STATE-CONSTRAINED CONTINUOUS-CONTROL ZERO-SUM DYNAMIC GAMES WITH ONE-SIDED PAYOFF INFORMATION

Anonymous authors
Paper under double-blind review

ABSTRACT

We study G, a K-stage zero-sum dynamic game with continuous and simultaneous controls, one-sided payoff information, and public state constraints. G arises in defense, sports, finance, among other agentic competitions. Nash equilibrium strategies in G may mix over a continuum of controls, producing an intractable game tree. To this end, we introduce a pair of information-relaxing Stackelberg variants of G where one player best responses to their opponent's next control, and prove that Nash leaders in these variants use at most I control atoms at any information states. Hence the game tree complexity of these variants are at most I^K. We then show that under sufficient game-theoretic and geometric assumptions, this pair of atomic leader strategies have vanishing exploitability in G as the time grid refines. This result justifies a computationally tractable approximation of the Nash of G via atomic strategy restriction and bounded control-atom insertion. Effectiveness of our method is demonstrated through an analytical Hexner's game and a car chasing game with nonlinear dynamics and nonconvex state constraints.

1 INTRODUCTION

Many real-world games combine hidden intent with hard physical constraints. An attacker in football knows the intended target and strategically plans their movement in order to exploit information advantage, while avoiding being tackled. A defender, on the other hand, must dynamically update risk estimation and defend against all possible consequences. Similar structures may arise in defense, finance, cybersecurity, among other agentic competitions.

Our reference game G is modeled to be two-player, zero-sum, and have K stages. Before the game starts, nature draws one of I possible payoff types and reveals it to Player 1 (P1). P2 only knows the prior $p_0 \sim \Delta(I)$. At the beginning of each stage, players simultaneously choose their constant stage-wise controls from continuous and admissible control spaces. The state then evolves according to known deterministic dynamics, and a payoff is computed based on the state trajectory. P1 minimizes the expected payoff over p_0 with respect to their strategy, while P2 maximizes.

G is different from standard two-player zero-sum imperfect-information extensive-form games (2p0s IIEFGs) due to the involvement of continuous controls and state constraints: (1) IIEFGs assume U discrete actions, resulting in a game tree complexity of U^{2K}. G with continuous controls has an infinite game tree complexity, i.e., infinitely many information states along the tree; (2) similar to IIEFGs, P1 may strategically randomize to conceal or reveal its type to exploit information asymmetry. However, due to continuous controls in G, behavior strategies at any information state have to be represented by probability distributions over continuous control spaces, rather than over finite action sets. (3) Finally, safety entangles with strategic planning: the behavioral strategy of P1 must ensure safety against any P2 strategy, yet may also manipulate information in specific ways in order to take advantage of P2's own constraints.

Overview of methods. Our method is inspired by studies on information-relaxation games around G: In the *upper relaxation* $G^{\uparrow}$, P2 best responses to P1's current control in every stage. In the *lower relaxation* $G^{\downarrow}$, P1 best responses to P2. We extend an existing atomicity result from unconstrained games: leaders in state-constrained $G^{\uparrow}$ and $G^{\downarrow}$ require at most I control atoms at any information

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.

state. We then prove that such an atomic Nash pair creates a vanishing exploitability in G along time grid refinement, under empirically testable geometric and game-theoretic assumptions. This result suggests that we can approximate Nash of G by restricting its strategies to be I-atomic, reducing the game-tree complexity from U^{2K} to I^K. To further accelerate the solve of P2's Nash, we propose a bounded control atom insertion mechanism.

Contributions. Our main contributions are:

- We extend value existence and Nash atomicity results from unconstrained zero-sum dynamic games with one-sided information to G, where *hard state constraints present*. For value existence, we provide a sufficient repairability assumption missing in Ghimire et al. (2024). We also identify a translation invariance property in P2's dual game formulation, reducing P2's game tree complexity from $(I+1)^K$ in Ghimire et al. (2026) to I^K.
- We prove that the exploitability of the resultant atomic strategy pair in G vanishes along time grid refinement, under nontrivial yet empirically testable assumptions. This is a nontrivial extension from the unconstrained setting in Cardaliaguet (2009).
- We develop a solver for G combining bounded control-atom insertion and active-set Newton correction. Compared with existing game solvers for continuous control, such as NXDO and APSRO (McAleer et al., 2021; 2024), our solver has two advantages unique to G: (1) Our atomicity result gives an a priori control support bound: each relaxed leader admits an optimal strategy with at most I control atoms per information state. This bound guides tree initialization and permits support growth to stop at I atoms. (2) The dual variables resulted from solving P1's security game is a complete and compact representation of history for deriving P2's Nash, removing the need of learning P2's history encoder.

2 Related Work

2p0s games with one-sided information. Harsanyi's Bayesian games and Aumann-Maschler's repeated games provide the classical foundations for strategic interaction with private information (Harsanyi, 1967; Aumann et al., 1995). De Meyer introduced the dual-game construction that is now standard for recovering a security strategy for the uninformed player (De Meyer, 1996). In continuous-time differential games, Cardaliaguet established value existence, characterized the value by a belief-dependent Hamilton-Jacobi equation, and developed complementary primal and dual approximations (Cardaliaguet, 2007; 2009). Ghimire et al. introduced state-constrained primal-dual recursions (Ghimire et al., 2024). **IIEFGs and action abstraction.** Finite-horizon POSGs and perfect-recall IIEFGs are closely related models of sequential interaction under private information (Kovařík et al., 2022). CFR and newer last-iterate or sampled variants solve *finite* IIEFGs (Zinkevich et al., 2007; Lee et al., 2021; Liu et al., 2023; Meng et al., 2025; Zhao et al., 2025). Their tabular formulations maintain quantities for each represented information state and action. Storage requirements and convergence bounds therefore depend on the size of this representation, and large or continuous action spaces require abstraction. Fixed and adaptive action abstractions have been developed for poker and other games (Hawkin et al., 2011; Sandholm, 2015; Brown & Sandholm, 2015; Kroer & Sandholm, 2018; Li et al., 2024). Double-oracle methods iteratively expand a restricted game by adding best-response strategies (McMahan et al., 2003; Adam et al., 2021; McAleer et al., 2021). XDO and NXDO allow strategies from a population to be mixed separately at each information state. Our support bound is structural: each relaxed Bellman problem admits an optimal strategy with at most I control atoms, without an ex ante action discretization. This restriction is exact for the relaxations, and their optimal leader pair has vanishing exploitability in G as the time grid refines under our assumptions.

3 Reference Game G

Fix decision times $0 = t_0 < \cdots < t_K = T$. The public state is $x \in \mathcal{X} \subset \mathbb{R}^{d_x}$. P1 and P2 choose controls $u \in \mathcal{U}$ and $v \in \mathcal{V}$, respectively. At t_0, nature draws a payoff type $i \in \mathcal{I} := \{1, \dots, I\}$ from the prior $p_0 \in \Delta(I)$ and reveals it only to P1. The information state is represented by $(k, x, p) \in$

$[0,T]\times\mathcal{X}\times\Delta(I)$[1]. P1's type-$i$ behavioral strategy $\eta_{i,k}$ is a probability distribution over $\mathcal{U}$, and P2's behavioral strategy ζ_k a probability distribution over $\mathcal{V}$. Let $\eta_k := \{\eta_{i,k}\}_{i=1}^{I}$, $\eta := \{\eta_k\}_{k=0}^{K-1}$ and $\zeta := \{\zeta_k\}_{k=0}^{K-1}$. At (k,x,p), players independently sample and announce their control (u,v) based on (η_k,ζ_k), and hold it till $k+1$. The state trajectory evolves by a deterministic ODE

$$\dot{X}_k(s) = f_{0,k}(X_k(s)) + f_{1,k}(X_k(s),u) + f_{2,k}(X_k(s),v), \quad X_k(0)=x, \tag{1}$$

with endpoint $F_k(x,u,v) := X_k(h_k;x,u,v)$ and $h_k := t_{k+1}-t_k$. Belief p is Bayes updated based on P1's announcement of (u,η_k).

Let $\mathcal{C}=\{x\in\mathcal{X} : c_j(x)\le 0,\ j=1,\dots,m\}$ be the set of state constraints. With $\Gamma_{j,k}(x,u,v) := \max_{0\le s\le h_k} c_j(X_k(s;x,u,v))$, set $\mathcal{Q}_K=\mathcal{C}$ and define backward the *safe-control set*

$$\mathcal{U}_k^{\mathrm{safe}}(x) := \{u\in\mathcal{U} :\ \Gamma_{j,k}(x,u,v)\le 0\ \forall(j,v),\ F_k(x,u,v)\in\mathcal{Q}_{k+1}\ \forall v\}. \tag{2}$$

The *viable set* $\mathcal{Q}_k := \{x\in\mathcal{X} : \mathcal{U}_k^{\mathrm{safe}}(x)\neq\varnothing\}$ is the set of states from which P1 can keep the game safe through the remaining stages [2]. At $x\in\mathcal{Q}_k$, $\mathcal{U}_k^{\mathrm{safe}}(x)$ contains controls that keep the entire stage safe for every P2 control and leave the next state in $\mathcal{Q}_{k+1}$.

The type-i loss incurred during the stage is $r_{i,k}(x,u,v) := \int_0^{h_k}\ell_{i,k}(X_k(s;x,u,v),u,v)\,\mathrm{d}s$. The terminal loss is $g_i(x_K)$. For a continuation game starting at (t_k,x,p), continuation loss is

$$J^G_{k,x,p}(\eta,\zeta) := \mathbb{E}_{i\sim p,\eta_i,\zeta}\left[\sum_{\ell=k}^{K-1} r_{i,\ell}(x_\ell,u_\ell,v_\ell) + g_i(x_K)\right]. \tag{3}$$

Let Σ_1^G and Σ_2^G denote the admissible behavioral strategy classes. P1 minimizes this loss with respect to η and P2 maximizes with respect to ζ. The lower and upper continuation values of G are

$$V_k^{G,-}(x,p) := \sup_{\zeta\in\Sigma_2^G}\inf_{\eta\in\Sigma_1^G} J^G_{k,x,p}(\eta,\zeta), \quad V_k^{G,+}(x,p) := \inf_{\eta\in\Sigma_1^G}\sup_{\zeta\in\Sigma_2^G} J^G_{k,x,p}(\eta,\zeta). \tag{4}$$

At the fixed initial state $x_0\in\mathcal{Q}_0$ we abbreviate $J_p^G := J^G_{0,x_0,p}$ and the two values by V_G^- and V_G^+. The exploitability of a candidate pair (η,ζ) is

$$\mathrm{Gap}_p^G(\eta,\zeta) := \sup_{\zeta'\in\Sigma_2^G} J_p^G(\eta,\zeta') - \inf_{\eta'\in\Sigma_1^G} J_p^G(\eta',\zeta). \tag{5}$$

Now let $\Pi := \{0=t_0<\cdots<t_{K_\Pi}=T\}$ be the time grid, $|\Pi| := \max_k(t_{k+1}-t_k)$, and G_Π the reference game specified by Π. Our goal is to find (η,ζ) with $\mathrm{Gap}_p^{G_\Pi}(\eta,\zeta)\approx 0$ for small enough $|\Pi|$ so that (η,ζ) can be deployed as secure and safe strategies to continuous-time real-world games.

We first present Theorem 1 that proves value existence of G, which relies on standard regularity and repairability conditions in Assumptions 1 and 2, respectively. Proof in Appendix A.

Assumption 1 (Primitive regularity and viability)**.** *$\mathcal{X},\mathcal{U},\mathcal{V}$ are compact metric spaces. $f_{0,k}$, $f_{1,k}$, and $f_{2,k}$ in Equation* (1) *are continuous in all arguments, and uniformly Lipschitz in the state, thus for every $x\in\mathcal{X}$ and $(u,v)\in\mathcal{U}\times\mathcal{V}$, Equation* (1) *has a unique solution on $[0,h_k]$. $\ell_{i,k}, g_i$ and c_j are bounded and continuous. $\mathcal{Q}_k$ is nonempty. Both players have perfect recall (Kuhn, 1953).*

Assumption 2 (Safe-control repairability)**.** *For $x\in\mathcal{Q}_k$ and $u\in\mathcal{U}$, define the robust safety residual*

$$\Phi_k^{\mathrm{safe}}(x,u) := \max\left\{\max_j \sup_{v\in\mathcal{V}}[\Gamma_{j,k}(x,u,v)]_+,\ \sup_{v\in\mathcal{V}}\mathrm{dist}(F_k(x,u,v),\mathcal{Q}_{k+1})\right\}, \tag{6}$$

where $[a]_+ := \max\{a,0\}$, $\mathrm{dist}(y,A) := \inf_{a\in A} d(y,a)$, and $d(\cdot,\cdot)$ is the Euclidean distance. For every stage k, there are constants $\varepsilon_k>0$, $\kappa_k<\infty$, and $\theta_k\in(0,1]$ such that, for every $x\in\mathcal{Q}_k$ and $u\in\mathcal{U}$,

$$\Phi_k^{\mathrm{safe}}(x,u)\le\varepsilon_k \quad\Longrightarrow\quad \mathrm{dist}\big(u,\mathcal{U}_k^{\mathrm{safe}}(x)\big)\le\kappa_k\big[\Phi_k^{\mathrm{safe}}(x,u)\big]^{\theta_k}. \tag{7}$$

[1] Because the payoff type is static, the dynamics and losses are Markov, and all realized states and controls are publicly observed, the current physical state and public posterior form a sufficient information state for the Bellman recursion (Aumann et al., 1995; Cardaliaguet, 2007; Ghimire et al., 2024).

[2] For conciseness, we keep the discussion to the setting where constraints only apply to P1. In Appendix J.2, we show that all theoretical results apply to the more general setting where P2 has its own state constraints.

In words, a constant stage control with safety violation at most ϵ can be corrected by $O(\epsilon^{\theta_k})$ to become safe. This is the sample-and-hold analogue of the neighboring-feasible-trajectory and metric-regularity conditions used in state-constrained control (Frankowska & Vinter, 2000; Bettiol et al., 2019; Dontchev & Rockafellar, 2014).

Theorem 1 (Behavioral value of G)**.** *Under Assumption 1 and Assumption 2, G has a value at every viable information state: $V_k^{G,-}(x,p) = V_k^{G,+}(x,p) = V_k^G(x,p)$. The value is bounded and continuous in (x,p), convex and uniformly Lipschitz in p, and both players have optimal admissible Borel security strategies. Their induced trajectories are safe almost surely.*

4 Atomic Strategy Restriction

We introduce two information relaxation games where Nash strategies of the leaders are atomic. We then show that this pair of leader strategies has a vanishing exploitability in G as time grid refines. This inspires an algorithm that tractably solves G by restricting strategy spaces to be atomic.

4.1 Upper relaxation game $G^{\uparrow}$

At any (k,x,p) of $G^{\uparrow}$, P1 announces a safe control u, P2 observes u, Bayes updates the belief from p to some q according to (u,η_k), and chooses a best response v. Therefore (u,q) represents a game branch from (k,x,p). Let $V_{k+1}^{\uparrow}(F_k(x,u,v),q)$ be the continuation value, the branch value is

$$\widetilde{M}_k^{\uparrow}(x,q) := \min_{u\in\mathcal{U}_k^{\mathrm{safe}}(x)} \max_{v\in\mathcal{V}} \left\{ \sum_i q_i r_{i,k}(x,u,v) + V_{k+1}^{\uparrow}(F_k(x,u,v),q) \right\}. \tag{8}$$

Introduce a Bayes-plausible distribution ν of posterior belief branches, whose mean must equal the current belief p. Pairing each posterior with a minimizing control gives Bellman recursion of $G^{\uparrow}$:

$$V_k^{\uparrow}(x,p) = \min_{\nu\in\mathcal{P}(\Delta(\mathcal{I})):\int q\nu(\,\mathrm{d}q)=p} \int \widetilde{M}_k^{\uparrow}(x,q)\nu(\,\mathrm{d}q). \tag{9}$$

Theorem 2 (Upper atomicity)**.** *Under Assumptions 1 and 2, an optimal split in Equation* (9) *has at most I atoms: $\nu^* = \sum_{r=1}^{R} \lambda_r \delta_{p^r}$, $R \le I$, and $\sum_r \lambda_r p^r = p$. Each p^r is paired with an attained robust-safe control u^r. For $p_i > 0$, type i selects u^r with probability $\alpha_i^r = \lambda_r p_i^r / p_i$. Backward recursion gives an optimal P1 leader with at most I controls at every information state.*

The distribution ν is constrained only by total probability and the $I-1$ independent coordinates of its mean. An extreme optimal split therefore needs at most I posterior branches. Proof in Appendix B.

4.2 Lower relaxation game $G^{\downarrow}$ and Its Dual

Typewise security guarantees. In $G^{\downarrow}$, P2 announces its control v and P1 responds $u \in \mathcal{U}_k^{\mathrm{safe}}(x)$. For a P2 strategy ζ, define its typewise guarantee vector by

$$\sigma_i^{\zeta}(k,x) := \inf_{\eta_i\in\Sigma_{1,i}^{\downarrow}} J_{i,k,x}^{\downarrow}(\eta_i,\zeta), \qquad i = 1,\dots,I,$$

where $J_{i,k,x}^{\downarrow}(\eta_i,\zeta)$ is the type-i continuation payoff from (k,x). $\sigma_i^{\zeta}(k,x)$ is the payoff that ζ guarantees when the realized type is i. The prior-weighted payoff guaranteed by ζ is thus $\sum_{i=1}^{I} p_i \sigma_i^{\zeta}(k,x)$. Let $\mathcal{S}_k(x)$ denote the set of typewise guarantee vectors secured by P2. The continuation value of $G^{\downarrow}$ can be written as $V_k^{\downarrow}(x,p) = \max_{\sigma\in\mathcal{S}_k(x)} \langle p,\sigma\rangle$.

From guarantee vectors to the dual game. Introduce $\widehat{p}\in\mathbb{R}^I$ and the convex conjugate of $V_k^{\downarrow}$: $V_k^*(x,\widehat{p}) := \sup_{p\in\Delta(\mathcal{I})} \left\{ \langle p,\widehat{p}\rangle - V_k^{\downarrow}(x,p) \right\}$. Applying the minimax theorem gives

$$V_k^*(x,\widehat{p}) = \sup_{p\in\Delta(\mathcal{I})} \min_{\sigma\in\mathcal{S}_k(x)} \langle p,\widehat{p}-\sigma\rangle = \min_{\sigma\in\mathcal{S}_k(x)} \sup_{p\in\Delta(\mathcal{I})} \langle p,\widehat{p}-\sigma\rangle = \min_{\sigma\in\mathcal{S}_k(x)} \max_{i=1,\dots,I} \{\widehat{p}_i - \sigma_i\}. \tag{10}$$

Equation (10) gives the dual variable $\widehat{p}$ a direct interpretation: $\widehat{p}_i$ is a target payoff for type i, while $\widehat{p}_i - \sigma_i$ is the shortfall between that target and the guarantee actually secured by P2. $V_k^*(x,\widehat{p})$

thus represents the value of a dual game $G^{\downarrow,*}$ where P2 minimizes the worst-type shortfall. At the terminal, the guarantee vector is $\sigma(x) = \big(g_1(x), \dots, g_I(x)\big)$, thus $V_K^*(x, \widehat{p}) = \max_i\{\widehat{p}_i - g_i(x)\}$. The proceeding dual Bellman backup at $(k, x, \hat{p})$ has two steps: First, for a fixed continuation target $y \in \mathbb{R}^I$, P2 announces v and P1 chooses a safe response u. The current typewise loss vector shifts the next dual state from y to $y - r_k(x, u, v)$, where $r_k(x, u, v) := \big(r_{1,k}(x, u, v), \dots, r_{I,k}(x, u, v)\big)$ is the vector of stage-wise running losses. Thus the branch dual value is:

$$\widetilde{N}_k(x, y) := \min_{v \in \mathcal{V}} \max_{u \in \mathcal{U}_k^{\mathrm{safe}}(x)} V_{k+1}^*\big(F_k(x, u, v),\ y - r_k(x, u, v)\big). \tag{11}$$

Second, P2's behavioral strategy allows them to randomize over continuation targets while preserving the mean, similar to P1's belief splitting in $G^{\uparrow}$. Let $\mathcal{P}_1(\mathbb{R}^I)$ denote the probability measures on $\mathbb{R}^I$ with finite first moment, the dual backup can be written as

$$V_k^*(x, \widehat{p}) = \min_{\substack{\mu \in \mathcal{P}_1(\mathbb{R}^I) \\ \int y\, \mu(\,\mathrm{d}y) = \widehat{p}}} \int_{\mathbb{R}^I} \widetilde{N}_k(x, y)\, \mu(\,\mathrm{d}y). \tag{12}$$

For a convex function f on the simplex, write $\partial_\Delta f(p) := \{z : f(q) \geq f(p) + \langle z, q - p\rangle\ \forall q \in \Delta(\mathcal{I})\}$ for its subdifferential. Theorem 3 states the lower atomicity result, and critically, an equivalence result between P2's Nash strategies for $G^{\downarrow,*}$ and $G^{\downarrow}$. Proof in Appendix C.

Theorem 3 (Lower atomicity)**.** *Under Assumptions 1 and 2, every $(k, x, \widehat{p})$ admits an optimal P2 leader with at most I current-control atoms. If $\widehat{p} \in \partial_\Delta V_k^{\downarrow}(x, p)$, the same strategy attains $V_k^{\downarrow}(x, p)$.*

4.3 Exploitability certificate and its vanishing

Given a time grid Π, the previous section constructs an optimal P1 leader $\eta_\Pi^{\uparrow}$ in $G_\Pi^{\uparrow}$ and an optimal P2 leader $\zeta_\Pi^{\downarrow}$ in $G_\Pi^{\downarrow}$. Both leaders can be played unchanged in the simultaneous game G_Π, yet they need not form an equilibrium of G_Π. This section turns their value difference into an exploitability certificate, and proves that the certificate vanishes as $|\Pi| \to 0$.

Denote the stage-k viable slice of G_Π by $\mathcal{Q}_{\Pi,k}$. At $(x, p) \in \mathcal{Q}_{\Pi,k}$, define the *relaxation width*

$$\Delta_{\Pi,k}(x, p) := V_{\Pi,k}^{\uparrow}(x, p) - V_{\Pi,k}^{\downarrow}(x, p). \tag{13}$$

At fixed initial state $x_0 \in \mathcal{Q}_{\Pi,0}$, we abbreviate it by $\Delta_\Pi(x_0, p)$. We prove in Appendix D:

Proposition 1 (Information-relaxation certificate)**.** $\mathrm{Gap}_{0,x_0,p}^{G_\Pi}(\eta_{\Pi,x_0,p}^{\uparrow}, \zeta_{\Pi,x_0,p}^{\downarrow}) \leq \Delta_\Pi(x_0, p)$.

The response advantage in either relaxation acts for less time as $|\Pi| \to 0$, suggesting that Δ_Π should vanish. However, state constraints make this conclusion nontrivial: Near a viability boundary, even a short response advantage may change which controls are safe. To accommodate this, our vanishing theorem directly works with the continuous-time robust viability tube $\mathscr{D} := \bigcup_{t \in [0,T]} \mathcal{Q}(t)$, i.e., the continuous-time limit of $\bigcup_{k \in [K]} \mathcal{Q}_{\Pi,k}$ as $|\Pi| \to 0$. Our proof in Appendix E adapts Koike's geometric and game-theoretic assumptions (A3-A5 in Koike (1995)) for state-constrained games to accommodate the additional incomplete-info and dynamic nature of G. We prove in Appendix E:

Theorem 4 (Vanishing relaxation width on the exact tube)**.** *Under Assumptions 1 and 2, and an adaptation of Koike's assumptions detailed in Appendix E.1,*

$$\sup_{0 \leq k \leq K_\Pi} \sup_{(x,p) \in \mathcal{Q}(t_k) \times \Delta(\mathcal{I})} \Delta_{\Pi,k}(x, p) \longrightarrow 0 \qquad \text{as } |\Pi| \to 0. \tag{14}$$

Consequently, at every fixed $x_0 \in \mathcal{Q}(0)$, $V_{\Pi,0}^{G}(x_0, p)$ converge to the same continuous limit as the two relaxed values. For each fixed (x_0, p), let $(\eta_{\Pi,x_0,p}^{\uparrow}, \zeta_{\Pi,x_0,p}^{\downarrow})$ be leaders optimized at that information state. Their exploitability in G_Π converges to zero.

Empirical diagnostics of Assumption 2 and Koike's assumptions are deferred to Appendix F for the car chasing studies in Sec. 6.3.

5 SECURITY GAME SOLVERS AND BOUNDED ATOM INSERTION

The atomicity and vanishing-width results motivate solving G directly using I-atomic strategy spaces. We compute a P1 security strategy against P2 best responses and, symmetrically, a P2 security strategy against type-dependent P1 best responses.

5.1 SECURITY PROBLEMS

Let Π denote the leader time grid and Γ a finer response grid containing all leader time steps. The lift $\mathcal{L}_\Gamma$ embeds a leader strategy into Γ: at a response-only time step $k \in \Gamma \setminus \Pi$, the leader holds its current control and introduces no new primal signal or dual split, whereas the responder may update using the information then available. Let $\Sigma_{j,\Lambda}$ denote the admissible behavioral strategies of player j on grid Λ. The superscript I-atom restricts each current leader law to at most I control atoms. For P2 leader, $\Sigma_{2,\Pi}^{I\text{-atom}}$ uses the mean-preserving dual-state realization developed in Section 4. We solve

$$\overline{V}_{\Pi|\Gamma} := \min_{\eta \in \Sigma_{1,\Pi}^{I\text{-atom}}} \max_{\zeta \in \Sigma_{2,\Gamma}} J_\Gamma^G(\mathcal{L}_\Gamma \eta, \zeta), \qquad \underline{V}_{\Pi|\Gamma} := \max_{\zeta \in \Sigma_{2,\Pi}^{I\text{-atom}}} \min_{\eta \in \Sigma_{1,\Gamma}} J_\Gamma^G(\eta, \mathcal{L}_\Gamma \zeta). \tag{15}$$

P1's strategy class is type contingent, so the second inner minimization already represents a separate response for every realized type. Although $\underline{V}_{\Pi|\Gamma}$ uses P2's primal max-min security objective, its mean-preserving architecture contains the behavioral realization of an optimal lower-relaxation leader. Likewise, $\Sigma_{1,\Pi}^{I\text{-atom}}$ contains an optimal upper-relaxation leader. Consequently, when $\Gamma = \Pi$, $V_\Pi^G \le \overline{V}_{\Pi|\Pi} \le V_\Pi^\uparrow$, $V_\Pi^\downarrow \le \underline{V}_{\Pi|\Pi} \le V_\Pi^G$, and hence $0 \le \overline{V}_{\Pi|\Pi} - \underline{V}_{\Pi|\Pi} \le \Delta_\Pi$. Therefore $\overline{V}_{\Pi|\Pi} - \underline{V}_{\Pi|\Pi}$ vanishes under the assumptions of Theorem 4.

Asymmetric grids. Best-response optimization is substantially cheaper than optimizing a branching leader strategy. We therefore evaluate each leader against responses on a finer grid Γ. Relative to an audit on the leader grid alone, this increases the P1 upper-security value and decreases the P2 lower-security value. The resulting gap is thus a more stringent discretized exploitability estimate. To control the leader's cost, physical controls may be updated on a relatively fine subgrid while new primal signals or dual splits are permitted only on a sparser subgrid. This preserves fine physical feedback without creating a new information branch at every time step. The formal decomposition of Π into signaling and control-update grids, together with the finer verification grid is in Appendix J.1.

Constrained feedback solver. Shared feedback networks parameterize the two atomic leaders. At a signaling time step, P1 maps (t, x, p) to a typewise signal law (an I-by-M conditional probability matrix) and at most M public controls, where $M \le I$; P2 maps $(t, x, \widehat{p})$ to at most I weights, mean-preserving dual children, and controls. Bayes' rule and the dual barycenter constraint are enforced explicitly throughout the tree. The resulting problems are constrained and nonconvex-nonconcave. For training, we apply safeguarded doubly smoothed gradient descent-ascent (DS-GDA) (Zheng et al., 2023) to mitigate rotational cycling. Once the positive-mass atoms and active constraints stabilize, a damped active-set Newton step solves the local KKT system to accelerate convergence. Because this procedure targets a local stationary saddle rather than a global equilibrium, we alternate leader updates with independently initialized best responses, retain the most adversarial responses in a finite archive. Final values are audited using fresh best responses on the fine verification grid. Network architectures, solver tolerances, constraint handling, and audit protocols in Appendix J.

P2-owned state constraints. Type-independent P2 constraints are handled by restricting P2 to a nonempty, compact, continuous viable-control correspondence and requiring P1 safety against every admitted control. The finite-grid value existence and atomicity results are unchanged. The exploitability vanishing result requires repairability, boundary, and Isaacs conditions for the joint viability problem. See Appendix J.2. Constraints coupling P2's feasibility to P1's private type or current control are not considered in this paper.

5.2 ADAPTIVE ATOM INSERTION

Theorems 2, 3 and 4 provide a bounded variant of an adaptive atom insertion algorithm (McAleer et al., 2021; 2024) specifically for G, see Algorithm 1. Without loss of generality, we describe the algorithm using P2's lower security game: At each P2 information state $n = (k, x, \widehat{p})$, a router

Algorithm 1 Adaptive atom insertion for P2's lower security solve

Require: Configured width cap I, birth mass ε
1: Solve the width-one P2 lower-security problem and audit its P1 best responses.
2: **for** $M = 1, \ldots, I-1$ **do**
3: Train multistart deterministic P2 continuation experts against the frozen P1 upper-security actor and select the best candidate.
4: Add candidate to every existing tree node with probability ε. All nodes share $M+1$ atoms.
5: Reoptimize the enlarged P2 strategy and exchange fresh typewise P1 feedback best responses.
6: **end for**
7: **return** the retained P2 policy and its lower security value.

selects among M deterministic feedback experts, where $M \leq I$ is the current support size. Expert a supplies a control $v^a(t, x, z)$ and a continuation from its child dual state z^a. Its selector weight $\lambda_a(n)$ and child state satisfy

$$\lambda(n) \in \Delta(M), \qquad \sum_{a=1}^{M} \lambda_a(n) z^a(n) = \widehat{p}. \tag{16}$$

The expert is a complete continuation policy: it remains defined on all tree nodes created by later splits. For a fixed width, the lower security solve optimizes router weights, child states, and expert parameters against typewise P1 feedback responses. The adaptive atom insertion algorithm solves P2's security game starting from $M = 1$. In each game solve, multistart candidate experts are trained against the frozen stochastic P1 upper-security actor. The highest-payoff candidate is then added to the current P2 security strategy as an additional atom at every information state with a small conditional birth probability mass ε and a branched dual $\widehat{p}$ that preserves the barycenter in Equation (16). The existing branches retain their duals. The expanded tree with $M+1$ atoms at all nodes constructs the next P2 security game to be re-optimized.

We highlight two key differences of our method from existing methods such as NXPO and APSRO: (1) Our insertion process is upper bounded at $M = I$ due to the one-sided information structure of G; and (2) we use the dual recursion as a principled Markov state of P2, circumventing the learning of a P2 history encoder. In Section 6.2, we report the empirical advantage of dual recursion against history encoders, through a long-horizon game with analytical P2 Nash.

6 Experiments

This section presents four experiments, across which, we isolate how atomic strategies, dual variables, and constraints shape security learning. In the unconstrained Hexner game, the value gap vanishes analytically; experiments further verify atomicity for P1 and confirm that P2 requires only one atom. The constrained Hexner game demonstrates that the constraint changes P1's information strategy and induces deceptive play. The car-chasing game extends the study to a more complex setting and further demonstrates how constraints alter information strategies. Lastly, the long-horizon game is the only setting we can find in which the dual variable is essential for compactly encoding P2's history during security learning, since in the other three games, P2 only needs one atom and the dual variable remains constant, making dual a redundant information.

6.1 Hexner's game

We consider Hexner's linear-quadratic homing game (Hexner, 1979; Ghimire et al., 2026). Each player follows $\dot{x}_j = A_j x_j + B_j u_j$, while P1 privately observes the target $\theta \sim p_0$. With public target direction z, P1 minimizes and P2 maximizes

$$J = \mathbb{E}_\theta \left[\frac{1}{2} \int_0^T \left(u_1^\top R_1 u_1 - u_2^\top R_2 u_2 \right) \mathrm{d}t + \|x_1(T) - z\theta\|_{K_1}^2 - \|x_2(T) - z\theta\|_{K_2}^2 \right]. \tag{17}$$

The analytical equilibrium pools the types until a critical time and then reveals the target; P2 uses a deterministic tracking strategy. For the two-type reduced game, Appendix G proves $V_\Pi^\downarrow = V_{G_\Pi}$ and

$$0 \leq \Delta_{\mathrm{LF}}(\Pi) \leq 4 p_0 (1 - p_0) \|d_2\|_\infty |\Pi| = O(|\Pi|).$$

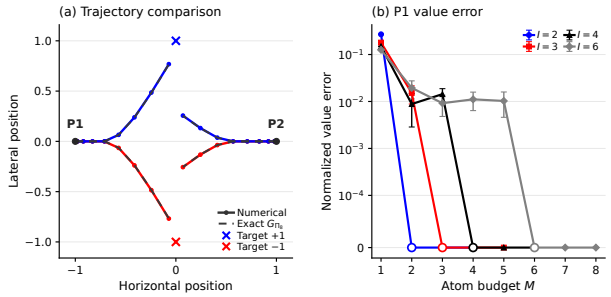


Figure 1: (a) Trajectory comparison for G_{Π_8} and $I = 2$. P1 approaches a private target (cross) and P2 follows. Solid curves: One-atom P2 policy and its P1 best responses. Dashed curves: GT equilibrium. (b) Normalized P1 upper-value error on G_{Π_3}, shown as mean $\pm$ 1 std over 20 seeds per atom budget. Open rings mark $M = I$.

We first test the I-atom upper-leader bound. We use $T = 1$, a uniform three-stage grid, equally weighted scalar types spaced uniformly on $[-1, 1]$, and $I \in \{2, 3, 4, 6\}$. P1 is allocated $M = 1, \ldots, I + 2$ control slots at each information state. We perform 20 independent upper-security solves for each (I, M) and audit every policy against independently optimized P2 responses. As shown in Figure 1(b), all 240 runs with $M \geq I$ recover the analytical upper value within 10^{-4} normalized error, whereas none of the 220 runs with $M < I$ does. Thus the theoretical budget I is sufficient in this benchmark, and allocating additional slots does not improve the recovered value.

We next test P2 insertion directly on G_{Π_8} with $I = 2$. P1 best responses are solved as finite-dimensional quadratic programs. For $M = 1$, the deterministic P2 feedback strategy attains a value error of 4.20×10^{-7}. We then insert a second atom with initial mass 0.05 and jointly optimize both atoms and their router. The resulting model improves the audited value by only 6.40×10^{-9}. This recovers the analytical fact that Hexner's P2 equilibrium is deterministic and shows that the insertion procedure rejects an unnecessary atom. Against the resulting one-atom P2 policy, P1 responses closely match the GT equilibrium trajectories: the position RMSE is 7.83×10^{-4} for P1 and 4.73×10^{-4} for P2 (Figure 1(a)). Full details are reported in Appendix H.1.

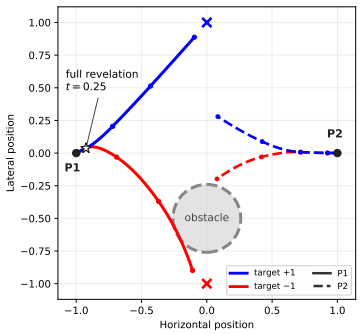


Figure 2: Constrained Hexner interaction. Solid and dashed curves show P1 upper-security and P2 lower-security strategies.

State constraints change information revelation. We add a circular obstacle to Hexner that blocks P1's unconstrained route. We solve P1 upper-security and $M = 1$ P2 lower-security problems on G_{Π_4}. As shown in Figure 2, P1 security strategy deceives at $t = 0$ by moving upward for both types and reveals at $t = 0.25s$. Thus state constraint affects P1's information strategy in addition to changing its trajectory.

6.2 Dual state versus public history for P2

From the discussion on the lower relaxation dual game $G^{\downarrow,*}$, we see that the dual variable $\widehat{p}$ is a Markov state in $G^{\downarrow,*}$, i.e., unlike standard IIEFG, P2's behavioral strategy in $G^{\downarrow}$ can be defined on $(k, x, \widehat{p})$ without full history, as long as $\widehat{p}$ is initialized in $\partial_{\Delta} V_0^{\downarrow}(x, p)$. Here we test if $\widehat{p}$ is indeed a compact sufficient memory for P2 in G. To this end, we introduce a constant-state game as follows: Before starting, nature draws P1's type $\theta \in \{-1, +1\}$ uniformly. The informed player P1 chooses and announces a sequence of L public signals $s_1, \ldots, s_L \in [-1, 1]$, after which P2 chooses $b \in [-1, 1]$. P1 minimizes and P2 maximizes

$$J_\theta(s_{1:L}, b) = \frac{1.25}{L} \sum_{\ell=1}^{L} (s_\ell - \theta)^2 - (b - \theta)^2. \tag{18}$$

The public physical state is constant, so it cannot distinguish histories. Nevertheless, P2's equilibrium action varies with the history through its mean $\bar{s} = L^{-1} \sum_\ell s_\ell$. The dual update $\widehat{p} = -5\bar{s}$ retains this information in one scalar: $b^\star = \mathrm{clip}(-\widehat{p}/4, -1, 1)$. The game has value zero. We solve P2's security game using three representations: $\widehat{p}$, the full P2 history, and a GRU that reads the same history sequentially. The learned policies had comparable numbers of parameters and were trained against adversarially optimized P1 histories under the same maximum optimization budget. Thus the dual model was given the correct summary, while the two history models had to discover

that summary and the action rule jointly. **Results.** All three informed representations were reliable through $L = 32$. At $L = 64$, the dual model passed the frozen security audit in 8s, while neither raw history or GRU passed after 25s and 141s, respectively. See Appendix I for details.

6.3 Constrained car chasing

Game settings. We consider a constrained pursuit-evasion game at a three-way intersection. P1 privately knows its intended route $\theta \in \{\text{north}, \text{left}, \text{right}\}$, with uniform prior, while P2 observes only the public vehicle state. Both players follow forward-only kinematic-bicycle dynamics, but their maneuvering capabilities are asymmetric: P1 has acceleration bound $|a_1| \le 8\,\text{m/s}^2$, drag coefficient 0.8, and curvature bound $|\kappa_1| \le 0.22\,\text{m}^{-1}$, while P2 has the larger acceleration bound $|a_2| \le 10\,\text{m/s}^2$ and smaller drag coefficient 0.7, but the much tighter curvature bound $|\kappa_2| \le 0.06\,\text{m}^{-1}$; both have maximum speed $10\,\text{m/s}$. Initially, P1 is at $(0, -9)$ m and Player 2 at $(0, -21)$ m, both facing north with speed 6 m/s. The interaction lasts 3 s, divided into six 0.5 s information stages and twelve 0.25 s control intervals. It is followed by a 6 s P2-only continuation used solely to certify whether P2's terminal state allows it to stay in lane in the future. Road, speed, control, collision, and exit constraints are enforced throughout. This dynamic asymmetry gives the candidate saddle point its characteristic timing structure: P1 can conceal its type by driving identically across types and reveal only near a critical decision time, while the less maneuverable P2 must decelerate before the reveal to preserve the option of completing a late turn, and then accelerate again after the revealed route determines the appropriate exit. See Figure 3.

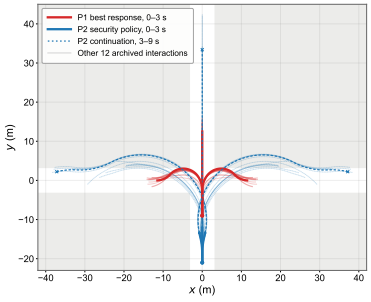


Figure 3: P2 security against archived P1 best responses. By delaying type revelation, P1 (red) forces P2 (blue) to move conservatively and preserve sufficient maneuvering room for any subsequent turn. Solid curves show the 0–3 s interaction; dotted blue curves show P2's 3–9 s continuation and expose the difficulty of restoring long-horizon safety under adversarial P1. Left/right symmetry enforced during P2 security solve.

Challenges in security solve and simplification. Even with atomic strategy restriction, the game tree grows rapidly because $I = 3$, resulting in minimax problems with a significant amount of hard constraints (e.g., about 25k constraints for P1) that keep on increasing as the opponent's best response archive expands (about 1.2k per best response). In addition, nonlinear dynamics, collision avoidance, road geometry make the problems highly nonconvex-nonconcave. We therefore restrict P1 to three physically meaningful routes-north, left, and right-with shared early actions. We use one P2 feedback strategy with six time-specific neural networks, shared actions whenever public information is identical, and speed-bounded outputs, in order to reduce P2 constraint violations due to adversarial P1 attacks. We sample safety at 1/8-second intervals while retaining the finer physical integration. We polish small network weight blocks during security optimization in order to manage all archive-induced constraints.

7 Conclusion

This paper considers a 2p0s dynamic game with simultaneous continuous controls, state constraints, and one-sided payoff information. We proved that the game attains a behavioral value under a repairability assumption. We then showed that in the upper and lower information-relaxation games, the optimal leaders use at most I control atoms, and that their embedded pair has vanishing exploitability in the original game. These results justified tractable computation of security strategies of both players within bounded strategy spaces. Empirical studies verify the use of restricted strategy spaces (unconstrained Hexner) and dual variable for history encoding (constant-state game), and the influence of state constraints to player strategies (constrained Hexner and car chasing).

Limitations. So far we have not identified a game where P2 requires more than one feedback atom, suggesting that under the studied games, the dual variables stay in a region where the dual value is convex. Sufficient game conditions for this property would be of interest, as it would guarantee a simple structure of P2's lower-security strategy. In addition, our ability at solving real-world G with hard state constraints, such as car chasing, is still limited, due to the exponentially growing tree complexity (along with constraints and decision variables) with respect to time grid size, and the nonconvexity of the minimax due to nonlinear dynamics and collision avoidance.

AI USE STATEMENT

GPT 5.6 Sol and Astra were used to assist theorem development and algorithm implementation. The authors independently reviewed the resulting text and take responsibility for all mathematical statements, proofs, citations, code, and experimental claims.

REPRODUCIBILITY STATEMENT

Complete assumptions and proofs are provided in the appendix. The anonymous repository contains the implementations, frozen configurations, registered experimental protocols, random seeds, archived response audits, and scripts needed to regenerate the reported tables and figures.

ETHICS STATEMENT

As increasingly capable AI agents act strategically under private objectives, asymmetric information, and physical constraints, principled defense strategies against game-theoretic AI are needed to anticipate manipulation and maintain safety against adaptive opponents. This work develops mathematical and computational tools for that *defensive purpose*: evaluating worst-case behavior and constructing safe security strategies in simulated zero-sum interactions. The same game-solving techniques could be repurposed for adversarial planning. We therefore restrict our study to abstract and simulated environments, and recommend domain-specific validation, human oversight, and applicable legal and safety review before deployment in consequential settings.

REFERENCES

Lukáš Adam, Rostislav Horčík, Tomáš Kasl, and Tomáš Kroupa. Double oracle algorithm for computing equilibria in continuous games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 5070–5077, 2021. doi: 10.1609/aaai.v35i6.16641.

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, 3 edition, 2006.

Robert J. Aumann. Mixed and behavior strategies in infinite extensive games. In Melvin Dresher, Lloyd S. Shapley, and Albert W. Tucker (eds.), *Advances in Game Theory*, volume 52 of *Annals of Mathematics Studies*, pp. 627–650. Princeton University Press, 1964.

Robert J. Aumann, Michael Maschler, and Richard E. Stearns. *Repeated Games with Incomplete Information*. MIT Press, 1995.

Martino Bardi, Shigeaki Koike, and Pierpaolo Soravia. Pursuit–evasion games with state constraints: Dynamic programming and discrete-time approximations. *Discrete and Continuous Dynamical Systems*, 6(2):361–380, 2000.

Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press, 1978.

Piernicola Bettiol, Marc Quincampoix, and Richard B. Vinter. Existence and characterization of the values of two player differential games with state constraints. *Applied Mathematics & Optimization*, 80:765–799, 2019. doi: 10.1007/s00245-019-09608-8.

Patrick Billingsley. *Convergence of Probability Measures*. Wiley, 2 edition, 1999.

Noam Brown and Tuomas Sandholm. Simultaneous abstraction and equilibrium finding in games. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, pp. 489–496, 2015. URL `https://www.ijcai.org/Abstract/15/075`.

Pierre Cardaliaguet. Differential games with asymmetric information. *SIAM Journal on Control and Optimization*, 46(3):816–838, 2007. doi: 10.1137/060654396.

Pierre Cardaliaguet. Numerical approximation and optimal strategies for differential games with lack of information on one side. In *Advances in Dynamic Games and Their Applications*, volume 10 of *Annals of the International Society of Dynamic Games*, pp. 159–176. Birkhäuser, 2009.

Bernard De Meyer. Repeated games, duality and the central limit theorem. *Mathematics of Operations Research*, 21(1):237–251, 1996. doi: 10.1287/moor.21.1.237.

Asen L. Dontchev and R. Tyrrell Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis*. Springer, 2 edition, 2014.

Hélène Frankowska and Richard B. Vinter. Existence of neighboring feasible trajectories: Applications to dynamic programming for state-constrained optimal control problems. *Journal of Optimization Theory and Applications*, 104:20–40, 2000.

Mukesh Ghimire, Lei Zhang, Zhe Xu, and Yi Ren. State-constrained zero-sum differential games with one-sided information. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 15512–15539. PMLR, 2024. URL https://proceedings.mlr.press/v235/ghimire24a.html.

Mukesh Ghimire, Lei Zhang, Zhe Xu, and Yi Ren. Solving football by exploiting equilibrium structure of 2p0s differential games with one-sided information. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=vRwuBOxbsJ.

John C. Harsanyi. Games with incomplete information played by "Bayesian" players, I–III: Part I. the basic model. *Management Science*, 14(3):159–182, 1967. doi: 10.1287/mnsc.14.3.159.

J. Hawkin, R. Holte, and D. Szafron. Automated action abstraction of imperfect information extensive-form games. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, pp. 681–687, 2011.

G. Hexner. A differential game of incomplete information. *Journal of Optimization Theory and Applications*, 28(2):213–232, 1979. doi: 10.1007/BF00933243.

Shigeaki Koike. On the state constraint problem for differential games. *Indiana University Mathematics Journal*, 44(2):467–487, 1995.

Viliam Kovařík, Martin Schmid, Neil Burch, Michael Bowling, and Viliam Lisý. Rethinking formal models of partially observable multiagent decision making. *Artificial Intelligence*, 303:103645, 2022. doi: 10.1016/j.artint.2021.103645.

Christian Kroer and Tuomas Sandholm. A unified framework for extensive-form game abstraction with bounds. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/aa942ab2bfa6ebda4840e7360ce6e7ef-Abstract.html.

Harold W. Kuhn. Extensive games and the problem of information. In Harold W. Kuhn and Albert W. Tucker (eds.), *Contributions to the Theory of Games II*, volume 28 of *Annals of Mathematics Studies*, pp. 193–216. Princeton University Press, 1953.

Kazimierz Kuratowski and Czesław Ryll-Nardzewski. A general theorem on selectors. *Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques, Astronomiques et Physiques*, 13:397–403, 1965.

Chung-Wei Lee, Christian Kroer, and Haipeng Luo. Last-iterate convergence in extensive-form games. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/77bb14f6132ea06dea456584b7d5581e-Abstract.html.

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*, volume 235 of *Proceedings of Machine Learning Research*, pp. 27752–27770. PMLR, 2024. URL https://proceedings.mlr.press/v235/li24t.html.

Mingyang Liu, Asuman E. Ozdaglar, Tiancheng Yu, and Kaiqing Zhang. The power of regularization in solving extensive-form games. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2206.09495.

Stephen McAleer, JB Lanier, Kevin A. Wang, Pierre Baldi, and Roy Fox. XDO: A double oracle algorithm for extensive-form games. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/c2e06e9a80370952f6ec5463c77cbace-Abstract.html.

Stephen McAleer, JB Lanier, Kevin A. Wang, Pierre Baldi, Tuomas Sandholm, and Roy Fox. Toward optimal policy population growth in two-player zero-sum games. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=J2TZgj3Tac.

H. Brendan McMahan, Geoffrey J. Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *Proceedings of the Twentieth International Conference on Machine Learning*, pp. 536–543, 2003.

Linjian Meng, Tianpei Yang, Youzhi Zhang, Zhenxing Ge, Shangdong Yang, Tianyu Ding, Wenbin Li, Bo An, and Yang Gao. Efficient last-iterate convergence in solving extensive-form games. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-3367. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/918c67cd7910578f1637beb56717f455-Abstract-Conference.html.

Ian M. Mitchell, Alexandre M. Bayen, and Claire J. Tomlin. A time-dependent Hamilton–Jacobi formulation of reachable sets for continuous dynamic games. *IEEE Transactions on Automatic Control*, 50(7):947–957, 2005. doi: 10.1109/TAC.2005.851439.

Tuomas Sandholm. Abstraction for solving large incomplete-information games. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pp. 4127–4131, 2015. doi: 10.1609/aaai.v29i1.9757.

Canzhe Zhao, Yutian Cheng, Jing Dong, Baoxiang Wang, and Shuai Li. Learning imperfect information extensive-form games with last-iterate convergence under bandit feedback. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 77405–77440. PMLR, 2025. URL https://proceedings.mlr.press/v267/zhao25h.html.

Taoli Zheng, Linglingzhi Zhu, Anthony Man-Cho So, José Blanchet, and Jiajin Li. Universal gradient descent ascent method for nonconvex-nonconcave minimax optimization. In *Advances in Neural Information Processing Systems*, volume 36, pp. 54075–54110, 2023.

Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. In *Advances in Neural Information Processing Systems*, volume 20, 2007.

APPENDIX CONTENTS

1. Appendix A: Proof of Value Existence of G (Theorem 1)
2. Appendix B: Proof of Upper Primal Atomicity (Theorem 2)
3. Appendix C: Proof of Lower Dual Atomicity (Theorem 3)
4. Appendix D: Information-Relaxation Certificates (Proposition 1)
5. Appendix E: State-Constrained Common-Limit Theorem (Theorem 4)
6. Appendix F: Empirical Koike Diagnostics
7. Appendix G: Exact Hexner Analysis
8. Appendix H: Numerical Hexner Study
9. Appendix I: Long-Horizon Dual Compression
10. Appendix J: Feedback-Security Solvers

A PROOF OF VALUE EXISTENCE OF G

Without hard state constraints, value existence and primal-dual dynamic programming for zero-sum games with one-sided information are classical (Aumann et al., 1995; Cardaliaguet, 2007). The additional difficulty here is state-dependent feasibility: the robust safe-control set $\mathcal{U}_k^{\mathrm{safe}}(x)$ may lose lower hemicontinuity at the viability boundary, preventing a direct application of Berge's theorem and measurable selection.

The proof proceeds in three steps. Section A.1 records the standard selection, disintegration, and behavioral-realization results used to turn pointwise Bellman optimizers into Borel strategies. Section A.2 contains the state-constraint-specific work: the closed viable recursion gives compact viable sets and a closed-graph safe-control correspondence; continuity of the safety residual together with Assumption 2 upgrades this correspondence to Hausdorff continuity; and the resulting admissible strategies are shown to be pathwise safe. With this regularity in hand, Section A.3 constructs the primal posterior-control Bellman recursion, proves exact implementability and P1 security, and then derives a Fenchel-dual recursion yielding P2 security. Fenchel equality matches the two security bounds and proves the value, existence of Borel optimal strategies, and almost-sure safety.

A.1 PREREQUISITES

The following three existing lemmas collect the measurable-selection tools used in this appendix: Lemma 1 provides Borel choices of stage-wise optimizers, Lemma 2 converts selected joint laws into measurable conditional kernels, and Lemma 3 converts the dual construction with a latent private state into an outcome-equivalent behavioral strategy.

Lemma 1 (Borel selection of optimizers (Kuratowski & Ryll-Nardzewski, 1965; Aliprantis & Border, 2006))**.** *Let S be a standard Borel space and A a Polish space. Suppose the feasible-set correspondence $\Gamma : S \rightrightarrows A$ has a Borel graph and nonempty compact values, and $f : S \times A \to \mathbb{R}$ is Borel with $f(s, \cdot)$ continuous. Then, for every $s \in S$, the minimum $\min_{a\in\Gamma(s)} f(s, a)$ is attained, and an optimizer $a^*(s)$ can be selected so that $s \mapsto a^*(s)$ is Borel measurable.*

Lemma 2 (Measurable conditional laws (Bertsekas & Shreve, 1978))**.** *Let S, V, Y be standard Borel spaces and let $s \mapsto \lambda_s \in \mathcal{P}(V \times Y)$ be a Borel probability kernel. Write $\beta_s = (\lambda_s)_V$ for its V-marginal. Then there is a jointly Borel conditional distribution $\kappa(\,\mathrm{d}y \mid s, v)$ such that, for all Borel sets $B \subseteq V$ and $C \subseteq Y$,*

$$\lambda_s(B \times C) = \int_B \kappa(C \mid s, v)\, \beta_s(\,\mathrm{d}v).$$

Lemma 3 (Behavioral realization (Kuhn, 1953; Aumann, 1964))**.** *In a finite-horizon perfect-recall game with standard-Borel histories and actions, every strategy generated by private randomization and Borel contingent decision rules is outcome-equivalent to a Borel behavioral strategy, i.e., the two strategies induce the same law of public histories and payoffs against every opponent strategy.*

All probability-measure spaces below carry the weak topology.

A.2 Trajectory Safety, Repairability, and Measurable Selection

Lemma 4 (Closed viable recursion). *Suppose $\mathcal{X}$, $\mathcal{U}$, and $\mathcal{V}$ are compact, $\mathcal{C} \subseteq \mathcal{X}$ is closed, and F_k and $\Gamma_{j,k}$ are continuous. Then every $\mathcal{Q}_k$ is compact, and $\mathcal{U}_k^{\mathrm{safe}}$ has a closed graph and nonempty compact values on $\mathcal{Q}_k$.*

Proof. We argue backward in k. The terminal viable set $\mathcal{Q}_K = \mathcal{C}$ is compact. Suppose that $\mathcal{Q}_{k+1}$ is compact and define

$$\mathcal{R}_k := \left\{ (x,u) : \begin{array}{ll} \Gamma_{j,k}(x,u,v) \le 0 & \forall j,\ v \in \mathcal{V}, \\ F_k(x,u,v) \in \mathcal{Q}_{k+1} & \forall v \in \mathcal{V} \end{array} \right\}.$$

Let $(x_n, u_n) \to (x,u)$ with $(x_n,u_n) \in \mathcal{R}_k$. For every fixed $v \in \mathcal{V}$, continuity gives

$$\Gamma_{j,k}(x,u,v) = \lim_{n\to\infty} \Gamma_{j,k}(x_n,u_n,v) \le 0 \qquad \forall j,$$

and

$$F_k(x,u,v) = \lim_{n\to\infty} F_k(x_n,u_n,v) \in \mathcal{Q}_{k+1},$$

where the last inclusion follows because $\mathcal{Q}_{k+1}$ is closed. Since v was arbitrary, $(x,u) \in \mathcal{R}_k$. Thus $\mathcal{R}_k$ is closed and, as a subset of the compact space $\mathcal{X} \times \mathcal{U}$, compact.

The projection of $\mathcal{R}_k$ onto $\mathcal{X}$ is precisely $\mathcal{Q}_k$, so $\mathcal{Q}_k$ is compact. For every $x \in \mathcal{Q}_k$, the fiber of $\mathcal{R}_k$ is $\mathcal{U}_k^{\mathrm{safe}}(x)$. This fiber is nonempty by the definition of $\mathcal{Q}_k$ and compact because it is a closed subset of $\mathcal{U}$. □

Corollary 1 (Pathwise safety). *If $x_0 \in \mathcal{Q}_0$, every admissible P1 strategy in G, $G^{\uparrow}$, or the robust-safe $G^{\downarrow}$ keeps the induced trajectory in $\mathcal{C}$ almost surely against every admissible P2 strategy. More strongly, every realization whose P1 controls belong to the supports of the admissible kernels is safe.*

Proof. We use induction over the stages. Suppose that the realized state satisfies $x_k \in \mathcal{Q}_k$. Admissibility gives

$$\eta_{i,k}\big(\mathcal{U}_k^{\mathrm{safe}}(x_k) \mid x_k, p_k\big) = 1 \qquad \forall i.$$

Hence $u_k \in \mathcal{U}_k^{\mathrm{safe}}(x_k)$ almost surely. For every possible $v_k \in \mathcal{V}$, the definition in Equation (2) then gives

$$c_j(X_k(s; x_k, u_k, v_k)) \le 0 \quad \forall j,\ s \in [0, h_k],$$

and

$$x_{k+1} = F_k(x_k, u_k, v_k) \in \mathcal{Q}_{k+1}.$$

Thus the current trajectory segment remains in $\mathcal{C}$, and the induction continues at the next stage. Starting from $x_0 \in \mathcal{Q}_0$ proves almost-sure safety through time T.

For the stronger support statement, Lemma 4 implies that $\mathcal{U}_k^{\mathrm{safe}}(x_k)$ is closed. A probability measure that assigns a closed set probability one has support contained in that set. The same induction therefore applies to every realization whose controls belong to the supports of the kernels. □

Lemma 5 (Continuity of the violation residual). *Under Assumption 1, Φ_k^{safe} is continuous on $\mathcal{Q}_k \times \mathcal{U}$ and therefore uniformly continuous.*

Proof. Uniform state-Lipschitz continuity of the dynamics gives continuous dependence of $X_k(s; x, u, v)$ on (s, x, u, v) over the compact stage interval. Consequently,

$$F_k(x,u,v) = X_k(h_k; x, u, v)$$

is continuous. Since each c_j is continuous, the function

$$(s, x, u, v) \longmapsto c_j(X_k(s; x, u, v))$$

is jointly continuous. Maximizing it over the compact interval $[0, h_k]$ shows that $\Gamma_{j,k}(x,u,v)$ is continuous.

Because $\mathcal{Q}_{k+1}$ is nonempty and closed, the map

$$y \longmapsto \mathrm{dist}(y, \mathcal{Q}_{k+1})$$

is 1-Lipschitz. Therefore the functions maximized over v in Equation (6) are jointly continuous in (x, u, v). Compactness of $\mathcal{V}$, the maximum theorem, and finiteness of the constraint index j imply continuity of Φ_k^{safe} in (x, u). Finally, $\mathcal{Q}_k \times \mathcal{U}$ is compact by Lemma 4, so this continuity is uniform. □

For two nonempty compact sets $A, B \subseteq \mathcal{U}$, define their directed excess and Hausdorff distance by

$$e(A, B) := \sup_{a \in A} \operatorname{dist}(a, B), \qquad d_H(A, B) := \max\{e(A, B), e(B, A)\}.$$

Thus $d_H(A_n, A) \to 0$ means both that every point of A_n lies near some point of A and that every point of A can be approximated by points of A_n.

Now we state and prove the safe-set regularity result necessary for value existence.

Lemma 6 (Safe-set regularity)**.** *Under Assumption 1, $\mathcal{U}_k^{\text{safe}}(x)$ is nonempty and compact for $x \in \mathcal{Q}_k$. Under Assumption 2, $\mathcal{U}_k^{\text{safe}}(x)$ varies continuously with x in Hausdorff distance.*

Proof of Lemma 6. Nonempty compact values and a closed graph follow from Lemma 4. Because $\mathcal{U}$ is compact, the closed-graph property also implies upper hemicontinuity.

It remains to show that every safe control can be approximated from nearby states. By Lemma 5, there is a modulus ω_k, with $\omega_k(r) \to 0$ as $r \downarrow 0$, such that

$$\left|\Phi_k^{\text{safe}}(x, u) - \Phi_k^{\text{safe}}(x', u)\right| \le \omega_k(d_{\mathcal{X}}(x, x'))$$

uniformly over $x, x' \in \mathcal{Q}_k$ and $u \in \mathcal{U}$. If $u \in \mathcal{U}_k^{\text{safe}}(x)$, then $\Phi_k^{\text{safe}}(x, u) = 0$, so

$$\Phi_k^{\text{safe}}(x', u) \le \omega_k(d_{\mathcal{X}}(x, x')).$$

For x' sufficiently close to x, the right-hand side is at most ε_k. Applying Assumption 2 at x' gives

$$\operatorname{dist}\big(u, \mathcal{U}_k^{\text{safe}}(x')\big) \le \kappa_k \big[\omega_k(d_{\mathcal{X}}(x, x'))\big]^{\theta_k}.$$

Taking the supremum over $u \in \mathcal{U}_k^{\text{safe}}(x)$ bounds

$$e\big(\mathcal{U}_k^{\text{safe}}(x), \mathcal{U}_k^{\text{safe}}(x')\big).$$

Repeating the argument with x and x' interchanged bounds the reverse excess. Therefore

$$d_H\big(\mathcal{U}_k^{\text{safe}}(x), \mathcal{U}_k^{\text{safe}}(x')\big) \le \kappa_k \big[\omega_k(d_{\mathcal{X}}(x, x'))\big]^{\theta_k}$$

for all sufficiently close $x, x' \in \mathcal{Q}_k$. Hence $x \mapsto \mathcal{U}_k^{\text{safe}}(x)$ is continuous in Hausdorff distance and, in particular, both upper and lower hemicontinuous. □

A.3 Proof of Theorem 1

Let $\mathcal{P}(A)$ denote the set of Borel probability measures on A, and write $r_k(x, u, v) := \big(r_{1,k}(x, u, v), \dots, r_{I,k}(x, u, v)\big) \in \mathbb{R}^I$. After P1's current control u becomes public, it may reveal information about P1's type and change the public belief from p to a posterior q. To write the Bellman recursion compactly, we represent P1's current prescription by a joint distribution ρ over the safe control u and the resulting posterior q:

$$\mathfrak{R}_k(x, p) := \left\{ \rho \in \mathcal{P}\big(\mathcal{U}_k^{\text{safe}}(x) \times \Delta(\mathcal{I})\big) : \int q \, \rho(\mathrm{d}u, \mathrm{d}q) = p \right\}. \tag{19}$$

The barycenter constraint is Bayes plausibility: the expected posterior must equal the current belief. For a bounded continuous continuation function $Z : \mathcal{Q}_{k+1} \times \Delta(\mathcal{I}) \to \mathbb{R}$, define the cost of a public branch (u, q, v) by

$$C_k^Z(x, u, q, v) := \langle q, r_k(x, u, v) \rangle + Z(F_k(x, u, v), q).$$

Let ρ and β be P1's and P2's current mixed control, respectively. The Bellman operator for G is

$$(\mathcal{T}_k^G Z)(x, p) := \min_{\rho \in \mathfrak{R}_k(x,p)} \max_{\beta \in \mathcal{P}(\mathcal{V})} \int_{\mathcal{U}_k^{\text{safe}}(x) \times \Delta(\mathcal{I})} \int_{\mathcal{V}} C_k^Z(x, u, q, v) \, \beta(\mathrm{d}v) \rho(\mathrm{d}u, \mathrm{d}q). \tag{20}$$

Starting from the expected terminal loss, define backward

$$V_K^G(x, p) := \sum_{i=1}^{I} p_i g_i(x), \qquad V_k^G := \mathcal{T}_k^G V_{k+1}^G. \tag{21}$$

Lemma 7 (Regularity of current prescription sets)**.** *For every (k, x, p), $\mathfrak{R}_k(x, p)$ is nonempty, compact, and convex, and the correspondence*

$$(x, p) \longmapsto \mathfrak{R}_k(x, p)$$

is continuous under the weak topology. If Z is bounded and continuous, then

$$L_k^Z(x, \rho, \beta) := \int_{\mathcal{U}_k^{\mathrm{safe}}(x)\times\Delta(\mathcal{I})} \int_{\mathcal{V}} C_k^Z(x, u, q, v)\, \beta(\,\mathrm{d}v)\rho(\,\mathrm{d}u,\,\mathrm{d}q)$$

is continuous and affine in each measure. Consequently, Equation (20) *has an attained saddle value, and $\mathcal{T}_k^G Z$ is bounded and continuous in (x, p) and convex in p.*

Proof. Write $A_k(x) := \mathcal{U}_k^{\mathrm{safe}}(x)$.

Nonemptiness, compactness, and convexity. For any $u \in A_k(x)$, the measure $\delta_{(u,p)}$ belongs to $\mathfrak{R}_k(x, p)$, proving nonemptiness. Because $A_k(x) \times \Delta(\mathcal{I})$ is compact, its probability-measure space is compact under weak convergence. The barycenter map

$$\rho \longmapsto \int q\, \rho(\,\mathrm{d}u,\,\mathrm{d}q)$$

is continuous and affine. Therefore $\mathfrak{R}_k(x, p)$ is a closed, convex subset of a compact probability-measure space.

Upper hemicontinuity. Let $(x_n, p_n) \to (x, p)$ and suppose

$$\rho_n \in \mathfrak{R}_k(x_n, p_n), \qquad \rho_n \Rightarrow \rho.$$

Hausdorff continuity of A_k implies that every weak limit of measures supported on $A_k(x_n) \times \Delta(\mathcal{I})$ is supported on $A_k(x) \times \Delta(\mathcal{I})$. Moreover, continuity of the barycenter map gives

$$\int q\, \rho(\,\mathrm{d}u,\,\mathrm{d}q) = \lim_{n\to\infty} \int q\, \rho_n(\,\mathrm{d}u,\,\mathrm{d}q) = \lim_{n\to\infty} p_n = p.$$

Hence $\rho \in \mathfrak{R}_k(x, p)$. Thus the prescription correspondence has a closed graph and, because its values lie in a compact ambient probability-measure space, is upper hemicontinuous.

Lower hemicontinuity. Let $(x_n, p_n) \to (x, p)$ and fix $\rho \in \mathfrak{R}_k(x, p)$. We construct $\rho_n \in \mathfrak{R}_k(x_n, p_n)$ such that $\rho_n \Rightarrow \rho$.

First transport the control coordinate from $A_k(x)$ to $A_k(x_n)$. Hausdorff continuity provides a Borel map

$$T_n : A_k(x) \to A_k(x_n)$$

such that

$$d_{\mathcal{U}}(T_n(u), u) \le d_H(A_k(x), A_k(x_n)) + \frac{1}{n}.$$

Let (U, Q) be a random pair with distribution ρ. Replace its control coordinate by the nearby feasible control

$$U_n := T_n(U),$$

while leaving its posterior coordinate unchanged, and let $\widetilde{\rho}_n$ denote the distribution of (U_n, Q). In other words, $\widetilde{\rho}_n$ assigns the same probability to each posterior-control branch as ρ, but moves its control from u to $T_n(u) \in A_k(x_n)$. Only the control coordinate is changed, so

$$\int q\, \widetilde{\rho}_n(\,\mathrm{d}u,\,\mathrm{d}q) = p.$$

Moreover, $T_n(u)$ converges uniformly to u because $d_H(A_k(x), A_k(x_n)) \to 0$. It follows that $\widetilde{\rho}_n \Rightarrow \rho$.

It remains to change the barycenter from p to p_n. If $p_n = p$, set $\rho_n := \widetilde{\rho}_n$. Otherwise, for all sufficiently large n, define

$$\epsilon_n := \max_{i:p_i>0} \frac{(p_i - p_{n,i})_+}{p_i}, \qquad r_n := \frac{p_n - (1-\epsilon_n)p}{\epsilon_n}.$$

Since $p_n \to p$, we have $\epsilon_n \downarrow 0$. The definition of ϵ_n ensures that every coordinate of $p_n - (1-\epsilon_n)p$ is nonnegative. Its coordinates sum to ϵ_n, so $r_n \in \Delta(\mathcal{I})$ and

$$p_n = (1-\epsilon_n)p + \epsilon_n r_n.$$

Choose any $u_n \in A_k(x_n)$ and set

$$\rho_n := (1-\epsilon_n)\widetilde{\rho}_n + \epsilon_n \delta_{(u_n, r_n)}.$$

Then ρ_n is supported on $A_k(x_n) \times \Delta(\mathcal{I})$, and its barycenter is

$$\int q \, \rho_n(\,\mathrm{d}u,\, \mathrm{d}q) = (1-\epsilon_n)p + \epsilon_n r_n = p_n.$$

Thus $\rho_n \in \mathfrak{R}_k(x_n, p_n)$. Since $\widetilde{\rho}_n \Rightarrow \rho$ and $\epsilon_n \to 0$, the additional atom has vanishing mass and therefore $\rho_n \Rightarrow \rho$. This proves lower hemicontinuity.

Continuity of the stage objective and attainment. Continuity of r_k, F_k, and Z makes $C_k^Z(x,u,q,v)$ continuous and bounded on the relevant compact domain. Weak convergence of ρ and β therefore implies convergence of its integral, so L_k^Z is continuous. It is affine in ρ and in β because integration is linear.

For fixed (x,p), both $\mathfrak{R}_k(x,p)$ and $\mathcal{P}(\mathcal{V})$ are compact and convex. Sion's minimax theorem consequently gives

$$\min_{\rho \in \mathfrak{R}_k(x,p)} \max_{\beta \in \mathcal{P}(\mathcal{V})} L_k^Z(x,\rho,\beta) = \max_{\beta \in \mathcal{P}(\mathcal{V})} \min_{\rho \in \mathfrak{R}_k(x,p)} L_k^Z(x,\rho,\beta).$$

Continuity and compactness imply that both extrema are attained.

To prove continuity of $\mathcal{T}_k^G Z$, first maximize L_k^Z over the fixed compact set $\mathcal{P}(\mathcal{V})$. Berge's theorem makes the resulting function continuous in (x,ρ). Minimizing this function over the continuous compact-valued correspondence $(x,p) \mapsto \mathfrak{R}_k(x,p)$ and applying Berge's theorem again proves continuity in (x,p).

Finally, let $p^\lambda := \lambda p^0 + (1-\lambda)p^1$ and choose optimal prescriptions $\rho^0 \in \mathfrak{R}_k(x,p^0)$ and $\rho^1 \in \mathfrak{R}_k(x,p^1)$. Their mixture

$$\rho^\lambda := \lambda\rho^0 + (1-\lambda)\rho^1$$

belongs to $\mathfrak{R}_k(x,p^\lambda)$. Hence

$$\begin{aligned}
(\mathcal{T}_k^G Z)(x,p^\lambda) &\le \max_\beta \left[\lambda L_k^Z(x,\rho^0,\beta) + (1-\lambda)L_k^Z(x,\rho^1,\beta)\right] \\
&\le \lambda \max_\beta L_k^Z(x,\rho^0,\beta) + (1-\lambda)\max_\beta L_k^Z(x,\rho^1,\beta) \\
&= \lambda(\mathcal{T}_k^G Z)(x,p^0) + (1-\lambda)(\mathcal{T}_k^G Z)(x,p^1).
\end{aligned}$$

Therefore $\mathcal{T}_k^G Z(x,\cdot)$ is convex. □

Lemma 8 (Exact implementability of posterior-control prescriptions)**.** *Suppose $Z(x,\cdot)$ is convex. For every $\rho \in \mathfrak{R}_k(x,p)$ there is an implementable $\bar{\rho} \in \mathfrak{R}_k(x,p)$ such that*

$$L_k^Z(x,\bar{\rho},\beta) \le L_k^Z(x,\rho,\beta) \qquad \forall \beta \in \mathcal{P}(\mathcal{V}).$$

Proof. Disintegrate $\rho(\,\mathrm{d}u,\,\mathrm{d}q) = m(\,\mathrm{d}u)\kappa(\,\mathrm{d}q \mid u)$ and replace the posterior labels at fixed u by their conditional mean $\bar{q}(u) = \int q\kappa(\,\mathrm{d}q \mid u)$. The branch payoff is convex in q, so conditional Jensen cannot increase P1's objective. For $p_i > 0$, define the type-i current law

$$\alpha_i(\,\mathrm{d}u) = \frac{\bar{q}_i(u)}{p_i} m(\,\mathrm{d}u).$$

The barycenter identity makes α_i a probability measure, and Bayes' rule recovers $\bar{q}(u)$ after observing u. Standard-Borel disintegration and measurable selection make the construction Borel in the state variables. For a zero-prior type $p_i = 0$, choose any Borel safe current law; it does not affect the outcome law at the current prior. □

Lemma 9 (P1 security)**.** *The recursion Equation* (21) *constructs an admissible P1 behavioral strategy satisfying*

$$\sup_{\zeta\in\Sigma_2^G} J^G_{k,x,p}(\eta^*,\zeta) \le V^G_k(x,p).$$

Proof. We proceed backward in k. At $k = K$, the terminal value equals the terminal loss, so the claim is immediate.

Suppose the claim holds from stage $k+1$. Fix a viable public information state (x,p) at stage k. Let

$$\bar\rho(\,\mathrm{d}u,\,\mathrm{d}q) = m(\,\mathrm{d}u)\delta_{q(u)}(\,\mathrm{d}q)$$

be an optimal implementable prescription in Equation (20). By Lemma 8, P1 can generate this prescription using type-dependent current laws $\{\alpha_i\}_{i=1}^I$.

Now fix an arbitrary behavioral P2 strategy. Conditional on the current public history, this strategy induces some current law $\beta \in \mathcal{P}(\mathcal{V})$. Because the controls are chosen simultaneously, β cannot depend on P1's realized current control u. Moreover, P2's current draw carries no information about P1's type. Consequently, after (u,v) becomes public, the posterior remains $q(u)$.

On the branch (u,v), P1 continues with the induction-stage security strategy at

$$\big(k+1, F_k(x,u,v), q(u)\big).$$

By the induction hypothesis, its continuation loss against any P2 continuation strategy is at most

$$V^G_{k+1}(F_k(x,u,v), q(u)).$$

Adding the current stage loss and averaging over the current controls gives

$$J^G_{k,x,p}(\eta^*,\zeta) \le L_k^{V^G_{k+1}}(x,\bar\rho,\beta).$$

Because $\bar\rho$ minimizes the worst-case one-stage expression,

$$L_k^{V^G_{k+1}}(x,\bar\rho,\beta) \le \max_{\widetilde\beta\in\mathcal{P}(\mathcal{V})} L_k^{V^G_{k+1}}(x,\bar\rho,\widetilde\beta) = V^G_k(x,p).$$

Therefore

$$J^G_{k,x,p}(\eta^*,\zeta) \le V^G_k(x,p)$$

for every P2 behavioral strategy ζ. Taking the supremum over ζ proves

$$\sup_{\zeta\in\Sigma_2^G} J^G_{k,x,p}(\eta^*,\zeta) \le V^G_k(x,p).$$

The construction is made at every viable information state using Borel selectors. Concatenating the selected current kernels with the induction-stage strategies therefore gives an admissible Borel P1 strategy secure against arbitrary history-dependent behavioral P2 deviations. □

Choose payoff bounds B_k satisfying

$$B_K \ge \max_{i,x}|g_i(x)|, \qquad B_k \ge B_{k+1} + \max_{i,x,u,v}|r_{i,k}(x,u,v)|.$$

Every typewise remaining payoff lies in $[-B_k, B_k]$. Consequently $V^G_k(x,\cdot)$ is B_k-Lipschitz on the simplex in the ℓ_1 metric. Define its convex conjugate

$$W_k(x,z) := \sup_{p\in\Delta(\mathcal{I})}\{\langle p,z\rangle - V^G_k(x,p)\}, \qquad z\in\mathbb{R}^I. \tag{22}$$

It satisfies

$$W_k(x,z+c\mathbf{1}) = W_k(x,z) + c, \qquad W_K(x,z) = \max_i\{z_i - g_i(x)\}.$$

Lemma 10 (Bounded normalized subgradients)**.** *For every (k, x, p) the relative subdifferential*

$$\partial_\Delta V_k^G(x,p) := \{z : V_k^G(x,q) \geq V_k^G(x,p) + \langle z, q-p\rangle \quad \forall q \in \Delta(\mathcal{I})\}$$

is nonempty. It contains a representative whose Ith coordinate is zero, that is, $z_I = 0$, and for which $|z|_\infty \leq 2B_k$. Hence, with

$$Y_k := \{z \in \mathbb{R}^I : z_I = 0,\ \|z\|_\infty \leq 2B_k\},$$
$$-V_k^G(x,p) = \min_{z \in Y_k}\{W_k(x,z) - \langle p, z\rangle\}. \tag{23}$$

Proof. On the affine hyperplane $\{\langle p, \mathbf{1}\rangle = 1\}$, let

$$\widetilde{V}_k(x,q) := \inf_{r \in \Delta(\mathcal{I})} \left\{V_k^G(x,r) + B_k\|q-r\|_1\right\}, \qquad \sum_i q_i = 1.$$

This is a finite convex B_k-Lipschitz extension that agrees with V_k^G on $\Delta(\mathcal{I})$. A finite convex function on a finite-dimensional affine space has a subgradient everywhere, and every subgradient of this extension has dual norm at most B_k. Adding a multiple of $\mathbf{1}$ does not change its restriction to the simplex. Normalizing the Ith coordinate to zero therefore gives $\|z\|_\infty \leq 2B_k$. Fenchel equality at this normalized subgradient proves Equation (23). □

For $\beta \in \mathcal{P}(\mathcal{V})$ define the compact convex coupling set

$$\mathfrak{C}_{k+1}(\beta) := \{\lambda \in \mathcal{P}(\mathcal{V} \times Y_{k+1}) : \lambda(B \times Y_{k+1}) = \beta(B) \text{ for every Borel } B \subseteq \mathcal{V}\}.$$

For $u \in A_k(x)$ put

$$D_k(x,z;\beta,u) := \min_{\lambda \in \mathfrak{C}_{k+1}(\beta)} \left\{ \int W_{k+1}(F_k(x,u,v),y)\lambda(\,\mathrm{d}v,\,\mathrm{d}y) + \max_i \left[z_i - \int \big(y_i + r_{i,k}(x,u,v)\big)\lambda(\,\mathrm{d}v,\,\mathrm{d}y)\right]\right\}. \tag{24}$$

Lemma 11 (Continuity of the coupling correspondence)**.** *The correspondence $\beta \mapsto \mathfrak{C}_{k+1}(\beta)$ has nonempty compact convex values and is continuous under weak convergence. Consequently, if the integrand in Equation* (24) *is continuous, then D_k is continuous in (x, z, β, u) and its minimizing-coupling correspondence has a Borel graph and nonempty compact values.*

Proof. Nonemptiness follows from $\beta \otimes \delta_{y_0}$ for any $y_0 \in Y_{k+1}$; compactness and convexity follow from compactness of $\mathcal{V} \times Y_{k+1}$ and closedness of the marginal constraint. The same closedness gives upper hemicontinuity. For lower hemicontinuity, let $\beta_n \Rightarrow \beta$ and disintegrate $\lambda(\,\mathrm{d}v,\,\mathrm{d}y) = \beta(\,\mathrm{d}v)\kappa(\,\mathrm{d}y \mid v)$. Choose couplings γ_n of (β, β_n) converging to the diagonal coupling. Such couplings exist because weak convergence on a compact metric space admits a Skorokhod coupling, or equivalently because the bounded Wasserstein-1 metric metrizes weak convergence (Billingsley, 1999). Set

$$\lambda_n(\,\mathrm{d}v',\,\mathrm{d}y) := \int \gamma_n(\,\mathrm{d}v,\,\mathrm{d}v')\kappa(\,\mathrm{d}y \mid v).$$

Then $(\lambda_n)_{\mathcal{V}} = \beta_n$ and $\lambda_n \Rightarrow \lambda$. Berge's theorem and the measurable minimum theorem give the final claims. □

Lemma 12 (Exact dual Bellman recursion)**.** *The conjugate of the primal recursion is*

$$W_k(x,z) = \min_{\beta \in \mathcal{P}(\mathcal{V})} \max_{u \in A_k(x)} D_k(x,z;\beta,u). \tag{25}$$

All displayed extrema are attained. Optimal β and optimal couplings λ can be selected Borel measurably in their state parameters.

Proof. Because the barycenter of ρ is p, conjugating Equation (20) gives

$$W_k(x,z) = \sup_{\rho\in\mathcal{P}(A_k(x)\times\Delta(\mathcal{I}))}\min_{\beta\in\mathcal{P}(\mathcal{V})}\int\left[\langle q, z - r_k(x,u,v)\rangle - V^G_{k+1}(F_k(x,u,v),q)\right]\,\mathrm{d}\rho\,\mathrm{d}\beta. \tag{26}$$

Both measure spaces are compact and convex and the integrand is continuous and affine in each measure. Sion's theorem therefore interchanges $\sup_\rho$ and $\min_\beta$; maximizing over ρ then reduces to a maximum over (u,q).

Fix (x,z,β,u). By Equation (23),

$$-V^G_{k+1}(F_k(x,u,v),q) = \min_{y\in Y_{k+1}}\{W_{k+1}(F_k(x,u,v),y) - \langle q,y\rangle\}.$$

For each fixed q, Lemma 1 provides a Borel selection of the point-wise minimizing dual state y as a function of v. Equivalently, these selections can be represented by couplings $\lambda \in \mathfrak{C}_{k+1}(\beta)$. A second application of Sion, now between $q\in\Delta(\mathcal{I})$ and $\lambda\in\mathfrak{C}_{k+1}(\beta)$, yields

$$\begin{aligned}&\max_{q\in\Delta(\mathcal{I})}\int\left[\langle q, z - r_k(x,u,v)\rangle - V^G_{k+1}(F_k(x,u,v),q)\right]\beta(\,\mathrm{d}v)\\ &\quad= \min_{\lambda\in\mathfrak{C}_{k+1}(\beta)}\left\{\int W_{k+1}(F_k(x,u,v),y)\,\mathrm{d}\lambda + \max_i\left[z_i - \int(y_i + r_{i,k})\,\mathrm{d}\lambda\right]\right\}.\end{aligned}$$

This is $D_k(x,z;\beta,u)$ and proves Equation (25). Compactness gives attainment. Lemma 11 shows that D_k is continuous and that the correspondence of optimal continuation couplings λ has a Borel graph and nonempty compact values. Because D_k is continuous, Berge's theorem gives the same properties for the correspondence of optimal current laws β in Equation (25). Finally, Lemma 1 provides Borel selections of both the optimal current law β and the optimal continuation coupling λ. □

Lemma 13 (Dual P2 security)**.** *For every (k,x,z), the dual recursion constructs an admissible P2 strategy $\zeta^{k,x,z}$ satisfying*

$$z_i - J^G_{i,k,x}(\eta_i,\zeta^{k,x,z}) \le W_k(x,z) \tag{27}$$

for every type i and every admissible type-i P1 strategy.

Proof. The claim is immediate at K from $W_K(x,z) = \max_i(z_i - g_i(x))$. Suppose it holds at $k+1$. At (k,x,z), P2 selects an optimal current law β^* in Equation (25). For each public P1 control u, select an optimal coupling $\lambda^*_u \in \mathfrak{C}_{k+1}(\beta^*)$ in Equation (24). Lemma 2 gives a jointly Borel kernel

$$\lambda^*_u(\,\mathrm{d}v,\,\mathrm{d}y) = \beta^*(\,\mathrm{d}v)\kappa^*(\,\mathrm{d}y \mid u,v).$$

P2 first draws $v\sim\beta^*$, independently of the simultaneous P1 draw. After (u,v) becomes public, it draws the next private dual state $y\sim\kappa^*(\cdot\mid u,v)$ and uses the induction-stage strategy at $(k+1, F_k(x,u,v), y)$.

Fix a type i and a current pure P1 action u. The induction hypothesis gives, conditional on (u,v,y),

$$y_i - J^G_{i,k+1} \le W_{k+1}(F_k(x,u,v),y).$$

Taking expectations under λ^*_u,

$$\begin{aligned}z_i - \mathbb{E}[r_{i,k} + J^G_{i,k+1}] &\le \int W_{k+1}(F_k(x,u,v),y)\,\mathrm{d}\lambda^*_u + z_i - \int(y_i + r_{i,k})\,\mathrm{d}\lambda^*_u\\ &\le D_k(x,z;\beta^*,u) \le \max_{\widetilde{u}\in A_k(x)} D_k(x,z;\beta^*,\widetilde{u}) = W_k(x,z).\end{aligned}$$

Integrating this point-wise inequality over any type-i current behavioral law and iterating over public histories proves Equation (27). The recursion first defines a measurable randomized contingent plan whose private seed records the latent dual state. Its kernels are Borel by Lemmas 1 and2. Lemma 3 then provides an outcome-equivalent behavioral P2 strategy. Hence the guarantee holds in the behavioral strategy class used to define $V^{G,-}_k$. □

Proof of Theorem 1. Backward induction from the terminal condition and Lemma 7 proves boundedness, continuity, convexity in p, and attainment. Lemma 9 gives $V_k^{G,+} \le V_k^G$. Choose $z \in \partial_\Delta V_k^G(x,p) \cap Y_k$, whose existence follows from Lemma 10, and use Fenchel equality

$$V_k^G(x,p) + W_k(x,z) = \langle p, z \rangle.$$

Weighting Equation (27) by p_i gives $J_{k,x,p}^G(\eta, \zeta^{k,x,z}) \ge V_k^G(x,p)$ for every P1 strategy. Hence $V_k^{G,-} \ge V_k^G$. Together with P1 security and the minimax inequality,

$$V_k^G \le V_k^{G,-} \le V_k^{G,+} \le V_k^G,$$

so all quantities are equal. Measurable selection yields Borel strategies, and Corollary 1 gives almost-sure safety. □

B Proof of Upper Primal Atomicity

For fixed (k,x) define

$$a_k(u,q) := \max_{v \in \mathcal{V}} \left\{ \sum_i q_i r_{i,k}(x,u,v) + V_{k+1}^{\uparrow}(F_k(x,u,v), q) \right\}. \tag{28}$$

Lemma 14 (Upper recursion regularity)**.** *For every k, $V_k^{\uparrow}$ is finite, continuous, and convex in p. The function a_k is continuous and convex in q for fixed u, and $\widetilde{M}_k^{\uparrow}(x,q) = \min_{u \in \mathcal{U}_k^{\mathrm{safe}}(x)} a_k(u,q)$ is continuous. All displayed extrema are attained and have Borel selectors.*

Proof. The claim holds at the terminal stage. If it holds at $k+1$, the branch expression is continuous and convex in q; maximizing over compact $\mathcal{V}$ preserves both. Berge's theorem and continuity of the safe correspondence give continuity and attainment of $\widetilde{M}_k^{\uparrow}$. The moment set in Equation (9) is compact, and its lower convex envelope is finite, continuous, and convex. Measurable selection completes the induction. □

Lemma 15 (Extreme Bayes-plausible splits)**.** *For every $p \in \Delta(\mathcal{I})$ and every continuous $m : \Delta(\mathcal{I}) \to \mathbb{R}$, the problem*

$$\min_{\nu \in \mathcal{P}(\Delta(\mathcal{I}))} \left\{ \int m(q)\nu(\,\mathrm{d}q) : \int q\nu(\,\mathrm{d}q) = p \right\}$$

has an optimal measure supported on at most I points.

Proof. The feasible moment set is compact and convex, and the objective is continuous and linear. Its minimizers form a nonempty compact face, so one may choose a minimizer that is extreme in the feasible moment set. If an extreme measure charged more than I disjoint measurable sets, the vectors formed by total mass and the $I-1$ independent conditional belief moments on those sets would be linearly dependent. The resulting nonzero signed perturbation preserves total mass and all belief moments. Sufficiently small positive and negative perturbations would then express the measure as the midpoint of two distinct feasible measures, contradicting extremality. Hence an extreme feasible measure has at most I support points. □

Let η_i be P1's current control law for type i and set $\bar{\eta} = \sum_i p_i \eta_i$. For $\bar{\eta}$-almost every u, define

$$q_i(u) := p_i \frac{\mathrm{d}\eta_i}{\mathrm{d}\bar{\eta}}(u).$$

Then $q(u) \in \Delta(\mathcal{I})$ and $\int q(u)\bar{\eta}(\,\mathrm{d}u) = p$.

Lemma 16 (Behavioral/splitting equivalence)**.** *The upper behavioral Bellman problem equals Equation* (9)*.*

Proof. After observing the announced control u, P2's posterior is $q(u)$ and P2 chooses a measurable maximizing response. Hence the behavioral objective is

$$\int_{\mathcal{U}_k^{\mathrm{safe}}(x)} a_k(u, q(u))\, \bar{\eta}(\,\mathrm{d}u).$$

Let ν be the distribution of $q(u)$ induced by $u \sim \bar{\eta}$. Since

$$\int q\, \nu(\,\mathrm{d}q) = \int q(u)\, \bar{\eta}(\,\mathrm{d}u) = p,$$

the measure ν is Bayes plausible. Moreover, $a_k(u, q) \geq \widetilde{M}_k^{\uparrow}(x, q)$, so

$$\int a_k(u, q(u))\, \bar{\eta}(\,\mathrm{d}u) \geq \int_{\Delta(\mathcal{I})} \widetilde{M}_k^{\uparrow}(x, q)\, \nu(\,\mathrm{d}q).$$

Thus every behavioral prescription induces a feasible posterior split with no larger split objective.

Conversely, by Lemma 15, the moment problem has an optimal finite split

$$\nu^* = \sum_{r=1}^{R} \lambda_r \delta_{p^r}, \qquad \sum_{r=1}^{R} \lambda_r p^r = p, \qquad R \leq I.$$

For each branch r, choose

$$u^r \in \operatorname*{arg\,min}_{u \in \mathcal{U}_k^{\mathrm{safe}}(x)} a_k(u, p^r).$$

If several branches choose the same control, merge them into one branch: its probability is the sum of their probabilities and its posterior is their probability-weighted average. Convexity of $a_k(u, \cdot)$ ensures that this merging does not increase P1's objective. We may therefore assume that each announced control identifies its associated posterior.

For every type with $p_i > 0$, let type i choose u^r with probability

$$\Pr(u^r \mid i) = \frac{\lambda_r p_i^r}{p_i}.$$

The barycenter identity shows that these probabilities sum to one, while Bayes' rule gives

$$\Pr(i \mid u^r) = \frac{p_i \Pr(u^r \mid i)}{\lambda_r} = p_i^r.$$

Types with $p_i = 0$ may use any safe current law. The resulting behavioral prescription therefore implements the finite split and attains no more than its objective value. Combined with the first inequality, this proves that the behavioral and moment formulations have the same value. □

Proof of Theorem 2. The feasible moment set

$$\mathfrak{M}(p) := \{\nu \in \mathcal{P}(\Delta(\mathcal{I})) : \int q\nu(\,\mathrm{d}q) = p\}$$

is compact and convex. Choose an extreme minimizer of the continuous linear objective. Unit mass plus $I - 1$ independent belief coordinates give I independent moment constraints, so an extreme feasible probability measure has at most I support points. Lemma 16 supplies the type-conditioned control probabilities, and every selected control is robustly safe. Backward measurable concatenation gives the full-horizon strategy. Apply Lemma 15 to $m(q) = \widetilde{M}_k^{\uparrow}(x, q)$. Lemma 16 supplies the type-conditioned control probabilities, and every selected control is robustly safe. Backward measurable concatenation gives the full-horizon strategy. □

C Proof of Lower Dual Atomicity

This appendix proves the exact Fenchel-dual recursion for the robust-safe lower relaxation and the I-atom P2 leader asserted in Theorem 3. The proof has four parts. First, we construct the compact convex set $\mathcal{S}_k(x)$ of typewise payoff vectors that one P2 strategy can secure simultaneously

and prove its recursive realization. This is the state-constraint-specific step: Hausdorff continuity of $\mathcal{U}_k^{\mathrm{safe}}(x)$ is needed to close the recursion under limits and to select robust-safe continuations measurably. Second, we express $V_k^{\downarrow}$ as the support function of $\mathcal{S}_k(x)$ and identify its Fenchel conjugate with the minimum worst-type security shortfall. Third, we derive the exact mean-preserving dual split. A preliminary Carathéodory argument gives a bounded $I+1$-atom optimizer; translation invariance then removes the payoff-irrelevant common coordinate, and extremality in the remaining $(I-1)$-dimensional space sharpens the bound to I atoms. Finally, we realize those atoms as a Borel P2 strategy and use Fenchel subgradient recovery to obtain an optimal leader in the primal lower game.

Throughout this appendix,

$$V_k^*(x,z) := \sup_{p\in\Delta(\mathcal{I})} \left\{ \langle p, z\rangle - V_k^{\downarrow}(x,p) \right\}. \tag{29}$$

denotes the conjugate of the lower-relaxation value. It is distinct from the conjugate of the simultaneous-game value used in Appendix A.

Choose bounds $B_k < \infty$ such that every typewise continuation loss from stage k lies in $[-B_k, B_k]$. All vector inequalities below are componentwise.

C.1 Typewise security sets

For a fixed state x, a vector $\sigma \in \mathbb{R}^I$ is a *typewise security vector* if one P2 strategy guarantees at least σ_i when P1's type is i. The same P2 strategy must provide all I guarantees, because P2 does not observe the type.

At the terminal stage, define

$$\mathcal{S}_K(x) := \{\sigma \in [-B_K, B_K]^I : \sigma_i \le g_i(x) \text{ for every } i\}. \tag{30}$$

Suppose that $\mathcal{S}_{k+1}$ has already been constructed. If P2 announces v at stage k, let

$$\mathcal{A}_k(x,v) := \{b \in [-B_k, B_k]^I : \text{for every } u \in \mathcal{U}_k^{\mathrm{safe}}(x), \text{ there exists } \sigma^u \in \mathcal{S}_{k+1}(F_k(x,u,v)) \text{ such that } b \le r_k(x,u,v) + \sigma^u\}. \tag{31}$$

Here $r_k = (r_{1,k}, \dots, r_{I,k})$. The continuation vector σ^u may depend on the publicly observed response u, but not on the hidden type. Thus $b \in \mathcal{A}_k(x,v)$ means that announcing v and then choosing a suitable continuation after every response u guarantees b.

P2 may randomize over announced controls. We therefore set

$$\mathcal{A}_k(x) := \bigcup_{v\in\mathcal{V}} \mathcal{A}_k(x,v), \qquad \mathcal{S}_k(x) := \operatorname{co} \mathcal{A}_k(x). \tag{32}$$

Lemma 17 (Security-set recursion and realization)**.** *For every stage k and viable state x, the set $\mathcal{S}_k(x)$ is nonempty, compact, convex, and downward closed relative to $[-B_k, B_k]^I$. The correspondence $x \mapsto \mathcal{S}_k(x)$ has a closed graph. Moreover,*

$$\mathcal{S}_k(x) = \left\{ \sigma \in [-B_k, B_k]^I : \begin{array}{c} \textit{there exists } \zeta \in \Sigma_2^{\downarrow} \textit{ such that} \\ \sigma_i \le \inf\limits_{\eta_i^{\downarrow} \in \Sigma_{1,i}^{\downarrow}} J_{i,k,x}^{\downarrow}(\eta_i^{\downarrow}, \zeta) \quad \textit{for every } i \end{array} \right\}. \tag{33}$$

The realizing decision kernels can be selected Borel measurably in the augmented state (k, x, σ).

Proof. We proceed backward in k. The terminal claim follows from Equation (30). Assume the result holds at stage $k+1$.

First consider $\mathcal{A}_k(x,v)$. It is downward closed by definition. It is also convex: if $b^1, b^2 \in \mathcal{A}_k(x,v)$, then for every safe response u there are continuation vectors $\sigma^{1,u}$ and $\sigma^{2,u}$ satisfying Equation (31). Convexity of $\mathcal{S}_{k+1}(F_k(x,u,v))$ shows that $\lambda\sigma^{1,u} + (1-\lambda)\sigma^{2,u}$ certifies $\lambda b^1 + (1-\lambda)b^2$.

We next verify closedness. Let $(x_n, v_n, b_n) \to (x, v, b)$ with $b_n \in \mathcal{A}_k(x_n, v_n)$. Fix any $u \in \mathcal{U}_k^{\mathrm{safe}}(x)$. Hausdorff continuity of the safe-control correspondence gives controls $u_n \in \mathcal{U}_k^{\mathrm{safe}}(x_n)$ with $u_n \to u$. For each n, choose

$$\sigma_n \in \mathcal{S}_{k+1}(F_k(x_n, u_n, v_n)), \qquad b_n \le r_k(x_n, u_n, v_n) + \sigma_n.$$

The vectors σ_n lie in a common compact box, so a subsequence converges to some σ. Continuity of F_k, r_k and the closed graph of $\mathcal{S}_{k+1}$ imply

$$\sigma \in \mathcal{S}_{k+1}(F_k(x,u,v)), \qquad b \le r_k(x,u,v) + \sigma.$$

Since u was arbitrary, $b \in \mathcal{A}_k(x,v)$. Thus the graph of $(x,v) \mapsto \mathcal{A}_k(x,v)$ is closed. Compactness follows from the compact domains and the bounding box.

The graph of $x \mapsto \mathcal{S}_k(x)$ is also closed. By Carathéodory's theorem, every $\sigma_n \in \mathcal{S}_k(x_n)$ has a representation

$$\sigma_n = \sum_{a=1}^{I+1} \lambda_{n,a} b_{n,a}, \qquad b_{n,a} \in \mathcal{A}_k(x_n, v_{n,a}),$$

where zero weights are allowed. Compactness permits passage to convergent subsequences of the weights, controls, and branch vectors. The closed-graph property just proved places each limiting branch vector in $\mathcal{A}_k(x, v_a)$, and the limiting convex combination equals σ. Hence $\sigma \in \mathcal{S}_k(x)$. Nonemptiness follows from viability and the induction hypothesis; convexity is built into Equation (32); and downward closure is inherited from the branch sets.

We now prove realization. For a branch triple (x,v,b) with $b \in \mathcal{A}_k(x,v)$, define

$$\mathcal{K}_k(x,v,b,u) := \{\sigma \in \mathcal{S}_{k+1}(F_k(x,u,v)) : b \le r_k(x,u,v) + \sigma\}. \tag{34}$$

For every safe u, this correspondence has nonempty compact values and a Borel graph. Lemma 1 therefore provides a Borel continuation choice.

The finite-dimensional correspondence of Carathéodory representations

$$\sigma = \sum_{a=1}^{I+1} \lambda_a b^a, \qquad b^a \in \mathcal{A}_k(x, v^a),$$

also has a Borel graph and nonempty compact values. Its weights and branches can therefore be selected measurably. P2 draws branch a with probability λ_a, announces v^a, observes P1's response u, selects a continuation vector from Equation (34), and uses the induction-stage strategy realizing that vector. If two selected branches announce the same v, replace them by their weight-averaged branch vector. Convexity of $\mathcal{A}_k(x,v)$ preserves feasibility, and the realized announcement now identifies the continuation branch. This produces a Borel P2 strategy realizing σ.

Conversely, fix any admissible Borel P2 strategy and condition on its current announced control v. For each public response u, let $\sigma(v,u) \in \mathcal{S}_{k+1}(F_k(x,u,v))$ be the typewise guarantee vector of the conditional continuation strategy. Define

$$b_i(v) := \inf_{u \in \mathcal{U}_k^{\mathrm{safe}}(x)} \big[r_{i,k}(x,u,v) + \sigma_i(v,u) \big]. \tag{35}$$

Then $b(v) \le r_k(x,u,v) + \sigma(v,u)$ for every safe u, so $b(v) \in \mathcal{A}_k(x,v)$. Lemma 2 gives a jointly Borel version of the conditional continuation strategy after the observed control v. Applying Lemma 1 typewise to the current P1 minimization then shows that $v \mapsto b(v)$ is Borel measurable. Because P1 observes v before responding, the current typewise guarantee is $\int b(v)\beta(\mathrm{d}v)$, where β is P2's current law. This vector belongs to $\operatorname{co} \mathcal{A}_k(x) = \mathcal{S}_k(x)$. The two directions prove Equation (33). □

C.2 Conjugate representation and dual recursion

Lemma 18 (Support-function and dual identities)**.** *For every viable* (k,x)*,*

$$V_k^{\downarrow}(x,p) = \max_{\sigma \in \mathcal{S}_k(x)} \langle p, \sigma \rangle, \tag{36}$$

$$V_k^{*}(x,\widehat{p}) = \min_{\sigma \in \mathcal{S}_k(x)} \max_{i \in \mathcal{I}} (\widehat{p}_i - \sigma_i). \tag{37}$$

Proof. Equation (36) follows from Lemma 17: for a fixed P2 strategy, P1's type-conditioned responses separate, and weighting the resulting typewise guarantees by p gives the expected guarantee.

Substituting this support representation into Equation (29) gives

$$V_k^*(x,\widehat{p}) = \sup_{p\in\Delta(\mathcal{I})} \min_{\sigma\in\mathcal{S}_k(x)} \langle p, \widehat{p}-\sigma\rangle.$$

The simplex and the security set are compact and convex, and the integrand is bilinear. Sion's theorem interchanges the supremum and minimum. Finally, $\sup_{p\in\Delta(\mathcal{I})}\langle p,a\rangle = \max_i a_i$, which proves Equation (37). □

For a continuation dual state $y\in\mathbb{R}^I$, the vector form of the nonrevealing value is

$$\widetilde{N}_k(x,y) := \min_{b\in\mathcal{A}_k(x)} \max_i (y_i - b_i). \tag{38}$$

Lemma 19 (Nonrevealing dual identity)**.** *For every* (k,x,y,v)*,*

$$\min_{b\in\mathcal{A}_k(x,v)} \max_i (y_i - b_i) = \max_{u\in\mathcal{U}_k^{\mathrm{safe}}(x)} V_{k+1}^*(F_k(x,u,v), y - r_k(x,u,v)). \tag{39}$$

Consequently, minimizing over v *shows that Equation* (38) *equals the control recursion in Equation* (11)*.*

Proof. Write the right-hand side of Equation (39) as M. Take any $b\in\mathcal{A}_k(x,v)$ and any safe response u. By definition, there exists $\sigma^u\in\mathcal{S}_{k+1}(F_k(x,u,v))$ with $b\le r_k(x,u,v)+\sigma^u$. Lemma 18 gives

$$\max_i(y_i-b_i) \ge \max_i(y_i - r_{i,k}(x,u,v) - \sigma_i^u) \ge V_{k+1}^*(F_k(x,u,v), y-r_k(x,u,v)).$$

Taking the maximum over u and then the minimum over b proves the "$\ge$" inequality.

For the reverse inequality, for every safe u choose a minimizer σ^u in the strategic representation of $V_{k+1}^*(F_k(x,u,v), y-r_k(x,u,v))$, and define

$$b_i := \inf_{u\in\mathcal{U}_k^{\mathrm{safe}}(x)} [r_{i,k}(x,u,v) + \sigma_i^u]. \tag{40}$$

The recursive choice of B_k places b in the payoff box, and every σ^u certifies that $b\in\mathcal{A}_k(x,v)$. Moreover,

$$y_i - r_{i,k}(x,u,v) - \sigma_i^u \le M \qquad \text{for every } i,u.$$

Taking the supremum over u gives $y_i - b_i \le M$ for every i, and hence $\max_i(y_i-b_i)\le M$. This proves the reverse inequality. □

Under Fenchel conjugacy, the convexification $\mathcal{S}_k(x) = \operatorname{co}\mathcal{A}_k(x)$ becomes a mean-preserving randomization over continuation dual states. The next lemma makes this correspondence exact and supplies a bounded optimal split.

Lemma 20 (Exact dual split and bounded attainment)**.** *For every* $(k,x,\widehat{p})$*,*

$$V_k^*(x,\widehat{p}) = \min_{\substack{\mu\in\mathcal{P}(\mathbb{R}^I)\\ \int y\,\mu(\mathrm{d}y)=\widehat{p}}} \int \widetilde{N}_k(x,y)\,\mu(\mathrm{d}y). \tag{41}$$

The minimum is attained by a measure with at most $I+1$ *atoms, which may be chosen to satisfy* $\|y-\widehat{p}\|_\infty \le 2B_k$*.*

Proof. Let μ be any feasible mean-preserving split. For each y, select a minimizing $b(y)\in\mathcal{A}_k(x)$, so that

$$\widetilde{N}_k(x,y) = \max_i (y_i - b_i(y)).$$

A Borel selection exists by the measurable minimum theorem. Since $\mathcal{S}_k(x) = \operatorname{co}\mathcal{A}_k(x)$, $\bar{b} := \int b(y)\mu(\mathrm{d}y)$ belongs to $\mathcal{S}_k(x)$. Jensen's inequality gives

$$\int \widetilde{N}_k(x,y)\mu(\mathrm{d}y) \ge \max_i(\widehat{p}_i - \bar{b}_i) \ge V_k^*(x,\widehat{p}).$$

Thus the right-hand side of Equation (41) cannot be smaller than the left-hand side.

Conversely, let σ^* minimize Equation (37). Carathéodory's theorem gives

$$\sigma^* = \sum_{a=1}^{I+1} \lambda_a b^a, \qquad b^a \in \mathcal{A}_k(x), \qquad \lambda_a \geq 0, \quad \sum_a \lambda_a = 1, \tag{42}$$

where zero-weight atoms are allowed. Define

$$y^a := b^a + (\widehat{p} - \sigma^*), \qquad \mu^* := \sum_{a=1}^{I+1} \lambda_a \delta_{y^a}. \tag{43}$$

Then $\sum_a \lambda_a y^a = \widehat{p}$. Testing the branch value with b^a gives

$$\widetilde{N}_k(x, y^a) \leq \max_i (y_i^a - b_i^a) = \max_i (\widehat{p}_i - \sigma_i^*) = V_k^*(x, \widehat{p}).$$

After averaging, the preceding lower bound forces equality. Thus μ^* is optimal and has at most $I+1$ atoms. Finally, $b^a, \sigma^* \in [-B_k, B_k]^I$, so

$$\|y^a - \widehat{p}\|_\infty = \|b^a - \sigma^*\|_\infty \leq 2B_k.$$

□

The resulting split is still posed in $\mathbb{R}^I$, although common translations of the dual state affect all typewise shortfalls equally. We next remove this redundant direction.

Lemma 21 (Translation gauge of the dual split)**.** *Let* $\mathbf{1} = (1, \ldots, 1)$ *and*

$$\pi(y) := y - y_I \mathbf{1}, \qquad H := \{\bar{y} \in \mathbb{R}^I : \bar{y}_I = 0\}.$$

Then

$$V_k^*(x, z + c\mathbf{1}) = V_k^*(x, z) + c, \tag{44}$$

$$\widetilde{N}_k(x, y + c\mathbf{1}) = \widetilde{N}_k(x, y) + c. \tag{45}$$

Consequently, Equation (12) *is equivalent to the* $(I-1)$*-dimensional moment problem*

$$V_k^*(x, z) = z_I + \min_{\substack{\bar{\mu} \in \mathcal{P}_1(H) \\ \int_H \bar{y}\, \bar{\mu}(\,\mathrm{d}\bar{y}) = \pi(z)}} \int_H \widetilde{N}_k(x, \bar{y})\, \bar{\mu}(\,\mathrm{d}\bar{y}). \tag{46}$$

Proof. The security-shortfall representation gives

$$V_k^*(x, z + c\mathbf{1}) = \min_{\sigma \in \mathcal{S}_k(x)} \max_i \{z_i + c - \sigma_i\} = V_k^*(x, z) + c.$$

Equation (45) then follows directly from Equation (11): translation invariance of V_{k+1}^* adds the same constant c before the outer minimum and maximum, so both preserve it.

For a feasible split μ in Equation (12), write $y = \pi(y) + y_I \mathbf{1}$, and for a feasible split μ, let $\bar{\mu}$ be the distribution of $\pi(y)$ when $y \sim \mu$; equivalently,

$$\bar{\mu}(B) := \mu(\pi^{-1}(B)), \qquad B \subseteq H \text{ Borel}.$$

Its mean is $\pi(z)$, while $\int y_I \mu(\,\mathrm{d}y) = z_I$. Hence

$$\int \widetilde{N}_k(x, y) \mu(\,\mathrm{d}y) = z_I + \int_H \widetilde{N}_k(x, \bar{y}) \bar{\mu}(\,\mathrm{d}\bar{y}).$$

Conversely, lifting any feasible $\bar{\mu}$ in Equation (46) by $\bar{y} \mapsto \bar{y} + z_I \mathbf{1}$ produces a feasible split of mean z with the same displayed objective. This proves the equivalence. □

Lemma 22 (Gauge-reduced atomic dual split)**.** *For every* (k, x, z)*, Equation* (12) *has an optimal measure supported on at most* I *points.*

Proof. We first construct one finite optimal split and then re-extremize it after removing the translation direction. Let $\mathcal{A}_k(x)$ be the compact set of one-stage typewise guarantee vectors and let $\sigma^* \in \operatorname{co} \mathcal{A}_k(x)$ minimize $\max_i\{z_i - \sigma_i\}$. An initial application of Carathéodory's theorem in $\mathbb{R}^I$ writes

$$\sigma^* = \sum_a \lambda_a b^a, \qquad b^a \in \mathcal{A}_k(x),$$

using at most $I + 1$ terms. Setting $y^a := b^a + (z - \sigma^*)$ gives $\sum_a \lambda_a y^a = z$ and

$$\widetilde{N}_k(x, y^a) \le \max_i\{y_i^a - b_i^a\} = \max_i\{z_i - \sigma_i^*\}.$$

Thus this split is optimal. Let $K := \operatorname{co}\{\pi(y^a)\}_a \subset H$, which is compact and contains $\pi(z)$.

Restrict Equation (46) to measures on K; the split just constructed shows that the restricted and unrestricted values agree. Its minimizer set is a nonempty compact face of the feasible moment set. Choose an extreme minimizer. The affine space H has dimension $I - 1$, so its probability and mean constraints have $1 + (I - 1) = I$ independent scalar coordinates. If the extreme measure charged more than I disjoint sets, their mass-and-mean vectors would be linearly dependent. A sufficiently small signed perturbation along that dependence would preserve positivity, total mass, and the mean, contradicting extremality. The extreme minimizer therefore has at most I support points. Lifting it by $\bar{y} \mapsto \bar{y} + z_I \mathbf{1}$ yields the claimed optimal split in Equation (12). □

Lemma 23 (Subgradient recovery)**.** *Let $p \in \Delta(\mathcal{I})$ and $\widehat{p} \in \partial_\Delta V_k^\downarrow(x, p)$. If σ^* minimizes Equation* (37)*, then*

$$\langle p, \sigma^* \rangle = V_k^\downarrow(x, p). \tag{47}$$

Hence any P2 strategy realizing σ^* is optimal in the lower relaxation at prior p.

Proof. The subgradient condition gives Fenchel equality,

$$V_k^*(x, \widehat{p}) = \langle p, \widehat{p} \rangle - V_k^\downarrow(x, p).$$

Since σ^* is dual-optimal,

$$V_k^*(x, \widehat{p}) = \max_i(\widehat{p}_i - \sigma_i^*) \ge \langle p, \widehat{p} - \sigma^* \rangle.$$

Combining the two displays gives $\langle p, \sigma^* \rangle \ge V_k^\downarrow(x, p)$. The reverse inequality follows from Equation (36), because $\sigma^* \in \mathcal{S}_k(x)$. □

Proof of Theorem 3. Fix a dual state (k, x, z). By Lemma 22, choose an optimal split

$$\mu^* = \sum_{a=1}^{R} \lambda_a \delta_{y^a}, \qquad R \le I, \qquad \sum_a \lambda_a y^a = z.$$

For each atom, choose $b^a \in \mathcal{A}_k(x)$ attaining

$$\widetilde{N}_k(x, y^a) = \max_i(y_i^a - b_i^a),$$

and choose $v^a \in \mathcal{V}$ such that $b^a \in \mathcal{A}_k(x, v^a)$. Let

$$\bar{\sigma} := \sum_a \lambda_a b^a \in \mathcal{S}_k(x).$$

Optimality of the split and convexity of the coordinatewise maximum give

$$\begin{aligned} V_k^*(x, z) &= \sum_a \lambda_a \max_i(y_i^a - b_i^a) \\ &\ge \max_i \left\{ \sum_a \lambda_a (y_i^a - b_i^a) \right\} \\ &= \max_i\{z_i - \bar{\sigma}_i\} \ge V_k^*(x, z). \end{aligned}$$

Hence all inequalities are equalities, and $\bar{\sigma}$ minimizes the security-shortfall representation of $V_k^*(x, z)$.

By Lemma 17, each branch vector b^a is realized by a P2 continuation strategy whose current announced control is v^a. Privately mixing these branch strategies with probabilities λ_a therefore realizes $\bar{\sigma}$ and uses at most $R \le I$ current controls. It is optimal in the dual game by the preceding equality.

The finite atomic representations, branch controls, and continuation strategies can be selected Borel measurably by parameterizing them through their weights and atoms and applying Lemma 1. Robust pathwise safety follows because every P1 response belongs to $\mathcal{U}_k^{\mathrm{safe}}(x)$.

Finally, if $z \in \partial_\Delta V_k^\downarrow(x, p)$, Lemma 23 shows that the same strategy secures $V_k^\downarrow(x, p)$ and is therefore optimal in the primal lower game. □

D Proof of the Information-Relaxation Certificates

The upper and lower relaxations differ from the simultaneous game only in the information available to the current follower. The underlying dynamics, losses, and robust safe-control constraints remain unchanged.

Fix an admissible P1 leader strategy η in the upper relaxation. Any simultaneous P2 strategy can be used in the upper relaxation by ignoring the additional observation of P1's current control. Therefore

$$\sup_{\zeta' \in \Sigma_2^G} J_p^G(\eta, \zeta') \le \sup_{\zeta' \in \Sigma_2^\uparrow} J_p^\uparrow(\eta, \zeta') = U^\uparrow(\eta). \tag{48}$$

Similarly, fix an admissible P2 leader strategy ζ in the lower relaxation. Any simultaneous P1 strategy can be used in the lower relaxation by ignoring the additional observation of P2's current control. Because both games use the same robust safe-control correspondence,

$$\inf_{\eta' \in \Sigma_1^G} J_p^G(\eta', \zeta) \ge \inf_{\eta' \in \Sigma_1^\downarrow} J_p^\downarrow(\eta', \zeta) = L^\downarrow(\zeta). \tag{49}$$

In both embeddings, the payoff of a fixed strategy profile is unchanged; only the follower's response set is enlarged.

Proof of Proposition 1. Taking the infimum over P1 leaders in Equation (48) gives

$$V_k^G(x, p) \le V_k^\uparrow(x, p).$$

Taking the supremum over P2 leaders in Equation (49) gives

$$V_k^\downarrow(x, p) \le V_k^G(x, p).$$

This proves the value sandwich

$$V_k^\downarrow(x, p) \le V_k^G(x, p) \le V_k^\uparrow(x, p).$$

Now let $\eta^\uparrow$ and $\zeta^\downarrow$ be optimal leaders of the two relaxations. Their embeddings satisfy

$$\sup_{\zeta' \in \Sigma_2^G} J_p^G(\eta^\uparrow, \zeta') \le V^\uparrow$$

and

$$\inf_{\eta' \in \Sigma_1^G} J_p^G(\eta', \zeta^\downarrow) \ge V^\downarrow.$$

Subtracting the second inequality from the first gives

$$\mathrm{Gap}_G(\eta^\uparrow, \zeta^\downarrow) \le V^\uparrow - V^\downarrow = \Delta_{\mathrm{LF}}.$$

□

E PROOF OF THE STATE-CONSTRAINED COMMON-LIMIT THEOREM

This appendix proves that the upper and lower information relaxations converge to a common value on the exact continuous-time viability tube. We establishe domain consistency (E.3), derive HJI inequalities for asymmetric half-relaxed limits over the exact and discrete viable tubes (E.4), and prove comparison (E.5). The central difficulty is the mismatch between the constrained boundary Hamiltonian H_R and the bulk Hamiltonian H, which is handled by Koike's boundary exclusion; a Cardaliaguet-type convex perturbation in belief completes the comparison argument. Together with finite-grid ordering, this comparison yields vanishing relaxation width (E.6).

E.1 THE AUGMENTED TUBE AND KOIKE-TYPE BOUNDARY HYPOTHESES

We use two assumptions adapted from Koike Koike (1995). Assumption 3 requires admissible controls to remain admissible at nearby states, and Assumption 4 requires a uniform inward direction at the viability boundary. Lemma 24 gives a sufficient condition for both.

Write $z = (t, x)$ and include time in the dynamics:

$$\widehat{f}(z, u, v) := (1, f(t, x, u, v)). \tag{50}$$

Let $\mathcal{Q}(t)$ be the exact robust viability set of the continuous-time safety problem and let

$$\mathscr{D} := \{(t, x) : 0 \le t \le T,\ x \in \mathcal{Q}(t)\}. \tag{51}$$

The terminal condition applies at $t = T$; Koike's construction concerns only the lateral boundary $\partial_{\mathrm{lat}}\mathscr{D}$. To exclude $t = 0$ from this boundary, we assume that the data and tube extend locally to $(-\delta, T] \times \mathcal{X}$ for some $\delta > 0$.

Assumption 3 (Uniform robust local admissibility)**.** *There exist constants $r_0, s_0 > 0$ and a nonempty compact-valued correspondence*

$$\mathcal{R} : \mathscr{D} \rightrightarrows \mathcal{U}$$

with $\mathcal{R}(z) = \mathcal{U}$ in the interior of $\mathscr{D}$, such that the following holds. For every $z \in \partial_{\mathrm{lat}}\mathscr{D}$, every $u \in \mathcal{R}(z)$, every $z' = (t', x') \in \mathscr{D} \cap B_{r_0}(z)$, and every constant $v \in \mathcal{V}$, the augmented trajectory generated from z' by holding (u, v) fixed remains in $\mathscr{D}$ on

$$[0, \min\{s_0, T - t'\}].$$

For $z \in \partial_{\mathrm{lat}}\mathscr{D}$ and $v \in \mathcal{V}$, define the convexified admissible-velocity set

$$\mathcal{G}(z, v) := \mathrm{co}\{\widehat{f}(z, u, v) : u \in \mathcal{R}(z)\}. \tag{52}$$

For a unit vector ν and $\theta \in (0, 1)$, define the open cone

$$\mathcal{C}_\theta(\nu) := \bigcup_{s>0} B_{s\theta}(s\nu). \tag{53}$$

Assumption 4 (Koike geometry on the augmented tube)**.** *For every $z_0 \in \partial_{\mathrm{lat}}\mathscr{D}$, there exist $r > 0$, $\theta \in (0, 1)$, and a unit vector ν such that*

$$\bigcup_{0<s<r} B_{s\theta}(z + s\nu) \subset \mathrm{int}\,\mathscr{D} \quad \textit{for every } z \in \mathscr{D} \cap B_r(z_0), \tag{54}$$

and, for every lateral $z \in \mathscr{D} \cap B_r(z_0)$ and every $v \in \mathcal{V}$,

$$\mathcal{G}(z, v) \cap \mathcal{C}_\theta(\nu) \neq \varnothing. \tag{55}$$

The following smooth inward condition offers sufficient verification of Assumptions 3 and 4.

Lemma 24 (A smooth sufficient condition)**.** *Suppose locally that*

$$\mathscr{D} = \{z : b(z) \le 0\}, \qquad b \in C^1, \qquad \|D_z b\| \ge \nu_0 > 0 \tag{56}$$

on the lateral boundary. If, at each boundary point z, a control $u_z \in \mathcal{R}(z)$ satisfies

$$\max_{v \in \mathcal{V}} D_z b(z) \widehat{f}(z, u_z, v) \le -\kappa \tag{57}$$

with one $\kappa > 0$, then Assumption 3 and Assumption 4 hold.

Proof. We prove local admissibility, the interior-cone condition, and the velocity-cone intersection in turn. The argument below is written for $\mathcal{R}(z) = \{u_z\}$. More generally, the same proof applies if Equation (57) holds uniformly for every $u \in \mathcal{R}(z)$.

Local admissibility. Fix a lateral-boundary point z and its selected control u_z. By Equation (57),

$$D_z b(z)\widehat{f}(z, u_z, v) \le -\kappa \qquad \text{for every } v \in \mathcal{V}.$$

Continuity of $D_z b$ and $\widehat{f}$, together with compactness of $\mathcal{V}$, implies that there is a neighborhood N_z of z such that

$$D_z b(y)\widehat{f}(y, u_z, v) \le -\frac{\kappa}{2} \qquad \text{for every } y \in N_z,\ v \in \mathcal{V}. \tag{58}$$

Let $z' = (t', x') \in \mathscr{D} \cap N_z$, fix $v \in \mathcal{V}$, and let $Z(s)$ be the augmented trajectory generated from z' by holding (u_z, v) fixed. As long as $Z(s) \in N_z$, the chain rule and Equation (58) give

$$\frac{\mathrm{d}}{\mathrm{d}s} b(Z(s)) = D_z b(Z(s))\widehat{f}(Z(s), u_z, v) \le -\frac{\kappa}{2}.$$

Since $z' \in \mathscr{D}$, we have $b(z') \le 0$. Therefore

$$b(Z(s)) \le b(z') - \frac{\kappa}{2} s \le 0.$$

Boundedness of $\widehat{f}$ allows the neighborhood and time horizon to be chosen so that $Z(s)$ remains in N_z throughout this argument. Hence the same control u_z keeps every sufficiently nearby viable initial point inside $\mathscr{D}$ for a uniform short time, against every constant $v \in \mathcal{V}$.

Because the lateral boundary is compact and the constants κ and ν_0 are uniform, a finite-cover argument yields common constants $r_0, s_0 > 0$. Thus Assumption 3 holds. In fact, the same proof applies to an arbitrary measurable opponent control $\beta(\cdot)$, since Equation (58) holds for every value $v \in \mathcal{V}$.

Interior-cone geometry. Fix an anchor $z_0 \in \partial_{\mathrm{lat}}\mathscr{D}$ and define its inward unit normal by

$$\nu := -\frac{D_z b(z_0)}{\|D_z b(z_0)\|}.$$

Since $b \in C^1$ and $\|D_z b\| \ge \nu_0 > 0$, the inward unit normal varies continuously near z_0. Consequently, after reducing the neighborhood radius, there are $r > 0$ and $\theta \in (0, 1)$ such that

$$\bigcup_{0<s<r} B_{s\theta}(z + s\nu) \subset \operatorname{int} \mathscr{D} \qquad \text{for every } z \in \mathscr{D} \cap B_r(z_0).$$

This is the uniform interior-cone condition Equation (54). Notice that θ may be chosen close to one, while r is then chosen sufficiently small; this permits the cone to contain inward velocities with substantial tangential components.

Velocity-cone intersection. For a lateral point z sufficiently close to z_0, let

$$\nu_z := -\frac{D_z b(z)}{\|D_z b(z)\|}$$

be the local inward unit normal. The strict inward condition gives

$$\nu_z^\top \widehat{f}(z, u_z, v) \ge \frac{\kappa}{\|D_z b(z)\|} \qquad \text{for every } v \in \mathcal{V}.$$

The gradients $D_z b$ are locally bounded, so the right-hand side is bounded below by a positive constant. Continuity of the normal field then gives, after possibly reducing r,

$$\nu^\top \widehat{f}(z, u_z, v) \ge c > 0 \qquad \text{for every lateral } z \in B_r(z_0),\ v \in \mathcal{V}.$$

Since $\widehat{f}$ is bounded, this is a uniform positive-angle condition: all the selected velocities make an angle strictly smaller than $\pi/2$ with the common axis ν. We may therefore choose the aperture $\theta < 1$ in the preceding interior-cone construction large enough that

$$\widehat{f}(z, u_z, v) \in \mathcal{C}_\theta(\nu) \qquad \text{for every lateral } z \in B_r(z_0),\ v \in \mathcal{V}.$$

Because $u_z \in \mathcal{R}(z)$,

$$\widehat{f}(z, u_z, v) \in \mathcal{G}(z, v) = \operatorname{co}\{\widehat{f}(z, u, v) : u \in \mathcal{R}(z)\}.$$

It follows that

$$\mathcal{G}(z, v) \cap \mathcal{C}_\theta(\nu) \neq \varnothing,$$

which proves Equation (55). □

Appendix F introduces the actionable diagnostics using this smooth sufficient condition.

E.2 Bulk and constrained Isaacs identities

Now we adapt standard Isaacs' conditions for 2p0s differential games to our state-constrained games. These conditions naturally hold for the case studies in this paper, with compact convex controls, control-affine dynamics, and convex-concave losses.

Let $\ell = (\ell_1, \dots, \ell_I)$ and define

$$\Lambda(t, x, p, \xi; u, v) := \langle p, \ell(t, x, u, v) \rangle + \xi^\top f(t, x, u, v). \tag{59}$$

The two ordered bulk Hamiltonians are

$$H^\uparrow(t, x, p, \xi) := \min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} \Lambda(t, x, p, \xi; u, v), \tag{60}$$

$$H^\downarrow(t, x, p, \xi) := \max_{v \in \mathcal{V}} \min_{u \in \mathcal{U}} \Lambda(t, x, p, \xi; u, v). \tag{61}$$

On the lateral boundary define

$$H^\uparrow_{\mathcal{R}}(t, x, p, \xi) := \min_{u \in \mathcal{R}(t,x)} \max_{v \in \mathcal{V}} \Lambda(t, x, p, \xi; u, v), \tag{62}$$

$$H^\downarrow_{\mathcal{R}}(t, x, p, \xi) := \max_{v \in \mathcal{V}} \min_{u \in \mathcal{R}(t,x)} \Lambda(t, x, p, \xi; u, v). \tag{63}$$

Assumption 5 (Bulk and constrained Isaacs identities)**.** *For every relevant* (t, x, p, ξ)*,*

$$H^\uparrow = H^\downarrow =: H, \qquad H^\uparrow_{\mathcal{R}} = H^\downarrow_{\mathcal{R}} =: H_{\mathcal{R}}. \tag{64}$$

The second identity is required only on the lateral boundary; in the interior it reduces to the first because $\mathcal{R} = \mathcal{U}$*.*

These identities are the tight operator assumptions. Compact convex controls together with convex-concave Hamiltonian structure are one standard sufficient regime. From regularity in Assumption 1, for bounded p,

$$|H(z, p, \xi) - H(z', p, \xi)| \le \omega_\ell(\|z - z'\|) + L_f \|\xi\| \, \|z - z'\|, \tag{65}$$

and H is uniformly Lipschitz in ξ. This is the structural modulus to be used in the doubling argument below. Equivalently, for

$$K(z, p, \tau, \xi; u, v) := -\tau - \langle p, \ell(z, u, v) \rangle - \xi^\top f(z, u, v), \tag{66}$$

Assumption 5 gives

$$-(\tau + H) = \min_{v \in \mathcal{V}} \max_{u \in \mathcal{U}} K(z, p, \tau, \xi; u, v), \tag{67}$$

$$-(\tau + H_{\mathcal{R}}) = \min_{v \in \mathcal{V}} \max_{u \in \mathcal{R}(z)} K(z, p, \tau, \xi; u, v). \tag{68}$$

This is the control order to be used in the boundary-exclusion argument.

E.3 Deriving the exact-flow tube approximation

We next derive the mesh/tube consistency. Let

$$\psi(x) := \max_{1 \le j \le m} c_j(x). \tag{69}$$

On a compact neighborhood containing all reachable trajectories, define the robust safety value

$$S(t, x) := \inf_{\alpha \in \mathfrak{A}_t} \sup_{v(\cdot)} \max_{s \in [t, T]} \psi\big(X^{t,x;\alpha,v}(s)\big),$$

where $\mathfrak{A}_t$ consists of causal state-feedback strategies with delay and $v(\cdot)$ ranges over measurable disturbances. The value S is the unique bounded uniformly continuous viscosity solution of

$$\max\{\psi(x) - S(t, x), S_t(t, x) + H_{\text{safe}}(t, x, D_x S(t, x))\} = 0, \quad S(T, x) = \psi(x), \tag{70}$$

where

$$H_{\text{safe}}(t,x,\xi) := \min_{u\in\mathcal{U}} \max_{v\in\mathcal{V}} \xi^\top f(t,x,u,v). \tag{71}$$

The exact viability set is therefore (Mitchell et al., 2005)

$$\mathcal{Q}(t) = \{x : S(t,x) \le 0\}. \tag{72}$$

For a partition Π, define the exact-flow obstacle scheme

$$S_{\Pi,K}(x) := \psi(x), \tag{73}$$

$$S_{\Pi,k}(x) := \min_{u\in\mathcal{U}} \max_{v\in\mathcal{V}} \max\left\{ \max_{0\le s\le h_k} \psi(X_k(s;x,u,v)), S_{\Pi,k+1}(F_k(x,u,v)) \right\}. \tag{74}$$

Lemma 25 (Exact-flow safety scheme and tube convergence)**.** *The interpolated functions S_Π converge locally uniformly to S as $|\Pi| \to 0$. Moreover,*

$$\mathcal{Q}_{\Pi,k} = \{x : S_{\Pi,k}(x) \le 0\}, \tag{75}$$

and therefore

$$\limsup_{|\Pi|\to 0} \{(t_k,x) : x \in \mathcal{Q}_{\Pi,k}\} \subseteq \mathscr{D}. \tag{76}$$

Under the local-persistence clause of Assumption 3, there is $h_0 > 0$ such that

$$\mathcal{Q}(t_k) \subseteq \mathcal{Q}_{\Pi,k} \quad \textit{whenever } |\Pi| \le h_0. \tag{77}$$

Proof. The exact flow and integrated path obstacle satisfy, uniformly on the compact domain and in (u,v),

$$F_h(t,x,u,v) = x + hf(t,x,u,v) + O(h^2), \tag{78}$$

$$\max_{0\le s\le h} \|X(s;t,x,u,v) - x\| \le Ch. \tag{79}$$

The scheme Equation (74) is monotone, stable, and consistent with Equation (70). Since the safety problem is posed on the ambient state neighborhood, the standard half-relaxed-limit argument and obstacle comparison yield $S_\Pi \to S$ locally uniformly; see Bardi et al. (2000) for the related discrete-game approximation.

We prove Equation (75) by backward induction. At $k = K$ it is the definition of $\mathcal{C}$. At stage k, compactness gives attainment in Equation (74). Thus $S_{\Pi,k}(x) \le 0$ exactly when one control u makes both the whole cell path safe and every successor satisfy $S_{\Pi,k+1} \le 0$. By the induction hypothesis, this is precisely the definition of $u \in \mathcal{U}^{\text{safe}}_{\Pi,k}(x)$ and $x \in \mathcal{Q}_{\Pi,k}$.

If $(t_{k_n}, x_n) \to (t,x)$ and $x_n \in \mathcal{Q}_{\Pi_n,k_n}$, then $S_{\Pi_n,k_n}(x_n) \le 0$. Local uniform convergence gives $S(t,x) \le 0$, proving Equation (76).

For the reverse inclusion, argue backward in k. At the terminal stage it is immediate. Cover the compact lateral boundary by finitely many local-persistence neighborhoods and shrink their common time horizon to respect $T - t$. The complement of a slightly smaller boundary neighborhood is a compact interior remainder with positive distance from the lateral boundary; bounded speed therefore gives one cell-length bound on which any fixed control stays inside the tube there. Taking the minimum of these finitely many bounds gives one h_0. At a boundary point, a held control in $\mathcal{R}(z)$ keeps every constant-v cell trajectory in $\mathscr{D}$ and ends in $\mathcal{Q}(t_{k+1})$. The induction hypothesis places the endpoint in $\mathcal{Q}_{\Pi,k+1}$. Hence the chosen control is discretely safe and $x \in \mathcal{Q}_{\Pi,k}$. This proves Equation (77). On compact sets, these slice inclusions and sequential outer convergence are the two domain facts used below; no stronger graph-convergence claim is needed. □

The lemma derives the domain consistency needed below.

Lemma 26 (Persistence on exact tube slices)**.** *Let grid points $z_n = (t_{k_n}, x_n) \in \mathscr{D}$ converge to $z = (t,x)$ and let $|\Pi_n| \to 0$. If z is interior and $u \in \mathcal{U}$, or if z is lateral and $u \in \mathcal{R}(z)$, then u is discretely safe from z_n for all sufficiently large n, and every one-cell successor lies in $\mathcal{Q}(t_{k_n+1}) \subseteq \mathcal{Q}_{\Pi_n,k_n+1}$.*

Proof. The interior claim follows from bounded speed. The boundary claim is the local-persistence property. Endpoint membership follows because the entire augmented trajectory remains in $\mathscr{D}$. The final inclusion is Equation (77). □

E.4 HALF-RELAXED LIMITS AND SCHEME CONSISTENCY

For brevity write $U_{\Pi,k} := V^{\uparrow}_{\Pi,k}$ and $L_{\Pi,k} := V^{\downarrow}_{\Pi,k}$. Let $r^h_k(t,x,u,v) = (r^h_{1,k},\dots,r^h_{I,k})(t,x,u,v)$ denote the vector of losses integrated over one grid cell and set $r^h_p := \langle p, r^h_k\rangle$. Uniformly on compact sets, $r^h_k = h\ell(t,x,u,v) + o(h)$.

All remaining losses are uniformly bounded, so the two relaxed grid values share a partition-independent sup-norm bound. They are also uniformly Lipschitz in p: for any fixed strategy pair the expected payoff is linear in p, with typewise coefficients bounded by the same remaining-payoff constant, and taking infima and suprema preserves this bound. No uniform state modulus is used.

For $z = (t,x) \in \mathscr{D}$, define the asymmetric half-relaxed limits

$$\overline{U}^{\mathscr{D}}(z,p) := \limsup_{\substack{|\Pi|\to 0,(t_k,x_\Pi,p_\Pi)\to(t,x,p)\\ x_\Pi\in\mathcal{Q}(t_k)}} V^{\uparrow}_{\Pi,k}(x_\Pi,p_\Pi), \tag{80}$$

$$\underline{L}^{\mathrm{full}}(z,p) := \liminf_{\substack{|\Pi|\to 0,(t_k,x_\Pi,p_\Pi)\to(t,x,p)\\ x_\Pi\in\mathcal{Q}_{\Pi,k}}} V^{\downarrow}_{\Pi,k}(x_\Pi,p_\Pi). \tag{81}$$

The upper limit is restricted to exact-tube points because Lemma 26 supplies all locally admissible branch controls there. The lower limit uses the full discrete tube because its supersolution proof must compare with every discrete successor. Lemma 25 places both limits on the same continuous domain.

Lemma 27 (Belief regularity and terminal trace)**.** *$\overline{U}^{\mathscr{D}}$ is bounded, upper semicontinuous, convex, and uniformly Lipschitz in p. The function $\underline{L}^{\mathrm{full}}$ is bounded, lower semicontinuous, and uniformly Lipschitz in p. Both have terminal trace $\langle p, g(x)\rangle$.*

Proof. Semicontinuity is built into the definitions, and the uniform payoff bound gives boundedness and the common belief-Lipschitz constant. To prove convexity of the upper limit, take one sequence realizing its limsup at $p_\lambda = \lambda p_1 + (1-\lambda)p_2$. Convexity of every finite-grid upper value, applied at the same state sequence, gives

$$V^{\uparrow}_{\Pi}(x,p_\lambda) \le \lambda V^{\uparrow}_{\Pi}(x,p_1) + (1-\lambda)V^{\uparrow}_{\Pi}(x,p_2).$$

Taking the limsup proves the claim; the common Lipschitz bound absorbs the small perturbation of the realizing beliefs. We do not claim that a lower half-relaxed limit of convex functions is convex. The terminal statement follows from uniform continuity of g_i, bounded speed, and the fact that the remaining running loss vanishes with $T - t$. □

For $p \in \Delta(\mathcal{I})$, let

$$T_p\Delta(\mathcal{I}) := \left\{ q \in \mathbb{R}^I : \sum_i q_i = 0, \quad q_i < 0 \Longrightarrow p_i > 0 \right\} \tag{82}$$

be the relative tangent cone and define

$$\lambda_{\min}(M,p) := \min\{q^\top M q : q \in T_p\Delta(\mathcal{I}),\ \|q\| = 1\}, \tag{83}$$

with value $+\infty$ when the unit tangent set is empty. A constrained subsolution is tested with p fixed and satisfies

$$\phi_t + H_{\mathcal{R}}(t,x,p,D_x\phi) \ge 0 \tag{84}$$

at relative upper contacts. A supersolution satisfies

$$\min\big\{\lambda_{\min}(D^2_{pp}\phi,p),\, \phi_t + H(t,x,p,D_x\phi)\big\} \le 0 \tag{85}$$

at relative lower contacts. Equation (84) is the definition used here even if $\mathcal{R}$ is not a continuous correspondence. Its validity for the upper half-limit is proved directly by transporting each fixed $u \in \mathcal{R}(z)$ through local persistence; we do not invoke a generic stability theorem for a discontinuous boundary Hamiltonian.

Lemma 28 (Upper consistency)**.** *$\overline{U}^{\mathscr{D}}$ satisfies Equation* (84)*.*

Proof. Fix p and let $\phi \in C^1$ touch $\overline{U}^{\mathscr{D}}(\cdot, p)$ strictly from above at $z = (t, x)$, $t < T$. The standard perturbed-maximum construction produces exact-tube grid points $z_n = (t_{k_n}, x_n) \to z$ at which the discrete upper value minus ϕ has an exact local maximum on the exact tube slices.

Fix an arbitrary $u \in \mathcal{U}$ if z is interior, or $u \in \mathcal{R}(z)$ if it is lateral. By Lemma 26, the same control is safe for all sufficiently large n and all successors remain on exact tube slices. The upper Bellman operator is no larger than the nonrevealing branch using this particular u. At the discrete maximum,

$$0 \le \max_{v \in \mathcal{V}} \big\{ r_p^{h_n}(t_{k_n}, x_n, u, v) + \phi(t_{k_n+1}, F_{k_n}(x_n, u, v)) - \phi(t_{k_n}, x_n) \big\} + o(h_n). \tag{86}$$

Uniform exact-flow and cost expansions give, after division by h_n,

$$\phi_t(z) + \max_{v \in \mathcal{V}} \Lambda(z, p, D_x\phi(z); u, v) \ge 0.$$

Taking the minimum over the applicable control family proves Equation (84). □

The lower consistency proof uses the dual recursion. For a convex function $W(p)$ on the simplex, write $W^*(y) := \sup_{p \in \Delta(\mathcal{I})} \{\langle p, y\rangle - W(p)\}$. At a boundary belief, the discrete contact must be kept on the same simplex face. Let

$$S(p) := \{i \in \mathcal{I} : p_i > 0\}$$

be the support of p. The minimal face of the simplex containing p and its relative interior are

$$F(p) := \{q \in \Delta(\mathcal{I}) : q_i = 0 \text{ for } i \notin S(p)\}, \qquad \operatorname{ri} F(p) := \{q \in F(p) : q_i > 0 \text{ for } i \in S(p)\}.$$

If L_p is the common belief-Lipschitz constant, replace a lower test ϕ by

$$\phi_M(z, q) := \phi(z, q) - M \sum_{i : p_i = 0} q_i, \qquad M > C_\Delta\big(L_p + \|D_p\phi\|_\infty\big),$$

where C_Δ is any constant dominating the norm change produced by moving unit mass between two simplex coordinates. For the norm used on the simplex,

$$C_\Delta := \max_{i,j \in \mathcal{I}} \|e_i - e_j\|.$$

For Euclidean norm, $C_\Delta = \sqrt{2}$. Moving any mass on a missing coordinate to an index with $p_i > 0$ decreases $W_n - \phi_M$; hence the perturbed discrete minima lie in $F(p)$. The large linear normal term also keeps the conjugate maximizers in that face under $O(h)$ perturbations. This standard face-locking device is needed because positive tangential curvature at p need not persist along sequences approaching from a larger face.

Lemma 29 (Discrete Fenchel contact)**.** *Let $W_n(z, p)$ be convex and uniformly Lipschitz in p, and suppose (z_n, p_n) is a global minimum of $W_n - \phi$ on the relevant discrete graph. Assume $(z_n, p_n) \to (z, p)$, that the minimum is isolated by the standard perturbed-test-function localization. Namely, for each $\delta > 0$ there is $c_\delta > 0$ such that, for all large n and every z_n' used below,*

$$\|q - p_n\| \ge \delta \Longrightarrow (W_n - \phi)(z_n', q) \ge (W_n - \phi)(z_n, p_n) + c_\delta,$$

that $p_n \in \operatorname{ri} F(p)$ and the conjugate maximizers remain in $F(p)$ under the perturbations below, and that

$$\lambda_{\min}(D_{pp}^2\phi(z, p), p) > 0. \tag{87}$$

Set $y_n = D_p\phi(z_n, p_n)$. If $z_n' - z_n = O(h_n)$ and $a_n = O(h_n)$, then

$$W_n^*(z_n', y_n - a_n) - W_n^*(z_n, y_n) \le -\langle p_n, a_n\rangle - \phi(z_n', p_n) + \phi(z_n, p_n) + o(h_n). \tag{88}$$

The remainder is uniform when the displayed $O(h_n)$ bounds are uniform.

Proof. The minimum condition and convexity imply $y_n \in \partial_\Delta W_n(z_n, p_n)$, hence Fenchel equality at the current point. On the fixed face $F(p)$, positive tangential curvature persists near (z, p). Taylor's formula and the minimum property first give a quadratic support near p_n; convexity propagates it along rays to a global support on the compact face. Thus there is one $\delta > 0$ such that

$$W_n(z_n', q) \ge W_n(z_n, p_n) + \phi(z_n', p_n) - \phi(z_n, p_n) + \langle y_n, q - p_n\rangle + \frac{\delta}{2}\|q - p_n\|^2 - Ch_n\|q - p_n\| - o(h_n) \tag{89}$$

for all $q \in F(p)$. The linear error occurs because $D_p\phi(z'_n, p_n) - D_p\phi(z_n, p_n) = O(h_n)$. Subtract the right side from $\langle q, y_n - a_n \rangle$. The face-locking normal term excludes $q \notin F(p)$. On the face, both a_n and the coefficient of the displayed linear error are $O(h_n)$; completing the square shows that every maximizing q lies $O(h_n)$ from p_n and that its excess over $q = p_n$ is $O(h_n^2)$. Fenchel equality at (z_n, p_n, y_n) then yields Equation (88). This is the finite-dimensional quadratic-support step of Cardaliaguet (2007) written in the form needed by the discrete dual backup. □

Lemma 30 (Lower consistency)**.** $\underline{L}^{\mathrm{full}}$ *satisfies Equation* (85)*.*

Proof. Let $\phi \in C^2$ touch $\underline{L}^{\mathrm{full}}$ strictly from below at (z, p), $t < T$. If $\lambda_{\min}(D^2_{pp}\phi(z, p), p) \le 0$, the assertion is immediate. Suppose it is positive.

Replace ϕ by the face-locking modification

$$\phi_M(z, q) := \phi(z, q) - M \sum_{i \notin S(p)} q_i, \qquad M > C_\Delta\big(L_p + \|D_p\phi\|_\infty\big),$$

defined above, and then apply the usual strict global localization. The perturbed-minimum construction gives global minimum points (z_n, p_n) of the localized discrete difference on the viable graphs of the discrete-time games, with

$$(z_n, p_n) \to (z, p).$$

We claim that $p_n \in F(p)$ for all sufficiently large n. Suppose instead that p_n assigns total mass

$$\eta_n := \sum_{j \notin S(p)} (p_n)_j > 0$$

outside $S(p)$. Since the belief domain is the full simplex, this mass may be transferred to any fixed coordinate $i_0 \in S(p)$, producing $\widetilde{p}_n \in F(p)$. By the definition of C_Δ,

$$\|\widetilde{p}_n - p_n\| \le C_\Delta \eta_n.$$

The belief-Lipschitz continuity of the discrete lower value and the local gradient bound for ϕ imply that transferring this mass changes the discrete value and the unmodified test by at most

$$C_\Delta\big(L_p + \|D_p\phi\|_\infty\big)\eta_n.$$

On the other hand, the face-locking term changes by exactly $M\eta_n$:

$$\phi_M(z_n, \widetilde{p}_n) - \phi_M(z_n, p_n) = \phi(z_n, \widetilde{p}_n) - \phi(z_n, p_n) + M\eta_n.$$

Hence, because

$$M > C_\Delta\big(L_p + \|D_p\phi\|_\infty\big),$$

the localized difference "discrete value minus test function" is strictly smaller at $(z_n, \widetilde{p}_n)$ than at (z_n, p_n), contradicting the minimality of (z_n, p_n). Therefore $p_n \in F(p)$.

It remains to show that p_n lies in the relative interior of this face. Since

$$m_p := \min_{i \in S(p)} p_i > 0$$

and $p_n \to p$, for all sufficiently large n,

$$(p_n)_i > \frac{m_p}{2} > 0 \qquad \text{for every } i \in S(p).$$

Together with $p_n \in F(p)$, this yields

$$S(p_n) = S(p), \qquad \text{hence} \qquad p_n \in \operatorname{ri} F(p).$$

Relabel the modified test ϕ_M as ϕ and set $y_n = D_p\phi(z_n, p_n)$.

The lower-game conjugate recursion permits the nonrevealing dual split δ_{y_n} and therefore gives

$$L^*_{\Pi_n, k_n}(x_n, y_n) \le \min_{v \in \mathcal{V}} \max_{u \in \mathcal{U}^{\mathrm{safe}}_{\Pi_n, k_n}(x_n)} L^*_{\Pi_n, k_n+1}\big(F_{k_n}(x_n, u, v), y_n - r_{k_n}(x_n, u, v)\big). \tag{90}$$

Apply Lemma 29 branchwise with $a_n = r_{k_n}(x_n, u, v) = O(h_n)$. After rearrangement and division by h_n,

$$0 \le -\phi_t(z,p) + \limsup_n \min_{v\in\mathcal{V}} \max_{u\in\mathcal{U}^{\mathrm{safe}}_{\Pi_n,k_n}(x_n)} -\Lambda(z_n, p_n, D_x\phi(z_n,p_n); u, v). \tag{91}$$

Enlarging the maximizing set from $\mathcal{U}^{\mathrm{safe}}_{\Pi_n,k_n}(x_n)$ to $\mathcal{U}$ can only increase the right-hand side. Uniform continuity and the bulk Isaacs identity give

$$0 \le -\phi_t(z,p) - H(z,p,D_x\phi(z,p)),$$

which is equivalent to the physical branch of Equation (85). No limit of the safe-control graphs is used. □

Together, Lemmas 27, 28, and 30 establish the regularity, terminal trace, and complementary viscosity inequalities needed for the comparison argument below.

E.5 Koike-Cardaliaguet comparison on the moving tube

Relation to Koike's comparison argument. After augmenting the state by time, Koike's spatial domain is replaced by the fixed space-time tube $\mathscr{D}$. Our local-persistence condition is the sample-and-hold counterpart of Koike's admissibility condition (A3), and Assumption 4 is the corresponding interior-cone and inward-velocity condition (A4). Koike's nondegeneracy condition (A5) is automatic here because every $\gamma \in \mathcal{G}(z,v)$ has time component one and hence $\|\gamma\| \ge 1$. The function Ψ and Equations (92), (93), and (94) are exactly Koike (1995, Lemma 6.1 and Section 7). The lateral-boundary contradiction follows the boundary step of Koike (1995, Theorem 6.2).

Geometric meaning of the tilted two-point penalty. Before we introduce Koike's method in full, we first use the following one-dimensional model to explain its core concept. Let the viable domain be $\mathscr{D} = [0,\infty)$, let $+1$ be an admissible inward velocity. Let w_1 be a bounded upper-semicontinuous constrained subsolution and w_2 a bounded lower-semicontinuous supersolution. In the context of the proof for Theorem 4, these functions are

$$w_1 = \overline{U}^{\mathscr{D}}, \qquad w_2 = \underline{L}^{\mathrm{full}}.$$

The goal is to prove $w_1 \le w_2$. and suppose for contradiction that

$$w_1(0) - w_2(0) = \delta > 0.$$

Consider the doubled functional

$$\Phi_{\alpha,\mu}(x,y) := w_1(x) - w_2(y) - P_{\alpha,\mu}(x,y), \qquad P_{\alpha,\mu}(x,y) := \alpha(x-y)^2 - \mu(x-y).$$

Because $P_{\alpha,\mu}(0,0) = 0$, the original gap remains an admissible competitor:

$$\max_{x,y\ge 0} \Phi_{\alpha,\mu}(x,y) \ge \delta.$$

Let (x_α, y_α) be a localized maximizer. Holding y_α fixed and defining

$$c_\alpha := w_1(x_\alpha) - w_2(y_\alpha) - P_{\alpha,\mu}(x_\alpha, y_\alpha), \qquad \phi_\alpha(x) := w_2(y_\alpha) + P_{\alpha,\mu}(x, y_\alpha) + c_\alpha,$$

the maximality of (x_α, y_α) gives $w_1(x) \le \phi_\alpha(x)$ near x_α, with equality at x_α. Hence ϕ_α genuinely touches w_1 from above. The slope introduced by the penalty is therefore deliberate, but the maximization certifies it as a legitimate viscosity supergradient; it is not asserted to be a classical derivative of w_1. If the upper contact were on the boundary, $x_\alpha = 0$, then $y_\alpha \ge 0$ and

$$\phi'_\alpha(0) = \partial_x P_{\alpha,\mu}(0, y_\alpha) = -2\alpha y_\alpha - \mu \le -\mu.$$

Thus $-\phi'_\alpha(0) \ge \mu$ is the rate at which this upper test function falls when one moves into the viable domain. Dynamic programming along the admissible inward velocity, or equivalently the constrained subsolution inequality, bounds this rate by the running-loss bound M:

$$-\phi'_\alpha(0) \le M.$$

Choosing $\mu > M$ rules out a boundary upper contact. The positive maximizer must therefore have $x_\alpha > 0$, where the constrained and bulk Hamiltonians coincide and the ordinary interior comparison

argument rules out the same positive maximum. This perturbation does not create the positive gap: writing $d = x - y$ gives

$$-\alpha d^2 + \mu d = -\alpha \left(d - \frac{\mu}{2\alpha} \right)^2 + \frac{\mu^2}{4\alpha}.$$

For fixed μ, the largest value contributed by the tilt is $O(\mu^2/\alpha)$, and the usual doubled-variable localization gives $x_\alpha - y_\alpha \to 0$ as $\alpha \to \infty$. The perturbation is therefore vanishing at the level of comparison values and contact locations while its slope contribution remains visible to the viscosity inequality. In the moving-tube proof, the inward direction $+1$ is replaced by $\gamma \in \mathcal{G}(z, \widehat{v}) \cap \mathcal{C}_\theta(\nu)$, and Koike's modified distance Ψ ensures $\langle \gamma, D_z \Psi \rangle \le 0$. Consequently,

$$-\langle \gamma, D_z P_{\alpha,\mu} \rangle = -\alpha \langle \gamma, D_z \Psi \rangle + \mu \langle \gamma, \nu \rangle \ge c\mu,$$

whereas the constrained subsolution inequality bounds the same directional decrease independently of μ. This is the boundary contradiction technique that Lemma 31 and Theorem 5 follow.

Lemma 31 (Koike two-point penalty)**.** *Under Assumption 4, for every lateral-boundary point z_0 and its cone data (r, θ, ν), there are $C_0, C_1 \ge 1$, $\sigma \in (0, 1-\theta)$, and $\Psi \in C^1((\mathscr{D} \cap B_r(z_0))^2)$ such that*

$$C_0^{-1} \|z - z'\|^2 \le \Psi(z, z') \le C_0 \|z - z'\|^2, \tag{92}$$

$$\|D_z \Psi(z, z')\| \le C_1 \|z - z'\|, \qquad D_z \Psi + D_{z'} \Psi = 0, \tag{93}$$

$$\langle \nu', D_z \Psi(z, z') \rangle \le 0 \tag{94}$$

whenever z is lateral and $\nu' \in B_{\theta+\sigma}(\nu)$. With

$$P_{\alpha,\mu}(z, z') := \alpha \Psi(z, z') - \mu \langle \nu, z - z' \rangle, \tag{95}$$

a positive doubled maximum cannot have its upper contact on the lateral boundary once μ and then α are sufficiently large.

Proof. The stated quadratic, oblique, and exact paired-gradient properties are precisely Koike's geometric penalty lemma under the local-uniform interior cone condition (Koike, 1995). We verify the boundary-exclusion step for our operator. If the upper contact z were lateral, the sign-transformed subsolution operator is, by the constrained Isaacs identity,

$$\min_{v \in \mathcal{V}} \max_{u \in \mathcal{R}(z)} K(z, p, \partial_t P_{\alpha,\mu}, D_x P_{\alpha,\mu}; u, v) \le 0.$$

Hence some $\widehat{v}$ makes the corresponding inequality hold for every $u \in \mathcal{R}(z)$. Taking a convex combination preserves it up to the bounded running-cost term. By Equation (55), choose $\gamma \in \mathcal{G}(z, \widehat{v})$ in the inward cone at the actual contact. Write $\gamma = s\nu'$ with $s > 0$ and $\nu' \in B_\theta(\nu) \subset B_{\theta+\sigma}(\nu)$. Continuity and Equation (94) give

$$-\langle \gamma, D_z P_{\alpha,\mu} \rangle = -\alpha \langle \gamma, D_z \Psi \rangle + \mu \langle \gamma, \nu \rangle \ge c\mu$$

for a local $c > 0$. The nondegeneracy condition and boundedness of the velocities keep the scale s, and hence c, uniformly away from zero. The viscosity inequality bounds the same quantity by a constant independent of μ, contradicting sufficiently large μ. This is Koike's A3-A5 boundary mechanism. Nondegeneracy prevents the cone velocity from collapsing. After fixing μ, the quadratic bounds send $z - z'$ to zero as $\alpha \to \infty$. □

Theorem 5 (State-constrained dual comparison)**.** *Let w_1 be bounded and upper semicontinuous, convex and uniformly Lipschitz in p, and suppose that it satisfies the constrained subsolution inequality in Equation* (84)*. Let w_2 be bounded and lower semicontinuous and suppose that it satisfies the supersolution inequality with the belief-convexity obstacle in Equation* (85)*. Under Assumptions 3 and5, if their terminal traces are uniformly attained and ordered, then*

$$w_1 \le w_2 \qquad \text{on } \mathscr{D} \times \Delta(\mathcal{I}).$$

Proof of Theorem 5. The comparison statement remains valid in the semicontinuous form needed here: let w_1 be bounded upper semicontinuous, convex and uniformly Lipschitz in p, and satisfy Equation (84); let w_2 be bounded lower semicontinuous and satisfy Equation (85). Assume terminal ordering and the uniform terminal trace used in Lemma 27.

Suppose $\sup(w_1 - w_2) > 0$. Standard localization produces a strict positive maximum near either an interior point or one lateral-boundary point z_0. Write s for the time component of z'. For $\sigma, \beta, \eta > 0$ and the penalty from Lemma 31, maximize

$$w_1(z,p) - w_2(z',p) - P_{\alpha,\mu}(z,z') - \sigma(T-s) + \beta(\|p\|^2 - 1) - \eta\big(\|z - z_0\|^2 + \|z' - z_0\|^2\big) \quad (96)$$

over the localized doubled domain. Choose the parameters in the order η, σ, β, then μ, and finally α. Choose σ and β smaller than the localized positive gap. Because $\|p\|^2 - 1 \le 0$ on the simplex, uniform terminal trace and terminal ordering exclude a contact on $t = T$; the quadratic localization keeps the contacts in the neighborhood where the cone data are valid. After fixing the Koike parameter μ, let $\alpha \to \infty$. Comparison with the diagonal competitor at the localized maximizer gives the standard strengthened doubling estimate

$$z_\alpha - z'_\alpha \to 0, \qquad \alpha\Psi(z_\alpha, z'_\alpha) \to 0, \qquad \alpha\|z_\alpha - z'_\alpha\|^2 \to 0.$$

Lemma 31 rules out a lateral upper contact. At the remaining interior upper contact, $\mathcal{R} = \mathcal{U}$ and $H_{\mathcal{R}} = H$.

It remains to force the physical branch of the supersolution obstacle. Choose a relative subgradient $q_\alpha \in \partial_\Delta w_1(z_\alpha, p_\alpha)$, which exists because w_1 is convex and Lipschitz in p. Maximality in the common belief variable implies that w_2 is touched from below by a function with p-dependence

$$\langle q_\alpha, p - p_\alpha\rangle + \beta(\|p\|^2 - \|p_\alpha\|^2).$$

Its relative tangential Hessian is $2\beta I$, so the convexity branch in Equation (85) is strictly positive. The supersolution must therefore satisfy its physical HJI inequality.

The derivative of $-\sigma(T-s)$ adds $+\sigma$ to the lower test's time derivative. Applying the two viscosity inequalities and subtracting gives

$$\sigma \le H(z_\alpha, p_\alpha, \xi_\alpha) - H(z'_\alpha, p_\alpha, \xi'_\alpha) + o(1). \quad (97)$$

Koike's paired-gradient identities Equation (93), the strengthened doubling estimate, and the structural modulus Equation (65) make the right-hand side $o(1)$ as $\alpha \to \infty$, contradicting $\sigma > 0$. Sending the remaining localization parameters to zero gives $w_1 \le w_2$ on $\mathscr{D} \times \Delta(\mathcal{I})$. □

E.6 Completion and precise scope

Apply Theorem 5 to $\overline{U}^{\mathscr{D}}$ and $\underline{L}^{\mathrm{full}}$. Lemmas 28 and 30 give

$$\overline{U}^{\mathscr{D}} \le \underline{L}^{\mathrm{full}}. \quad (98)$$

Choose an exact-tube sequence realizing $\overline{U}^{\mathscr{D}}(z,p)$. The finite-grid ordering $L_{\Pi,k} \le U_{\Pi,k}$ and the fact that the lower half-limit ranges over all discrete-tube sequences give the reverse inequality

$$\underline{L}^{\mathrm{full}}(z,p) \le \overline{U}^{\mathscr{D}}(z,p).$$

Hence the two half-relaxed limits agree. Their common value is continuous, and the usual compactness argument upgrades subsequential convergence to uniform convergence on compact exact-tube slices. More explicitly, along every exact-tube sequence converging to (z,p),

$$\underline{L}^{\mathrm{full}}(z,p) \le \liminf L_{\Pi,k} \le \limsup L_{\Pi,k} \le \limsup U_{\Pi,k} \le \overline{U}^{\mathscr{D}}(z,p) \le \underline{L}^{\mathrm{full}}(z,p).$$

The third inequality is the finite-grid ordering; the outer inequalities are the definitions of the two half-relaxed limits and comparison.

Suppose, to the contrary, that there are partitions Π_n, exact-tube grid points $x_n \in \mathcal{Q}(t_{k_n})$, and beliefs p_n such that

$$V^{\uparrow}_{\Pi_n,k_n}(x_n,p_n) - V^{\downarrow}_{\Pi_n,k_n}(x_n,p_n) \ge \epsilon > 0. \quad (99)$$

After extracting a subsequence, the arguments and the two payoff sequences converge. By the definitions of the half-relaxed limits,

$$\limsup_n V^{\uparrow}_{\Pi_n,k_n}(x_n,p_n) \le \overline{U}^{\mathscr{D}}(z,p),$$
$$\underline{L}^{\mathrm{full}}(z,p) \le \liminf_n V^{\downarrow}_{\Pi_n,k_n}(x_n,p_n).$$

Together with Equation (98), these inequalities contradict Equation (99). Therefore

$$\sup_k \sup_{(x,p)\in\mathcal{Q}(t_k)\times\Delta(\mathcal{I})} \left[V^{\uparrow}_{\Pi,k}(x,p) - V^{\downarrow}_{\Pi,k}(x,p)\right] \longrightarrow 0. \tag{100}$$

In particular, the conclusion holds at every fixed $x_0 \in \mathcal{Q}(0)$. The finite-grid sandwich then gives convergence of the simultaneous value at $(0, x_0, p)$ and vanishing exploitability of the embedded leader pair there.

Equation (100) does not take the supremum over all $x \in \mathcal{Q}_{\Pi,k}$. Grid-only viable states can form a shrinking boundary layer in which one-cell-safe controls have no limiting locally admissible counterpart. Neither Koike's condition nor fixed-grid repairability rules out a value boundary layer there. The stronger conclusion

$$\sup_k \sup_{(x,p)\in\mathcal{Q}_{\Pi,k}\times\Delta(\mathcal{I})} [V^{\uparrow}_{\Pi,k} - V^{\downarrow}_{\Pi,k}] \to 0 \tag{101}$$

follows if one additionally assumes a no-upper-boundary-layer property, equivalently that the upper half-relaxed limit over the full discrete graphs does not exceed $\overline{U}^{\mathscr{D}}$. A mesh-uniform state modulus is one sufficient condition, but it is not imposed here.

Finally, Assumption 4 is substantive. A maximal viability tube can have characteristic or merely tangential moving fronts, at which the strict cone-intersection condition may fail. The empirical diagnostics in Appendix F are designed to reveal such regions rather than silently treating the condition as automatic.

F Empirical Diagnostics for the Smooth Koike Condition

Lemma 24 provides a smooth sufficient condition for Assumptions 3 and 4. We test its pointwise ingredients on the learned zero surface, conditional on local C^1 regularity; numerical gradient agreement does not prove that regularity or identify the learned surface with the exact viability boundary.

The archived panel contains 53 anchors: 10 legacy development and 43 random confirmation samples. The latter comprise 22 exact-safe-prefix and 21 adversarial-bank samples. The planned primitive/time coverage panel selected no anchors, and the random panel lacked its requested boundary-neighborhood stratum. At the projected points, 49 have speed-lower and four have lane as the largest normalized primitive. These labels do not imply active physical constraints; none meets the recorded primitive near-boundary threshold. All 53 archived screen records were reconciled without missing inputs or integrity errors.

At a projected anchor $z_j = (t_j, x_j)$, let u_j be the selected P1 control and define

$$m_j := -\max_{v\in\mathcal{V}} D_z\widehat{b}_\vartheta(z_j)\widehat{f}(z_j, u_j, v), \qquad \widehat{f} = (1, f).$$

For the car's control-affine dynamics and independent box controls, the pointwise control extrema are evaluated by exact box support. The archived checks assess projection accuracy, nonzero normals, and multiscale finite-difference agreement. Reconciliation keeps u_j fixed and recomputes P2's box extremum for each saved finite-difference normal.

If $m_j > 0$ on a C^1 regular patch, continuity and compactness of $\mathcal{V}$ imply that some $r_j > 0$ satisfies

$$\max_{v\in\mathcal{V}} D_z\widehat{b}_\vartheta(z)\widehat{f}(z, u_j, v) \le -m_j/2 \qquad \text{whenever } \|z - z_j\| < r_j.$$

Thus strict pointwise robust inwardness supplies local common-control persistence without a numerical radius sweep. Extending this conclusion to the uniform hypotheses requires regularity and appropriate uniform control of the relevant boundary, which finite samples do not establish. At HJI continuation points, the optimal margin may be zero; the direct admissibility and cone conditions must then be assessed instead.

The reconciled outcomes are:

HJI classification	Pointwise margin	Count
Consistent	Positive	34
Inconsistent	Positive	13
Inconsistent	Negative	4
Inconsistent	Geometry inconclusive	2

All 47 positive margins reproduce under the archived numerical checks. No HJI-consistent sampled pointwise counterexample was observed. HJI consistency and the margin use the same PDE expression and are not independent confirmations. The 34 consistent positive rows therefore provide conditional sampled support for regular-point inwardness, not global verification of Assumptions 3 and 4. The incomplete neighborhood-suffix run produced no complete saved radius result.

For Assumption 2, we supplemented the earlier primitive-constraint MFCQ tests with 12 predetermined state/hold cases. We included primitive safety along 32-step RK4 trajectories and learned endpoint viability, $\widehat{b}_\vartheta(t_{k+1}, F_k(x,u,v)) \le 0$. Both players hold their controls constant within the interval. Continuous P2 searches augment the finite scenario set before it is frozen. A box-relative MFCQ linear program then tests for one control adjustment that strictly decreases all nearly active constraints.

Ten cases passed with the endpoint condition nearly active, including five with both path and endpoint constraints nearly active. Normalized MFCQ margins ranged from 0.002341 to 0.430244. Passing cases also passed multiscale derivative checks, independent P2 attacks, and 64-step replay. One case exhausted the disturbance-search budget and one had no strictly feasible control found; both remain unresolved. These results support sampled repairability of the learned successor surrogate, not a uniform error bound for the exact recursive safe-control correspondence.

Reproducibility. The audit code, frozen configurations, sample-selection records, and row-level results, including failed and inconclusive cases, are provided in the anonymous repository: `https://anonymous.4open.science/r/iclr_36156`. The repository includes commands for reproducing the reported summaries from the archived evidence.

G Hexner's Game: Exact Relaxation Width and Refinement Oracle

This appendix starts from the reduced information-control form used in the Hexner benchmark of Ghimire et al. (2026). The derivation of the three discrete announcement-order operators and their relaxation width is specific to the present paper.

G.1 Reduced stage game

Let the private type be

$$\theta \in \{-1, 1\}, \qquad p = \Pr(\theta = -1).$$

After the LQ state/control reduction, the information-dependent stage payoff is

$$A_k(a-\theta)^2 - B_k(b-\theta)^2, \qquad A_k, B_k \ge 0, \tag{102}$$

where a and b are scalar information controls. In the continuous-time where $a, b \in [-1, 1]$ are scalar information controls. The reduced benchmark has zero terminal continuation payoff. In the continuous-time parameterization,

$$A_k = \int_{t_k}^{t_{k+1}} d_1(s)\,\mathrm{d}s, \qquad B_k = \int_{t_k}^{t_{k+1}} d_2(s)\,\mathrm{d}s.$$

Define the posterior mean and variance

$$m(p) := \mathbb{E}_p[\theta] = 1 - 2p, \qquad \nu(p) := \mathrm{Var}_p(\theta) = 4p(1-p). \tag{103}$$

For every scalar z,

$$\mathbb{E}_p[(z-\theta)^2] = (z - m(p))^2 + \nu(p). \tag{104}$$

Hence the minimizing P1 action and maximizing P2 action on a known posterior branch are both $m(p)$, and the residual expected squared error is $\nu(p)$.

Let $Z(p)$ be a continuation value. A Bayes-plausible split is a finite or measure-valued family $\{(\lambda_r, q_r)\}$ with $\sum_r \lambda_r q_r = p$.

G.2 Upper operator

In the upper relaxation, P2 sees the realized P1 branch. On branch r, both current controls equal $m(q_r)$, so the optimized stage contribution is $(A_k - B_k)\nu(q_r)$. Minimizing over posterior splits yields

$$\boxed{(\mathcal{T}_k^{\uparrow} Z)(p) = \mathrm{vex}_q \left[Z(q) + 4(A_k - B_k)q(1-q) \right](p).} \tag{105}$$

Here vex is the lower convex envelope in the belief.

G.3 Simultaneous reference operator

In the simultaneous game, P2 chooses one current b before observing the realized branch. Bayes plausibility gives

$$\sum_r \lambda_r \mathbb{E}_{q_r}[(b-\theta)^2] = \mathbb{E}_p[(b-\theta)^2]. \tag{106}$$

Thus P2's common best response is $b = m(p)$ and contributes $-B_k \nu(p)$. P1 still chooses $a_r = m(q_r)$ branchwise. Therefore

$$\boxed{(\mathcal{T}_k^{G} Z)(p) = -4B_k p(1-p) + \mathrm{vex}_q \left[Z(q) + 4A_k q(1-q) \right](p).} \tag{107}$$

G.4 Robust-lower operator

In the lower relaxation, P2 announces b. For any Bayes-plausible P1 split, the P2 term remains

$$-B_k \mathbb{E}_p[(b-\theta)^2],$$

independent of the split. P1's branchwise action and continuation optimization are independent of b. Observing b therefore does not change P1's reduced problem, and P2 again chooses $b = m(p)$. Hence

$$\boxed{\mathcal{T}_k^{\downarrow} Z = \mathcal{T}_k^{G} Z} \tag{108}$$

for every continuation Z. Consequently,

$$V_{\Pi}^{\downarrow} = V_{G\Pi} \tag{109}$$

for every partition. The entire relaxation width is the upper current-response advantage.

G.5 Closed-form coefficient recursion

All three values are multiples of $\nu(p)$. Write Because the terminal continuation is zero, all three values are multiples of $\nu(p)$. Write

$$V_k^{\uparrow}(p) = \nu(p) u_k, \qquad V_k^{G}(p) = V_k^{\downarrow}(p) = \nu(p) c_k. \tag{110}$$

The lower convex envelope of $4Cq(1-q)$ is itself when $C \le 0$ and is zero when $C \ge 0$. Therefore, with $u_K = c_K = 0$,

$$u_k = \min\{u_{k+1} + A_k - B_k, 0\}, \tag{111}$$

$$c_k = -B_k + \min\{c_{k+1} + A_k, 0\}. \tag{112}$$

These scalar recursions are exact finite-grid oracles for the numerical solver.

G.6 Reveal-time formula and linear mesh-width bound

Define

$$S_r := \sum_{k=0}^{r-1} (A_k - B_k), \qquad S_0 = 0. \tag{113}$$

If the upper game pools before stage r and fully reveals at stage r, its coefficient is S_r. Hence

$$u_0 = \min_{0 \le r \le K} S_r. \tag{114}$$

In the simultaneous/lower game, revelation at $r < K$ leaves P2 uninformed for the current stage, adding the coefficient $-B_r$. The option of never revealing has coefficient S_K, so

$$c_0 = \min \left\{ S_K,\ \min_{0 \le r < K} (S_r - B_r) \right\}. \tag{115}$$

Therefore

$$\boxed{\Delta_{\mathrm{LF}}(\Pi) = 4p_0(1-p_0)(u_0 - c_0).} \tag{116}$$

If the minimum in Equation (115) is S_K, then $u_0 \le S_K$. Otherwise, if it is $S_r - B_r$, then $u_0 \le S_r$. In either case,

$$0 \le u_0 - c_0 \le \max_k B_k, \tag{117}$$

and thus

$$\boxed{0 \le \Delta_{\mathrm{LF}}(\Pi) \le 4p_0(1-p_0) \max_k B_k.} \tag{118}$$

For a bounded rate d_2, $B_k \le \|d_2\|_\infty h_k$, so $\Delta_{\mathrm{LF}} = O(|\Pi|)$.

When the upper and simultaneous games choose the same reveal cell r, the width is exactly

$$\Delta_{\mathrm{LF}} = 4p_0(1-p_0)B_r. \tag{119}$$

For $p_0 = 1/2$, this reduces to B_r.

H Numerical Studies on Hexner's Game

H.1 Hexner numerical protocol and finite-grid reference

Game and registered runs. The support sweep uses the original $T = 1$ double-integrator game with $K = 3$ uniform intervals and discrete-time dynamics. The initial position-velocity states are $(-1, 0, 0, 0)$ and $(1, 0, 0, 0)$, and the targets are $(0, \theta_i)$, where

$$\theta_i = -1 + \frac{2(i-1)}{I-1}, \qquad p_{0,i} = \frac{1}{I}.$$

Both control boxes are $[-16, 16]^2$, with no state or terminal-velocity constraints. Running matrices are $R_1 = \mathrm{diag}(0.05, 0.025)$ and $R_2 = \mathrm{diag}(0.05, 0.10)$; running cost has the factor $1/2$ in Equation (17), and terminal position squares have full weight. We register $I \in \{2, 3, 4, 6\}$, $M = 1, \ldots, I+2$, and random seeds $0, \ldots, 19$.

Exact finite-grid benchmark. For a belief p, define the posterior mean and variance of the scalar type by

$$m(p) := \sum_i p_i \theta_i, \qquad \nu(p) := \sum_i p_i \theta_i^2 - m(p)^2.$$

A standard finite-horizon linear-quadratic Riccati recursion gives the certainty-equivalent control

$$-L_{j,k} x_{j,k} - f_{j,k} m(p),$$

where $L_{j,k}$ is the state-feedback gain and $f_{j,k}$ is the coefficient multiplying the posterior mean. Let $P_{j,k+1}$ be the complete-information value Hessian and let B_k be the exact input matrix for a control held constant over interval k. The remaining coefficient multiplying posterior variance is

$$D_{j,k} := \tfrac{1}{2} f_{j,k}^\top S_{j,k} f_{j,k}, \qquad S_{j,k} := h_k R_j + \mathsf{B}_k^\top P_{j,k+1} \mathsf{B}_k.$$

After subtracting the prior-averaged complete-information value, the upper and simultaneous/lower values reduce to

$$V_k^{\uparrow}(p) = \nu(p)u_k, \qquad V_k^{G}(p) = V_k^{\downarrow}(p) = \nu(p)c_k,$$

where

$$u_K = c_K = 0, \qquad u_k = \min\{0, u_{k+1} + D_{1,k} - D_{2,k}\}, \qquad c_k = -D_{2,k} + \min\{0, c_{k+1} + D_{1,k}\}.$$

To see why these scalar recursions are exact, let $\{(\lambda_r, q^r)\}$ be a Bayes-plausible posterior split. The law of total variance gives

$$\sum_r \lambda_r \nu(q^r) = \nu(p) - \mathrm{Var}_{r\sim\lambda}(m(q^r)).$$

Pooling leaves the expected posterior variance equal to $\nu(p)$, whereas full revelation reduces it to zero. Therefore minimizing $C\sum_r \lambda_r \nu(q^r)$ selects pooling when $C \le 0$ and full revelation when $C \ge 0$. For the reported instances, this recursion pools through the first two intervals and fully reveals at $t = 2/3$.

Atoms and observation. At each public node, an $I \times M$ row-stochastic matrix specifies the type-conditioned probabilities of M shared current-control atoms. Zero-weight atoms are allowed. All allocated control slots satisfy $\|u^m - u^{m'}\|_2 \ge \varepsilon$ for $m \ne m'$, with $\varepsilon = 1$. Pooling is represented by one positive-weight action and inactive padding; it does not require distinct actions for the different types. Exactly identical observed actions must share one posterior and continuation. The implementation separates all slots, including inactive ones, rather than conditioning separation on posterior differences. It therefore imposes a stronger strategy restriction than separation of informative actions alone. Inactive padding is feasible for every registered M and does not alter the physical value. This restriction leaves the delayed-revelation reference feasible and permits multistage revelation.

Optimization and independent response audits. All computations use float64. Initialization screens four seeded stage/partition candidates, fits bounded physical controls by L-BFGS, and refits P1 controls by separation-constrained SLSQP. The existing first-order schedule uses up to 80 outer steps and 10 inner steps, with local tangent-halfspace projections for control separation. Analytical values, controls, and revelation times are used only for evaluation. Five fresh continuous-response starts audit both the initialization and the final iterate, independently of the leader's atom budget; the better feasible audited candidate is retained. Each run has a 30-minute cap including initialization and audit, with no outcome-dependent extension. All values, residuals, effective supports, runtimes, and termination states are retained.

Summary of results. For each seed we compute

$$e_{I,M} := \frac{\widehat{V}_{0,I,M}^{\uparrow} - V_{0,I}^{\uparrow}}{S_I}, \qquad S_I := \max\{1, |V_{0,I}^{\uparrow}|, |V_{0,I}^{\downarrow}|\},$$

Figure 1(a) reports the arithmetic mean and sample standard deviation of $e_{I,M}$ across the 20 seeds. All 240 outcomes with $M \ge I$ recover the reference, whereas all 220 with $M < I$ remain above tolerance.

I LONG-HORIZON DUAL COMPRESSION CONTROL

I.1 GAME AND EXACT BENCHMARK

Nature draws $\theta \in \{-1, +1\}$ with probability $1/2$. P1 observes θ and selects $s_{1:L} \in [-1, 1]^L$; P2 observes the public signal history and selects $b \in [-1, 1]$. The payoff is

$$J_\theta(s_{1:L}, b) = \frac{A}{L}\sum_{\ell=1}^{L}(s_\ell - \theta)^2 - B(b - \theta)^2, \qquad A = 1.25, \quad B = 1, \tag{120}$$

with P1 minimizing and P2 maximizing. The strict choice $A > B$ avoids the uninformative nonunique equilibria that occur at $A = B$.

Let

$$\bar{s} := \frac{1}{L}\sum_{\ell=1}^{L} s_\ell, \qquad \sigma_s^2 := \frac{1}{L}\sum_{\ell=1}^{L}(s_\ell - \bar{s})^2.$$

For the two types, write the dual target as $\widehat{p} = (\widehat{p}_-, \widehat{p}_+)$. Translation invariance makes only their difference relevant, so define

$$d_\ell := \widehat{p}_{-,\ell} - \widehat{p}_{+,\ell}.$$

At the uniform prior, symmetry permits equal initial targets. Subtracting their common value fixes the translation gauge and gives $d_0 = 0$.

A signal s_ℓ contributes the type-loss difference

$$\frac{A}{L}\big[(s_\ell + 1)^2 - (s_\ell - 1)^2\big] = \frac{4A}{L} s_\ell.$$

Hence the dual difference updates as

$$d_\ell = d_{\ell-1} - \frac{4A}{L} s_\ell, \qquad d_L = -4A\bar{s}.$$

At P2's final decision, the two typewise dual shortfalls are equalized when

$$d_L + B\big[(b+1)^2 - (b-1)^2\big] = d_L + 4Bb = 0.$$

Therefore

$$b^\star(s_{1:L}) = \mathrm{clip}\left(-\frac{d_L}{4B}, -1, 1\right) = \mathrm{clip}\left(\frac{A}{B}\bar{s}, -1, 1\right),$$

where

$$\mathrm{clip}(a, -1, 1) := \min\{1, \max\{-1, a\}\}.$$

We now verify that this feedback guarantees value zero. The identity

$$\frac{1}{L}\sum_{\ell=1}^{L}(s_\ell - \theta)^2 = \sigma_s^2 + (\bar{s} - \theta)^2$$

separates variation within the history from its mean. In the region $|A\bar{s}/B| \le 1$, substitution of $b^\star = A\bar{s}/B$ gives

$$J_\theta = A\sigma_s^2 + (A - B)\left(1 - \frac{A}{B}\bar{s}^2\right) \ge 0.$$

In the region $A\bar{s}/B \ge 1$, $b^\star = 1$. For $\theta = 1$,

$$J_1 = A\sigma_s^2 + A(\bar{s} - 1)^2 \ge 0.$$

For $\theta = -1$,

$$J_{-1} = A\sigma_s^2 + A(\bar{s} + 1)^2 - 4B.$$

Over $\bar{s} \ge B/A$, the final two terms are minimized at $\bar{s} = B/A$, where they equal

$$\frac{(A - B)^2}{A} \ge 0.$$

The region $A\bar{s}/B \le -1$ is symmetric. Thus P2 guarantees $J_\theta \ge 0$ for both types and every signal history. Conversely, P1 can choose $s_\ell = \theta$ for every ℓ; then $\bar{s} = \theta$, $b^\star = \theta$, and $J_\theta = 0$. The game value is therefore zero.

If P2 observes only the constant physical state, its action cannot depend on the signal history. The constant action $b = 0$ guarantees $J_\theta \ge -B$ because the first term is nonnegative. Conversely, if P1 chooses $s_\ell = \theta$, the prior-averaged payoff against a constant b is

$$-\frac{B}{2}\big[(b+1)^2 + (b-1)^2\big] = -B(1 + b^2) \le -B.$$

Hence the best state-only action is $b = 0$, with restricted value $-B = -1$. The construction therefore has genuine state aliasing: all histories reach the same physical state, yet histories with different means require different P2 actions. The scalar dual is an exact sufficient memory. Perfect history is also sufficient because it can reconstruct $\bar{s}$; the experiment tests the difficulty of learning this compression, not whether history contains the information.

I.2 Representations and policy classes

P2 never observes θ. The compared inputs are

$$o_{\text{state}} = \varnothing, \qquad o_{\text{dual}} = d_L, \qquad o_{\text{raw}} = s_{1:L}, \qquad o_{\text{GRU}} = (s_1, \dots, s_L). \tag{121}$$

The raw baseline contains exactly the L varying signals: constant physical state columns, padding, and an all-one validity mask are omitted. The GRU receives the same one-scalar sequence and has hidden width 16. This keeps the history baselines from carrying redundant inputs.

All policies use two SiLU hidden layers followed by a scalar head and a $\tanh$ output, so $b \in [-1, 1]$. A common sign-equivariant construction is imposed by antisymmetrizing the pre-$\tanh$ output g_ϕ:

$$\pi_\phi(o) = \tanh\left(\frac{g_\phi(o) - g_\phi(-o)}{2}\right), \qquad \pi_\phi(-o) = -\pi_\phi(o). \tag{122}$$

For the state mode the input is empty, so this construction fixes $\pi_\phi = 0$. Dense weights use Glorot-uniform initialization, with the final weight matrix multiplied by 0.1 and biases initialized to zero. The GRU uses the standard reset, update, and candidate equations, Glorot-uniform weights, zero reset/candidate biases, update bias one, and zero initial hidden state. All calculations use 64-bit floating point. Architectures were chosen deterministically to be nearest to $16{,}384$ trainable parameters:

Representation	Horizon L	Dense widths	GRU width	Parameters
state	all	$(127, 127)$	–	16,511
dual	all	$(126, 126)$	–	16,381
raw history	8	$(123, 123)$	–	16,483
raw history	16	$(119, 119)$	–	16,423
raw history	32	$(112, 112)$	–	16,465
raw history	64	$(99, 99)$	–	16,435
GRU history	all	$(115, 115)$	16	16,275

I.3 Payoff-only objective and DS-GDA settings

For R response histories per type and representation χ, the optimized finite-response game is

$$\min_{\phi} \max_{\{s^{\theta,r}\}} -\frac{1}{2R} \sum_{\theta \in \{-1,+1\}} \sum_{r=1}^{R} J_\theta\left(s^{\theta,r}, \pi_\phi^\chi(o_\chi(s^{\theta,r}))\right), \qquad s^{\theta,r} \in [-1, 1]^L. \tag{123}$$

Thus the policy minimizes negative payoff and each type-specific history maximizes it. Each complete solve uses one projected gradient step per DS-GDA iteration, $12{,}000$ iterations, proximal weights $(r_1, r_2) = (1, 2)$, and anchor rates $(\beta, \mu) = (0.1, 0.1)$. The actor step is 10^{-3} for state, dual, and raw history and 3×10^{-3} for the GRU.

I.4 Results

All models learn the correct P2 strategy at $L = 16$ and $L = 32$, confirming that the raw and GRU baselines can represent and optimize useful security strategies. At $L = 64$, the dual passes every seed, raw history fails every seed, and the GRU fails $11/20$. Median resources before the two final audits were:

Representation	Passes	DS-GDA gradient evaluations	Wall time (s)	Memory
dual	$20/20$	12,061	7.86	1
raw history	$0/20$	48,244	25.39	64
GRU history	$9/20$	48,244	140.81	16

J Constrained Feedback-Security Solvers

This appendix specifies the implementation of Equation (15). We first define the asymmetric time grids and the extension to P2-owned constraints, then describe the policy parameterizations, response exchange, optimization, and independent audits. All computations use `float64`. Unless

stated otherwise, the reported tolerances diagnose a local feedback-security solution; they do not certify global optimality of the nonconvex actor problems or the learned viability tube.

J.1 Asymmetric time grids and leader lifting

Write

$$\Pi = (\Pi^{\mathrm{sig}}, \Pi^{\mathrm{ctl}}), \qquad \Pi^{\mathrm{sig}} \subseteq \Pi^{\mathrm{ctl}} \subseteq \Gamma,$$

where Π^{sig} contains the leader's signaling knots, Π^{ctl} contains all knots at which it may change its physical control, and Γ is the finer response grid. At a knot in $\Pi^{\mathrm{ctl}} \setminus \Pi^{\mathrm{sig}}$, the leader may change its control but introduces no primal signal or dual split. At a response-only knot in $\Gamma \setminus \Pi^{\mathrm{ctl}}$, it holds its preceding control, while the responder may update using the information then available. Controls remain simultaneous at every shared knot. The lift $\mathcal{L}_\Gamma$ implements these holds exactly.

Using a finer grid for the responder makes the best-response audit more stringent and closer to the intended fine-grid game. Keeping signaling and most leader updates on the sparser grids controls the cost of the leader tree, while response solves remain comparatively inexpensive. All security values reported in the experiments use fixed actors and independently recomputed responses on the stated audit grid.

J.2 P2-owned state constraints

Suppose P2 owns a type-independent state x^2 whose within-stage dynamics depend only on (x^2, v), with constraint set

$$\mathcal{C}^2 := \left\{x^2 : d_\ell(x^2) \le 0, \quad \ell = 1, \dots, m_2\right\}.$$

Let $X_k^2(s; x^2, v)$ and $F_k^2(x^2, v)$ denote its exact within-stage flow and endpoint, respectively, and define

$$\Gamma_{\ell,k}^2(x^2, v) := \max_{0 \le s \le h_k} d_\ell\big(X_k^2(s; x^2, v)\big).$$

Starting from $\mathcal{Q}_K^2 = \mathcal{C}^2$, define backward

$$\begin{aligned} \mathcal{V}_k^{\mathrm{own}}(x^2) &:= \left\{v \in \mathcal{V} : \Gamma_{\ell,k}^2(x^2, v) \le 0 \ \forall \ell, \quad F_k^2(x^2, v) \in \mathcal{Q}_{k+1}^2\right\}, \\ \mathcal{Q}_k^2 &:= \left\{x^2 : \mathcal{V}_k^{\mathrm{own}}(x^2) \neq \varnothing\right\}. \end{aligned} \tag{124}$$

P1 must remain safe against every control that P2 may admissibly select:

$$\mathcal{U}_k^{\mathrm{rob}}(x) := \left\{ u \in \mathcal{U} : \begin{array}{ll} \Gamma_{j,k}(x, u, v) \le 0 & \forall j, \quad v \in \mathcal{V}_k^{\mathrm{own}}(x^2), \\ F_k(x, u, v) \in \mathcal{Q}_{k+1}^{\mathrm{rob}} & \forall v \in \mathcal{V}_k^{\mathrm{own}}(x^2) \end{array} \right\}, \tag{125}$$

with

$$\mathcal{Q}_K^{\mathrm{rob}} := \left\{x \in \mathcal{C} : x^2 \in \mathcal{Q}_K^2\right\}, \qquad \mathcal{Q}_k^{\mathrm{rob}} := \left\{x : x^2 \in \mathcal{Q}_k^2,\ \mathcal{U}_k^{\mathrm{rob}}(x) \neq \varnothing\right\}.$$

We assume that $\mathcal{V}_k^{\mathrm{own}}$ and the resulting robust safe-control correspondence are nonempty, compact valued, and continuous on their viable domains. Under these conditions, the finite-grid value, atomicity, and security arguments apply with $\mathcal{V}$ replaced by $\mathcal{V}_k^{\mathrm{own}}(x^2)$. The continuous-time limit additionally requires the corresponding repairability, boundary-geometry, and Isaacs assumptions for the joint viability problem.

Numerically, both the P2 actor and every P2 response are restricted to $\mathcal{V}_k^{\mathrm{own}}(x^2)$, while every P1 actor control and typewise P1 best response is projected into $\mathcal{U}_k^{\mathrm{rob}}(x)$. The present experiments impose no P2-owned constraint, so

$$\mathcal{V}_k^{\mathrm{own}}(x^2) = \mathcal{V}.$$

Constraints depending on P1's private type, state, or simultaneous control would change the information and viability structure and are outside the present formulation.

J.3 Feedback parameterizations

P1 actor. The P1 actor is one shared MLP evaluated at every P1 information state. Its input is the game-provided normalized physical feature map concatenated with the public belief p. At a signal stage with I types it outputs $I \times I$ logits and I P1 controls. Row-wise softmax gives the typewise signal law, with probability floor 10^{-6}, and Bayes masses are propagated exactly. At a control-only leader stage the actor outputs one public control and no signal. At response-only knots this control is held. Physical controls are mapped into their primitive boxes by a hyperbolic tangent and then passed through the robust safe-control projection.

P2 actor. The P2 actor is a shared MLP with input

$$\big(\phi(t,x),\ (z_1 - z_I, \dots, z_{I-1} - z_I)/s_z\big), \qquad s_z = 8.$$

The root is normalized by $z_I = 0$ and its $I - 1$ free coordinates are optimized jointly with the network. For each of at most I active atoms, the output contains one logit, $I - 1$ raw dual-memory offsets, and one P2 control. The offsets are bounded by s_z, augmented with a zero final coordinate, and centered by their atom-weighted mean. Therefore

$$\sum_a \lambda_a z^a = z$$

holds to floating-point precision. Inactive heads create no edge, response variable, probability-floor mass, or objective term; the active weights have zero probability floor. P2 controls are mapped into their primitive box and, when present, projected into $\mathcal{V}_k^{\mathrm{own}}(x^2)$.

Both actors use two hidden layers of width 64, SiLU activation, and a linear output layer; the output weights are initialized at scale 10^{-3}. For the constrained car study, $\phi(t,x)$ has 13 entries:

$$\frac{t}{T},\quad \frac{T-t}{T},\quad 1,\quad \frac{x_1^1}{L},\ \frac{x_2^1}{L},\quad \sin\theta^1,\ \cos\theta^1,\ \frac{s^1}{s^1_{\max}},\quad \frac{x_1^2}{L},\ \frac{x_2^2}{L},\quad \sin\theta^2,\ \cos\theta^2,\quad \frac{s^2}{\max\{s^2_{\max}, a^2_{\max}/\kappa^2\}},$$

where L is the street half-length and κ^2 is P2's drag coefficient. The game adapter, rather than the generic solver, owns this feature map and the control dimensions.

The P2 root memory is initialized from the gauge-fixed centered vector of typewise payoffs produced by the P1 solve. This is a warm start only. The physical master in Equation (15) may leave several dual-memory representations equivalent, and neither the learned root nor its descendants are reported as Fenchel subgradients unless an independent contact condition is verified.

J.4 Response exchange and finite-archive masters

For a P1 actor η_θ and retained P2-response archive $\mathcal{A}_2$, the exact finite master is

$$\min_\theta \max_{r \in \mathcal{A}_2} J_\Gamma^G(\mathcal{L}_\Gamma \eta_\theta, r). \tag{126}$$

For a P2 actor $\zeta_{\vartheta,z}$ and one retained P1 archive $\mathcal{A}_{1,i}$ per type, it is

$$\max_{\vartheta,z} \sum_i p_i \min_{r \in \mathcal{A}_{1,i}} J_{i,\Gamma}^G(r, \mathcal{L}_\Gamma \zeta_{\vartheta,z}). \tag{127}$$

Positive-prior types have separate response solves. The smooth actor phase uses log-sum-exp for the maximum in Equation (126) and negative log-sum-exp for each typewise minimum in Equation (127), both at temperature 10^{-3}. It is followed by an exact nonsmoothed epigraph or hypograph SLSQP polish. The reported stationarity residual is the minimum-norm convex combination of the gradients of responses within $\max\{10^{-7}, 5 \times 10^{-3}\}$ of the corresponding active envelope.

Each exchange cycle performs 600 Adam steps with learning rate 10^{-3} and global gradient clipping at 20, followed by at most 100 L-BFGS-B iterations with relative function tolerance 10^{-13} and gradient tolerance 10^{-9}. The nonsmoothed SLSQP polish uses at most 100 iterations and function tolerance 10^{-11}. Actor exchange runs for at least 8 cycles and may continue to 16 in blocks of 4. Each physical response is solved by box-constrained L-BFGS-B for at most 500 iterations, using the incumbent and 3 random starts during exchange and 5 random starts in the final audit; its projected box-gradient residual is stored.

At most 16 responses per archive are retained, and at most 16 are active in an actor master. After generating a fresh response, the incumbent and the proposal are evaluated on the same enlarged, nonsmoothed archive. A P1 proposal is accepted only if that archive maximum does not increase; a P2 proposal is accepted only if its prior-weighted archive minimum does not decrease. Only after this comparison is the archive pruned, retaining the most adversarial responses under the accepted actor. The stopping tolerance for a fresh response is $0.005S_{\text{task}}$ and the actor stationarity target is 10^{-3}. Production runs use seeds $0, 1, 2$. The final stored upper and lower objectives are the nonsmoothed finite-archive envelopes, accompanied by their log-sum-exp error bounds and independently seeded fresh responses.

J.5 Separate neural reachability solve

The P1 safe-control layer is trained before either game solver. Let $\ell_\epsilon(t, x)$ be a smooth conservative aggregation of the normalized primitive P1 constraints, and let

$$\widehat{B}_\vartheta(t, x) = \ell_\epsilon(t, x) + (T - t)\,\mathrm{softplus}(N_\vartheta(t, x)). \tag{128}$$

This architecture enforces $\widehat{B}_\vartheta(T, x) = \ell_\epsilon(T, x)$ and $\widehat{B}_\vartheta \geq \ell_\epsilon$. For control-affine dynamics $f = f_0 + F_1 u + F_2 v$, the training Hamiltonian is

$$H(t, x, D_x B) = D_x B^\top f_0 + \min_{u \in \mathcal{U}} D_x B^\top F_1 u + \max_{v \in \mathcal{V}_k^{\text{own}}(x^2)} D_x B^\top F_2 v. \tag{129}$$

For the current car study $\mathcal{V}_k^{\text{own}} = \mathcal{V}$ is a box, so both control optimizations in Equation (129) are evaluated by exact box support. A non-box P2-owned viable-control set would require its own exact support oracle. The residual is the HJI variational inequality

$$\max\{\partial_t \widehat{B}_\vartheta + H, \ell_\epsilon - \widehat{B}_\vartheta\} = 0. \tag{130}$$

The production reachability network has three sine-activated hidden layers of width 256. Training uses Adam at learning rate 10^{-4}, global gradient clip 1, batch size 8192, and backward curriculum fractions $(.05, .1, .2, .4, .7, 1)$. Each curriculum stage receives 5000 steps, giving 30,000 initial steps; a seed missing its validation gate resumes exactly to 60,000 steps. Every batch contains 50% uniform-domain, 25% primitive-boundary, and 25% learned-zero-level samples. The loss combines the mean residual with the largest 5% residuals. Every 250 steps, 512 high-residual counterexamples are selected from a pool four times larger. Validation uses 4096 held-out points, checkpoints are written every 1000 steps and at curriculum boundaries, and seeds $0, 1, 2$ use disjoint validation and final audit splits. Seed selection uses only the validation score; the selected seed must then pass the immutable audit rather than being replaced post hoc.

For the constrained car, the primitive P1 inequalities are lower and upper speed, street signed distance, and disk separation from P2. They are scaled, respectively, by the speed range, street half-width, and squared disk radius before smooth aggregation. The frozen game has horizon $T = 3$, three intents, uniform prior, no running loss, and a terminal position/velocity-vector loss. The pre-registered initial P1-P2 gaps are $4, 5, 6$; the pre-game procedure uses the first gap passing the reachability, safety, and complete-information gates.

J.6 Robust projection and safety audits

At every actor or response evaluation, the learned barrier generates a local affine inequality in P1's two-dimensional control. The implementation retains exact support over P2's control box and adds sample-and-hold inequalities at the interval midpoint and endpoint. It evaluates 4 midpoint scenarios and 16 endpoint vertex-pair scenarios, intersects the resulting half-spaces with P1's primitive control box, and projects the raw P1 control onto the polygon by active-set enumeration. An empty polygon terminates the rollout as infeasible. The default optimization derivative is a projected-control semi-gradient: it differentiates through the active projection with respect to the raw control while stopping the locally rebuilt polygon coefficients. Every forward rollout and safety audit nevertheless recomputes the full polygon. Full differentiation through its geometry is available only as an ablation.

The minimum pre-solve diagnostic is $\widehat{B}_\vartheta(0, x_0) < -\delta_{\text{safe}}$. Additional gates evaluate normalized primitive residuals, projection feasibility, relative robust MFCQ near active constraints, sampled

neighborhood inward selectors, adversarial safety-only rollouts, the inner 80% of the HJI training box, and dense-RK4 reconstruction discrepancy. These are empirical reached-set audits; without a uniform approximation error bound for $\widehat{B}_\vartheta$, they are not proof-level safety certificates.

The leader has K signal intervals; its control grid and the training response grid both have $2K$ intervals, while the immutable verification response grid has $4K$ intervals. Thus the development run uses $K = 6$, Γ_{12}, and immutable Γ_{24}, while the target run uses $K = 8$, Γ_{16}, and immutable Γ_{32}. Verification may evaluate fixed actors and generate fresh responses, but may not train an actor, accept an atom, or rewrite a checkpoint.