row strategy; pairing any candidate with this face equilibrium excludes both, so no equilibrium places mass on them. Therefore

$$\|\hat{x} - x^\star(\epsilon)\|_\infty = \frac{1}{2} - \frac{2\epsilon}{4\epsilon + \eta} = \frac{\eta}{2(4\epsilon + \eta)},$$

and $\|M + \epsilon\Delta - \widehat{W}\|_{\max} = \eta$. For $\eta \le 4\epsilon$ the per-player ℓ_∞ displacement is at least $\eta/(16\epsilon)$; as $\eta/\epsilon \to \infty$ it tends to $1/2$. Both players move by this amount, so the product-ℓ_1 distance of the statement equals 4 times the per-player ℓ_∞ displacement, i.e. $2\eta/(4\epsilon + \eta)$: each player contributes ℓ_1 mass twice its ℓ_∞ displacement. The experiment of Section B.5 reproduces this displacement to LP accuracy in all 20 settings.

A.3 Deployment certificates

The full-game gap $\epsilon_{\mathrm{full}}(z) = \max_i (My)_i - \min_j (x^\top M)_j = \mathrm{Gap}_M(z)$ costs one matrix-vector product per side and requires no support oracle, margin, non-degeneracy constant, or temperature. Compression supplies the candidate; the certificate is evaluated on the original game.

A.4 Proofs of Propositions 5.5 and 5.6

Let z_S be the unique equilibrium of $M[S_0, S_1]$ and let its embedding be an equilibrium of M. The exact restricted solve returns z_S with $g_S = 0 \le \epsilon_{\mathrm{solve}}$. Full-game optimality of z_S gives $\max_i (My)_i = v(M)$; equilibrium complementarity makes every action of $\mathrm{supp}(x_S) \subseteq S_0$ attain this maximum, so $\max_{i \in S_0} (My)_i = v(M)$ and $r_x = 0 \le \epsilon_{\mathrm{cert}}$; symmetrically $r_y = 0$. The first certificate check passes.

Flat carriers need no uniqueness. In the flat-carrier model of Theorem 4.3, Proposition 5.5 extends without uniqueness, under a support condition alone: if M has flat carrier $S^\star_x \times S^\star_y$ with entrywise margin γ and the anchor meets it on both sides, $S_0 \cap S^\star_x \neq \emptyset$ and $S_1 \cap S^\star_y \neq \emptyset$, then every exact equilibrium of $M[S_0, S_1]$ passes the first certificate. The restricted game is itself a flat-carrier game with carrier $(S_0 \cap S^\star_x) \times (S_1 \cap S^\star_y)$ and the same margin, so Lemma 4.2 with $\rho = 0$ confines its equilibria to that carrier; the embedding lies in $\Delta(S^\star_x) \times \Delta(S^\star_y) = \mathrm{NE}(M)$, and the computation above gives $g_S = r_x = r_y = 0$. An anchor disjoint from the carrier on either side has no such

bound: the first restricted solve is then governed by the corner block, and Algorithm 1 may need several expansions.

Proof of Proposition 5.6. If a round does not return, then $r_x > \epsilon_{\mathrm{cert}}$ or $r_y > \epsilon_{\mathrm{cert}}$; a best response already contained in the current pool contributes a residual of zero, so the added policy is new and the pool grows strictly, which can happen at most $|\mathcal{P}_0| + |\mathcal{P}_1|$ times. Because the payoff of every existing pair is unchanged under pool growth, each round's restricted game is the restriction of the same fixed u, and on return Theorem 5.3 bounds the gap of the output mixture against pure deviations in $\mathcal{P}_0 \cup \mathcal{P}_1$; bilinearity extends the bound to mixed deviations over $\Delta(\mathcal{P}_0) \times \Delta(\mathcal{P}_1)$.

A.5 Proof of Theorem 4.5

We use zero-sum equilibrium interchangeability and complementary slackness (Nisan et al., 2007; von Stengel, 2007). First fix M alone. Every $(x_F, y_F) \in Z_F$ is a restricted equilibrium: the uniform margin makes any $i \notin S_x$ and any $j \notin S_y$ strictly suboptimal against it, so (x_F, y_F) is a saddle point of the full game and $Z_F \subseteq \mathrm{NE}(M)$ with $v(M) = v_F$. Conversely, take $(x, y) \in \mathrm{NE}(M)$ and pair it with any $(x_F, y_F) \in Z_F$. Since y_F is a minimax strategy of the full game and x is a maximin strategy, $My_F \le v_F \mathbf{1}$ and $x^\top M y_F \ge v_F$; hence $x^\top M y_F = v_F$. But condition (ii) gives $(My_F)_i \le v_F - \gamma$ for every $i \notin S_x$, so positive mass outside S_x would make the inner product strictly less than v_F, a contradiction. Thus $x \in \Delta(S_x)$; symmetrically $y \in \Delta(S_y)$, and $(x, y) \in Z_F$. Hence $\mathrm{NE}(M) = Z_F$.

Face exactness makes the two restricted games identical, so $\widehat{M}$ shares the value v_F and the restricted equilibrium set Z_F. For $(x_F, y_F) \in Z_F$, condition (iii) gives, for every $i \notin S_x$,

$$(\widehat{M} y_F)_i \le (My_F)_i + \eta_F \le v_F - (\gamma - \eta_F) < v_F,$$

and inside rows keep their restricted payoff; the column side is symmetric. Hence $Z_F \subseteq \mathrm{NE}(\widehat{M})$. Repeating the pairing argument with γ replaced by $\gamma - \eta_F$ shows $\mathrm{NE}(\widehat{M}) \subseteq Z_F$.

For the approximate direction, take $z \in \mathrm{NE}_\epsilon(M)$. Its player-wise errors sum to at most 2ϵ, so $\mathrm{Gap}_M(z) \le 2\epsilon \le 2(\epsilon+\eta) \le r$, and condition (iv) yields $\mathrm{dist}(z, Z_F) \le 2\kappa\epsilon$. Since $Z_F \subseteq \mathrm{NE}_\epsilon(\widehat{M})$, this bounds $\mathrm{dist}(z, \mathrm{NE}_\epsilon(\widehat{M}))$. Conversely, for $z \in \mathrm{NE}_\epsilon(\widehat{M})$, Lemma 4.6 gives

$$\mathrm{Gap}_M(z) \le \mathrm{Gap}_{\widehat{M}}(z) + 2\eta \le 2(\epsilon + \eta) \le r,$$

and condition (iv) applied to M yields $\mathrm{dist}(z, Z_F) \le 2\kappa(\epsilon+\eta)$. Because $Z_F \subseteq \mathrm{NE}_\epsilon(M)$, the same quantity bounds $\mathrm{dist}(z, \mathrm{NE}_\epsilon(M))$. Taking the maximum over both directions gives the stated bound in the product ℓ_1 distance; since the Euclidean distance of Section 3 never exceeds it ($\|z - z'\|_2 \le \|z - z'\|_1$ componentwise), the Euclidean Hausdorff bound is the same number.

Proof of Lemma 4.6. Both the best row payoff and the best column deviation change by at most η. The difference of the two bounds gives the result.

Polyhedral interpretation of the local error bound. For a fixed zero-sum matrix game, the exact player-wise solution sets are polyhedra:

$$X_0(M) = \{x \in \Delta_m : M^\top x \ge v(M)\mathbf{1}\}, \qquad Y_0(M) = \{y \in \Delta_n : My \le v(M)\mathbf{1}\}.$$

Hoffman error bounds for finite systems of linear equalities and inequalities (Hoffman, 2003) imply that distance to each set is bounded by a constant times the norm of its constraint violations. Combining the two player-wise systems gives Definition 3.1. The constant depends on the active constraint geometry and the selected norm. Two quantities are the natural diagnostics: the smallest singular value σ of the face block on the sum-zero subspace and the smallest equilibrium probability p_0 (both defined in Section A.6); a vanishing σ or p_0 signals a poorly conditioned active system. Proposition A.1 turns them into a closed-form constant when the face equilibrium is unique and completely mixed and the outside margin holds entrywise; when the face carries a non-singleton equilibrium set, as in Leduc, κ remains an explicit conditioning assumption.

Conditions (i)–(iii) in Theorem 4.5 identify the same exact equilibrium set Z_F for M and $\widehat{M}$. Face exactness gives a common restricted game, and condition (ii) uniformly protects all its equilibria from outside deviations; condition (iii) preserves the separation for $\widehat{M}$, because outside actions are

evaluated only against face strategies and are therefore exposed to the face-direction residual η_F. If a row strategy places mass a outside S_x, pairing it with any restricted equilibrium column strategy lowers its payoff by at least $a\gamma$ in M and $a(\gamma - \eta_F)$ in $\widehat{M}$; an equilibrium requires $a = 0$. The column argument is symmetric. This explains why exact-set identification uses the face-direction threshold $\eta_F < \gamma$; the full residual η enters only through gap propagation in the approximate-set bound. The experimental envelope η_{face} of Section 5 is a different, operator-level quantity that dominates η_F: for $j \in S_y$, $\|(M - \widehat{M})e_j\|_\infty = \|(I - Q_x Q_x^\top)Me_j\|_\infty \le \|(I - Q_x Q_x^\top)M\|_{2\to\infty}$ because $Q_y Q_y^\top e_j = e_j$, and the column side is symmetric, so $\eta_F \le \eta_{\text{face}}$. The experimental diagnostic $\eta_{\text{face}} < \hat{\gamma}/2$ measures $\hat{\gamma}$ at one solved equilibrium, and $\hat{\gamma} \ge \gamma$ because γ is the infimum over Z_F; the diagnostic therefore certifies condition (iii) when Z_F is a single point ($\hat{\gamma} = \gamma$), and is a heuristic when the face carries several equilibria, where a margin measured at one of them can exceed the uniform margin by an arbitrary factor (a 3×2 example: face row 1 with zero payoffs against both columns, outside rows $(-0.1, -10)$ and $(-0.5, -10.4)$, so $\gamma = 0.1$ while $\hat{\gamma} = 10$ at the solved equilibrium $y = e_2$). Throughout, the constant κ of Definition 3.1 is taken in the product-ℓ_1 metric, and the Euclidean Hausdorff bound of Theorem 4.5 follows from $\|z - z'\|_2 \le \|z - z'\|_1$; the lower-bound probes of Section B measure Euclidean distances, which are lower bounds for the product-ℓ_1 constant as well.

One-sided stability for arbitrary anchors. Theorem 4.5 is two-sided: its set-level equality needs the candidate face to protect the equilibria of M (condition (ii)), which the non-oracle seeds of the Leduc study are not shown to satisfy: their recall of 0.94–0.98 (Table 5) is measured against one oracle corner of the equilibrium set and does not verify coverage of the carrier. The one-sided direction needs no anchor assumption, because exploitability propagates independently of the face structure: for any orthogonal projection $\widehat{M} = Q_x Q_x^\top M Q_y Q_y^\top$ and any pair z, Lemma 4.6 gives $\mathrm{Gap}_M(z) \le \mathrm{Gap}_{\widehat{M}}(z) + 2\eta$. Every $z \in \mathrm{NE}_\epsilon(\widehat{M})$ has player-wise residuals at most ϵ, hence $\mathrm{Gap}_{\widehat{M}}(z) \le 2\epsilon$ and $\mathrm{Gap}_M(z) \le 2(\epsilon + \eta)$. If M satisfies Definition 3.1 around its full equilibrium set $\mathrm{NE}(M)$ with radius $r \ge 2(\epsilon + \eta)$, then

$$\mathrm{dist}\big(z, \mathrm{NE}(M)\big) \le 2\kappa(\epsilon + \eta) \qquad \text{for every } z \in \mathrm{NE}_\epsilon(\widehat{M}).$$

Compression therefore cannot manufacture an ϵ-equilibrium of $\widehat{M}$ whose exploitability in M exceeds $2(\epsilon + \eta)$, whatever the anchor. This bound is all that holds without an anchor assumption. It does not say that an anchor which omits carrier actions removes the corresponding equilibria from $\mathrm{NE}(\widehat{M})$ (for $M = 0$ every projection returns $\widehat{M} = 0$ and $\mathrm{NE}(\widehat{M}) = \mathrm{NE}(M)$ for any anchor), and the certificate of Theorem 5.3 does not detect missing set coverage: it tests one returned pair for profitable deviations in M. If that pair is already an equilibrium of M, Algorithm 1 stops, whether or not the working set has reached the carrier; on the flat-carrier family of Proposition 4.4, the seed $S_0 = S_1 = \{1\}$ returns the pure pair $(1, 1)$, which is a full-game equilibrium, and the loop terminates without ever adding action 2. Expansion repairs a failed point certificate; it does not repair a missing face. Carrier recovery is neither guaranteed nor needed for the output guarantee.

A.6 An explicit local constant

Write $M_F = M[S_x, S_y]$ with $|S_x| = p$, $|S_y| = q$, and $p, q \ge 2$ throughout the proposition (a pure strict saddle point, $p = q = 1$, has σ and c undefined and is treated separately after the proof), and let $\mathbf{1}^\perp$ denote the sum-zero subspace of the relevant dimension. Set

$$\sigma \;=\; \min\Big\{ \min_{u \in \mathbf{1}^\perp,\ \|u\|_2 = 1} \|M_F u\|_2,\ \min_{w \in \mathbf{1}^\perp,\ \|w\|_2 = 1} \|M_F^\top w\|_2 \Big\}.$$

Proposition A.1 (Explicit constant for a unique completely mixed face equilibrium)**.** *Suppose the face game M_F has a unique equilibrium $z^\star = (x^\star, y^\star)$ with every coordinate positive, value v_F, and $p_0 = \min\{\min_i x_i^\star, \min_j y_j^\star\}$, and that the outside margin holds entrywise with $\gamma > 0$: $M_{ij} \le v_F - \gamma$ for $i \notin S_x$, $j \in S_y$, and $M_{ij} \ge v_F + \gamma$ for $i \in S_x$, $j \notin S_y$, with $M[S_x^c, S_y^c]$ arbitrary. Then $\sigma > 0$, $\mathrm{NE}(M) = \{z^\star\}$, and for every $z \in \Delta_m \times \Delta_n$,*

$$\mathrm{Gap}_M(z) \;\ge\; \min\Big\{c, \frac{\gamma}{2}\Big\} \|z - z^\star\|_1, \qquad c = \frac{p_0\,\sigma}{(1 - p_0)\sqrt{pq}},$$

so M satisfies Definition 3.1 around $Z_F = \{z^\star\}$ with $\kappa = \max\{1/c,\, 2/\gamma\}$ and every radius r. Without outside actions ($m = p$, $n = q$) the bound reads $\mathrm{Gap}_{M_F}(z) \ge c\,\|z - z^\star\|_1$.

Proof. Step 1 ($\sigma > 0$). If $M_F u = 0$ for some $u \in \mathbf{1}^\perp \setminus \{0\}$, then $y^\star + tu \in \Delta(S_y)$ for small $|t|$ because $y^\star$ is interior, and $M_F(y^\star + tu) = M_F y^\star = v_F \mathbf{1}$, so $(x^\star, y^\star + tu)$ is a second equilibrium of M_F; the column side is symmetric. Uniqueness therefore forces $\sigma > 0$.

Step 2 (weighted zero-sum vectors). Since $y^\star$ is completely mixed, complementarity makes every face column a best response to $x^\star$: $x^{\star\top} M_F = v_F \mathbf{1}^\top$; symmetrically $M_F y^\star = v_F \mathbf{1}$. For any $\tilde{y} \in \Delta(S_y)$ the vector $w = M_F \tilde{y} - v_F \mathbf{1} = M_F(\tilde{y} - y^\star)$ satisfies $\sum_i x_i^\star w_i = 0$. If $\|w\|_\infty = -w_{i_0}$ is attained at a negative coordinate, then $0 = \sum_i x_i^\star w_i \le -x_{i_0}^\star \|w\|_\infty + (1 - x_{i_0}^\star) \max_i w_i$, so $\max_i w_i \ge \frac{x_{i_0}^\star}{1 - x_{i_0}^\star} \|w\|_\infty \ge \frac{p_0}{1 - p_0} \|w\|_\infty$; if it is attained at a positive coordinate the same inequality holds because $p_0 \le 1/2$. With $\tilde{y} - y^\star \in \mathbf{1}^\perp$,

$$\max_{i \in S_x} w_i \;\ge\; \frac{p_0}{1 - p_0} \cdot \frac{\|w\|_2}{\sqrt{p}} \;\ge\; \frac{p_0\, \sigma}{(1 - p_0)\sqrt{p}} \|\tilde{y} - y^\star\|_2 \;\ge\; c\, \|\tilde{y} - y^\star\|_1.$$

Step 3 (outside mass). Fix $y \in \Delta_n$ with outside mass $b = \sum_{j \notin S_y} y_j$ and write $y = (1 - b)\tilde{y} + b\, y_{\mathrm{out}}$ with $\tilde{y} \in \Delta(S_y)$, $y_{\mathrm{out}} \in \Delta(S_y^c)$. For $i \in S_x$ the entrywise margin gives $(My)_i \ge (1 - b)(v_F + w_i) + b(v_F + \gamma)$, hence

$$\max_i (My)_i - v_F \;\ge\; (1 - b) \max_{i \in S_x} w_i + b\gamma \;\ge\; c(1 - b)\|\tilde{y} - y^\star\|_1 + \frac{\gamma}{2} \cdot 2b \;\ge\; \min\Big\{c, \frac{\gamma}{2}\Big\} \|y - y^\star\|_1,$$

using $\|y - y^\star\|_1 \le (1 - b)\|\tilde{y} - y^\star\|_1 + 2b$. The same argument on the column side gives $v_F - \min_j (x^\top M)_j \ge \min\{c, \gamma/2\} \|x - x^\star\|_1$. Both left-hand sides are non-negative (the first is at least $x^{\star\top} M y - v_F \ge b\gamma \ge 0$), and their sum is $\mathrm{Gap}_M(z)$, which proves the bound; $\mathrm{Gap}_M(z) = 0$ then forces $z = z^\star$, so $\mathrm{NE}(M) = \{z^\star\}$. Without outside actions Step 3 is void and Step 2 alone gives the constant c. □

Pure saddle points. If $p = q = 1$ with outside actions and the entrywise margin, then $\|y - y^\star\|_1 = 2b$ and Step 3 alone gives $\mathrm{Gap}_M(z) \ge (\gamma/2)\|z - z^\star\|_1$, so $\kappa = 2/\gamma$ with $r = \infty$; without outside actions the game is 1×1 and the distance is zero.

The constant is attained. On matching pennies, $\sigma = 2$, $p_0 = 1/2$, $p = q = 2$ give $c = 1$, which is the value $\kappa_\Delta = 1$ used in Proposition 4.4; the selection game of Theorem 4.3 therefore has $\kappa_\Delta \le (1 - p_0)\sqrt{pq}/(p_0 \sigma_\Delta)$ with $r_\Delta = \infty$ whenever its unique saddle point z_Δ is completely mixed, where σ_Δ and p_0 are computed from Δ_F and z_Δ. On the 6×6 game of Section B ($p = q = 3$, uniform $z^\star$, $\sigma = \sqrt{3}$ for the rock–paper–scissors core), $c = \sqrt{3}/6$ and $2/\gamma$ dominates for $\gamma < \sqrt{3}/3$; the face-outer family of that section has product-ℓ_1 ratio exactly $2/\gamma$, so the upper bound $\kappa = 2/\gamma$ matches the measured lower bound (its Euclidean value $\sqrt{4/3}/\gamma$ is the same family in the other norm). The entrywise margin cannot be weakened to the $x^\star$-averaged margin of condition (ii) in Theorem 4.5: with the matching-pennies face, $\gamma = 0.1$, and an outside column $(\gamma - 3, \gamma + 3)$ whose $x^\star$-average is exactly $v_F + \gamma$, the pair $z = (x^\star, 0.75\, e_1 + 0.25\, e_3)$ has $\mathrm{Gap}_M(z) = 0.025$ against $\min\{c, \gamma/2\}\|z - z^\star\|_1 = 0.05$. These identities and 4,000 random instances of the bound are checked in `exp15_explicit_kappa_sanity.py`.

A.7 Proofs of Propositions 5.1 and 5.2

For $i \in S_x$, the coordinate vector e_i lies in $\mathrm{range}(Q_x)$, so $Q_x Q_x^\top e_i = e_i$. Likewise, $Q_y Q_y^\top e_j = e_j$ for $j \in S_y$. Therefore

$$e_i^\top \widehat{M} e_j = (Q_x Q_x^\top e_i)^\top M (Q_y Q_y^\top e_j) = e_i^\top M e_j.$$

Rank cost of anchoring. Face exactness is bought at a controlled rank cost. Since $\widehat{M} = Q_x Q_x^\top M Q_y Q_y^\top$ factors through two orthogonal subspaces of dimensions $\dim Q_x \le \min(m, k + |S_x|)$ and $\dim Q_y \le \min(n, k + |S_y|)$, the anchored output satisfies $\mathrm{rank}(\widehat{M}) \le \min\{\dim Q_x, \dim Q_y\}$: anchoring adds at most $|S_x|$ (respectively $|S_y|$) anchor coordinates to the k-dimensional spectral subspace. In the Leduc diagnostic (Section B) the anchored factors span 56–85% of the 337-dimensional sequence space, so anchoring exchanges storage for face exactness and adds no spectral work.

Proof of Proposition 5.2. Since $\mathrm{supp}(y) \subseteq S_y$, the column projector fixes y, so $(\widehat{M}y)_i = e_i^\top Q_x Q_x^\top My$. For $i \in S_x$, $Q_x Q_x^\top e_i = e_i$, giving (i); (ii) is the same identity evaluated outside the face, and the column side is symmetric.

A.8 Certified-output and termination

Insert the restricted maximum and minimum into the definition of the full gap. The resulting three terms are exactly r_x, g_S, and r_y, so their respective bounds add.

Proof of Lemma 5.4. Suppose round t of Algorithm 1 does not return, so $r_x > \epsilon_{\mathrm{cert}}$ or $r_y > \epsilon_{\mathrm{cert}}$. If $r_x > \epsilon_{\mathrm{cert}}$, then $(My)_{i^\star} > \max_{i \in S_0}(My)_i$, which forces $i^\star \notin S_0$: any action inside S_0 satisfies $(My)_i \le \max_{i \in S_0}(My)_i$. Hence the assignment $S_0 \leftarrow S_0 \cup \{i^\star\}$ adds an action that was absent in every previous round, since the supports only grow. The case $r_y > \epsilon_{\mathrm{cert}}$ is symmetric. Each non-returning round therefore strictly increases $|S_0| + |S_1|$, which never exceeds $m + n$, so the loop returns after at most $m + n$ rounds.

The sequence-form implementation evaluates the same objects through best-response LPs: the restricted solve returns a saddle pair (x, y) of $M[S_0, S_1]$, and the global best responses are the primal realization plans maximizing $x'^\top My$ and minimizing $x^\top My'$ over the full sequence polytopes. If the full-game gap $\mathrm{Gap}_M(x, y) > 2\epsilon_{\mathrm{cert}}$, then at least one of the two best responses is strictly better than the restricted value, and a strictly improving realization plan cannot be supported entirely inside the current support (any such plan is feasible for the restricted solve and would already achieve the restricted optimum). The expansion step adds the support of that primal best response, so a new action enters and the same counting argument holds.

A.9 The regularized-point guarantee and the three levels of claim

Remark A.2 (Regularized-point stability). The entropy-regularized saddle point, anchored at fixed full-support strategies x^0, y^0, is the anchored quantal response equilibrium (McKelvey & Palfrey, 1995), unique for $\tau > 0$ by the strong convexity of entropy (Hoda et al., 2010). With natural-log entropy, comparing the two games at $\eta = \|M - \widehat{M}\|_{\max}$ gives $\|z_\tau(M) - z_\tau(\widehat{M})\|_1 \le 2\eta/\tau$ in the product ℓ_1 norm.

Three levels of claim appear in the main text. Theorem 4.5 controls an *equilibrium set* under face exactness, a uniform margin, and a local error bound. Remark A.2 controls one *regularized point* with sensitivity $1/\tau$, and we use it only to reason about the regularized support estimator of Section 6. Theorem 5.3 controls one *deployed candidate* by direct evaluation on M. Of the set-level conditions, face exactness and the margin are LP-checkable once a candidate face is supplied, while κ is instance-dependent and explicit only for a unique completely mixed face equilibrium (Proposition A.1). The margin-residual diagnostic of Section 5 and Appendix B.10 screens conditions (ii) and (iii) heuristically when the restricted equilibrium set is not a singleton; the deployed candidate is certified pointwise by best responses on the original game, and the diagnostic plays no part in that certificate.

A.10 Proof of Remark A.2

Let $z = z_\tau(M)$ and $\widehat{z} = z_\tau(\widehat{M})$ use the same full-support anchors, natural-log entropy, and temperature $\tau > 0$, and write $\eta = \|M - \widehat{M}\|_{\max}$. Write F_M and $F_{\widehat{M}}$ for their saddle-point variational-inequality operators, including the entropic gradients. The standard variational-inequality optimality conditions (Facchinei & Pang, 2003) give

$$\langle F_M(z), \widehat{z} - z\rangle \ge 0, \qquad \langle F_{\widehat{M}}(\widehat{z}), z - \widehat{z}\rangle \ge 0.$$

Adding the inequalities and separating the payoff perturbation yields

$$\langle F_M(z) - F_M(\widehat{z}), z - \widehat{z}\rangle \le \langle (F_{\widehat{M}} - F_M)(\widehat{z}), z - \widehat{z}\rangle.$$

Per-component Pinsker bounds bound each symmetric Kullback–Leibler divergence below by the square of that component's ℓ_1 distance (Cover et al., 1991), so the left side is at least $\tau(\|x - \widehat{x}\|_1^2 + \|y - \widehat{y}\|_1^2)$, and the AM–QM inequality gives $\frac{\tau}{2}\|z - \widehat{z}\|_1^2$. By the definition of η, the right side is at most $\eta\|z - \widehat{z}\|_1$. Cancelling the distance, unless it is zero, gives

$$\|z_\tau(M) - z_\tau(\widehat{M})\|_1 \le \frac{2\eta}{\tau}.$$

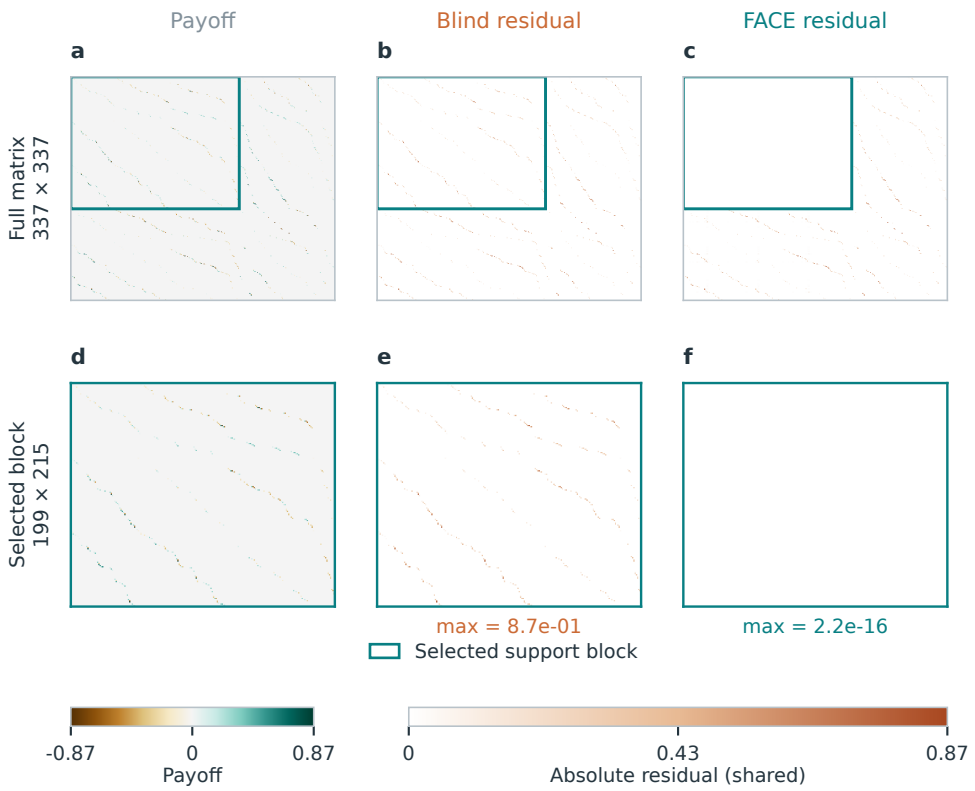


Figure 2: Support-block preservation in the perturbed Leduc sequence-form matrix (seed 1, $\epsilon = 0.1$, base rank 1). Columns show payoffs, blind-truncation residuals, and FACE residuals; rows show the full matrix and the selected 199×215 oracle support block. All full matrices use the same stable row and column permutations to place the selected indices first; teal frames mark the exact block extracted below. Payoffs share a zero-centered scale, and all residuals share one absolute scale. Residuals at or below 8.70×10^{-7} render white; the unthresholded block maxima are 0.870 for Blind and 2.22×10^{-16} for FACE. Anchoring augments the two basis dimensions to 200 and 216; the illustration establishes preservation of the selected block.

A.11 A 4×4 ILLUSTRATION OF GRADIENT SEPARATION

With the face $\{0, 1\} \times \{0, 1\}$ of a 4×4 game given by a matching-pennies core and generic outside entries

$$M = \begin{bmatrix} 1 & -1 & 0.4 & -0.3 \\ -1 & 1 & -0.2 & 0.6 \\ 0.5 & 0.3 & 0.1 & -0.5 \\ -0.4 & 0.2 & 0.7 & 0 \end{bmatrix},$$

the rank-1 FACE projection is face-exact to 2.2×10^{-16}, and the restricted equilibrium $x^\star = y^\star = (1/2, 1/2, 0, 0)$ solves both $M[S_x, S_y]$ and $\widehat{M}[S_x, S_y]$. Yet at $y^\star$ the outside row payoffs are $(My^\star)_2, (My^\star)_3 = (0.4, -0.1)$ versus $(\widehat{M}y^\star)_2, (\widehat{M}y^\star)_3 = (0.2748, -0.2421)$, and one projected-gradient step of size 0.1 from $x_0 = (0.45, 0.45, 0.05, 0.05)$, $y_0 = (0.4, 0.4, 0.1, 0.1)$ differs between the two matrices by 1.1×10^{-2} in sup norm. Compression changes where a numerical solver goes while leaving the exact restricted equilibrium fixed.

B Additional Experimental Protocol

B.1 The 3×3 support-family probe

A 3×3 probe family with rows $(0, 0, 1)$, $(c, 0, 1)$, $(b, b, 0)$ over $b \in \{0.01, 0.05, 0.1, 0.3\}$, $c \in \{0, 0.001, 0.01\}$, and ranks $k \in \{1, 2, 3\}$ separates point fidelity from set fidelity. Support families are enumerated in rational arithmetic, exactly for the original matrices and under separated tolerances for the floating-point compressed matrices (Section B). Blind rank one creates support patterns absent from the original game. Blind rank two still loses a member of the true support family at spectral error 5×10^{-4} and gap 1.2×10^{-5}, so preserving every equilibrium is strictly stronger than keeping one solution accurate. Given the true support as an oracle, FACE passes the tolerance-dependent support diagnostic at rank one (Table 3).

B.2 Logged settings and interquartile means

The synthetic sweep contains 72 logged rows, including blind ranks, face-anchored variants, and per-setting summaries. The Leduc sequence-form diagnostic contains 120 per-instance rows: six seeds, four perturbation levels $\epsilon \in \{10^{-3}, 10^{-2}, 10^{-1}, 3 \times 10^{-1}\}$, and five ranks, giving 24 instances per rank. The matrix-game diagnostic repeats this protocol on each of three 250×250 families (Section B.3), for 360 per-instance rows. Reported interquartile means sort the 24 values at each rank, remove the lowest and highest quarters, and average the remaining 12 (Agarwal et al., 2021). The margin-residual diagnostic contains 96 Leduc, 12 random full-rank, and 12 biased low-rank settings. The non-oracle support study contains five margin settings, four temperature settings, four controlled initialization scenarios, and 72 Leduc expansion rows (24 instances by three seed types); the matrix-game expansion repeats the three seed types on the same 24 instances of each family, adding 216 expansion rows. All rows are deterministic. The rank-k spectral errors of Table 1 and the initial-support precision/recall of Table 5 are recomputed from the same deterministic per-seed instances. Every seed fixes one deterministic perturbation instance: the Rademacher draws are produced by a fixed NumPy generator seeded by the integer seed, so the 24 instances are fully reproducible; the interquartile means summarize the spread over these instances, and determinism removes rerun noise; sampling variability across instances remains. Support families in the synthetic diagnostic are enumerated in rational arithmetic: every matrix entry is read as the dyadic rational it represents, the value $v(M)$ is computed exactly, and a nonempty support S is accepted for the row player when the linear program $\max\{\delta : x \in \Delta,\ x_i \geq \delta\ (i \in S),\ x_i = 0\ (i \notin S),\ M^\top x \geq (v(M) - \tau_v)\mathbf{1}\}$, solved by exact vertex enumeration, has optimal value $\delta^\star > 0$ and $\delta^\star \geq \tau_\delta$ (the column side is symmetric). For the original matrices $\tau_v = \tau_\delta = 0$, so the test is exact: $\delta^\star > 0$ is a proof and $\delta^\star = 0$ a refutation that S is an equilibrium support of the stored matrix. For the floating-point compressed matrices, whose entries carry machine-precision noise, exact arithmetic on the stored entries would refute every indifference that rounding has broken, so we use $\tau_v = 10^{-12}$ and $\tau_\delta = 10^{-6}$; this diagnostic depends on the tolerances and does not enumerate exactly, and the reference family of each original matrix is checked to be the same under both settings (asserted in the code). The relaxation can lend mass to a coordinate that every exact equilibrium forces to zero: maximizing that mass under the relaxed constraints in exact arithmetic over the eight such coordinates of the family gives at most 1.01×10^{-7} (at $b = 0.01$, $c = 0.001$), one decade below τ_δ; a bound of the form τ_v / c is too optimistic, since the mass scales with $\tau_v(1 + 1/b)/c$. This audit excludes revival of the original family's globally forbidden coordinate at the default thresholds; it does not certify exact support-family preservation for compressed matrices, whose rounded entries can break a different indifference. The diagnostic is therefore reported with its margin: moving τ_δ to 10^{-7} or τ_v to 10^{-11} flips the verdict of three of the 60 compressed candidates, all on the $b = 0.01$, $c = 0.001$ matrix, and no verdict changes at $\tau_\delta = 10^{-5}$ or $\tau_v = 10^{-13}$ (`exp17_support_tolerance_audit.py`). (A single-tolerance test that accepts mass 10^{-9} against a payoff slack 10^{-6} does accept false supports here, e.g. the full column support at $c > 0$; the reported families do not use it.) The family criterion compares the resulting collections of supports for set equality.

B.3 Matrix games: constructions and results

Matrix-game constructions. The three 250×250 families isolate distinct mechanisms; their constructions and the oracle-face definition are in Section 6. Each family constructs one base matrix (the constructor seed is fixed at 8) and the six instance seeds perturb that same matrix by masked

Table 3: Synthetic support diagnostic at $b = 0.01, c = 0.001$. Gap is evaluated on the original game. "Family" requires equality of the row and column equilibrium-support families, enumerated exactly for the original matrix and under the tolerances $\tau_v = 10^{-12}$, $\tau_\delta = 10^{-6}$ for the compressed matrix (Section B); both missing true supports and additional false supports count as failures. FACE-1 denotes spectral rank 1 plus the coordinate anchors of the oracle support; on this 3×3 game the anchored reconstruction reproduces M to 3×10^{-16}, so the row tests face exactness, and storage reduction is untested here.

Method	$\|M - \widehat{M}\|_2$	Gap_M	Family
Blind-1	1.42×10^{-2}	1.00×10^{-2}	no
Blind-2	5.00×10^{-4}	1.24×10^{-5}	no
Blind-3	8.67×10^{-16}	1.02×10^{-16}	yes
FACE-1	3.33×10^{-16}	3.47×10^{-18}	yes

Table 4: Non-oracle QRE support estimation at $\tau_{\mathrm{lo}} = 10^{-4}$; a softmax heuristic predicts diffuse outside mass when $\gamma < 6.9 \times 10^{-4}$ at this temperature. Each path uses six temperatures, tolerance 10^{-6}, and at most 50,000 iterations per temperature. $|S|$ is the estimated support size on each side; P/R is support precision/recall; wall-clock time is the recorded QRE estimation time per setting in a single CPU process on the evaluation machine.

γ	QRE iterations	wall-clock (s)	converged	$\|S\|$	P/R
10^{-1}	3,160	0.13	yes	3	1.0/1.0
10^{-2}	25,320	1.12	yes	3	1.0/1.0
10^{-3}	115,400	4.97	no	3	1.0/1.0
10^{-4}	112,000	4.79	no	6	0.5/1.0
10^{-5}	108,560	4.67	no	6	0.5/1.0

Rademacher draws of level ϵ, so the seeds vary the perturbation and hold the base fixed. Repeated coordinates overwrite earlier values, so the realized nonzero fraction falls slightly below the nominal 0.12, and the label of the second family names the added noise: no pair of actions is constructed to be a near duplicate. The two-stage design bounds the rank of the base matrix by $4 + 125 = 129$, but the instance perturbation is a masked Rademacher matrix of Frobenius norm ϵ supported on every nonzero position, so the perturbed matrix is numerically full rank and the rank-128 truncation removes a small residual direction. The observed failure at spectral error 2.9×10^{-3} is consistent with strategically important tail directions, and we do not isolate which discarded direction causes the displacement. The value of the wide-face base is exactly 0 in every instance and any strategy supported on its zero actions is optimal, so the construction fixes the dimension of its equilibrium set without support counts; the oracle pair is returned by the HiGHS LP. The $|S_0|/|S_1|$ columns of Table 2 report the sizes of these faces; the LP support sizes are a different quantity. On the wide-face family the LP supports themselves are small, as they must be for a basic solution of a game with 243 identical zero actions per player: over the 24 instances each returned strategy has at most 3 coordinates above 10^{-8} and at most one of them on a zero action, while the best-response faces contain 247–250 rows and 247–250 columns because every zero action is indifferent at value 0. The audit reads the faces with the same rule as the stored pipeline (`exp4_cond._nf_shared`, called by `exp10_multigame.py`) and asserts index-set equality on all 24 reconstructed instances (`exp16_wideface_oracle_audit.py`); the audit re-derives the instances from their seeds and compares the two current extraction paths, it does not reload masks saved by the stored run, so it corroborates the face sizes of Table 2 (247.5 rows, 250 columns) under the shared rule; it does not authenticate the stored outputs themselves. The five oracle-seeded misses of Table 2 are one instance (seed 2, $\epsilon = 10^{-3}$) at all five ranks, since the oracle-seeded solve does not depend on the rank. We separated the candidate causes on that instance. Thresholding is not one of them: the face read off at payoff tolerances 10^{-8} and 10^{-6} is the same (250 rows, 248 columns; the two excluded columns lose 7.4×10^{-5} and 8.7×10^{-6} against the oracle pair), and the oracle pair itself has restricted gap 8×10^{-22}, so it is an equilibrium of the restricted game, which is degenerate: the LP returns a different one of its equilibria; that pair has restricted gap 0 but full-game gap 1.1×10^{-4} in the stored run (`figure_data/exp10_diag.csv`, column `gap_face`) and 1.9×10^{-4} when the

audit repeats the interior-point face solve (restricted gap 0), a different restricted equilibrium each time, and since every row is in the face the profit can only come from an excluded column. The miss is therefore selection among multiple restricted equilibria; numerical truncation is not the cause: the excluded columns have small positive slacks against the oracle pair (8.7×10^{-6} and 7.4×10^{-5}), but these are pointwise oracle-pair slacks, weaker than the uniform outside margin of Theorem 4.5, and the restricted face fails that uniform-margin hypothesis because one of its restricted equilibria is exploited by an excluded column. The expansion loop certifies a point after adding the exploiting column.

B.4 Leduc sequence-form setup and evaluation

The Leduc instance is converted to its sequence form with suit isomorphism: the sequence spaces of the row and column players index a 337×337 payoff matrix with 966 structurally nonzero entries. A perturbation is masked to the nonzero pattern: each structurally nonzero entry receives an independent Rademacher draw in $\{\pm 1\}$, and the result is rescaled to unit Frobenius norm before multiplication by the perturbation level ϵ. Masking preserves the zero pattern, so the perturbation awards payoff mass only to sequence pairs the game can reach; an unmasked perturbation of the same scale shifts the game value by more than an order of magnitude beyond the intended level.

Point-to-set equilibrium distance. For sequence-form candidates, d_{set} is the sum of two LP distances. The row player's equilibrium set $X(M)$ consists of all realization plans x with $F_0 x = f_0$, $x \geq 0$, whose value against every column realization plan is at least $v(M)$; by linear programming duality the exact value condition holds exactly when there is a dual vector u with $F_1^\top u \leq A^\top x$ and $f_1^\top u \geq v(M)$, where the sequence-form payoff matrix A equals M, and the implemented distance relaxes the last inequality to $f_1^\top u \geq v(M) - 10^{-6}$. The distance of a candidate $\widehat{x}$ to the relaxed set $X_{10^{-6}}(M)$ is $\min \|x - \widehat{x}\|_1$ over such pairs (x, u); the column side is symmetric, and $d_{\text{set}}(\widehat{x}, \widehat{y})$ is the sum of the two minima. A candidate is counted as a hit when $d_{\text{set}} \leq 10^{-5}$ and its full-game NashConv is at most 10^{-6}. The NashConv of a pair adds the best-response gains of the two players against it (Lanctot et al., 2017), which in a zero-sum game equals the pair's gap. Because the value constraint is relaxed by 10^{-6}, d_{set} is the distance to the 10^{-6}-relaxed optimal-strategy polytopes $X_{10^{-6}}(M) \times Y_{10^{-6}}(M)$; the distance to the exact sets $X(M) \times Y(M)$ is a different quantity. A zero distance certifies membership in the relaxed set, and the hit criterion pairs it with the exact full-game gap. The metric is one-sided: it locates the returned pair relative to the original equilibrium set and says nothing about equilibria of the compressed game away from the returned point. For scale, feasible Leduc realization plans have ℓ_1 mass between 25 and 43 (row player) and 25 and 91 (column player), so distances above 100 rearrange a constant fraction of the plan mass.

Table 5: Non-oracle support expansion on the Leduc sequence form, medians over 24 instances per seed. $|S_0|/|S_1|$: behavioral-support sizes; P/R: support precision/recall against the oracle carrier face, the best-continuation-value carrier of the oracle pair (post-hoc only). "Initial": seed support before expansion; "final": output of Algorithm 1. Expansion raises recall on every seed. All 24 instances of each seed terminate certified with residual-verified gap within 2×10^{-8}.

	initial				final		
seed	$\lvert S_0\rvert/\lvert S_1\rvert$	$P/R\ S_0$	$P/R\ S_1$	solves	$\lvert S_0\rvert/\lvert S_1\rvert$	$P/R\ S_0$	$P/R\ S_1$
base	171/190	.97/.83	.98/.85	8	228/224.5	.84/.96	.91/.94
rank-1	42/45	.74/.15	.73/.15	22	284.5/292.5	.69/.98	.72/.96
rank-16	24/63	.63/.07	.81/.23	20	276.5/292	.70/.97	.71/.96

B.5 Selection-displacement protocol

The explicit family of Proposition 4.4 is instantiated at $\gamma = 1$ with the matrices displayed in its proof. Over the 20 settings $\epsilon \in \{10^{-4}, \ldots, 10^{-1}\}$ and $\eta \in \{0, \ldots, 10^{-1}\}$, the LP equilibrium of the compressed perturbed game matches $\eta/(2(4\epsilon + \eta))$ to 10^{-6} with zero outside mass. At fixed residual $\eta = 10^{-3}$ the displacement grows from 0.0012 to 0.357 as ϵ falls from 10^{-1} to 10^{-4}, and it saturates toward $1/2$ once η exceeds ϵ. SVD truncation of the same game erases the selection signal: at rank 2 the spectral error is about 2ϵ, as small as 2×10^{-4}, yet the truncated game's unique equilibrium is a pure point at displacement $1/2$ for every ϵ. Aggregated by level, the

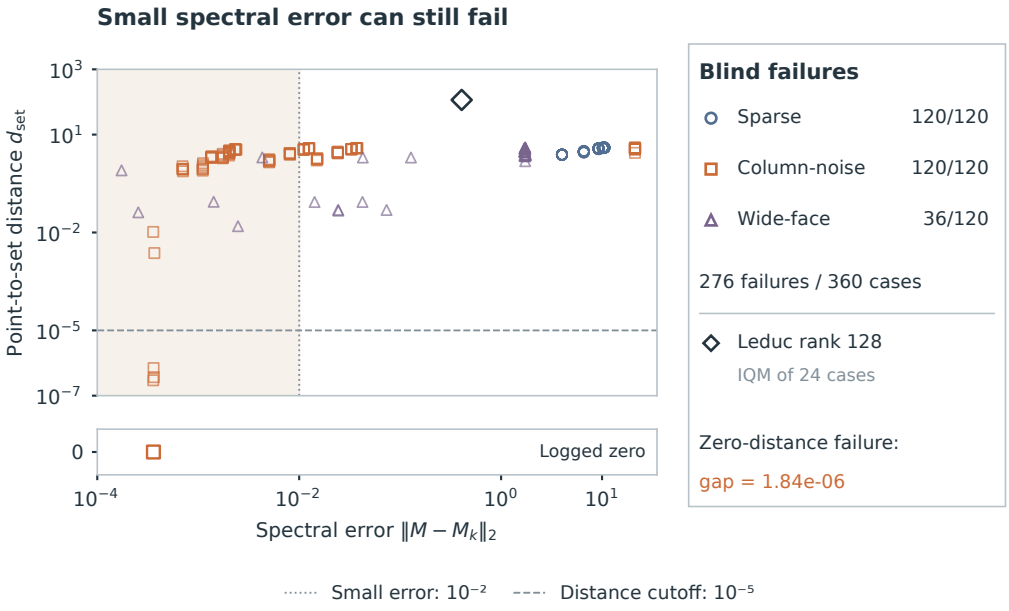


Figure 3: Blind failures in the spectral-error/point-to-set-distance plane. The figure accounts for all 276 failures among 360 matrix-game cases; the side panel gives each family's count. The column-noise family fails at every rank, including small-error cases with substantial displacement; the wide-face family fails only at ranks 1 and 4. Positive distances appear on the logarithmic axis. The separate zero strip retains one column-noise case with logged distance 0 and gap 1.84×10^{-6}: a hit requires both $d_{\text{set}} \leq 10^{-5}$ and gap $\leq 10^{-6}$. The dotted line marks spectral error 10^{-2}, and the dashed line marks the distance cutoff. The diamond gives the rank-128 Leduc interquartile means over 24 cases, recomputed from the paired logs.

sparse and column-noise families miss $30/30$ at every ϵ down to 10^{-3} and the Leduc point-to-set distance stays above 120, so failure persists at every scale of the entry perturbation. The sweep crosses $\epsilon \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ with $\eta \in \{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$, giving 20 settings; each compressed perturbed game is solved exactly by the Koller–Megiddo–von Stengel (KMvS) sequence-form LP (Koller et al., 1996), and the reported check compares the solved displacement against the closed form to an absolute tolerance of 10^{-6} while verifying zero mass on the outside actions. The SVD arm is fully specified: starting from the $\gamma = 1$ base game, the four margin entries (third row against the face columns, third column against the face rows) are shifted by 0.3 times standard-normal draws from a fixed generator (NumPy `default_rng(0)`) so the top singular directions are not axis-aligned, the in-face perturbation is ϵ times matching pennies, and $M + \epsilon\Delta$ is truncated at ranks one and two; the log records the spectral error, the face residual, and the displacement of the truncated game's equilibrium from the perturbed game's equilibrium. Uniqueness at rank two is certified analytically: the returned equilibrium is pure at some (i, j), and the truncated matrix satisfies the strict saddle inequalities $\widehat{M}_{ij} > \widehat{M}_{rj}$ for $r \neq i$ and $\widehat{M}_{ij} < \widehat{M}_{ic}$ for $c \neq j$ with margin at least 9.0×10^{-6}, scaling as $\approx 0.09\,\epsilon$ across the four levels. A strict pure saddle is the unique equilibrium of a zero-sum matrix game: if x' is optimal then $x'^{\top}\widehat{M}e_j \geq v = \widehat{M}_{ij}$, while strictness of the column maximum forces $x'^{\top}\widehat{M}e_j < \widehat{M}_{ij}$ unless $x' = e_i$; the column side is symmetric. A secondary perturb-and-resolve check (20 generic payoff perturbations of size 10^{-9}) agrees; at rank one the equilibrium is degenerate at $\epsilon \leq 10^{-3}$ and no saddle certificate exists, so only the rank-two displacement is quoted in the main text. The rank-two truncation is itself well determined at every level: the spectra of $M + \epsilon\Delta$ have $\sigma_2 - \sigma_3$ between 1.22 and 1.42, so the rank-two truncated SVD is unambiguous, and $\|M + \epsilon\Delta - \widehat{M}_2\|_2 = \sigma_3 \approx 2\epsilon$ confirms the quoted spectral error (Table 6). The per-level re-aggregation groups the 360 matrix-game rows and the 120 Leduc rows of the main diagnostic by perturbation level and reports the IQM of the blind point-to-set distance and the miss count at each level. A separate sanity sweep of 55 random flat-carrier instances satisfying

the condition of Theorem 4.3 confirms support containment, the selection identity, and the proof's inequality $\mathrm{Gap}_{\Delta}^{\mathrm{sel}}(\hat{z}) \le 2\eta/\epsilon$ in every instance.

Table 6: SVD arm of the selection experiment: spectrum of $M + \epsilon\Delta$, the cutoff gap $\sigma_2 - \sigma_3$, the spectral error of the rank-two truncation, and the certified strict-saddle margins (saddle at row 2, column 2). `gesdd` of NumPy/SciPy LAPACK, single precision not involved; all values from the fixed-generator construction of this section.

ϵ	σ_1	σ_2	σ_3	$\sigma_2 - \sigma_3$	row margin	col margin
10^{-4}	1.576419	1.416622	1.99×10^{-4}	1.416	9.01×10^{-6}	1.30×10^{-5}
10^{-3}	1.576419	1.416622	1.99×10^{-3}	1.415	9.01×10^{-5}	1.30×10^{-4}
10^{-2}	1.576420	1.416622	1.99×10^{-2}	1.397	9.01×10^{-4}	1.30×10^{-3}
10^{-1}	1.576486	1.416643	1.99×10^{-1}	1.217	9.21×10^{-3}	1.32×10^{-2}

B.6 Double-oracle and support-expansion protocols

Both policy pools start from the uniform policy. Each round builds the exact meta-game by evaluating every policy pair with OpenSpiel's (Lanctot et al., 2019) expected game score, then adds the fixed perturbation $\epsilon\delta(i, j)$ with $\epsilon = 10^{-2}$, where $\delta(i, j) \in [-1, 1]$ is drawn from a generator keyed by the two policies' rounded action tables and the run seed; entries of existing pairs are therefore unchanged when the pool grows. The blind arm truncates the perturbed meta-game at rank $\lceil T/2 \rceil$ of the current pool size before the LP; the certified arm runs Algorithm 1 with $\epsilon_{\mathrm{solve}} = \epsilon_{\mathrm{cert}} = 10^{-9}$, seeded from the previous round's final working set; that working set contains the previous meta-support and, as recorded below, equals the previous pool in every logged round. All arms compute exact best responses in the true game against the aggregated meta-mixture, deduplicate new policies by their action tables, and record NashConv of the aggregate on the unperturbed game. The perturbation keyed to policy pairs keeps existing entries fixed as the pool grows, but a true-game best response is not in general a best response for the perturbed payoff $u + \epsilon\delta$, so the runs are not a literal instance of Proposition 5.6, whose oracle and certificate share one payoff function. What transfers is the pool-level bound: since $|\epsilon\delta| \le \epsilon$, a mixture certified to gap g on the perturbed current pool has true-game gap at most $g + 2\epsilon$ over the same pool; deviations outside the pool are controlled only by the outer loop's own convergence. Per-round diagnostics evaluate the returned meta-mixture on the uncompressed perturbed meta-game: its exploitability there, whether its support (indices with mass above 10^{-6}) differs from that of an exact LP solution of the same current pool, the certificate status, and the number of restricted solves; the support comparison is within-pool and does not compare the pools grown by different arms. The runs cover $12/40/14$ iterations on Kuhn poker, Leduc poker, and one-die liar's dice, six seeds per game and three arms, for $1{,}188$ logged rounds. An instrumented replay of the certified arm logs the seed working set: it equals the current pool on all three games in every logged round, because the retained working set already covers the previous pool and the first expansion adds the newly entered policy. The support of the certified mixture is much smaller than the working set (for example 13 of 40 pool policies per player at Leduc seed 1, round 39), so the full working set reflects retention history; the meta-equilibrium is far from fully mixed. The two restricted solves per round act at nearly the pool dimension, and the certified arm buys per-round certification and performs no dimension reduction. (Replay trajectories can differ from the main runs in best-response tie-breaking; the working-set observation is identical across both.)

Table 7: Double oracle with compressed meta-solvers, six seeds per game ($12/40/14$ iterations). "gap": median exploitability on the uncompressed perturbed meta-game (equilibrium-invariant); "flip": rounds whose meta-support differs from the reference LP's; "min NC": seed IQM of best NashConv; the exact-LP arm matches the Algorithm 1 columns.

	Blind-$\lceil T/2 \rceil$			Algorithm 1		
game	gap	flip	min NC	cert.	solves/round	min NC
Kuhn	1.5×10^{-2}	49/72	0.018	72/72	1.6	0.006
Leduc	3.3×10^{-1}	210/240	0.707	240/240	2.0	0.666
liar's dice	2.3×10^{-1}	64/84	1.416	84/84	1.9	1.254

Compressed meta-candidate on a shared pool. A second diagnostic separates certification from the pool itself. One pool trajectory per seed, driven by the exact-LP arm, serves every meta-solver, so all of them see the same perturbed meta-game and the same reference support. Truncated at rank $\lceil T/4 \rceil$, far below the pool size, the candidate's support differs from the exact LP solution in 361 of 396 rounds, so the candidate is not a disguised exact solve. Seeding Algorithm 1 from that candidate returns the exact LP support in all 396 rounds and certifies in every round, at a median of 3 to 8 restricted solves against 2 for a solve seeded from the previous round's final working set, which equals the previous pool in every logged round (the seed-size columns of `exp20_psro_shared_pool.csv` coincide with the previous pool dimensions in all 378 rounds after the first). Support agreement is not an exploitability certificate: the blind candidate's median gap on the perturbed pool is 0.20, 1.23, and 0.56 on the three games, and in 7 rounds it agrees in support with the exact LP while its gap exceeds 10^{-6}, whereas both certified arms stay below 5×10^{-14}. The pool rank T is the larger pool dimension. The 72 Kuhn rounds of each arm contain only 44 distinct seed-and-pool-size combinations, since Kuhn pools stall once best responses repeat; the Leduc and liar's dice rounds are all distinct, and the round counts are not independence claims (Table 8).

Table 8: Compressed meta-candidate on one shared pool trajectory per seed, six seeds per game ($12/40/14$ iterations). The exact-LP arm drives pool growth and supplies the reference support, so every row compares meta-solvers on the same perturbed meta-game. "blind-$\lceil T/4 \rceil$": LP on the rank-$\lceil T/4 \rceil$ truncation; "seeded": Algorithm 1 started from the blind candidate's support; "working set": started from the previous round's final working set, which equals the previous pool in every logged round. "flip": rounds whose meta-support differs from the exact LP solution of the same pool; "gap": median exploitability of the returned mixture on the perturbed pool; "solves": median restricted solves. Both certified arms certify and match the reference support in every round.

	blind-$\lceil T/4 \rceil$		seeded			working set
game	flip	gap	flip	gap	solves	solves
Kuhn	54/72	0.20	0/72	2.7×10^{-17}	3	2
Leduc	234/240	1.23	0/240	3.9×10^{-16}	8	2
liar's dice	73/84	0.56	0/84	6.6×10^{-17}	4	2

Best-response expansion and oracle-face extraction. The main-text statements are written for simplex strategy sets; the Leduc experiments run on realization-plan polytopes. The certificate and the termination argument transfer verbatim:

Proposition B.1 (Sequence-form certificate and termination)**.** *Consider a game in extensive form where recall is perfect and two players have zero-sum payoffs. For each player p, let $P_p = \{x \geq 0 : F_p x = f_p\}$ be the sequence-form realization-plan polytope over n_p sequences (nonempty and compact), and write M for the sequence-form payoff matrix. (i) For any $(x, y) \in P_0 \times P_1$, the quantity $\mathrm{Gap}_M(x, y) = \max_{x' \in P_0} x'^\top M y - \min_{y' \in P_1} x^\top M y'$, computed by two LPs on the original M, bounds the gain of every deviation of either player. (ii) Call $S_p \subseteq [n_p]$* closed *if it contains the parent of each of its sequences and a continuation at every information set it reaches, and let $P_p[S_p] = \{x \in P_p : x_j = 0,\ j \notin S_p\}$; every nonempty closed set contains the root sequence, which carries the initial flow f_p, so closedness makes $P_p[S_p]$ nonempty. Run Algorithm 1 with restricted solves over $P_0[S_0] \times P_1[S_1]$ and expansion by the support of a violating best-response realization plan (a closed set). Every non-returning round grows $|S_0| + |S_1|$ strictly, so the loop returns after at most $n_0 + n_1$ rounds, and the returned pair satisfies $\mathrm{Gap}_M(x, y) \leq \epsilon_{\mathrm{solve}} + 2\epsilon_{\mathrm{cert}}$ with deviations ranging over $P_0 \times P_1$.*

Proof. (i) is the definition of exploitability for bilinear payoffs on compact polytopes; both extrema are LPs over P_0 and P_1. (ii) The restricted solve returns (x, y) with restricted gap $g_S \leq \epsilon_{\mathrm{solve}}$, where g_S takes both extrema over $P_0[S_0] \times P_1[S_1]$; the residuals $r_x = \max_{x' \in P_0} x'^\top M y - \max_{x' \in P_0[S_0]} x'^\top M y$ and r_y (symmetric) are evaluated on the full polytopes. If the round returns, $\mathrm{Gap}_M(x, y) = g_S + r_x + r_y \leq \epsilon_{\mathrm{solve}} + 2\epsilon_{\mathrm{cert}}$ exactly as in Theorem 5.3, whose proof uses only this decomposition and never the simplex structure. If it does not return, say $r_x > \epsilon_{\mathrm{cert}}$, the maximizing best-response plan $x^\star$ satisfies $x^{\star\top} M y > \max_{x' \in P_0[S_0]} x'^\top M y$, so $\mathrm{supp}(x^\star) \not\subseteq S_0$ and the expansion adds at least one new sequence; supports of realization plans are closed by the flow constraints, so the invariant is maintained. Each of the $n_0 + n_1$ sequences is added at most once. □

In the implementation, the restricted solve and the global best responses are evaluated through sequence-form LPs; the expansion step adds the support of the primal best-response realization plans, which contains a violating sequence whenever the certificate fails, so the count of Lemma 5.4 carries over with actions replaced by sequences.

The oracle-face pipeline embeds the oracle behavioral support as a set of sequences closed under ancestry, solves the restricted game, and then alternates global best responses on the full matrix with support expansion until the full-game NashConv drops to 2×10^{-8}, following Algorithm 1. Each global best response is evaluated through its dual sequence LP at a tight HiGHS feasibility tolerance (10^{-9}) (Huangfu & Hall, 2018). A returned dual is not trusted as-is: the certificate measures the largest constraint violations $\delta_0 = \max(My - F_0^\top w, 0)$ of the row best-response dual w and $\delta_1 = \max(F_1^\top u - M^\top x, 0)$ of the column best-response dual u, and corrects the values with the sequence L1-mass bound C_p (the number of rows of the realization matrix F_p, one for the root plus one per information set of player p). The certified gap is the residual-verified upper bound $(f_0^\top w + C_0\delta_0) - (f_1^\top u - C_1\delta_1) \geq \mathrm{Gap}_M(x, y)$, clamped below zero; the primal LP values are kept only as a sanity/sandwich log and never enter the certificate. The candidate itself is checked before certification: the realization-plan residuals $\|F_0x - f_0\|_\infty$, $\|F_1y - f_1\|_\infty$ and the negative parts of x, y must be at most 10^{-6} (measured values are about 10^{-15}), otherwise the pair is rejected. The correction is exact for a feasible pair when the residuals and objectives are evaluated exactly; in the implementation they are evaluated in double precision without outward rounding and without repairing the 10^{-15}-scale realization residuals, so every reported certified gap is a numerically checked corrected-gap estimate; it is not a rigorously rounded upper bound: the stopping threshold 2×10^{-8} is seven orders above the observed residuals, but we do not compute a roundoff budget for the inner products, and the values recorded at termination (10^{-14}–10^{-13}) are floating-point evaluations of the corrected bound. The oracle pair is a corner point of the equilibrium sets: the game value is computed by the sequence-form LP, and x° and y° are the corner points of two zero-objective feasibility LPs over the 10^{-6}-relaxed sets, solved through SciPy's linear-programming interface with HiGHS (its default `highs` method); the oracle face of the Leduc diagnostic is the behavioral support $\{(I, a) : \beta(I, a) > 10^{-9}\}$ of that pair, while the non-oracle study of Table 5 scores against its best-continuation-value carrier. The LP returns one point of the equilibrium polytope, so when the equilibrium set is degenerate the extracted face is the sub-face selected by the solver, in general a proper sub-face of the full carrier, with no guarantee of minimality or maximality. The expansion loop then certifies a point of the original game from that seed, as the non-oracle study confirms; it does not recover the carrier coordinates the corner omits. Over the 24 Leduc instances the loop certifies after 6–10 rounds (median 8), each round one restricted solve plus two global best-response LPs; the returned pair's behavioral support differs from the seeding corner by up to seven actions per side (typically smaller, since the corner can carry actions the certified pair does not use), so Table 1's certified columns report the post-expansion output of Algorithm 1; a raw corner-face solve would be a different pair. For scale, the face-anchored projector of Proposition 5.1 on the seed-1, $\epsilon = 10^{-3}$ instance has dimensions $\dim Q_x$ rising from 190 to 281 and $\dim Q_y$ from 196 to 288 over nominal ranks 1–128, i.e. 56–85% of the 337-dimensional ambient space; storing the two orthonormal factors and the dense core exceeds the original sparse matrix, so the anchored representation buys face exactness and costs storage.

The non-oracle study seeds Algorithm 1 without any oracle: the *base* seed is the equilibrium behavioral support of the unperturbed base game (171/190 actions on average); the *rank*-1 and *rank*-16 seeds are the equilibrium behavioral supports of the spectral truncations of the perturbed matrix at those ranks (42/45 and 24/63). Seed construction never reads the perturbed oracle pair, which is used only for post-hoc precision/recall. Restricted games are solved twice when needed: the default HiGHS presolve can spuriously flag a degenerate restricted KMvS LP as infeasible under a nonzero objective, and retrying without presolve recovers a valid restricted equilibrium. All 72 expansion rows terminate certified, each with a non-negative residual-verified corrected gap, a numerically checked estimate of the full-game gap evaluated in double precision as described above (median 9.5×10^{-14}). Wall-clock times in Table 4 are measured in a single CPU process on the evaluation machine.

The residual correction is substantive: it changes every reported value. Re-running the 72 expansion loops while recording the final-round certificate shows dual residuals δ_0, δ_1 with medians 4.6×10^{-16} and 1.1×10^{-16} (maxima 2.3×10^{-15} and 7.6×10^{-16}), about six orders of magnitude below the 10^{-9} dual-feasibility tolerance; the raw dual endpoint is slightly negative ($\approx -10^{-15}$) in 38 of 72

instances and the clamp maps it to zero. The reported certified gap is then dominated by the explicit correction term $C_0\delta_0$ (median 6.6×10^{-14}); the clamp contributes less. The bound $C_p = \mathrm{rows}(F_p)$ is deliberately conservative: it assumes all realization constraints can be violated simultaneously, and the data show the resulting conservatism is a few times 10^{-13} on Leduc.

Support estimation and cost. A support estimate is an efficiency hint, supplied by a previous meta-solver solution, a basic LP solve, or an anchored-QRE (McKelvey & Palfrey, 1995) path. The certificate decides validity, and the estimate only affects how many expansion rounds the loop needs. In the Leduc diagnostic the oracle supports contain about 190 behavioral actions per player in a sequence space of 337. In the matrix games of Table 2 they span 117 to 250 of 250 actions per side, so the dimension reduction comes almost entirely from support restriction.

Remark B.2 (Point certificate, empirical support recovery). Algorithm 1 certifies its output per strategy on the original matrix, and it achieves high recall against the selected reference carrier face. Recall rises to 0.94 to 0.98 over the 72 Leduc instances of Table 5 from initial medians as low as 0.07, and reaches 0.98 to 1.00 on the sparse and column-noise games of Table 2. The wide-face game is the exception, where a point certificate covers one point of a large face and recall is correspondingly low.

B.7 Local constant and runtime scope

The local constant of Theorem 4.5 is probed on the 6×6 game of the support study, whose equilibrium set Z_F is the single point with both players uniform on $\{0, 1, 2\}$. Two probes return only lower bounds on the constant; neither certifies an upper bound. First, softmax perturbations anchored at the uniform point with the outside coordinates held at $\log 0$ confine every sample to the face (the largest outside probability is below 10^{-6}); the maximum ratio $\mathrm{dist}(z, Z_F)/\operatorname{Gap}_M(z)$ over gaps $[10^{-5}, 10^{-1}]$ is ≈ 1.3, probing only face-internal directions. Second, the face-outer family $(x^\star,\ (1-p)y^\star + pe_j)$ has gap $p\gamma$ and distance $p\sqrt{4/3}$, hence ratio $\sqrt{4/3}/\gamma$, growing without bound as $\gamma \to 0$; this closed form is verified to machine precision. The honest conclusion is $\kappa \geq \sqrt{4/3}/\gamma$ on this family, and the apparent stability of $\widehat{\kappa} \approx 1.3$ across margins was vacuous because the Monte-Carlo grid never leaves the face.

Resource-matched benchmark. To test whether a compressed candidate saves computation, `exp18_resource_matched.py` times certified paths on each of the 24 perturbed Leduc instances in one process. Two baselines solve the full matrix: the exact KMvS LP pair followed by its certificate alone (two best-response duals on the LP pair, no further solve), and the exact LP pair followed by one full-support round of Algorithm 1, which re-solves the full game once before the same two duals; the logged `n_solves` counts Algorithm 1 rounds and excludes the seed LP. The compressed path takes the rank-k truncated SVD, the exact LP pair on the truncated matrix, and Algorithm 1 on the original matrix seeded with that solution's behavioral support, for $k \in \{1, 16, 64, 128\}$. All 144 runs certify. The LP-plus-certificate baseline takes a median of 0.036 s (range 0.032–0.050) and the LP-plus-one-round baseline 0.066 s (range 0.061–0.103), both over the 672 non-root sequence coordinates; the compressed paths take medians of 0.62–0.69 s with 20–22 restricted solves (range 16–25) and 8.0×10^3–8.6×10^3 summed restricted sequence-space sizes. Every compressed path is slower than both baselines on its instance, by a factor of 12 to 27 (median 18) against the LP-plus-certificate baseline and 5.5 to 14 (median 9.3) against the LP-plus-one-round baseline, and no tested rank recovers the cost. The SVD itself takes 0.018–0.028 s, about half of the LP-plus-certificate baseline but under 5% of the compressed path; the cost is the expansion rounds that the truncated support forces. The benchmark does not pin BLAS or HiGHS thread counts; all paths ran under the same library defaults in one process. For the two full-matrix baselines the logged `size_sum` is the full dimension 672 by convention; it is not a sum over expansion rounds. This benchmark tests one SVD-seeded certification pipeline on one 337×337 sequence-form game; low-rank compression may still help meta-solving in other pipelines. *Runtime scope*. The benchmark covers the solver path: SVD, seed LP, restricted solves, best-response duals, and certificate evaluation, all on the stored perturbed matrix; it excludes payoff construction, which every path shares. The Leduc study, using an oracle equilibrium support, establishes face-exactness behavior and conditional support sizes; the support-estimation experiment records QRE iterations and wall-clock time.

B.8 Information abstraction as an alternative

Information abstraction solves a smaller game over coarser information states and leaves the payoff matrix unchanged (Sandholm & Singh, 2012; Kroer & Sandholm, 2014). We compare the equity and WEVA abstractions against the oracle-seeded certified solve and blind truncation on the same 24 Leduc instances, with one yardstick, the exploitability on the full game. The abstractions build a restricted sequence-form game on the same perturbed payoff matrix, solve it exactly, lift the equilibrium back to the 337×337 full game, and evaluate gap there; the Blind and certified rows are the gaps on the perturbed matrix; the certified row is a restricted solve of the original matrix and uses no compressor.

Equity bucketing (Gilpin & Sandholm, 2007) groups private cards by showdown equity, so $K{=}3$ separates every rank and $K{=}1$ merges them all. The WEVA abstraction (Li & Huang, 2026b) learns the buckets from a warm-up instead: W iterations of vanilla counterfactual regret minimization (CFR) (Zinkevich et al., 2007), per-information-set expected values weighted by the depth-attention ratio of that warm-up, then k-means++ (Arthur & Vassilvitskii, 2007) clustering into K buckets. Leduc has three private ranks, so $K{=}3$ merges no cards (the identity abstraction) and necessarily returns the original game; the saturation at $K{=}3$ follows from the game's small card space. Table 9 reports interquartile means over the 24 instances.

Table 9: Information abstraction on the Leduc sequence form, interquartile means over 24 instances. All columns are exploitability on the full 337×337 sequence-form game, in chips. Blind and the oracle-seeded certified solve evaluate their output on the perturbed matrix; equity and WEVA lift their abstract equilibrium to the full game. "≈ 0" rows are below 10^{-7}, the LP solve residual.

method	full-game gap
Blind-1	9.04
Blind-16	9.01
Blind-128	1.89
Oracle-seeded certified solve	≈ 0
equity $K{=}1$	4.35
equity $K{=}2$	1.54
equity $K{=}3$	≈ 0
WEVA $K{=}1$	4.35
WEVA $K{=}2$	1.53
WEVA $K{=}3$	≈ 0

Blind truncation fails at every tested rank budget, leaving interquartile-mean gaps of 1.9 to 9.0 chips across ranks, 1.89 at $k = 128$: singular energy carries no information about which payoffs determine equilibrium indifference. The abstractions read the strategic value of each hand and are exact as soon as no two ranks are merged. WEVA's buckets coincide for $W{=}10$ and $W{=}50$: the warm-up is far from converged (its own profile gap is 2.5 chips at $W{=}10$ and 0.6 at $W{=}50$), yet the per-infoset expected values already separate the three ranks, so the learned clustering reproduces the equity split (WEVA $K{=}2$ gives 1.53 versus 1.54 for equity). The certified row carries a numerically checked corrected-gap estimate against the original matrix (Theorem 5.3), while the abstraction columns are empirical point values.

B.9 Controlled initialization and higher-temperature support estimation

The margin sweep is complemented by a controlled initialization study. Starting from a 2×2 subset of the core requires two restricted solves and one expansion round, which adds one action per player; a singleton requires three solves and two expansion rounds. Both terminate at the true 3×3 support with gap at machine precision. Starting from all six actions terminates in one solve with gap 1.1×10^{-16} but keeps the full 6×6 game. Full-matrix best responses therefore repair missing strategically relevant actions; extra actions are safe and stay in place under this expansion-only loop.

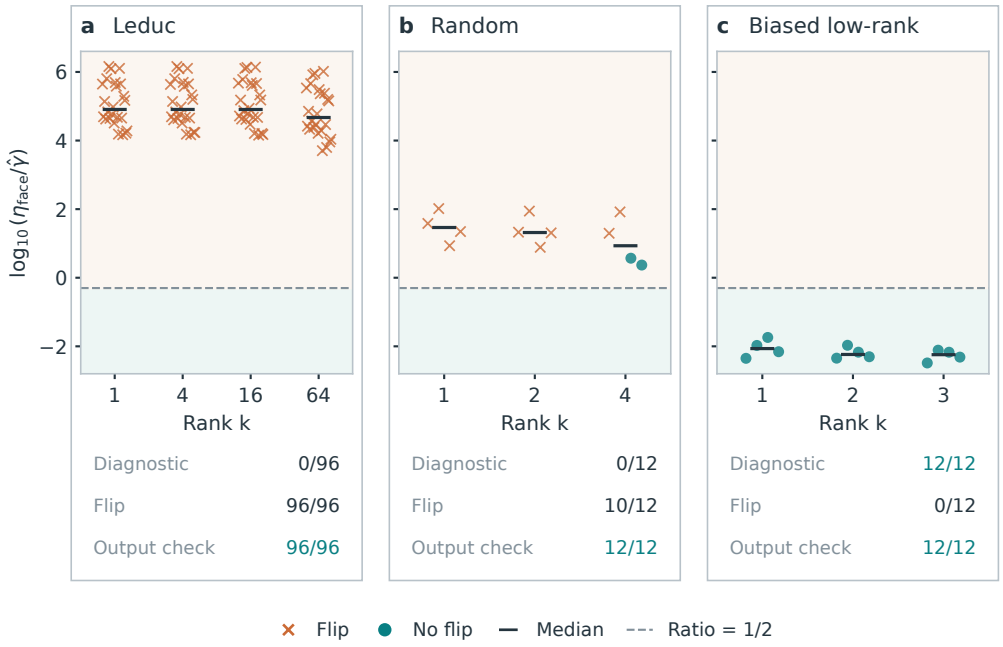


Figure 4: Local margin-residual diagnostic across three game families. All 120 logged settings are shown on a shared vertical scale; rank groups are spaced categorically, with deterministic horizontal offsets to separate observations. The dashed line is $d_{\text{diag}}/\hat{\gamma} = 1/2$, with $\hat{\gamma}$ evaluated at the solved equilibrium. The pale teal band satisfies this local diagnostic and the pale orange band does not. Crosses mark an outside-action flip after compression; circles mark no flip, and dark horizontal marks are within-rank medians. The aligned summaries report diagnostic satisfaction, flips, and the logged legacy output flag. The flag is true in all 120 settings, while the diagnostic holds in the 12 biased low-rank settings.

B.10 Margin-residual diagnostic

The biased low-rank family of the margin-residual diagnostic (Figure 4) is a 14×14 matrix UV whose four support rows are copies of a shared positive vector d and whose ten outside rows are $0.05\,d$, with independent Gaussian noise of scale 2×10^{-3} confined to outside rows. Figure 4 shows the sampled row-side diagnostic $d_{\text{diag}} < \hat{\gamma}/2$ of Section 5 on this family.

On 120 settings we measure $\hat{\gamma}$, the sampled row-side residual d_{diag}, and outside-action flips against the line $d_{\text{diag}} < \hat{\gamma}/2$. The residual is the largest change in the row player's best-response values under compression, $d_{\text{diag}} = \max_i |r_i(\widehat{M}) - r_i(M)|$ for the normal-form families and the maximum change of the row player's best-continuation values over reachable information sets for Leduc; it is one-sided and does not estimate the two-sided operator envelope η_{face} of Section 5. A biased low-rank family satisfies the line in all 12 settings with no flips. The random and Leduc families violate it in all 108 and flip in 106. The logged output flag records $\texttt{gapM} < 10^{-8}$ on the stored runs and is true in all 120 settings; 40 of the 120 stored $\texttt{gapM}$ values are negative (minimum -7.3×10^{-8}), so the flag is a legacy numerical output check of the stored runs; it is not a certified gap bound.

Support estimation at a higher low temperature. The same margin becomes tractable when the low temperature rises. At $\gamma = 10^{-4}$ with $\tau_{\text{lo}} = 10^{-2}$, the path converges in 5,800 iterations, yet the retained support is still the full six actions on each side. Diffuse support and non-convergence are therefore distinct failure modes: the relative threshold $\rho = 10^{-3}$ retains all six actions even for a converged path, independent of whether the residual tolerance is met under the iteration budget.

C Reproducibility Statement

The perturbation instances of Section B are deterministic: each seed fixes one instance through a fixed NumPy generator seeded by the integer seed, so every instance is exactly reproducible, and the interquartile means of Section 6 aggregate these per-instance rows. The double-oracle entry perturbations of Section B.6 are keyed by the two policies' signatures and the run seed through Python's built-in hash, which keeps existing entries fixed as the pool grows within a run but is randomized across processes, so the same seed and policy pair can receive a different perturbation in another process; the protocol of the released runs can therefore be rerun, but the stored aggregate statistics need not reproduce exactly, since a replay in another process can differ from a stored run in the payoff matrix itself, beyond tie-breaking. Future runs should key the perturbation by a stable digest of the serialized policies. Solver outputs are reproducible up to LP and best-response tie-breaking, which is fixed within a run of the released code but can differ across solver versions; the double-oracle replay of Section B.6 records one such difference. The core implementation is distributed as compiled Python extension modules, with recorded reference results and separate output directories for new runs. Table 10 collects the hyperparameters of every experiment.

Equilibria are computed with HiGHS through SciPy's linear-programming interface: SciPy's default `highs` method on the KMvS sequence-form LP of the Leduc experiments, for the oracle faces, the certified solves, and the double-oracle meta-games; and, for the restricted face solves of the three matrix-game families (`exp10`, `exp16`), that solution followed by an interior-point polish (`highs-ipm`, primal and dual feasibility tolerances 10^{-10}), which is why those runs can report a different restricted equilibrium of a degenerate face than the first solve alone. Singular-value decompositions use the `gesdd` driver of NumPy/SciPy LAPACK; game environments use OpenSpiel. Throughout, "exact LP" means the linear program on the uncompressed matrix solved in floating point; exact rational arithmetic is used only for the support-family enumeration of the original synthetic matrices in Section B. All experiments run on the evaluation machine; each instance is solved in a single process, and the drivers of the Leduc diagnostic, the non-oracle expansion study, and the margin-residual sweep parallelize independent instances across at most eight worker processes. Wall-clock entries in the tables are indicative only.

D Extended Related Work

Abstraction and equilibrium sensitivity. Abstraction solves a smaller game and bounds the lifted solution (Sandholm & Singh, 2012; Kroer & Sandholm, 2014; 2018; 2016); refinement can be non-monotone (Waugh et al., 2009), and learning-based abstraction and pruning select which actions and information states to keep (Li et al., 2024; Li & Huang, 2025; 2026b). Strategic stability asks which equilibria survive perturbation classes (Kohlberg & Mertens, 1986; Van Damme, 1991). We bound how compression moves the equilibrium set; prior work bounds the support size of such a set (Lipton et al., 2003; Schmid et al., 2014).

Regularized targets and certificates. Quantal response equilibrium replaces the set-valued correspondence with a unique regularized point (McKelvey & Palfrey, 1995; Hoda et al., 2010). Last-iterate results for regret matching (RM) and CFR characterize a solver on a fixed game (Cai et al., 2025; Meng et al., 2025b;a), whereas compression changes the game itself. Small certificates make partial verification practical (Zhang & Sandholm, 2020b), anytime DO variants give stop-anytime exploitability bounds (McAleer et al., 2022), and a deployed strategy can carry a certificate for its own exploitability (Li & Huang, 2026a).

Conclusion

We studied when low-rank payoff compression preserves the strategic structure of a zero-sum game. Spectral accuracy alone cannot answer this question: arbitrarily small truncation error can leave the original and compressed equilibrium sets a constant distance apart. Our selection analysis identifies the relevant scale. When a perturbation of size ϵ selects an equilibrium on a flat carrier face, a compression residual of size η produces displacement controlled by η/ϵ, under support containment and local regularity. The explicit family attains coordinate displacement $\eta/(2(4\epsilon+\eta))$ (Proposition 4.4).

Thus, a direction that contributes little to the matrix norm can still determine the equilibrium selected by the game.

We introduced face-anchored compression (FACE) to retain these strategically relevant directions. By augmenting the spectral subspace with candidate support coordinates, FACE preserves payoffs exactly on the selected face. Together with a uniform outside-action margin that survives compression, face exactness identifies a common equilibrium set; a local gap-to-distance bound then controls nearby approximate equilibria (Theorem 4.5). For an unknown support, Algorithm 1 uses original-matrix best responses to test and expand the candidate. Its returned strategy satisfies $\mathrm{Gap}_M \leq \epsilon_{\mathrm{solve}} + 2\epsilon_{\mathrm{cert}}$ (Theorem 5.3). This separates the structural conditions for preserving an entire equilibrium set from the directly checkable guarantee for one returned strategy.

The experiments connect these guarantees to observed behavior. Blind truncation misses at every tested rank on the 337×337 Leduc sequence form, and the column-noise matrix family exhibits failure at spectral error 2.9×10^{-3}. Oracle-seeded certification succeeds on all 24 Leduc instances at each rank, using expansion to repair the seeded support. Across three non-oracle initializations of the same 24 instances, certification succeeds in all 72 runs and raises median support recall to 0.94–0.98 (Table 5). In double-oracle runs on three OpenSpiel games, the certified meta-solver returns the exact meta-solution's support in all 396 logged rounds under fixed solver and best-response tie-breaking. Blind truncation changes that support in 68–88% of rounds (Table 7).

The contribution is a criterion for deciding what a compressed game can be trusted to preserve. Selection analysis explains why spectral approximation can fail; face anchoring retains the selected payoff structure; and original-game certification checks the strategy used by the solver or population update. Support estimates determine the work needed for certification, while the certificate determines the validity of the output. Together, these results make equilibrium geometry and original-game exploitability explicit requirements for reliable payoff compression.

Table 10: Hyperparameters. All values as used in the experiments of Section 6; the surrounding protocol details are in Section B.

experiment	hyperparameter	value
synthetic probe	rows b	$\{0.01, 0.05, 0.1, 0.3\}$
	rows c	$\{0, 0.001, 0.01\}$
	ranks k	$\{1, 2, 3\}$
	support test	rational-arithmetic LP; exact ($\tau_v = \tau_\delta = 0$) for original matrices, $\tau_v = 10^{-12}$, $\tau_\delta = 10^{-6}$ for compressed matrices
Leduc diagnostic	perturbation ϵ	$\{10^{-3}, 10^{-2}, 10^{-1}, 3 \times 10^{-1}\}$
	seeds	6
	ranks k	$\{1, 4, 16, 64, 128\}$
	hit criterion	$d_{\text{set}} \le 10^{-5}$, gap $\le 10^{-6}$
QRE support estimation	low temperature τ_{lo}	10^{-4} (10^{-2})
	temperatures	6
	solve tolerance	10^{-6}
	iteration cap	50,000 per temperature
	retention ρ	10^{-3}
	margin γ	$\{10^{-1}, \dots, 10^{-5}\}$
non-oracle expansion	seeds	base, rank-1, rank-16
	certificate tolerance ϵ_{cert}	10^{-8}; stop when certified full gap $\le 2\epsilon_{\text{cert}} = 2 \times 10^{-8}$
	LP feasibility tolerance (HiGHS)	10^{-9}
matrix games	protocol	as Leduc
	sparse density	0.12
	column-noise δ	10^{-3}
	wide-face noise	10^{-4}
selection displacement	γ	1
	ϵ	$\{10^{-4}, \dots, 10^{-1}\}$
	η	$\{0, \dots, 10^{-1}\}$
	LP tolerance	10^{-6}
	SVD shift	0.3 (`default_rng(0)`)
	saddle margin	$\ge 9.0 \times 10^{-6}$
double oracle	perturbation ϵ	10^{-2}
	solver tolerances	10^{-9}
	iterations	$12/40/14$
	seeds	6
	blind rank	$\lceil T/2 \rceil$
	shared-pool blind rank	$\lceil T/4 \rceil$, $T = \max(m, n)$
	certificate tolerance	$\epsilon_{\text{solve}} = \epsilon_{\text{cert}} = 10^{-9}$
resource benchmark	ranks k	$\{1, 16, 64, 128\}$
	baselines	full LP + one full-support round of Algorithm 1; full LP + certificate only
	certificate tolerance ϵ_{cert}	10^{-8}
information abstraction	buckets K	$\{1, 2, 3\}$
	WEVA CFR iterations W	$\{10, 50\}$
margin-residual	matrix size	14×14
	outside scale	$0.05\,d$
	noise	2×10^{-3}