CERTIFYING GLOBAL OPTIMALITY OF STATIONARY POINTS IN IMPERFECT-RECALL GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Single-player decision problems with imperfect recall model a player who may forget information she once had. They arise wherever a strategy must stay compact: the absent-minded driver, imperfect-recall abstractions of poker, and team games compressed into a single forgetful player. The ex-ante optimal behavior strategy of such a problem maximizes a polynomial over a product of simplices. Deciding optimality is $\exists\mathbb{R}$-complete, near-optimality is NP-hard, and the degree of the polynomial grows with the number of visits to an information set. Practice therefore relies on first-order methods: regret matching and its predictive-plus variant reach stationary points at scale. Such a point comes with no guarantee about its distance from the global optimum. We add a certificate layer that runs after the solver. A support-pruned sum-of-squares feasibility program at level $\lceil \deg u/2 \rceil$ bounds how far any strategy can beat the reported value. A posterior step turns the floating-point solution into a bound re-verified in exact and interval arithmetic, and the same certificate decomposes the residual gap across information sets. On 77 instances regenerated from the generators of the reference benchmark, the layer certifies the PRM$^+$ stationary point as globally optimal within relative tolerance 10^{-4} on 70, or 90.9%. The verdict is invariant across 10 seed groups and 4 of 5 first-order methods. The global solver SCIP corroborates it to the same tolerance on all 77; on one instance SCIP reaches that tolerance only at its one-hour limit, where the layer needs 68 seconds. Support pruning is lossless at level one and wherever strict complementarity holds. The one certified loss is an interior cubic whose gap narrows $33\times$ one level up. On these instances, the stationary points of the standard solvers are almost always global optima, and a sum-of-squares certificate makes that fact checkable in seconds and provable in exact arithmetic in minutes.

1 INTRODUCTION

A single-player decision problem with imperfect recall asks for an ex-ante optimal behavior strategy in a game tree where the player may forget information she once had (Piccione & Rubinstein, 1997; Aumann et al., 1997). The setting contains the absent-minded driver (Piccione & Rubinstein, 1997; Aumann et al., 1997), the imperfect-recall abstractions that make large poker games solvable (Waugh et al., 2009; Ganzfried & Sandholm, 2014; Kroer & Sandholm, 2016; Li & Huang, 2025; 2026b), and the compressed representations through which adversarial team games are computed (Celli & Gatti, 2018; Zhang et al., 2023; Carminati et al., 2026). Through the bridge of Gimbert *et al.* (Gimbert et al., 2020), ex-ante optimality is polynomial optimization over a product of simplices. Deciding optimality is $\exists\mathbb{R}$-complete, even near-optimality is NP-hard (Tewolde et al., 2023), and the degree of the objective grows with the number of visits to an information set, a set of decision nodes the player cannot tell apart.

In practice the field runs on first-order methods. Regret matching (Hart & Mas-Colell, 2000; Zinkevich et al., 2007; Lanctot et al., 2012) and its RM$^+$ successors (Tammelin, 2014; Farina et al., 2023) scale to large games, including predictive regret matching$^+$, or PRM$^+$ (Farina et al., 2021), the method used by the reference benchmark (Tewolde et al., 2026). Under alternating updates RM$^+$ provably reaches ε-KKT points (Anagnostides et al., 2025). These methods stop at stationarity. A user who runs PRM$^+$ learns that no local first-order improvement exists at the returned point. Whether that point is globally optimal stays open, and the algorithms give no bound on the gap.

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.

We regenerate 77 instances whose base-level semidefinite relaxation fits the basis-size cap of §4, using the generators of the reference benchmark (Tewolde et al., 2026). On 70 of them a sum-of-squares certificate labels the PRM$^+$ stationary point certified-global, or 90.9%. Certified-global means a posterior gap below 10^{-4} relative, as §3 defines. Labels are invariant across ten seed groups, and RM, RM$^+$, PRM$^+$, and projected gradient agree on the certified-global rate *exactly*. Optimistic gradient descent (Rakhlin & Sridharan, 2013) is the one method that disagrees, at 88.3%. It stalls at worse non-stationary points on two otherwise-certified instances. On the extended absent-minded driver, the cubic instance of §5 whose stationary point is interior, it collapses to value zero.

The finding survives cross-validation with the global solver SCIP (Bestuzheva et al., 2023). Judged by the same 10^{-4} tolerance, its proven bound corroborates all 77 PRM$^+$ points as optimal: 52 of the 70 certified instances within one minute and 69 within an hour. On the last instance SCIP reaches the tolerance only at its one-hour limit, with a dual bound 4×10^{-5} relative above the stationary value, where our layer needs 68 seconds. Both solvers work on the same instances in floating point, SCIP on a bilinear reformulation with one variable per leaf and our layer on the utility polynomial.

The certificate is a feasibility program in the SOS hierarchy of §4: a Handelman representation (Handelman, 1988) of a polynomial nonnegative on Δ, pruned by the support of the stationary point (the actions it plays) and strengthened by a block of monomial multipliers. Its existence at level $d_0 = \lceil \deg u/2 \rceil$ certifies that no strategy beats the reported value by more than ε. Two design choices carry the method. First, posterior *verification* (Lemma 1) converts the floating-point semidefinite-programming (SDP) solution into a scalar gap bound in well under a second, with no second solve and 6×10^{-9} relative inflation on the median. Second, support pruning removes the blocks that complementary slackness prices at zero at the stationary point, shrinking the kept basis by 39% on the pool. Proposition 3 covers quadratic utilities exactly, where the pruned and full cones coincide. On the 57 vertex and mixed instances that CLARABEL (Goulart & Chen, 2026) solves accurately (those with at least one infoset on a single action), the posterior gaps agree to 1.8×10^{-7} on the median. The theory exhibits the failure of those conditions in Example 1 of §5: an interior stationary point with a cubic utility, gap 0.172, closing $33\times$ one level up.

The certificate also *decomposes* the suboptimality gap. One certificate per stationary point attributes the gap to each information set through a nonnegative, exactly-consistent decomposition (Proposition 5). A second solve of the same size selects the decomposition (§5), and no separate solve is needed per infoset. Across the 567 perturbations that shift mass onto unplayed actions, on 69 of the 70 certified instances, the perturbed block receives the largest share in 100% of cases, with 99.6% of the gap on the median.

Contributions.

- **We find that on the generated pool, stationary points are almost always global optima.** Certified-global PRM$^+$ points are 70 of 77, seed-invariant, stable across four of five first-order methods, and corroborated by SCIP on all 77 instances under a common tolerance.
- **We build a certificate layer that runs after the solver with no second solve.** A cone reduced by the candidate's support lifts the floating-point solution to a certificate of the same cone and level (Corollary 1), bounded in under a second and re-verified in exact and interval arithmetic once per instance. It also attributes the gap to each infoset (Proposition 5); in every covered perturbation the perturbed infoset receives the largest share.
- **We prove support pruning lossless at level one and under strict complementarity, and exhibit where it fails.** At level one the pruned and full cones coincide (Proposition 3), covering all 22 quadratic instances. Beyond it, an interior cubic instance (Example 1) is an exact obstruction; it suffices that the multipliers vanishing at the optimum vanish identically (Proposition 2), and the loss vanishes up the hierarchy (Proposition 4).

2 Related Work

Imperfect recall. Piccione and Rubinstein (Piccione & Rubinstein, 1997) identified the interpretational puzzles of imperfect recall; the absent-minded driver (Aumann et al., 1997) is its canonical instance. Solving single-player imperfect-recall games is $\exists\mathbb{R}$-hard (Gimbert et al., 2020; Tewolde et al., 2023), equilibrium concepts are mapped in (Tewolde et al., 2024), and the value of recall is

quantified in (Berker et al., 2025). Imperfect recall has also served as a deliberate abstraction device in large games (Waugh et al., 2009), and no-regret dynamics over such trees were studied early (Lanctot et al., 2012). Our setting is narrower: we take the solver output as given and certify it a posteriori. Appendix C discusses exact algorithms for two-player imperfect-recall games.

Sum-of-squares for imperfect recall. The closest reference is Zheng *et al.* (Zheng et al., 2026), who solve imperfect-recall games through sum-of-squares optimization, proving convergence on the SOS-concave and SOS-monotone subclasses and solving them from scratch. We start from *any* stationary point returned by a first-order method and cover all utility subclasses. Our program also minimizes a global upper bound over a certificate cone, with $\gamma = \hat{u} + \varepsilon$. The cone is reduced by the candidate's support (§4), and the floating-point solution is lifted to a certificate of the same cone at the same level (Corollary 1). The underlying moment–SOS hierarchy is Lasserre's (Lasserre, 2001), and we follow the SDP–game connection of Laraki and Lasserre (Laraki & Lasserre, 2012) at a candidate-dependent cone. Recovering a verifiable certificate from a floating-point SOS solve is classical (Peyrl & Parrilo, 2008); Lemma 1 recovers one by pure evaluation, and the repaired certificate stays in the pruned cone. §5 establishes sufficient conditions for lossless reduction and exact obstructions to it, and Proposition 5 reads the certificate out per infoset. The polynomial-optimization and certificate literature behind the method is surveyed in Appendix C.

First-order methods and their guarantees. Regret matching has been a standard method for game solving since Zinkevich et al. (2007); under alternating updates, RM$^+$ provably converges to ε-KKT points in constrained smooth optimization (Anagnostides et al., 2025). These guarantees cover stationarity alone. Appendix C surveys how the same family scales to large imperfect-information games. Our layer checks global optimality of the stationary point a posteriori.

3 PRELIMINARIES

A single-player extensive-form decision problem is a finite rooted game tree. The leaves form a set Z, chance nodes carry fixed distributions, and decision nodes are divided into information sets, or *infosets*, indexed by $j = 1, \dots, \ell$. Each infoset j has an action set A_j, and each leaf carries a utility $U(z) \in \mathbb{R}$. We allow *imperfect recall* in full generality: distinct nodes on the same root–leaf path may belong to the same infoset, as in the absent-minded driver (Piccione & Rubinstein, 1997; Aumann et al., 1997). A *behavioral strategy* assigns each infoset j a distribution over its actions. Writing $n = \sum_j |A_j|$ and identifying variables with infoset–action pairs $v = (j, a)$, the strategy space is the product of simplices

$$\Delta \;=\; \Big\{\mu \geq 0 \,:\, \textstyle\sum_{a \in B_j} \mu_a = 1,\; j = 1, \dots, \ell\Big\} \;\subset\; [0,1]^n,$$

where B_j is the variable block of infoset j and μ_a abbreviates $\mu_{(j,a)}$.

Expected utility is a polynomial. Under a behavioral strategy μ, the probability of reaching leaf z factorizes over its path, so the expected utility

$$u(\mu) \;=\; \sum_{z \in Z} p_c(z)\, U(z) \prod_{(j,a) \in \mathrm{path}(z)} \mu_{(j,a)} \tag{1}$$

is a polynomial indexed by leaves. Here $p_c(z)$ collects the chance probabilities, and (j, a) appears with multiplicity equal to the number of times the path to z traverses infoset j choosing a. Hence $\deg u \;\leq\; \max_z |\mathrm{path}(z)|$, and absentmindedness is what makes $\deg u$ exceed 1 per infoset. We write $|\alpha| = \sum_v \alpha_v$ and $\mu^\alpha = \prod_v \mu_v^{\alpha_v}$ for a multi-index $\alpha \in \mathbb{N}^n$, and e_v for the multi-index with a single 1 at position v. The coefficients of u are computed from the tree in exact rational arithmetic when chance probabilities and payoffs are rational (true of every benchmark file). A broad class of polynomial optimization problems on a product of simplices can be encoded this way (Gimbert et al., 2020).

Solution concepts and stationary points. The *ex-ante optimal* strategies are the global maximizers of u on Δ. First-order methods instead target stationary points. For $\mu \in \Delta$ define the block-wise stationarity gap

$$\mathrm{gap}_{\mathrm{kkt}}(\mu) \;=\; \max_{j \leq \ell} \Big(\max_{a \in B_j} \partial_a u(\mu) \;-\; \sum_{a \in B_j} \mu_a\, \partial_a u(\mu) \Big), \tag{2}$$

where $\partial_a u$ is the partial derivative of u with respect to μ_a. This is the largest first-order improvement available by moving mass within one infoset. A point with $\mathrm{gap}_{\mathrm{kkt}}(\mu) = 0$ satisfies the Karush–Kuhn–Tucker (KKT) conditions for $\max_\Delta u$. These points coincide with the causal-decision-theoretic equilibria of Tewolde et al. (2023). Regret-matching-type dynamics dominate in practice (Tewolde et al., 2026). Under alternating updates, RM^+ provably converges to ε-KKT points in constrained smooth optimization (Anagnostides et al., 2025). Throughout, $\hat{\mu}$ is a candidate stationary point returned by such a method, with $\hat{u} = u(\hat{\mu})$ and support $\mathrm{supp}(\hat{\mu}) = \{v : \hat{\mu}_v > \tau\}$. We use PRM^+ in the experiments, with four alternatives for robustness, and set $\tau = 10^{-6}$ there and $\tau = 0$ in the statements of Section 5.

Certified global optimality. Our object of study is an *a posteriori certificate*: a bound, computed after the run, that $u(\mu) \le \hat{u} + \varepsilon$ for all $\mu \in \Delta$, in which case we call $\hat{\mu}$ *ε-certified globally optimal*. The certificate applies to any candidate $\hat{\mu}$. Its validity reduces to finitely many explicit linear-algebra quantities, stated in Lemma 1 and evaluated in double precision for the headline labels. For one certificate per instance these quantities are re-verified using exact rational and interval arithmetic (Remark 2), which proves the bound for the game specified in the benchmark file.

Moment–SOS background. A polynomial σ of degree $\le 2d$ is *sum-of-squares* (SOS) iff $\sigma = b_d^\top G\, b_d$ for some $G \succeq 0$, the *Gram matrix* of the representation, where b_d lists the monomials of degree $\le d$, of size $\mathcal{B} = \binom{n+d}{d}$; the truncated basis b_{d-1} has size $\mathcal{L} = \binom{n+d-1}{d-1}$. The Lasserre/moment hierarchy (Lasserre, 2001) bounds $\max_\Delta u$ by semidefinite programs (SDPs) indexed by a *level* d, with base level $d_0 = \max\{1, \lceil \deg u/2 \rceil\}$, which equals 1 for constant u. On the moment side the variable is a truncated moment vector $y = (y_\alpha)_{|\alpha| \le 2d}$ with $y_0 = 1$, and Λ_y is the linear functional $\mu^\alpha \mapsto y_\alpha$ on polynomials. The constraints are $M_d(y) \succeq 0$ and its *localizing* versions $M_{d-1}(\mu_v y) \succeq 0$, the moment matrices weighted by the constraint $\mu_v \ge 0$, where $M_d(y)$ is the $\mathcal{B}\times\mathcal{B}$ moment matrix with entries $y_{\alpha+\beta}$, together with the simplex equalities $\Lambda_y\big((\sum_{a\in B_j} \mu_a - 1)\,\mu^\beta\big) = 0$ for $|\beta| \le 2d-1$. The dual certificate view is given in Appendix A. Section 4 defines the exact cone, strengthened by a nonnegative monomial block essential at the base level (Remark 1).

4 The Certificate Layer

The layer takes a utility polynomial u, a stationary point $\hat{\mu}$ with value $\hat{u} = u(\hat{\mu})$, and a level $d \ge d_0$. It returns a scalar bound $\varepsilon_{\mathrm{rig}} \ge 0$ that satisfies $u \le \hat{u} + \varepsilon_{\mathrm{rig}}$ on Δ; we define it in Lemma 1 and evaluate it in double precision, as Remark 2 details. The layer has three parts: a strengthened certificate cone whose base level is well-posed, a support-pruned feasibility program that searches this cone, and a posterior numerical verification step that converts the numerical solution into a scalar upper bound.

Definition 1 (Truncated certificate cone)**.** *For the product of simplices $\Delta \subset [0,1]^n$ with blocks $B_1, \dots, B_\ell$ and level $d \ge 1$, let*

$$\mathcal{C}_d = \Big\{ \sigma_0 + \textstyle\sum_{v=1}^n \mu_v \sigma_v + \sum_{j=1}^\ell c_j\big(\sum_{a\in B_j} \mu_a - 1\big) + \sum_{|\alpha|\le 2d} t_\alpha \mu^\alpha \Big\},$$

where σ_0 is SOS of degree $\le 2d$, each σ_v is SOS of degree $\le 2d-2$ and $\mu_v\sigma_v$ is the localizing multiplier *of v, $c_j \in \mathbb{R}[\mu]_{\le 2d-1}$ is a free polynomial, and $t_\alpha \ge 0$. Here $\mathbb{R}[\mu]_{\le k}$ denotes the polynomials in μ of degree at most k. On the moment side, the dual of the t-block is the componentwise constraint $y \ge 0$ on the truncated moment vector y, which is how the strengthening enters the relaxation.*

Proposition 1 (Validity and asymptotic exactness)**.** *(i) Every $p \in \mathcal{C}_d$ is nonnegative on Δ; consequently $\inf\{\gamma : \gamma - u \in \mathcal{C}_d\}$ upper-bounds $\max_\Delta u$, and the strengthened moment relaxation remains valid. (ii) If p is strictly positive on Δ, then $p \in \mathcal{C}_d$ for all sufficiently large d; the hierarchy of bounds converges to $\max_\Delta u$ as $d \to \infty$.*

Proof. See Appendix A: the argument reduces Δ to a full-dimensional simplex product, invokes Handelman's theorem (Handelman, 1988) on the affine functions defining it, and lifts the representation back through the forms $\sum_{a\in B_j} \mu_a - 1$ that define Δ. The ideal remainders obey the degree bound $2d \ge \max\{D, \deg p\}$, with D the maximal degree of the Handelman representation. □

Remark 1 (Why the strengthening matters at the base level). Without the t-block, the level-d_0 relaxation can be unbounded. For $\deg u = 2$ the localizing blocks degenerate to scalar inequalities, and

boundedness then depends on the utility: maximizing p^2 over the two-action simplex $p+q=1$ gives an unbounded level-1 moment relaxation, whereas the absent-minded driver $u = 4pq$ stays bounded and exact at level 1. The t-block adds $y_\alpha \geq 0$ for every truncated moment. The simplex moment equalities give $y_\beta = \sum_{a \in B_j} y_{\beta+e_a}$ within each block j, and nonnegativity then forces $y_\alpha \in [0,1]$ by recursion. This restores boundedness at every level at negligible cost. On the certificate side it widens the cone at fixed level, which makes the support-pruned feasibility program feasible at level d_0 where the same cone without the t-block is not, as Example 1 shows.

The moment-side check. Minimizing γ over $\gamma - u \in \mathcal{C}_d$ is a single SDP with optimal value $\gamma_d \geq \max_\Delta u$. We report $\gamma_{d_0} - \hat{u}$ as a non-constructive baseline. The paper builds on the explicit certificate below, which the posterior step bounds numerically as Lemma 1 states.

The support-pruned certificate program. Fix the level $d = d_0$ unless stated otherwise, and let $V = \{1,\dots,n\} \setminus \mathrm{supp}(\hat{\mu})$ be the *off-support* variables of the stationary point. The layer solves

$$\begin{aligned} \min_{\varepsilon,\, G_0,\, \{H_v\},\, \{c_j\},\, t} \quad & \varepsilon \\ \text{s.t.} \quad & \hat{u} + \varepsilon - u \;=\; b_d^\top G_0\, b_d + \sum_{v \in V} \mu_v\, b_{d-1}^\top H_v\, b_{d-1} \\ & \qquad + \sum_{j=1}^{\ell} c_j\big(\textstyle\sum_{a \in B_j} \mu_a - 1\big) + \sum_{|\alpha| \leq 2d} t_\alpha\, \mu^\alpha, \\ & G_0 \succeq 0, \qquad H_v \succeq 0 \;\; (v \in V), \qquad t \geq 0, \end{aligned} \tag{P_{cert}}$$

where the polynomial identity is enforced coefficient-wise. Program (P_{cert}) deletes the localizing multipliers of every *supported* variable, compared with searching all of $\mathcal{C}_d$. If $\hat{\mu}$ is optimal and an exact certificate exists, the multipliers of supported variables are priced at zero there and vanish (Lemma 2), so they are candidates for removal. The removal leaves the bound valid and can only enlarge ε; Section 5 characterizes when it is lossless.

The payoff is computational and structural. The number of positive semidefinite (PSD) blocks drops from n to $|V|$, and every surviving block is tied to off-support mass of one infoset, which makes the attribution of Definition 3 meaningful. By Proposition 1(i), any feasible point establishes $u \leq \hat{u} + \varepsilon$ on Δ. Writing $\varepsilon^\star$ for the optimal value of (P_{cert}), a small $\varepsilon^\star$ certifies $\hat{\mu}$ to be $\varepsilon^\star$-certified globally optimal in the sense of Section 3, and a positive $\varepsilon^\star$ is a certified *upper bound on the optimality gap* of $\hat{\mu}$. A lossy cone can certify $\varepsilon^\star > 0$ even when $\hat{\mu}$ is optimal, as Section 5 shows. The coefficient matching is enforced by one sparse linear system on the vectorized Gram variables, reused unchanged in the posterior step below, with no second solve.

Posterior numerical verification. An SDP solver returns (P_{cert}) only up to feasibility tolerances. The identity holds only to within a coefficient residual: the Gram matrices can acquire slightly negative eigenvalues, and t can fall below zero. The following lemma converts such numerical output into a scalar upper bound by pure evaluation, with no second solve.

Lemma 1 (Posterior numerical certification of a solver output)**.** *Let $\Delta \subset [0,1]^n$ be the product of simplices over blocks $B_1,\dots,B_\ell$, let $d \geq 1$, let $u \in \mathbb{R}[\mu]$ with $\deg u \leq 2d$, and let $\hat{u} \in \mathbb{R}$. Write $\mathbb{S}^k$ for the set of symmetric $k \times k$ matrices. Suppose a numerical solver returns $\varepsilon \in \mathbb{R}$, symmetric matrices $G_0 \in \mathbb{S}^{\mathcal{B}}$ and $H_v \in \mathbb{S}^{\mathcal{L}}$ for $v \in V \subseteq \{1,\dots,n\}$, free coefficient vectors $c_1,\dots,c_\ell$, and $t = (t_\alpha)_{|\alpha| \leq 2d}$, intended to satisfy the certificate identity of (P_{cert}),*

$$\begin{aligned} \hat{u} + \varepsilon - u(\mu) \;=\; & b_d(\mu)^\top G_0\, b_d(\mu) \;+\; \sum_{v \in V} \mu_v\, b_{d-1}(\mu)^\top H_v\, b_{d-1}(\mu) \\ & + \sum_{j=1}^{\ell} c_j(\mu)\big(\textstyle\sum_{a \in B_j} \mu_a - 1\big) \;+\; \sum_{|\alpha| \leq 2d} t_\alpha\, \mu^\alpha, \end{aligned} \tag{3}$$

where $b_k = (\mu^\alpha)_{|\alpha| \leq k}$ is the standard monomial basis of Section 3, so that $\|b_d(\mu)\|^2 \leq \mathcal{B}$ on Δ; $c_j \in \mathbb{R}[\mu]$ has degree $\leq 2d-1$; and b_d has size $\mathcal{B}$, b_{d-1} size $\mathcal{L}$. In floating point, the identity (3) *and the constraints $G_0 \succeq 0$, $H_v \succeq 0$, $t \geq 0$ need not hold exactly:*

- $t_\alpha^+ = \max(t_\alpha, 0)$, *penalty* $\pi_t = \sum_\alpha \max(-t_\alpha, 0)$;

- *the eigenvalue penalty* $\pi_\lambda = \big|\lambda_{\min}(G_0)\big|_- \cdot \mathcal{B} + \sum_{v\in V} \big|\lambda_{\min}(H_v)\big|_- \cdot \mathcal{L}$*, where* $|x|_- = \max(-x, 0)$*;*
- *the residual polynomial* r *collecting the coefficient-wise defect of* (3) *with* t^+*, with coefficient norm* $\rho = \sum_\alpha |r_\alpha|$*.*

Then, with $\varepsilon_{\mathrm{rig}} := \varepsilon + \rho + \pi_\lambda + \pi_t$*,* $u(\mu) \le \hat{u} + \varepsilon_{\mathrm{rig}}$ *for all* $\mu \in \Delta$*. If* $\hat{u} = u(\hat{\mu})$ *for some* $\hat{\mu} \in \Delta$*, then* $\hat{\mu}$ *is* $\varepsilon_{\mathrm{rig}}$*-certified globally optimal.*

Proof. See Appendix A: the Gram terms are lower-bounded by $-\pi_\lambda$ via $\|b_d(\mu)\|^2 \le \mathcal{B}$ and the localizing factor $\mu_v \in [0, 1]$; the simplex terms vanish on Δ; the monomial multipliers are lower-bounded by 0; and the residual is absorbed through $|r(\mu)| \le \rho$. □

Remark 2 (Numerical scope of the posterior bound). Lemma 1 is exact given exact values of ρ, π_λ, π_t, and the pipeline evaluates them in double precision. For each of the 77 instances, one certificate is then re-verified offline: the seed-42 re-solve, whose double-precision verdict coincides with the ten-seed label. The residual ρ is recomputed in exact rational arithmetic, and each $\lambda_{\min}$ is bounded below by an interval-arithmetic Cholesky factorization of the Gram matrix shifted by that bound, with the utility polynomial re-extracted from the game file in exact rational arithmetic. The repaired polynomial is itself a certificate in the same cone (Corollary 1). All 77 verdicts are reproduced with gaps enlarged by at most 5.4×10^{-8}, so these certificates are proven bounds for the game as specified; the check costs 24 CPU-minutes for the pool, at most 94 seconds per instance. All other certificates in the paper are double-precision posterior bounds, listed in Table 7.

The basis size $\mathcal{B} = \binom{n+d_0}{d_0}$ governs cost and memory. We solve $\mathcal{B} \le 300$ with the interior-point solver CLARABEL (Goulart & Chen, 2026) and up to 1500 with the first-order solver SCS (O'Donoghue et al., 2016), and report every instance above the hard cap as *skipped*. Table 3 in Appendix B lists the coverage.

5 Analysis: Lossless Pruning and Certificate Attribution

Program (P_{cert}) deletes one PSD block per supported variable. Write $\varepsilon_d^{\mathrm{pruned}}$ and $\varepsilon_d^{\mathrm{full}}$ for the optimal values, or infima, of (P_{cert}) and of the same program with all n localizing blocks; both are finite at every level (Corollary 1). This section answers the two questions pruning raises. When is the deletion *lossless*, meaning $\varepsilon_d^{\mathrm{pruned}} = \varepsilon_d^{\mathrm{full}}$? And what interpretable structure do the surviving blocks carry? All proofs appear in Appendix A.

Lemma 2 (Certificate-side complementary slackness)**.** *Suppose the full (unpruned) certificate is exact at level* d*: there exist SOS* $\sigma_0, \{\sigma_v\}_{v=1}^n$*, free* c_j*, and* $t \ge 0$ *with*

$$u(\hat{\mu}) - u = \sigma_0 + \sum_{v=1}^{n} \mu_v \sigma_v + \sum_j c_j \big(\textstyle\sum_{a\in B_j} \mu_a - 1\big) + \sum_\alpha t_\alpha \mu^\alpha, \tag{4}$$

where $\hat{\mu} \in \Delta$ *is a global maximizer. Then evaluating at* $\hat{\mu}$ *forces every summand to vanish there:* $\sigma_v(\hat{\mu}) = 0$ *for every* $v \in \mathrm{supp}(\hat{\mu})$*,* $\sigma_0(\hat{\mu}) = 0$*, and* $t_\alpha\, \hat{\mu}^\alpha = 0$ *for all* α*.*

Lemma 2 says the supported multipliers are *pointwise* useless at $\hat{\mu}$. Pruning asks for more: that some exact certificate does not use them *as polynomials*.

Definition 2 (Certificate-side strict complementarity)**.** *An exact certificate* (4) *satisfies* strict complementarity *if* $\sigma_v \equiv 0$ *for every* $v \in \mathrm{supp}(\hat{\mu})$*.*

Throughout, the *pruned* program (P_{cert}) fixes $\sigma_v \equiv 0$ for $v \in \mathrm{supp}(\hat{\mu})$ and keeps every other variable and degree bound unchanged. Its feasible set is therefore contained in that of the full program, so $\varepsilon_d^{\mathrm{pruned}} \ge \varepsilon_d^{\mathrm{full}} \ge 0$. The last inequality follows by evaluating the identity at $\hat{\mu} \in \Delta$, where the right side is nonnegative and the left side equals ε.

Proposition 2 (Strict complementarity implies lossless pruning)**.** *If some exact level-*d *certificate satisfies strict complementarity, then the support-pruned program is exact at level* d*:* $\varepsilon_d^{\mathrm{pruned}} = \varepsilon_d^{\mathrm{full}} = 0$*.*

At level 1 the question does not even arise, since the localizing blocks are redundant (Proposition 3). At higher levels strict complementarity can fail, at interior points (Example 1) and at mixed points (Remark 5 in Appendix A). The cost of full pruning vanishes up the hierarchy (Proposition 4).

Proposition 3 (Level one: pruning is lossless unconditionally)**.** *Let $d = 1$, which is the base level d_0 whenever $\deg u \le 2$. Then the pruned program (P_{cert}) and the full program over $\mathcal{C}_1$ have the same set of representable polynomials, so for every $\hat{\mu}$, optimal or not, $\varepsilon_1^{\mathrm{pruned}} = \varepsilon_1^{\mathrm{full}}$.*

Example 1 (Pruning can be lossy at an interior point)**.** *The extended absent-minded-driver instance has $\deg u = 3$, so $d_0 = 2$, and its PRM$^+$ stationary point $\hat{\mu} = (2/3, 1/3)$ lies in the* relative interior *of its single simplex block. Complementary slackness therefore prunes* every *localizing block. An exact level-d_0 full certificate exists, namely the explicit identity of §A.9 in Appendix A, so the level-d_0 moment bound equals $u(\hat{\mu}) = 4$ exactly. The fully pruned program (P_{cert}) has no exact certificate at level d_0 or at any level: the argument in §A.9 is level-free, so level-up closes this gap only to tolerance. The level-d_0 optimum of (P_{cert}) is bounded away from zero. In §A.9 an explicit pseudo-moment functional, nonnegative on the pruned cone, gives $\varepsilon_2^{\mathrm{pruned}} \ge 343/2000 = 0.1715$, matching the solver value 0.172; the gap contracts $33\times$ to 0.0052 at $d_0{+}1$ and $23\times$ further to 2.2×10^{-4} at $d_0{+}2$. Full pruning is thus lossy at an interior point of degree $\deg u \ge 3$ even where an exact full certificate exists; at $\deg u \le 2$ it is never lossy (Proposition 3). The distinction is whether the multipliers vanishing at $\hat{\mu}$ vanish identically, which is strict complementarity as in Definition 2. Remark 5 embeds this instance in a larger game so that the same loss appears at a mixed point.*

Proposition 4 (Asymptotic losslessness)**.** *Fix the polynomial u and the domain Δ. For every $\eta > 0$ there is a level $d = d(u, \Delta, \eta)$ at which the* fully pruned *certificate (no localizing blocks at all) certifies $u \le u(\hat{\mu}) + \eta$ whenever $\hat{\mu}$ is a global maximizer.*

Infoset-level attribution. Beyond its scalar ε, identity (3) *decomposes* the certified gap $\hat{u}+\varepsilon-u(\mu)$ of every strategy μ into per-infoset pieces. After pruning, each surviving localizing block is tied to the off-support actions of one infoset. On exact certificates the pieces are nonnegative and exactly consistent, as Proposition 5 shows. On solver outputs they are empirical diagnostics: a solver output does not satisfy (3) exactly, so each piece inherits the coefficient residual and the Gram eigenvalue error, a piece may be slightly negative, and $\arg\max_j \Gamma_j$ can flip when the two largest attributions are within the identity error of each other. Either way the certificate says which infoset's deviation is charged for the gap.

Definition 3 (Certificate attribution)**.** *Let a (pruned) certificate for $\hat{u} = u(\hat{\mu})$ with gap ε be given as in* (3)*, with kept multiplier set V (the off-support variables) and $\sigma_v = b_{d-1}^\top H_v b_{d-1}$. For a monomial α write $\mathrm{off}(\alpha) = \{j : \exists v \in B_j \cap V,\ \alpha_v > 0\}$, the set of blocks whose off-support variables appear in α. For any $\mu \in \Delta$ define the* per-infoset attribution $\Gamma_j(\mu) = \sum_{v \in B_j \cap V} \mu_v\, \sigma_v(\mu) + \sum_{\alpha:\, \mathrm{off}(\alpha) = \{j\}} t_\alpha\, \mu^\alpha$ *for $j = 1, \dots, \ell$, the* cross term $\Gamma_\times(\mu) = \sum_{\alpha:\, |\mathrm{off}(\alpha)| \ge 2} t_\alpha \mu^\alpha$*, and the* residual $\mathrm{res}(\mu) = \sigma_0(\mu) + \sum_{\alpha:\, \mathrm{off}(\alpha) = \emptyset} t_\alpha \mu^\alpha$*. The t-terms with $\mathrm{off}(\alpha) = \{j\}$ price off-support mass of block j exactly as the σ_v terms do. For $\alpha = e_v + 2\beta$ they are rank-one localizing multipliers $t_\alpha\, \mu_v \mu^{2\beta}$, so they are charged to block j and not to the residual.*

Proposition 5 (Attribution is a valid decomposition)**.** *For every $\mu \in \Delta$: (i) $\Gamma_j(\mu) \ge 0$ for all j, $\Gamma_\times(\mu) \ge 0$, and $\mathrm{res}(\mu) \ge 0$; (ii) $\sum_j \Gamma_j(\mu) + \Gamma_\times(\mu) + \mathrm{res}(\mu) = \hat{u} + \varepsilon - u(\mu)$, i.e. the attributions, the cross term, and the residual exactly split the certified gap of μ; (iii) computing them requires only evaluations of the certificate polynomials, with no additional solve. A single certificate therefore attributes the certified gap of* each *strategy μ simultaneously.*

The per-infoset pieces carry a multiplier interpretation: $\mu_v\, \sigma_v(\mu)$ is the certified price charged to the probability mass that infoset j places on action v outside the support of $\hat{\mu}$, a polynomial analogue of the KKT multiplier of $\mu_v \ge 0$. Two boundary cases are inherent. Deviations *within* the support of an infoset, including every deviation at an interior infoset whose localizing blocks are pruned entirely, are charged to the residual res; and the split of a joint deviation across infosets depends on which certificate is chosen. The faithfulness experiment of §6 measures both, reporting the in-support perturbations separately as residual-charged.

To make the attribution meaningful we select the certificate in two stages, since the exact certificate is not unique and a generic solver dumps most of the explanation into σ_0, which carries no infoset structure. After minimizing ε we pin $\varepsilon \le \varepsilon^\star + \delta$ with $\delta = 10^{-7} \max(|u(\hat{\mu})|, 1)$ and minimize $\mathrm{tr}(G_0) + \mathbf{1}^\top t$. The penalty drives explanation mass out of the unstructured σ_0 block and the t-block;

Table 1: Certified-global verdicts on the 77 instances, by family; family names are the benchmark's generators (Appendix B). A row labelled "5 revisits" adds five further draws to the size tier above it. *One simulation-c instance reaches the tolerance only at the one-hour limit, where our layer certifies it in 68 s.

		certified-global		RM / RM$^+$ / PGD	OGD	SCIP	median
family	**inst.**	**PRM$^+$**	**rate**	**certified**	**certified**	**confirms**	**rel. $\varepsilon_{\mathrm{rig}}$**
sim-a	6	6	100%	6	6	6	4.6×10^{-8}
sim-a, 5 revisits	30	30	100%	30	30	30	1.3×10^{-8}
sim-b, 5 revisits	18	18	100%	18	16	18	3.2×10^{-8}
simple	4	4	100%	4	4	4	4.6×10^{-10}
sim-c, 5 revisits	18	12	66.7%	12	12	18*	6.6×10^{-5}
interior cubic	1	0	0%	0	0	1	4.3×10^{-2}
all	77	70	90.9%	70	68	77	4.4×10^{-8}

Table 2: Pruned vs. full certificates by location and solver, all 77 instances, median absolute posterior gaps. Location: vertex, every infoset on one action; mixed, some mixing; interior, none.

location	**solver**	n	**med. $\varepsilon_{\mathrm{rig}}$**		**pruned vs. full**
			pruned	**full**	
vertex	CLARABEL	27	2.5×10^{-8}	6.9×10^{-9}	consistent, $\lvert\Delta\varepsilon\rvert \le 1.3 \times 10^{-5}$
vertex	SCS	5	2.3×10^{-5}	1.4×10^{-6}	consistent, both certified
mixed	CLARABEL	30	9.0×10^{-7}	4.5×10^{-7}	consistent, $\lvert\Delta\varepsilon\rvert \le 2.5 \times 10^{-5}$
mixed	SCS	13	3.9×10^{-3}	1.6×10^{-4}	undecided, SCS-limited
interior, $\deg u{=}2$	CLARABEL	1	1.1×10^{-8}	1.3×10^{-8}	theorem, lossless
interior, $\deg u{=}3$	CLARABEL	1	1.7×10^{-1}	5.4×10^{-8}	theorem, lossy: exact $343/2000$, certified 0.1718

in our experiments that mass lands in the localizing multipliers. At a full-support point there are no localizing blocks to receive it, and the mass stays in the residual. Pinning ε keeps the certificate's tightness, and the second stage has the same cone and size as the first.

6 Experiments

We instantiate the layer on the benchmark family of Tewolde *et al.* (Tewolde et al., 2026) at the base level $d_0 = \lceil \deg u/2 \rceil$. Table 1 covers the 77 main-pool instances; the 13 others with moment basis at most 1500 form the scale set of Figure 2. Stationary points come from PRM$^+$ over ten seed groups, and a point is certified-global when the posterior gap of Lemma 1 is at most $10^{-4} \max(|u(\hat{\mu})|, 1)$; every verdict is re-verified in exact and interval arithmetic. Appendix B.2 gives the census, protocol, solver settings, and baselines; every instance receives a verdict.

Certified-global stationary points are the norm. The PRM$^+$ point is certified-global on 70 of 77 instances. The label is seed-invariant, with per-instance posterior-gap standard deviation 0 on the median and at most 2.4×10^{-6}, and stable across four orders of magnitude of the tolerance, from 74 certified at 10^{-3} to 60 at 10^{-6}. Of the seven uncertified instances, the extended absent-minded driver is the structural interior cubic counterexample of §5, whose pruned optimum is exactly $343/2000$, certified at 0.1718. The six simulation-c instances belong to the 18 solved by SCS at basis 330, and Figure 1 places them where SCS inflation crosses the threshold. Whether their exact pruned optima vanish remains open (Appendix B.3).

The verdict is independent of the solver and confirmed by SCIP. RM, RM$^+$, and projected gradient descent (PGD) reproduce the certified set exactly over three seed groups (Table 1). Optimistic gradient descent (OGD) certifies two fewer: it stops non-stationary on two simulation-b instances, and it collapses to 0 on the extended absent-minded driver, already uncertified (Appendix B.1). SCIP (Bestuzheva et al., 2023), run for one hour on the benchmark's bilinear reformulation of the

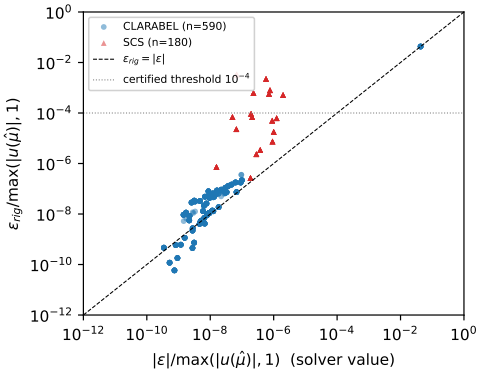


Figure 1: Solver gap $|\varepsilon|$ vs. posterior gap ε_{rig}, one point per (instance, seed). The posterior step leaves CLARABEL gaps at solver accuracy; SCS solutions inflate by up to 3.4×10^{-3}, and every uncertified verdict besides the structural counterexample is an SCS instance.

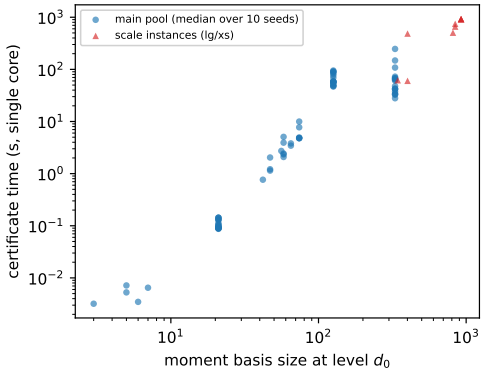


Figure 2: Certificate wall-clock vs. moment basis size at level d_0 (log-log, single core). Blue: main pool, median over ten seeds. Red: completed scale instances (lg: degree-2 detection games; xs: simulation); lg40 at basis 1250–1396 exceeds the practical limit.

same instances at our tolerance, confirms all 77 PRM$^+$ points optimal. Its optimal values match $u(\hat{\mu})$ to relative 1.7×10^{-16} on the median and 1.3×10^{-6} on the seven that our layer withholds, and its median time is 0.6 s on the certified set. On one simulation-c instance SCIP reaches our tolerance only at the one-hour limit, logged at 3612 s with the dual bound still 4×10^{-5} above $u(\hat{\mu})$ and no optimality declared; our layer certifies it in 68 s.

Pruning is lossless where the theory says so. Table 2 classifies the pool by location and solver; the pattern follows the degree split of §5. Proposition 3 settles all 22 quadratic instances before any solve, as Table 2 records by location. On the CLARABEL vertex and mixed instances the pruned and full gaps agree, and both programs certify the SCS vertex instances. The SCS mixed instances are the exception: there solver accuracy dominates in both directions, so they cannot separate the cones. The theory also settles the cubic interior point, which is lossy by Example 1 and closes to tolerance only after a level-up. Pruning also cuts certificate time: total wall-clock drops $1.49\times$ over the pool, the pruned solve is faster on 67 of 77 instances (median $1.53\times$ on those), and the kept basis drops by 39%. Certificate cost follows the moment basis (Figure 2), and the posterior step takes under a second per instance (Appendix B.5, B.3).

Attribution faithfulness. The decomposition is exact (Proposition 5), so the experiment asks which infoset the certificate charges for the gap. For each certified instance, each infoset block j, and each $\delta \in \{0.1, 0.3, 0.5\}$, we shift mass δ within block j onto off-support actions, and check whether the decomposition assigns the largest share of the induced gap to block j, using the single certificate solved once at $\hat{\mu}$, with no re-solving. This covers 567 off-support single-block perturbations over 69 of the 70 certified instances; the perturbed block receives the largest share in 100% of cases at every δ, with median share 0.996 and identity error below 5.3×10^{-6} (Appendix B.4).

7 Conclusion

On 77 instances of the reference benchmark, the stationary points found by the RM-family solvers are globally optimal almost always. The verdict is checkable in under a second and provable in exact arithmetic within two minutes per instance: it certifies 70/77 instances globally optimal within 10^{-4} relative tolerance. The label is seed-invariant, stable across four of five first-order methods, and corroborated by SCIP on all 77 under the same tolerance. The certificate also attributes the gap per infoset. Where SCIP reaches the same tolerance only at its one-hour limit, the layer needs 68 seconds. A first-order run thus returns a certified gap and an account of where it sits, not a bare stationary point. Extending the layer to the semidefinite treatment of two-player maxmin equilibria (Laraki & Lasserre, 2012) is a natural next step.

AI USE STATEMENT

We used generative AI tools to polish the prose of this paper and to write code for the experiments reported in this work. We have reviewed all AI-assisted work, and we take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work concerns the certification of stationary points in imperfect-recall games and involves no human subjects, private data, or deployed systems. We see no material ethical concerns beyond standard academic practice.

REPRODUCIBILITY STATEMENT

The benchmark families, generator parameters, and the experimental protocol appear in Appendix B.2. The 77 certificates behind Table 1 ship with their exact game files and a solver-free standalone verifier that rechecks them in about 24 CPU-minutes. The implementation is included as an anonymous code package in the supplementary material.

REFERENCES

Ioannis Anagnostides, Emanuel Tewolde, Brian Hu Zhang, Ioannis Panageas, Vincent Conitzer, and Tuomas Sandholm. Convergence of regret matching in potential games and constrained optimization. *arXiv preprint arXiv:2510.17067*, 2025.

Robert J Aumann, Sergiu Hart, and Motty Perry. The absent-minded driver. *Games and Economic Behavior*, 20(1):102–116, 1997.

Ratip Emin Berker, Emanuel Tewolde, Ioannis Anagnostides, Tuomas Sandholm, and Vincent Conitzer. The value of recall in extensive-form games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 13631–13640, 2025.

Ksenia Bestuzheva, Mathieu Besançon, Wei-Kun Chen, Antonia Chmiela, Tim Donkiewicz, Jasper van Doornmalen, Leon Eifler, Oliver Gaul, Gerald Gamrath, Ambros Gleixner, et al. Enabling research through the SCIP optimization suite 8.0. *ACM Transactions on Mathematical Software*, 49(2):1–21, 2023.

Branislav Bosansky, Jiri Cermak, Karel Horak, and Michal Pechoucek. Computing maxmin strategies in extensive-form zero-sum games with imperfect recall. In *International Conference on Agents and Artificial Intelligence*, volume 2, pp. 63–74. SCITEPRESS, 2017.

Luca Carminati, Brian Hu Zhang, Federico Cacciamani, Junkang Li, Gabriele Farina, Nicola Gatti, and Tuomas Sandholm. Efficient representations for team and imperfect-recall equilibrium computation. *Artificial Intelligence*, 356:104529, 2026.

Andrea Celli and Nicola Gatti. Computational results for extensive-form adversarial team games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Jirí Cermák, Branislav Bosanský, and Michal Pechoucek. Combining incremental strategy generation and branch and bound search for computing maxmin strategies in imperfect recall games. In *AAMAS*, pp. 902–910. ACM, 2017.

Jirí Cermák, Branislav Bosanský, Karel Horák, Viliam Lisý, and Michal Pechoucek. Approximating maxmin strategies in imperfect recall games using a-loss recall property. *Int. J. Approx. Reason.*, 93:290–326, 2018.

Gabriele Farina, Christian Kroer, and Tuomas Sandholm. Faster game solving via predictive Blackwell approachability: Connecting regret matching and mirror descent. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 5363–5371, 2021.

Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, and Haipeng Luo. Regret Matching+: (In)Stability and Fast Convergence in Games. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=nYgs0qZJ97.

Sam Ganzfried and Tuomas Sandholm. Potential-aware imperfect-recall abstraction with earth mover's distance in imperfect-information games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

Hugo Gimbert, Soumyajit Paul, and B. Srivathsan. A bridge between polynomial optimization and games with imperfect recall. In *AAMAS*, pp. 456–464. International Foundation for Autonomous Agents and Multiagent Systems, 2020.

Paul J Goulart and Yuwen Chen. Clarabel: An interior-point solver for conic programs with quadratic objectives. *Mathematical Programming Computation*, pp. 1–83, 2026.

David Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific Journal of Mathematics*, 132(1):35–62, 1988.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Daphne Koller and Nimrod Megiddo. The complexity of two-person zero-sum games in extensive form. *Games and economic behavior*, 4(4):528–552, 1992.

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.

Marc Lanctot, Richard G. Gibson, Neil Burch, Martin Zinkevich, and Michael Bowling. No-regret learning in extensive-form games with imperfect recall. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, pp. 65–72. icml.cc / Omnipress, 2012.

Rida Laraki and Jean B Lasserre. Semidefinite programming for min–max problems and games. *Mathematical programming*, 131(1):305–332, 2012.

Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.

Vincent Léon, Iosif Sakos, Ryann Sim, and Antonios Varvitsiotis. Certifying concavity and monotonicity in games via sum-of-squares hierarchies. *CoRR*, abs/2512.10292, 2025.

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.

Boning Li, Yu Chen, and Longbo Huang. Correlated chance sampling for monte carlo counterfactual regret minimization. *arXiv preprint arXiv:2607.27035*, 2026.

Jiawang Nie and Xindong Tang. Nash equilibrium problems of polynomials. *Mathematics of Operations Research*, 49(2):1065–1090, 2024.

Brendan O'Donoghue, Eric Chu, Neal Parikh, and Stephen Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding, 2016. URL https://arxiv.org/abs/1312.3039.

Pablo A Parrilo. Polynomial games and sum of squares optimization. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pp. 2855–2860. IEEE, 2006.

Helfried Peyrl and Pablo A Parrilo. Computing sum of squares decompositions with rational coefficients. *Theoretical Computer Science*, 409(2):269–281, 2008.

Michele Piccione and Ariel Rubinstein. On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3–24, 1997.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Conference on Learning Theory*, pp. 993–1019. PMLR, 2013.

Oskari Tammelin. Solving large imperfect information games using CFR+. *arXiv preprint arXiv:1407.5042*, 2014.

Emanuel Tewolde, Caspar Oesterheld, Vincent Conitzer, and Paul W Goldberg. The computational complexity of single-player imperfect-recall games. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pp. 2878–2887, 2023.

Emanuel Tewolde, Brian Hu Zhang, Caspar Oesterheld, Manolis Zampetakis, Tuomas Sandholm, Paul Goldberg, and Vincent Conitzer. Imperfect-recall games: equilibrium concepts and their complexity. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 2994–3004, 2024.

Emanuel Tewolde, Brian Hu Zhang, Ioannis Anagnostides, Tuomas Sandholm, and Vincent Conitzer. Decision making under imperfect recall: Algorithms and benchmarks. *arXiv preprint arXiv:2602.15252*, 2026.

Kevin Waugh, Martin Zinkevich, Michael Johanson, Morgan Kan, David Schnizlein, and Michael H Bowling. A practical use of imperfect recall. In *SARA*, 2009.

Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. Team belief dag: generalizing the sequence form to team games for fast computation of correlated team max-min equilibria via regret minimization. In *International Conference on Machine Learning*, pp. 40996–41018. PMLR, 2023.

Rui Zheng, Ryann Sim, and Antonios Varvitsiotis. Solving imperfect-recall games via sum-of-squares optimization. In *Forty-third International Conference on Machine Learning*, 2026.

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 MOMENT–SOS BACKGROUND

For a degree bound d, let $b_d(\mu)$ be the vector of monomials μ^α with $|\alpha| \le d$, of size $\mathcal{B} = \binom{n+d}{d}$. A polynomial σ is a *sum of squares* (SOS) of degree $\le 2d$ iff $\sigma = b_d^\top G\, b_d$ for some $G \succeq 0$. The Lasserre/moment hierarchy (Lasserre, 2001) bounds $\max_\Delta u$ from above by semidefinite programs indexed by the level d; the smallest admissible level for (1) is the base level $d_0 = \max\{1, \lceil \deg u/2 \rceil\}$, with $d_0 = 1$ for constant u (then $\varepsilon = 0$ is an exact certificate at any level). On the moment side the variable is a pseudo-moment vector y constrained by $y_0 = 1$, $M_d(y) \succeq 0$, and localizing matrices $M_{d-1}(\mu_v y) \succeq 0$ for $\mu_v \ge 0$; on the dual side a bound $u \le \gamma$ is witnessed by an explicit identity expressing $\gamma - u$ inside a cone of manifestly nonnegative polynomials. It is this dual view, the *certificate*, that the main text constructs, prunes, verifies numerically, and reads attributions from.

A.2 TWO EXTENDED REMARKS ON PRUNING

Pruning is lossless whenever an exact certificate satisfies strict complementarity and unconditionally at level 1; it is asymptotically lossless in general and possibly lossy at interior or mixed points of degree ≥ 3 (Propositions 2–4, Example 1, Remark 5). In generic position the stationary-point identities enter the base level only for odd degree: for even degree $\deg q_{a,a'} = \deg u + 1$ exceeds $2d_0 = \deg u$, while for odd degree the unmultiplied equalities already fit at d_0 (Remark 4).

A.3 PROOF OF PROPOSITION 1

Proof. (i) On Δ: SOS terms are nonnegative, $\mu_v \ge 0$ makes the localizing terms nonnegative, the simplex terms vanish identically, and $t_\alpha \mu^\alpha \ge 0$ on the nonnegative orthant.

(ii) We reduce to a full-dimensional simplex product before applying Handelman. Fix a distinguished action $a_j^\star$ of each block B_j and collect the remaining actions into $W_j = B_j \setminus \{a_j^\star\}$, with $N = \sum_j |W_j| = n - \ell$; a block with a single action has $W_j = \emptyset$ and is treated by the same substitution below, its Handelman factor being the constant 1 (if $N = 0$ the domain is a point, $\mathcal{P} = \mathbb{R}^0$, and the substitution below already yields $p = \tilde{p} + \sum_j c_j h_j$ with $\tilde{p}$ a positive constant). For $\mu \in \Delta$ set $x_v = \mu_v$ for $v \in \cup_j W_j$; the simplex equality forces $\mu_{a_j^\star} = 1 - \sum_{v \in W_j} x_v$, so the map $\mu \mapsto x$ identifies Δ with the compact polytope $\mathcal{P} = \prod_j \{\, x_v \ge 0 \ (v \in W_j),\ 1 - \sum_{v \in W_j} x_v \ge 0 \,\}$, which has nonempty interior in $\mathbb{R}^N$. Substituting the distinguished variables turns p into a polynomial $\tilde{p}$ in x, and $\tilde{p} > 0$ on $\mathcal{P}$ because $p > 0$ on Δ. Since $\mathcal{P}$ is full-dimensional, Handelman's theorem (Handelman, 1988) applies directly to its defining affine functions $\{x_v\}_v$ and $\{1 - \sum_{v \in W_j} x_v\}_j$: there are nonnegative coefficients $d_{\kappa m}$ and a finite maximal degree D with

$$\tilde{p}(x) \;=\; \sum_{\kappa, m} d_{\kappa m} \prod_v x_v^{\kappa_v} \prod_j \Big(1 - \textstyle\sum_{v \in W_j} x_v\Big)^{m_j}.$$

We now lift this representation to $\mathbb{R}[\mu]$. Write $h_j = \sum_{a \in B_j} \mu_a - 1$ for the simplex linear form of block j; $\tilde{p}$ is obtained from p by substituting $\mu_{a_j^\star} = 1 - \sum_{v \in W_j} x_v$, so substituting back and using $\sum_{a \in B_j} \mu_a = 1 + h_j$ shows that the difference $p - \tilde{p}(x(\mu))$ lies in the simplex ideal $\langle h_1, \dots, h_\ell \rangle \subset \mathbb{R}[\mu]$; more precisely, since each h_j is monic of degree 1 in the distinguished variable $\mu_{a_j^\star}$, division by these linear forms yields $p - \tilde{p}(x(\mu)) = \sum_j q_j h_j$ with $\deg q_j \le \deg p - 1$. Inside the ring, the equality $1 - \sum_{v \in W_j} \mu_v = \mu_{a_j^\star} - h_j$ holds identically, so expanding $\prod_j (1 - \sum_{v \in W_j} \mu_v)^{m_j} = \prod_j (\mu_{a_j^\star} - h_j)^{m_j}$ separates each Handelman term into its leading monomial $\prod_j \mu_{a_j^\star}^{m_j}$ (coefficient $d_{\kappa m} \ge 0$) plus a remainder in which every summand carries at least one factor h_j and hence lies in the simplex ideal. Therefore the expansion lifts to the identity

$$p \;=\; \underbrace{\textstyle\sum_{\kappa, m} d_{\kappa m} \prod_v \mu_v^{\kappa_v} \prod_j \mu_{a_j^\star}^{m_j}}_{Q} \;+\; \sum_{j=1}^{\ell} c_j\, h_j, \qquad \deg Q \le D', \quad \deg c_j \le D' - 1,$$

where $D' = \max\{D, \deg p\}$ combines the maximal degree D of the Handelman representation with the degree of p itself, Q is a nonnegative combination over monomials of degree at most D' (the t-block), and each c_j collects the two ideal remainders above. Taking $2d \geq D'$ places Q in the t-block and c_j in $\mathbb{R}[\mu]_{2d-1}$, so $p \in \mathcal{C}_d$ without SOS terms or localizing multipliers. □

A.4 Proof of Lemma 1

Proof. Fix $\mu \in \Delta$ and write $s(\mu)$ for the right-hand side of (3) evaluated with the numerical data $(G_0, \{H_v\}, \{c_j\}, t^+)$. By definition of r,

$$\hat{u} + \varepsilon - u(\mu) \;=\; s(\mu) - r(\mu). \tag{5}$$

We lower-bound each term of $s(\mu)$ on Δ.

Gram terms. Every monomial satisfies $\mu^\alpha \in [0,1]$ on $\Delta \subset [0,1]^n$, hence $\|b_d(\mu)\|^2 = \sum_{|\alpha|\leq d} \mu^{2\alpha} \leq \mathcal{B}$. For any symmetric G with basis size $\mathcal{B}_G$, $b^\top G\, b \geq \lambda_{\min}(G)\,\|b\|^2 \geq -\big|\lambda_{\min}(G)\big|_-\,\mathcal{B}_G$. For the localizing terms this bound is multiplied by $\mu_v \in [0,1]$, which can only shrink its magnitude. Summing, $b_d^\top G_0 b_d + \sum_v \mu_v\, b_{d-1}^\top H_v b_{d-1} \geq -\pi_\lambda$.

Simplex terms. $\sum_{a\in B_j} \mu_a - 1 = 0$ on Δ, so these terms vanish exactly, regardless of the numerical c_j.

Monomial multipliers. $t_\alpha^+ \geq 0$ and $\mu^\alpha \geq 0$, so $\sum_\alpha t_\alpha^+ \mu^\alpha \geq 0$.

Therefore $s(\mu) \geq -\pi_\lambda$, and $|r(\mu)| \leq \sum_\alpha |r_\alpha|\,|\mu^\alpha| \leq \rho$ since $|\mu^\alpha| \leq 1$. Substituting into (5),

$$\hat{u} + \varepsilon - u(\mu) \;\geq\; -\pi_\lambda - \rho, \qquad \text{i.e.} \qquad u(\mu) \;\leq\; \hat{u} + \varepsilon + \rho + \pi_\lambda \;\leq\; \hat{u} + \varepsilon_{\mathrm{rig}}. \qquad \square$$

A.5 The posterior bound is itself a certificate in the same cone

Lemma 1 proves a pointwise bound. The next statement shows that the repaired polynomial $\hat{u} + \varepsilon_{\mathrm{rig}} - u$ also lies in the *pruned* cone with the same kept set V and the same level, so $\varepsilon_{\mathrm{rig}}$ is the objective value of a feasible point of (P_{cert}); this is what makes the numerical output an upper bound on the pruned optimum (used in §A.9) and shows that both programs are feasible at every level. Write $h_j = \sum_{a\in B_j} \mu_a - 1$ and let $\mathcal{T}_d \subset \mathcal{C}_d$ denote the sub-cone of polynomials $\sum_{|\alpha|\leq 2d} t_\alpha \mu^\alpha + \sum_j c_j h_j$ with $t \geq 0$ and $\deg c_j \leq 2d-1$ (no SOS and no localizing terms).

Lemma 3 (Complements of monomials)**.** *For every monomial with $1 \leq |\alpha| \leq 2d$, $1 - \mu^\alpha \in \mathcal{T}_d$, with ideal multipliers of degree $\leq |\alpha| - 1$. For every variable v in any block and every $|\beta| \leq 2d-1$, $\mu_v(1 - \mu^\beta) \in \mathcal{T}_d$.*

Proof. For a single variable v in block j, $1 - \mu_v = \sum_{a\in B_j\setminus\{v\}} \mu_a - h_j$, a nonnegative combination of monomials plus a constant multiple of h_j. Writing $\mu^\alpha = \mu_{v_1}\mu_{v_2}\cdots\mu_{v_k}$ with repetitions,

$$1 - \mu^\alpha \;=\; \sum_{i=1}^{k} \mu_{v_1}\cdots\mu_{v_{i-1}}\,(1 - \mu_{v_i}),$$

and each summand is a nonnegative combination of monomials of degree $\leq i \leq 2d$ minus $\mu_{v_1}\cdots\mu_{v_{i-1}} h_{j(v_i)}$, an ideal term whose multiplier has degree $i-1 \leq |\alpha|-1$. For $\beta = 0$ the second claim is trivial ($\mu_v(1-\mu^0) = 0$); otherwise multiplying the representation of $1 - \mu^\beta$ by μ_v raises every degree by one: monomials of degree $\leq 2d$ and ideal multipliers of degree $\leq |\beta| \leq 2d-1$. □

Corollary 1 (Realizability of the posterior bound)**.** *In the setting of Lemma 1, let $\lambda_0^- \leq \lambda_{\min}(G_0)$ and $\lambda_v^- \leq \lambda_{\min}(H_v)$ be any lower bounds and put $\pi'_\lambda = |\lambda_0^-|_-\mathcal{B} + \sum_{v\in V} |\lambda_v^-|_-\mathcal{L} \geq \pi_\lambda$ and $\varepsilon_{\mathrm{rig}}(\lambda^-) = \varepsilon + \rho + \pi'_\lambda + \pi_t$, which equals $\varepsilon_{\mathrm{rig}}$ when λ_0^-, λ_v^- are the exact minimal eigenvalues and is the quantity the offline verification computes (Appendix B.3). Then $\hat{u} + \varepsilon_{\mathrm{rig}}(\lambda^-) - u$ admits a representation* (3) *with PSD Gram matrices, $t \geq 0$, kept set V, and level d. In particular $\varepsilon_d^{\mathrm{pruned}} \leq \varepsilon_{\mathrm{rig}}(\lambda^-)$ for every numerical output and every choice of lower bounds, and both (P_{cert}) and the full program are feasible at every level $d \geq d_0$, with $\varepsilon = \sum_\alpha |u_\alpha| - \hat{u}$ feasible for both.*

Proof. Set $G_0' = G_0 + |\lambda_0^-|_- I \succeq 0$, $H_v' = H_v + |\lambda_v^-|_- I \succeq 0$, and $t' = t^+$. By the definition of r,

$$\hat{u} + \varepsilon - u = b_d^\top G_0' b_d + \sum_{v \in V} \mu_v \, b_{d-1}^\top H_v' b_{d-1} + \sum_j c_j h_j + \sum_\alpha t_\alpha^+ \mu^\alpha - |\lambda_0^-|_- \|b_d\|^2 - \sum_{v \in V} |\lambda_v^-|_- \, \mu_v \|b_{d-1}\|^2 - r,$$

so $\hat{u} + \varepsilon_{\mathrm{rig}}(\lambda^-) - u$ equals the first four terms plus

$$E \;=\; \big(\rho - r\big) + |\lambda_0^-|_- \sum_{|\alpha| \le d} \big(1 - \mu^{2\alpha}\big) + \sum_{v \in V} |\lambda_v^-|_- \Big(\mathcal{L}\,(1 - \mu_v) + \mu_v \sum_{|\beta| \le d-1} \big(1 - \mu^{2\beta}\big) \Big) + \pi_t,$$

using $\mathcal{B} = \sum_{|\alpha| \le d} 1$ and $\mathcal{L} = \sum_{|\beta| \le d-1} 1$. Each piece lies in $\mathcal{T}_d$: $\rho - r = \sum_{r_\alpha > 0} r_\alpha (1 - \mu^\alpha) + \sum_{r_\alpha < 0} |r_\alpha| (1 + \mu^\alpha)$ by Lemma 3 (the constant coefficient contributes $|r_0| - r_0 \ge 0$); the remaining sums are nonnegative multiples of $1 - \mu^{2\alpha}$, $1 - \mu_v$, and $\mu_v(1 - \mu^{2\beta})$ with $2|\alpha| \le 2d$ and $2|\beta| \le 2d - 2$, covered by the lemma; $\pi_t \ge 0$ is a constant. Adding the t- and ideal parts of E to t' and c_j keeps $t' \ge 0$ and $\deg c_j \le 2d - 1$, which is the claimed representation. For the feasibility statement, $\hat{u} + \varepsilon - u = \sum_\alpha |u_\alpha| - u = (|u_0| - u_0) + \sum_{u_\alpha > 0,\, \alpha \ne 0} u_\alpha (1 - \mu^\alpha) + \sum_{u_\alpha < 0,\, \alpha \ne 0} |u_\alpha| (1 + \mu^\alpha) \in \mathcal{T}_d$, again by Lemma 3, since $\deg u \le 2d$. □

A.6 Proofs for losslessness (Section 5)

Proof of Lemma 2. The left side of (4) vanishes at $\hat{\mu}$, the simplex terms vanish identically on Δ, and every remaining summand is nonnegative at $\hat{\mu}$ (SOS values, $\hat{\mu}_v \ge 0$, $t_\alpha \ge 0$, $\hat{\mu}^\alpha \ge 0$). A zero sum of nonnegative terms vanishes termwise; for $v \in \mathrm{supp}(\hat{\mu})$ the factor $\hat{\mu}_v > 0$ can be divided out. □

Proof of Proposition 2. Deleting identically-zero blocks from (4) leaves a feasible point of (P_{cert}) with $\varepsilon = 0$; conversely the pruned cone is contained in the full cone, so $\varepsilon_d^{\mathrm{pruned}} \ge \varepsilon_d^{\mathrm{full}} \ge 0$. □

Proof of Proposition 3. At level $d = 1$ each localizing multiplier σ_v is an SOS of degree $2(d-1) = 0$, that is a constant $c_v \ge 0$, and the t-block ranges over all monomials of degree at most $2d = 2$, which include every μ_v. Given a full certificate with multipliers $\{c_v\}$ and t-block t, set $t'_{e_v} = t_{e_v} + c_v$ for every v and $t'_\alpha = t_\alpha$ otherwise; then $t' \ge 0$ and $\sum_v c_v \mu_v + \sum_\alpha t_\alpha \mu^\alpha = \sum_\alpha t'_\alpha \mu^\alpha$, so the same polynomial is represented with $\sigma_v \equiv 0$ for all v; this representation is a feasible point of (P_{cert}) with the same ε. The reverse containment is the inclusion of the pruned cone in the full cone. Hence both programs have the same feasible values of ε and the same optimum, for every $\hat{\mu}$ and without any exactness assumption. Strict complementarity (Definition 2) is then automatic at level 1. □

Proof of Proposition 4. Since $\hat{\mu}$ is a global maximizer, $u(\hat{\mu}) + \eta - u \ge \eta > 0$ on Δ. By Proposition 1(ii) (Handelman), that polynomial lies in the certificate cone at some finite level using only the t-block and the simplex multipliers; no σ_v and no SOS term are needed in the containment. The pruned cone contains this representation. □

A.7 Proof of Proposition 5

Proof. (i) Each σ_v is SOS, $\mu_v \ge 0$ on Δ, $t \ge 0$, and monomials are nonnegative on Δ; every group in the definition is a sum of such terms. (ii) Evaluate the certificate identity at μ; the simplex terms vanish on Δ and the remaining terms are exactly $\sum_j \Gamma_j + \Gamma_\times + \mathrm{res}$. (iii) Immediate. □

A.8 Extended remarks on pruning

Remark 3 (The range of pruning outcomes). Propositions 3 and 4 bracket the phenomenon: pruning is lossless at level 1 for every stationary point, and lossless in the limit for all utilities. Example 1 shows loss at an *interior* stationary point, where complementary slackness deletes every multiplier and the certified price at level d_0 is large when $\deg u \ge 3$; Remark 5 covers the mixed point.

The ablation reported in §6 matches this picture. On the 57 vertex- and mixed-support points outside the 18 SCS-solved c-family instances (basis size 330, where solver accuracy dominates), the solver-level gap difference $|\varepsilon_{d_0}^{\mathrm{pruned}} - \varepsilon_{d_0}^{\mathrm{full}}|$ never exceeds 2.5×10^{-5} (27 vertex + 30 mixed). The interior

quadratic instance is lossless, as Proposition 3 predicts. The one lossy case in the pool is the interior cubic of Example 1, which one level of the hierarchy closes $33\times$ in the experiment, consistent with the asymptotic losslessness of Proposition 4.

Structurally, the number of deleted blocks equals the support size. Boundary points, which first-order methods gravitate toward, therefore prune the least and are the safest in our experiments; fully interior points delete everything and rely on the Handelman-type remainder cone alone.

Remark 4 (Adding the stationary-point identities at the base level). An alternative to pruning is to *add* structure. At any stationary point, complementarity across two actions a, a' in the same block yields the multiplier-free polynomial identities

$$q_{a,a'}(\mu) \;=\; \mu_a\,\mu_{a'}\big(\partial_a u(\mu) - \partial_{a'} u(\mu)\big) \;=\; 0, \qquad \deg q_{a,a'} \;\le\; \deg u + 1.$$

The bound is sharp in generic position but can drop under cancellation: for $u = f(\mu_a + \mu_{a'})$ the factor $\partial_a u - \partial_{a'} u$ vanishes identically and $q_{a,a'} \equiv 0$. A global maximizer $\mu^\star$ is a KKT point: for a, a' in one block with $\mu^\star_{a'} > 0$, moving mass from a' to a is a feasible direction, so $\partial_a u(\mu^\star) \le \partial_{a'} u(\mu^\star)$, and equality holds whenever both are supported; hence $q_{a,a'}(\mu^\star) = 0$. Adding the moment equalities $\Lambda_y(q_{a,a'} \cdot m) = 0$ keeps the relaxation valid. However, these equalities require $2d \ge \deg q_{a,a'} + \deg m$ to admit a multiplier monomial m, while the base level satisfies $2d_0 = \deg u$ for even $\deg u \ge 2$, and $2d_0 = \deg u + 1$ for odd degree. Hence for even-degree utilities in generic position ($\deg q_{a,a'} = \deg u + 1$) no constraint enters at level d_0: no multiplier monomial, not even $m = 1$, fits. For odd degree the situation differs: $2d_0 = \deg u + 1$, so the *unmultiplied* equalities $\Lambda_y(q_{a,a'}) = 0$ already fit at d_0 and, whenever $q_{a,a'} \not\equiv 0$, are nontrivial constraints of degree at most $2d_0$. In generic position $\deg q_{a,a'} = 2d_0$, and then no nonconstant multiple fits at level d_0: those first appear at $d_0 + 1$. If cancellation lowers $\deg q_{a,a'}$, every multiplier of degree at most $2d_0 - \deg q_{a,a'}$ already fits at the base level. In our experiments, the strengthened arm nonetheless performs on par with the bare relaxation at the base level. The unmultiplied equalities act on a low-dimensional moment subspace at d_0, and the surrounding SOS and monomial constraints appear to absorb their effect.

A.9 The interior-lossy example is exact

Both claims about Example 1 are established here in exact arithmetic. The instance has one block with two actions; write μ_c for the action whose mass enters u quadratically, so on the simplex $\mu_c + \mu_e = 1$ we have $u = 27\,\mu_c^2\mu_e$. At the interior point $\hat\mu = (\mu_c, \mu_e) = (2/3, 1/3)$ this gives $u(\hat\mu) = 4$.

An exact full certificate exists. The identity

$$4 - u \;=\; \mu_e\,(3\mu_c - 2)^2 \;+\; \mu_c\,(6\mu_c - 4)^2 \;+\; \big(12\mu_c - 36\mu_c^2 - 4\big)\,(\mu_c + \mu_e - 1) \tag{6}$$

holds by expansion. The first two terms are localizing multipliers $\mu_v\sigma_v$ with $\sigma_c = (6\mu_c - 4)^2$ and $\sigma_e = (3\mu_c - 2)^2$ squares of degree-1 polynomials, admissible at level $d_0 = 2$; the last term is a simplex-ideal multiplier of degree $2 \le 2d_0 - 1$. Hence $\gamma = 4$ is feasible in the *full* (unpruned) level-2 certificate program with $\varepsilon = 0$; both localizing terms are used, so this is not a point of (P_{cert}), whose kept set is empty at $\hat\mu$. Since $\hat\mu$ is feasible and attains $u(\hat\mu) = 4$, the level-d_0 moment bound equals 4 exactly; the 6×10^{-8} residual of the solver is approximation error.

No exact pruned certificate exists. Deleting both localizing blocks (the support is the full simplex) leaves the cone of $\sigma_0 + c(\mu_c + \mu_e - 1) + \sum_{|\alpha|\le 4} t_\alpha \mu^\alpha$ with σ_0 an SOS of degree ≤ 4 and $t_\alpha \ge 0$. Restricting to the simplex line $(\mu_c, \mu_e) = (x, 1 - x)$, on which the ideal term vanishes, suppose $4 - u = 27x^3 - 27x^2 + 4$ admitted such a representation. At $x = 2/3$ the left side is zero, while each term on the right is nonnegative there ($\sigma_0 \ge 0$ everywhere, and each monomial $x^{\alpha_c}(1 - x)^{\alpha_e}$ is positive at $x = 2/3 \in (0, 1)$ with $t_\alpha \ge 0$). A vanishing sum of nonnegative terms vanishes termwise, so $t_\alpha = 0$ for every α and $\sigma_0(x, 1 - x)$ vanishes at $x = 2/3$; the identity then reads $\sigma_0(x, 1-x) = 27x^3 - 27x^2 + 4$ for all $x \in \mathbb{R}$. But an SOS polynomial is nonnegative on all of $\mathbb{R}^2$, hence on this affine line, while the cubic on the right diverges to $-\infty$ as $x \to -\infty$. Contradiction. So $\varepsilon = 0$ is infeasible in the level-d_0 pruned program. Nothing in this argument uses the level: at any level d the fully pruned cone still consists of an SOS σ_0, an ideal multiple, and nonnegative monomial multiples, all positive at $x = 2/3$, so $\varepsilon = 0$ is infeasible at every level. Proposition 4 is therefore tight in the sense that raising the level closes this gap to any $\varepsilon > 0$ but never to 0.

The pruned optimum is bounded away from zero. Infeasibility of $\varepsilon = 0$ does not by itself exclude an infimum of 0, since the truncated cone need not be closed. A dual functional settles this. Let Λ be the linear functional on $\mathbb{R}[x]_{\le 4}$ with $\Lambda(x^k) = y_k$ for the pseudo-moments

$$(y_0, y_1, y_2, y_3, y_4) = \Big(1,\ \tfrac{3191}{5000},\ \tfrac{2073}{5000},\ \tfrac{2601}{10000},\ \tfrac{1747}{10000}\Big),$$

and define $\Lambda_2(p) = \Lambda\big(p(x, 1-x)\big)$ for $p \in \mathbb{R}[\mu_c, \mu_e]_{\le 4}$. Three facts are checked by exact rational arithmetic. (a) The Hankel matrix $(y_{i+j})_{0 \le i,j \le 2}$ has leading principal minors 1, $182519/(25 \times 10^6)$, and $13129/(5 \times 10^{10})$, so it is positive definite, and hence $\Lambda_2(\sigma_0) \ge 0$ for every SOS σ_0 of degree ≤ 4. Indeed $\sigma_0(x, 1-x)$ is a univariate SOS $\sum_i q_i(x)^2$ with $\deg q_i \le 2$, and $\Lambda(q_i^2)$ is the Hankel quadratic form in the coefficients of q_i. (b) $\Lambda\big(x^a(1-x)^b\big) \ge 0$ for all $a + b \le 4$, so $\Lambda_2(t_\alpha \mu^\alpha) \ge 0$; the fifteen values, indexed by (a, b), are $(0,0)\colon 1$, $(0,1)\colon \frac{1809}{5000}$, $(0,2)\colon \frac{691}{5000}$, $(0,3)\colon \frac{691}{10000}$, $(0,4)\colon \frac{691}{10000}$, $(1,0)\colon \frac{3191}{5000}$, $(1,1)\colon \frac{559}{2500}$, $(1,2)\colon \frac{691}{10000}$, $(1,3)\colon 0$, $(2,0)\colon \frac{2073}{5000}$, $(2,1)\colon \frac{309}{2000}$, $(2,2)\colon \frac{691}{10000}$, $(3,0)\colon \frac{2601}{10000}$, $(3,1)\colon \frac{427}{5000}$, $(4,0)\colon \frac{1747}{10000}$. (c) Λ_2 annihilates the ideal term, as $\mu_c + \mu_e - 1$ vanishes identically on the line. Applying Λ_2 to the pruned identity $4 + \varepsilon - u = \sigma_0 + c\,(\mu_c + \mu_e - 1) + \sum_\alpha t_\alpha \mu^\alpha$ of any feasible point, and using $\Lambda_2(1) = 1$ and $\Lambda_2(u) = 27\,(y_2 - y_3)$, gives

$$\varepsilon \ \ge\ 27\,(y_2 - y_3) - 4 \ =\ \frac{343}{2000} \ =\ 0.1715.$$

The re-verified pruned certificate of the experiment (Appendix B.3) has certified posterior gap 0.1718. By Corollary 1 that value comes from a feasible point of the same pruned program, so the level-d_0 pruned optimum lies between $343/2000$ and 0.1718. The gap reported as 0.172 in the main text is this certified value rounded to three decimals.

Remark 5 (The same loss at a mixed point). The loss of Example 1 reappears once the interior block is embedded in a mixed stationary point. Add to the extended absent-minded driver a second infoset with two actions and let $U(\mu_c, \mu_e, \mu_z, \mu_w) = 27\mu_c^2\mu_e - \mu_z$ on $\{\mu_c + \mu_e = 1\} \times \{\mu_z + \mu_w = 1\}$. The maximizer is $(2/3,\ 1/3,\ 0,\ 1)$ with value 4: the first block is interior and the second sits at a vertex, so the point is mixed and μ_z is its only off-support variable. A full exact level-2 certificate is the identity of this section for $4 - 27\mu_c^2\mu_e$ plus the single t-block term μ_z, so $\varepsilon_2^{\text{full}} = 0$. Pruning keeps only the multiplier of μ_z. Restrict any pruned certificate to the face $\mu_z = 0$, $\mu_w = 1$: the μ_z multiplier and the second simplex term vanish, and what remains is a fully pruned level-2 certificate for the cubic on the first simplex with the same ε. By the argument above no such certificate is exact, and the functional Λ_2 extended by $\mu_z \mapsto 0$, $\mu_w \mapsto 1$ gives $\varepsilon_2^{\text{pruned}} \ge 343/2000$ again. The loss is produced by the interior block alone, and adding vertex blocks cannot remove it.

B Experiment Details

B.1 Census and coverage tables

Table 1 and Table 2 in the main text aggregate the pool by instance family and by stationary-point location; the two tables below, Tables 3 and 4, account for the full 192-instance census by generator family. Tewolde *et al.* (Tewolde et al., 2026) report 61 released benchmark instances. Their files are not reachable from our compute environment, so we regenerate the simulation, detection, and random instances locally from the same generators under fixed seeds, and use the five canonical simple examples of the benchmark family. The census is therefore an independently regenerated pool whose instances need not coincide with the released files.

Across the five solvers of Table 1, RM and RM$^+$ reproduce the PRM$^+$ certified set exactly, as does projected gradient descent in the label sense. On the extended absent-minded driver its iterate hits the iteration budget at a non-stationary point (KKT gap 1.39; value 3.97 against 4), but PRM$^+$ already withholds that instance, so no label changes. OGD deviates from the certified set on two otherwise-certified instances, the row-4 and row-5 simulation-b revisits. There its best iterate stops at a non-stationary point (KKT gaps 0.60 and 0.37), with values 1.2% and 0.7% below the PRM$^+$ optimum, so the label changes there. On the extended absent-minded driver it collapses to value 0 against 4, already gap-unclosed under PRM$^+$, so the rate is unchanged there.

SCIP (Bestuzheva et al., 2023), run by spatial branch and bound on the bilinear reach-variable reformulation of the benchmark's exact baseline with a one-hour limit, declares optimality of 52

Table 3: Census and coverage. An instance is skipped when its level-d_0 moment basis exceeds 1500 (the utility degree, and hence the basis, grows with the number of rounds; random and most detection instances exceed the cap already at d_0). Coverage is reported explicitly.

family	instances	CLARABEL ($\mathcal{B} \leq 300$)	SCS ($\mathcal{B} \leq 1500$)	skipped
simulation	59	36	22	1
detection	72	18	9	45
random	56	0	0	56
simple	5	5	0	0
total census	192	59	31	102
feasible: main pool	77	59	18	0
feasible: scale (lg/xs)	13	0	13	0

Table 4: Per-family accounting. The pool rate is conditional on main-pool membership (77 of the 90 instances within the cap; the other 13 form the scale set); the census coverage is unconditional over all 192 instances. random and most detection instances exceed the 1500 basis cap already at d_0 (Table 3), so their certified counts reflect the method's coverage at this cap.

family	census	pool	certified	pool rate	census coverage
simulation	59	54	48	88.9%	81.4%
detection	72	18	18	100%	25.0%
random	56	0	0	n/a	0%
simple	5	5	4	80%	80.0%
total census	192	77	70	90.9%	36.5%

certified-global instances within one minute and 69 within one hour (median 0.6s), and of all 7 gap-unclosed instances; only one simulation-c instance hits the limit. There, at 3612s, its proven dual bound is 7.39483 against $u(\hat{\mu}) = 7.39454$, a relative excess of 3.95×10^{-5}, below the 10^{-4} tolerance of §6 (at 3575s the excess was still 3.9×10^{-4}), while SCIP's own zero-gap termination is not reached. The *SCIP confirms* column of Table 1 therefore applies our tolerance to SCIP's proven bound, the criterion under which the two instruments are comparable; SCIP's bounds are floating-point and not enclosed (Remark 2).

B.2 Instance manifest and run protocol

The 192-instance census is the union of the five simple examples and of simulation, random, and detection instances regenerated locally from the same generators under fixed seeds (Section B.1). A regeneration seeds the generator with one integer and leaves it otherwise untouched, so a game file is a deterministic function of its seed and generator flags, and every game file records its full generator parameter block in its header. The seed schedule over the type–size grid is $1000\,\tau + 100\,s + i$, for type index τ, size index s, and game index $i \in \{1, 2, 3\}$; Table 5 lists the parameters per size tier.

Every size tier also has five *revisit* instances: further seeded draws under the same parameters, draw r using seed $10000\,r + 1000\,\tau + 100\,s + i$. A simulation revisit redraws the payoffs on the same tree shape, and a detection or random revisit redraws the graph or tree as well, so their utility degrees and moment bases match the base tier.

An instance is *feasible* when its level-d_0 moment basis is at most 1500 (Table 3): 90 of the 192 pass, split into the 77 of the main pool (the 59 CLARABEL instances plus the 18 SCS instances) and the 13 of the scale set. On the 59 CLARABEL instances the unpruned moment-side bound, the tightest numerical baseline, matches $u(\hat{\mu})$ to 1.5×10^{-7} with median 1.3×10^{-9}. In Table 1 the simple row covers the absent-minded driver; the extended absent-minded driver is listed separately.

For the main pool, every method runs three restarts per seed and returns the best-valued of the three final iterates as its stationary point. Restart 0 starts from the uniform strategy, and restarts $1, 2$ from independent exponential draws at seeds $b + 1$ and $b + 2$, where b is the run's base seed. Each restart runs at most 2×10^5 iterations or 200 seconds, and stationarity means the block-wise gap of (2)

Table 5: Generator parameters per size tier, as recorded in the instance file headers. The seed schedule appears in the text below, and Table 3 gives the per-family counts. Degrees and level-d_0 bases are those of the resulting utilities; the revisit tiers share the row of their base tier. lg instances are degree-2 detection games with 12 community samples on graphs of 20 to 50 nodes.

type	tier	structural parameters	tree nodes	$\deg u$	$\mathcal{B}$
simulation	a	scenarios 1; 5–7 tests, 3–5 deployments	245–759	10	21
	b	scenarios 2	3302–12844	10	126
	c	scenarios 2	33648–124550	14	330
	xs	scenarios 3	105388–440718	12	924
detection	a	10 nodes; 1 sample; 2 rounds	86–108	2	42–74
	b	10 nodes; 2 samples; 3 rounds	1251–1933	3	$\geq$ 19900
	c	13 nodes; 5 samples; 3 rounds	8271–10799	3	$\geq$ 170820
	lg	20–50 nodes; 12 samples; 2 rounds	4973–30531	2	344–2060
random	a–c	depth 4–10; branching 3–5	6026–183221	7–10	$\geq$ 91390
simple		five canonical examples		2–3	3–7

Table 6: Certified counts on the main pool as the tolerance $10^{-4}\max(|u(\hat{\mu})|, 1)$ is relaxed and tightened. The verdicts of Table 1 use the 10^{-4} column.

tolerance	10^{-3}	10^{-4}	10^{-5}	10^{-6}
certified (of 77)	74	70	63	60

below 10^{-9}. Two OGD points stop non-stationary and a third converges to a worse stationary point; all three are listed in Section B.1.

The main pool uses seed groups $42, 52, \ldots, 132$. The multi-algorithm comparisons of Table 1 use groups $42, 72, 102$, where RM, RM$^+$, and PRM$^+$ are the three settings of the regret-matching implementation released with the benchmark (Tewolde et al., 2026). They differ in whether negative cumulative regrets are clipped to zero (RM$^+$ and PRM$^+$) and whether the next-period regret is predicted (PRM$^+$). PGD and OGD use a constant step of 0.1, and OGD adds the optimistic term.

Certificates solve the program of Section 4 at the base level d_0, with CLARABEL when the basis is ≤ 300 and SCS otherwise, at solver default tolerances on one core. A coordinate below the threshold $\tau = 10^{-6}$ is declared off-support; it is independent of the solvers' feasibility tolerances.

B.3 Posterior verification, certificate cost, and level-up

The verdict is stable in the aggregate across four orders of magnitude of the tolerance (Table 6), with a median relative gap of 4.4×10^{-8}. Up to 10 of the 70 labels flip as the threshold is tightened, so a small minority of instances are numerically marginal.

The seven gap-unclosed instances split into two regimes. The structural gap of the extended absent-minded driver is 0.172 at d_0, from pruning at an interior point. It contracts $33\times$ to 0.0052 at $d_0{+}1$ and a further $23\times$ to 2.2×10^{-4} at $d_0{+}2$, i.e. 5.5×10^{-5} relative to $u(\hat{\mu}) = 4$, already below the 10^{-4} threshold, consistent with the asymptotic losslessness of Proposition 4 (a statement about exact optima). Closure here is closure within tolerance: by §A.9 no fully pruned certificate attains $\varepsilon = 0$ at any level.

The six simulation-c gaps are different. They sit at 3×10^{-3} to 2×10^{-2} and are consistent with *solver precision* loss of SCS on basis-330 problems, since the moment bound matches $u(\hat{\mu})$ to 10^{-6} on these six. Whether their exact pruned optima are zero remains open, and the ablation of Appendix B.5 cannot settle it. On these instances the full-cone SCS solve is sometimes far less accurate than the pruned one and sometimes closes where the pruned one does not, so the numerical pruning comparison excludes them. Raising the level helps only when the extra accuracy survives the larger SDP. One of six closes at $d_0{+}1$ ($3.9 \times 10^{-3} \to 8.6 \times 10^{-5}$), and the others stagnate or degrade as the basis grows to 495 and SCS precision drops. Structural loss cannot be separated from SCS accuracy in this regime, so only the first regime is informative about the relaxation.

Table 7: Which certificates are enclosed. Every certificate in the paper carries the double-precision posterior bound of Lemma 1 (under one second each); only the first row is additionally re-verified using exact rational and interval arithmetic. Counts are certificates, not instances.

certificate set	count	enclosed	used for
seed-42 re-solve, one per instance	77	yes	Remark 2; Table 1 verdicts
PRM$^+$ runs, 10 seed groups	770	no	seed invariance, median $\varepsilon_{\mathrm{rig}}$
other first-order methods, 3 seed groups	924	no	RM, RM$^+$, PGD, OGD columns
pruning ablation (pruned and full re-solves)	154	no	Table 2
level-up certificates	15	no	pruning paragraph of §6, this section
scale-set certificates (lg, xs)	10	no	Figure 2 timings, §B.3
stage-2 attribution certificates	70	no	attribution faithfulness, §B.4

The posterior step of Lemma 1 adds less than one second per instance and inflates the gap by a solver-dependent amount: 2.1×10^{-8} relative on the median for CLARABEL, 6.5×10^{-5} for SCS (90th percentile 2.2×10^{-3}). Every gap-unclosed verdict in Table 1 is either the structural counterexample or an SCS-solved instance whose gap cannot be separated from solver inflation; on the CLARABEL instances the posterior gap sits at solver-level accuracy. Figure 1 shows the full picture: the CLARABEL cloud sits on the diagonal, the SCS cloud lifts off it, and exactly the points pushed across the 10^{-4} threshold are the gap-unclosed verdicts.

Scope of the rigorous re-verification. Table 7 states which certificates are enclosed and which are double-precision posterior bounds. For one certificate per instance, the seed-42 re-solve, all three ingredients of Lemma 1 are enclosed.

The residual ρ is recomputed in exact rational arithmetic. The coefficient-matching matrices are integer, and every solver entry, utility coefficient, and $\hat{u}$ is a double, hence a dyadic rational, so each residual coefficient is exact.

Each Gram block G receives a rational lower bound $\lambda_G^- < \lambda_{\min}(G)$, proven by an interval-arithmetic Cholesky factorization of $G - \lambda_G^- I$ whose pivots are all strictly positive. The verified penalty uses $|\lambda_G^-|_-$ in place of $|\lambda_{\min}(G)|_-$ in π_λ, i.e. the certified gap is $\varepsilon_{\mathrm{rig}}(\lambda^-)$ of Corollary 1. The shift λ_G^- is placed $10^{-10}\max(1, \|G\|_{\max})$ below the double-precision eigenvalue estimate, and no block required a looser shift once the working precision was raised from 40 to 100 digits on the basis-330 blocks. Naive interval Cholesky overestimates pivot widths on near-rank-deficient Gram matrices.

The utility polynomial entering the residual is re-extracted from the game file with chance probabilities and payoffs parsed as exact rationals. The double-precision extraction the solver saw differs from it by at most 7.6×10^{-13} in coefficient ℓ_1 norm, and this difference is charged to ρ, so the bound holds for the game as specified. The stored t is already clipped, and the stationary point is re-normalized onto the product of simplices in exact arithmetic; the exact $u(\hat{\mu})$ of the re-normalized point differs from the stored $\hat{u}$ by at most 4.0×10^{-15}, and this difference is added to the certified gap. That gap is then tested against the threshold $10^{-4}\max(|\hat{u}|, 1)$ of §6.

All 70 certified-global and all 7 gap-unclosed verdicts are reproduced. The certified bound exceeds the double-precision posterior bound by at most 5.4×10^{-8} absolute (8.2×10^{-9} relative, median 1.5×10^{-9}), and the closest certified call sits at 9.3×10^{-5} relative, below the 10^{-4} threshold. The check costs 24 CPU-minutes for the pool, at most 94 seconds per instance. The certificates are the outputs of one re-solve per instance at the seed-42 stationary point, whose double-precision verdicts coincide with the ten-seed labels. The per-certificate record for all 77 instances, including the interior-cubic counterexample and the closest certified call, is released with the code.

On the main pool the median certificate time is 2.4s with CLARABEL and 54s with SCS, on a single core. Figure 2 shows how cost grows with basis size: the basis-344 and basis-401 instances close in 1–8 minutes, the basis-814 and basis-843 in 8–13 minutes, and the basis-924 xs instances already exceed the fifteen-minute budget, returning an inaccurate solution whose posterior gap stays positive. Cost tracks the moment basis $\mathcal{B}$: basis-924 xs ($n_{\mathrm{var}} = 6$) costs more than the basis-843 30-node detection instance ($n_{\mathrm{var}} = 842$). The 40-node detection instances ($n_{\mathrm{var}} = 1395$, basis

1250–1396) exceed the benchmark's practical limit: the level-d_0 formulation alone costs 9 minutes and 15 GB to construct, and the SCS solve does not return within the budget.

B.4 Attribution faithfulness protocol

The protocol of §6 covers 567 perturbations (69 of the 70 certified instances, 3–24 per instance; the fully interior absent-minded driver has no off-support action and contributes none). The argmax block is correct in 100% of cases at every δ, at the instance level as well. The attributed share of the gap is 0.996 on the median, and the identity $\sum_j \Gamma_j + \Gamma_\times + \mathrm{res} = \text{gap}$ holds to below 5.3×10^{-6} absolute error (median 2.4×10^{-7}), the approximation discussed in §5. Perturbations inside a full-support block carry no localizing multiplier and land in the residual res (60 of 627); this is the documented boundary of localizing-multiplier attribution, and such rows are reported separately.

B.5 Pruning ablation details

The main text reports the medians in Table 2; this section gives the per-regime detail. On the 57 vertex and mixed instances solved by CLARABEL, the pruned and full solver gaps ε agree to 2.5×10^{-5} at worst (7×10^{-8} on the median), and the posterior gaps to 1.8×10^{-7} on the median. On the 18 SCS instances the comparison cannot separate the cones: both programs certify the five vertex instances, but on the 13 mixed ones the full-cone posterior gap is far worse on four. On three of the six gap-unclosed instances, all simulation-c revisits, the full-cone solve lands within tolerance while the pruned one does not. Since the pruned cone is contained in the full cone, this direction is either pruning loss or pruned-solve inaccuracy. Over all 75 vertex and mixed instances the median posterior-gap difference is 9×10^{-7}.

Fixed-level closure is decided by the theory at 23 of the 77 instances: all 22 quadratic instances of the pool, including the absent-minded driver, are lossless by Proposition 3, while the extended absent-minded driver loses 0.172, the interior-lossy counterexample of Example 1. This is an observation about the pool; Remark 5 shows the same loss at a mixed point.

Pruning also improves numerics. On the four SCS instances where the larger *full*-cone SDP solves least accurately, its posterior gap inflates to 0.29–1.4 while the pruned certificate stays at 3×10^{-3} to 4×10^{-2}. Deleting supported blocks thus helps the solver as well as the certificate, and it cuts the certificate cost: total wall-clock drops from 17601s to 11842s (1.49×), the pruned solve is faster on 67 of 77 instances (median 1.53× on those), and the kept basis drops by 39% (1283 → 780).

C Additional Related Work

Solving imperfect-recall games. The computational literature on imperfect recall has largely targeted two-player zero-sum maxmin equilibria, whose hardness dates to (Koller & Megiddo, 1992), via branch-and-bound and incremental strategy generation (Bosansky et al., 2017; Cermák et al., 2017; 2018). Team games have been handled through compressed representations (Carminati et al., 2026). Those methods solve a two-player zero-sum problem against a perfect-recall adversary. Our problem is a single-player ex-ante optimization, to which the polynomial optimization bridge of (Gimbert et al., 2020) applies. Those methods compute strategies in exponential worst-case time; we take a stationary point of a scalable first-order method and test it for global optimality.

Scaling first-order methods. The regret-matching family has scaled to large imperfect-information games through action abstraction, correlated chance sampling, and self-certified exploitation (Li et al., 2024; 2026; Li & Huang, 2026a); those systems produce the stationary points our layer is designed to check.

Polynomial optimization & certificates. Our certificate is a membership test in the moment-SOS hierarchy of Lasserre (Lasserre, 2001), using the Handelman representation over the product of simplices (Handelman, 1988) with a monomial multiplier block that makes the base level useful (Remark 1). Sum-of-squares certificates have also moved from optimization to equilibrium problems, beginning with polynomial games (Parrilo, 2006). Nash equilibrium problems of polynomials

are analyzed in (Nie & Tang, 2024), and hierarchies are used to certify concavity and monotonicity in games in (Léon et al., 2025).

D Additional Remarks

Remark 6 (Conservativeness of the implementation). The bound holds already with $\varepsilon + \rho + \pi_\lambda$. The reported value adds the clipping penalty π_t on top: the residual accounts for replacing t by t^+, so π_t only loosens the certificate, and the clipped mass it counts sits at solver tolerance. All quantities are computed by one sparse matrix–vector product and one eigendecomposition per Gram block, so no second solve is required and the posterior step costs well under a second per instance.