SAFE SUBGAME RESOLVING IN MULTIPLAYER GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Poker and other games with hidden information are too large to solve in one pass. Systems that play them at a superhuman level instead compute a complete strategy offline, the *blueprint*, and then re-solve the part of the game actually reached during play, refining the blueprint in real time inside that *subgame*. For two players in a zero-sum game this second stage, subgame resolving, comes with a safety theorem: refinement never increases how much an adversary can gain by exploiting the strategy. With three or more players acting independently (uncorrelated profiles) no such guarantee exists. Multiplayer systems resolve heuristically, and resolving can increase a profile's vulnerability, the largest gain any single player obtains by deviating. We prove the first conditional safety theorem for uncorrelated multiplayer subgame resolving, a post-hoc certificate for any delivered update. The vulnerability of the resolved profile σ^1 exceeds that of the blueprint σ^0 by at most four terms. The first is the activation term A_i^S, the subgame value the blueprint promised but did not realize. The other three are the structural tax $\Theta \leq 2n\,\delta_\infty$, the summed best-response margin deficits $\sum_q M_q$, and the resolver's blueprint-relative subgame suboptimality ε_i. The tax is an oscillation of the total utility, bounded by $2n$ times the distance δ_∞ to the class of constant-sum polymatrix (CSP) games, whose payoffs decompose into pairwise constant-sum interactions. Constructions attain each term, and both branches of the maximum in the sharper bound, with equality; a certified adversarial search over 162 instances finds no violation. Every rake-free poker variant is pointwise constant-sum, so the sharp structural tax vanishes there.

1 INTRODUCTION

Imperfect-information extensive-form games (IIEFGs) model sequential decision making with hidden information, from poker (Bowling et al., 2015; Brown & Sandholm, 2018) to Mahjong (Li et al., 2020) and Stratego (Perolat et al., 2022). Counterfactual regret minimization (CFR) (Zinkevich et al., 2007) and its variants (Lanctot et al., 2009; Tammelin, 2014; Brown & Sandholm, 2019a; Brown et al., 2019) are the standard route to a two-player zero-sum equilibrium, but their cost grows with the game tree. In large games the full game is therefore solved only once, offline and under abstraction, a coarsening of the game that groups similar states (Brown et al., 2015). The resulting full-game strategy, the *blueprint*, is then refined in real time inside the subgame actually reached (Burch et al., 2014; Brown et al., 2018). In two-player zero-sum games subgame resolving is safe (Burch et al., 2014; Moravcik et al., 2016; Brown & Sandholm, 2017): refinement never increases exploitability, the amount an adversary gains by best-responding, and the pipeline underpins DeepStack (Moravčík et al., 2017), Libratus (Brown & Sandholm, 2018), and Student of Games (Schmid et al., 2023).

Beyond two players no such guarantee is known. Pluribus plays six-player poker with heuristic resolving and no safety guarantee (Brown & Sandholm, 2019b). Two-player safety rests on the identity that an opponent's gain equals the resolver's loss (Von Neumann & Morgenstern, 2007). For $n > 2$ this identity disappears, and even its per-opponent weakening fails. Value can transfer across opponents, leaving the harmed opponent's margin silent while the beneficiary's margin fires; Example 2 instantiates this. Under a near-equilibrium blueprint, current resolving implementations also misreport their margins, at -0.46 and -0.74 even at zero deviation (Proposition 5), because their root measure omits the other opponents' reach.

We therefore condition safety on two computable quantities. The structural tax Θ is an oscillation of the total utility, computable before resolving and bounded by $2n$ times the distance δ_∞ to the

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.

constant-sum polymatrix (CSP) class. The per-opponent best-response margin deficit M_q is obtained one subgame best-response pass later. Margins and solve accuracy alone cannot certify safety for the profile vulnerability Vul_G, the largest gain a player obtains by deviating. Example 1 raises Vul_G with nonnegative margins and zero suboptimality, already in two-player zero-sum games. Theorem 1 states the post-hoc bound: for any resolving update σ^1 of a blueprint σ^0 by resolver i in subgame S,

$$\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0) + 2n\,\delta_\infty + \textstyle\sum_q M_q + \varepsilon_i,$$

where $A_i^S(\sigma^0)$ is the activation term, the subgame value the blueprint promised but did not realize, and ε_i is the resolver's suboptimality against the blueprint opponents inside S. A sharper structural tax built from the constant-sum defects replaces $2n\,\delta_\infty$, and the display above is its corollary (Corollary 5). Figure 1 previews the pipeline and the bound.

Our experiments tie these terms to realized resolving outcomes. Numerical checks corroborate every inequality of Theorem 1, and a certified adversarial search over 162 instances finds no theorem-level counterexample: every increase beyond tolerance is explained by a large δ_∞, a margin deficit, A_i^S, or ε_i (Section 7). The certificate is also a decision rule, and reading the blueprint's stored advantages lifts zero-tolerance acceptance from 2 to 37 of the 162 delivered updates (Table 1).

Our contributions:

- We prove the first conditional safety theorem for uncorrelated multiplayer subgame resolving. For any update σ^1 of a blueprint σ^0 by a resolver i in a subgame S, vulnerability rises by at most $A_i^S(\sigma^0) + 2n\,\delta_\infty + \sum_q M_q + \varepsilon_i$ (Theorem 1). A sharper Θ-form replaces $2n\,\delta_\infty$ and vanishes on pointwise constant-sum games (Corollary 4). An exact form charges the realized total-utility leak (Corollary 1). The bound separates each channel of harm: activation, structural distance, and per-opponent margin deficit.
- We construct instances pinning the bound's form. The activation term is exact already in two-player CFR-D (CFR with decomposition; Example 1); the per-opponent margin term is forced by a three-player transfer (Example 2); realized negative margins carry harm (Proposition 3); and a certificate too quiet to explain a realized increase defeats margin-only bounds (Proposition 4). We certify δ_∞ lower bounds for poker-like games with linear-programming (LP) dual certificates (Proposition 2) and exhibit natural rules with $\delta_\infty = 0$ (Section 6).
- We validate the certificate pipeline by a certified adversarial search over 162 instances, which finds no violation, and by numerical regression checks of the theorem's inequalities on finite configurations (Section 7).

2 Related Work

Two-player safe resolving. CFR-D introduced game decomposition, the resolving gadget, and the first safety guarantee for two-player zero-sum subgame refinement (Burch et al., 2014); maxmargin resolving made the margin the optimization objective (Moravcik et al., 2016); safe-and-nested solving made the guarantees practical at scale (Brown & Sandholm, 2017) via depth-limited solving (Brown et al., 2018). The framework extends to endgame solving (Ganzfried & Sandholm, 2015a), limited lookahead (Kroer & Sandholm, 2020), subgame solving without common knowledge (Zhang & Sandholm, 2021), continual resolving (Šustr et al., 2019), sound online solving (Šustr et al., 2021), and efficient refinement (Ge et al., 2023). These guarantees are intrinsically two-player zero-sum: the opponent's gain is the resolver's loss. We carry their margin-certificate language to the original multiplayer game, where the bound lands on Vul_G itself.

Multiplayer resolving in practice and refinement theory. Pluribus showed that heuristic real-time resolving yields strong six-player play with no safety guarantee (Brown & Sandholm, 2019b), and no multiplayer analogue of solved heads-up limit hold'em (Bowling et al., 2015) exists. Provable multiplayer guarantees appear where two-sided or correlated structure is restored: in zero-sum team games (Zhang & An, 2020; Zhang et al., 2021), and for correlated play. There, resolving a correlation plan is safe in two-player general-sum games (Ling & Fang, 2022). Which equilibrium concept subgame solving should target is itself under study (Kubicek et al., 2026). Our setting is general

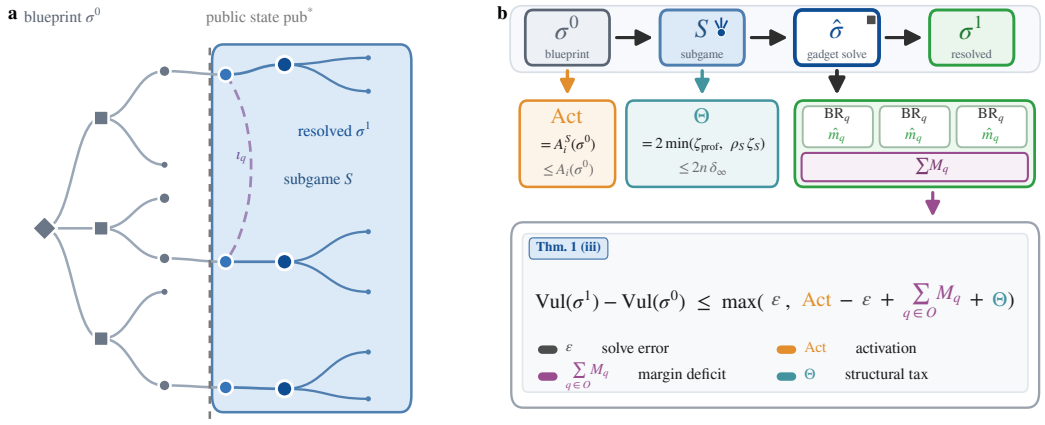


Figure 1: Subgame resolving and its certified guarantee. **(a)** Three-player game tree. The blueprint σ^0 (CFR) is kept outside the public state pub^*; the resolver updates it to σ^1 inside the subgame S, where the entries E are grouped into entry information sets ι_q. **(b)** The certificate pipeline and the bound of Theorem 1(iii): $\text{Vul}_G(\sigma^1) - \text{Vul}_G(\sigma^0) \leq \max\big(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \sum_{q\in O} M_q + \Theta\big)$. Each term is a distinct failure channel, matched by color in the bound: $A_i^S(\sigma^0)$ the blueprint's unkept promise, ε_i the solve error, $\sum_q M_q$ the margin deficit, Θ the structural tax.

n-player play with no team structure and no correlation device. The guarantee is conditional on the distance to constant-sum structure and requires no two-sided reduction.

Safe exploitation, adversarial team games, blueprint computation, and the polymatrix games that supply our distance measure are discussed in Appendix A.

3 Preliminaries

We consider a finite n-player extensive-form game (EFG) G, a game tree with chance moves and hidden information, with perfect recall (no player forgets what it knew or did): $N = \{1, \ldots, n\}$, Z the terminal histories, and $u_p(z)$ player p's payoff at z. A behavioral strategy σ_p assigns a distribution over actions to each of p's information sets (the sets of histories p cannot tell apart); a profile σ induces expected payoff $u_p(\sigma)$. For a history h, $\pi_p^\sigma(h)$ is the product of p's action probabilities on the path to h. The others' joint reach is $\pi_{-p}^\sigma(h) := \prod_{q\neq p} \pi_q^\sigma(h)$, with chance kept separate. Every non-terminal history carries a public label $\text{pub}(h)$, the information visible to all players; we write $(x)_+ := \max(x, 0)$.

Resolving, gadget games, and best-response (BR) certified margins. Subgame resolving starts from a *blueprint* σ^0 computed offline and refines the strategy of one player, the *resolver* i, inside a subgame reached during play; the remaining players are the *opponents* $O := N \setminus \{i\}$. A *public subgame* is specified by a public state pub^*: its *entry set* E collects the earliest histories e with $\text{pub}(e) = \text{pub}^*$, each with chance probability c_e; the subgame S consists of the descendants of E, and the rest of the game is the *trunk*. Entries are grouped into *entry information sets* $\iota_q(e) := (\text{priv}_q(e), \text{pub}^*)$, collected in I_q; $Z_S \subseteq Z$ denotes the terminals inside S. Each set carries counterfactual entry mass $\bar{W}_q(\iota) := \sum_{e\in\iota} c_e\, \pi_{-q}^{\sigma^0}(e)$ and blueprint counterfactual value $\text{CFV}_q(\iota) := \bar{W}_q(\iota)^{-1} \sum_{e\in\iota} c_e\, \pi_{-q}^{\sigma^0}(e)\, \text{EV}_q^{\sigma^0}(e)$, where $\text{EV}_q^{\sigma^0}(e)$ is q's expected payoff from entry e under σ^0 (counterfactual: weighted by the others' reach, never by q's own). Resolving solves a CFR-D/maxmargin-style *gadget game* (Burch et al., 2014; Moravcik et al., 2016; Brown & Sandholm, 2017): a root chance node enters $e \in E$ with probability proportional to $c_e \pi_i^{\sigma^0}(e)$; each opponent j then chooses Terminate or Follow at an information set keyed by ι_j: terminating pays $\text{CFV}_j(\iota)$, following lets j play freely. The output is an update σ^1 *confined to the resolver's subgame*: $\sigma_p^1 = \sigma_p^0$ for $p \neq i$, and σ_i^1 differs from σ_i^0 only on i's information sets inside S.

The certificate is a *margin* against an exact subgame best response (the payoff-maximizing reply to the others' fixed strategies): for $j \in O$, $\iota \in I_j$, and a subgame strategy x_i of the resolver, $\mathrm{SBV}_j(\iota;\, x_i) := \max_{\tau_j^S} \sum_{e\in\iota} c_e\, \pi^{\sigma^0}_{-j}(e)\ \mathbb{E}[u_j \mid e,\ (\tau_j^S, x_i, \sigma^0_{-ij}) \text{ inside } S]$ is j's best-response value with the resolver playing x_i and all others, σ^0_{-ij}, following the blueprint.

Definition 1 (BR-certified margin)**.** For an opponent $j \in O$ and entry information set $\iota \in I_j$, the *BR-certified margin* of the update σ^1 is $\hat{m}_j(\iota) := \mathrm{CFV}_j(\iota) - \mathrm{SBV}_j(\iota;\, \sigma_i^1)/\bar{W}_j(\iota)$.

Definition 2 (Margin deficit)**.** The *margin deficit* of opponent $q \in O$ under σ^1 is the mass-weighted sum $M_q := \sum_{\iota\in I_q} \bar{W}_q(\iota)\,\big(-\hat{m}_q(\iota)\big)_+$.

Only entry groups with $\bar{W}_q(\iota) > 0$ contribute; zero-mass groups weigh zero (Appendix J).

We call a *caliber* a choice of definition or measurement convention (Appendix J). First, M_q is a *non-normalized, mass-weighted sum* of per-group deficits $(-\hat{m}_q(\iota))_+$; it is small when the mass spreads thin, even if one entry has a large negative margin. When every opponent's own action sequence coincides across E, $\sum_\iota \bar{W}_q(\iota) \le 1$ and $M_q \le (-\min_\iota \hat{m}_q(\iota))_+$ (Appendix B). This holds when all actions before pub^* are public, as in poker-style games. Second, the margin that the implementations report *overstates* Definition 1. It values the Follow branch under the opponent's gadget average strategy, which is optimistic. It exceeds the BR margin on 33 of 40 entry information sets (Appendix J). Solve quality enters through one term $\varepsilon_i \ge 0$, defined in Section 4 with A_i^S.

Definition 3 (Vulnerability)**.** For a profile σ in G, player p's best-response gain is $A_p(\sigma) := \max_{\tau_p} u_p(\tau_p, \sigma_{-p}) - u_p(\sigma)$, and the *vulnerability* of σ is $\mathrm{Vul}_G(\sigma) := \max_{p\in N} A_p(\sigma)$.

In two-player zero-sum games safety is phrased as exploitability, the distance to the interchangeable equilibrium set (Von Neumann & Morgenstern, 2007). Beyond two players equilibria are not interchangeable, their values are not comparable, and even a refinement is intractable (Daskalakis et al., 2009; Hansen et al., 2010); multiplayer exploitability has no canonical definition (Cai & Daskalakis, 2011; Gibson & Szafron, 2011). Vul_G is defined for every profile and exactly computable by one best-response pass per player; at $n = 2$ it is the larger of the two best-response advantages, close to but distinct from the common exploitability convention $\frac{1}{2}(A_1 + A_2)$.

The structural knob is the distance from G's utility table to the class of constant-sum polymatrix (CSP) games, on the *pure-strategy profile table*. Let $s = (s_1, \ldots, s_n)$ range over profiles of reduced pure strategies (one action per reachable information set), $u_p(s)$ be p's payoff after the expectation over chance, and $U(s) := \sum_p u_p(s)$ the player sum.

Definition 4 (δ_∞-CSP distance)**.** Let V denote the set of profile-utility tables v admitting a *constant-sum polymatrix decomposition*: pairwise tables $\{v_{pq}\}_{p\ne q}$ and edge constants $\{c_{pq}\}_{p<q}$ such that $v_p(s) = \sum_{q\ne p} v_{pq}(s_p, s_q)$ for every profile s, and $v_{pq}(a,b) + v_{qp}(b,a) = c_{pq}$ for every edge $\{p,q\}$ and all pure strategies a of p and b of q. The δ_∞*-CSP distance* of G is $\delta_\infty(G) := \min_{v\in V} \|u - v\|_\infty = \min_{v\in V} \max_{p,s} |u_p(s) - v_p(s)|$.

The class is the *per-edge constant-sum* class of MacQueen & Wright (2023) and Cai et al. (2016), which contains the antisymmetric subclass; our proofs use only that $\sum_p v_p$ is constant on V. Defining δ_∞ on profile tables, not on terminals, is what lets a pointwise zero-sum game have $\delta_\infty > 0$ (three-player Kuhn, Section 6). V is a linear subspace, so δ_∞ is an exact LP.

Assumption 1 (Public-state subgame regularity)**.** The subgame S satisfies: (P1) every information set of every player is either entirely contained in S or disjoint from it; (P2) the gadget's entry grouping $\iota_q = (\mathrm{priv}_q, \mathrm{pub}^*)$ coincides with q's actual information at entry, and q's own reach $\pi_q^\tau(e)$ is constant across $e \in \iota$ for every strategy τ of q.

Assumption 1 restricts the subgame's geometry. (P1) holds automatically for subgames cut at public states, and (P2) holds under perfect recall whenever entries follow a public signal; the scope of the assumption is discussed in Appendix G.

4 Main Theorem and How to Read It

We state the main result; the proof is in Section 5. Three quantities beyond Definitions 1 and 2 appear. The *structural tax* is built from constant-sum defects of the utility table: with $U(s) = \sum_p u_p(s)$ and

$U(z) := \sum_p u_p(z)$, let

$$\zeta_{\text{prof}} := \min_c \max_s |U(s) - c|, \qquad \zeta_S := \min_c \max_{z \in Z_S} |U(z) - c|, \qquad \Theta := 2\,\min(\zeta_{\text{prof}},\ \rho_S\,\zeta_S),$$

where $\rho_S := \sum_{e \in E} c_e \prod_p \pi_p^{\sigma^0}(e)$ is the blueprint's probability of reaching the subgame. The *activation term* $A_i^S(\sigma^0) := \max_{\tau_i:\, \tau_i = \sigma_i^0 \text{ outside } S} u_i(\tau_i, \sigma_{-i}^0) - u_i(\sigma^0)$ is the subgame payoff the blueprint promises the resolver but does not realize; it is computable from σ^0 alone and satisfies $A_i^S(\sigma^0) \le A_i(\sigma^0) \le \text{Vul}_G(\sigma^0)$. The *blueprint-relative suboptimality* $\varepsilon_i := A_i^S(\sigma^0) - [u_i(\sigma^1) - u_i(\sigma^0)] \ge 0$, called the solve error for short, measures the suboptimality of σ_i^1 against the blueprint opponents inside S; it is evaluated on the delivered σ^1. We call σ^1 *margin-certified* when $\hat{m}_q(\iota) \ge 0$ for all q and all ι with $\bar{W}_q(\iota) > 0$; the theorem charges violations through M_q. Throughout, $\Delta\text{Vul} := \text{Vul}_G(\sigma^1) - \text{Vul}_G(\sigma^0)$.

Theorem 1 (Conditional safety of multiplayer resolving)**.** *Let G be an n-player perfect-recall EFG satisfying Assumption 1, let σ^0 be an arbitrary blueprint, and let σ^1 be any update confined to resolver i's subgame strategy ($\sigma_p^1 = \sigma_p^0$ for all $p \ne i$, and σ_i^1 differs from σ_i^0 only on i's information sets inside S). Let $\hat{m}$, M_q, $A_i^S(\sigma^0)$, and ε_i be defined with respect to the delivered σ^1 (Definitions 1 and 2 and the displays above), and let $\Theta = 2\min(\zeta_{\text{prof}}, \rho_S \zeta_S)$. Then:*

(i) $A_i(\sigma^1) - A_i(\sigma^0) \;=\; \varepsilon_i - A_i^S(\sigma^0) \;\le\; \varepsilon_i$ *;*

(ii) for every opponent $j \ne i$, $A_j(\sigma^1) - A_j(\sigma^0) \le M_j + [A_i^S(\sigma^0) - \varepsilon_i + \sum_{k \in O \setminus \{j\}} M_k] + \Theta$*;*

(iii) $\text{Vul}_G(\sigma^1) - \text{Vul}_G(\sigma^0) \;\le\; \max_p\big[A_p(\sigma^1) - A_p(\sigma^0)\big] \;\le\; \max\Big(\varepsilon_i,\quad A_i^S(\sigma^0) - \varepsilon_i + \sum_{q \in O} M_q + \Theta\Big)$.

Corollary 1 (Exact post-hoc form)**.** *Let $\Lambda := \sum_p u_p(\sigma^0) - \sum_p u_p(\sigma^1)$ be the realized total-utility leak of the update. Then Theorem 1(ii) and (iii) hold with Θ replaced by Λ, and $\Lambda \le |\Lambda| \le \Theta$ (Lemma 2). Since Λ is signed it stays inside the maximum: for the sequential resolves of Corollary 6 the bound is $\sum_t \max(\varepsilon_i[t],\ A_i^S[t] - \varepsilon_i[t] + \sum_q M_q[t] + \Lambda[t])$.*

Corollary 2 (Direct per-player form)**.** *Write $g_p := u_p(\sigma^1) - u_p(\sigma^0)$ for the realized gain of player p. Then*

$$\text{Vul}_G(\sigma^1) - \text{Vul}_G(\sigma^0) \;\le\; B_{\text{dir}} := \max\big(-g_i,\ \max_{j \ne i}[M_j - g_j]\big), \tag{1}$$

and B_{dir} is pointwise at most the bound of Corollary 1 (Appendix C).

Corollary 3 (Baseline-aware form)**.** *With the blueprint advantages $A_p^0 := A_p(\sigma^0)$, $\text{Vul}_G(\sigma^1) - \text{Vul}_G(\sigma^0) \le B_{\text{cache}} := \max\big(A_i^0 - g_i,\ \max_{j \ne i}[A_j^0 + M_j - g_j]\big) - \max_p A_p^0 \le B_{\text{dir}}$. On a constant-sum game $B_{\text{dir}} \le 0$ forces $g_p = 0$ and $M_j = 0$ for all p, $j \ne i$ (Appendix C); B_{cache} does not.*

B_{dir} needs only the BR-certified margins and the n expected utilities; B_{cache} also reads the stored A_p^0. Theorem 1 decomposes B_{dir} into activation, margin, and structural channels, and Θ envelops Λ a priori. A tighter blueprint-specific quantity is L_i: the total-utility leak of one best response, confined to S, to the negated total utility $-\sum_p u_p$. It satisfies $\Lambda \le L_i \le \Theta$ and costs one best response.

Lemma 1 (Section 5, Appendix B) proves the theorem at $\delta_\infty = 0$; Lemma 2 (Appendix C) replaces the one step where $\delta_\infty = 0$ entered, bounding $\sum_p[u_p(\sigma^1) - u_p(\sigma^0)]$ by $\Theta \le 2n\delta_\infty$; the a priori $2n\delta_\infty$ instantiation is Corollary 5 (Appendix C).

5 Proof of the Main Theorem

The exact case ($\delta_\infty = 0$). We first prove the main bound at $\delta_\infty = 0$, where the tax vanishes. Throughout, G satisfies Assumption 1 and σ^1 is any update of the form in Section 4, with $\hat{m}$, M_q, $A_i^S(\sigma^0)$, and ε_i defined for the delivered σ^1 (Definitions 1 and 2).

Lemma 1 (Safety of joint-gadget resolving when $\delta_\infty = 0$)**.** *Let G satisfy Assumption 1 and $\delta_\infty(G) = 0$ (LP semantics of Definition 4), and let σ^1 be any update of the above form. Then:*

(i) (resolver) $A_i(\sigma^1) - A_i(\sigma^0) \;=\; \varepsilon_i - A_i^S(\sigma^0) \;\le\; \varepsilon_i$.

(ii) (each opponent $j \neq i$) $A_j(\sigma^1) - A_j(\sigma^0) \;\le\; M_j + [A_i^S(\sigma^0) - \varepsilon_i + \sum_{k\in O\setminus\{j\}} M_k] \;\le\; A_i^S(\sigma^0) + \sum_{q\in O} M_q$.

(iii) (total) $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \le \max\big(\varepsilon_i,\; A_i^S(\sigma^0) - \varepsilon_i + \sum_{q\in O} M_q\big)$.

If $\hat{m}_q(\iota) \ge 0$ for all q, ι (a margin-certified update) and $\varepsilon_i \le \varepsilon_{\mathrm{solve}}$, the solve-error budget, then $\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0) + \max\big(\varepsilon_{\mathrm{solve}},\; A_i^S(\sigma^0)\big)$.

The full proof, with its 97-item numerical regression check, is in Appendix B. Perfect recall splits any deviation of opponent j into a trunk part and independent subgame continuations, so j's best-response value rises by at most M_j. No gadget step uses the CSP structure. The assumption $\delta_\infty = 0$ enters exactly once, where it forces $\sum_p g_p = 0$ for the realized gains $g_p := u_p(\sigma^1) - u_p(\sigma^0)$; that identity is what re-prices a victim's on-path loss into $A_i^S(\sigma^0) - \varepsilon_i$ plus the other deficits.

Two examples show the shape of Lemma 1 is forced: the activation term survives a certificate with every margin nonnegative, even in two-player zero-sum games.

Example 1 (Activation: margin-certified, $\varepsilon = 0$, yet $\Delta\mathrm{Vul} = A_i^S$)**.** Two players $\{i, j\}$, zero-sum (so $\delta_\infty = 0$). Player j moves first: action d ends the game with $(-1, +1)$; *enter* reaches the subgame S, where i alone acts: a' gives $(0, 0)$, a gives $(+1, -1)$. The blueprint is $\sigma^0 = (a', \textit{enter})$: $A_i(\sigma^0) = A_i^S(\sigma^0) = 1$, so $\mathrm{Vul}_G(\sigma^0) = 1$. Resolving delivers $\sigma_i^1 = a$: $\mathrm{CFV}_j = 0$, $\mathrm{SBV}_j(\sigma_i^1) = -1$, so $\hat{m}_j = +1 > 0$ and $\varepsilon_i = 0$; yet $A_j(\sigma^1) = 2$, so $\mathrm{Vul}_G(\sigma^1) = 2 = \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0)$.

Resolving to a realizes the unit of subgame value the blueprint left unrealized; no margin flags the move ($\hat{m}_j = +1 > 0$), and Vul_G rises $1 \to 2$ while the half-sum convention $\frac{1}{2}(A_1 + A_2)$ stays at 1. What moves is Vul_G, which prices what two-player exploitability leaves out. CFR-D and maxmargin resolving (Burch et al., 2014; Moravcik et al., 2016) bound that quantity for the resolver alone. The coefficient 1 on A_i^S is optimal.

Second, the per-player granularity of the margin term is forced: a per-pair guarantee keyed to the victim's own margin is false.

Example 2 (Cross-opponent transfer: the victim's margin is silent)**.** Three players $\{i, j, k\}$, one simultaneous move (the subgame is the whole game): i chooses in $\{a', a\}$, j in $\{\textit{stay}, d\}$, k has no action. Take the pair-local witness $u_{ij}(a, \textit{stay}) = +1$, $u_{ij}(a, d) = 0$, $u_{ij}(a', \cdot) = 0$; $u_{ik}(a, \cdot) = -1$, $u_{ik}(a', \cdot) = 0$; $u_{jk} \equiv 0$; so $\delta_\infty = 0$. The blueprint is $\sigma^0 = (a', \textit{stay})$: $\mathrm{Vul}_G(\sigma^0) = A_i^S(\sigma^0) = 0$. Resolving delivers $\sigma_i^1 = a$: $\hat{m}_j = 0$, $\hat{m}_k = -1$, $\varepsilon_i = 0$. Then $\Delta\mathrm{PairVul}_{ji} = 1 > 0 = (-\hat{m}_j)_+$, where $\mathrm{PairVul}_{ji}$ is j's best-response gain on the pair utility u_{ji} (Appendix H), refuting the per-pair form, and $A_j(\sigma^1) = 1$, so $\Delta\mathrm{Vul} = 1 = \sum_q M_q$.

i's own total payoff is unchanged by the update, yet value moves from the (i, k) to the (i, j) pair table. The victim's margin stays silent, and the alarm fires at the *beneficiary* k. A per-pair certificate keyed to the victim's margin is therefore unobtainable, which forces the per-player granularity of Lemma 1 and pins the coefficient 1 on the margin term as optimal. The per-pair statement that survives appears in Appendix H. The other two terms of Theorem 1 are attained as well.

Example 3 (Pure leak and self-harm)**.** *Leak:* two players; j moves first, d ends the game with $(0, 0)$, *enter* reaches S, where i alone acts: a' gives $(0, 0)$, a gives $(0, -1)$. From $\sigma^0 = (a', \textit{enter})$ the update $\sigma_i^1 = a$ has $A_i^S = \varepsilon_i = 0$ and $\hat{m}_j = +1$, yet $A_j(\sigma^1) = 1$: $\Delta\mathrm{Vul} = 1 = \Theta$ ($\zeta_{\mathrm{prof}} = \zeta_S = \frac{1}{2}$, $\rho_S = 1$). *Self-harm:* two players, zero-sum, i alone acts: a gives $(1, -1)$, a' gives $(0, 0)$. From $\sigma_i^0 = a$ the update $\sigma_i^1 = a'$ has $A_i^S = 0$, $\varepsilon_i = 1$, $M_j = 1$, and $\Delta\mathrm{Vul} = A_i(\sigma^1) = 1 = \varepsilon_i$ (Appendix B.5).

In the leak the resolver destroys one unit of j's utility at no cost to itself; no margin sees it and only the total-utility oscillation prices it. With Examples 1 and 2, all four terms of Theorem 1(iii) and both branches of its maximum are attained with equality.

The transfer step. $\delta_\infty = 0$ entered Lemma 1 at one step (S7 in Appendix B), which used only that $\sum_p u_p$ is constant across profiles. When $\delta_\infty > 0$ this constancy can fail, by a bounded amount, and the transfer lemma controls exactly that drift through the *distance to constant-sum* structure:

Lemma 2 (Transfer)**.** *Let $\delta := \delta_\infty(G)$ in the LP profile-table semantics of Definition 4. Then:*

(a) (profile level) $\zeta_{\mathrm{prof}} \le n\,\delta_\infty$.

(b) (update level) *For any update σ^1 as in Lemma 1 (Assumption 1; only the resolver's strategy inside S changes),* $\left|\sum_p g_p\right| \le \Theta := 2\min(\zeta_{\mathrm{prof}}, \rho_S\,\zeta_S) \le 2\zeta_{\mathrm{prof}} \le 2n\delta_\infty$.

Proof sketch and the $2n$ accounting. For (a), summing an optimal LP witness's residuals $|e_p(s)| \le \delta$ collapses pairwise terms through the per-edge constant-sum constraint, leaving $|U(s) - c^*| \le n\delta$. For (b), both σ^0 and σ^1 lie within ζ_{prof} of the optimal center, and $P^{\sigma^1} - P^{\sigma^0}$ has zero total mass on Z_S, giving $|\sum_p g_p| \le 2\rho_S\,\zeta_S$. Substituting Θ in the zero-sum step of Lemma 1 yields Theorem 1 (Appendix C); the n of leg (a) is instance-level tight ($\zeta_{\mathrm{prof}}/(3\delta) = 1.000000$ on screening instances gid49 and gid41, Appendix I).

The $\min$ inside Θ is essential. On the pairpot family of Proposition 1, $\Theta = 0$ although $\delta_\infty = \lambda/16 > 0$, and $\zeta_{\mathrm{prof}} \le n\delta_\infty$ rules out the reverse, so the Θ-form is strictly stronger than the $2n\delta_\infty$ corollary. On the screening set of Appendix I, $\Theta < 2n\delta_\infty$ on 266 of 296 points, with ζ_{prof} the smaller leg on 204 and $\rho_S\zeta_S$ on 92 (Appendix F).

Corollary 4 (Constant-sum games pay no structural tax)**.** *If G is pointwise constant-sum ($\sum_p u_p(z) = c$ for every $z \in Z$), then $\zeta_{\mathrm{prof}} = \zeta_S = 0$ and hence $\Theta = 0$, so Theorem 1 reduces exactly to the Lemma 1 form* $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \le \max(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \sum_{q\in O} M_q)$.

Proof. $U(s) = \mathbb{E}[U(z)] = c$ for every profile and $Z_S \subseteq Z$, so $\zeta_{\mathrm{prof}} = \zeta_S = 0$. □

The class is large: three-player Kuhn, rake-free Hold'em, and zero-sum-normalized Goofspiel are all pointwise constant-sum, so they pay no structural tax; two of them have $\delta_\infty \ge 1$ (Proposition 2).

6 Structure and Tightness of the Bound

The $2n\delta_\infty$ corollary has force when δ_∞ is small; the sharp form of Theorem 1 needs only a small structural tax. The CSP class (MacQueen & Wright, 2023) (see also (Cai et al., 2016; Cai & Daskalakis, 2011)) consists of tables admitting a decomposition with per-edge constant-sum constraints, as in Definition 4.

Lemma 3 (Linear perturbation law)**.** *The set V of CSP-decomposable profile utility tables is a linear subspace of the space of utility tables. Consequently, for every $u^0 \in V$, every utility table C, and every $\lambda \ge 0$, $\delta_\infty(u^0 + \lambda C) = \lambda\,\delta_\infty(C)$.*

Proof sketch (Appendix D): V is a subspace, so the sup-distance is translation-invariant and positively homogeneous; this upgrades the measured pairpot slope $1/16$ to a necessity (Appendix F).

The *pairpot family* is $u_\lambda := u^0 + \lambda C$. Here u^0 is a three-player table of independent per-pair side pots, which is CSP. The coupling tensor C adds a shared pot that is a pure transfer between the players, so every u_λ is pointwise zero-sum on terminals (rules in Appendix D).

Proposition 1 (Explicit slope constant for the pairpot family)**.** *For the coupling tensor C, $K := \max_{p,s} |C_p(s)| = 1/4$, and hence $\delta_\infty(u^0 + \lambda C) \le \lambda/4$ for all $\lambda \ge 0$. The constant is attained: an explicit pure profile realizes $|C_p(s)| = 1/4$.*

The proof is a hand case analysis (Appendix D), reproduced by exhaustive enumeration over all 4096 pure profiles; since $\delta_\infty(C) = 1/16$ by exact LP, Proposition 1 covers the true slope with a $4\times$ gap.

A positive class with natural rules. The condition $\delta_\infty = 0$ holds on a robust family: the pairwise-showdown games. Exact LP over all 4096 pure profiles gives $\delta_\infty = 0$, persisting under correlated dealing, because each edge utility enters expectations only through that pair's marginal of the deal. Payoff coupling through opponents' *actions* breaks CSP structure (Appendix E.5).

Proposition 2 (Poker-style games are far from CSP)**.** *The following lower bounds hold, each certified by an LP dual certificate: three-player Kuhn poker with a four-card deck has $\delta_\infty \ge 1.000$ ($\max|u| = 4.0$, relative distance ≥ 0.250); blind Goofspiel with K=3 prizes has $\delta_\infty = 2.500$*

*exactly (*216 *pure profiles, relative* 0.500*); blind Goofspiel with* K=4 *has* $\delta_\infty = 4.625$ *exactly (*13824 *pure profiles, relative* 0.514*); observed Goofspiel with* K=3 *has* $\delta_\infty \geq 3.000$ *(relative* ≥ 0.500*).*

The relative distance does not shrink between the two Goofspiel sizes, 0.500 at K=3 and 0.514 at K=4. Each of the four lower bounds comes with an LP dual certificate (Appendix E).

For poker-style games, Corollary 5 is valid but vacuous. Since $\delta_\infty \geq 1$ (Proposition 2), the structural term $2n\,\delta_\infty$ exceeds the utility scale of three-player Kuhn, whereas the sharp Θ-form charges poker no tax (Corollary 4). Pointwise constant-sumness is a sufficient condition for $\Theta = 0$, and the pairpot family shows that $\delta_\infty = 0$ is not a necessary one. One obstacle for poker is the gadget-margin proxy that implementations report, which misreads near equilibrium (Proposition 5); the BR certificate of Definition 1 prices the same entries with one best-response pass per opponent.

Tightness and negative results. The analytic constructions of Section 5 attain each term with equality. This section shows the channels on realized instances with exact numerical constructions (Appendix E). The instances realize a large margin deficit (Proposition 3), a certificate too quiet to explain a realized increase (Proposition 4), and implementation margins that misread near equilibrium (Proposition 5).

Proposition 3 (Negative margins are realized)**.** *There exist three-player games and resolving updates satisfying the quantifiers of Theorem 1 on which strongly negative BR-certified margins are matched by a vulnerability increase of the same order. Screening instance gid78 (identifiers follow Appendix I) has* $\Delta\mathrm{Vul} = 0.2164$ *with minimum margin* $\hat{m} = -1.0559$*; a locally optimized variant of gid90 has* $\Delta\mathrm{Vul} = 0.2498$ *with minimum margin* -0.5788 *and exact* $\delta_\infty = 0.3033$*. The constants* $(0.03, 0.03)$ *fitted on the pilot run are falsified by the gid90 variant. The screening objective,* $\Delta\mathrm{Vul} - 0.03\,\delta_\infty - 0.03(-\hat{m})_+$*, is* $+0.223$ *at its exact* δ_∞*. On gid78 it is* $+0.178$*, a screening score under a* δ_∞ *lower bound; the BR-certified deficit there aggregates to* $\sum_q M_q = 0.252$ *with* $A_i^S \approx 0$*,* $\varepsilon_i = 0.002$*,* $\Theta = 0.017$ *and* $\Lambda = -0.003$*; with the margin term removed, the bound of Theorem 1 is* 0.015 *against* $\Delta\mathrm{Vul} = 0.216$*, so the increase is carried by the margin channel alone.*

Proposition 4 (Margin-only bounds are insufficient)**.** *There is a three-player instance, gid174, from the arm in which the resolver's blueprint subgame strategy is reset to a Dirichlet draw before resolving (Appendix I), whose resolving update has* $\Delta\mathrm{Vul} = 0.0383$ *above tolerance* 0.02 *with* $\delta_\infty = 0.43$*, while its BR-certified margin deficit aggregates to only* $\sum_q M_q = 0.0111$ *(*$A_i^S = 0.0382$*,* $\varepsilon_i = 0.0379$*). Any margin-only bound* $\Delta\mathrm{Vul} \leq c_2 \sum_q M_q + \mathrm{tol}$ *needs* $c_2 \approx 1.65$*, five times the fitted* $c_2 = 0.31$ *(Table 5).*

A margin-only bound fails here by more than slack: at tolerance 0.02 it leaves the realized increase of 0.0383 uncovered, since $\sum_q M_q = 0.0111$ is small. The solve-error term carries most of the increase, so the construction isolates the non-margin channels jointly, and Theorem 1 covers the instance through ε_i and the δ_∞ term.

Proposition 5 (Structural margin mismatch near equilibrium for $n > 2$)**.** *There exist three-player games and blueprints with full-game vulnerability* 1.4×10^{-3} *for which the implementation-caliber gadget margins are strictly negative at zero deviation: in the pilot construction, the opponents' margins evaluate to* -0.46 *and* -0.74 *on single entry information sets although the update deviates nowhere. These are gadget margins (Appendix J). The BR-certified margins of Definition 1 behave differently: at zero deviation they are nonpositive at every entry set, and the mass-weighted deficits are* 1.8×10^{-4} *and* 3.5×10^{-4}*.*

The mechanism is a weighting defect. Correct beliefs for opponent q need the reach of all players other than q; the gadget's root measure $c_e\, r_i(e)$ omits the other opponents' reach (Appendix J). For $n > 2$ an alternative is overweighted with no deviation present; only the BR certificate supplies the counterfactual weight $c_e\, \pi_{-q}(e)$ at unreached entry sets.

7 Adversarial Empirical Evaluation

Numerical regression checks corroborate every inequality of Theorem 1 (Appendix F). This section searches for theorem-level counterexamples, instances where vulnerability rises while every term of the bound is silent. It also measures how small the coefficients of a two-term $(\delta_\infty, \sum_q M_q)$ law can

Table 1: Acceptance of the 162 delivered (nominal) M1 updates; a form accepts an update when $B \leq \mathrm{tol}$. Rows 1–2 bound $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0)$; rows 3–6 the per-player increase of Theorem 1(iii). Full metrics in Appendix I.5.

Certificate	Acc. 0	Acc. 0.02
B_{dir} (Eq. (1))	2	73
B_{cache} (Corollary 3)	37	104
$X = \Lambda$ (realized leak)	0	72
$X = L_i$ (confined bound)	0	45
$X = \Theta$ (sharp tax)	0	19
$X = 2n\delta_\infty$ (Corollary 5)	0	2

Table 2: Trunk scaling (Appendix I.6): median wall-clock over six seeds, subgame fixed at 6 information sets; the local certificate is constant at 2.0 ms. Ratio is the median of the paired per-seed ratios.

K	Cached exact (ms)	Local / cached
1	2.1	0.96×
4	8.1	0.25×
16	32.8	0.061×
64	145	0.014×

be made. Three instruments: a certified cutting-plane best-response search over small n-player EFGs (M1: 162 instances at $n = 3$, every point evaluated without sampling); a heuristic sweep over 433 screened points; and LP fits, the theorem-level counterexample filter with $\mathrm{tol} = 0.02$ (Appendix I).

Result 1: zero theorem-level counterexamples. On the full 162-instance M1 set the strict filter fires *zero* times under both margin calibers. Every point where $\Delta\mathrm{Vul}$ exceeds tolerance is explained by a large δ_∞, a margin deficit, A_i^S, or ε_i. Relaxing the filter to ($\delta_\infty \leq 0.15$, $\sum_q M_q \leq 0.05$) then ($\delta_\infty \leq 0.30$, $\sum_q M_q \leq 0.10$) yields 3 and 6 hits, none theorem-level; the screening sweep agrees ($\max \Delta\mathrm{Vul} = 0.039$ at $\delta_\infty \leq 0.1$, 156 configurations).

Result 2: a two-term ablation. An LP over all M1 measurements finds the minimum (c_1, c_2) with $\Delta\mathrm{Vul} \leq c_1\delta_\infty + c_2 \sum_q M_q + \mathrm{tol}$ on every point (Table 5): $(1.76, 0.31)$ at $\mathrm{tol} = 0.02$, $(2.35, 0.17)$ at $\mathrm{tol} = 0$; within this law both terms are needed ($c_2 = 0$ forces $c_1 \geq 2.64$; $c_1 = 0$ is infeasible). The pair measures a different object than the coefficients $(2n, 1)$ of Corollary 5, whose bound also carries A_i^S and ε_i.

Result 3: scaling in the number of players. An exploratory n-sweep fits the c_1 coordinate of the two-term law at 0.162, 0.563, and 0.916 for $n = 3, 4, 5$, a slope of 0.377 per player (Appendix I.3).

Result 4: which terms carry the bound. Split the bound into $B = \max(\varepsilon_i,\ A_i^S - \varepsilon_i + \sum_q M_q)$ and the δ_∞ term: nominal-caliber B alone covers all 162 points; the pipeline caliber covers 89.3% of 150, the rest by δ_∞ (Table 5).

Result 5: what the certificate does for delivered updates. On the 162 delivered M1 updates every form of Theorem 1 and Corollary 2 covers $\Delta\mathrm{Vul}$, and at tolerance 0.02 every form flags all 19 harmful updates; at tolerance 0 the structural forms accept none of the 143 safe ones (Table 1; full table in Table 6). A form is useful when its bound is small enough to *accept* a safe update, and the baseline-aware form B_{cache} accepts 37 of them at zero tolerance, at 1.5× the cost of a cached exact evaluation that already accepts every safe update (Appendix I.5; trunk scaling in Table 2).

8 Conclusion

This paper gives the first safety guarantee for subgame resolving with three or more independent players, where the two-player identity that one player's gain is the other's loss fails. We replace it with a certificate: any delivered update raises the profile's vulnerability over the blueprint by at most four computable terms, the blueprint's unkept promise A_i^S, the structural tax $\Theta \leq 2n\,\delta_\infty$, the best-response margin deficits $\sum_q M_q$, and the solve error ε_i (Theorem 1). Each term is a distinct channel of harm, explicit constructions attain all four with equality (Sections 5 and 6), and on pointwise constant-sum games, which include every rake-free poker variant, the structural tax vanishes (Corollary 4). The certificate is post-hoc and covers CFR-D-style and sequential resolving; it runs at 1.5× a cached exact evaluation and, in its baseline-aware form, lifts zero-tolerance acceptance from 2 to 37 of 162 delivered updates. The n factor of the transfer lemma is attained on screening instances, but the end-to-end response of $\Delta\mathrm{Vul}$ to δ_∞ is not proved tight in n (Appendix C). Multiplayer resolving now rests on conditions an implementer can check before and after each update.

AI USE STATEMENT

We used generative AI for language polishing and for writing code.

REPRODUCIBILITY STATEMENT

Theory results come with complete human-verifiable proofs (Appendices B and C) and numerical regression suites checking every stated inequality on finite configurations (97 and 6,157 checks, zero violations; Appendix F). The protocol, attack encoding, n-sweep, fitting LP, and numerical tolerances are in Appendix I; all experiments are deterministic and sampling-free. The constructions behind the tightness, necessity, and δ_∞ lower-bound results are in Appendix E.

REFERENCES

Nicholas Abou Risk, Duane Szafron, et al. Using counterfactual regret minimization to create competitive multiplayer poker agents. In *AAMAS*, pp. 159–166, 2010.

Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015.

Noam Brown and Tuomas Sandholm. Safe and nested subgame solving for imperfect-information games. *Advances in neural information processing systems*, 30, 2017.

Noam Brown and Tuomas Sandholm. Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

Noam Brown and Tuomas Sandholm. Solving imperfect-information games via discounted regret minimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1829–1836, 2019a.

Noam Brown and Tuomas Sandholm. Superhuman ai for multiplayer poker. *Science*, 365(6456): 885–890, 2019b.

Noam Brown, Sam Ganzfried, and Tuomas Sandholm. Hierarchical abstraction, distributed equilibrium computation, and post-processing, with application to a champion no-limit texas hold'em agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pp. 7–15, 2015.

Noam Brown, Tuomas Sandholm, and Brandon Amos. Depth-limited solving for imperfect-information games. *Advances in neural information processing systems*, 31, 2018.

Noam Brown, Adam Lerer, Sam Gross, and Tuomas Sandholm. Deep counterfactual regret minimization. In *International conference on machine learning*, pp. 793–802. PMLR, 2019.

Noam Brown, Anton Bakhtin, Adam Lerer, and Qucheng Gong. Combining deep reinforcement learning and search for imperfect-information games. *Advances in neural information processing systems*, 33:17057–17069, 2020.

Neil Burch, Michael Johanson, and Michael Bowling. Solving imperfect information games using decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

Neil Burch, Martin Schmid, Matej Moravcik, Dustin Morill, and Michael Bowling. Aivat: A new variance reduction technique for agent evaluation in imperfect information games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Yang Cai and Constantinos Daskalakis. On minmax theorems for multiplayer games. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete algorithms*, pp. 217–234. SIAM, 2011.

Yang Cai, Ozan Candogan, Constantinos Daskalakis, and Christos Papadimitriou. Zero-sum polymatrix games: A generalization of minmax. *Mathematics of Operations Research*, 41(2):648–655, 2016.

Luca Carminati, Federico Cacciamani, Marco Ciccone, and Nicola Gatti. A marriage between adversarial team games and 2-player games: Enabling abstractions, no-regret learning, and subgame solving. In *International Conference on Machine Learning*, pp. 2638–2657. PMLR, 2022.

Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.

Sam Ganzfried. Fast complete algorithm for multiplayer nash equilibrium. In *International Conference on Decision and Game Theory for Security*, pp. 109–123. Springer, 2024.

Sam Ganzfried. Projected exploitability descent for nash equilibrium computation in multiplayer imperfect-information games. *arXiv preprint arXiv:2606.29169*, 2026a.

Sam Ganzfried. Variable bound tightening for nash equilibrium computation in multiplayer imperfect-information games. *arXiv preprint arXiv:2606.25997*, 2026b.

Sam Ganzfried and Tuomas Sandholm. Endgame solving in large imperfect-information games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pp. 37–45, 2015a.

Sam Ganzfried and Tuomas Sandholm. Safe opponent exploitation. *ACM Transactions on Economics and Computation (TEAC)*, 3(2):1–28, 2015b.

Zhenxing Ge, Zheng Xu, Tianyu Ding, Wenbin Li, and Yang Gao. Efficient subgame refinement for extensive-form games. *Advances in Neural Information Processing Systems*, 36:8280–8291, 2023.

Zhenxing Ge, Zheng Xu, Tianyu Ding, Linjian Meng, Bo An, Wenbin Li, and Yang Gao. Safe and robust subgame exploitation in imperfect information games. In *Forty-first International Conference on Machine Learning*, 2024.

Richard Gibson and Duane Szafron. Regret minimization in multiplayer extensive games. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, pp. 2802, 2011.

Kristoffer Arnsfelt Hansen, Peter Bro Miltersen, and Troels Bjerre Sørensen. The computational complexity of trembling hand perfection and other equilibrium refinements. In *International Symposium on Algorithmic Game Theory*, pp. 198–209. Springer, 2010.

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

Christian Kroer and Tuomas Sandholm. Limited lookahead in imperfect-information games. *Artificial Intelligence*, 283:103218, 2020.

Ondrej Kubicek, Viliam Lisy, and Tuomas Sandholm. Equilibrium refinements improve subgame solving in imperfect-information games. *arXiv preprint arXiv:2601.17131*, 2026.

Harold W Kuhn. Extensive games and the problem of information. *Contributions to the Theory of Games*, 2(28):193–216, 1953.

Marc Lanctot, Kevin Waugh, Martin Zinkevich, and Michael Bowling. Monte carlo sampling for regret minimization in extensive games. *Advances in neural information processing systems*, 22, 2009.

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. Real-time parallel counterfactual regret minimization. *arXiv preprint arXiv:2605.19928*, 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. Av-aivat: 74x cheaper agent evaluation with certified anytime-valid stopping in imperfect-information games. *arXiv preprint arXiv:2608.06362*, 2026a.

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

Junjie Li, Sotetsu Koyamada, Qiwei Ye, Guoqing Liu, Chao Wang, Ruihan Yang, Li Zhao, Tao Qin, Tie-Yan Liu, and Hsiao-Wuen Hon. Suphx: Mastering mahjong with deep reinforcement learning. *arXiv preprint arXiv:2003.13590*, 2020.

Chun Kai Ling and Fei Fang. Safe subgame resolving for extensive form correlated equilibrium. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 5116–5123, 2022.

Viliam Lisỳ and Michael Bowling. Equilibrium approximation quality of current no-limit poker bots. In *AAAI Workshops*, 2017.

Mingyang Liu, Chengjie Wu, Qihan Liu, Yansen Jing, Jun Yang, Pingzhong Tang, and Chongjie Zhang. Safe opponent-exploitation subgame refinement. *Advances in Neural Information Processing Systems*, 35:27610–27622, 2022.

Weiming Liu, Haobo Fu, Qiang Fu, and Yang Wei. Opponent-limited online search for imperfect information games. In *International Conference on Machine Learning*, pp. 21567–21585. PMLR, 2023.

Revan MacQueen and James Wright. Guarantees for self-play in multiplayer games via polymatrix decomposability. *Advances in Neural Information Processing Systems*, 36:20230–20259, 2023.

David Milec, Vojtěch Kovařík, and Viliam Lisỳ. Adapting beyond the depth limit: Counter strategies in large imperfect information games. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, pp. 2675–2677, 2025.

Matej Moravcik, Martin Schmid, Karel Ha, Milan Hladik, and Stephen Gaukrodger. Refining subgames in large imperfect information games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisỳ, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

Julien Perolat, Bart De Vylder, Daniel Hennes, Eugene Tarassov, Florian Strub, Vincent de Boer, Paul Muller, Jerome T Connor, Neil Burch, Thomas Anthony, et al. Mastering the game of stratego with model-free multiagent reinforcement learning. *Science*, 378(6623):990–996, 2022.

Martin Schmid, Matej Moravčík, Neil Burch, Rudolf Kadlec, Josh Davidson, Kevin Waugh, Nolan Bard, Finbarr Timbers, Marc Lanctot, G Zacharias Holland, et al. Student of games: A unified learning algorithm for both perfect and imperfect information games. *Science Advances*, 9(46): eadg3256, 2023.

Michal Šustr, Vojtěch Kovařík, and Viliam Lisỳ. Monte carlo continual resolving for online strategy computation in imperfect information games. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 224–232, 2019.

Michal Šustr, Martin Schmid, Matej Moravčík, Neil Burch, Marc Lanctot, and Michael Bowling. Sound algorithms in imperfect information games. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1674–1676, 2021.

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

John Von Neumann and Oskar Morgenstern. Theory of games and economic behavior: 60th anniversary commemorative edition. In *Theory of games and economic behavior*. Princeton university press, 2007.

Brian Zhang and Tuomas Sandholm. Subgame solving without common knowledge. *Advances in Neural Information Processing Systems*, 34:23993–24004, 2021.

Brian Zhang, Luca Carminati, Federico Cacciamani, Gabriele Farina, Pierriccardo Olivieri, Nicola Gatti, and Tuomas Sandholm. Subgame solving in adversarial team games. *Advances in Neural Information Processing Systems*, 35:26686–26697, 2022.

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, 2023a.

Youzhi Zhang and Bo An. Computing team-maxmin equilibria in zero-sum multiplayer extensive-form games. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 2318–2325, 2020.

Youzhi Zhang, Bo An, and Jakub Černỳ. Computing ex ante coordinated team-maxmin equilibria in zero-sum multiplayer extensive-form games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 5813–5821, 2021.

Youzhi Zhang, Bo An, and Venkatramanan Subrahmanian. Computing optimal nash equilibria in multiplayer games. *Advances in Neural Information Processing Systems*, 36:21230–21254, 2023b.

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 Additional Related Work

Safe exploitation and margin certificates. Safe exploitation bounds the value lost in deviating from an equilibrium to exploit an assumed-weak opponent (Ganzfried & Sandholm, 2015b), and it certifies the deployed strategy against a reference before play (Li & Huang, 2026a). Recent work brings the same discipline into subgame refinement. The refined strategy must never worsen guaranteed value (Liu et al., 2022) or lose robustness across opponent classes (Ge et al., 2024). The subgame itself can shrink through opponent-limited search (Liu et al., 2023) and adaptation beyond the depth limit (Milec et al., 2025). Our margin certificate plays a different role. It is a post-hoc verifier of a chosen update, built from one exact subgame best response per opponent, pricing every alarm by counterfactual mass. Practical exploitability measures for large games give only lower bounds (Lisý & Bowling, 2017); our Vul_G is exactly computable.

Adversarial team games. Adversarial team games bridge multiplayer play to solvable structure: a coordinating team faces a single adversary, and only the team's internal coordination prevents two-sided reasoning. Subgame solving extends to this setting (Zhang et al., 2022), and the team belief DAG enables correlated team-maxmin equilibria (Carminati et al., 2022; Zhang et al., 2023a); we keep the complementary setting, no team and self-interested players.

Blueprint computation and multiplayer equilibria. Resolving presupposes a blueprint, itself a solver output: CFR converges to an equilibrium in two-player zero-sum games (Zinkevich et al., 2007), scaled by Monte Carlo variants (Lanctot et al., 2009), discounted regret schedules (Brown & Sandholm, 2019a), and neural approximation (Brown et al., 2019), the structure preserved in search-plus-RL systems (Brown et al., 2020). Solver engineering keeps scaling this pipeline through parallel regret minimization (Li & Huang, 2026b), correlated chance sampling (Li et al., 2026b), and action abstraction (Li et al., 2024; Li & Huang, 2025). For multiplayer games the equilibrium target deteriorates: computing an equilibrium is PPAD-complete (Daskalakis et al., 2009), equilibria are not interchangeable, and multiplayer exploitability admits no canonical definition, hence profile-level vulnerability. Multiplayer regret minimization is more recent (Abou Risk et al., 2010; Gibson & Szafron, 2011), with Kuhn poker (Kuhn, 1953) as the standard small testbed and direct multiplayer

Nash algorithms (Ganzfried, 2024; 2026b;a; Zhang et al., 2023b). None addresses the safety of a refinement update.

Polymatrix games. Zero-sum polymatrix games form a tractable island, with polynomial-time equilibria (Cai et al., 2016), minmax-type theorems beyond two players (Cai & Daskalakis, 2011), and a distance-to-CSP measure under which self-play guarantees degrade gracefully (MacQueen & Wright, 2023). Here that distance serves as an a priori upper bound on the structural tax in a safety bound for resolving.

Positioning. Each nearest neighbor supplies one axis. Two-player safe resolving guarantees safety only in two-player zero-sum games (Burch et al., 2014; Brown & Sandholm, 2017). Pluribus operates in multiplayer games without a safety guarantee (Brown & Sandholm, 2019b). Adversarial team games obtain multiplayer safety under a team structure we exclude (Zhang et al., 2022). Correlated resolving certifies a correlation plan (Ling & Fang, 2022). Safe exploitation conditions on an assumed opponent model (Ganzfried & Sandholm, 2015b). Polymatrix decomposability supplies the distance measure (MacQueen & Wright, 2023). Our result joins these axes: a safety guarantee for multiplayer play, applied to the delivered profile, with the distance measure tied to resolving. Section 4 occupies the intersection.

B Proof of the Exact Case

This appendix contains the full proof of Lemma 1, the complete constructions behind Examples 1 and 2, and the two further constructions of Example 3 that attain the bound of Theorem 1 on its structural-tax and solve-error terms. Notation follows Section 3; we first restate all objects at the level of detail the proof consumes, then give the proof in numbered steps, and close with the four equality constructions. Throughout, $(x)_+ := \max\{x, 0\}$.

B.1 Setting, Objects, and Assumptions

Game, subgame, entries. G is an n-player finite extensive-form game with perfect recall, chance acting first, and a public label $\mathrm{pub}(h)$, the information visible to all players, on every non-terminal history. Players are $N = \{0, \dots, n-1\}$; the resolver is i and the opponents are $O = N \setminus \{i\}$. Terminals are Z with utilities $u_p(z)$. Fix a public state pub^*. The entry set E consists of the earliest histories e with $\mathrm{pub}(e) = \mathrm{pub}^*$; each entry carries its chance probability c_e and the per-player blueprint reaches $r_q(e) = \pi_q^{\sigma^0}(e)$. The subgame S is the set of descendants of E; $Z_S \subset Z$ are the terminals passing through S and Z_{out} the rest.

Entry information sets and counterfactual entry mass. For each player q, group E by $\iota_q(e) := (\mathrm{priv}_q(e), \mathrm{pub}^*)$ and write $\iota \in I_q$ for the resulting entry information sets; these are exactly the information sets at which the gadget places q's T/F decision. The counterfactual entry mass is

$$\bar{W}_q(\iota) := \sum_{e \in \iota} c_e \, \pi_{-q}^{\sigma^0}(e), \qquad \pi_{-q}^{\sigma^0}(e) := \prod_{p \neq q} r_p(e). \tag{2}$$

The resolved profile σ^1 satisfies $\sigma^1_p = \sigma^0_p$ for all $p \neq i$, and σ^1_i differs from σ^0_i only at information sets of i inside S. The lemma holds for *any* such update: every certificate quantity below is defined post hoc for the delivered σ^1.

Per-player gains. For a profile σ let $A_p(\sigma) := \max_{\tau_p} u_p(\tau_p, \sigma_{-p}) - u_p(\sigma)$, so that $\mathrm{Vul}_G(\sigma) = \max_p A_p(\sigma)$ (Definition 3). Write $\mathrm{gain}_p := u_p(\sigma^1) - u_p(\sigma^0)$ and $\mathrm{drop}_p := -\mathrm{gain}_p$.

We use the CFR-D-style joint gadget (Burch et al., 2014; Moravcik et al., 2016; Brown & Sandholm, 2017):

- **(G1) Joint gadget.** A root chance node enters entry e with probability proportional to $c_e \cdot r_i(e)$ (opponent reach excluded, as in robust/maxmargin entry). Below each entry, every opponent j receives, in a fixed order, a binary T/F decision node with information set $\iota_j(e)$. An opponent choosing T is frozen to its blueprint behavior inside the subgame copy and its terminal payoff coordinate is replaced by the constant $\mathrm{CFV}_j(\iota)$; an opponent choosing F acts freely and receives its true payoff; the resolver always acts freely.

- **(G2) Blueprint counterfactual values.**

$$\mathrm{CFV}_j(\iota) := \frac{\sum_{e\in\iota} c_e\, \pi^{\sigma^0}_{-j}(e)\cdot \mathrm{EV}^{\sigma^0}_j(e)}{\bar{W}_j(\iota)}, \tag{3}$$

where $\mathrm{EV}^{\sigma^0}_j(e)$ is j's expected value from e under the all-blueprint profile. CFV_j and $\hat{m}_j$ are defined only at groups with $\bar{W}_j(\iota) > 0$; a zero-mass group has $c_e\,\pi^{\sigma^0}_{-j}(e) = 0$ at each of its entries, so it contributes 0 to every entry sum below and is omitted from M_j and from the margin-certified condition (Appendix J).

- **(G4) Subgame best-response values (game-side measure).** For $j \in O$, $\iota \in I_j$, and a subgame strategy x_i of the resolver,

$$\mathrm{SBV}_j(\iota;\, x_i) := \max_{\tau^S_j} \sum_{e\in\iota} c_e\, \pi^{\sigma^0}_{-j}(e)\cdot \mathbb{E}\big[\, u_j \mid e,\ (\tau^S_j,\, x_i,\, \sigma^0_{-ij}) \text{ inside } S\,\big], \tag{4}$$

the maximum ranging over j's behavioral strategies inside S; the environment inside S is x_i for the resolver and blueprint for everyone else (σ^0_{-ij} denotes the blueprint of all players other than i and j).

- **(G5) BR-certified margin and margin deficit** (Definitions 1 and 2):

$$\hat{m}_j(\iota) := \mathrm{CFV}_j(\iota) - \frac{\mathrm{SBV}_j(\iota;\, \sigma^1_i)}{\bar{W}_j(\iota)}, \qquad M_j := \sum_{\iota\in I_j} \bar{W}_j(\iota)\,\big(-\hat{m}_j(\iota)\big)_+. \tag{5}$$

The deficit is weighted by $\bar{W}_j$, never divided by it. Its normalization depends on the entry structure: $\sum_\iota \bar{W}_j(\iota) = \sum_{e\in E} c_e\,\pi^{\sigma^0}_{-j}(e)$ omits j's own reach, so if j has a private pre-entry action whose branches all reach pub^*, each branch contributes its full environmental mass and the sum can exceed 1. If instead j follows one action sequence into every $e \in E$ (as when all actions before pub^* are public), let τ_j play that sequence; then $\pi^\tau_j(e) = 1$ for all $e \in E$ and $\sum_\iota \bar{W}_j(\iota) = \Pr_{(\tau_j,\sigma^0_{-j})}[\text{reach } E] \le 1$, since entries are pairwise non-nested histories; hence $M_j \le \big(-\min_\iota \hat{m}_j(\iota)\big)_+$ and the weighted-sum caliber is directly comparable to a min-margin caliber.

- **(G7) Activation and solve error.**

$$\begin{aligned} A^S_i(\sigma^0) &:= \max_{\tau_i\,:\,\tau_i=\sigma^0_i \text{ outside } S} u_i(\tau_i,\sigma^0_{-i}) - u_i(\sigma^0) \quad \big(\le A_i(\sigma^0) \le \mathrm{Vul}_G(\sigma^0)\big), \\ \varepsilon_i &:= A^S_i(\sigma^0) - \mathrm{gain}_i. \end{aligned} \tag{6}$$

Here $\varepsilon_{\mathrm{solve}}$ is a bound on the solve error the delivered resolver may carry: $\varepsilon_i \le \varepsilon_{\mathrm{solve}}$ by assumption, and the corollary below replaces ε_i by this budget when only the solver's accuracy class is known. Because σ^1_i is a feasible point of the maximization defining A^S_i, $\varepsilon_i \ge 0$ *always holds*. The quantity ε_i measures the suboptimality of σ^1_i against blueprint opponents inside S: it is 0 for an exact gadget solve in a Terminate-compatible environment whose root measure induces the resolver's game-side beliefs at each of its entry sets (all opponents choosing T makes the gadget environment coincide with the blueprint, and $c_e\,\pi_{-i}(e)$ or the full blueprint reach both give the correct beliefs because i's own reach is constant on each of its information sets; the implementation measure $c_e\, r_i(e)$ of Appendix J does not in general), while for iterative implementations it comprises solve error plus gadget-environment drift, and can be measured exactly post hoc.

Remark 1 (Margin caliber; implementation deviation). The margins in Eq. (5) are *BR-certified*: the F-branch is evaluated by an exact subgame best response of j, under the game-side measure $c_e\,\pi^{\sigma^0}_{-j}(e)$ and the environment $(\sigma^1_i, \sigma^0_{-ij})$. Common implementations instead evaluate the F-branch with j's own gadget average strategy under the gadget-internal measure. That average-strategy margin exceeds $\hat{m}$ on 33 of 40 entry information sets in our numerical regression check, with worst overstatement $+1.667$ on an off-path branch; the theorem is stated with $\hat{m}$ (Appendix J).

Assumption 1 consists of: *(P1)* every information set of every player is either entirely inside S or disjoint from it (the standard property of public-state subgames); *(P2)* the gadget's entry grouping

$\iota_q = (\mathrm{priv}_q, \mathrm{pub}^*)$ agrees with q's true information at entry, and q's own reach $\pi_q^\tau(e)$ is constant across each ι (automatic under perfect recall with public actions).

(F2, exactness) $\delta_\infty(G) = 0$ in the LP profile-table semantics of Definition 4: there exists a table $\{u_{pq}\}$ on reduced pure-strategy pairs with

$$\begin{aligned} u_p(\rho) &= \sum_{q \neq p} u_{pq}(\rho_p, \rho_q) \ \text{ for every pure profile } \rho, \\ u_{pq}(a,b) + u_{qp}(b,a) &= c_{pq} \ \text{ for edge constants } c_{pq}, \end{aligned} \tag{7}$$

i.e. G is exactly a constant-sum polymatrix (CSP) game (MacQueen & Wright, 2023).

Remark 2 (Two definitional conventions). (i) The CSP class is taken with the *per-edge constant-sum* constraint of Eq. (7) (with the c_{pq} free), following MacQueen & Wright (2023) and Cai et al. (2016). This class contains the antisymmetric special case $u_{pq} + u_{qp} = 0$. The only property the proof uses is that $\sum_p u_p$ is constant on the class, and constants cancel in the difference taken in Step S7 below. (ii) The tables u_{pq} are defined on reduced pure-strategy pairs (LP profile-table semantics): a literal terminal-level reading, in which $u_{pq}(z)$ may depend arbitrarily on the full terminal history, would make every pointwise zero-sum game trivially CSP via $u_{pq} := (u_p - u_q)/n$, contradicting the measured lower bound $\delta \geq 1$ for three-player Kuhn poker. Dependence on third-party randomness is allowed only through the chance expectation; dependence on third-party strategies is excluded (Appendix J).

B.2 Proof of Lemma 1

We prove Lemma 1: under Assumption 1 and $\delta_\infty(G) = 0$, for any blueprint σ^0 and any update σ^1 that changes only i's strategy inside S,

$$\text{(i)} \quad A_i(\sigma^1) - A_i(\sigma^0) = \varepsilon_i - A_i^S(\sigma^0) \leq \varepsilon_i; \tag{8}$$

$$\text{(ii)} \quad A_j(\sigma^1) - A_j(\sigma^0) \leq M_j + \Big[A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{k \in O \setminus \{j\}} M_k\Big] \tag{9}$$

$$\leq A_i^S(\sigma^0) + \textstyle\sum_{q \in O} M_q \qquad \forall j \in O; \tag{10}$$

$$\text{(iii)} \quad \mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \leq \max\Big(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{q \in O} M_q\Big). \tag{11}$$

In particular, if $\hat{m}_j(\iota) \geq 0$ for all j, ι (margin-certified) and $\varepsilon_i \leq \varepsilon_{\mathrm{solve}}$, then $\mathrm{Vul}_G(\sigma^1) \leq \mathrm{Vul}_G(\sigma^0) + \max(\varepsilon_{\mathrm{solve}}, A_i^S(\sigma^0))$. The constants are: margin coefficient $c_2 = 1$ (sum over all opponents, weighted-sum caliber, no inverse reach); blueprint coefficient 1 on $A_i^S(\sigma^0)$ (equivalently, coefficient 2 on $\mathrm{Vul}_G(\sigma^0)$ after weakening via $A_i^S \leq \mathrm{Vul}_G(\sigma^0)$); and $D = 1$ on the solve-error term, since only the resolver's solve quality enters (the opponent-side certificates hold exactly for the delivered σ^1).

Proof. Fix σ^0, σ^1, and pub^*; all maxima below are over finite sets and are attained. Write σ^x for σ^0 if $x = 0$ and σ^1 if $x = 1$.

Step S0 (terminal decomposition; uses P1). For any profile π the expected utility splits over Z_{out} and Z_S; a path through S determines its entry $e \in E$ uniquely. For $z \in Z_S$, $P^\pi(z) = c(z) \prod_p \pi_p(z)$, and the trunk segment (before e) of every player's behavior is identical under σ^0 and σ^1, because σ^1 changes only i's information sets inside S.

Step S1 (trunk–subgame decomposition of opponent strategies; uses P1, P2, perfect recall). For $j \in O$ and any strategy τ_j, decompose $\tau_j = (\tau_j^{\mathrm{tr}}, \{\tau_j^S(\iota)\}_{\iota \in I_j})$, its trunk behavior (everything outside S) and its continuations inside S: perfect recall lets j choose its continuation inside S independently at each ι, and P2 makes j's own reach constant on ι, written $\pi_j^\tau(\iota) \in [0,1]$. Then for $x \in \{0,1\}$,

$$u_j(\tau_j, \sigma_{-j}^x) = \mathrm{OUT}_j(\tau_j^{\mathrm{tr}}) + \sum_{\iota \in I_j} \pi_j^\tau(\iota) \cdot G_j^x\big(\iota;\, \tau_j^S(\iota)\big), \tag{12}$$

where OUT_j is the payoff collected on terminals outside S and

$$G_j^x(\iota;\, \tau^S) := \sum_{e \in \iota} c_e\, \pi_{-j}^{\sigma^0}(e) \cdot \mathbb{E}\big[\, u_j \mid e,\ (\tau^S,\, \sigma_i^x,\, \sigma_{-ij}^0)\,\big], \tag{13}$$

and OUT_j and the $\pi_j^\tau(\iota)$ do not depend on x (the trunk and Z_{out} are identical under both profiles).

Step S2 (blueprint-value identity; by definition G2). $G_j^0(\iota;\, \sigma_j^0|_S) = \bar{W}_j(\iota) \cdot \mathrm{CFV}_j(\iota)$, by expanding Eq. (3) entry by entry.

Step S3 (BR-certificate identity; by definitions G4, G5).

$$\max_{\tau^S} G_j^1(\iota;\, \tau^S) = \mathrm{SBV}_j(\iota;\, \sigma_i^1) = \bar{W}_j(\iota) \cdot \big[\mathrm{CFV}_j(\iota) - \hat{m}_j(\iota)\big]. \tag{14}$$

Step S4 (opponent BR-value change is at most M_j). Let $\mathrm{BRval}_j(x) := \max_{\tau_j} u_j(\tau_j, \sigma_{-j}^x)$. Take τ_j optimal for $x = 1$. By S1,

$$\begin{aligned}
\mathrm{BRval}_j(1) &= \mathrm{OUT}_j + \sum_\iota \pi_j(\iota)\, G_j^1\big(\iota;\, \tau^S(\iota)\big) \\
&\le \Big[\mathrm{OUT}_j + \sum_\iota \pi_j(\iota)\, \bar{W}_j(\iota)\, \mathrm{CFV}_j(\iota)\Big] \\
&\qquad + \sum_\iota \pi_j(\iota)\, \bar{W}_j(\iota)\, \big(-\hat{m}_j(\iota)\big)_+ \qquad\qquad \text{(S3; } (-\hat{m})_+ \ge -\hat{m}\text{)} \\
&\le \mathrm{BRval}_j(0) + M_j,
\end{aligned} \tag{15}$$

where the last inequality uses $\pi_j(\iota) \le 1$ and the fact that the bracketed term is, by S2, the value of the feasible strategy $(\tau_j^{\mathrm{tr}}, \text{blueprint continuations})$ against σ_{-j}^0, hence at most $\mathrm{BRval}_j(0)$. Thus $\mathrm{BRval}_j(1) - \mathrm{BRval}_j(0) \le M_j$. No CSP structure is used; this step holds for arbitrary δ.

Step S5 (opponent on-path gain is at most M_k). Here on-path gain means the change in k's expected payoff under the profile actually played, with k following the blueprint. For $k \in O$, expand $\mathrm{gain}_k = u_k(\sigma^1) - u_k(\sigma^0)$ via S1 with $x = 1$ versus $x = 0$ and $\tau_k = \sigma_k^0$ (the OUT and $\pi_k(\iota)$ terms cancel):

$$\begin{aligned}
\mathrm{gain}_k &= \sum_\iota \pi_k^{\sigma^0}(\iota)\big[G_k^1(\iota;\, \sigma_k^0|_S) - G_k^0(\iota;\, \sigma_k^0|_S)\big] \\
&\le \sum_\iota \pi_k^{\sigma^0}(\iota)\big[\mathrm{SBV}_k(\iota;\, \sigma_i^1) - \bar{W}_k(\iota)\, \mathrm{CFV}_k(\iota)\big] \qquad\qquad \text{(S2)} \\
&= \sum_\iota \pi_k^{\sigma^0}(\iota)\, \bar{W}_k(\iota)\, \big(-\hat{m}_k(\iota)\big) \;\le\; M_k.
\end{aligned} \tag{16}$$

The inequality steps: the blueprint continuation is a feasible point in the maximization defining $\mathrm{SBV}_k(\iota; \sigma_i^1) = \max_{\tau^S} G_k^1(\iota; \tau^S)$, which gives the first inequality; the bracket then equals $-\bar{W}_k(\iota)\, \hat{m}_k(\iota)$ by S3 (it may have either sign), and $x \le (x)_+$ with $\pi_k^{\sigma^0}(\iota) \le 1$ gives the bound by M_k. This is the giveaway-detection mechanism: if i cedes value to k inside S, then k's Follow-and-best-respond value collects at least the ceded amount and $\hat{m}_k$ turns negative accordingly; cross-opponent transfers always leave a margin trace.

Step S6 (resolver identity; by definition G7). $\mathrm{gain}_i = A_i^S(\sigma^0) - \varepsilon_i$ with $\varepsilon_i \ge 0$. Moreover the maximization term in A_i does not depend on i's own strategy and $\sigma_{-i}^1 = \sigma_{-i}^0$, so $A_i(\sigma^1) - A_i(\sigma^0) = -\mathrm{gain}_i = \varepsilon_i - A_i^S(\sigma^0)$. This proves Eq. (8).

Step S7 (the only use of $\delta = 0$: the profile sum is constant). By F2, for every pure profile ρ,

$$\sum_p u_p(\rho) = \sum_p \sum_{q \ne p} u_{pq}(\rho_p, \rho_q) = \sum_{p<q} \big[u_{pq}(\rho_p, \rho_q) + u_{qp}(\rho_q, \rho_p)\big] = \sum_{p<q} c_{pq} =: c^*, \tag{17}$$

a constant independent of ρ; taking expectations, $\sum_p u_p(\sigma) = c^*$ for every mixed profile (a product of distributions over pure strategies) as well, hence for every behavioral profile, since under perfect recall each behavioral strategy is realization-equivalent to a mixed one (Kuhn, 1953). Hence the constant cancels in the difference and

$$\sum_p \mathrm{gain}_p = 0. \tag{18}$$

Table 3: Payoffs of the activation construction (Example 1). Player j acts in the trunk; player i acts inside S.

j (trunk)	i (inside S)	u_i	u_j
d	$\cdot$	-1	$+1$
enter	a'	0	0
enter	a	$+1$	-1

Step S8 (pricing the victim's on-path loss). For $j \in O$, by Eq. (18),

$$\mathrm{drop}_j = -\mathrm{gain}_j = \sum_{p \neq j} \mathrm{gain}_p = \mathrm{gain}_i + \sum_{k \in O \setminus \{j\}} \mathrm{gain}_k \leq \big(A_i^S(\sigma^0) - \varepsilon_i\big) + \sum_{k \in O \setminus \{j\}} M_k, \quad (19)$$

using S5 and S6.

Step S9 (total change for each opponent). Since $A_j(\sigma) = \mathrm{BRval}_j(\sigma) - u_j(\sigma)$,

$$A_j(\sigma^1) - A_j(\sigma^0) = \big[\mathrm{BRval}_j(1) - \mathrm{BRval}_j(0)\big] + \mathrm{drop}_j \leq M_j + \big(A_i^S(\sigma^0) - \varepsilon_i\big) + \sum_{k \neq j} M_k, \quad (20)$$

by S4 and S8; the further weakening to $A_i^S(\sigma^0) + \sum_{q \in O} M_q$ uses $\varepsilon_i \geq 0$. This proves Eq. (10).

Step S10 (total bound). A difference of maxima is at most the maximum of differences:

$$\begin{aligned} \mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) &\leq \max_p \big[A_p(\sigma^1) - A_p(\sigma^0)\big] \\ &\leq \max\Big(\varepsilon_i - A_i^S(\sigma^0),\ \max_{j \in O}\big[M_j + A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{k \neq j} M_k\big]\Big) \\ &\leq \max\Big(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{q \in O} M_q\Big). \end{aligned} \quad (21)$$

This proves Eq. (11). □

Dependency accounting. Steps S1–S4 use only the entry decomposition, the blueprint counterfactual values, and the game-side best responses defining $\hat{m}$ (Definition 1), with no CSP assumption and without executing the gadget; S5 uses the same objects plus blueprint feasibility (no CSP); S6 is definitional; S7 is the *only* step using $\delta = 0$; S8–S10 are algebra. This accounting is what Appendix C consumes: the $\delta > 0$ extension replaces S7 alone.

B.3 Example 1: the Activation Term Cannot Be Removed

We give the full construction behind Example 1: a two-player zero-sum game (hence $\delta_\infty = 0$) with all margins certified strictly positive and $\varepsilon_i = 0$, in which $\Delta\mathrm{Vul} = A_i^S(\sigma^0) > 0$ with equality in Eq. (11).

Construction. Players $\{i, j\}$. Player j moves first at the root: action d ends the game with payoffs $(u_i, u_j) = (-1, +1)$; action *enter* enters the subgame S. Inside S, only i acts: a' yields $(0, 0)$ and a yields $(+1, -1)$. Table 3 lists all terminals. The blueprint is $\sigma^0 = (a', \mathit{enter})$.

Blueprint quantities. Under σ^0, payoffs are $(0, 0)$. Player i's unrealized subgame value is a: $A_i(\sigma^0) = A_i^S(\sigma^0) = 1$. Player j's unrealized deviation is d: $A_j(\sigma^0) = 1$. Hence $\mathrm{Vul}_G(\sigma^0) = 1$.

The update and its certificates. Take $\sigma_i^1 = a$. Then $\mathrm{CFV}_j = 0$ (blueprint value of the entry for j), and $\mathrm{SBV}_j(\sigma_i^1) = -1$ (j has no action inside S), so

$$\hat{m}_j = 0 - (-1) = +1 > 0 \qquad \text{(certified)}, \qquad \varepsilon_i = A_i^S(\sigma^0) - \mathrm{gain}_i = 1 - 1 = 0.$$

Outcome and equality. Under σ^1, payoffs are $(+1, -1)$, so $A_j(\sigma^1) = \max(+1, -1) - (-1) = 2$ and $A_i(\sigma^1) = 0$, giving $\mathrm{Vul}_G(\sigma^1) = 2 = \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0)$. The bound Eq. (11) evaluates to $\max(\varepsilon_i,\ A_i^S - \varepsilon_i + \sum_q M_q) = \max(0, 1 + 0) = 1$: *equality holds*, so the coefficient 1 on $A_i^S(\sigma^0)$ is

Table 4: Total payoffs induced by the witness Eq. (22) (Example 2); k has no action. Every row sums to 0.

i	j	u_i	u_j	u_k
a'	$stay$	0	0	0
a'	d	0	0	0
a	$stay$	0	-1	$+1$
a	d	-1	0	$+1$

optimal. All-positive margins cannot prevent the effect: the resolver realizes its own blueprint slack, and the victim's best-response gain rises by exactly that amount. The same phenomenon occurs in two-player CFR-D and maxmargin resolving, whose theorems bound the resolver's exploitability and leave the profile vulnerability out of their scope. The activation term is therefore intrinsic to the profile-vulnerability measure of Definition 3, and it already appears in two-player zero-sum games.

B.4 Example 2: Cross-Opponent Transfer and the Necessity of the Sum

We give the full construction behind Example 2: a three-player one-shot game with $\delta_\infty = 0$, $\varepsilon_i = 0$, and the deviating opponent's margin exactly zero, in which the literal per-pair bound fails while the per-player bound Eq. (11) holds with equality.

Construction. Players $\{i, j, k\}$ act simultaneously once: $i \in \{a', a\}$, $j \in \{stay, d\}$, k has no action. The game is defined through a pair-local CSP witness (all edges antisymmetric, edge constants $c_{pq} = 0$):

$$u_{ij}(a, stay) = +1, \quad u_{ij}(a, d) = 0, \quad u_{ij}(a', \cdot) = 0, \quad u_{ik}(a, \cdot) = -1, \\ u_{ik}(a', \cdot) = 0, \quad u_{jk} \equiv 0, \tag{22}$$

so $\delta_\infty = 0$ exactly (and the pair-locality assumptions of the per-pair corollaries in Appendix H hold as well). The induced total utilities are listed in Table 4.

Blueprint quantities. The blueprint is $\sigma^0 = (a', stay)$, with payoffs $(0, 0, 0)$. Every player's best-response gain is 0, so $\mathrm{Vul}_G(\sigma^0) = 0$; and $A_i^S(\sigma^0) = 0$, since switching to a changes i's own total by $+1 - 1 = 0$ (i has no incentive).

The update and its certificates. Take $\sigma_i^1 = a$, a legal update (the subgame here is the whole one-shot interaction, with a single trivial entry of mass $\bar{W}_q = 1$ for each opponent). Then $\varepsilon_i = 0$, and:

$$\hat{m}_j = \mathrm{CFV}_j - \mathrm{SBV}_j(\sigma_i^1) = 0 - \max(-1, 0) = 0 \qquad (j\text{'s subgame BR is } d, \text{ worth } 0 = \mathrm{CFV}_j),$$
$$\hat{m}_k = \mathrm{CFV}_k - \mathrm{SBV}_k(\sigma_i^1) = 0 - (+1) = -1, \qquad M_k = 1, \quad M_j = 0.$$

Failure of the literal per-pair form. The pair vulnerability of j against i on the (j, i) pair table jumps: $\Delta\mathrm{PairVul}_{ji} = \max(0, -1) - (-1) = 1 > 0 = (-\hat{m}_j)_+$. The mechanism: i's update transfers value across boards (taking 1 from j on the (i, j) board while giving 1 to k on the (i, k) board, leaving i's own total unchanged). The victim j's margin stays silent (j's Follow-and-best-respond is blocked back exactly to CFV_j); the alarm fires at the *beneficiary* k ($\hat{m}_k = -1$). A per-pair certificate keyed to the victim's own margin is therefore unobtainable under the joint gadget; the theorem prices the trade at the per-player level, through the beneficiary's margin and the resolver's blueprint slack in Eq. (10).

Equality of the per-player bound. $A_j(\sigma^1) = \max(0, -1) - (-1) = 1$, so $\Delta\mathrm{Vul} = 1 = \sum_q M_q$ (from $M_k = 1$), while the bound Eq. (11) evaluates to $\max(0,\ 0 - 0 + 1) = 1$: *equality holds again*. This shows the sum-over-all-opponents structure in Eqs. (10) and (11) cannot be weakened to the victim's own margin (which would report nothing here), and pins the margin coefficient $c_2 = 1$.

B.5 Example 3: the Tax and Solve-Error Terms Are Attained

The two constructions of Example 3 complete the picture: together with Examples 1 and 2, every term of Theorem 1(iii) and both branches of its maximum are attained with equality. Both games

satisfy Assumption 1 trivially (one entry, no private information), and in both the certificate quantities are computed for the delivered update.

Leak construction ($\Delta\mathrm{Vul} = \Theta = \Lambda$). Players $\{i, j\}$. Player j moves first: d ends the game with $(0, 0)$; *enter* reaches the single entry of S, where i alone acts: a' yields $(0, 0)$ and a yields $(0, -1)$. The blueprint is $\sigma^0 = (a', \mathit{enter})$ and the update is $\sigma_i^1 = a$. ***Blueprint quantities:*** all payoffs are 0 under σ^0, both actions of i are worth 0 to i, and both actions of j are worth 0 to j, so $A_i(\sigma^0) = A_j(\sigma^0) = 0$, $\mathrm{Vul}_G(\sigma^0) = 0$, and $A_i^S(\sigma^0) = 0$. ***Certificates:*** $\mathrm{gain}_i = 0$, so $\varepsilon_i = 0$; $\bar{W}_j = c_e \pi_i^{\sigma^0}(e) = 1$, $\mathrm{CFV}_j = 0$, $\mathrm{SBV}_j(\sigma_i^1) = -1$ (j has no action inside S), so $\hat{m}_j = +1$ and $M_j = 0$: the update is margin-certified. ***Structure:*** the four pure profiles have $U(s) \in \{0, 0, 0, -1\}$, so $\zeta_{\mathrm{prof}} = \frac{1}{2}$; Z_S has $U(z) \in \{0, -1\}$, so $\zeta_S = \frac{1}{2}$; $\rho_S = 1$; hence $\Theta = 2\min(\frac{1}{2}, \frac{1}{2}) = 1$, and the realized leak is $\Lambda = \sum_p u_p(\sigma^0) - \sum_p u_p(\sigma^1) = 0 - (-1) = 1$. ***Outcome:*** under σ^1 the payoffs are $(0, -1)$, so $A_i(\sigma^1) = 0$ and $A_j(\sigma^1) = \max(0, -1) - (-1) = 1$: $\mathrm{Vul}_G(\sigma^1) = 1$. The bound evaluates to $\max(\varepsilon_i,\ A_i^S - \varepsilon_i + \sum_q M_q + \Theta) = \max(0,\ 0 + 0 + 1) = 1 = \Delta\mathrm{Vul}$, ***equality***, and the same holds with Λ in place of Θ (Corollary 1). The mechanism is a transfer with no counterparty: the resolver destroys one unit of j's utility at no cost to itself, so no margin, activation, or solve-error term can see it, and only the total-utility oscillation prices it. Since the game has two players, its δ_∞-CSP distance is the distance to a constant-sum table, $\delta_\infty = \zeta_{\mathrm{prof}}/2 = \frac{1}{4}$, so $2n\,\delta_\infty = 1 = \Theta$: at $n = 2$, Corollary 5 is attained with equality as well.

Self-harm construction ($\Delta\mathrm{Vul} = \varepsilon_i$). Players $\{i, j\}$, zero-sum (so $\delta_\infty = 0$ and $\Theta = 0$); the subgame is the whole game and i alone acts: a yields $(+1, -1)$ and a' yields $(0, 0)$. The blueprint is $\sigma_i^0 = a$ and the update is $\sigma_i^1 = a'$. ***Blueprint quantities:*** a is optimal for i, so $A_i(\sigma^0) = A_i^S(\sigma^0) = 0$; j has no action, so $A_j \equiv 0$; $\mathrm{Vul}_G(\sigma^0) = 0$. ***Certificates:*** $\mathrm{gain}_i = -1$, so $\varepsilon_i = A_i^S(\sigma^0) - \mathrm{gain}_i = 1$; $\bar{W}_j = 1$, $\mathrm{CFV}_j = -1$, $\mathrm{SBV}_j(\sigma_i^1) = 0$, so $\hat{m}_j = -1$ and $M_j = 1$. ***Outcome:*** $A_i(\sigma^1) = \max(1, 0) - 0 = 1$, so $\mathrm{Vul}_G(\sigma^1) = 1$ and $\Delta\mathrm{Vul} = 1 = \varepsilon_i$. Claim (i) holds with equality ($A_i(\sigma^1) - A_i(\sigma^0) = 1 = \varepsilon_i - A_i^S$), claim (ii) holds with equality ($0 \le M_j + A_i^S - \varepsilon_i = 0$), and the bound (iii) evaluates to $\max(\varepsilon_i,\ A_i^S - \varepsilon_i + M_j + \Theta) = \max(1,\ 0) = 1$: the ε_i branch of the maximum is attained. The resolver's own loss is what raises the vulnerability; the opponent's margin deficit records the value handed over but is cancelled inside the second branch by $-\varepsilon_i$.

C Proof of the Transfer Lemma and Main Theorem

This appendix proves Lemma 2 (the transfer lemma, both legs) and assembles it with the exact-case skeleton of Appendix B into Theorem 1 and Corollary 5. It closes with the reason δ enters only through the total utility rather than through a proxy constant-sum polymatrix (CSP) game. All objects of Appendix B.1 are carried over unchanged: entries E, subgame S, terminals Z_S/Z_{out}, entry information sets ι, masses $\bar{W}$, values CFV, margins $\hat{m}$, deficits M_q, activation A_i^S, solve error ε_i, gains gain_p and drop_p, best-response gains A_p, and Vul_G.

C.1 Additional Objects

Profile tables and the CSP subspace. Let $s = (s_1, \ldots, s_n)$ range over reduced pure-strategy profiles and let $u_p(s)$ denote p's utility after taking the expectation over chance (the entries of the profile utility tensor). Write

$$U(s) := \sum_p u_p(s), \qquad U(z) := \sum_p u_p(z) \ \text{ for terminals } z. \tag{23}$$

Let V be the set of profile utility tables admitting a decomposition $u_p(s) = \sum_{q \neq p} u_{pq}(s_p, s_q)$ with per-edge constant-sum constraints $u_{pq}(a, b) + u_{qp}(b, a) = c_{pq}$ for edge constants c_{pq} (MacQueen & Wright, 2023); see also Cai et al. (2016) for the constant-sum polymatrix caliber. V is a ***linear subspace***: the decomposition is componentwise linear and the constraints are homogeneous linear (a sum or scalar multiple of two decompositions is again a decomposition, with the edge constants adding or scaling accordingly). By Definition 4, $\delta_\infty(u) = \min_{v \in V} \|u - v\|_\infty$ is the sup-distance to V in the LP profile-table semantics.

Distances to constant-sum. Define

$$\begin{aligned}\zeta_{\mathrm{prof}} &:= \min_{c\in\mathbb{R}} \max_{s} |U(s)-c| = \tfrac{1}{2}\big(\max_{s} U(s) - \min_{s} U(s)\big),\\ \zeta_{\mathrm{term}} &:= \min_{c} \max_{z\in Z} |U(z)-c|, \qquad \zeta_S := \min_{c} \max_{z\in Z_S} |U(z)-c|.\end{aligned} \tag{24}$$

Clearly $\zeta_{\mathrm{prof}} \le \zeta_{\mathrm{term}}$ (each $U(s)$ is a convex combination of the $U(z)$) and $\zeta_S \le \zeta_{\mathrm{term}}$.

$\rho_S := \sum_{e\in E} c_e \prod_p \pi_p^{\sigma^0}(e) \in [0,1]$. Since σ^1 changes strategy only inside S, the reach probability of Z_S is the same ρ_S under σ^1.

Legality of the update is as in Appendix B.1: $\sigma_p^1 = \sigma_p^0$ for all $p \neq i$ and σ_i^1 differs only at i's information sets inside S; all certificate quantities are defined post hoc for the delivered σ^1. Assumption 1 (P1, P2) is assumed; no exactness assumption is made (δ_∞ is arbitrary).

C.2 Proof of the Transfer Lemma (Lemma 2)

We prove Lemma 2: with $\delta := \delta_\infty(G)$,

(a) *(profile level)* $\zeta_{\mathrm{prof}} \le n \cdot \delta$;

(b) *(update level)* for any legal update σ^1,

$$\begin{aligned}\Big|\sum_p \mathrm{gain}_p\Big| = \Big|\sum_p \big[u_p(\sigma^1) - u_p(\sigma^0)\big]\Big| \;&\le\; \Theta := 2\cdot\min\big(\zeta_{\mathrm{prof}},\ \rho_S\cdot\zeta_S\big)\\ &\le 2\,\zeta_{\mathrm{prof}} \;\le\; 2n\,\delta.\end{aligned} \tag{25}$$

Proof of (a). Take an optimal decomposition $\{u_{pq}\},\{c_{pq}\}$ attaining δ. For each pure s set $e_p(s) := u_p(s) - \sum_{q\neq p} u_{pq}(s_p,s_q)$, so $|e_p(s)| \le \delta$ for every p. Summing over players,

$$U(s) = \sum_p \sum_{q\neq p} u_{pq}(s_p,s_q) + \sum_p e_p(s) = \sum_{p<q} c_{pq} + \sum_p e_p(s), \tag{26}$$

where the pairwise terms are eliminated by pairing each edge and applying the constant-sum constraint. Hence $|U(s)-c^*| \le n\delta$ with $c^* := \sum_{p<q} c_{pq}$, and taking the minimum over centers, $\zeta_{\mathrm{prof}} \le n\delta$. □

Proof of (b), first leg (ζ_{prof}). Under perfect recall, behavioral and mixed strategies are realization-equivalent (Kuhn, 1953) (against every opposing profile they reach each terminal history with the same probability), so the expected utility of any behavioral profile σ is a mixture of pure profiles under a product distribution: $\mathbb{E}_\sigma[U] = \sum_s P_\sigma(s)\,U(s) \in [c-\zeta_{\mathrm{prof}},\ c+\zeta_{\mathrm{prof}}]$, where c is the optimal center attaining ζ_{prof}. Both σ^1 and σ^0 lie in this interval, so their difference is at most $2\,\zeta_{\mathrm{prof}}$. □

Proof of (b), second leg ($\rho_S\cdot\zeta_S$). Let $P^x(z)$ be the terminal distribution under σ^x ($x\in\{0,1\}$, $\sigma^0=\sigma^0$, $\sigma^1=\sigma^1$). Since σ^1 changes strategy only inside S, we have $P^1 = P^0$ pointwise on Z_{out}, and $P^1(Z_S) = P^0(Z_S) = \rho_S$ (trunk and chance are unchanged). Therefore

$$\sum_p \mathrm{gain}_p = \sum_{z\in Z_S} \big[P^1(z) - P^0(z)\big]\,U(z) = \sum_{z\in Z_S} \big[P^1 - P^0\big](z)\,\big[U(z)-c\big], \tag{27}$$

where the second equality holds for *any* constant c because the signed measure P^1-P^0 has total mass 0 on Z_S. Choosing c as the optimal center attaining ζ_S, and writing $\|P^1-P^0\|_1(Z_S) := \sum_{z\in Z_S}|P^1(z)-P^0(z)|$,

$$\Big|\sum_p \mathrm{gain}_p\Big| \le \big\|P^1-P^0\big\|_1(Z_S)\cdot\zeta_S \le 2\,\rho_S\cdot\zeta_S. \qquad \square$$

Explicit accounting for $c_1 = 2n$. The chain from δ to the structural tax is: the factor n from part (a) (the n players' decomposition errors add, while the pairwise terms are eliminated by the constant-sum constraint), times the factor 2 from part (b) (two evaluations, one under σ^1 and one under σ^0). Hence $c_1 = 2n$.

Remark 3 (The tax can be far below $n\delta$). The pairpot family is pointwise zero-sum on terminals, so $\zeta_{\text{term}} = 0$ and $\Theta = 0$ for the whole family, even though $\delta_\infty = \lambda/16 > 0$ (the coupling term is a pure transfer payment that preserves the total); games with $\delta_\infty > 0$ can thus carry zero structural tax. The converse is impossible: $\zeta_{\text{prof}} \le n\delta$ always (part (a)). On the adversarial random family of Section 7, $2\,\zeta_{\text{prof}}$ is often smaller than $2n\delta$ (for instance 0.058 versus 0.064 on one instance with $n = 3$) and $\rho_S\zeta_S$ and ζ_{prof} alternate as the smaller leg. The numerical regression check (Appendix F) found instances of the adversarial random family (Appendix I) on which part (a) is attained exactly ($\zeta_{\text{prof}}/(n\delta) = 1.000000$ at $n = 3$).

C.3 Proof of the Main Theorem (Theorem 1 and Corollary 5)

We prove Theorem 1: under Assumption 1, for any blueprint σ^0 and any legal update σ^1, with $\hat{m}$, M_q, A_i^S, ε_i defined for σ^1 as in Appendix B.1 and Θ as in Eq. (25),

$$\text{(i)}\quad A_i(\sigma^1) - A_i(\sigma^0) = \varepsilon_i - A_i^S(\sigma^0) \le \varepsilon_i \qquad (\varepsilon_i \ge 0 \text{ always}); \tag{28}$$

$$\text{(ii)}\quad A_j(\sigma^1) - A_j(\sigma^0) \le M_j + \Big[A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{k\in O\setminus\{j\}} M_k\Big] + \Theta \qquad \forall j \ne i; \tag{29}$$

$$\text{(iii)}\quad \mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \le \max_p\big[A_p(\sigma^1) - A_p(\sigma^0)\big] \tag{30}$$

$$\le \max\Big(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{q\in O} M_q + \Theta\Big). \tag{31}$$

Proof. Steps S0–S6 of the proof in Appendix B.2 are used *unchanged*. By the dependency accounting stated there, S0 and S1 use only Assumption 1 and perfect recall; S2 and S3 are definitional identities; S4 (opponent BR-value change at most M_j) and S5 (opponent on-path gain at most M_k) use only the entry decomposition, the blueprint counterfactual values, and the game-side best responses that define $\hat{m}$ (Definition 1), together with blueprint feasibility; the gadget is never executed in the proof, which therefore applies to any legal update σ^1; S6 is the definition of A_i^S and ε_i. None of these steps involves δ. The remaining steps are replaced as follows.

Step S7′ (structural tax; replaces S7). By Lemma 2(b), $\big|\sum_p \mathrm{gain}_p\big| \le \Theta$. (When $\delta = 0$, part (a) gives $\zeta_{\text{prof}} = 0$, so $\Theta = 0$ and $\sum_p \mathrm{gain}_p = 0$, recovering the original S7.)

Step S8′ (pricing the victim's on-path loss). For $j \in O$,

$$\mathrm{drop}_j = -\mathrm{gain}_j = \sum_{p\ne j}\mathrm{gain}_p - \sum_p \mathrm{gain}_p \le \big(A_i^S(\sigma^0) - \varepsilon_i\big) + \sum_{k\in O\setminus\{j\}} M_k + \Theta, \tag{32}$$

by S5, S6, and S7′.

Step S8″ (exact form; proves Corollary 1). The first line of Eq. (32) is an identity, and $-\sum_p \mathrm{gain}_p = \sum_p u_p(\sigma^0) - \sum_p u_p(\sigma^1) = \Lambda$. Bounding only the first two groups by S5 and S6 gives $\mathrm{drop}_j \le (A_i^S(\sigma^0) - \varepsilon_i) + \sum_{k\in O\setminus\{j\}} M_k + \Lambda$, and S9′–S10′ below go through verbatim with Λ in place of Θ. S7′ is exactly the statement $|\Lambda| \le \Theta$, so the Λ-form implies the Θ-form pointwise. Because Λ may be negative, it enters S10′ inside the maximum and cannot be moved outside it; the sequential form is discussed after Corollary 6.

Step S9′. $A_j(\sigma^1) - A_j(\sigma^0) = \big[\mathrm{BRval}_j(1) - \mathrm{BRval}_j(0)\big] + \mathrm{drop}_j \le M_j + \big(A_i^S(\sigma^0) - \varepsilon_i\big) + \sum_{k\ne j} M_k + \Theta$, by S4 and S8′. This is Eq. (29).

Step S10′. Claim (i) is exactly S6 ($\sigma^1_{-i} = \sigma^0_{-i}$ leaves the maximization term of A_i unchanged). A difference of maxima is at most the maximum of differences, so combining (i) and (ii),

$$\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \le \max\Big(\varepsilon_i - A_i^S(\sigma^0),\ \max_j\big[M_j + A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_{k\ne j} M_k + \Theta\big]\Big)$$

$$\le \max\Big(\varepsilon_i,\ A_i^S(\sigma^0) - \varepsilon_i + \textstyle\sum_q M_q + \Theta\Big). \qquad \square$$

The corollary form. By Lemma 2, $\Theta \le 2n \cdot \delta_\infty$, so Eq. (31) weakens to Corollary 5:

$$\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0) + 2n \cdot \delta_\infty + \sum_{q \in O} M_q + \varepsilon_i, \tag{33}$$

with $c_1 = 2n$ (equal to 6 at $n = 3$), $c_2 = 1$ (weighted-sum caliber, summed over all opponents, no inverse reach), and $D = 1$ (only the resolver's solve quality enters; under an exact solve in a Terminate-compatible environment whose root measure induces the resolver's game-side beliefs at each of its entry sets, $\varepsilon_i \le \varepsilon_{\text{solve}}$, inheriting the argument of Appendix B.1). A weaker form follows from $A_i^S(\sigma^0) \le \mathrm{Vul}_G(\sigma^0)$: $\mathrm{Vul}_G(\sigma^1) \le 2\,\mathrm{Vul}_G(\sigma^0) + c_1\delta_\infty + c_2 \sum_q M_q + \varepsilon_i$. If in addition all margins are nonnegative and the solve is exact, then $\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0) + 2n \cdot \delta_\infty$.

A blueprint-specific a priori bound on the leak. Let $U := \sum_p u_p$ and $L_i := U(\sigma^0) - \min_{\tau_i \,:\, \tau_i = \sigma_i^0 \text{ outside } S} U(\tau_i, \sigma^0_{-i})$. The blueprint is feasible, so $L_i \ge 0$; every update $\sigma^1 = (\sigma_i^1, \sigma^0_{-i})$ with $\sigma_i^1 = \sigma_i^0$ outside S is feasible, so $\Lambda = U(\sigma^0) - U(\sigma^1) \le L_i$; and L_i is itself the leak of one such update, so $L_i \le \Theta$ by Lemma 2. Hence $\Lambda \le L_i \le \Theta$. Computing L_i is one best response of player i confined to S against σ^0_{-i} with payoff $-U$, and needs no knowledge of σ^1. Θ is a structural upper envelope for L_i: its coefficient is attained on some games (Example 3), but it can stay loose even after maximizing L_i over blueprints and resolvers (with $u_p = a_p$ for independent binary actions and S the whole game, $\Theta = n$ while every unilateral update moves the total by at most one), and it depends on the blueprint only through ρ_S.

The direct per-player form. Writing $g_p := u_p(\sigma^1) - u_p(\sigma^0)$, claim (i) reads $A_i(\sigma^1) - A_i(\sigma^0) = -g_i$, and S4 gives $A_j(\sigma^1) - A_j(\sigma^0) \le M_j - g_j$ for $j \ne i$; hence $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0) \le \max_p[A_p(\sigma^1) - A_p(\sigma^0)] \le B_{\text{dir}}$, which is Eq. (1) (Corollary 2). Since $A_i^S - \varepsilon_i + \sum_q M_q + \Lambda = \sum_{j \ne i}(M_j - g_j)$ with every summand nonnegative by S5, and $-g_i = \varepsilon_i - A_i^S \le \varepsilon_i$, B_{dir} is pointwise at most the Λ form of Corollary 1; it needs the BR-certified margins and the n expected utilities of σ^0 and σ^1, and no activation term or structural quantity. The instantiations of Theorem 1 do not exhaust the useful post-hoc certificates (Appendix I.5).

The a priori instantiation of Theorem 1(iii) through Lemma 2 is stated here for reference.

Corollary 5 ($2n \cdot \delta_\infty$ form)**.** *Since $\Theta \le 2\zeta_{\text{prof}} \le 2n \cdot \delta_\infty$ (Lemma 2), Theorem 1(iii) implies $\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0) + A_i^S(\sigma^0) + 2n\,\delta_\infty + \sum_{q \in O} M_q + \varepsilon_i$; that is, constants $(c_1, c_2, D) = (2n, 1, 1)$; for $n = 3$, $c_1 = 6$.*

The baseline-aware form. Write $A_p^0 := A_p(\sigma^0)$. Claim (i) gives $A_i(\sigma^1) = A_i^0 - g_i$ exactly, and S4 gives $A_j(\sigma^1) \le A_j^0 + M_j - g_j$ for $j \ne i$. Taking the maximum over p and subtracting $\mathrm{Vul}_G(\sigma^0) = \max_p A_p^0$ yields the bound of Corollary 3. Since $\max_p(A_p^0 + x_p) - \max_p A_p^0 \le \max_p x_p$ for any reals x_p, applying it with $x_i = -g_i$ and $x_j = M_j - g_j$ gives $B_{\text{cache}} \le B_{\text{dir}}$. B_{cache} bounds $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0)$ itself, not the middle term $\max_p[A_p(\sigma^1) - A_p(\sigma^0)]$ of claim (iii), which every other form of this appendix bounds; the two differ whenever the player whose advantage rises most is not the most exploitable one after the update.

Zero tolerance on constant-sum games. If $\sum_p u_p$ is constant on the game then $\sum_p g_p = 0$. Suppose $B_{\text{dir}} \le 0$. Then $-g_i \le 0$ and $M_j - g_j \le 0$ for every $j \ne i$, so $g_i \ge 0$ and $g_j \ge M_j \ge 0$; with $\sum_p g_p = 0$ this forces $g_p = 0$ for every p and $M_j = 0$ for every $j \ne i$. Conversely these equalities give $B_{\text{dir}} = 0$. Hence on a constant-sum game $B_{\text{dir}} \le 0$ holds exactly when the update changes no expected utility and leaves every opponent's certified deficit at zero, and every form of Theorem 1 inherits the restriction, since each is pointwise at least B_{dir}. B_{cache} does not: an update with $g_i > 0$, $g_j = -g_i$ for one opponent j, $g_k = 0$ for the others, every $M_q = 0$, and $A_j^0 + g_i \le \max_p A_p^0$ has $B_{\text{cache}} \le 0$ while $B_{\text{dir}} = g_i > 0$.

Degeneration at $\delta = 0$. If $\delta_\infty = 0$ then $\zeta_{\text{prof}} \le n \cdot 0 = 0$, hence $\Theta = 0$, and claims (i)(ii)(iii) reduce term by term to Lemma 1(i)(ii)(iii), verbatim. The numerical regression check (Appendix F) confirms this degeneration numerically: on all exact-case configurations, the bound of Theorem 1 equals the recorded exact-case bound pointwise.

Tightness status. The coefficient 1 on A_i^S is optimal (Example 1, equality); the sum-over-opponents structure cannot be weakened and a single-margin per-pair certificate is unobtainable (Example 2, equality); $c_2 = 1$ is pinned by the same two equality constructions. The coefficient 1 on Θ is optimal (Example 3, leak: $\Delta\mathrm{Vul} = \Theta = \Lambda = 1$ with every other term zero), and the ε_i branch of the maximum is attained (Example 3, self-harm); see Appendix B.5. The leak construction has $n = 2$ and $\delta_\infty = \frac{1}{4}$, so $2n\,\delta_\infty = \Theta$ there and Corollary 5 is attained at $n = 2$. For $n \geq 3$ the n factor inside Lemma 2(a) is attained exactly on adversarial random instances (Remark 3), but the end-to-end response of $\Delta\mathrm{Vul}$ to δ is not proved tight in n; note that the single-instance tax can be 0 while $\delta > 0$ (the pairpot family has $\Theta \equiv 0$), so a sup-δ lower bound needs a different construction from the leak.

C.4 Why δ Enters Only Through the Total Utility

An alternative argument would transfer the analysis to the nearest CSP game: establish $|\mathrm{Vul}_G(\sigma) - \mathrm{Vul}_{\tilde{G}}(\sigma)| \leq 2\delta$ for the projection $\tilde{G}$ of G onto V, run the exact-case argument inside $\tilde{G}$, and transfer back. Two obstacles block it.

First, $\tilde{G}$ is only a profile utility table, not an extensive-form game. Steps S1–S6 of Appendix B.2 (the trunk–subgame decomposition and the gadget certificates) require tree structure and are undefined on $\tilde{G}$.

Second, even if $\tilde{G}$ is realized as an EFG on the same tree with modified terminal utilities, the entry-level certificate quantities (the margins $\hat{m}$ and the values CFV) are defined on G, not on $\tilde{G}$, and the profile-table distance δ does not control their transport (the activation A_i^S, a game-side quantity, moves by at most 2δ). The entry-level quantities are conditional on entering S, so their discrepancy between G and $\tilde{G}$ can be of order δ/ρ_S: with chance entering S with probability r and the resolver alone acting inside ($a \mapsto (0,0)$, $b \mapsto (0,-1)$), one has $\delta = r/4$, while a realization of the nearest constant-sum table on the same tree shifts CFV_j by $1/4$ and $\hat{m}_j$ by $1/2$ for every r. Without restrictions on how the proxy table is realized on terminals, the discrepancy is not controlled even by that ratio: shifting j's payoff by H on the terminals of S and by $-rH/(1-r)$ outside leaves every profile utility, hence δ and ρ_S, unchanged while moving CFV_j by H. The profile-table distance therefore does not transport the entry-level certificate quantities uniformly in the blueprint; a transfer of the aggregate certificate would need a separate argument.

The proof above lets δ enter at exactly one step (S7′), through the constant-sum defect. This is also the mechanism behind the gap between the two taxes, Θ and $2n\delta_\infty$, noted in Section 5: the structural tax is carried by the distance to constant-sum, which can be far smaller than $n \cdot \delta_\infty$.

D Linearity and the Explicit Constant

This appendix proves Lemma 3 (the exact linearity of δ_∞ along affine perturbation families anchored in the constant-sum polymatrix (CSP) class) and Proposition 1 (the explicit constant $K = 1/4$, defined below, for the pairpot family, by hand case analysis, with an equality construction). Notation for the CSP subspace V is as in Appendix C.1: V is the set of profile utility tables admitting a decomposition $u_p(s) = \sum_{q \neq p} u_{pq}(s_p, s_q)$ with per-edge constant-sum constraints $u_{pq}(a,b) + u_{qp}(b,a) = c_{pq}$, and $\delta_\infty(u) = \mathrm{dist}_\infty(u, V) = \min_{v \in V} \|u - v\|_\infty$ in the LP profile-table semantics of Definition 4.

D.1 Proof of the Linearity Lemma (Lemma 3)

The pairpot game has three players. Chance deals each player p a private card t_p, i.i.d. uniform on $\{0,1\}$, and a public signal X uniform on $\{A,B\}$; the effective card is $h_p := t_p \oplus [X = B]$ (exclusive or with the indicator of $X = B$), so $X = B$ reverses the card order. Each player sees only (t_p, X) and either checks or bets; the three actions are mutually unobserved, so a pure strategy is a map $\{0,1\} \times \{A,B\} \to \{\mathrm{check}, \mathrm{bet}\}$ and there are $16^3 = 4096$ pure profiles. The base payoff u^0 is one independent side pot per pair $\{p,q\}$: if both bet, the higher h receives 1 from the other (a tie pays nothing); if exactly one bets, the bettor receives $1/2$ from the checker; if both check, nothing is paid. The coupling C is a shared pot among the set $\mathcal{B}$ of players who bet: if $|\mathcal{B}| \geq 2$ and one member holds the unique highest h within $\mathcal{B}$, that member collects $1/2$ from every other member of $\mathcal{B}$; a tie for the highest card pays nothing.

The perturbation family is $u_\lambda = u^0 + \lambda \cdot C$, where u^0 is the base table of the pairpot family (independent per-pair side pots, which is exactly CSP: $\delta_\infty(u^0) = 0$, verified by exact LP) and C is the shared-pot coupling tensor (the difference between the $\lambda = 1$ and $\lambda = 0$ profile utility tables; the family is affine in λ by construction, verified numerically to residual 10^{-12} in Appendix F).

Proof of Lemma 3. V is a linear subspace: the decomposition map is componentwise linear and the per-edge constant-sum constraints are homogeneous linear, so a sum or a scalar multiple of two decomposable tables is again decomposable, with the edge constants adding or scaling accordingly. Since $u^0 \in V$, for any $\lambda \geq 0$,

$$\delta_\infty(u^0 + \lambda C) = \mathrm{dist}_\infty(u^0 + \lambda C,\, V) = \mathrm{dist}_\infty(\lambda C,\, V) = \lambda \cdot \mathrm{dist}_\infty(C,\, V) = \lambda \cdot \delta_\infty(C). \quad (34)$$

The second equality is translation invariance of the distance along the subspace ($u^0 \in V$); the third is positive homogeneity of the sup distance to a linear subspace. □

Consequently, the measured pairpot behavior, a slope of exactly $1/16$ with zero intercept, matching pointwise across all recorded λ values (Section 6), follows from the theorem: $\delta_\infty(u_\lambda) = \lambda \cdot \delta_\infty(C)$ with $\delta_\infty(C) = 1/16$ computed by exact LP. Lemma 3 applies to any affine family with base point in V; only the numerical constants below are specific to the pairpot family.

D.2 Proof of the Explicit Constant (Proposition 1)

Proof of Proposition 1. Define $K := \max_{p,s} |C_p(s)|$. Since $0 \in V$, the zero-decomposition bound gives $\delta_\infty(C) = \mathrm{dist}_\infty(C, V) \leq \|C - 0\|_\infty = K$, hence by Lemma 3,

$$\delta_\infty(u_\lambda) = \lambda \cdot \delta_\infty(C) \leq \lambda \cdot K. \quad (35)$$

It remains to show $K = 1/4$.

$C_p(s)$ is player p's expected coupling payoff at $\lambda = 1$ under the pure profile s: among the set of bettors $\mathcal{B}$ (coupling active only when $|\mathcal{B}| \geq 2$), the holder of the unique highest X-adjusted card collects $1/2$ from each other member of $\mathcal{B}$. Fix the public signal X and take conditional expectations. A pure strategy is a function of (t_p, X), so conditionally on X the betting action of each player is a function of $h_p := t_p \oplus [X = B]$, and (h_p, h_q, h_r) are i.i.d. uniform on $\{0, 1\}$.

Receipt side. Player p collects if and only if $h_p = 1$, $p \in \mathcal{B}$, and every other bettor has $h = 0$: since h takes only the values $0/1$, a player with $h_p = 0$ can never be the unique highest within any bettor set of size at least 2. The amount collected is $(|\mathcal{B}| - 1)/2 \leq 1$. Bounding case by case over (h_q, h_r): $(0, 0)$ gives at most 1 (both opponents may bet); $(0, 1)$ and $(1, 0)$ give at most $1/2$ each (an opponent holding 1 destroys p's uniqueness the moment it bets, so p collects from at most the one opponent holding 0); $(1, 1)$ gives 0. Hence

$$\mathbb{E}[\text{receipt} \mid X] \leq \Pr(h_p = 1) \cdot \tfrac{1}{4}\left(1 + \tfrac{1}{2} + \tfrac{1}{2} + 0\right) = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4}. \quad (36)$$

Payment side. Player p pays if and only if $h_p = 0$, $p \in \mathcal{B}$, and some opponent with $h = 1$ is the unique highest within $\mathcal{B}$; the amount is $1/2$. Case by case over (h_q, h_r): $(0, 0)$ gives 0; each of $(0, 1)$, $(1, 0)$, $(1, 1)$ gives at most $1/2$. Hence

$$\mathbb{E}[\text{payment} \mid X] \leq \tfrac{1}{2} \cdot \tfrac{3}{4} \cdot \tfrac{1}{2} = \tfrac{3}{16}. \quad (37)$$

Averaging over X leaves both bounds unchanged, so $|C_p(s)| \leq \max(1/4,\, 3/16) = 1/4$, i.e. $K \leq 1/4$.

Equality construction. Let p bet if and only if $h_p = 1$, and let q and r bet if and only if their $h = 0$. Then p's receipt is

$$\tfrac{1}{8} \cdot 1 + \tfrac{1}{8} \cdot \tfrac{1}{2} + \tfrac{1}{8} \cdot \tfrac{1}{2} = \tfrac{1}{4},$$

(the three cases $h = (1, 0, 0), (1, 0, 1), (1, 1, 0)$, with bettor sets of sizes $3, 2, 2$ respectively; in the case $h = (1, 1, 1)$ only p bets and the coupling is inactive), and p's payment is 0. Hence $K = 1/4$. □

Exhaustive enumeration over all 4096 pure profiles and all 3 players reproduces $\max_{p,s} |C_p(s)| = 0.25$ exactly (Appendix F).

Remark 4 (The factor-4 gap). The exact LP value is $\delta_\infty(C) = 1/16 = 0.0625$, which is also the measured slope of the pairpot family (Section 6); the analytic bound $K = 1/4$ therefore covers the truth with a gap of a factor 4. The bound $K = 1/4$ is a zero-decomposition bound: it does not use the freedom of choosing a nonzero element of V, whereas the LP exploits the optimal decomposition to compress the constant from $1/4$ to $1/16$. We do not have an analytic characterization of that optimal decomposition; the factor-4 gap is left open.

E Construction Details for the Tightness and Necessity Results

This appendix documents the constructions behind the negative results of Section 6 (Propositions 3 to 5) and the poker-class δ_∞ lower bounds of Section 6 (Proposition 2), at the level of detail needed to reproduce them. Two single-core checks (runtimes 1 min 57 s and under 1 s) re-derive every number quoted here (tolerance 10^{-9}; pilot figures at their recorded two-decimal precision). Constant-sum polymatrix (CSP) is used in the sense fixed in Section 3.

E.1 Margin-term necessity: the gid78 and gid90 instances (Proposition 3)

Both instances come from the adversarial screen of Section 7: random small extensive-form games (depth ≤ 3, 2–3 actions per node) put through the pipeline blueprint $\to$ corrupt $\to$ resolve $\to$ exact best-response evaluation, with screening objective $\Delta\mathrm{Vul}_{\max} - 0.03\,\delta_\infty - 0.03\,(-\mathrm{margin})_+$ (the 0.03 constants are the pilot's benign-family fit).

gid78 (scan original). Sampled δ_∞ lower bound 0.2188, realized $\Delta\mathrm{Vul} = 0.2164$, minimum margin -1.0559. The scan-maximal objective over all 433 points is $+0.178$, attained here. Under a deep re-solve ($7.5\times$ blueprint and $5\times$ resolve iterations), the screening objective persists and slightly grows ($\Delta\mathrm{Vul}$ $0.216 \to 0.224$; the top objectives move $0.216 \to 0.235$, $0.208 \to 0.222$, $0.225 \to 0.255$), and companion originals behave the same way (gid58: $0.148 \to 0.193$; gid90: $0.147 \to 0.151$; gid48: $0.194 \to 0.152$), with margins stably and strongly negative throughout (-0.27 to -1.07). Because gid78's utility tensor exceeds the exact-enumeration cap, its δ_∞ certificate is a restricted-support *lower bound* (same support and value as the screening figure). A lower bound on δ_∞ shrinks the right-hand side of any bound with a positive coefficient on δ_∞, so the objective is a screening score, not a certified violation of the $2n\delta_\infty$ corollary or of the fitted c=0.03 law.

What the instance does certify is a different claim: the structural theorem with its margin term deleted fails, because the sharp tax is exact. The sharp tax needs only the range of total utility over the same $1 \times 64 \times 256$ profile table and is exact: $\zeta_{\mathrm{prof}} = 0.817$, $\rho_S = 0.0057$, $\zeta_S = 1.458$, hence $\Theta = 2\rho_S\zeta_S = 0.0166$; the realized leak is $\Lambda = -0.0025$ (the update slightly raises total utility). With $A_i^S = 0.00005$ and $\varepsilon_i = 0.0019$, the bound of Theorem 1 with $\sum_q M_q$ deleted is $\max(\varepsilon_i, A_i^S - \varepsilon_i + \Theta) = 0.015$, and 0.002 in the Λ form of Corollary 1, both far below $\Delta\mathrm{Vul} = 0.216$; the full bound $A_i^S - \varepsilon_i + \sum_q M_q + \Lambda = 0.248$ covers it.

gid90 (hill-climbed variant). Starting from the scan point, 25 Gaussian leaf-utility mutation steps (9 accepted) produce the variant gid90: $\Delta\mathrm{Vul} = 0.249829$, minimum margin -0.578777. The objective 0.224797 was evaluated at the climb-time restricted-support lower bound $\delta_\infty \geq 0.255645$; the certificate δ_∞ is computed by *exact* full profile enumeration, $\delta_\infty = 0.303275$, matching the deep-dive value within 10^{-7}, and at this exact value the objective is 0.223367, still positive.

In both instances the margin certificate fires with strongly negative $\hat{m}$, matched by a real vulnerability increase of the same order: setting $c_2 = 0$ in the bound of Theorem 1 is refuted on gid78, where the sharp tax is measured (see also the fitted-constant evidence in Section 7: $c_2 = 0$ forces $c_1 \geq 2.64$ on the full 162-instance set).

E.2 Margin-only bounds are insufficient: the gid174 instance (Proposition 4)

The complementary construction shows that a margin-only bound is false. Instance gid174 (from the arm in which the resolver's blueprint subgame strategy is reset to a Dirichlet draw before resolving, Appendix I) carries an exact full-enumeration δ_∞ certificate: $\delta_\infty = 0.4292$, the resolve realizes $\Delta\mathrm{Vul} = 0.0383$ above the 0.02 noise floor, and the BR-certified margin deficit aggregates to only $\sum_q M_q = 0.0111$ with $A_i^S = 0.0382$, $\varepsilon_i = 0.0379$, and minimum margin -0.0446. Any margin-

only bound $\Delta\mathrm{Vul} \le c_2 \sum_q M_q + \mathrm{tol}$ needs $c_2 \approx 1.65$, five times the fitted $c_2 = 0.31$ (Table 5). The solve-error term carries most of the increase, so the construction isolates the non-margin channels jointly. Theorem 1 covers the point through ε_i and the δ_∞ term.

E.3 Structural margin mismatch near equilibrium (Proposition 5)

The pilot mismatch experiment resolves from a CFR blueprint with *zero* deviation in a three-player game (3p Kuhn, deck 4, 6000 iterations) and still observes gadget margins of -0.4606 and -0.7448 (reported as $-0.46/-0.74$ in the main text). The run records the blueprint strategy, its per-player full-game best-response gains $0.0012/0.0014/0.0004$ (so $\mathrm{Vul}_G(\sigma^0) = 1.4 \times 10^{-3}$: the blueprint is near-equilibrium, not exact), and the complete margins table. The BR-certified margins of Definition 1 at the same zero deviation are nonpositive at all 16 entry information sets of the two pilot public states; at the two entries where the gadget reports -0.46 and -0.74, the certified margins are -0.093 and -0.0026, and the mass-weighted deficits are 1.8×10^{-4} and 3.5×10^{-4}. The mechanism is a weighting defect in the gadget's root measure $c_e\, r_i(e)$, which omits the other opponents' reach (Appendix J) and can mis-weight opponents' beliefs when $n > 2$ (the two measures coincide only in the special entry structures of Appendix J); the certificate can therefore fire in the false-positive direction even when nothing has changed. A negative margin in multiplayer resolving is accordingly an alarm; it does not by itself establish damage. Appendix E.1 shows the converse regime where the alarm is real.

E.4 Certificate robustness: the gid183 case

Under the *fixed-certificate* caliber (certificates frozen before a utility perturbation, so σ^1 is *not* the resolve of the perturbed game; this is outside the quantifier range in which the certificates of Theorem 1 are defined), exactly one strict-filter candidate appears in the full 162-instance attack campaign: gid183, with $\delta_\infty = 0.063$, $\hat{m} \approx 0$, $\Delta\mathrm{Vul} = 0.077$. Its $2n\delta_\infty = 0.38 > 0.077$ dominates the observed increase, so re-measuring $A_i^S \ge 0$ under the pipeline caliber places the point under the theorem. The practical reading is that certificates must be issued under the same caliber as the deployed pipeline.

E.5 Pairwise showdown: a positive class with $\delta_\infty = 0$

This appendix records the analysis behind the positive class of Section 6. In the pairwise-showdown family, three players are dealt distinct private cards from a shared deck and act once; at showdown (when the cards are compared) each *pair* of players settles an independent side pot (both bet: ± 2; one bets: ± 1; both check: 0). Exact LP over all 4096 pure profiles gives $\delta_\infty = 0$ exactly, and the value persists under correlated dealing (without replacement). The structural reason is that each edge utility u_{ij} enters the profile table only through the (i, j) marginal of the deal: any third-party card is integrated out in expectation, so the pairwise witness $\{u_{pq}\}$ may absorb it through the chance expectation without depending on any third-party *strategy*. This is exactly the freedom allowed by Definition 4: third-party randomness is in scope, third-party strategies are not. The same computation isolates what breaks CSP structure: a shared pot whose settlement depends on the joint ranking *and* the actions is a coupling passing through (private type, i.e. card, action) pairs, so a player's utility depends on an opponent's strategy-dependent behaviour. An early variant whose shared pot depended only on the joint ranking and not on any action has δ_∞ identically zero (exact LP measures 0 even at $\lambda = 0.5$), confirming that the obstruction is action-passing coupling, not correlated dealing. The resulting checkable design criterion: if every settlement is pairwise and third parties enter a pair's payoff only through chance, then $\delta_\infty = 0$.

E.6 Poker-class δ_∞ lower bounds: LP dual certificates (Proposition 2)

The four lower bounds quoted in Section 6 are established by explicit dual solutions of the δ_∞ LP; each certificate records the utility tensor, the LP primal–dual pair, the claimed value, and the list of verification conditions:

Game	Claimed bound	δ_∞ caliber	Relative scale
3-player Kuhn (deck 4)	$\delta_\infty \geq 1.0$	restricted-support lower bound	≥ 0.250
blind Goofspiel K=3	$\delta_\infty = 2.5$	exact (216 profiles)	0.500
blind Goofspiel K=4	$\delta_\infty = 4.625$	exact (13,824 profiles)	0.514
observed Goofspiel K=3	$\delta_\infty \geq 3.0$	restricted-support lower bound	≥ 0.500

The Kuhn certificate is small enough to state in full. Its dual places weight $\frac{1}{4}$ on four pure profiles that fix player 3's strategy and vary players 1 and 2 over two strategies each, with the signs of a mixed second difference; player 3's payoffs at the four profiles are 2, -1.0833, -0.6667, and 0.25 (in the order $(s_1, s_2) = (a,a), (a,b), (b,a), (b,b)$), so the mixed second difference $u_3(a,a) - u_3(a,b) - u_3(b,a) + u_3(b,b)$ equals 4.0. For every $v \in V$ the pairwise tables give $v_3(s) = v_{31}(s_3, s_1) + v_{32}(s_3, s_2)$, whose mixed second difference in (s_1, s_2) is zero, so $4.0 \leq 4\,\|u - v\|_\infty$ and $\delta_\infty \geq 1.0$ for the full game, beyond the restricted support on which the LP was solved. The certificate's verification conditions (dual weights nonnegative and summing to one, aggregated pair tables symmetric and summing to zero, dual objective equal to the claimed value) reduce here to this one identity.

Goofspiel conventions: the unique highest bid wins the prize, ties void it, and utilities are zero-sum-ized as $u_i = \mathrm{score}_i - \frac{1}{2}\sum_{j \neq i} \mathrm{score}_j$; "blind" means opponents' bids are unobserved, "observed" means bids are revealed after each round, and the prize order is fixed and known. Verification of each certificate requires only arithmetic, never re-solving the LP: nonnegativity of the dual variables, sum-to-one normalization, aggregate symmetry, aggregate zero-sum, and dual-objective match against the claimed value; the Kuhn certificate has dual support of size 4. Each utility tensor is rebuilt from the game definition before the check. For restricted-support entries the certified quantity is a lower bound by LP duality.

F Numerical Regression and Property Tests

Every proof in this paper is accompanied by a numerical check. These suites verify the stated inequalities on finite configurations; the proofs are in Appendices B and C. Headline result: **97 checks** for the exact-case lemma (Lemma 1) and **6,157 checks** for the transfer lemma and main theorem (Lemma 2 and Theorem 1), with **0 violations** in both. All runs use seed 42 and a single core; best responses, δ_∞ LPs, and utility tensors are computed *exactly*, and CFR is used only to generate the input profiles σ^0, σ^1, so the checked inequalities are exact up to a 10^{-9} floating-point tolerance.

Hard-gate semantics. A *hard gate* is an inequality or identity that must hold on every checked configuration; any excess beyond the 10^{-9} tolerance counts as a violation and fails the suite.

F.1 The 97-check suite for the exact case (Lemma 1)

Configurations: the two $\delta_\infty = 0$ families of Section 6 (pairwise-showdown, deck 4, public state "o", resolver $\in \{0,1,2\}$; pairpot at $\lambda = 0$, public signal A with resolver 0 and public signal B with resolver $\in \{0,1,2\}$), giving 7 CFR configurations (blueprint CFR 500 iterations, resolve 1600), plus the two hand-built equality constructions Examples 1 and 2. Gates, mirroring the proof steps:

- **G1**: exact-LP recomputation $\delta_\infty = 0$ for both families (4,096 pure profiles each) and pointwise terminal zero-sum $\sum_p u_p(z) = 0$ (the premise of the single step where $\delta_\infty = 0$ is used).
- **G2** (each opponent j): best-response value change $\leq M_j$.
- **G3** (each opponent k): on-path gain (change in k's expected payoff under the profile actually played) $\leq M_k$.
- **G4**: $\varepsilon_i \geq 0$.
- **G5** (each j): $\Delta A_j \leq M_j + (A_i^S - \varepsilon_i) + \sum_{k \neq j} M_k$, i.e. Lemma 1(ii).
- **G6**: $\Delta\mathrm{Vul} \leq \max(\varepsilon_i,\ A_i^S - \varepsilon_i + \sum_q M_q)$, i.e. Lemma 1(iii).
- **G7** (each j): the accounting identity $\mathrm{drop}_j = \sum_{p \neq j} \mathrm{gain}_p$.

- **G8**: the two constructions reproduce their claimed values exactly (Example 1: $\hat{m}_j = +1$, $\Delta\mathrm{Vul} = 1 = A_i^S$, equality; Example 2: $\hat{m}_j = 0$, $\hat{m}_k = -1$, pairwise vulnerability jump $= 1 > (-\hat{m}_j)_+$, $\Delta\mathrm{Vul} = 1 = \sum_q M_q$, equality).

Result (runtime 170 s): 97 checks, 0 violations. On the 7 CFR configurations the realized $\Delta\mathrm{Vul} \in [-6.0 \times 10^{-5}, 0]$ against bound values in $[3.4 \times 10^{-4}, 6.1 \times 10^{-2}]$. The suite additionally reports, as a diagnostic with no gate attached, the implementation-margin versus BR-margin comparison quantified in Appendix J (D-b).

F.2 The 6,157-check suite for the general case (Lemma 2 and Theorem 1)

The suite is layered into five parts, from theory constants to full experimental replay:

- **Part A: perturbation law and explicit constant** (Lemma 3 and Proposition 1; 2.3 s). Gates: the pairpot family is affine in λ (residual $\leq 10^{-12}$); $\delta_\infty(u^0) = 0$; enumerated $\max_{p,\rho} |C_p(\rho)| = 0.25$ exactly; $\delta_\infty(C) = 0.0625$ by exact LP; exact linearity $|\delta_\infty(\lambda) - \lambda\,\delta_\infty(C)| \leq 10^{-9}$ on all 5 recorded λ points; coverage $K_{\mathrm{theory}} \geq K_{\mathrm{measured}}$ with gap $4.0\times \leq 10\times$.
- **Part B: $\delta_\infty = 0$ degeneration hard gate** (99 s). All 7 L3 CFR configurations and both examples are re-run under the Theorem 1 machinery: $\Theta = 0$ on every configuration, and the Theorem 1 bound equals the recorded Lemma 1 bound *pointwise* within 10^{-9}, as do the vulnerability differences. The general proof degenerates to the exact one term by term.
- **Part C: fresh $\delta_\infty > 0$ resolves** (141 s). 16 new configurations (pairpot $\lambda \in \{0.1, 0.25, 0.5, 1\}$ crossed with 4 resolve configurations; CFR 500/1600). All theorem gates pass, plus the transfer gate $\zeta_{\mathrm{prof}} \leq n\delta_\infty$ and the composite $2n\delta_\infty$ gate; recomputed δ_∞ matches the recorded measurement 4/4. Observed mechanism: ζ_S and Θ are numerically zero throughout (the coupling is a transfer payment preserving the utility sum), an entire family with $\delta_\infty > 0$ but zero structural tax.
- **Part D: attack-campaign points**. All 312 theorem-caliber points from the certified adversarial search (Section 7) are checked against the composite bound $\Delta\mathrm{Vul} \leq \max(\varepsilon_i,\ A_i^S - \varepsilon_i + \sum_q M_q + 2n\delta_\infty)$; all pass, minimum slack 7.4×10^{-5} (gid12, no-corruption arm). The 155 fixed-certificate-caliber points are skipped, as their A_i^S/ε_i were not recorded and the caliber lies outside the theorem's quantifier range (see Appendix E.4).
- **Part E: full screening replay** (1,665 s). All **420** stage-1 screening configurations (210 games $\times$ 2 corruption arms) are re-run by replicating the original random-number stream, with a *fidelity gate*: replayed $\Delta\mathrm{Vul}$, margins, δ_∞, and the exact/lower-bound caliber flags must match the recorded screen values pointwise within 10^{-9}. Result: 420/420 match (the original screen itself had zero replay deviation). On top of the replay, all theorem gates pass (core-bound minimum slack 3.4×10^{-5}, gid7/uniform).
- **Part E (exact-δ_∞ subset)**. For the 148 games (296 points) whose tensors admit exact δ_∞ enumeration, the transfer gate passes everywhere and is *attained*: $\max \zeta_{\mathrm{prof}}/(n\delta_\infty) = 1.000000$ at gid49 and gid41, so the n factor in Lemma 2(a) is instance-level tight; the composite gate passes with minimum slack 7.5×10^{-3} (gid188). On these points ζ_{prof} is the smaller leg on 204 of the 296 points and $\rho_S\zeta_S$ on 92, and $\Theta < 2n\delta_\infty$ on 266/296 (relative slack beyond 10^{-9}); the composite inequality $\Theta \leq 2n\delta_\infty$ holds on all 296 points with 0 violations, and is *attained* ($\Theta = 2n\delta_\infty$) on the remaining 30 points (15 instances $\times$ 2 corruption arms), each with $\zeta_{\mathrm{prof}} = n\delta_\infty$ and $\rho_S\,\zeta_S \geq n\delta_\infty$, so the $\min$ binds on the ζ_{prof} leg there.

Total: 6,157 checks, 0 violations (runtime 1,908 s). Coverage. The 13 mutated hill-climb/deep-dive points are not replayed; the 420 stage-1 base points are. For the 62 games (124 points) whose tensors exceed the enumeration cap, only the Θ-form core gate is checked; the composite and transfer gates are unverified there.

Sample-free measurement. Every theorem-level quantity in the main text is measured without sampling. Each negative result is established on an exact finite construction, eight such constructions in total (Appendix E), computed by full-table LP or enumeration, so the estimators for sampled

evaluation, variance reduction (Burch et al., 2018) and anytime-valid stopping (Li et al., 2026a), do not enter.

G Nested Resolving: a Telescoping Corollary and Supporting Data

Scope of Assumption 1. (P1) asks that no information set straddle the boundary of S. It holds for any subgame cut at a public state, since every player then knows whether play has entered S; it can fail when S is cut below a private signal, because a player who has not observed that signal may own information sets with nodes on both sides of the cut. (P2) asks that the gadget's entry grouping $\iota_q = (\mathrm{priv}_q, \mathrm{pub}^*)$ be q's actual information at entry and that q's own reach be constant on each group. Under perfect recall this holds whenever the entries follow a public signal and q's private history is fully recorded in priv_q; it fails when two entries with the same $(\mathrm{priv}_q, \mathrm{pub}^*)$ differ in an earlier action of q, because q can then reach them with different probabilities and the margin of Definition 1 would average over strategies q can tell apart. Both conditions are properties of the cut, not of the blueprint.

The main theorem bounds a *single* resolve. Practical systems resolve repeatedly as play descends the public tree (the tree of public states). Under Assumption 1, repeated resolving composes by telescoping, because Theorem 1 holds for *any* update that only touches the resolver's strategy inside the current subgame, with all certificate quantities defined a posteriori for the delivered profile.

Corollary 6 (Telescoping bound for sequential resolves)**.** *Let $\sigma^{(0)} = \sigma^0$ and let $\sigma^{(t)}$, $t = 1, \dots, T$, be a sequence of profiles where each $\sigma^{(t)}$ is obtained from $\sigma^{(t-1)}$ by an update that changes only resolver i's strategy inside a public subgame S_t, with Assumption 1 holding for each $(\sigma^{(t-1)}, S_t)$. Let $A_i^S[t]$, $M_q[t]$, $\varepsilon_i[t]$, and $\Theta[t]$ denote the certificate quantities of Theorem 1 evaluated at step t (blueprint role played by $\sigma^{(t-1)}$). Then*

$$\mathrm{Vul}_G(\sigma^{(T)}) \;\le\; \mathrm{Vul}_G(\sigma^0) \;+\; \sum_{t=1}^{T} B_t \;+\; \sum_{t=1}^{T} \Theta[t],$$

$$B_t := \max\Big(\varepsilon_i[t],\; A_i^S[t] - \varepsilon_i[t] + \textstyle\sum_q M_q[t]\Big),$$

and in the $2n\delta_\infty$ caliber of Corollary 5 the tax sum is at most $T \cdot 2n\,\delta_\infty$.

Proof. Apply Theorem 1(iii) at each step with $\sigma^{(t-1)}$ as blueprint and $\sigma^{(t)}$ as resolved profile, and sum the resulting one-step inequalities; the intermediate $\mathrm{Vul}_G(\sigma^{(t)})$ terms cancel. The tax bound $\Theta[t] \le 2n\delta_\infty$ is Lemma 2(b) applied at each step (the game, and hence δ_∞, is fixed across steps). □

In the exact form of Corollary 1 each one-step bound becomes $\max(\varepsilon_i[t],\; A_i^S[t] - \varepsilon_i[t] + \sum_q M_q[t] + \Lambda[t])$ and the sum of these maxima bounds $\mathrm{Vul}_G(\sigma^{(T)}) - \mathrm{Vul}_G(\sigma^0)$. The signed leak cannot be pulled outside the maximum: with one acting player, blueprint $a \mapsto (1, -1)$ and update $a' \mapsto (0, \frac{1}{2})$ give $B_1 = 1$, $\Lambda[1] = -\frac{1}{2}$, yet $\Delta\mathrm{Vul} = 1 > B_1 + \Lambda[1]$. The leaks $\Lambda[t]$ do telescope to $\sum_p u_p(\sigma^0) - \sum_p u_p(\sigma^{(T)})$, but only their positive parts $\Lambda[t]_+$ can be separated out, and those do not telescope.

The corollary is linear in T; it does not by itself exclude that the *certificate quantities* degrade as resolving proceeds (e.g. margins drifting negative after repeated splicing, i.e. inserting the local update into the current profile), which would inflate the B_t. The depth sweep below measures this. The corollary covers CFR-D-style nested resolving in which each step edits only the resolver's strategy inside the current subgame; variants that rewrite the blueprint elsewhere fall outside Assumption 1.

Supporting data. We ran a depth sweep on 14 instances (7 random, 7 near constant-sum polymatrix (CSP), same generator family and seed as the adversarial screen), resolver fixed, with $T = 3$ sequential resolves: the depth-1 public subgame, then a depth-2 subgame below it, then a sibling depth-2 subgame. At each step we recorded B_t under the Lemma 1 caliber and the cumulative vulnerability change. The check

$$\mathrm{cum}_T \;\le\; \sum_{t \le T} B_t + T \cdot c_1 \delta_\infty + \mathrm{tol}, \qquad \mathrm{tol} = 0.02,$$

was evaluated for all 42 (instance, T) pairs and for $c_1 \in \{0, 0.34, 1.0\}$: **0 failures**; the pure-margin caliber $c_1 = 0$ passes everywhere, so the δ_∞ leg was not needed on this family. Single-step violations ($\Delta\mathrm{Vul}_t > B_t + c_1\delta_\infty + \mathrm{tol}$) were also zero. No superlinear accumulation was observed: per-step ΔVul has a median around 10^{-5} with no systematic growth in t, and the one large single step (gid13, $t = 3$, $\Delta\mathrm{Vul} = 0.230$) is covered by its own concurrent certificate ($B_3 = 0.303$); the certificate quantities (M_q, A_i^S, ε_i) show no drift amplification when later resolves start from the previously spliced profile. One boundary case: gid9 has $\mathrm{cum}_3 = 0.0163$ against $\sum_t B_t = 0.0050$ and is absorbed by the tolerance at $c_1 = 0$ (any $c_1 \geq 0.04$ covers it without tolerance); the other 13 instances are net-covered by the margin term alone. The δ_∞ values used are sampled *lower* bounds.

Caliber note. The instance family has public-tree depth ≤ 3 (decision depths 0–2), so no depth-3 subgames exist; "depth $1/2/3$" is realized as a depth ladder $1/2$ plus sequential step counts $T = 1/2/3$. The data support Corollary 6 on trees of this depth.

H Per-Pair Statements: What Survives and What Fails

The main theorem is stated at per-player granularity, with the margin deficits of *all* opponents summed. This appendix asks whether, under stronger structural conditions, a *per-pair* safety statement holds: a bound on each pairwise vulnerability by that pair's own certificate. Under two additional assumptions, half of the literal per-pair form is a theorem and the other half is false, with Example 2 as the minimal counterexample.

Additional assumptions. Fix a constant-sum polymatrix (CSP) witness $\{u_{pq}\}$ attaining $\delta_\infty = 0$.

- **(F3) Pair-local witness:** the witness can be chosen so that u_{pq} is a function $f_{pq}(\omega_c, s_p, s_q)$ of the chance outcome and the (private type, i.e. card, action sequence) of players p and q only.
- **(F4) Pair blindness:** for every pair (p, q) not containing the resolver i, the joint distribution of (ω_c, s_p, s_q) does not depend on i's strategy inside S (sufficient condition: p and q cannot observe i's in-subgame actions before completing their own, and no chance move that enters ω_c follows an in-subgame action of i).

Both $\delta_\infty = 0$ families used in this paper (pairpot at $\lambda = 0$ and pairwise-showdown) satisfy F3 and F4: pairwise settlement is per-edge independent and actions are mutually unobserved.

Pairwise vulnerability. For a witness $\{u_{pq}\}$, define

$$\mathrm{PairVul}_{pq}^{\sigma_{-(p,q)}}(\sigma_p, \sigma_q) \;:=\; \max_{\tau_p} \mathbb{E}_{(\tau_p, \sigma_q, \sigma_{-(p,q)})}[u_{pq}] \;-\; \mathbb{E}_{(\sigma_p, \sigma_q, \sigma_{-(p,q)})}[u_{pq}],$$

and write $X_{ij} := \mathbb{E}_{\sigma^1}[u_{ij}] - \mathbb{E}_{\sigma^0}[u_{ij}]$ for the resolver's on-path value change on board (i, j) (the *board* (p, q) is the pair table u_{pq}). Throughout, σ^1 is a legal update in the sense of Theorem 1 (only resolver i's strategy inside S changes), $\delta_\infty = 0$, and the certificate quantities $\hat{m}$, M_q, A_i^S, ε_i are as in Section 3.

Proposition 6 (Per-pair statements under F3 and F4)**.** *Under Assumption 1, $\delta_\infty = 0$, F3, and F4:*

(C1) ***Identity.*** *For every opponent j: $\mathrm{gain}_j = -X_{ij}$ (boards not containing i are frozen by F3/F4, and the (i, j) board is per-edge constant-sum under the $\delta_\infty = 0$ witness), and $\sum_j X_{ij} = \mathrm{gain}_i = A_i^S - \varepsilon_i$. The edge constants c_{pq} may be nonzero; they cancel in every difference used below, so the identities are unaffected.*

(C2) ***Deviator-i pairs: the literal form holds with constant 1.*** *$\Delta\mathrm{PairVul}_{ij} = -X_{ij} \leq (X_{ij})_- \leq M_j$: the max term of $\mathrm{PairVul}_{ij}$ does not depend on σ_i, and j's pairwise on-path gain equals its full-utility on-path gain under F3/F4, which is priced by j's own margin deficit.*

(C3) ***Telescoping identity.*** *$\Delta\mathrm{PairVul}_{ij} + \Delta\mathrm{PairVul}_{ji}$ equals the change in j's pair best-response value, $\max_{\tau_j} U_{ji}(\tau_j, \sigma_i^1) - \max_{\tau_j} U_{ji}(\tau_j, \sigma_i^0)$.*

(C4) ***Deviator-j pairs: the literal form is false.*** *There is a game satisfying all the hypotheses above (Example 2) with $\varepsilon_i = 0$ and $\hat{m}_j(\iota) = 0$ for all ι, yet $\Delta\mathrm{PairVul}_{ji} = 1 > 0 = (-\hat{m}_j)_+$. No bound of the form $\Delta\mathrm{PairVul}_{ji} \leq C_m(-\hat{m}_j)_+ + C_\varepsilon\,\varepsilon_i$ can hold with any finite constants.*

The mechanism behind (C4) is cross-opponent value transfer: in Example 2 the resolver's update moves value from the (i, k) board to the (i, j) board while leaving its own total unchanged, so the *victim* j's margin stays silent (j's follow-best-response is held exactly at its blueprint counterfactual value), and the alarm fires at the *beneficiary* k instead ($\hat{m}_k = -1$). A per-pair bound keyed to the victim's own certificate is therefore unobtainable; Theorem 1(ii) prices the trade at the per-player level, through the cross terms $\sum_{k \neq j} M_k + A_i^S$. A joint gadget cannot issue single-margin per-pair certificates; it issues additively priced per-player certificates, and every cross-opponent trade is charged in full to the beneficiary's margin plus the resolver's blueprint slack.

Numerically, (C2) and (C4) are exercised by the 97-check suite of Appendix F: Example 2 reproduces $\hat{m}_j = 0$, $\hat{m}_k = -1.0$, $\Delta\mathrm{PairVul}_{ji} = 1.0 > (-\hat{m}_j)_+ = 0$, and the aggregate bound of Lemma 1(iii) holds with equality ($\Delta\mathrm{Vul} = 1.0 = \sum_q M_q$).

I Experimental Details

This appendix collects the protocol details behind Section 7: the search procedure and the affine encoding of the attack, the kill-filter tiers, the n-sweep and its restricted caliber, and the fitting LP used for the minimal-constant analysis. Environment throughout: single-core CPU, seed 42; payoffs and best responses are evaluated without sampling, LPs are solved on the full table to the tolerances stated in Appendix I.1, and restricted-support distances are reported as lower bounds.

I.1 Heuristic screen and certified attack campaign (M1)

Screen. 210 random small extensive-form games (three families $\times$ 70: random, near-CSP with coupling $\lambda \sim U[0, 0.3]$, and synthetic with $\lambda \sim U[0, 0.3]$; depth ≤ 3, 2–3 actions per node, random utilities), crossed with 2 corruption arms (none / uniform-or-Dirichlet reset of the resolver's subgame strategy), for 420 configurations. Each runs blueprint CFR (400 iterations) $\rightarrow$ corrupt $\rightarrow$ gadget resolve (800 iterations) $\rightarrow$ *exact* best-response measurement of $\Delta\mathrm{Vul}$. The screening objective is $\Delta\mathrm{Vul}_{\max} - 0.03\,\delta_\infty - 0.03\,(-\mathrm{margin}_{\min})_+$ with tolerance 0.02 (the pilot's CFR noise floor); the 0.03 constants are the pilot's benign-family fit. The top-5 points get hill climbing (Gaussian leaf-utility mutation, 25 steps) and the top-3 a deep re-solve ($7.5\times$ blueprint and $5\times$ resolve iterations, $\times 3$), for 433 data points in total. δ_∞ is an exact LP for small games and a restricted-support *lower bound* for large ones (flagged per point; a lower bound overestimates the objective). All 420 stage-1 records were completed by deterministic replay with zero replay deviation (see also Appendix F).

Affine encoding and certified search. The exact campaign replaces heuristic best responses by exactly evaluated ones. Margins enter the model through an affine encoding whose values were cross-checked against brute-force enumeration with maximum deviation 1.1×10^{-16} on the smoke-test instance, and by spot checks on a perturbation grid where the mixed-integer linear program (MILP) optimum coincided with exact brute force. On the full campaign (162 instances, 0 errors) the encoding-consistency check reports maximum absolute deviation 9.1×10^{-10}, and the attack statuses over the 486 (instance, player) searches split as CERTIFIED 380 / VIOLATION 78 / BUDGET 28; the search objective uses the $(c_1, c_2) = (1, 0.5)$ caliber. VIOLATION here means the searched bound with those trial constants is exceeded, not a theorem violation; every such point is then run through the kill filters below. BUDGET means the cutting-plane attack exhausted its δ_∞-oracle budget (60 exact LP calls) before closing the interval between its incumbent (the best value the search has reached) and its upper bound; for those 28 searches only the exactly evaluated incumbent enters any table, and the attack interval above it is open, so the campaign is complete as a certificate of every reported point and incomplete as a search on those 28. Terminology used throughout: payoffs, reach probabilities, and best responses are evaluated without sampling (backward induction over the enumerated pure strategies, ≤ 128 per player, in floating point); δ_∞ is the optimum of the full-table LP where the profile table fits the enumeration cap and is otherwise reported as a restricted-support *lower bound*; each attack search is bounded per pure strategy by an LP over globally valid δ_∞ cuts and closed at absolute gap 10^{-7} (CONVERGED) or by an upper bound below tolerance (CERTIFIED). All LPs are solved with HiGHS (Huangfu & Hall, 2018) at its default primal and dual feasibility tolerances (10^{-7}), and "exact" in this paper means full-table and sample-free to those tolerances, a weaker requirement than exact arithmetic. The reported numbers come from the certified cutting-plane search. The replay run behind Appendix I.5 used Python 3.12.3, NumPy 1.26.4, and SciPy 1.11.4,

and it reproduces the campaign values exactly: all 162 ΔVul values and the minimum zero-tax slack 4.49×10^{-6} coincide. The library versions of the original cluster run were not logged.

I.2 KILL-FILTER TIERS

A *theorem-level* counterexample must combine small δ_∞, a silent margin certificate, and a real vulnerability increase not explained by A_i^S or ε_i. The screen used four nested threshold tiers on its 433 points:

Tier	Margin condition	δ_∞ condition	Hits ($\Delta\text{Vul} > 0.02$)
1 (strict)	$\text{margin} \geq 0$	$\delta_\infty \leq 0.10$	**0**
2	$\text{margin} \geq -0.05$	$\delta_\infty \leq 0.10$	**0**
3	$\text{margin} \geq -0.05$	$\delta_\infty \leq 0.15$	**0**
4 (very loose)	$\text{margin} \geq -0.10$	$\delta_\infty \leq 0.30$	3 (each with $\delta_\infty \geq 0.05$ or margin ≤ -0.09)

In the small-δ_∞ region ($\delta_\infty \leq 0.1$, 156 configurations) the maximum ΔVul is 0.039, with only 2 points above tolerance, both with margin ≤ -0.09; across all 56 above-tolerance points, the smallest δ_∞ is 0.051 and the largest implementation-caliber margin is $+0.031$ (that point having $\delta_\infty = 0.43$; it is the gid174 construction of Appendix E.2).

The certified attack campaign uses the corresponding certified filters, in two margin calibers (BR margin $\hat{m}$ and implementation margin): *strict* = small $\delta_\infty \wedge$ quiet $\hat{m} \wedge \Delta\text{Vul} > \text{tol}$, excluding points explained by A_i^S; *relaxed* = $\delta_\infty \leq 0.15 \wedge \sum_q M_q \leq 0.05$; *loose* = $\delta_\infty \leq 0.30 \wedge \sum_q M_q \leq 0.10$. On the full 162-instance set: strict **0 hits** in both margin calibers; relaxed 3 and loose 6 hits, all accompanied by a firing $\sum_q M_q$ or an A_i^S/ε_i explanation, hence no theorem-level candidates. The single strict candidate under the (non-theorem) fixed-certificate caliber is gid183, discussed in Appendix E.4.

I.3 n-SWEEP: PROTOCOL AND RESTRICTED CALIBER

The sweep runs the M1-style pipeline at $n \in \{3, 4, 5\}$ (target 32 instances per n) and fits, per n, the c_1 coordinate of the minimum-sum fit (see Appendix I.4). Landed coverage: $n = 3$: 32/32 instances (98 points); $n = 4$: 31/32 (91 points; one worker process died early); $n = 5$: **15/32** (36 points, 7 binding), closed out at a 24-hour compute cutoff. The $n = 5$ column reflects the 15 fastest instances; the missing slow instances could be more adversarial, so the $n = 5$ figures carry a fast-instance selection bias.

δ_∞ caliber under the enumeration cap. The exact-δ_∞ routine enumerates pure-strategy profiles up to $\text{CAP} = 60{,}000$; instances whose profile table exceeds the cap are *skipped entirely*, never degraded to restricted-support lower bounds. All 15 landed $n = 5$ records have exactly 1,024 profiles ($\ll$ CAP), so every $n = 5$ δ_∞ in the sweep is an exact LP value.

Results. Main caliber (joint LP fit): $c_1 = (0.162,\ 0.563,\ 0.916)$ at $n = 3/4/5$, slope 0.377 per player, second difference -0.049. Pinned $c_2 = 0$: $c_1 = (0.190,\ 0.563,\ 0.916)$, second difference -0.020. Pinned $c_2 = 0.5$ (secondary caliber): $c_1 = (0.162,\ 0.260,\ 0.606)$, second difference $+0.247$. These are exploratory ablation fits: the law omits A_i^S and ε_i, the fitted c_1 is a necessity estimate on the points found, and the $n = 5$ column carries the fast-instance selection above, so none of the three sequences measures the scaling of the theorem's coefficient and no comparison with $2n$ is drawn. The $n = 3$ ablation on this family: margin-only ($c_1 = 0$) is infeasible within the two-term law (the uncovered points can be carried by δ_∞, or by the A_i^S and ε_i terms the law omits), and $c_2 = 0$ inflates the required c_1 by ratio 1.17 (0.190 vs 0.162); this family is less adversarial than the attack family of the MILP campaign, so the two sets of numbers are reported separately.

I.4 FITTING LP FOR MINIMAL CONSTANTS

Given a point set with per-point data $(\Delta\text{Vul}_r, \delta_r, M_r)$, where δ_r is the instance's δ_∞ and M_r its margin-deficit aggregate, the minimal-constant analysis solves the LP

$$\min\ c_1 + c_2 \quad \text{s.t.} \quad \Delta\text{Vul}_r\ \leq\ c_1\,\delta_r + c_2\,M_r + \text{tol}\ \ \forall r, \qquad c_1, c_2 \geq 0,$$

with $\mathrm{tol} = 0.02$ (and a $\mathrm{tol} = 0$ variant), plus pinned solves at $c_2 = 0$, $c_2 = 0.5$, and $c_1 = 0$. The result is a *necessity lower-bound estimate* on the constants needed for the hardest points found, not a sufficiency certificate. Headline values: full attack campaign (162 instances, pipeline/$\sum_q M_q$ caliber, tol 0.02): $(c_1, c_2) = (1.76, 0.31)$ with $c_1 + c_2$ minimum 2.07 and 17 binding points; $c_2 = 0$ forces $c_1 \geq 2.64$; $\mathrm{tol} = 0$ gives $(2.35, 0.17)$; the earlier 118-instance snapshot gave $(0.34, 0.81)$, and the trial caliber $(1, 0.5)$ leaves 2 points uncovered at full volume. The heuristic screen (433 points, a differently adversarial family, reported only as a footnote in Section 7): $(0.39, 0.30)$ with tolerance, $(0.11, 0.69)$ without; $c_2 = 0$ forces $c_1 \geq 1.33$ (tol) / 1.99 (no tol); $(1, 0.5)$ covers all 433 points. The two-term law omits the A_i^S and ε_i terms of Corollary 5, so the fitted pair is not a tightness measurement of the proven coefficients $(2n, 1) = (6, 1)$.

Table 5: Minimum feasible constants of the two-term ablation $\Delta\mathrm{Vul} \leq c_1\delta_\infty + c_2 \sum_q M_q + \mathrm{tol}$, reproduced from Section 7; the first row lists the coefficients of the same two terms in Corollary 5, whose bound also carries A_i^S and ε_i. Fits use the full 162-instance M1 set, pipeline/$\sum_q M_q$ caliber, $n = 3$. Fitted values are lower-bound estimates, not proven guarantees.

Source	Condition	c_1	c_2
Corollary 5	coefficients of δ_∞, $\sum_q M_q$ ($n = 3$)	6	1
Fitted (LP)	$\mathrm{tol} = 0.02$	1.76	0.31
Fitted (LP)	$\mathrm{tol} = 0$	2.35	0.17
Fitted (LP)	$c_2 = 0$ pinned	≥ 2.64	0
Fitted (LP)	$c_1 = 0$ pinned (margin-only)	infeasible	

I.5 Certificate usefulness on the gadget's own outputs

The campaigns above test the inequality; this experiment asks what the certificate does for the updates the gadget actually delivers. **Crosswalk to M1.** The M1 campaign records, per instance, one *nominal* update (the delivered update $\sigma^0 \to \sigma^1$ on the unperturbed utility table: CFR blueprint 400 iterations, gadget resolve 800 iterations, arms *none*/*corrupt*) and, where the attack found an argmax, one *pipeline* update (a fresh blueprint and resolve on the perturbed table). Result 4 covers both sets (162 nominal and 150 pipeline points, the 312 theorem-caliber points of Appendix F).

This experiment re-runs exactly the 162 nominal updates, same seeds, same game, arm, table, blueprint, and update, and recomputes $\Delta\mathrm{Vul}$ by exact best response; all 162 values coincide with the recorded ones, and the minimum slack of the zero-tax bound (4.49×10^{-6}) is identical to the recorded value, so the two sets of measurements are the same points. Coverage is tested at an absolute tolerance of 10^{-9}. The 30 pool records that M1 skipped are set aside and are not used here.

For each update we evaluate the four instantiations of Theorem 1,

$$B_X \;=\; \max\big(\varepsilon_i,\; A_i^S - \varepsilon_i + \textstyle\sum_q M_q + X\big), \qquad X \in \{\Lambda,\; L_i,\; \Theta,\; 2n\delta_\infty\},$$

the zero-tax bound $X = 0$ (the B of Result 4, which drops the structural term but keeps activation and solve error), and the direct per-player form $B_{\mathrm{dir}} = \max(-g_i, \max_{j \neq i}[M_j - g_j])$ of Eq. (1), all with BR-certified margins.

Λ costs two expected-utility evaluations; L_i one best response of the resolver confined to S on the total utility (Appendix C); $\Theta = 2\min(\zeta_{\mathrm{prof}}, \rho_S\zeta_S)$ the full pure-profile table (at most 6,561 profiles here); $2n\delta_\infty$ the exact δ_∞ LP; B_{dir} the margins and the n expected utilities only.

Table 6 reports coverage of $\Delta\mathrm{Vul}$, the slack $B - \Delta\mathrm{Vul}$, how many updates each form *accepts* at a tolerance ($B \leq \mathrm{tol}$), and wall-clock. The chain $\Lambda \leq L_i \leq \Theta \leq 2n\delta_\infty$ holds on all 162 instances, $L_i \geq 0$ throughout, $\Lambda < 0$ on 91 (the update raised total utility), and $B_{\mathrm{dir}} \leq B_\Lambda$ on all 162 (strictly on 147). Because every form is a valid upper bound, no accepted update has $\Delta\mathrm{Vul} > \mathrm{tol}$ and every harmful update (19 with $\Delta\mathrm{Vul} > 0.02$) is flagged by every form; the forms differ only in how many of the 143 safe updates they let through, from 73 (B_{dir}) down to 2 ($2n\delta_\infty$). The zero-tax bound is not a theorem for $\delta_\infty > 0$ but covers all 162 nominal points (Result 4); on the 150 pipeline points it does not.

Runtime is measured on the same replay (single core, $n = 3$, small trees), as the total wall-clock of each certificate (margins, activation identity where used, and the structural term) against the exact

recomputation of Vul_G for both profiles (six full-game best responses, median 2.2 ms). The ratio reported is the median over the 162 instances of the paired per-instance ratio (certificate time divided by recomputation time on the same instance), so a ratio and a median time need not multiply out.

B_{dir} costs $0.57\times$ at the median (1.3 ms; $9.3\times$ worst); the Λ form, which adds the confined activation best response, $1.05\times$ (2.3 ms; $11.9\times$ worst); the L_i form $1.6\times$ (3.7 ms); the Θ form $28\times$ (123 ms; $68\times$ worst) because it enumerates the pure-profile table; the $2n\delta_\infty$ form $381\times$ (2.16 s; the exact δ_∞ LP, replayed here and matching the recorded value on all 162 instances) so Θ and δ_∞ are design-time quantities. The zero-tax bound shares the margin and activation computation of the Λ form and costs $0.82\times$ (1.8 ms).

The baseline for B_{dir} is a cached exact evaluation. Because the update is unilateral, $\mathrm{BRval}_i(\sigma^1_{-i}) = \mathrm{BRval}_i(\sigma^0_{-i})$, so the exact $\Delta\mathrm{Vul}$ follows from the blueprint's stored best-response values plus $n-1=2$ full-game best responses on σ^1, not six. Re-timing the three on one process over the same 162 updates (five repeats, minimum taken): the cached evaluation costs $0.37\times$ the full recomputation, and B_{dir} costs $1.49\times$ the cached evaluation (paired median; worst $2.16\times$; $0.55\times$ the full recomputation in this re-timing, against $0.57\times$ in the original measurement) and is faster than it on none of the 162 updates. On trees this small the subgame-confined best responses save little, so the saving B_{dir} offers is locality, not raw time: it never leaves S and its entry sets, whereas the cached baseline still needs full-game best responses for the $n-1$ opponents. Appendix I.6 measures when that locality pays by growing the trunk around a fixed subgame. The cached baseline also exposes the cost of zero tolerance: 73 of the 162 updates have $\max_p[A_p(\sigma^1) - A_p(\sigma^0)] \le 0$, yet at tolerance 0 B_{dir} accepts 2 and every form of Theorem 1 accepts none, the behaviour Appendix C derives for constant-sum games (none of the 162 games is exactly constant-sum; the median $|\Lambda|$ is 8×10^{-4}). The baseline-aware form B_{cache} of Corollary 3, replayed on the same 162 updates with the same cached A^0_p, covers $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0)$ on all 162, is at most B_{dir} on all 162 (strictly on 112), accepts 37 updates at tolerance 0 and 104/122/143 at 0.02/0.05/0.10 (against 2/73/85/113 for B_{dir}), never accepts an update with $\mathrm{Vul}_G(\sigma^1) > \mathrm{Vul}_G(\sigma^0) + \mathrm{tol}$, and costs $1.00\times$ B_{dir} and $1.47\times$ the cached evaluation (paired medians; worst $2.06\times$). At tolerance 0 it accepts 8 updates whose per-player maximum increase is positive while Vul_G itself does not rise, which is the gap between the two calibers of the caption. The gadget's own proxy margin is negative on 148 of 162 updates while the BR-certified deficit $\sum_q M_q$ is positive on 145; the two disagree on 5 (four proxy alarms with zero certified deficit, one certified deficit the proxy misses), consistent with Proposition 5.

Table 6: Certificate usefulness on the 162 delivered (nominal) M1 updates (Appendix I.5). Cover: instances with $B \ge \Delta\mathrm{Vul}$. Slack: $B - \Delta\mathrm{Vul}$. Acc. t: instances with $B \le t$ (every accepted update has $\Delta\mathrm{Vul} \le t$; all 19/12/5 updates with $\Delta\mathrm{Vul} > 0.02/0.05/0.10$ are flagged by every form). Time: median total wall-clock of the certificate and, in parentheses, the median paired per-instance ratio to the exact recomputation of Vul_G for σ^0 and σ^1 (median 2.2 ms). $\Delta\mathrm{Vul}$ is the per-player maximum increase $\max_p[A_p(\sigma^1) - A_p(\sigma^0)]$ of Theorem 1(iii), which every form above the rule bounds. B_{cache} bounds $\mathrm{Vul}_G(\sigma^1) - \mathrm{Vul}_G(\sigma^0)$ instead, and its Cover, Slack, and Acc. entries are measured against that quantity (105 of the 162 updates have $\mathrm{Vul}_G(\sigma^1) \le \mathrm{Vul}_G(\sigma^0)$, 73 have $\max_p[A_p(\sigma^1) - A_p(\sigma^0)] \le 0$); its time is the paired ratio to B_{dir} in the same process. The last row is the exact value; the other rows are bounds. It is the runtime baseline for the direct forms.

Form	Cover	Slack med. / q90	Acc. 0 / 0.02 / 0.05 / 0.10	Time (ratio)
B_{dir} (Eq. (1))	162/162	0.018 / 0.22	2 / 73 / 85 / 113	1.3 ms ($0.57\times$)
B_{cache} (Corollary 3)	162/162	0.020 / 0.22	37 / 104 / 122 / 143	1.2 ms ($1.0\times B_{\mathrm{dir}}$)
$X = \Lambda$ (realized leak)	162/162	0.031 / 0.31	0 / 72 / 82 / 105	2.3 ms ($1.05\times$)
$X = L_i$ (confined bound)	162/162	0.094 / 0.55	0 / 45 / 57 / 78	3.7 ms ($1.6\times$)
$X = \Theta$ (sharp tax)	162/162	0.273 / 1.22	0 / 19 / 29 / 42	123 ms ($28\times$)
$X = 2n\delta_\infty$ (Corollary 5) (zero-tax B	162/162	1.007 / 2.34	0 / 2 / 3 / 6	2.16 s ($381\times$)
$X = 0$ of Result 4)	162/162	0.037 / 0.32	0 / 70 / 80 / 99	1.8 ms ($0.82\times$)
$\Delta\mathrm{Vul}$ (cached exact)	162/162	0 / 0	73 / 143 / 150 / 157	0.83 ms ($0.37\times$)

I.6 TRUNK SCALING: CERTIFICATE COST WITH A FIXED SUBGAME IN A GROWING GAME

The cached exact evaluation needs $n-1$ best responses on the *whole* game, whereas B_{dir} needs best responses confined to S and its entry sets. On the small trees of Table 6 the two are within a factor of 1.5, so the comparison is repeated with the subgame held fixed and the rest of the game grown around it. The base game is the synthetic family of Section 7 ($\lambda = 0.1$, six seeds, 12 information sets, resolver $i =$ seed $\bmod 3$). A public chance move at the root with K equiprobable outcomes selects one of K copies of the base game, each with its own information sets; S is the $X = B$ public subtree of the first copy, so S has 6 information sets and 8 entry histories for every K, while the game has $12K$ information sets. The blueprint is the 400-iteration CFR average of the base game, replicated to every copy; the update is a Dirichlet corruption of the resolver's blueprint inside S followed by the 800-iteration gadget resolve, with the same corruption draw for every K.

Four judges of the same update are timed in one process (seven repeats, median; $K \in \{1, 2, 4, 8, 16, 32, 64\}$): the full recomputation of Vul_G ($2n$ full-game best responses); the cached exact evaluation ($n-1$ full-game best responses and one expected-utility pass); B_{dir} as timed in Table 6, which locates the entry histories by a pass over the whole tree and evaluates the n expected utilities on the whole game; and a local B_{dir} that reads the blueprint's entry reaches from the cache, as the cached baseline reads its stored best-response values, and evaluates the realized gains $g_p = u_p(\sigma^1) - u_p(\sigma^0)$ inside S only (exact, because the update is confined to S). The two B_{dir} judges agree to 10^{-9} on all 42 (seed, K) pairs, the two exact judges agree likewise, and $B_{\mathrm{dir}} \geq \Delta\mathrm{Vul}$ on all 42. Table 7 reports the medians over the six seeds. The local certificate is flat at 2.0 ms for every K, while the cached evaluation grows linearly with the game, from 2.1 ms at $K = 1$ to 145 ms at $K = 64$: the paired ratio of local to cached falls from $0.96\times$ to $0.014\times$, and the cached evaluation is faster than the local certificate on none of the 42 instances. The whole-tree B_{dir} also grows linearly, because its entry search and expected-utility passes visit the whole game, but at a smaller slope ($0.76\times$ the cached evaluation at $K = 64$, against $1.35\times$ at $K = 1$), since an expected-utility pass is cheaper than a best response. The base game keeps the subgame at $1/(2K)$ of the blueprint's reach, so $\Delta\mathrm{Vul}$ and every bound shrink with K as well; the table therefore reports runtime only.

Table 7: Trunk scaling (Appendix I.6): median wall-clock over six seeds of four judges of the same update, with the subgame fixed (6 information sets) and the game grown to $12K$ information sets. Ratios are medians of paired per-seed ratios to the cached exact evaluation.

K	Infosets	Full	Cached exact	B_{dir}, whole tree	B_{dir}, local	Whole tree / cached	Local / cached
1	12	5.5 ms	2.1 ms	2.8 ms	2.0 ms	1.35×	0.96×
2	24	11.1 ms	4.1 ms	4.3 ms	2.0 ms	1.06×	0.49×
4	48	22.4 ms	8.1 ms	7.6 ms	2.0 ms	0.93×	0.25×
8	96	46.4 ms	17.1 ms	13.8 ms	2.0 ms	0.83×	0.12×
16	192	90.9 ms	32.8 ms	27.6 ms	2.0 ms	0.82×	0.061×
32	384	189 ms	70.6 ms	54.4 ms	2.0 ms	0.77×	0.029×
64	768	401 ms	145 ms	111 ms	2.0 ms	0.76×	0.014×

J DEFINITIONAL CALIBERS: IMPLEMENTATION VERSUS THEOREM

The resolving implementation used for the experiments and the theorem-side definitions of Section 3 differ in six conventions, listed in Table 8. Re-implementations of the certificates and of δ_∞ follow the theorem column; comparisons against the implementation follow the implementation column.

D-a (normalization). The implementation divides counterfactual action values by the weight sum. Per infoset the sign of the margin is unaffected (T and F branches are divided by the same positive number), but the cross-infoset aggregate differs. M_q is the unnormalized $\bar{W}_q(\iota)$-weighted sum; when each opponent's own action sequence is identical across entries (all pre-entry actions public), $\sum_\iota \bar{W}_q(\iota) \leq 1$ and $M_q \leq (-\min_\iota \hat{m}_q(\iota))_+$ connects it to the min-margin caliber (Appendix B); private pre-entry branching can push the sum above 1.

D-b (average-strategy margins are optimistic). The implementation values an opponent's F branch under that opponent's CFR *average strategy*, while the certificate requires the F-branch *best-response*

ID	Implementation	Theorem caliber	Consequence
D-a	Counterfactual action values *normalized* by the weight sum	M_q aggregates with unnormalized weights $\bar{W}_q(\iota)$	Per-infoset margin signs agree; cross-infoset aggregation differs
D-b	Margin's F-branch valued under the opponent's own *average strategy*	Margin $\hat{m}$ valued against the exact subgame *best response*	Implementation margin optimistic on 33/40 infosets, worst overestimate $+1.667$
D-c	F-branch evaluated under the gadget-internal measure, which depends on the opponent *ordering* in the gadget chain (opponents' gadgets nested one inside the next)	Both branches evaluated under the game-side measure $c_e\,\pi^{\sigma^0}_{-j}(e)$ and environment $(\sigma^1_i, \sigma^0_{-ij})$	Measures coincide only when entries follow a public chance node, there are no trunk actions, and the other opponents play T purely (the benchmark families); differ in general
D-d	Zero-denominator entry infosets fall back to a raw mean (flagged)	Assumes $\bar{W}_q(\iota) > 0$	Such ι carry weight 0 in M_q
D-e	δ_∞ LP defined on the reduced pure-strategy *profile table*	Same	A literal terminal-history-dependent witness would force $\delta_\infty \equiv 0$ on all zero-sum games, contradicting the Kuhn bound
D-f	LP constraint is per-edge *constant-sum*: $u_{pq}(a,b) + u_{qp}(b,a) = c_{pq}$	Same (constant-sum, not antisymmetric)	Proofs use only that $\sum_p u_p$ is constant on the class

Table 8: Definitional conventions of the implementation and of the theorem.

value. With environment and measure fixed, the average-strategy value can only understate the best response, biasing the implementation margin optimistic; the effect is worst when the opponent plays T almost purely in the gadget, leaving its F-branch behavior off-path and essentially untrained. Measured on the benchmark families: the implementation margin is at least the BR margin on 33 of 40 entry infosets, with worst overestimate $+1.667$ (the information set holding the highest card in pairwise-showdown); the 7 reversals, caused by the D-c measure gap, all have magnitude $\leq 1.2 \times 10^{-3}$, so the net direction is not pointwise constant. $\hat{m}$ is recomputed a posteriori by one exact subgame best response per opponent.

D-c (gadget-internal measure depends on opponent order). The implementation evaluates F-branch values under the gadget's internal measure: root chance proportional to $c_e \cdot r_i(e)$ (excluding the other opponents' trunk reach) times the T/F mixing probabilities of the opponents earlier in the gadget's nesting chain, which depends on the order in which opponents are chained; other opponents' in-subgame behavior is their gadget strategy. The theorem measure is the game-side $c_e\,\pi^{\sigma^0}_{-j}(e)$ with environment $(\sigma^1_i, \sigma^0_{-ij})$. The two coincide when entries immediately follow a public chance node, there are no trunk actions, and the other opponents play T purely (the benchmark families); they differ in general. The implementation's T-branch value (the blueprint counterfactual value) uses the game-side measure; $\hat{m}$ puts both branches on that measure.

D-d (zero-denominator entries). Entry infosets with zero counterfactual weight fall back to a raw unweighted mean and are flagged. They contribute weight 0 to M_q.

D-e (δ_∞ is defined by LP profile-table semantics). If the pairwise witness u_{ij} were allowed to depend arbitrarily on the terminal history z, every pointwise zero-sum game would admit $u_{ij} := (u_i - u_j)/n$ and $\delta_\infty \equiv 0$, contradicting the 3-player Kuhn bound $\delta_\infty \geq 1$. Definition 4 places the witness table on reduced pure-strategy *pairs* (ρ_p, ρ_q) (after taking expectation over chance and third parties): the witness may depend on third-party randomness through the chance expectation, but not on third-party *strategies*.

D-f (per-edge constant-sum, not antisymmetric). The LP implements the constant-sum polymatrix (CSP) class of the literature, $u_{pq}(a,b) + u_{qp}(b,a) = c_{pq}$ with the edge constants c_{pq} free LP variables, which contains the antisymmetric class $u_{pq} + u_{qp} = 0$. The proofs use only that $\sum_p u_p$ is constant

on the class; at $\delta_\infty = 0$ the conclusion $\sum_p \mathrm{gain}_p = 0$ follows because the constant cancels in differences. The benchmark $\delta_\infty = 0$ families are also pointwise zero-sum on terminals.