Continuous-Time Stochastic Games with Imperfect Public Monitoring*

Benjamin Bernard†

This paper characterizes the set of perfect public equilibrium payoffs, the set of Markov perfect equilibrium payoffs, and a simple class of state-order dependent equilibrium payoffs in continuous-time stochastic games with finitely many states and a publicly observable Brownian signal. Contrary to many discrete-time methods, the characterization does not rely on a convergence to a stationary distribution of the underlying state process. As a consequence, (i) the correspondence from initial state to equilibrium payoffs is preserved, which allows us to address a richer set of research questions, (ii) the characterization is possible for any level of discounting, allowing us to study non-Markovian behavior away from the patient limit, and (iii) the characterization applies to non-irreducible games, opening up a wide range of applications to economic analysis.

Keywords: Stochastic games, continuous time, imperfect monitoring, perfect public equilibrium, Markov-perfect equilibrium, computation of equilibria.

1 Introduction

Stochastic games encompass a wide range of economics applications, in which the environment responds stochastically to the agents' actions.[1] Despite their broad applicability, our understanding of equilibrium behavior in discrete-time stochastic games is largely confined to two special classes of games. The first class are *absorbing games*, in which all but the initial

*I am grateful for comments and insightful discussions with Andrew Clausen, Eduardo Faingold, Johannes Hörner, Michihiro Kandori, Marcin Pęski, Yuliy Sannikov, and many participants of conferences and seminars. Financial support from the Ministry of Science and Technology in Taiwan through grant 108-2410-H-002-249 and CRETA by the Ministry of Education in Taiwan through grant 109L900203 is gratefully acknowledged.

†Department of Economics, University of Wisconsin–Madison; bbernard3@wisc.edu.

[1]Stochastic games have been used to model dynamic competition with inventories (e.g., Kirman and Sobel (1974)) or entry and exit dynamics (e.g., Ericson and Pakes (1995)), common resource extraction (e.g., Levhari and Mirman (1980)), economic growth (e.g., Bernheim and Ray (1989)), strategic pricing (e.g., Bergemann and Välimäki (1996)), social uprisings (e.g., Acemoglu and Robinson (2001)), research and development (e.g., Grossman and Shapiro (1987)), and strategic experimentation (e.g., Keller, Rady, and Cripps (2005)).

state are absorbing. Characterizing equilibria in these games is tractable because continuation values in absorbing states can be determined separately, hence the analysis reduces to a single incentive constraint in the initial state. The second class are *irreducible games*, in which the state process under any stationary strategy profile is an irreducible Markov chain. The distribution over states thus converges to a steady-state distribution, which makes the limit equilibrium payoff set independent of the initial state as players become arbitrarily patient. This steady-state structure allows one to adapt linear programming techniques from repeated games to characterize limit equilibrium payoffs. However, many economically relevant games are neither absorbing nor irreducible. When the game is not irreducible or when players are impatient, continuation values may vary substantially across states, hence the discrete-time techniques do not generalize well to such cases. This paper shows that in continuous time, it is possible to characterize equilibrium payoffs in substantially more general stochastic games.

We study a class of two-player stochastic games with finitely many states and imperfect public monitoring in a continuous-time setting. The state process is a publicly observed Markov process whose state transition intensity may depend on the players' actions. In addition, players observe a Brownian public signal whose expected change depends on the current actions. We extend the techniques in Sannikov (2007) to this class of stochastic games and to behavior strategies, yielding a differential characterization of extremal perfect public equilibrium (PPE) payoffs. Because players are impatient, the initial state plays a non-negligible role. Our main results characterizes the PPE payoff correspondence $y \mapsto \mathcal{E}(y)$, which maps each initial state y to the set of all PPE payoffs that are attainable from that state. Because the continuation value after a transition to state y' must lie in $\mathcal{E}(y')$, the characterizations across states are generally coupled. If they are, we show how the PPE payoff correspondence can be approximated through an iterative procedure for numerical implementation. To the best of our knowledge, this paper is the first to extend the techniques in Sannikov (2007) to behavior strategies. The ability to mix is particularly important in applications where stationary or Markovian strategies are of interest, such as, for example, in models of strategic experimentation. We characterize the PPE payoff correspondence for both pure and behavior strategies, and provide sufficient conditions for the two to coincide.

The boundary of $\mathcal{E}(y)$ may contain four types of payoffs, depending on which information is used locally to provide incentives. At *stationary payoffs*, players condition only on state transitions, hence actions and continuation values are locally constant. Extremal stationary pay-

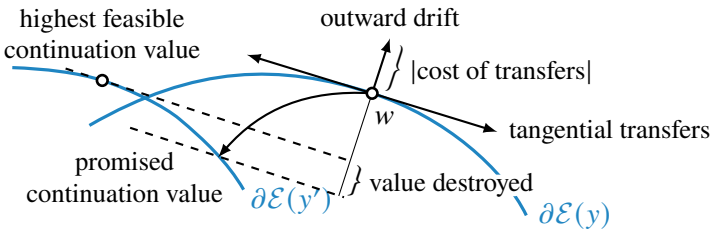

Figure 1: At a non-stationary boundary payoff w, incentives can be provided by destroying value after a state transition or through tangential transfers. The Brownian transfers generate an outward drift, whose strength is the effective cost of providing incentives through those transfers.

offs are characterized through a local inclusion property that resembles a differential inclusion. If players condition on state transitions and the passage of time, the boundary takes two different forms, depending on which feasibility constraint binds. At *continuation-binding payoffs*, feasibility in the continuation state binds, and the boundary is characterized via a linear program that resembles the characterization of PPE payoffs in discrete time. The similarity stems from the inherent discreteness of state transitions in games with finitely many states. If feasibility in the current state binds, the boundary solves a first-order ordinary differential equation (ODE), called the *state-transition optimality equation*. Lastly, if players condition on all available information, including the public signal, the boundary solves a second-order ODE simply called the *optimality equation*. The two optimality equations extend analogous equations in Sannikov (2007) and Bernard (2024) to stochastic games and behavior strategies.

Stationary payoffs that attain extremal equilibrium payoffs are constrained-efficient if the highest-feasible continuation values after a state transition provide sufficient incentives; see Figure 1. All other boundary payoffs are constrained-inefficient. The optimality equations capture how players balance instantaneous payoffs, the impact of the chosen actions on future opportunities, and the cost of providing incentives. Incentives are provided either by destroying value after state transitions or by transferring value tangentially to $\mathcal{E}(y)$ based on realizations of the public signal. Because the players are impatient and the public signal is noisy, relying on tangential transfers is not costless. The optimality equation shows that the cost of tangential transfers is proportional to their quadratic variation and the curvature of $\partial\mathcal{E}(y)$. The boundary is characterized by the state-transition optimality equation when such transfers are too costly to be used effectively. A third form of inefficiency, which is exclusive to non-irreducible stochastic games, arises at continuation-binding payoffs. If the payoff difference between states is large relative to the players' patience, maximizing the probability

of transitions to favorable states is the players' primary concern. In extreme cases, only a single action profile may be enforceable even if it is constrained-inefficient in the current state.

In applied work, simpler classes of equilibria are often of interest. Our methodology also yields a characterization of payoffs attainable in *state-order dependent* PPE, which condition on the public history only through the order of states visited, but neither on the times of state transitions nor on the time spent in each state. A special case of state-order dependent PPE are Markov-perfect equilibria (MPE), for which we provide a simple verification result.

We apply our framework to a winner-take-all, two-stage patent race, similar in spirit to Grossman and Shapiro (1987). Two competing research labs choose how many resources to invest, which determines the arrival rate of breakthroughs. Each lab must achieve two breakthroughs to secure the patent. While intermediate breakthroughs are publicly observed, the investment levels are private. Instead, the labs observe a noisy Brownian signal that is informative about the chosen research effort. Our main insight is that when the marginal cost of research increases with intensity, the two labs engage in a tacit non-compete agreement along the efficient frontier of the PPE correspondence. They take turns exerting minimal effort, using the public signal to monitor compliance. This is in stark contrast to the MPE in Grossman and Shapiro (1987), where research intensities increase with research progress.

On the technical side, this paper extends the continuous-time methods of Sannikov (2007) and Bernard (2024) to stochastic games and behavior strategies. Both papers characterize the PPE payoff set in repeated games; the former with Brownian information, the latter with both Poisson and Brownian information. Since the state process in our model generates Poisson-type information, the signal structure matches Bernard (2024). However, the players face different constraints on how they can use this information, hence most proofs do not readily extend to our setting. While players are free to ignore Poisson signals in a repeated game, players in a stochastic game must react to state transitions. As mentioned above, this may restrict the set of enforceable action profiles and cause inefficiencies. The extension to behavior strategies has a twofold impact on enforceability. If a player mixes, the set of their own incentives shrink because they have to be indifferent between all pure actions in the support. However, mixing may enlarge the set of enforceable responses by their opponents by smoothing the impact of the opponent's deviations on public signals. The latter property implies that enforceability for behavior strategies is a set of bilinear constraints rather than a set of affine constraints as it is the case for pure strategies. This poses a significant challenge

to establishing regularity of the optimality equations—a key ingredient in the proof—which has to be established jointly in actions and incentives. We show that the solutions to the optimality equations are sufficiently well-behaved if either (i) players are restricted to pure strategies, (ii) state transitions satisfy a product structure, or (iii) each state has a unique successor state and each stage game has two pure actions. These conditions guarantee that if two pure actions of a player are statistically distinguishable against some action of their opponent, they can be distinguished also against nearby mixed actions. In addition, we derive monotonicity results with respect to the game primitives not present in earlier work.

The most closely related paper in discrete time is Pęski and Wiseman (2015), which establishes a folk theorem in a related model as the length of the time period shrinks to zero. In their framework, the intensity of state transitions and the discount rate are fixed per unit of time, while the informativeness of the public signal is fixed per period. As the period length vanishes, the state transitions converge to Poisson processes and the public signal becomes perfectly informative. In this limit, providing incentives through the public signal is costless, hence the only binding constraints are feasibility and individual rationality. This is different from our setting where providing incentives through the public signal is costly.

Aside from the folk theorems in Dutta (1995) and Pęski and Wiseman (2015), the techniques used to study stochastic games in discrete time often rely on a very specific form of the underlying state process. The analysis of irreducible games often relies on convergence to a steady state distribution (e.g., Hörner, Sugaya, Takahashi, and Vieille (2011) and Fudenberg and Yamamoto (2011)), whereas the analysis of absorbing games often hinges on the fact that there is only a single non-absorbing state (e.g., Kohlberg (1974), Mertens and Neyman (1981), and Mertens, Neyman, and Rosenberg (2009)). By contrast, our characterization is valid for rather general state processes. For pure-strategy PPE, the only restriction is that the state space is finite. For behavior-strategy PPE, we impose additional restrictions on how the players' actions affect transition intensities, but not on which state transitions are feasible.

The techniques in this paper do not readily extend to infinite state spaces or to games with perfect monitoring. If the state process follows a diffusion process, techniques to characterize PPE have been developed in concurrent work by Faingold and Sannikov (2020). For continuous-time stochastic games with perfect monitoring, see Neyman (2017).

Because differential equations lend themselves well to numerical methods, our ODE characterization also contributes to the literature on computing equilibrium payoffs. Existing

algorithms focus on settings with either patient players (e.g., Hörner et al. (2011)) or perfect monitoring (e.g., Yeltekin, Cai, and Judd (2017) or Abreu, Brooks, and Sannikov (2020)). Our techniques provide an algorithm for impatient players with imperfect public monitoring.

The remainder of the paper is organized as follows. The model is presented in Section 2 and illustrated with a simple running example of a political regime-change game in Section 3. Section 4 states our assumptions and describes enforceability and self-generation in our setting. Section 5 provides a heuristic derivation of the PPE payoff correspondence. Section 6 characterizes stationary, MPE, and state-order dependent PPE payoffs. Section 7 completes the argument sketched in Section 5 and it contains the main result. Section 8.1 applies the techniques to a two-stage patent race, Section 8.2 discusses computational aspects, and Section 8.3 studies monotonicity of the PPE payoff correspondence. Section 9 concludes. The proofs are contained in Appendices A–C and Online Appendices F–I.

2 Model

Two players $i = 1, 2$ play a continuous-time stochastic game with a finite state space $\mathcal{Y}$. In each state $y \in \mathcal{Y}$, each player i has a finite set $\mathcal{A}^i(y)$ of pure actions available with typical element a^i. We denote by $\Delta\mathcal{A}^i(y)$ the set of player i's mixed actions with typical element α^i. We denote by $\mathcal{A}(y) := \mathcal{A}^1(y) \times \mathcal{A}^2(y)$ the set of all pure action profiles $a = (a^1, a^2)$. If pure action profile a is played in state y, each player i receives an expected flow payoff $g^i(y, a)$, extended to mixed actions by bilinearity. The expected flow payoff is the analogue of the ex-ante stage game payoff in discrete time and is unobserved by the players. The state evolves according to a Markov process $S = (S_t)_{t\geq 0}$ that transitions from state y to y' with instantaneous intensity $\lambda_{y,y'}(a) \geq 0$ when pure action profile a is played, extended to mixed actions by bilinearity. It will be convenient to denote by $\lambda(y, \alpha) := (\lambda_{y,y_1}(\alpha), \ldots, \lambda_{y,y_{|\mathcal{Y}|}}(\alpha))^\top$ the column vector of all transition intensities under action profile α.

Actions are only imperfectly observable through their impact on the state process S and a publicly observable signal X. In state y, the signal X is a $d(y)$-dimensional drifted Brownian motion with volatility matrix $\sigma(y)$ and drift rate $\mu(y, a)$, again extended to mixed actions by bilinearity. Without loss of generality, the signal is driven by a state-independent Brownian motion Z of dimension $d \geq \max_y d(y)$ and $\sigma(y) \in \mathbb{R}^{d(y)\times d}$ is of full rank $d(y)$. The full rank condition guarantees that $\sigma(y)$ admits a right inverse, which we denote by $\sigma^{-1}(y)$ with slight abuse of notation. The public information $\mathcal{F}_t$ at time t is a σ-algebra that contains the

history of the processes S and Z up to time t, together with sufficiently rich information for public randomization. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ the filtration of public information.

Definition 2.1. A *(public) strategy* A^i of player i is an $\mathbb{F}$-predictable process with values in $\Delta\mathcal{A}^i(S)$ or the limit of a sequence of such processes.[2] A *pure strategy* takes values in $\mathcal{A}^i(S)$.

The public signal is defined pathwise via $\mathrm{d}X_t = \sigma(S_t)\,\mathrm{d}Z_t$, and the impact of a strategy profile $A = (A^1, A^2)$ is modeled through a change of probability measure to $Q^A = (Q^A_t)_{t\geq 0}$, under which players observe the game. Under Q^A, the public signal takes the form

$$X_t = \int_0^t \mu(S_s, A_s)\,\mathrm{d}s + \int_0^t \sigma(S_s)\,\mathrm{d}Z^A_s, \tag{1}$$

where $Z^A = Z - \int \sigma^{-1}(S_s)\mu(S_s, A_s)\,\mathrm{d}s$ is a Q^A-Brownian motion by Girsanov's theorem. Moreover, under Q^A, state transitions to state y at time t occur with instantaneous intensity $\lambda_{S_{t-},y}(A_t)$. The details of the mathematical foundation are relegated to Appendix A.

Remark 2.1. Simon and Stinchcombe (1989) and Neyman (2017) demonstrate that continuous-time strategies may not have well-defined outcomes if they depend on actions chosen in the immediate past. The requirement that strategies condition only on public information resolves this issue, hence each strategy is admissible in the sense of Neyman (2017).

Both players discount future payoffs at discount rate $r > 0$. Player i's discounted expected future payoff or player i's *continuation value* under strategy profile A at time $t \geq 0$ is

$$W^i_t(S_t, A) = \int_t^\infty r\mathrm{e}^{-r(s-t)}\mathbb{E}_{Q^A_s}\left[g^i(S_s, A_s)\,\middle|\,\mathcal{F}_t\right]\mathrm{d}s. \tag{2}$$

The factor r normalizes the total discounting weights to 1. As a consequence, all feasible payoffs lie in the convex hull of stage-game payoffs $\mathcal{V}_0 := \mathrm{conv}\{g(y,a) \mid a \in \mathcal{A}(y), y \in \mathcal{Y}\}$.

Definition 2.2. A strategy profile A is a *perfect public equilibrium (PPE)* for initial state S_0 if for each player i and all possible deviations $\tilde{A}^i$,

$$W^i(y, A) \geq W^i(y, (\tilde{A}^i, A^{-i})) \tag{3}$$

holds almost everywhere on the set $\{S = y\}$ for each state y.[3]

[2] Because the times of state transitions are totally inaccessible, they cannot be approximated by a sequence of predictable times. Thus, no strategy can anticipate state changes. Allowing the limit of a sequence of $\mathbb{F}$-predictable processes closes the strategy space by allowing non-predictable public randomization.

[3] Formally, (3) holds for $P \otimes Lebesgue$-almost every $(\omega, t) \in \{S = y\}$, where P is a reference measure with

We denote by $\mathcal{E}(r; y)$ the set of all payoff pairs that are achievable by a PPE with initial state y. For most of the paper, the discount rate is fixed and we may simply write $\mathcal{E}(y)$. We denote by $\mathcal{E}(r)$ or simply by $\mathcal{E}$ the family $(\mathcal{E}(r; y))_{y\in\mathcal{Y}}$ of equilibrium payoff sets. We denote by $\mathcal{E}^p(r; y)$ and $\mathcal{E}^p(r)$ the corresponding pure-strategy PPE payoff sets.

By deviating to the strategy of myopic best replies to their opponent's strategy, each player i can ensure an equilibrium payoff of at least their lowest minmax payoff

$$\underline{v}^i = \min_{y\in\mathcal{Y}} \min_{\alpha^{-i}\in\Delta\mathcal{A}^{-i}(y)} \max_{a^i\in\mathcal{A}^i(y)} g^i(y, (a^i, \alpha^{-i})).$$

Let $\underline{v}^i_p$ denote the lowest pure-action minmax payoff, and denote by $\mathcal{V}_0^*$ and $\mathcal{V}_0^{p*}$ the set of all payoff pairs $w \in \mathcal{V}_0$ where w^i is at least $\underline{v}^i$ and $\underline{v}^i_p$, respectively. The subscript 0 indicates that $\mathcal{V}_0$ and $\mathcal{V}_0^*$ are rather naive bounds for the sets of feasible and individually rational payoffs because they do not condition on the intensity of state transitions. Section D.2 contains tighter bounds for the sake of computation, but the exposition in the main text is simpler with $\mathcal{V}_0^*$.

In certain applications, simpler classes of equilibria may be of interest, in which players condition on only a subset of the available information. Markov strategies condition only on the current state, and state-order dependent strategies condition only on the order of states visited, but neither on the time of state transitions nor on the time spent in each state.

Definition 2.3. PPE A is a *Markov perfect equilibrium (MPE)* if there exist $\alpha^i_*(y) \in \Delta\mathcal{A}^i(y)$ for each state y such that $A^i = \alpha^i_*(S)$ for $i = 1, 2$. PPE A is *state-order dependent* if there exist $\hat{\alpha}^i(y_1, \dots, y_k) \in \Delta\mathcal{A}^i(y_k)$ for each sequence of states $(y_1, \dots, y_k)$ such that $A^i = \hat{\alpha}^i(\hat{S})$ for $i = 1, 2$, where $\hat{S}_t$ is the sequence of states visited by S up to and including time t.

Let $\mathcal{E}_m(r; y)$ and $\mathcal{E}_s(r; y)$ denote the set of payoffs achievable in MPE and state-order dependent PPE, respectively. Denote by $\mathcal{E}_r(r; y)$ the state-order dependent PPE payoffs with public randomization. We use superscript p to denote their restrictions to pure strategies.

3 Example of a regime-change game

Consider a political regime-change game in which two ideological groups $i = 1, 2$ compete for power. State y_i indicates that group i holds office. The incumbent i chooses between a benevolent rule with a relatively fair division of the economic surplus and an authoritarian

respect to which all Q^A are defined; see Appendix A. Each family Q^A is absolutely continuous with respect to P, hence (3) also holds for $Q^A \otimes Lebesgue$-almost every $(\omega, t) \in \{S = y\}$: no player has a profitable deviation on the path. The $P \otimes Lebesgue$ qualifier implies that no player has a profitable deviation off the path either.

	S	I
B	4, 2	2, −3
A	3, 0	1, −5

$g(y_1, a)$

	S	I
B	3, −1	1, −1
A	3, −3	1, −3

$\mu(y_1, a)$

	S	I
B	1	9
A	0.5	4.5

$\lambda_{y_1,y_2}(a)$

Figure 2: Game primitives in state y_1 in the regime-change game; primitives in y_2 are symmetric.

rule that appropriates a large share of the surplus, denoted by $\mathcal{A}_i(y_i) = \{B, A\} \triangleq \{0, 1\}$. The non-incumbent $-i$ decides whether to submit to the incumbent or to instigate a revolution, denoted by $\mathcal{A}_{-i}(y_i) = \{S, I\} \triangleq \{0, 1\}$. Unlike many related models in the literature, the division of surplus is stochastic and the implemented policies are only imperfectly observed. Specifically, consider an orthogonal decomposition of welfare $X = (X^1, X^2)$ into components X^1, stemming from issues on which the two groups agree on, and X^2, stemming from issues on which the two groups have diametrically opposite views. Group 1 enjoys negative changes in X^2 and group 2 enjoys positive changes in X^2 so that the ex-post payoffs of group i are $\mathrm{d}V_t^i = \mathrm{d}X_t^1 + (-1)^i \mathrm{d}X_t^2 - 3A_t^i \,\mathrm{d}t$, where we normalize the flow cost of the cooperative action to 0 and the non-cooperative action has a flow cost of 3. The two groups' actions affect the drift rates of X^1 and X^2 via $\mu^1(y_i, a) = 3 - 2a^{-i}$ and $\mu^2(y_i, a) = (-1)^i(1 + 2a^i)$. The expected flow payoffs are then $g^i(y, a) = \mu^1(y, a) + (-1)^i \mu^2(y, a) - 3a^i$. Aggregate welfare $2X^1$ depends only on the non-incumbent's actions, and the expected division of total welfare is the incumbent's choice. The frequency of state transitions is

$$\lambda(y_i, a) = \frac{1}{2}(2 - a_i)(1 + 8a_{-i});$$

see Figure 2 for the game primitives in matrix form. An authoritarian regime halves the frequency of revolutions by maintaining an oppressive state apparatus at a flow cost of 3. Both groups discount the future at discount rate $r = 1$. Either group can ensure a payoff of at least 0 by choosing to cooperate when they are not in charge. The set $\mathcal{V}_0^*$ is thus the set of all payoff pairs of $\mathrm{conv}(g(\mathcal{Y}, \mathcal{A}))$ in the first quadrant; see Figure 3.

Each stage game has a unique Nash equilibrium in which the two groups cooperate and both receive their highest-possible stage-game payoff. If a single group were to rule indefinitely, the efficient PPE would involve permanent cooperation. However, the possibility of a regime change undermines this outcome: the opposition is tempted to instigate a revolution in order to seize power, and the incumbent is tempted to install an authoritarian regime to

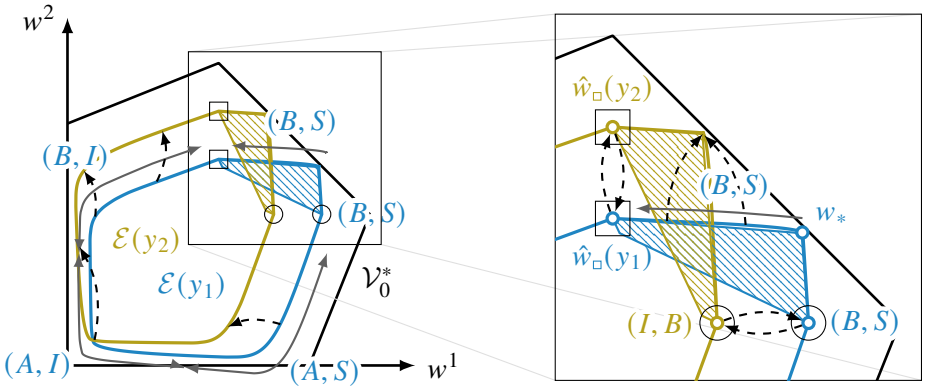

Figure 3: The PPE payoff correspondence $\mathcal{E}$in the regime-change game. The hatched areas are $\mathcal{E}_r$. Payoffs $\mathcal{E}_m$ attained in the two Markov-perfect equilibria are highlighted with squares and circles, respectively. Grey arrows indicate the directions, in which the continuation value can move in absence of revolutions. Dashed arrows show the change in continuation value when a revolution occurs.

remain in power. Figure 3 shows the three different equilibrium payoff correspondences $\mathcal{E}_m$, $\mathcal{E}_r$, and $\mathcal{E}$ computed with the main results of the paper for discount rate $r = 1$. There are two Markov-perfect equilibria, in which one group always cooperates and the other group cooperates only when in power. The revolutions instigated by one group reduce the long-run value of incumbency for both, thereby reducing the profitability of deviations. However, an unnecessarily high amount of value is destroyed by instigating revolutions in every second term. In a state-order dependent PPE, a higher sum of payoffs can be attained through less frequent revolutions; see Figure 3. In the most efficient such PPE, the groups randomize every third term between absorption in the MPE or continued cooperation for another three terms.

Outside the set of state-order dependent PPE payoffs, the boundary is characterized by the optimality equations. At such boundary payoffs, the generically unique action profile that solves the optimality equation must be played in equilibrium; see Figure 3. The efficient frontier is a solution to the state-transition optimality equation, which means that the group's strategies locally condition only on the occurrence of revolutions and time, not on welfare realizations. The PPE that attains the payoff pair w_* with the highest sum of payoffs begins with a cooperative phase, during which the continuation value drifts along the boundary towards the MPE payoff $\hat{w}_\square(y_1)$. In absence of revolutions, $\hat{w}_\square(y_1)$ is reached at a deterministic time t_0, at which point group 2 begins to instigate a revolution and players follow the MPE from there on. While the non-incumbent is willing to cooperate initially, doing so forever is not incentive compatible. If a state transition occurs before time t_0, the continuation value jumps

to the symmetric segment on the boundary of $\mathcal{E}(y_2)$, where both groups cooperate until $\hat{w}_\circ(y_2)$ is reached, and so on. Either group's utility is higher than in any state-order dependent PPE payoff in $\mathcal{E}_r(y_1)$ because cooperation is entirely front-loaded. Other extremal PPE payoffs are solutions to the optimality equation, where the strategies locally do condition on welfare.

4 Construction of equilibria

To construct perfect public equilibria, we first describe how the players' continuation value depends on the publicly available information. For that purpose, define the processes

$$J_t^y = \sum_{0<s\leq t} 1_{\{S_s=y, S_{s-}\neq y\}} \tag{4}$$

that count the number of transitions to state y. Its instantaneous intensity is $\lambda_{S_{t-},y}(A_t)$. The increment $\mathrm{d}J_t^y$ is 1 if a state transition to state y occurs at time t and it is 0 otherwise. We obtain the following stochastic differential representation of the continuation value.

Lemma 4.1. *A semimartingale W is the continuation value of a strategy profile A for initial state y_0 if and only if W is bounded and for $i = 1, 2$, it holds that*

$$\begin{aligned} \mathrm{d}W_t^i = {} & r(W_t^i - g^i(S_t, A_t))\,\mathrm{d}t + r\beta_t^i\,(\sigma(S_t)\,\mathrm{d}Z_t - \mu(S_t, A_t)\,\mathrm{d}t) \\ & + r\sum_{y\in\mathcal{Y}} \delta_t^i(y)(\mathrm{d}J_t^y - \lambda_{S_{t-},y}(A_t)\,\mathrm{d}t) + \mathrm{d}M_t^i \end{aligned} \tag{5}$$

for a martingale M^i orthogonal to $\int_0^\cdot \sigma(S_t)\,\mathrm{d}Z_t$ and $(J^y)_{y\in\mathcal{Y}}$ with $M_0^i = 0$, predictable and locally square-integrable processes β^i and $\delta^i(y)$ for $y \in \mathcal{Y}$, and S induced by A with $S_0 = y_0$.

The first term in (5) is a drift term that captures the expected change in the continuation value. If player i extracts a flow payoff that exceeds their continuation value, their continuation value has to decrease in return. Geometrically, this means W moves away from the flow payoff $g(y, A_t)$ in expectation. The second term is a diffusion term that captures how the continuation value responds to the public signal. In a continuous-time setting, rewards or punishments $r\beta_t^i$ are provided proportionally to the difference between the observed change $\sigma(S_t)\,\mathrm{d}Z_t$ in the signal and its expected change $\mu(S_t, A_t)\,\mathrm{d}t$. The third term captures the impact of state transitions. Since $\mathrm{d}J^y = 1$ if and only if the state transitions to state y, the term $r\delta_t^i(y)$ is player i's reward or punishment when such a transition occurs. If a state transition to state y is good news for player i, then its absence is bad news. Therefore, the continuation

value experiences a drift of $-r\delta_t^i(y)\lambda_{S_{t-},y}(A_t)\,\mathrm{d}t$ in the opposite direction. The final term is a martingale that reflects the use of public randomization. Changes in the continuation value due to public randomization average out in expectation and $M = 0$ if the players do not use it.

4.1 Enforceability

A one-shot deviation is a deviation for one pair (ω, t), where ω is an element of the underlying probability space, determining the realization of the publicly observable processes. An action profile is enforceable if no one-shot deviation increases the drift of the continuation value.

Definition 4.2. An action profile α is *enforceable* in state y if there exists a *continuation promise* (β, δ) with $\beta = (\beta^1, \beta^2)^\top$ and $\delta = (\delta^1, \delta^2)^\top$ such that for $i = 1, 2$ and each $\tilde{a}^i \in \mathcal{A}^i(y)$,

$$g^i(y,\alpha) + \beta^i\mu(y,\alpha) + \delta^i\lambda(y,\alpha) \geq g^i(y,\tilde{a}^i,\alpha^{-i}) + \beta^i\mu(y,\tilde{a}^i,\alpha^{-i}) + \delta^i\lambda(y,\tilde{a}^i,\alpha^{-i}). \quad (6)$$

A strategy profile A is *enforceable* for initial state y_0 if there exist processes $(\beta_t)_{t\geq 0}$, $(\delta_t)_{t\geq 0}$ such that (6) is satisfied a.e. for the induced state process S.

Contrary to discrete-time games, the existence of profitable one-shot deviation does not preclude a strategy profile from being a PPE—as long as such deviations do not exist on a set of positive measure. We obtain the following positive-measure deviation principle, stating that the local incentive constraint (6) is sufficient if it is satisfied almost everywhere.

Lemma 4.3. *A strategy profile A is a PPE for initial state y_0 if and only if (β, δ) related to A by* (5) *enforces A for the induced state process S with $S_0 = y_0$.*

For the characterization of PPE payoffs—but not Markov-perfect and state-order dependent PPE payoffs—we need to make some assumption on the public signal in any given state y to ensure that the ODEs arising in the characterization are sufficiently well-behaved.

Definition 4.4. For any y, any $a \in \mathcal{A}(y)$, and any player $i = 1, 2$, let $M_y^i(a)$ denote the matrix with column vectors $\mu(y, \tilde{a}^i, a^{-i}) - \mu(y, a)$ for any deviation $\tilde{a}^i \in \mathcal{A}^i(y) \setminus \{a^i\}$.

(i) Action profile a has *individual full rank* in state y if $\operatorname{rank} M_y^i(a) = |\mathcal{A}^i(y)| - 1$ for $i = 1, 2$.

(ii) The public signal has a *product structure* in state y if $\operatorname{span} M_y^1(a) \perp \operatorname{span} M_y^2(a)$ for all a.

The individual full rank condition guarantees that $M_y^i(a)$ is invertible, hence there exists β^i that solves (6) with equality for any a and δ. If an action profile has individual full rank, then it is enforceable and incentives from state transitions can be compensated through incentives within the state. It turns out that the latter condition is relevant to establish regularity of

the optimality equations; see Assumption 1.(ii.b) below. If the public signal has a product structure, incentives for the two players can be provided along separate dimensions.

Assumption 1. In each state y, exactly one of the following conditions is satisfied:

(i) The public signal is uninformative in state y, i.e., $\mu(y, a)$ is constant across a.

(ii) All of the following conditions are satisfied, where α refers to a mixed action profile in general, or to a pure action profile if players are restricted to use pure strategies:

(a) The public signal has a product structure in state y.

(b) Incentives from state transitions can be compensated within the state: for any enforceable action profile α and any δ, there exists β such that (β, δ) enforces α.

(c) Best responses are enforceable: for any enforceable action profile α, any δ, and any player i, there exist $\tilde{\alpha}^i$ and β with $\beta^i = 0$ such that (β, δ) enforces $(\tilde{\alpha}^i, \alpha^{-i})$.

The following are sufficient conditions on game primitives for Assumption 1 to hold.

Lemma 4.5. *Suppose that Assumption 1.(ii.a) holds in state y. Then Assumptions 1.(ii.b) and (ii.c) also hold if for both players i, either of the following conditions are satisfied:*

(i) $M^i_y(a)$ has individual full rank for any action profile $a \in \mathcal{A}(y)$.

(ii) There exists an embedding of $\mathcal{A}^i(y)$ into $\mathbb{R}$ such that for any a^{-i}, both $\lambda(y, a)$ and $\mu(y, a)$ are affine in a^i, with $\mu(y, a)$ non-constant, and $g^i(y, a)$ is strictly concave in a^i.

For the characterization of the PPE payoff correspondence in behavior strategies, but not for pure strategies, we make the following assumption about state transitions.

Assumption 2. Suppose that in each state y, either of the following conditions is satisfied:

(i) State transitions satisfy a product structure: $\mathcal{Y} \setminus \{y\}$ can be partitioned into $\mathcal{Y}_1$ and $\mathcal{Y}_2$ such that player i's actions affect the frequency of transitions only to states in $\mathcal{Y}_i$.

(ii) There is a single successor state y', $\left|\mathcal{A}^i(y)\right| = \left\{a^i_1, a^i_2\right\}$ for $i = 1, 2$, and λ is monotone: $\lambda_{y,y'}(a^i_1, a^{-i}_1) - \lambda_{y,y'}(a^i_2, a^{-i}_1)$ and $\lambda_{y,y'}(a^i_1, a^{-i}_2) - \lambda_{y,y'}(a^i_2, a^{-i}_2)$ have the same sign.

4.2 Self-generation

To motivate the local construction of PPE in our model, it will be useful to recap its analogue in discrete time, introduced in Abreu, Pearce, and Stacchetti (1990). In each period, incentives for the chosen action profile are provided through the promise of a continuation value from

some family of payoff sets $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$, where the continuation value lies in $\mathcal{W}_y$ if the next state is y. The concept of self-generation ensures that these promises are credible by imposing that any payoff pair $w \in \mathcal{W}_y$ is itself attainable through a current-period action profile α, enforced by a continuation promise on $\mathcal{W}$. Then an infinite concatenation of these action profiles delivers the promised values in an enforceable way: it is a PPE. In continuous time, this local construction corresponds to finding enforceable solutions to SDE (5) that remain within the same family of payoff sets at all times. Contrary to discrete time, there is no minimal time increment, at which we can concatenate the solution. There is, however, a suitable sequence of stopping times that permits an iterative construction of equilibria. Let τ_n denote the time of the n^{th} state transition. Since interarrival times of state changes are exponentially distributed, it follows that $\tau_n \to \infty$ a.s. Thus, if we can construct locally enforceable solutions to (5) between state transitions, a countable concatenation is a PPE.

Definition 4.6.

(i) A family of payoff sets $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$ *generates* a family of payoff sets $\mathcal{X} = (\mathcal{X}_y)_{y\in\mathcal{Y}}$ if for each y and $w \in \mathcal{X}_y$, there exists $(W, S, A, \beta, \delta, M)$ on a probability space $(\Omega, \mathcal{F}, P)$ with $S_0 = y$ and $W_0 = w$ such that on $[\![0, \tau_1]\!] := \{(\omega, t) \in \Omega \times [0,\infty) \mid 0 \le t \le \tau_1(\omega)\}$, the processes satisfy (5), (β, δ) enforces A, $W \in \mathcal{X}_y$ before time τ_1 and $W_{\tau_1} \in \mathcal{W}_{S_{\tau_1}}$.

(ii) $\mathcal{W}$ *pure-strategy generates* $\mathcal{X}$ if for any y and $w \in \mathcal{X}_y$, a solution exists in pure strategies.

(iii) A family $\mathcal{W}$ is *(pure-strategy) self-generating* if it (pure-strategy) generates itself.

$\mathcal{W}$ generates $\mathcal{X}$ if continuation values in $\mathcal{W}$ at time τ_1 are able to enforceably sustain any payoffs in $\mathcal{X}$. If $\mathcal{W}$ is self-generating, those continuation values can themselves be sustained through another promise in $\mathcal{W}$ at the time of the next state transition, and so on. A countable concatenation thus yields a PPE by Lemma 4.3, giving rise to the following lemma.

Lemma 4.7. *$\mathcal{E}$ is the largest bounded family of self-generating payoff sets.*

For the local construction of equilibria and the characterization of $\mathcal{E}$, it will be helpful to study the largest family of bounded payoff sets $\mathcal{B}(\mathcal{W}) = (\mathcal{B}_y(\mathcal{W}))_{y\in\mathcal{Y}}$ that is generated by $\mathcal{W}$.[4] As usual, let $\mathcal{B}^p(\mathcal{W})$ denote its pure-strategy analogue. The following lemma relates the operators $\mathcal{B}$ and $\mathcal{B}^p$ to self-generation and the family of PPE payoffs.

[4] If $\mathcal{X}$ and $\mathcal{X}'$ are generated by $\mathcal{W}$, their state-wise union is as well, hence the largest family is well defined.

Lemma 4.8. *The operator $\mathcal{B}$ has the following properties:*

(i) A bounded family $\mathcal{W}$ is self-generating only if $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for each state y.

(ii) If $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for every state y, then $\mathcal{B}(\mathcal{W})$ is self-generating.

(iii) $\mathcal{B}_y(\mathcal{E}) = \mathcal{E}(y)$ for any state y.

(iv) If y is absorbing, then $\mathcal{B}_y(\mathcal{W}) = \mathcal{E}(y)$ for any $\mathcal{W}$.

Moreover, the operator $\mathcal{B}^p$ satisfies the analogue properties with $\mathcal{E}$ replaced by $\mathcal{E}^p$.

Statement (i) is simply a restatement that $\mathcal{B}(\mathcal{W})$ is the largest bounded family generated by $\mathcal{W}$. Statement (ii) permits the concatenation procedure described above: any $w \in \mathcal{B}_y(\mathcal{W})$ can be attained by an enforceable strategy profile with a continuation value in $\mathcal{W}_{S_{\tau_1}}$ at the time τ_1. If $\mathcal{W}_{S_{\tau_1}} \subseteq \mathcal{B}_{S_{\tau_1}}(\mathcal{W})$, the strategies can be extended in an enforceable way until the second state transition, and so on. Statements (i) and (iii) together imply that $\mathcal{E}$ is the largest bounded fixed point of $\mathcal{B}$. The first three statements of Lemma 4.8 highlight the similarities of $\mathcal{B}$ with the standard set operator from Abreu et al. (1990). These similarities arise because incentives at state transitions have an inherently discrete character in games with finitely many states. However, $\mathcal{B}$ must also capture incentives between state transitions, which give rise to a local description of its boundary. As statement (iv) highlights, those characterize equilibrium payoffs explicitly in absorbing states where no more state transitions occur.

Our main goal in Sections 5–7 is to characterize the boundary of $\mathcal{B}_y(\mathcal{W})$ for an arbitrary family of payoff sets $\mathcal{W}$. In some games, this yields an explicit description of the PPE payoff correspondence. For instance, in repeated games we recover the main result of Sannikov (2007) by statement (iv) of Lemma 4.8. In absorbing games with initial state y_0, statement (iv) applies to all absorbing states, and statement (iii) characterizes $\mathcal{E}(y_0)$ by applying $\mathcal{B}_{y_0}$ to $(\mathcal{E}(y))_{y \neq y_0}$. More generally, when each state is visited at most once, the characterization of $\mathcal{B}$ yields an explicit characterization of $\mathcal{E}$. When each state can be visited more than once, we obtain a fixed-point characterization, describing the boundary of each $\mathcal{E}(y)$ as ODEs that depend on $\mathcal{E}(y')$ for states $y' \neq y$. For such games, the following algorithm, based on the fact that $\mathcal{E}$ is the largest bounded fixed point of $\mathcal{B}$, allows us to approximate $\mathcal{E}$. Replacing $\mathcal{B}^p$ in the iterative procedure below yields an approximation of $\mathcal{E}^p$ instead.

Proposition 4.9. *Let $\mathcal{W}_0 = (\mathcal{W}_{0,y})_{y \in \mathcal{Y}}$ be the family of payoff sets with $\mathcal{W}_{0,y} = \mathcal{V}_0^*$ for every $y \in \mathcal{Y}$. Define the sequence $(\mathcal{W}_n)_{n \geq 0}$ iteratively via $\mathcal{W}_n = \mathcal{B}(\mathcal{W}_{n-1})$ for $n \geq 1$. Then $(\mathcal{W}_{n,y})_{n \geq 0}$ is decreasing in the set-inclusion sense for every $y \in \mathcal{Y}$ with $\bigcap_{n \geq 0} \mathcal{W}_{n,y} = \mathcal{E}(y)$.*

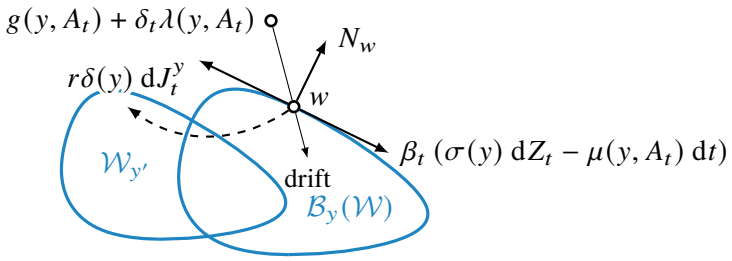

Figure 4: Restrictions on incentives at extremal payoff pairs $w \in \partial \mathcal{B}_y(\mathcal{W})$.

5 Heuristic derivation

This section provides a heuristic derivation of $\partial \mathcal{B}_y(\mathcal{W})$ in order to motivate the four types of boundary payoffs that may arise and to highlight the connection among them. The characterization of stationary payoffs and continuation-binding payoffs are deferred to Sections 6 and 7, respectively, because they depend on which specific assumptions are satisfied.

Throughout the remainder of the paper, we fix a state y and a family $\mathcal{W} = (\mathcal{W}_{y'})_{y' \in \mathcal{Y}}$ of compact and convex payoff sets $\mathcal{W}_{y'}$. For the heuristic derivation only, suppose that $\mathcal{B}_y(\mathcal{W})$ is closed so that any boundary payoff pair w is attainable by an enforceable solution to (5). This places the following restrictions on the continuation promise (β, δ) used to incentivize action profile α at boundary payoff w, where N is an arbitrary outward normal vector at w:

(I1) The drift points inwards: $N^\top(g(y, \alpha) + \delta\lambda(y, \alpha) - w) \geq 0$.

(I2) The public signal is used to transfer values tangentially to the set: $N^\top \beta = 0$.

(I3) δ is $\mathcal{W}$-*feasible*: $w + r\delta(y') \in \mathcal{W}_{y'}$ for any state y'.

See Figure 4 for an illustration of these restrictions. Restriction (I2) arises because the public signal has unbounded variation: if the continuation value moved non-tangentially with the public signal, it would escape $\mathcal{B}_y(\mathcal{W})$ immediately. On the boundary, the public signal is thus used exclusively for tangential value transfers. Restriction (I3) is feasibility in the continuation state. Lemma 5.2 shows that restriction (I1) corresponds to feasibility in the current state.

Definition 5.1. Payoff w is $\mathcal{W}$-*feasible* in state y if there exist α and δ with $w + r\delta(y') \in \mathcal{W}_{y'}$ for each $y' \in \mathcal{Y}$ such that $w = g(y, \alpha) + \delta\lambda(y, \alpha)$. We say such (α, δ) *decomposes* w on $\mathcal{W}$ in state y. Denote by $\mathcal{V}_y(r; \mathcal{W})$ or $\mathcal{V}_y(\mathcal{W})$ the convex hull of all $\mathcal{W}$-feasible payoffs in state y.

Lemma 5.2. *A set $\mathcal{X}_y$ is contained in $\mathcal{V}_y(\mathcal{W})$ if and only if for every $w \in \partial \mathcal{X}_y$ and outward normal N, there exist (α, δ) that satisfy (I1) and (I3). In particular, $\mathcal{B}_y^p(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W}) \subseteq \mathcal{V}_y(\mathcal{W})$.*

Non-stationary strategies cannot attain payoffs outside of $\mathcal{V}_y(\mathcal{W})$ because (I1) and (I3) are at odds with each other there: the larger $\delta(y')$ has to be to reach $\mathcal{W}_{y'}$ (in some direction), the stronger is the drift in the opposite direction. Thus, (I1) can be interpreted as feasibility in the current state. We refer to continuation promises that satisfy all restrictions as follows:

- $\Xi_{y,\alpha}(w, N, \mathcal{W})$ is the set of all continuation promises (β, δ) that enforce α in state y that satisfy (I1)–(I3). Such (β, δ) are said to $\mathcal{W}$-*restricted-enforce* α at (w, N).
- $\Psi_{y,\alpha}(w, \mathcal{W})$ is the set of all $\mathcal{W}$-feasible δ that enforce action profile α with $\beta = 0$.

We may omit the dependence on y and $\mathcal{W}$ if they are clear from context.

5.1 The four kinds of boundary payoffs

Because the public signal is driven by Brownian motion, players transfer value very rapidly. At payoff pairs where the boundary is strictly curved, these transfers lie above the tangent at every instant, causing the continuation value to drift away from the boundary. It follows from Itō's formula that the strength of this outward drift is $\frac{r^2}{2}\kappa_y(w)\|T_w{}^\top\beta\sigma(y)\|^2$, proportional to the curvature $\kappa_y(w)$ at w. For the continuation value to remain within $\mathcal{B}_y(\mathcal{W})$, this outward drift has to be compensated by the inward drift characterized in (I1), i.e., it is necessary that there exist some action profile α and continuation promise $(\beta, \delta) \in \Xi_{y,\alpha}(w, N, \mathcal{W})$ that satisfy

$$\frac{r^2}{2}\kappa_y(w)\|T_w{}^\top\beta\sigma(y)\|^2 \leq rN^\top(g(y,\alpha) + \delta\lambda(y,\alpha) - w). \tag{7}$$

Suppose first that (7) binds, i.e., there are no α and $(\beta, \delta) \in \Xi_{y,\alpha}(w, N, \mathcal{W})$ for which the inequality is strict. The inequality may bind in one of three ways. First, w may decomposed by some (α, δ), hence the continuation value has no drift and (7) implies $\beta = 0$. In particular, the continuation value is locally constant, and we call such payoffs *stationary*.

Definition 5.3. A payoff pair w is $\mathcal{W}$-*stationary* in state y if there exist α and $\delta \in \Psi_{y,\alpha}(w, \mathcal{W})$ with $w = g(y, \alpha) + \delta\lambda(y, \alpha)$. We denote by $\mathcal{S}_y(\mathcal{W})$ the set of all such payoff pairs, by $\mathcal{S}^p_y(\mathcal{W})$ its restriction to pure actions, and $\mathcal{S}(\mathcal{W}) = (\mathcal{S}_y(\mathcal{W}))_{y\in\mathcal{Y}}$ as well as $\mathcal{S}^p(\mathcal{W}) = (\mathcal{S}^p_y(\mathcal{W}))_{y\in\mathcal{Y}}$.

Second, the continuation value may have a non-zero tangential drift, hence (7) still requires that incentives are provided through state transitions exclusively. Rewriting the binding inequality (7) with $\beta = 0$ yields the *state-transition optimality equation*

$$N_w{}^\top w = \max_\alpha \max_{\delta\in\Psi_{y,\alpha}(w,\mathcal{W})} N_w{}^\top(g(y,\alpha) + \delta\lambda(y,\alpha)), \tag{8}$$

where the maximization is taken over all action profiles α for which $\Psi_{y,\alpha}(w, \mathcal{W})$ is non-empty. Strategy profiles that attain boundary payoffs characterized by (8) condition on state transitions and time. If feasibility were the only concern, the highest weighted sum of payoffs $N_w{}^\top w$ would be attained for the highest-feasible continuation promise δ_w that maximizes $N_w{}^\top \delta_w(y')$ in each successor state y' for the same payoff weights N_w. Equation (8) shows that when incentives are taken into account, the weighted sum of payoffs is maximized in α and δ that maximize three equally weighted objectives: the expected flow payoff received $N_w{}^\top g(y, \alpha)$, the expected value of future opportunities $N_w{}^\top \delta_w \lambda(y, \alpha)$, and (the negative of) the expected cost of incentives $N_w{}^\top(\delta - \delta_w)\lambda(y, \alpha)$; see also Figure 1.

Third, incentives may be provided through both state transitions and the public signal. Solving the binding inequality (7) for the curvature yields the *optimality equation*

$$\kappa_y(w) = \max_{\alpha} \max_{(\beta,\delta)\in\Xi_{y,\alpha}(w,N_w,\mathcal{W})} \frac{2N_w{}^\top(g(y,\alpha) + \delta\lambda(y,\alpha) - w)}{r\|T_w{}^\top\beta\sigma(y)\|^2}. \tag{9}$$

Equation (9) characterizes the optimal trade-off between incentives provided through state transitions and the public signal. Due to Assumption 1.(ii.b), it is always possible to attach the highest-feasible continuation promise δ_w to state transitions and provide all incentives through tangential transfers. The optimality equation shows that this may not be optimal for impatient players due to the noise in the signal: the tangential transfers entail a cost equal to the inward drift needed to offset them, which scales quadratically in the size of the transfers.

If the public signal is used to provide incentives, the boundary must be smooth because (I2) has to hold for each outward normal vector, hence $\beta \neq 0$ implies that the outward normal vector N_w to $\mathcal{B}_y(\mathcal{W})$ at w is unique. The optimality equation is thus a second-order ODE in the state w, its first derivative N_w, and its second derivative $\kappa_y(w)$. We show in Appendix F.2 that its solutions are unique and continuous in initial conditions. The use of the notation N_w in the state-transition optimality equation (8), suggesting uniqueness of the normal vector, requires some justification. We show in Section 7 that the set of restricted-enforceable action profiles changes discontinuously only for payoffs in a set $\mathcal{K}_y(\mathcal{W})$ of measure zero, the so-called continuation-binding payoffs, defined formally in Section 7. Outside of this set, solutions to the first-order ODE (8) are smooth and continuous in initial conditions.

Solutions to either optimality equation have the following very desirable property that helps us characterize the boundary of $\mathcal{B}_y(\mathcal{W})$ even when (7) does not bind.

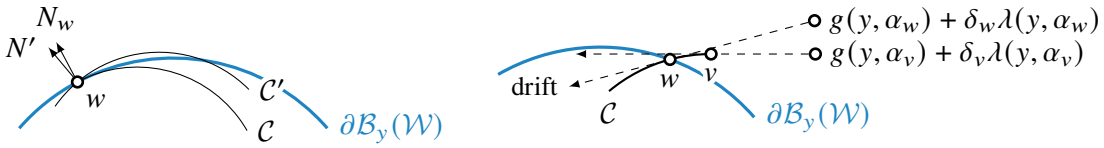

Figure 5: Perturbation arguments that motivate the optimality equations.

Lemma 5.4. *Let $\mathcal{C}$ be a smooth solution to either* (8) *or* (9)*, and let $\alpha_*(w)$, $\beta_*(w)$, and $\delta_*(w)$ denote the corresponding maximizers at w. Call an end point v of $\mathcal{C}$ reachable if the solution W to SDE* (5) *with $A = \alpha_*(W)$, $\beta = \beta_*(W)$, and $\delta = \delta_*(W)$ reaches v with positive probability.*[5] *If all reachable endpoints of $\mathcal{C}$ are in $\mathcal{B}_y(\mathcal{W})$, then $\mathcal{C} \subseteq \mathcal{B}_y(\mathcal{W})$.*

The solution W remains on $\mathcal{C}$ because the inward and outward drift (relative to N_w) precisely offset each other. Thus, W is a restricted-enforceable solution to (5) that remains on $\mathcal{C}$ until either an end point of $\mathcal{C}$ is reached or a state transition occurs. If an end point $v \in \mathcal{B}_y(\mathcal{W})$ is reached before a state transition occurs, then by definition it is attainable by a restricted-enforceable solution $(W', A', \beta', \delta')$ to (5) that remains in $\mathcal{B}_y(\mathcal{W})$ until the next state transition. Thus, the concatenation of the solutions attains w and remains in $\mathcal{C} \cup \mathcal{B}_y(\mathcal{W})$. In particular, $\mathcal{C} \cup \mathcal{B}_y(\mathcal{W})$ is generated by $\mathcal{W}$, hence maximality of $\mathcal{B}_y(\mathcal{W})$ implies $\mathcal{C} \subseteq \mathcal{B}_y(\mathcal{W})$.

Finally, consider the case where (7) does not bind at a boundary payoff w, i.e., there exist some outward normal vector N, an action profile α_w, and incentives $(\beta_w, \delta_w) \in \Xi_{\alpha_w}(w, N)$ for which the inequality is strict. Suppose first that all such incentives satisfy $\beta_w \neq 0$. Then the curvature of $\partial\mathcal{B}_y(\mathcal{W})$ at w is strictly smaller than $\kappa_y(w)$ given by the optimality equation. A solution $\mathcal{C}$ to (9) starting at (w, N_w) must therefore fall into the interior of $\mathcal{B}_y(\mathcal{W})$. Consequently, a solution $\mathcal{C}'$ for slightly rotated initial conditions leaves and re-enters $\mathcal{B}_y(\mathcal{W})$ as illustrated in the left panel of Figure 5, a contradiction to Lemma 5.4. Suppose next that incentives through state transitions are sufficient to enforce α_w with strict inward drift. Then a solution $\mathcal{C}$ to the state-transition optimality equation (8) started at w must fall into the interior of $\mathcal{B}_y(\mathcal{W})$. Because (8) has smooth solutions outside of $\mathcal{K}_y(\mathcal{W})$, this can happen only at $w \in \mathcal{K}_y(\mathcal{W})$. Otherwise, there is a smooth extension $\mathcal{C}'$ of $\mathcal{C}$ that solves (8) and attains payoffs slightly outside of $\mathcal{B}_y(\mathcal{W})$ as illustrated in the right panel of Figure 5, again a contradiction to Lemma 5.4. We conclude that (7) must bind everywhere on the boundary of $\mathcal{B}_y(\mathcal{W})$ except at payoffs in $\mathcal{K}_y(\mathcal{W})$, where a solution to (8) fails to be smooth.

[5] Only a single end point of solutions to (8) is reachable because (5) is deterministic if $\beta_* \equiv 0$ and $M \equiv 0$ while no state transition occurs. Both end points of solutions to (9) are reachable with positive probability.

5.2 Conditioning on the support of mixed actions

To fill in the details to Section 5.1, we will condition on the support of the players' mixed actions in Sections 6 and 7. We introduce the relevant notation here because it will be used in both sections. Throughout the rest of the paper, $\mathcal{A}_\diamond = \mathcal{A}^1_\diamond \times \mathcal{A}^2_\diamond$ will refer to a product set with $\mathcal{A}^i_\diamond \subseteq \mathcal{A}^i(y)$. Continuation promise δ enforces α with support $\mathcal{A}_\diamond$ if and only if

$$g^i(y, a^i, \alpha^{-i}) + \delta^i \lambda(y, a^i, \alpha^{-i}) \geq g^i(y, \tilde{a}^i, \alpha^{-i}) + \delta^i \lambda(y, \tilde{a}^i, \alpha^{-i}) \tag{10}$$

for all $a^i \in \mathcal{A}^i_\diamond$, all $\tilde{a}^i \neq a^i$, and $i = 1, 2$. Given δ^i, player i must be indifferent between all actions in $\mathcal{A}^i_\diamond$, and must weakly prefer them over all actions outside of the support. Note that the set of player i's incentives depends on the precise mixture α^{-i} of their opponent, but it depends on their own mixture only through its support. We will use the following notation:

- $\Psi_{y,\mathcal{A}_\diamond}(\alpha)$ and $\Psi_{y,\mathcal{A}_\diamond}$ denotes the set of all δ that satisfy (10) for both players. The former for given α, the latter for some α supported on $\mathcal{A}_\diamond$.
- $\Upsilon_{y,\mathcal{A}_\diamond}(w, \mathcal{W})$ is the set of all (α, δ) with $\operatorname{supp}(\alpha) \subseteq \mathcal{A}_\diamond$ and $\mathcal{W}$-feasible $\delta \in \Psi_{y,\mathcal{A}_\diamond}(\alpha)$.

We again omit the dependence on y and $\mathcal{W}$ if they are clear from context. While mixed actions break the polygon structure of solutions to (10) in general, it holds under Assumption 2.

Lemma 5.5. *If $\mathcal{A}_\diamond$ is a singleton or Assumption 2 holds, $\Psi_{\mathcal{A}_\diamond}$ is a closed convex polyhedron.*

6 State-order dependent strategies and stationary payoffs

6.1 Stationary payoffs with a single successor state

Fix a state y with a unique successor state y' as in the regime-change game of Section 3, which we will use for illustration. A key simplification in this setting is that $\delta\lambda(y, a) = \delta(y')\lambda_{y,y'}(a)$, hence the provided incentives $\delta^i(y')$ are one-dimensional. As in Section 5.2, we will condition on the support of the chosen action profiles. For any support $\mathcal{A}_\diamond \subseteq \mathcal{A}(y)$, let $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ denote the set of all payoff pairs w that can be decomposed by some pair $(\alpha, \delta) \in \Upsilon_{\mathcal{A}_\diamond}(w)$.

We begin with pure action profiles. Since a state transition has a payoff impact of $r\delta(y')$, the set of continuation payoffs in $\mathcal{W}_{y'}$ reachable from some $w \in \mathcal{S}_{y,a}(\mathcal{W})$ is the set

$$\mathcal{X}_{y,a}(\mathcal{W}) := \mathcal{X}_{y,a} \cap \mathcal{W}_{y'}, \qquad \text{where} \qquad \mathcal{X}_{y,a} := g(y, a) + (\lambda_{y,y'}(a) + r)\Psi_{y,a}.$$

Because $\Psi_{y,a}$ is an intersection of half-spaces, $\mathcal{X}_{y,a}(\mathcal{W})$ is easily computed as the intersection

of two convex sets. Payoffs in $\mathcal{S}_{y,a}(\mathcal{W})$ are the expected values of future payoffs in $\mathcal{X}_{y,a}(\mathcal{W})$ and the current flow payoff $g(y,a)$, hence $\mathcal{S}_{y,a}(\mathcal{W})$ is the convex combination

$$\mathcal{S}_{y,a}(\mathcal{W}) = \frac{r}{r+\lambda_{y,y'}(a)} g(y,a) + \frac{\lambda_{y,y'}(a)}{r+\lambda_{y,y'}(a)} \mathcal{X}_{y,a}(\mathcal{W}).$$

Geometrically, this corresponds to shrinking $\mathcal{X}_{y,a}(\mathcal{W})$ towards $g(y,a)$; see the left panel in Figure 6. The set of all pure-strategy stationary payoffs is then $\mathcal{S}^p_y(\mathcal{W}) = \bigcup_{a\in\mathcal{A}(y)} \mathcal{S}_{y,a}(\mathcal{W})$.

Example 1. We illustrate this construction by computing $\mathcal{S}^p_{y_1}(\mathcal{V}^*_0)$ in the regime-change game. We first compute $\mathcal{S}_{y_1,a}(\mathcal{W})$ for the cooperative action profile $a = (B,S)$. The incumbent is willing to play B as long as state transitions are not too costly; otherwise a deviation to A would be profitable to reduce their frequency. The non-incumbent is willing to play S as long as state transitions are not too rewarding; otherwise a deviation to I would be profitable to increase their frequency. Specifically, $\Psi_{y_1,(B,S)}$ is the set of all δ with $\delta^1(y_2) \geq -2$ and $\delta^2(y_2) \leq 5/8$. To compute $\mathcal{X}_{y_1,(B,S)}$, we scale this set by a factor of $\lambda_{y_1,y_2}(a) + r = 2$ and shift it by $g(y_1,a) = (4,2)$; see the left panel in Figure 6. We then shrink $\mathcal{X}_{y_1,(B,S)} \cap \mathcal{V}^*_0$ towards $g(y_1,a)$ by a factor of $\lambda_{y_1,y_2}(a)/(\lambda_{y_1,y_2}(a)+r) = 1/2$. Repeating this construction for all pure action profiles and states yields $\mathcal{S}^p(\mathcal{V}^*_0)$ shown in the middle panel of Figure 6.

Next, we characterize the mixed-action stationary payoffs under Assumption 2.(ii), that is, when $\left|\mathcal{A}^i(y)\right| = 2$ and λ is monotone. Monotonicity allows us to label $\mathcal{A}^i(y) = \left\{a^i_1, a^i_2\right\}$ so that $\lambda(y, a^i_1, \alpha^{-i}) \geq \lambda(y, a^i_2, \alpha^{-i})$ for any α^{-i}. We say that λ is *strictly monotone* for player i if the inequality is strict. Promise δ^i makes player i indifferent among both pure actions if

$$g^i(y, a^i_1, \alpha^{-i}) + \delta^i \lambda(y, a^i_1, \alpha^{-i}) = g^i(y, a^i_2, \alpha^{-i}) + \delta^i \lambda(y, a^i_2, \alpha^{-i}). \tag{11}$$

Strict monotonicity implies that there is a unique $\delta^i_*(\alpha^{-i})$ such that (11) holds. If λ is not strictly monotone for player i under Assumption 2.(ii), then player i has no impact on state transitions. In that case, α is enforceable if and only if α^i is a best response to α^{-i}.

Suppose first that λ is strictly monotone for both players. An upper bound for $\mathcal{S}_{y,\mathcal{A}}(\mathcal{W})$ is the set $\mathcal{S}^0_{y,\mathcal{A}}$ of all payoffs that can be decomposed by (α,δ) without requiring that δ be feasible. It follows from (11) that $g^i(y,\alpha) + \delta^i_*(\alpha^{-i})\lambda(y,\alpha)$ depends on α only through α^{-i}, and it does so monotonically. Thus, $\mathcal{S}^0_{y,\mathcal{A}}$ is a rectangle whose corners are attained by the pure action profiles; see the right panel in Figure 6. Let $\mathcal{X}_{y,\mathcal{A}}$ denote the set of all $v_{y'}$ that satisfy

$$v_{y'} = g(y,\alpha) + \delta_*(\alpha)\lambda(y,\alpha) + r\delta_*(\alpha, y') \tag{12}$$

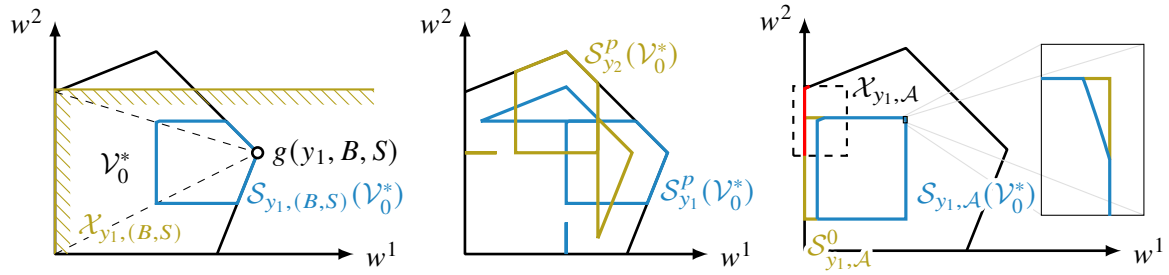

Figure 6: Construction of $\mathcal{S}^p(\mathcal{V}_0^*)$ and $\mathcal{S}(\mathcal{V}_0^*)$ in the regime-change game.

for some mixed action profile α, i.e., the set of continuation payoffs $v_{y'}$ reachable from $\mathcal{S}^0_{y,\mathcal{A}}$. Note that $\mathcal{X}_{y,\mathcal{A}}$ is again a rectangle, hence the feasible continuation set $\mathcal{X}_{y,\mathcal{A}}(\mathcal{W}) = \mathcal{X}_{y,\mathcal{A}} \cap \mathcal{W}_{y'}$ is easily computed; see the right panel in Figure 6. Different from the pure-strategy case, the map from $\mathcal{S}_{y,\mathcal{A}}(\mathcal{W})$ to $\mathcal{X}_{y,\mathcal{A}}(\mathcal{W})$ is not affine because $\delta_*(\alpha)$ depends on α. Instead, extremal payoffs in $\mathcal{S}_{y,\mathcal{A}}$ are either extremal in $\mathcal{S}^0_{y,\mathcal{A}}(\mathcal{W})$ or they satisfy (12) for some $v \in \partial\mathcal{W}_{y'}$. Since (12) is a quadratic equation in each coordinate, it can be solved explicitly for boundary segments in $\partial\mathcal{W}_{y'}$ to trace out the boundary of $\mathcal{S}_{y,\mathcal{A}}(\mathcal{W})$. Given that (12) is a quadratic, a single segment in $\partial\mathcal{W}_{y'}$ may correspond to multiple boundary segments in $\mathcal{S}_{y,\mathcal{A}}(\mathcal{W})$.

Example 2. To compute $\mathcal{S}_{y_1,\mathcal{A}}(\mathcal{V}_0^*)$ in the regime-change game, let α^1 and α^2 denote the probabilities assigned to actions A and I, respectively. A brief computation yields

$$\delta^1_*(\alpha^2) = \frac{-2}{1+8\alpha^2}, \quad \delta^2_*(\alpha^1) = \frac{5}{8-4\alpha^1}, \quad g(y_1,\alpha) + \delta_*(\alpha)\lambda(y_1,\alpha) = \begin{bmatrix} 2-2\alpha^2 \\ 2.625-2\alpha^1 \end{bmatrix}. \tag{13}$$

An upper bound for $\mathcal{S}_{y_1,\mathcal{A}}(\mathcal{V}_0^*)$ is thus certainly the square $\mathcal{S}^0_{y_1,\mathcal{A}}$ with vertices $(0, 0.625)$, $(0, 2.625)$, $(2, 0.625)$, and $(2, 2.625)$, indicated in yellow in the right panel of Figure 6. To find the extremal payoffs of $\mathcal{S}_{y_1,\mathcal{A}}(\mathcal{V}_0^*)$, we solve the quadratic equation (12) for v in the intersection $\mathcal{X}_{y_1,\mathcal{A}} \cap \partial\mathcal{V}_0^*$ indicated in red. The quadratic has two solutions in this example; see the zoom-in. Lemma 6.1 below implies that $\mathcal{S}_y(\mathcal{V}_0^*) = \mathcal{S}^p_y(\mathcal{V}_0^*) \cup \mathcal{S}_{y,\mathcal{A}}(\mathcal{V}_0^*)$.

Lemma 6.1. *Suppose Assumption 2.(ii) holds with strictly monotone λ. Then $\mathcal{S}_{y,\mathcal{A}^i\times\{a^{-i}\}}(\mathcal{W})$ is* $\operatorname{conv}\{w \in \mathcal{S}_{y,\mathcal{A}}(\mathcal{W}) \cup \mathcal{S}_{y,(1,a^{-i})}(\mathcal{W}) \cup \mathcal{S}_{y,(0,a^{-i})}(\mathcal{W}) \mid w^i = g^i(y,a) + \delta^i_*(a^{-i})\lambda(y,a)\}$.

Suppose next that λ is strictly monotone for player 1 but not for player 2, and consider first mixtures with full support. Since no dynamic incentives can be provided to player 2, this requires that player 2 does not have a strictly dominant action. Player 2 is indifferent only against one specific action α^1_*, hence enforceability fixes α^1 but places no constraints

on $\delta^2(y')$.[6] Thus, (12) is instead an affine equation

$$v^2_{y'} = g^2(y, \alpha^1_*, \alpha^2) + \delta^2 \lambda(y, \alpha^1_*) + r\delta^2(y') \tag{14}$$

in the free variable $\delta^2(y')$ for player 2. As above, we first compute $\mathcal{S}^0_{y,\mathcal{A}}$, and solve (12) for player 1 and (14) for player 2 for any boundary segments $\partial\mathcal{W}$ that intersect $\mathcal{S}^0_{y,\mathcal{A}}$.

The same approach works for one-sided mixtures with support $\mathcal{A}_\diamond = \mathcal{A}^1 \times \{a^2_0\}$. We solve equations (12) and (14) for α^1 and δ^2, subject to the additional constraint that a^2_0 is a best response to α^1. Finally, for one-sided mixtures with support $\mathcal{A}_\diamond = \{a^1_0\} \times \mathcal{A}^2$, incentives δ^1 are no longer unique. While both α^2 and δ^1 affect the decomposition of player 1's payoffs, equation (14) for player 1 is affine in (α^2, δ^1) because player 2 has no impact on state transitions. Thus, we solve a pair of affine equations (14) as in the pure-action case.

Finally, if λ is constant, only static Nash profiles are enforceable, hence

$$\mathcal{S}_y(\mathcal{W}) = \bigcup_{\alpha \text{ static Nash}} \frac{r}{r + \lambda_{y,y'}(\alpha)} g(y, \alpha) + \frac{\lambda_{y,y'}(\alpha)}{r + \lambda_{y,y'}(\alpha)} \mathcal{W}_{y'}.$$

6.2 Stationary payoffs with multiple successor states

For a pure action profile a, the construction of $\mathcal{S}_{y,a}(\mathcal{W})$ remains similar, except that we now work in the space of incentives rather than in the payoff space. Let $F_{y,a}(y', \mathcal{W})$ denote the set of all δ that satisfy $g(y, a) + \delta\lambda(y, a) + r\delta(y') \in \mathcal{W}_{y'}$. Note that $F_{y,a}(y', \mathcal{W})$ is convex as the pre-image of a convex set under an affine map, hence so is

$$\mathcal{S}_{y,a}(\mathcal{W}) = g(y, a) + \Big(\Psi_{y,a} \cap \textstyle\bigcap_{y'} F_{y,a}(y', \mathcal{W})\Big)\lambda(y, a). \tag{15}$$

This set can be computed efficiently through a linear program. Alternatively, for a polygon approximation $\mathcal{W}'$ of $\mathcal{W}$, each set $F_{y,a}(y', \mathcal{W}')$ is a polyhedron, hence $\mathcal{S}_{y,a}(\mathcal{W}') - g(y, a)$ is a polygon whose vertices are projections of the vertices of $\Psi_{y,a} \cap \bigcap_{y'} F_{y,a}(y', \mathcal{W}')$.

Equation (15) remains valid for mixed action profiles.[7] While one can compute $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ through (15) for a sufficiently fine grid of mixed actions, doing so is inefficient unless $|\mathcal{A}_\diamond| \leq 2$. The following fixed-point characterization is computationally more efficient when the support

[6] If player 2 has a weakly dominant action in the stage game, then α^1_* is a pure action.

[7] Specifically, let $\mathcal{S}_{y,\mathcal{A}_\diamond}(\alpha_0, \mathcal{W})$ denote the set of all payoffs w that are decomposed by $(\alpha_0, \delta) \in \Upsilon_{y,\mathcal{A}_\diamond}(w)$. Then $\mathcal{S}_{y,\mathcal{A}_\diamond}(\alpha, \mathcal{W}) = g(y, \alpha) + \big(\Psi_{\mathcal{A}_\diamond}(\alpha) \cap \bigcap_{y'} F_{y,\alpha}(y', \mathcal{W})\big)\lambda(y, \alpha)$ for $F_{y,\alpha}(y', \mathcal{W})$ defined as above.

is large. Any $w \in \mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ is a fixed point of the *decomposition operator*

$$\mathcal{D}_{y,\mathcal{A}_\diamond}(w) := \left\{ v \in \mathbb{R}^2 \;\middle|\; v = g(y,\alpha) + \delta\lambda(y,\alpha) \text{ for some } (\alpha,\delta) \in \Upsilon_{y,\mathcal{A}_\diamond}(w) \right\}.$$

We will omit the dependence on y if it is clear from context. The set $\mathcal{D}_{\mathcal{A}_\diamond}(w)$ is easily computed as a bilinear program over a set of convex constraints. $\mathcal{D}_{\mathcal{A}_\diamond}$ has the following properties.

Lemma 6.2.

(i) $\mathcal{D}_{\mathcal{A}_\diamond}(w)$ *is compact. Moreover, if* $\mathcal{A}_\diamond$ *is a singleton, then* $\mathcal{D}_{\mathcal{A}_\diamond}(w)$ *is convex.*

(ii) If $\mathcal{A}_\diamond$ *is a singleton or Assumption 2 holds, then* $w \mapsto \mathcal{D}_{\mathcal{A}_\diamond}(w)$ *is continuous.*

(iii) $\mathcal{D}_{\mathcal{A}_\diamond}(w)$ *generally moves in the opposite direction of* w*: for any* w*,* N*, and* N' *with* $N^\top N' > 0$*, the function* $f(\varepsilon) = \max_{v \in \mathcal{D}_{\mathcal{A}_\diamond}(w+\varepsilon N')} N^\top v$ *is eventually non-increasing in* ε*.*

The motivation for property (iii) is the following: a change of $\varepsilon N'$ in the payoff pair shifts the set of feasible continuation promises $\delta(y')$ by $-\varepsilon/rN'$. For any direction N with $N^\top N' > 0$, let v_N denote the payoff that maximizes $N^\top v$ over $v \in \mathcal{D}_{\mathcal{A}_\diamond}(w)$. If only feasibility constraints bind at v_N, then the supporting tangent shifts by $-\varepsilon/rN^\top N'$, i.e., they decrease in ε. If only enforceability constraints bind at v_N, the boundary is locally constant. If both types of constraints bind at v_N, the supporting tangent may shift in either direction depending on how the constraints interact, but eventually the feasibility constraints will dominate.

To compute $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$, we find some stationary payoff w_0 with a grid search; see Section D.3 for alternatives. By property (iii), we can find an extremal stationary payoff by shifting w_0 in an arbitrary direction N until $w_0 + \varepsilon N$ is extremal in $\mathcal{D}_{\mathcal{A}_\diamond}(w_0 + \varepsilon N)$. From there, we can solve for $\partial\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ with the following differential-inclusion type result.

Lemma 6.3. *The exterior boundary of any connected component of* $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ *is locally the largest closed curve* $\mathcal{C}$ *that satisfies* $w \in \partial\mathcal{D}_{y,\mathcal{A}_\diamond}(w)$ *for all* $w \in \mathcal{C}$*.*

Lemma 6.3 establishes that the tangent direction $\dot{w}$ is the extremal direction for which w remains on the boundary of $\mathcal{D}_{\mathcal{A}_\diamond}(w)$. We elaborate how to implement it numerically in Section D.3. The following lemma provides sufficient conditions for $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ to be simply connected, i.e., for it to consist of a single connected component with no holes. Then its boundary is a single closed curve that solves the inclusion.

Lemma 6.4. *Suppose that either Assumption 2.(ii) is satisfied or Assumption 2.(i) and Condition (ii) of Lemma 4.5 are satisfied such that $\lambda(y,a)$ is linear in a^i for any a^{-i}. Then $\mathcal{D}_{y,\mathcal{A}_\circ}(w)$ and $\mathcal{S}_{y,\mathcal{A}_\circ}(\mathcal{W})$ have convex coordinate sections, hence they are simply connected.*

6.3 State-order dependent equilibria

In a state-order dependent PPE, the continuation value after a state transition must itself be attainable by a locally constant strategy profile, i.e., it must again be stationary. The restriction of self-generation to state-order dependent strategies is the following notion.

Definition 6.5. A family of payoff sets $\mathcal{W}$ is *mutually stationary* if for each state $y \in \mathcal{Y}$ and each payoff pair $w \in \mathcal{W}_y$, there exist α and $\delta \in \Psi_{y,\alpha}(w,\mathcal{W})$ with $w = g(y,\alpha) + \delta\lambda(y,\alpha)$. We say that $\mathcal{W}$ is *mutually stationary in pure strategies* if, in addition, $\alpha \in \mathcal{A}(y)$.

Lemma 6.6. *Fix a family $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$ of payoff sets.*

(i) $\mathcal{W}$ is mutually stationary only if $\mathcal{W}_y \subseteq \mathcal{S}_y(\mathcal{W})$ for each state y.

(ii) If $\mathcal{W}_y \subseteq \mathcal{S}_y(\mathcal{W})$ for every state y, then $\mathcal{S}(\mathcal{W})$ is mutually stationary.

The same statements hold for mutual stationarity in pure strategies if $\mathcal{S}$ is replaced with $\mathcal{S}^p$.

It follows immediately from the definition of mutual stationarity that $\mathcal{S}_y(\mathcal{E}_s) = \mathcal{E}_s(y)$: any payoff pair in $\mathcal{S}_y(\mathcal{E}_s)$ can be attained by a locally enforceable strategy profile whose continuation values after a state transition are attainable by a state-order dependent equilibrium. Conversely, any equilibrium payoff in $\mathcal{E}_s(y)$ must be attainable by such a strategy profile.

Lemma 6.7. *$\mathcal{E}_s$ is the largest family of bounded mutually stationary payoff sets and $\mathcal{E}_s^p$ is the largest family of bounded payoff sets that are mutually stationary in pure strategies.*

Because $\mathcal{E}_s$ and $\mathcal{E}_r$ are the largest fixed points of the monotone operators $\mathcal{S}$ and $\operatorname{conv}\mathcal{S}$, respectively, we obtain the following iterative procedure to approximate $\mathcal{E}_s$ and $\mathcal{E}_r$.

Proposition 6.8. *Set $\mathcal{W}_y^0 = \mathcal{W}_{p,y}^0 = \mathcal{V}_0^*$ for each y and iteratively define $\mathcal{W}_y^n := \mathcal{S}_y(\mathcal{W}^{n-1})$ and $\mathcal{W}_{p,y}^n := \mathcal{S}_y^p(\mathcal{W}_p^{n-1})$ for $n \geq 1$. Then $(\mathcal{W}_y^n)_{n\geq 0}$ and $(\mathcal{W}_{p,y}^n)_{n\geq 0}$ are decreasing in the set-inclusion sense for each state y with $\bigcap_{n\geq 0}\mathcal{W}_y^n = \mathcal{E}_s(y)$ and $\bigcap_{n\geq 0}\mathcal{W}_{p,y}^n = \mathcal{E}_s^p(y)$.*

Moreover, if $\mathcal{S}_y$ is replaced by $\operatorname{conv}\mathcal{S}_y$, then the limits are $\mathcal{E}_r(y)$ and $\mathcal{E}_r^p(y)$ instead.

Example 3. The left panel of Figure 7 shows the convergence of the algorithm in Proposition 6.8 to $\mathcal{E}_s^p$ in the regime-change game of Section 3. Even though $\mathcal{S}_y(\mathcal{V}_0^*)$ is strictly larger than its pure-strategy counterpart $\mathcal{S}_y^p(\mathcal{V}_0^*)$, an iteration of these operators converges to the

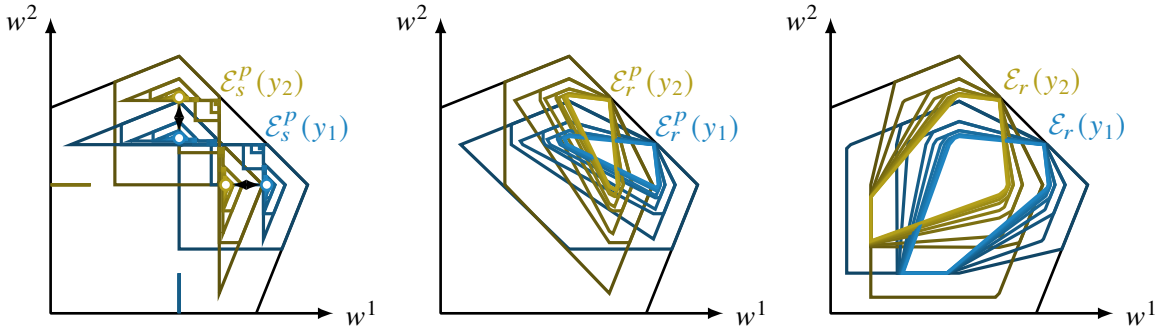

Figure 7: From left to right: convergence to $\mathcal{E}_s^p = \mathcal{E}_s$, $\mathcal{E}_r^p$, and $\mathcal{E}_r$ in the regime-change game.

same limit (not shown in Figure 7). For each state y_i, the sets $\mathcal{E}_s^p(y_i) = \mathcal{E}_s(y_i)$ contain only two payoff pairs corresponding to the players' payoffs in the two Markov-perfect equilibria.

If, additionally, the players have access to a public randomization device, then the ability to use behavior strategies enlarges the state-order dependent PPE correspondence. The middle and right panels of Figure 7 shows the convergence to $\mathcal{E}_r^p$ and $\mathcal{E}_r$, respectively.

6.4 Markov-perfect equilibria

Markov-perfect equilibria are a special case of state-order dependent PPE, in which the strategies depend on the sequence of observed states only through the current state. The following is a verification result for whether a given stationary strategy profile is an MPE.

Proposition 6.9. *A stationary strategy profile $A = \alpha_*(S)$ satisfies SDE (5) with $W = w_*(S)$, $\beta = M = 0$, and $\delta(y') = \delta_*(S, y')$ given by*

$$w_* := rG_*(\operatorname{diag}(r\mathbf{1} + \mathbf{1}\Lambda_*) - \Lambda_*)^{-1}, \qquad \delta_*(y, y') := \frac{w_*(y') - w_*(y)}{r}, \tag{16}$$

where we denote by Λ_* the matrix with entries $\lambda_{y,y'}(\alpha_*(y))$ in row y' and column y, by G_* and w_* the matrices with entries $g^i(y, \alpha_*(y))$ and $w_*^i(y)$, respectively, in row i and column y, and by $\mathbf{1}$ the row vector containing all ones. Moreover, $\alpha_*(S)$ is an MPE if and only if

$$w_*^i(y) \geq g^i(y, a^i, \alpha_*^{-i}(y)) + \sum_{y' \in \mathcal{Y}} \delta_*^i(y, y')\lambda_{y,y'}(a^i, \alpha_*^{-i}(y))$$

holds for each deviation to $a^i \in \mathcal{A}^i(y)$, each state y, and each player $i = 1, 2$.

Note that for a given map α_*, matrices w_* and δ_* are determined explicitly by the respective expressions in (16). A naive algorithm to find the set of pure-strategy MPE when each $|\mathcal{A}(y)|$ is small is thus to verify the conditions of all such pure-strategy maps a_*.

7 Characterization of the PPE correspondence

Let us consider a restriction of the state-transition optimality equation (8) to action profiles with support in some fixed set $\mathcal{A}_\diamond \subseteq \mathcal{A}(y)$. For Lemma 5.4 to apply, it is imperative that we restrict (α, δ) to a closed set. Instead of conditioning on action profiles with exact support $\mathcal{A}_\diamond$, we condition on its closure, the set of all α with $\operatorname{supp}(\alpha) \subseteq \mathcal{A}_\diamond$. However, as the support of α^i shrinks, enforceability imposes fewer constraints of the form (10), hence the set of incentives changes discontinuously. That is why in the definition of $\Upsilon_{\mathcal{A}_\diamond}(w)$, we condition explicitly on the set of binding enforceability constraints, and we consider the following restriction of (8):

$$N^\top w = \max_{(\alpha,\delta)\in\Upsilon_{\mathcal{A}_\diamond}(w)} N^\top(g(y,\alpha) + \delta\lambda(y,\alpha)), \tag{17}$$

where N is an arbitrary direction in $\mathbb{R}^2$ for now. The maximization in (17) is taken over a non-empty set for any payoff pair w in $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W}) := \left\{w \in \mathcal{V}_y(\mathcal{W}) \;\middle|\; \Upsilon_{\mathcal{A}_\diamond}(w) \neq \emptyset\right\}$. The following lemma is the key ingredient to establish regularity of solutions to ODE (17).

Lemma 7.1. *If $\mathcal{A}_\diamond$ is a singleton or if Assumption 2 is satisfied, then $w \mapsto \Upsilon_{\mathcal{A}_\diamond}(w)$ is continuous on $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ and locally Lipschitz continuous on the interior of $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$.*

We can now rewrite the state-transition optimality equation as

$$N_w{}^\top w = \max_{\mathcal{A}_\diamond\subseteq\mathcal{A}(y)} \max_{(\alpha,\delta)\in\Upsilon_{\mathcal{A}_\diamond}(w)} N_w{}^\top(g(y,\alpha) + \delta\lambda(y,\alpha)). \tag{18}$$

Note that we may maximize over the same action profile α multiple times in (18), once for each $\mathcal{A}_\diamond \supseteq \operatorname{supp}(\alpha)$. However, since $(\alpha,\delta) \in \Upsilon_{\mathcal{A}_\diamond}(w) \subseteq \Upsilon_{\operatorname{supp}(\alpha)}(w)$ for each such set $\mathcal{A}_\diamond$, equation (18) indeed coincides with (8). Because the outer maximum is taken over finitely many sets $\mathcal{A}_\diamond$, ODE (18) inherits all regularity properties from (17) except, possibly, on the boundaries of any of the sets $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$. Solutions to (18) for behavior strategies and pure strategies, respectively, are unique and smooth where they exist outside of the sets

$$\mathcal{K}_y(\mathcal{W}) := \bigcup_{\mathcal{A}_\diamond\subseteq\mathcal{A}(y)} \partial\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W}), \qquad \mathcal{K}^p_y(\mathcal{W}) := \bigcup_{a\in\mathcal{A}(y)} \partial\mathcal{K}_{y,a}(\mathcal{W}).$$

We call $\mathcal{K}_y(\mathcal{W})$ the set of *continuation-binding payoffs* because feasibility must bind in a continuation state. To see this, fix any $w \in \partial\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ and $(\alpha,\delta) \in \Upsilon_{\mathcal{A}_\diamond}(w)$. Then $w + r\delta(y')$ must lie on the boundary of at least one set $\mathcal{W}_{y'}$, otherwise δ would remain feasible for any v sufficiently close to w, contradicting $w \in \partial\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$. In addition, at least one enforceability constraint must bind. Otherwise, continuation promise δ_v, defined by $v + r\delta_v(y') = w + r\delta(y')$,

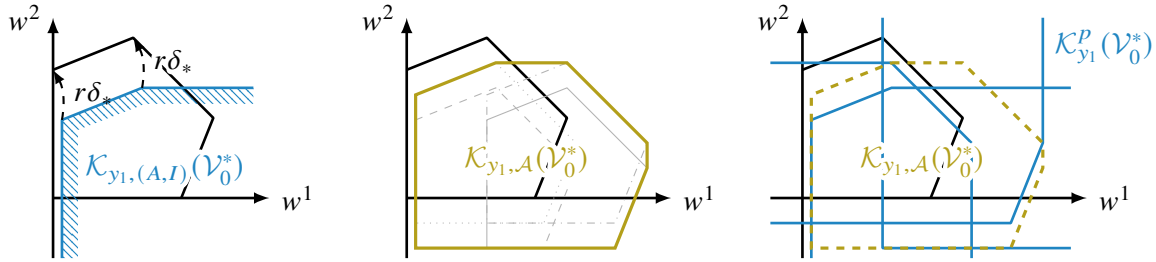

Figure 8: Left: construction of $\mathcal{K}_{y_1,(A,I)}(\mathcal{V}_0^*)$. Middle: construction of $\mathcal{K}_{y,\mathcal{A}}(\mathcal{V}_0^*)$ as the convex hull of the four translates $\mathcal{V}_0^* - r\delta_*(a)$ in grey. Right: $\mathcal{K}^p_y(\mathcal{V}_0^*)$ and $\mathcal{K}_y(\mathcal{V}_0^*) = \mathcal{K}^p_y(\mathcal{V}_0^*) \cup \mathcal{K}_{y,\mathcal{A}}(\mathcal{V}_0^*)$.

enforces α for v sufficiently close to w, again contradicting $w \in \partial\mathcal{K}_{y\mathcal{A}_\diamond}(\mathcal{W})$. For mixed actions, this does not require that $\tilde{a}^i \notin \mathcal{A}^i_\diamond$ becomes a better response than actions in $\mathcal{A}^i_\diamond$, it may simply be that it is no longer possible to feasibly support indifference among actions in $\mathcal{A}^i_\diamond$.

7.1 Continuation-binding payoffs

Consider first the case where the players' actions affect the intensity of state transitions to only a single state y'. Then incentives to player i are provided along a single dimension through $\delta^i(y')$, hence $\Psi_{\mathcal{A}_\diamond}$ is a (possibly empty or unbounded) rectangle by Lemma 5.5. If it is non-empty, denote its four (possibly infinite) corners by $\delta_1(y'), \ldots, \delta_4(y')$. Then

$$\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W}) = \operatorname{conv}\left(\bigcup_{k=1}^{4} \mathcal{W}_{y'} - r\delta_k(y')\right). \tag{19}$$

Example 4. Let us construct $\mathcal{K}_{y_1}(\mathcal{V}_0^*)$ and $\mathcal{K}^p_{y_1}(\mathcal{V}_0^*)$ in the regime-change game, beginning with action profile (A, I). For the non-incumbent to play the costly action I, it is necessary that a revolution is sufficiently rewarding. Similarly, the incumbent is willing to play the costly action A only if revolutions are sufficiently costly. Specifically, δ enforces (A, I) if and only if $\delta^1 \leq -\frac{2}{9} =: \delta^1_*$ and $\delta^2 \geq \frac{5}{4} =: \delta^2_*$. Since the continuation value comes from $\mathcal{V}_0^*$, player 1's utility must be at least $\frac{2}{9}r$—otherwise no punishment in $\mathcal{V}_0^*$ would be severe enough to deter deviations. Similarly, player 2's utility must be no more than $4 - \frac{5}{4}r$ for continuations in $\mathcal{V}_0^*$ to be sufficiently rewarding; see the left panel of Figure 8. Incentives δ_* bind for both players simultaneously along the slanted segment of $\partial\mathcal{K}_{y_1,(A,I)}(\mathcal{V}_0^*)$. Repeating this procedure for all pure action profiles yields $\mathcal{K}^p_{y_1}(\mathcal{V}_0^*)$, shown in the right panel of Figure 8.

Next, we construct $\mathcal{K}_{y_1,\mathcal{A}}(\mathcal{W})$. It follows from (13) that the four extremal incentives are $\delta_*(B, S) = (-2, \frac{5}{8})$, $\delta_*(B, I) = (-\frac{2}{9}, \frac{5}{8})$, $\delta_*(A, S) = (-2, \frac{5}{4})$, and $\delta_*(A, I) = (-\frac{2}{9}, \frac{5}{4})$. The center panel of Figure 8 shows $\mathcal{K}_{y_1,\mathcal{A}}(\mathcal{V}_0^*)$ constructed via (19). It follows from Lemma 7.2

below that $\mathcal{K}_{y_1}(\mathcal{V}_0^*)$ is is the union of $\mathcal{K}^p_{y_1}(\mathcal{W})$ and $\partial\mathcal{K}_{y_1,\mathcal{A}}(\mathcal{V}_0^*)$.

The following lemma establishes that under Assumption 2.(ii) with strictly monotone λ, the boundaries of $\Psi_{y,\mathcal{A}_\diamond}$ overlap for the different supports $\hat{\mathcal{A}}$ so that $\mathcal{K}_y(\mathcal{W})$ can either be computed for all one-sided mixtures or for all pure and fully mixed action profiles.

Lemma 7.2. *Suppose that Assumption 2.(ii) is satisfied with strictly monotone λ. Then*

$$\mathcal{K}_y(\mathcal{W}) = \mathcal{K}^p_y(\mathcal{W}) \cup \partial\mathcal{K}_{y,\mathcal{A}}(\mathcal{W}) = \textstyle\bigcup_{i,a^i,\mathcal{A}_\diamond^{-i}} \partial\mathcal{K}_{y,a^i,\mathcal{A}_\diamond^{-i}}(\mathcal{W}).$$

Consider next the case where the players' actions affect the intensities of state transitions to multiple states and Assumption 2.(i) is satisfied, i.e., the successor states can be partitioned into $\mathcal{Y}_1$ and $\mathcal{Y}_2$ such that player i affects the intensity of transitions only to states $y_i \in \mathcal{Y}_i$. Let $\underline{w}^i_{y_i}$ and $\bar{w}^i_{y_i}$ denote player i's lowest and highest payoff in $\mathcal{W}_{y_i}$, respectively. Under Assumption 2.(i), $\mathcal{W}$-feasibility reduces to $\underline{w}^i_{y_i} \leq w + r\delta^i(y_i) \leq \bar{w}^i_{y_i}$ for all $y_i \in \mathcal{Y}_i$: because transitions to y_i provide no incentives to player $-i$, for any such $\delta^i(y_i)$ there exists $\delta^{-i}(y_i)$ with $w + r\delta(y_i) \in \mathcal{W}_{y_i}$. It follows that $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W})$ is the rectangle of all w that satisfy

$$\min_{\delta\in\Psi_{y,\mathcal{A}_\diamond}} \max_{y_i\in\mathcal{Y}_i} \underline{w}^i_{y_i} - r\delta^i(y_i) \leq w^i \leq \max_{\delta\in\Psi_{y,\mathcal{A}_\diamond}} \min_{y_i\in\mathcal{Y}_i} \bar{w}^i_{y_i} - r\delta^i(y_i).$$

To characterize $\mathcal{K}^p_y(\mathcal{W})$ when Assumption 2 is not satisfied, we rewrite the enforceability constraints in the following matrix form. Let $\Lambda^i_y(a)$ denote the matrix containing the column vectors $\lambda(y,\tilde{a}^i,a^{-i}) - \lambda(y,a)$ for $\tilde{a}^i \neq a^i$ and let $G^i_y(a)$ denote the row vector with entries $g(y,\tilde{a}^i,a^{-i}) - g(y,a)$ for $\tilde{a}^i \neq a^i$. Then $\Psi_{y,a}$ is the set of all δ that satisfy $\delta^i\Lambda^i_y(a) \leq -G^i_y(a)$ for $i = 1,2$. Let us reparametrize incentives via $\tilde{\delta} = h(w) + r\delta$, where $h(w) \in \mathbb{R}^{2\times|\mathcal{Y}|}$ is the matrix containing w in each column. Then $\delta \in \Psi_{y,a}(w,\mathcal{W})$ if and only if $\tilde{\delta} \in \bigtimes_{y'} \mathcal{W}_{y'}$ and

$$\tilde{\delta}^i\Lambda^i_y(a) - h^i(w)\Lambda^i_y(a) \leq -rG^i_y(a) \tag{20}$$

for $i = 1,2$, where $h^i(w) = (w^i,\dots,w^i)$ is row i of $h(w)$. Let $\mathcal{Z}_{y,a}$ denote the polyhedron of all $(w,\tilde{\delta})$ that satisfy (20). Then $\mathcal{K}_{y,a}(\mathcal{W})$ is the projection of $\mathcal{Z}_{y,a}\cap\mathbb{R}^2\times\bigtimes_{y'}\mathcal{W}_{y'}$ onto the first two dimensions. For any polygon approximation $\mathcal{W}'$ of $\mathcal{W}$, extremal points of $\mathcal{K}_{y,a}(\mathcal{W}')$ can be computed as the projection of the finitely many extremal points of $\mathcal{Z}_{y,a}\cap\mathbb{R}^2\times\bigtimes_{y'}\mathcal{W}'_{y'}$.

7.2 Regularity of the state-transition optimality equation

To discuss regularity of the state-transition optimality equation, consider first its restriction (17) to a fixed support $\mathcal{A}_\diamond$. Let us rewrite (17) as $N^\top w = \max_{v\in\mathcal{D}_{\mathcal{A}_\diamond}(w)} N^\top v$, and con-

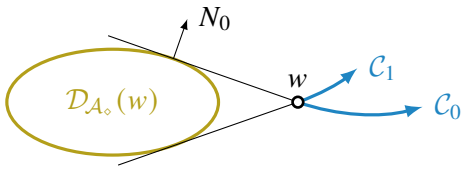

Figure 9: There are two solutions to (17) for every $w_0 \notin \mathcal{S}_{y,\mathcal{A}_\circ}(\mathcal{W})$.

sider first a payoff pair w that does not lie in the convex hull of $\mathcal{D}_{\mathcal{A}_\circ}(w)$ as in Figure 9. Then solutions to (17) evolve tangentially to $\mathcal{D}_{\mathcal{A}_\circ}(w)$. Since $\mathcal{D}_{\mathcal{A}_\circ}(w)$ is closed, it is strictly separated from w, hence there are two tangents to $\operatorname{conv}\mathcal{D}_{\mathcal{A}_\circ}(w)$ that go through any such w. Because $\mathcal{D}_{\mathcal{A}_\circ}(w)$ is bounded and continuous in w by Lemma 6.2, the direction of its two tangents through w change continuously. We call a solution $\mathcal{C}$ to (17) an *oriented solution* if N_w always points towards the same side of $\mathcal{C}$. Any oriented solution is continuously differentiable, and (17) has two such solutions for any $w_0 \notin \operatorname{conv}\mathcal{D}_{\mathcal{A}_\circ}(w_0)$. Initial conditions for the ODE (17) thus consist of a an initial point w_0 and a binary choice over starting directions.

When players are restricted to pure strategies, $\mathcal{D}_{\mathcal{A}_\circ}$ is convex-valued by Lemma 6.2, hence $w_0 \in \operatorname{conv}\mathcal{D}_{\mathcal{A}_\circ}(w_0)$ if and only if w_0 is stationary. Thus, oriented solutions to (17) exist and are continuously differentiable outside of $\mathcal{S}_{y,\mathcal{A}_\circ}(\mathcal{W})$. This is convenient for a numerical implementation of $\mathcal{B}_y(\mathcal{W})$: we can compute $\mathcal{S}_y(\mathcal{W})$ with the methods in Section 6, and need not worry about existence of solutions to (17) outside of it. For behavior strategies, $\mathcal{D}_{\mathcal{A}_\circ}$ need not be convex-valued, hence (17) may have no solutions at some non-stationary payoffs. We establish in Lemma 7.4 below that all such payoffs lie in the interior of $\mathcal{B}_y(\mathcal{W})$. Consider now (8) without restriction on the support, and write it as $N^\top w = \max_{x\in\mathcal{D}_y(w)} N^\top x$, where

$$\mathcal{D}_y(w) := \operatorname{conv}\big(\textstyle\bigcup_{\mathcal{A}_\circ} \mathcal{D}_{y,\mathcal{A}_\circ}(w)\big).$$

Let us denote by $\bar{\mathcal{K}}_y(\mathcal{W}) := \bigcup_{\mathcal{A}_\circ\subseteq\mathcal{A}(y)} \mathcal{K}_{y,\mathcal{A}_\circ}(\mathcal{W})$ the effective domain of $\mathcal{D}_y$, i.e., the set of all w where $\mathcal{D}_y(w) \neq \emptyset$. Since each $\mathcal{D}_{y,\mathcal{A}_\circ}$ is continuous by Lemma 6.2, $\mathcal{D}_y$ is continuous on $\bar{\mathcal{K}}_y(\mathcal{W}) \setminus \mathcal{K}_y(\mathcal{W})$ as the convex hull of a finite union of such maps. Let us denote by $\bar{\mathcal{S}}_y(\mathcal{W})$ the set of all w with $w \in \mathcal{D}_y(w)$. Outside of $\bar{\mathcal{S}}_y(\mathcal{W})$, a solution to (8) again evolves in the direction of one of the two tangents to $\mathcal{D}_y(w)$. There are two oriented solutions at each w, hence initial conditions of (8) consist of a payoff pair w and the selection of an orientation.

Proposition 7.3. *For any $w \in \bar{\mathcal{K}}_y(\mathcal{W}) \setminus \bar{\mathcal{S}}_y(\mathcal{W})$, there are two oriented solutions to* (8)*. If players are restricted to pure strategies or Assumption 2 holds, then oriented solutions are unique, continuously differentiable, and continuous in initial conditions outside of $\mathcal{K}_y(\mathcal{W})$.*

Finally, the following lemma establishes that existence of solutions outside of $\bar{\mathcal{S}}_y(\mathcal{W})$ is sufficient to characterize the boundary of $\mathcal{B}_y(\mathcal{W})$.

Lemma 7.4. *$\mathcal{S}_y(\mathcal{W}) \subseteq \bar{\mathcal{S}}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W})$ and any $w \in \partial\mathcal{B}_y(\mathcal{W}) \cap \bar{\mathcal{S}}_y(\mathcal{W})$ solves* (8).

7.3 Main result

To aggregate the results discussed so far, note that the state process is sufficient to provide incentives for (w, N) in $\Gamma_y(\mathcal{W}) := \left\{(w, N) \in \mathbb{R}^2 \times S^1 \;\middle|\; \exists\, \alpha \text{ and } (0, \delta) \in \Xi_{y,\alpha}(w, N, \mathcal{W})\right\}$. We denote by $\Gamma_y^p(\mathcal{W})$ the subset where such α exists in pure actions.

Theorem 7.5. *Suppose that Assumptions 1 and 2 hold. For any family $\mathcal{W}$ of compact and convex payoff sets and any state y, the set $\mathcal{B}_y(\mathcal{W})$ is the largest closed convex subset of $\mathcal{V}_y(\mathcal{W})$ that contains $\mathcal{S}_y(\mathcal{W})$ such that its boundary is continuously differentiable outside of $\mathcal{S}_y(\mathcal{W}) \cup \mathcal{K}_y(\mathcal{W})$ and (w, N_w) solves* (8) *within $\Gamma_y(\mathcal{W})$ and* (9) *outside of $\Gamma_y(\mathcal{W})$.*

Moreover, $\mathcal{B}_y^p(\mathcal{W})$ satisfies the same characterization under Assumption 1 for the pure-strategy restrictions of the optimality equations and the sets $\mathcal{V}$, $\mathcal{S}$, $\mathcal{K}$, and Γ.

Because we have discussed the optimality equations, $\mathcal{S}$, and $\mathcal{K}$ in the preceding subsections, it remains to provide an intuition for $\Gamma_y(\mathcal{W})$. Incentives through the public signal are used for two reasons. First, it may simply not be possible to provide sufficient incentives without it. Second, even if it is possible, the marginal cost of tangential transfers at $\beta = 0$ may be lower than the marginal cost of incentives through the destruction of value after state transitions.[8] A sufficient statistic for the public signal to be used is that $g(y, \alpha) + \delta\lambda(y, \alpha)$ lies below the tangent for any action profile α enforced by $\mathcal{W}$-feasible δ. Thus, as a general rule, the public signal is more likely to be used if stage-game payoffs in state y are high in direction N relative to its successor states. We illustrate this below for the regime-change game.

Example 5. The left panel in Figure 10 shows that the boundary of $\mathcal{B}_{y_1}(\mathcal{V}_0^*)$ is given by smooth solutions to the optimality equations outside the set of stationary payoffs. Along the efficient frontier, incentives are provided through state transitions exclusively because transitions to state y_2 weakly increase the weighted sum of payoffs $N^\top w$. Extremal stationary payoffs are constrained efficient, but other extremal payoffs are not: the continuation value $\hat{w} + r\hat{\delta}$ lies

[8] Formally, let $\beta(\alpha, \delta, N)$ be the lowest-variance transfers orthogonal to N that enforce α, given δ. Let α_* and δ_* maximize $f(\alpha, \delta) := N^\top(g(y, \alpha) + \delta\lambda(y, \alpha))$ among all $\mathcal{W}$-feasible δ that enforce α. Tangential transfers are too costly if all feasible directional derivatives of $\frac{r^2}{2}\kappa_y(w)\|T^\top\beta(\alpha, \delta, N)\sigma(y)\|^2$ at (α_*, δ_*) exceed those of f.

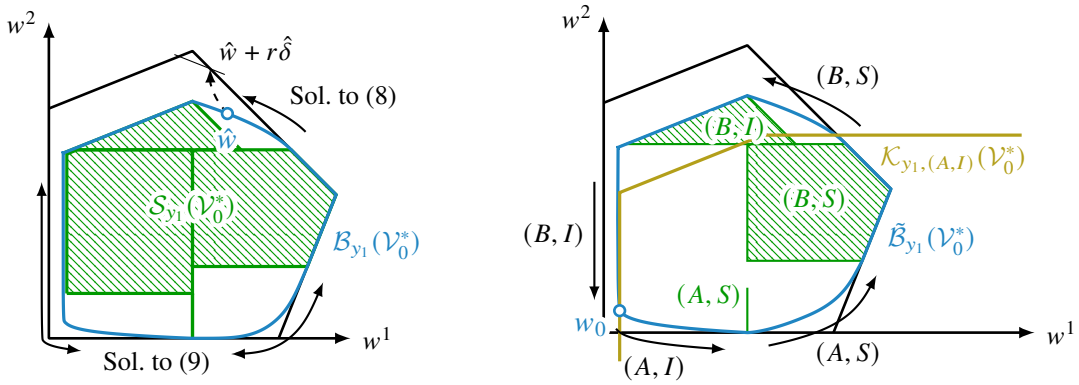

Figure 10: The left panel shows $\mathcal{B}_{y_1}(\mathcal{V}_0^*)$ in the regime-change game. The right panel shows $\tilde{\mathcal{B}}_{y_1}(\mathcal{V}_0^*)$ of all payoffs that can be attained with incentives from the state process only.

below the $\mathcal{V}_0^*$-feasible level in direction $N_{\hat{w}}$ because rewards exceeding $5/4$ would encourage player 2 to instigate a revolution. The remaining parts of the boundary are solutions to (9), where incentives are provided through both channels. An iterative application of $\mathcal{B}$ yields $\mathcal{E}$ in Figure 3. In this example, $\mathcal{B}$ coincides with $\mathcal{B}^p$ and, hence, $\mathcal{E}^p(r) = \mathcal{E}(r)$.

Example 6. For the sake of comparison, the right panel in Figure 10 shows the set $\tilde{\mathcal{B}}_{y_1}(\mathcal{V}_0^*)$ of all payoffs generated by $\mathcal{V}_0^*$ with incentives provided exclusively through the state process. Equivalently, this is the set $\mathcal{B}_{y_1}(\mathcal{V}_0^*)$ in a version of the regime-change game with a completely uninformative public signal. The solution to the state-transition optimality equation (8) has a kink at $w_0 \in \partial\mathcal{K}_{y_1,(A,I)}(\mathcal{V}_0^*)$ because (A, I) is not enforceable by $\mathcal{V}_0^*$-feasible incentives alone at any payoff w with $w^1 < w_0^1$. The kink disappears when the residual incentives can be provided through an informative public signal as seen in the left panel of Figure 10.

In Example 6, $\mathcal{K}_{y_1}(\mathcal{V}_0^*)$ affects only the smoothness of solutions to (8). Online Appendix E provides an example where $\mathcal{E}(y)$ contains a segment of positive length in $\mathcal{K}_y(\mathcal{E})$.

We conclude this section with the following sufficient condition for the pure- and behavior-strategy PPE payoff correspondences to coincide. This is the case in games where g, λ, and μ all have a product structure, hence players exclusively compete over which continuation states are reached. A patent race is an example of such a game; see Section 8.1 below.

Lemma 7.6. *If Assumptions 1 and 2.(i) hold in state y and $g^i(y, a)$ depends only on player i's actions, then $\mathcal{B}_y(\mathcal{W}) = \mathcal{B}_y^p(\mathcal{W})$. If these assumptions hold in all states, then $\mathcal{E}(r) = \mathcal{E}^p(r)$.*

8 Discussion

8.1 Example of a two-stage patent race

An example of a stochastic game that is neither absorbing nor irreducible is a multi-stage patent race. On the path to developing a patentable technology, two competing labs $i = 1, 2$ have to achieve several intermediate milestones that are publicly osberved. For instance, a vaccine must pass multiple trial phases, and a self-driving car requires the development of various auxiliary technologies. For specificity, consider a winner-take-all two-stage patent race. The value of the patent is 1, accruing to the first lab to complete both stages of research. There are two absorbing states y_i in which lab i holds the patent, and four intermediate states y_{s_1,s_2} for $s_i \in \{0, 1\}$ indicating that lab i has completed s_i stages of research. The initial state is $y_{0,0}$, and the feasible state transitions are indicated in Figure 11. We parametrize the game through the research intensity $A_t^i \in \{0, \dots, 5\}$ chosen by each lab i, i.e., state transitions are given by

$$\lambda_{y_{0,0},y_{1,0}}(A_t) = \lambda_{y_{1,s_2},y_1}(A_t) = A_t^1, \qquad \lambda_{y_{0,0},y_{0,1}}(A_t) = \lambda_{y_{s_1,1},y_2}(A_t) = A_t^2.$$

Each lab i incurs a flow cost of $(A_t^i)^2\,\mathrm{d}t$, that is, $g^i(y, a) = -(a^i)^2$ in any non-absorbing state y. The flow cost is convex because good research ideas and personnel are scarce, hence additional research effort is more costly at the margin. The stock price X^i of each lab i reflects market beliefs about the chances of i's success, and it is publicly observed. We suppose that $\mathrm{d}X_t^i = \gamma A_t^i\,\mathrm{d}t + \mathrm{d}Z_t^i$ for a constant $\gamma \geq 0$ and two independent Brownian motions Z^1 and Z^2. The parameter γ captures the informativeness of the stock prices: if $\gamma > 0$, market beliefs about the chance of success carry some information about research efforts, whereas if $\gamma = 0$, movements in the stock price are driven purely by speculation. Both labs discount the future at rate $r = 1/8$. Each lab can ensure a payoff of 0 by never carrying out any research, hence $\mathcal{V}_0^*$ is the triangle with extremal points $(1, 0)$, $(0, 1)$, and $(0, 0)$; see the right panel of Figure 11.

Because state transitions and flow payoffs have a product structure, Lemma 7.6 implies that all PPE are attained by pure strategies. We begin by computing the Markov-perfect equilibria via Proposition 6.9. Independent of γ, there are two MPE $a_{*,k}$ for $k = 1, 2$, given by

$$a_{*,k}(y_{0,0}) = (k, k), \quad a_{*,k}(y_{1,0}) = (2, 1), \quad a_{*,k}(y_{0,1}) = (1, 2), \quad a_{*,k}(y_{1,1}) = (3, 3).$$

To compute the full PPE correspondence efficiently, we proceed backwards through the graph in the left panel of Figure 11. First, the payoff sets $\mathcal{E}(y_1) = \{(1, 0)\}$ and $\mathcal{E}(y_2) = \{(0, 1)\}$ in

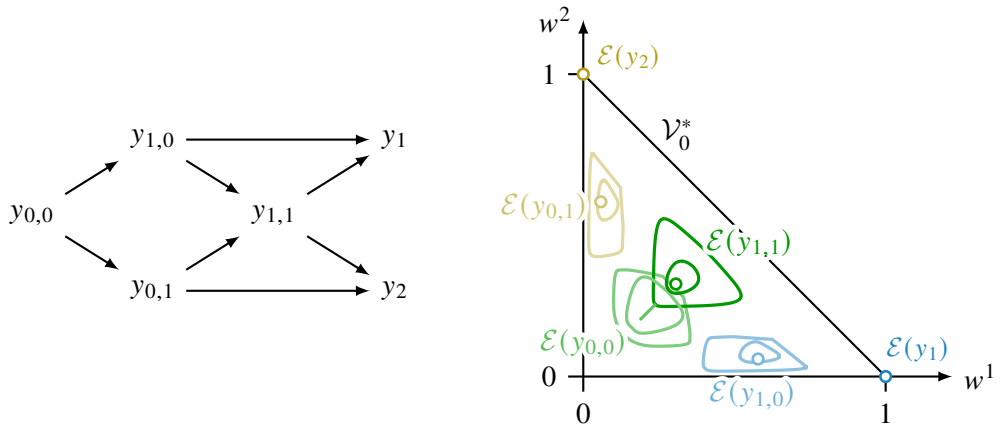

Figure 11: Left panel: feasible state transitions in the two-stage patent race. Right panel: PPE payoff correspondence for $\gamma = 0$ (small), $\gamma = 2$ (intermediate), and $\gamma = 5$ (large).

the terminal states are simply given by the value of the patent. Second, because only states y_1 and y_2 can be reached from $y_{1,1}$, a single application of $\mathcal{B}$ yields $\mathcal{E}(y_{1,1}) = \mathcal{B}_{y_{1,1}}(\mathcal{E}(y_1), \mathcal{E}(y_2))$. Similarly, the PPE payoffs with initial states $y_{1,0}$, $y_{0,1}$, and finally $y_{0,0}$ follow from a single application of $\mathcal{B}$ after the PPE payoffs in their continuation states have been computed. In this example, the differential characterizations decouple entirely across states. We discuss how the computation across states decouples more generally in Section 8.2. The right panel of Figure 11 shows the PPE payoff correspondence for the three values $\gamma = 0$, $\gamma = 2$, and $\gamma = 5$. The family of PPE payoffs is increasing in γ. We show in Section 8.3 that such a monotonicity result holds more generally. For $\gamma = 0$, the only PPE are public randomizations of the two Markov-perfect equilibria. We can read off equilibrium behavior from the PPE payoff correspondence through the action profiles and continuation promises that attain the maximum in the optimality equations. Those are illustrated in Figures 12 and 13 for $\gamma = 5$.

Let us first analyze equilibrium behavior in state $y_{1,1}$, equivalent to a one-stage patent race. The entire boundary of $\mathcal{E}(y_{1,1})$ solves the optimality equation (9), hence incentives are provided through the public signal at all times. Because research success is linear in effort but the cost is convex, low research intensities are chosen along the efficient frontier. PPE on the efficient frontier can be interpreted as a tacit non-compete agreement, where compliance is monitored via stock prices. Outside of payoffs $\hat{w}_1$ and $\hat{w}_2$ only a single lab engages in research; the region between $\hat{w}_1$ and $\hat{w}_2$ corresponds to a transitional phase where the labs switch roles. An increase of $X^1 - X^2$ leads to a counterclockwise change in the continuation value. Above $\hat{w}_2$, where only lab 2 is supposed to engage in research, this may indicate that

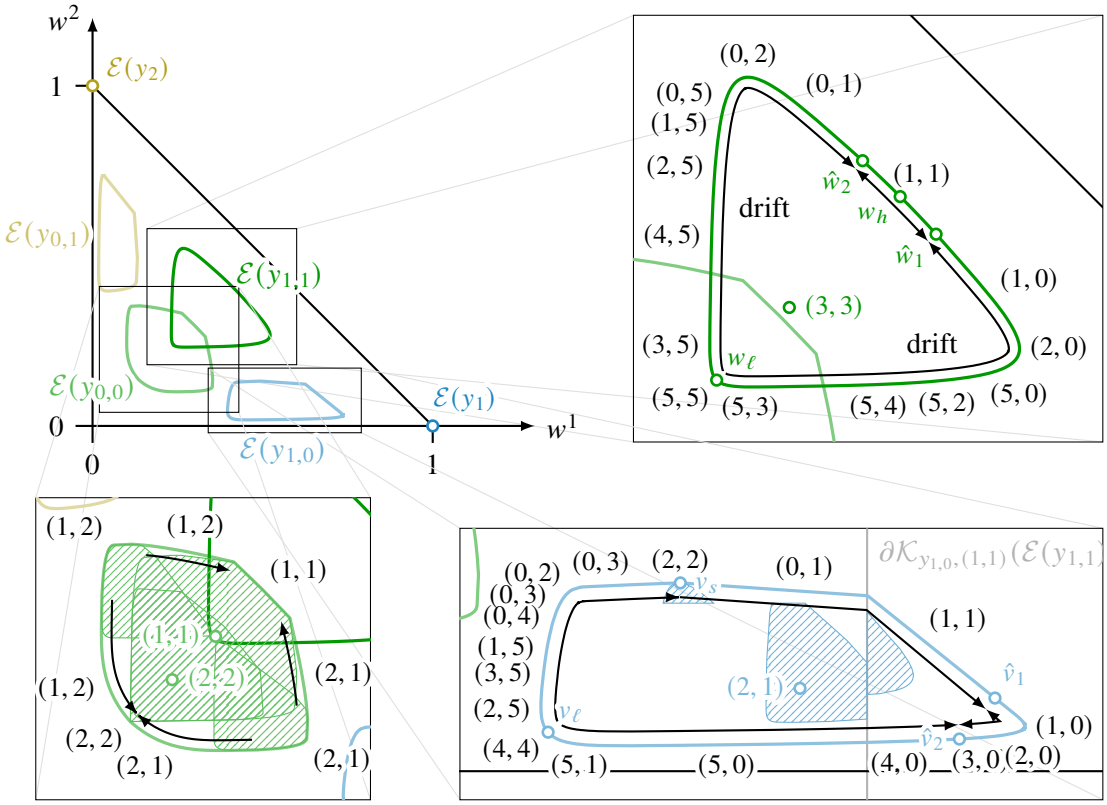

Figure 12: PPE payoff correspondence for $\gamma = 5$. The hatched ares correspond to stationary payoffs. The generically unique actions profiles played on the boundary are indicated outside of the payoff sets. Arrows in the interior indicate the direction of the drift. Finally, MPE payoffs and action profiles are indicated in color in the interior of the payoff sets.

lab 1 has violated the agreement. If $X^1 - X^2$ increases sufficiently, lab 2 increases its research efforts to make such violations by lab 1 unattractive. Below $\hat{w}_1$, the same phenomenon suggests that lab 2 has complied, hence lab 2 is rewarded with the permission to carry out research when the continuation value crosses $\hat{w}_1$. The continuation value drifts away from payoff pairs w_ℓ and w_h towards $\hat{w}_1$ and $\hat{w}_2$. The drift between w_h and $\hat{w}_i$ is moderate so that a switch of roles is not uncommon along an equilibrium path. The drift between w_ℓ and $\hat{w}_i$, however, is rather strong, hence a lab rarely increases their research intensity beyond 2: it will have to see very strong evidence that their competitor has violated the agreement before carrying out such a costly punishment.

The PPE that attains the lowest symmetric payoff pair w_ℓ begins with maximal research effort $(5, 5)$. If X^1 initially outperforms X^2, the continuation value again moves counterclockwise and lab 1 emerges as the market favorite. The drift from w_ℓ to $\hat{w}_1$ is strong, hence lab 2

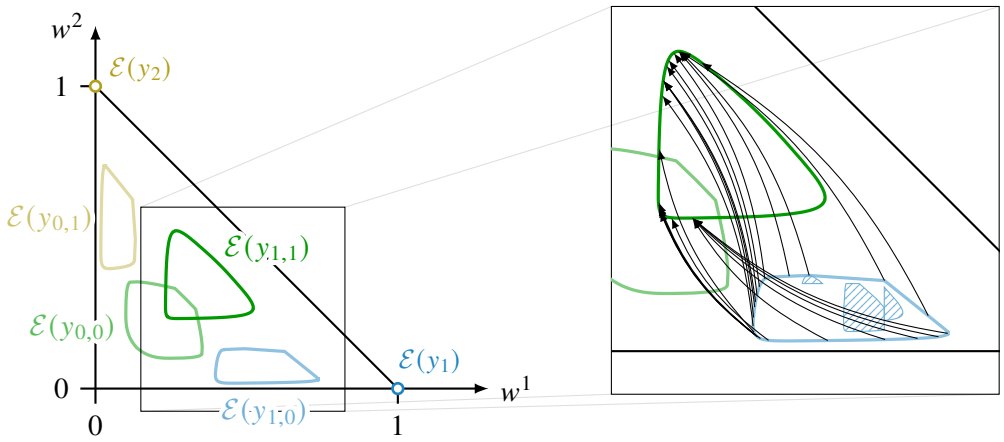

Figure 13: State transitions if the follower catches up.

reduces its research intensity rather quickly until lab 1 is the sole active researcher. Lab 1 then also gradually lowers its research effort and the labs enter the tacit non-compete regime described above. The period of high research activity is short because the prize at the end is fixed. High symmetric research effort beyond the Markov-perfect level $(3, 3)$ is not sustainable.

In state $y_{1,0}$, the action profiles along the boundary yield the equilibrium behavior until the next breakthrough and the continuation promise $\delta(y_{1,1})$ indicates what continuation PPE is played if lab 2 happens to catch up; see Figure 13. If lab 2 exerts a positive amount of effort at a boundary payoff w, transitions to state $y_{1,1}$ happen on the path and, generally, a continuation PPE is played that maximizes $\mathcal{E}(y_{1,1})$ for payoff weights close to N_w. Along the lower boundary of $\mathcal{E}(y_{1,0})$, lab 2 is not supposed to conduct research, hence transitions to state $y_{1,1}$ lie off the path. Violations by lab 2 are deterred most effectively by a continuation PPE A^* that yields the lowest utility for lab 2. Because A^* is also a poor continuation PPE for the leading lab 1, Figure 13 illustrates that the leading lab's continuation value is always low if the follower catches up—there is momentum to research and it shifts away from the former leader.

The efficient frontier between v_s and $\hat{v}_2$ again corresponds to a tacit non-compete agreement, though the asymmetry between the labs affect its structure. The leading lab 1 is willing to enter a non-compete agreement only if the agreement is highly advantageous. The drift rate indicates that the continuation value fluctuates around the region between $\hat{v}_1$ and $\hat{v}_2$, where lab 1 invests a small amount of resources and lab 2 drops in and out of research. Because off-path continuations are detrimental to lab 1, it increases its research intensity for much lower evidence that lab 2 has cheated than in state $y_{1,1}$. Research intensities of 3 or 4 are not uncom-

mon on the path. As in state $y_{1,1}$, the lowest-utility payoff pair v_ℓ is attained by a high initial research effort and a quick reduction of research effort, first by the perceived underdog, then by the market favorite. If lab 1 becomes the market favorite, the labs reach the non-compete region, whereas if lab 2 initially becomes the market favorite, the labs eventually play the stationary profile $(2, 2)$ when v_s is reached. Because the labs are further away from the patent, research intensities both at v_ℓ and at stationary payoffs are lower than in state $y_{1,1}$.

Finally, the boundary of $\mathcal{E}(y_{0,0})$ consists stationary payoffs and solutions to (8) exclusively. Notably, stock price information is completely ignored. Because the patent is even more distant, research efforts in the MPE are weakly lower than in state $y_{1,0}$.

Even though the PPE payoff correspondence is monotone in γ, which type of information is used along the boundary is not. A more informative public signal is not necessarily used more in all states. For $\gamma = 2$, the boundary of $\mathcal{E}(y_{0,0})$ is entirely a solution to (9), hence the public signal is used everywhere along the boundary. By contrast, it is completely ignored if $\gamma = 0$ or $\gamma = 5$. Driving this non-monotonicity is the fact that $\mathcal{E}(y_{1,1})$ is smaller for $\gamma = 2$ than for $\gamma = 5$, hence fewer incentives can be provided through state transitions alone in earlier states.

Grossman and Shapiro (1987) characterize the MPE of this model and extensions to arbitrary increasing and convex cost functions. Their main conclusions are "that the leader always devotes more resources to R&D than does the follower, but that if the latter happens to catch up, both firms intensify their effort." Our model allows us to expand on their results:

(i) If an informative pubic signal is available, labs engage in a tacit non-compete agreement on the efficient frontier of the equilibrium correspondence to keep research costs low.

(ii) The leader devotes more resources than the follower in the non-compete agreement because the leader is in a better position to negotiate.

(iii) Since research intensities remain low on the efficient frontier, our analysis suggests that research intensities in the MPE increase when the follower catches up not because the labs' dynamics shift, but because the discounted expected prize of the patent increases.

8.2 Minimizing state interdependencies

In the patent race, we computed $\mathcal{E}(y)$ sequentially, each with a single application of $\mathcal{B}_y$. This is not possible in the regime-change game because the states *communicate*, i.e., both states can be reached from each other. If there are communicating states, we approximate $\mathcal{E}$ through Proposition 4.9, but we do not have to approximate all states simultaneously.

A set of states $\mathcal{Y}_0 \subseteq \mathcal{Y}$ is a *communicating class* if each state within $\mathcal{Y}_0$ can be reached from any other state in $\mathcal{Y}_0$ with positive probability. Communicating classes of a Markov process form a directed acyclic graph; similarly to the left panel in Figure 11. The PPE payoff correspondence can be computed sequentially across communicating classes by starting with a terminal communicating class and proceeding backwards through the graph. Because the graph is acyclic, at any point in this procedure, PPE payoffs of all reachable states outside the current communicating class have already been computed. PPE payoffs within each communicating class $\mathcal{Y}_0$ are given by a single application of $\mathcal{B}$ if and only if $\mathcal{Y}_0$ is a singleton as in the patent race. Otherwise, we apply Proposition 4.9 to $\mathcal{Y}_0$.

8.3 Monotonicity in game primitives

In this section, we perform comparative statics on the game primitives r, λ, μ, and σ. We begin with a monotonicity result in the informativeness of the public signal.

Lemma 8.1. *Let $\tilde{\mu}$ and $\tilde{\sigma}$ be such that for each state y, there exists a diagonal matrix D_y with diagonal elements in $[-1,1] \setminus \{0\}$ such that $\tilde{\sigma}^{-1}(y) D_y \tilde{M}^i_y(a) = \sigma^{-1}(y) M^i_y(a)$ for each $a \in \mathcal{A}(y)$.*[9] *Then $\mathcal{E}(\tilde{\mu}, \tilde{\sigma}; y) \supseteq \mathcal{E}(\mu, \sigma; y)$ and $\mathcal{E}^p(\tilde{\mu}, \tilde{\sigma}; y) \supseteq \mathcal{E}^p(\mu, \sigma; y)$ for each state y.*

The existence of such a diagonal matrix D_y means that the signal-to-noise ratio of each dimension of the public signal is at least as high under $(\tilde{\mu}, \tilde{\sigma})$ as it is under (μ, σ). Some dimensions may be scaled more than others, but scaling is independent of the actions taken. The key idea is that the signal with the higher signal-to-noise ratio $\left|\tilde{\sigma}^{-1}(y)\tilde{\mu}(y, \cdot)\right|$ can be garbled with public randomization until it is as informative as under parameters (μ, σ). A PPE A under (μ, σ) gives rise to a PPE under $(\tilde{\mu}, \tilde{\sigma})$ that is the same function from histories to actions, except that the public signal is replaced with the garbled signal.

The same garbling logic cannot be applied to derive a monotonicity result in λ: while garbling would make state transitions less informative, players cannot ignore state transitions between non-equivalent states.[10] For the same reason, the PPE payoff correspondence need not vary monotonically in the discount rate since that affects the relative weight of current and future states. We provide a counterexample to both in Online Appendix E.

While the PPE payoff correspondence does not change monotonically in λ and r individ-

[9] Recall here that $\sigma^{-1}(y)$ is used to denote the right-inverse $\sigma(y)^\top(\sigma(y)\sigma(y)^\top)^{-1}$.

[10] Two states y and y' are equivalent if $\mathcal{A}$, g, μ, and σ coincide in the two states and $\lambda_{y,y'}(a) = \lambda_{y',y}(a)$ as well as $\lambda_{y,\tilde{y}} = \lambda_{y',\tilde{y}}$ for $\tilde{y} \notin \{y, y'\}$. A transition from state y to y' or vice versa is then just a public signal.

ually, the following time-change result establishes that a decrease of both by the same factor $\eta \in (0,1)$ has the same effect as an increase of the signal-to-noise ratio of the public signal because players get to observe more signal realizations in the same discounted time.

Lemma 8.2. *Let $\tilde{\mu}$, $\tilde{\sigma}$ and $\eta > 0$ be such that $\tilde{\sigma}^{-1}(y)\tilde{\mu}(y) = \sqrt{\eta}\sigma^{-1}(y)\mu(y)$ for each y. Then a strategy profile A is a PPE for the game primitives $(r, \mu, \sigma, \lambda)$ if and only if $(A_{\eta t})_{t\geq 0}$ is a PPE with respect to the game primitives $(\eta r, \tilde{\mu}, \tilde{\sigma}, \eta\lambda)$. Moreover, for every $t \geq 0$,*

$$\tilde{W}_t((A_{\eta s})_{s\geq 0}; \eta r, \tilde{\mu}, \tilde{\sigma}, \eta\lambda) = W_{\eta t}(A; r, \mu, \sigma, \lambda) \quad a.s.,$$

where $\tilde{W}$ is the continuation value (2) *with respect to the time-changed filtration $(\mathcal{F}_{\eta t})_{t\geq 0}$.*

Lemma 8.2 shows that the family of PPE payoffs depends on the game primitives r, μ, λ, and σ only through the quantities λ/r and $\sigma^{-1}(y)\mu(y)/\sqrt{r}$. Similar square-root laws for repeated games with Brownian information, but not for Poisson information or stochastic games, have been established in Faingold and Sannikov (2011) and Bernard and Frei (2016). Together with Lemma 8.1, we obtain the following joint monotonicity result in λ and r.

Corollary 8.3. *$\mathcal{E}(\eta\lambda, \eta r; y) \supseteq \mathcal{E}(\lambda, r; y)$ and $\mathcal{E}^p(\eta\lambda, \eta r; y) \supseteq \mathcal{E}^p(\lambda, r; y)$ for any $\eta \in (0,1)$.*

9 Conclusion

This paper characterizes perfect public equilibria, state-order dependent equilibria, and Markov-perfect equilibria in continuous-time stochastic games with imperfect public monitoring. Crucially, the analysis holds for any discount rate $r > 0$, which preserves the correspondence from initial states to equilibrium payoffs. The central restriction—and only restriction when players use pure strategies—on the state process is that the state space is finite. In particular, the methodology is not limited to irreducible games or absorbing games.

There are four avenues for future research that appear promising. First, the two-player restriction and the product structure are satisfied in dynamic principal-agent models. This framework could be used to extend the dynamic contracting literature initiated by Sannikov (2008) to models, in which the agent's wage level additionally depends on a discrete state variable, which may or may not be affected by the agent's actions. Second, identifiability assumptions are key to folk theorems and limit results as players get arbitrarily patient. The techniques in the paper may thus help establish a folk-theorem for non-irreducible stochastic games where, unlike in Pęski and Wiseman (2015), the public signal remains imperfectly

informative in the limit. Third, it would be interesting to extend the framework to deterministic state transitions so that the model can accommodate quarterly reports by companies or regular political elections. Fourth, and more on the applied side, the fact that the correspondence of initial states to PPE payoffs is preserved allows us to draw conclusions about the values the players or a designer should assign to different initial states. For example, if a policy is implemented that requires players to take on two asymmetric roles in alternating fashion, the imbalance created by the initial role assignment can be properly compensated.

References

Abreu, Dilip, Benjamin Brooks, and Yuliy Sannikov. 2020. "Algorithms for Stochastic Games With Perfect Monitoring." *Econometrica*, **88(4)**: 1661–1695.

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, **58(5)**: 1041–1063.

Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review*, **91(4)**: 938–963.

Bergemann, Dirk, and Juuso Välimäki. 1996. "Learning and Strategic Pricing." *Econometrica*, **64(5)**: 1125–1149.

Bernard, Benjamin. 2024. "Continuous-Time Games with Imperfect and Abrupt Information." *The Review of Economic Studies*, **91(4)**: 1988–2052.

Bernard, Benjamin, and Christoph Frei. 2016. "The folk theorem with imperfect public information in continuous time." *Theoretical Economics*, **11(2)**: 411–453.

Bernheim, B. Douglas, and Debraj Ray. 1989. "Markov perfect equilibria in altruistic growth economies with production uncertainty." *Journal of Economic Theory*, **47(1)**: 195–202.

Dutta, Prajit K. 1995. "A Folk Theorem for Stochastic Games." *Journal of Economic Theory*, **66(1)**: 1–32.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *The Review of Economic Studies*, **62(1)**: 53–82.

Faingold, Eduardo, and Yuliy Sannikov. 2011. "Reputation In Continuous-Time Games." *Econometrica*, **79(3)**: 773–876.

Faingold, Eduardo, and Yuliy Sannikov. 2020. "Moral Hazard in Stochastic Differential Games: Beyond Markov Equilibrium." working paper.

Fudenberg, Drew, and Yuichi Yamamoto. 2011. "The folk theorem for irreducible stochastic games with imperfect public monitoring." *Journal of Economic Theory*, **146(4)**: 1664–1683.

Grossman, Gene M., and Carl Shapiro. 1987. "Dynamic R & D Competition." *The Economic Journal*, **97(386)**: 372–387.

Hörner, Johannes, Takuo Sugaya, Satoru Takahashi, and Nicolas Vieille. 2011. "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem." *Econometrica*, **79(4)**: 1277–1318.

Karatzas, I., and S. Shreve. 2017. *Methods of Mathematical Finance. Probability Theory and Stochastic Modelling*, Springer New York.

Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica*, **73(1)**: 39–68.

Kirman, Alan P., and Matthew J. Sobel. 1974. "Dynamic Oligopoly with Inventories." *Econometrica*, **42(2)**: 279–287.

Kohlberg, Elon. 1974. "Repeated Games with Absorbing States." *The Annals of Statistics*, **2(4)**: 724–738.

Levhari, David, and Leonard J. Mirman. 1980. "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution." *The Bell Journal of Economics*, **11(1)**: 322–334.

Mertens, Jean-François, and Abraham Neyman. 1981. "Stochastic games." *International Journal of Game Theory*, **10(1)**: 53–66.

Mertens, Jean-François, Abraham Neyman, and Dinah Rosenberg. 2009. "Absorbing Games with Compact Action Spaces." *Mathematics of Operations Research*, **34(2)**: 257–262.

Neyman, Abraham. 2017. "Continuous-time stochastic games." *Games and Economic Behavior*, **104**: 92–130.

Pęski, Marcin, and Thomas Wiseman. 2015. "A folk theorem for stochastic games with infrequent state changes." *Theoretical Economics*, **10(1)**: 131–173.

Sannikov, Yuliy. 2007. "Games with Imperfectly Observable Actions in Continuous Time." *Econometrica*, **75(5)**: 1285–1329.

Sannikov, Yuliy. 2008. "A Continuous-Time Version of the Principal–Agent Problem." *The Review of Economic Studies*, **75(3)**: 957–984.

Simon, Leo K., and Maxwell B. Stinchcombe. 1989. "Extensive Form Games in Continuous Time: Pure Strategies." *Econometrica*, **57(5)**: 1171–1214.

Yeltekin, Sevin, Yongyang Cai, and Kenneth L. Judd. 2017. "Computing Equilibria of Dynamic Games." *Operations Research*, **65(2)**: 337–356.

Outline of the appendices

Appendix A contains the mathematical foundation of the model and the proofs of those results that rely on it with the exception of Lemmas 4.1, 4.3, and 4.7, as well as Proposition 4.9—their proofs are similar to proofs in Sannikov (2007) and Bernard (2024) and are deferred to Appendix H in the supplemental information found online.[11] Appendix B contains the proofs related to feasibility and state-order dependence. Appendix C contains the proofs of Lemmas 4.5, 5.5, and 7.6 that relate game primitives and enforceability. Online Appendix F establishes regularity of the optimality equations. The results that characterize $\mathcal{S}_y(\mathcal{W})$ and $\mathcal{K}_y(\mathcal{W})$ or simplify their computation are proven in Online Appendix G. Finally, the perturbation argument in Section 5.1 is proven in Appendix I in the supplemental information online.[11]

A Mathematical foundation

Let $\mathcal{Z}$ denote the set of ordered pairs (y, y'), for which there exists a pure action profile $a \in \mathcal{A}(y)$ with $\lambda_{y,y'}(a) > 0$. Let $(\Omega, \mathcal{F}, P)$ be a probability space containing a d-dimensional Brownian motion Z and a Poisson process $J^{y,y'}$ with intensity 1 for each $(y, y') \in \mathcal{Z}$ such that $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ and Z are mutually independent. The state process S is defined by

$$S_0 = y_0, \qquad S_t := \begin{cases} S_{t-} & \text{if } \Delta J_t^{S_{t-},y} = 0 \text{ for all } y \text{ with } (S_{t-}, y) \in \mathcal{Z}, \\ y & \text{if } \Delta J_t^{S_{t-},y} = 1, \end{cases} \tag{21}$$

i.e., S is piecewise constant and jumps to y at time t if and only if $\Delta J_t^{S_{t-},y} = 1$. The family of probability measures $Q^A = (Q_t^A)_{t\geq 0}$ is defined via its density process

$$\frac{\mathrm{d}Q_t^A}{\mathrm{d}P} = \mathcal{E}_t\left(\int_0^{\cdot} \mu(S_s, A_s)^\top (\sigma^{-1}(S_s))^\top \mathrm{d}Z_s + \sum_{(S_{s-},y)\in\mathcal{Z}} \int_0^{\cdot} (\lambda_{S_{s-},y}(A_{s-}) - 1)\left(\mathrm{d}J_s^{S_{s-},y} - \mathrm{d}s\right)\right), \tag{22}$$

where $\mathcal{E}_t(X) = \exp(X_t - X_0 - \frac{1}{2}\langle X^c\rangle_t) \prod_{0<s\leq t}(1+\Delta X_s)\mathrm{e}^{-\Delta X_s}$ is the Doléans-Dade exponential of process X. By Girsanov's theorem, the public signal indeed takes the form (1) under Q^A and state transitions occur with instantaneous intensities $\lambda_{S_{t-},y}(A_{t-})$.

$\mathbb{F}$ contains the history of all processes $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$, not just of S, so that the density process is adapted to $\mathbb{F}$. Under Q^A, the instantaneous intensities of the processes $J^{y,y'}$ are equal to 1 on $\{S \neq y\}$ so that players learn nothing from these processes. A mathematical foundation

[11] See here: https://benjamin-bernard.com/research/stochastic-games_SI.pdf.

based on processes $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ ensures that each Q_t^A is absolutely continuous with respect to some reference measure P. Then, demanding that a statement holds $P \otimes Lebesgue$-almost everywhere ensures that it holds almost everywhere both on and off the path.

A.1 Monotonicity

Proof of Lemma 8.1. For any y_0 and $w_0 \in \mathcal{E}(\mu, \sigma; y_0)$, let A be a PPE attaining w_0 for parameters (μ, σ) on a stochastic framework $\left(\Omega, \mathcal{F}, \mathbb{F}, P, Z, (J^{y,y'})_{(y,y')\in\mathcal{Z}}\right)$ with induced state process S. Abbreviate $W = W(y_0, A)$ and let β, δ, and M be the processes that satisfy (5). By assumption on $\tilde{\mu}$ and $\tilde{\sigma}$, there exist constants c_y such that $\tilde{\sigma}^{-1}(y)D_y(\tilde{\mu}(y,a) + c_y) = \sigma^{-1}(y)\mu(y,a)$ for each action profile $a \in \mathcal{A}(y)$. By an equivalent change of the reference measure P, we may assume without loss of generality that $c_y = 0$. Moreover, by Lemma 8.2 it is sufficient to show that w_0 can be attained by a PPE for parameters μ and $D_y\sigma(y)$. Here, it is relevant that D_y has non-zero eigenvalues so that $D_y\sigma(y)$ has a right-inverse $\sigma^{-1}(y)D_y^{-1}$ and the density process in (22) with respect to $(\mu, D_y\sigma)$ is well-defined. The goal is to split the Brownian motion into two orthogonal components $Z = R_y\hat{Z} + R_y^{\perp}\hat{Z}^{\perp}$, where R_y captures the relevant information in D_y and the second term corresponds to public randomization.

Abbreviate $\sigma_y := \sigma(y)$ and define $R_y := \sigma_y{}^{\top}(\sigma_y\sigma_y{}^{\top})^{-1}D_y\sigma_y$ so that $\sigma_y R_y = D_y\sigma_y$. Let v be an eigenvector of R_y with eigenvalue η. If $v \in \ker(\sigma_y)$, then $\eta = 0$. If $v \notin \ker(\sigma_y)$, then $D_y\sigma_y v = \sigma_y R_y v = \eta\sigma_y v$, hence η is the eigenvalue of eigenvector $\sigma_y v$ of D_y. This implies that R_y has eigenvalues in $[-1, 1]$ and $\ker(\sigma_y R_y) = \ker(\sigma_y)$. Symmetry of $\sigma_y\sigma_y{}^{\top}$ implies

$$R_y{}^{\top} = \sigma_y{}^{\top}D_y(\sigma_y\sigma_y{}^{\top})^{-1}\sigma_y = \sigma_y{}^{\top}(\sigma_y\sigma_y{}^{\top})^{-1}D_y\sigma_y = R_y,$$

i.e., R_y is symmetric. Thus, R_y allows a decomposition into $R_y = Q_y{}^{\top}E_yQ_y$, where Q_y is orthogonal and E_y is a diagonal matrix with the eigenvalues of R_y on the diagonal. Set $E_y^{\perp} := \sqrt{\mathrm{I} - E_y^2}$, where I is the identity matrix, and define $R_y^{\perp} := Q_y{}^{\top}E_y^{\perp}Q_y$. Since $R_y^2 + (R_y^{\perp})^2 = \mathrm{I}$, it follows from Lévy's characterization theorem that $\hat{Z} := R_yZ + R_y^{\perp}Z^{\perp}$ is a standard Brownian motion with respect to $\hat{\mathbb{F}}$. Define the diagonal matrix $\hat{E}_y$ with diagonal elements

$$\hat{E}_y^{kk} = \begin{cases} 0 & \text{if } (E_y^{\perp})^{kk} = 0, \\ 1/(E_y^{\perp})^{kk} & \text{otherwise,} \end{cases}$$

and set $\hat{R}_y := Q_y{}^{\top}\hat{E}_yQ_y$. Note that $\Pi_y := \hat{E}_yE_y^{\perp}$ has diagonal elements $\Pi_y^{kk} = 1_{\{(E_y^{\perp})^{kk}>0\}}$,

i.e., it is the projection onto the image of $E_y^\perp$. It follows that $\Pi_y E_y^\perp = E_y^\perp$ and, hence,

$$Q_y{}^\top \Pi_y Q_y R_y^\perp = Q_y{}^\top \Pi_y E_y^\perp Q_y = Q_y{}^\top E_y^\perp Q_y = R_y^\perp. \tag{23}$$

Let $\hat{Z}^\perp := \hat{R}_y(Z - R_y\hat{Z})$. It follows from $\mathrm{d}\langle \hat{Z}^\perp, \hat{Z}\rangle_t = 0$ that $\hat{Z}^\perp$ is orthogonal to $\hat{Z}$. Moreover,

$$\mathrm{d}\langle \hat{Z}^\perp\rangle_t = (\hat{R}_y)^2\,\mathrm{d}\langle Z - R_y\hat{Z}\rangle_t = (\hat{R}_y)^2(\mathrm{I} + R_y^2 - 2R_y^2)\,\mathrm{d}t = \hat{R}_y R_y^\perp\,\mathrm{d}t$$

shows that $\hat{Z}^\perp$ is a Brownian motion with correlation matrix $\hat{R}_y R_y^\perp$ by Lévy's characterization theorem. In particular, $\hat{Z}^\perp$ is a martingale. Finally, it follows from (23) that $R_y\hat{Z} + R_y^\perp\hat{Z}^\perp = Z$.

Substituting this decomposition of Z into (5) for the PPE A yields

$$\begin{aligned}\mathrm{d}W_t^i = {} & r(W_t^i - g^i(S_t, A_t))\,\mathrm{d}t + r\beta_t^i\,(\sigma(S_t)R_{S_t}\,\mathrm{d}\hat{Z}_t - \mu(S_t, A_t)\,\mathrm{d}t) \\ & + r\sum_{y\in\mathcal{Y}} \delta_t^i(y)\,(\mathrm{d}J_t^y - \lambda_{S_{t-},y}(A_t)\,\mathrm{d}t) + \mathrm{d}M_t^i + r\beta_t^i\sigma(S_t)R_{S_t}^\perp\,\mathrm{d}\hat{Z}_t^\perp.\end{aligned} \tag{24}$$

Because $\int_0^\cdot \sigma(S_t)\,\mathrm{d}Z_t$ is strongly orthogonal to M^i by Lemma 4.1 and we have chosen $Z^\perp$ strongly orthogonal to M^i, it follows from

$$\sigma(S_t)R_{S_t}\,\mathrm{d}\hat{Z}_t = D_{S_t}^2\sigma(S_t)\,\mathrm{d}Z_t + \sigma(S_t)R_{S_t}R_{S_t}^\perp\,\mathrm{d}Z_t^\perp$$

that the diffusion term in (24) is strongly orthogonal to M^i and, hence, also to the martingale $\hat{M}^i$ defined by $\mathrm{d}\hat{M}^i := \mathrm{d}M^i + r\beta_t^i\sigma(S_t)R_{S_t}^\perp\,\mathrm{d}\hat{Z}_t^\perp$. It follows that $(W, A, \beta, \delta, \hat{Z}, \hat{M}, S)$ is an enforceable solution to (5) and A is a PPE for game parameters μ and $\sigma(y)R_y = D_y\sigma(y)$. □

Proof of Lemma 8.2. For any constant $\eta > 0$ and any strategy profile A on a stochastic framework $(\Omega, \mathcal{F}, \mathbb{F}, P, Z, (J^{y,y'})_{(y,y')\in\mathcal{Z}})$, define the time-changed processes

$$\tilde{A}_t := A_{\eta t}, \quad \tilde{Z}_t := \frac{1}{\sqrt{\eta}}Z_{\eta t}, \quad \tilde{J}_t^{y,y'} := J_{\eta t}^{y,y'} \text{ for } (y, y') \in \mathcal{Z}, \quad \frac{\mathrm{d}\tilde{P}_t}{\mathrm{d}P} := \prod_{(y,y')\in\mathcal{Z}} \eta^{-J_{\eta t}^{y,y'}}\,\mathrm{e}^{(\eta-1)t}.$$

Moreover, define $\tilde{S}$ based on $(\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}}$ as in (21) and observe that $\tilde{S}_t = S_{\eta t}$. Since A is the limit of a sequence of $\mathbb{F}$-predictable processes, $\tilde{A}$ is the limit of a sequence of predictable processes with respect to the time-changed filtration $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq 0}$. For any $T > 0$, $\tilde{P}_T$ defines a probability measure equivalent to P such that on $[0, T]$, $(\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}}$ are Poisson processes with intensity 1 under $\tilde{P}_T$. Arguing as in Proposition I.7.4 of Karatzas and Shreve (2017), there exists a unique probability measure $\tilde{P}$ that agrees with $\tilde{P}_t$ on $\tilde{\mathcal{F}}_t$ for every $t > 0$. Thus, $\tilde{P}$

will serve as new reference probability measure, under which $(\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}}$ have intensity 1. By the scaling property of Brownian motion, $\tilde{Z}$ is an $\tilde{\mathbb{F}}$-Brownian motion under both P and $\tilde{P}$. Define a family $(\tilde{Q}^{\tilde{A}}_t)_{t\geq 0}$ of probability measures induced by $\tilde{A}$ with respect to $\tilde{\mu}$, $\tilde{\sigma}$, $\tilde{\lambda} = \eta\lambda$, $\tilde{Z}$, $(\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}}$ and $\tilde{P}$ analogously as in (22). Because $\tilde{P}$ is equivalent to P on $\mathcal{F}_t$ for any $t > 0$, $(\tilde{Q}^{\tilde{A}}_t)_{t\geq 0}$ can also be represented via a density process with respect to P. Since $(\tilde{J}^{y,y'}_t)_{t\geq 0}$ are Poisson processes with intensity η under P, it follows that

$$\begin{aligned}\frac{\mathrm{d}\tilde{Q}^{\tilde{A}}_t}{\mathrm{d}P} &= \mathcal{E}_t\left(\int_0^{\cdot} \tilde{\mu}(\tilde{S}_s, \tilde{A}_s^{\top})(\tilde{\sigma}^{-1}(\tilde{S}_s))^{\top}\,\mathrm{d}\tilde{Z}_s + \sum_{(\tilde{S}_{s-},y)\in\mathcal{Z}} \int_0^{\cdot} \left(\frac{\tilde{\lambda}_{\tilde{S}_{s-},y}(\tilde{A}_s)}{\eta} - 1\right)\left(\mathrm{d}\tilde{J}^{\tilde{S}_{s-},y}_s - \eta\,\mathrm{d}s\right)\right)\\ &= \mathcal{E}_t\left(\int_0^{\cdot} \tilde{\mu}(S_{\eta s}, A_{\eta s})^{\top}(\tilde{\sigma}^{-1}(S_{\eta s}))^{\top}\frac{\mathrm{d}Z_{\eta s}}{\sqrt{\eta}} + \sum_{(S_{\eta s-},y)} \int_0^{\cdot} (\lambda_{S_{\eta s-},y}(A_{\eta s}) - 1)\left(\mathrm{d}J^{S_{\eta s-},y}_{\eta s} - \eta\,\mathrm{d}s\right)\right).\end{aligned}$$

Since $\tilde{\sigma}^{-1}(y)\tilde{\mu}(y)/\sqrt{\eta} = \sigma^{-1}(y)\mu(y)$ for each y by assumption, substituting $\mathrm{d}\tilde{s} = \eta\,\mathrm{d}s$ yields

$$\frac{\mathrm{d}\tilde{Q}^{\tilde{A}}_t}{\mathrm{d}P} = \mathcal{E}_{\eta t}\left(\int_0^{\cdot} \mu(S_{\tilde{s}}, A_{\tilde{s}})^{\top}(\sigma^{-1}(S_{\tilde{s}}))^{\top}\,\mathrm{d}Z_{\tilde{s}} + \sum_{(S_{\tilde{s}-},y)\in\mathcal{Z}} \int_0^{\cdot} (\lambda_{S_{\tilde{s}-},y}(A_{\tilde{s}}) - 1)\,(\mathrm{d}J^{S_{\tilde{s}-},y}_{\tilde{s}} - \mathrm{d}\tilde{s})\right) = \frac{\mathrm{d}Q^A_{\eta t}}{\mathrm{d}P},$$

hence $\tilde{Q}^{\tilde{A}}_t$ coincides with $Q^A_{\eta t}$ on $\tilde{\mathcal{F}}_t = \mathcal{F}_{\eta_t}$. By substituting $\mathrm{d}\tilde{s} = \eta\,\mathrm{d}s$, we obtain for every $t \geq 0$,

$$\begin{aligned}\tilde{W}^i_t(\tilde{S}, \tilde{A}; \eta r, \tilde{\mu}, \tilde{\sigma}, \eta\lambda) &= \int_t^{\infty} \eta r \mathrm{e}^{-\eta r(s-t)}\,\mathbb{E}_{\tilde{Q}^{\tilde{A}}_s}\left[g^i\left(\tilde{S}_s, \tilde{A}_s\right)\middle|\,\tilde{\mathcal{F}}_t\right]\mathrm{d}s \qquad (25)\\ &= \int_{\eta t}^{\infty} r\mathrm{e}^{-r(\tilde{s}-\eta t)}\,\mathbb{E}_{Q^A_{\tilde{s}}}\left[g^i(S_{\tilde{s}}, A_{\tilde{s}})\,\middle|\,\mathcal{F}_{\eta t}\right]\mathrm{d}\tilde{s} = W^i_{\eta t}(S, A; r, \mu, \sigma, \lambda),\end{aligned}$$

where we have used that $\tilde{Q}^{\tilde{A}}_s$ and $Q^A_{\tilde{s}}$ coincide on $\tilde{\mathcal{F}}_s$. Because there is a one-to-one correspondence of deviations of $(A_{\eta t})_{t\geq 0}$ to deviations of A, it follows from (25) that $(A_{\eta t})_{t\geq 0}$ is a PPE with respect to $(\eta r, \tilde{\mu}, \tilde{\sigma}, \eta\lambda)$ if and only if A is a PPE with respect to $(r, \mu, \sigma, \lambda)$. □

Proof of Corollary 8.3. Let $w_0 \in \mathcal{E}(r, \mu, \sigma, \lambda; y)$ be attained by a PPE A. By Lemma 8.2, there exists a time-changed PPE with respect to parameters $(\eta r, \mu, \sigma/\sqrt{\eta}, \eta\lambda)$ that attains w_0. Finally, the result follows from an application of Lemma 8.1 to $D_y = \sqrt{\eta}\,\mathrm{I}$ for each state y. □

A.2 Markov-perfect equilibria

Proof of Proposition 6.9. Fix a stationary strategy profile $A = \alpha_*(S)$. For any state y, let S^y denote the state process defined in (21) with initial state y and set $w_y := W_0(y, \alpha_*(S^y))$. Fix an arbitrary stopping time τ and define the times-shifted processes $\tilde{J}^{y,y'} := J^{y,y'}_{\cdot+\tau} - J^{y,y'}_{\tau}$ for every

$(y, y') \in \mathcal{Z}$. Let $\tilde{S}$ be defined as in (21) from processes $(\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}}$ with initial state S_τ. Since $J^{y,y'}$ is a Lévy process for every $(y, y') \in \mathcal{Z}$, the process $\tilde{J}^{y,y'}$ is identically distributed as $J^{y,y'}$. In particular, on $\{S_\tau = y\}$, the process $\tilde{S}$ is identically distributed as S^y. This implies that

$$\tilde{A} := A_{\cdot+\tau} = \alpha_*(S_{\cdot+\tau}) = \alpha_*(\tilde{S}) = \sum_{y\in\mathcal{Y}} \alpha_*(\tilde{S})1_{\{S_\tau=y\}} \overset{\mathrm{d}}{=} \sum_{y\in\mathcal{Y}} \alpha_*(S^y)1_{\{S_\tau=y\}}.$$

Consequently, it follows that

$$W_\tau(S_\tau, A) = W_0(S_\tau, \alpha_*(\tilde{S})) \overset{\mathrm{d}}{=} \sum_{y\in\mathcal{Y}} W_0(y, \alpha_*(S^y))1_{\{S_\tau=y\}} = \sum_{y\in\mathcal{Y}} w_y 1_{\{S_\tau=y\}}.$$

Since τ was arbitrary, this shows that $W(S_0, A) = w_*(S)$, where $w_*(y) = w_y$ for each state y. In particular, $W(S_0, A)$ is locally constant where S is constant. Together with Lemma 4.1, this implies that $\beta = 0$, $M = 0$, and $\delta(y') = \delta_*(S_-, y')$ for every y' such that:

(i) $\Delta W = w_*(y') - w_*(y)$ when the state changes from y to y', hence $\delta_*(y, y')$ satisfies (16).

(ii) The drift rate of (5) is 0 in every state y, implying the decomposition

$$w_*(y) = g(y, \alpha_*(y)) + \sum_{y'\in\mathcal{Y}} \delta_*(y, y')\lambda_{y,y'}(\alpha_*(y)). \tag{26}$$

Substituting the expression for $\delta_*(y, y')$ into (26) and solving for rG_* yields

$$rG_* = w_*(\mathrm{diag}(r\mathbf{1} + \mathbf{1}\Lambda_*) - \Lambda_*). \tag{27}$$

Abbreviate $D = \mathrm{diag}(r\mathbf{1} + \mathbf{1}\Lambda_*)$ and observe that $D^{-1}\Lambda_*$ is substochastic since the multiplication with D^{-1} divides each row of Λ_* by its sum plus r. It follows that $D^{-1}(D - \Lambda_*) = \mathrm{I} - D^{-1}\Lambda_*$ is diagonally dominant, hence invertible. Thus, $D - \Lambda$ is invertible and (16) follows from (27). The second statement now follows from Lemma 4.3. □

B Iterative construction of PPE

B.1 Feasibility

Proof of Lemma 5.2. Fix a strategy profile A with continuation value W and (β, δ) given by (5) such that $W + r\delta(y') \in \mathcal{W}_{y'}$. Suppose towards a contradiction that $W_0 \notin \mathcal{V}_y(\mathcal{W})$. Because $\mathcal{V}_y(\mathcal{W})$ is convex, there exists some N such that $N^\top(W_0 - v) > 0$ for all $v \in \mathcal{V}_y(\mathcal{W})$. For any α, let v_α be the payoff pair that maximizes $\mathcal{V}_y(\mathcal{W})$ in direction N among those that are

decomposable by α and some δ_α. Since $v_\alpha + r\delta_\alpha(y')$ maximizes $\mathcal{W}_{y'}$ in direction N, we obtain

$$rN^\top(\delta_{A_t}(y') - \delta_t(y')) = N^\top(v_{A_t} + r\delta_{A_t}(y') - W_t - r\delta_t(y')) + N^\top(W_t - v_{A_t}) > 0.$$

It follows that the drift rate of W points away from $\mathcal{V}_y(\mathcal{W})$ because

$$N^\top(W_t - g(y, A_t) - \delta_t\lambda(y, A_t)) > N^\top(W_t - g(y, A_t) - \delta_{A_t}\lambda(y, A_t)) = N^\top(W_t - v_{A_t}) > 0.$$

Since the sensitivity to the public signal is a martingale, the distance of W from $\mathcal{V}_y(\mathcal{W})$ increases superlinearly in expectation on $[\![0, \tau_1)\!)$. Because τ_1 is unbounded, W escapes $\mathcal{V}_0$ with positive probability, violating feasibility of W_0. □

B.2 Self-generation and mutual stationarity

Proof of Lemma 4.8.

(i) If $\mathcal{W}$ generates itself, it must satisfy $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ by maximality of $\mathcal{B}_y(\mathcal{W})$.

(ii) By definition, if a set $\mathcal{X}$ is generated by $\mathcal{W}$, it is generated by any superset of $\mathcal{W}$.

(iii) Since $\mathcal{E}$ is self-generating by Lemma 4.7, (i) implies that $\mathcal{E}(y) \subseteq \mathcal{B}_y(\mathcal{E})$. Therefore, (ii) implies $\mathcal{B}_y(\mathcal{E})$ is self-generating and, hence, $\mathcal{B}_y(\mathcal{E}) \subseteq \mathcal{E}(y)$ by maximality of $\mathcal{E}$.

(iv) If y is absorbing, then $\tau_1 = \infty$ a.s., hence the definition of a set being generated by $\mathcal{W}$ reduces to the definition of self-generation in Sannikov (2007). □

Proof of Lemma 6.6.

(i) If any $w \in \mathcal{W}_y$ is $\mathcal{W}$-stationary, then $W_y \subseteq \mathcal{S}_y(\mathcal{W})$ by maximality of $\mathcal{S}_y(\mathcal{W})$.

(ii) If $\mathcal{X}$ is stationary with respect to $\mathcal{W}$, it is stationary for any superset of $\mathcal{W}$. □

Proof of Lemma 6.7. We first show that any payoff pair in a family $\mathcal{W}$ of bounded mutually stationary payoff sets can be attained by a state-order dependent PPE. Fix initial state y_0 and $w_0 \in \mathcal{W}_{y_0}$. By definition of mutual stationarity, there exists (α_0, δ_0) decomposing w_0 on $\mathcal{W}$ in state y such that $(0, \delta_0)$ enforces α_0. A solution to (5) for $A \equiv \alpha_0$, $\beta \equiv 0$, $\delta \equiv \delta_0$, and $M \equiv 0$ starting at w_0 thus remains in w_0 until the time τ_1 of the first state transition. The choice of δ_0 implies that $W_{\tau_1} \in \mathcal{W}_{S_{\tau_1}}$. Moreover, $W_{\tau_1} = w_0 + r\delta_0(y')$ is constant on each event $\{S_{\tau_1} = y'\}$. By definition of mutual stationarity, for each y' there exist $(\alpha_1(y'), \delta_1(y'))$ decomposing $w_0 + r\delta_0(y')$ on $\mathcal{W}$ such that $(0, \delta_1(y'))$ enforces $\alpha_1(y')$. Thus, W_{τ_1} can be attained by a locally constant solution to (5) that remains at W_{τ_1} until time τ_2 with $W_{\tau_2} \in$

$\mathcal{W}_{\tau_2}$, etc. Since there are countably many state transitions, a concatenation will yield an enforceable solution to (5) on $[0, \infty)$, which is a PPE by Lemma 4.3. Moreover, it is state-order dependent because α_k is a function of $(y_0, S_{\tau_1}, \ldots, S_{\tau_k})$.

To show that $\mathcal{E}_s$ is mutually stationary, fix y_0, $w_0 \in \mathcal{E}_s(y_0)$, and a state-order dependent PPE A attaining w_0. Fix an arbitrary stopping time τ. It follows along the same lines as in the proof of Proposition 6.9 that $W_\tau(y_0, A) = w_0$ on the event $\{\tau < \tau_1\}$. Since τ was arbitrary, W has to be locally constant on $[\![0, \tau_1)\!)$. Lemma 4.1 thus implies that $\beta \equiv 0$, $M \equiv 0$, and $\delta \equiv \delta_0$ for some δ_0 that satisfies $w_0 = g(y_0, A_0) + \delta_0 \lambda(y_0, A_0)$. State-order dependence implies that also the continuation profile after τ_1 is state-order dependent, hence $w_0 + r\delta_0(y) = W_{\tau_1} \in \mathcal{E}_s(y')$ on each event $\{S_{\tau_1} = y'\}$. Finally, since A is a PPE, Lemma 4.3 implies that $(0, \delta_0)$ enforces A_0. Because y_0 and w_0 were arbitrary, it follows that $\mathcal{E}_s$ is mutually stationary. □

B.3 Convergence of the algorithms

Proof of Proposition 6.8. We show an iterated application of $\operatorname{conv} \mathcal{S}$ to $\widetilde{\mathcal{W}}^0 = (\mathcal{V}_0^*, \ldots, \mathcal{V}_0^*)$ converges to $\mathcal{E}_r$. Convergence of an iterated application of $\mathcal{S}$ to $\mathcal{E}_s$ follows along the same lines without the public randomization. Since $B_y(\mathcal{V}_0^*)$ is convex, Lemma 7.4 implies $\mathcal{S}_y(\mathcal{V}_0^*) \subseteq \operatorname{conv} \mathcal{S}_y(\mathcal{V}_0^*) \subseteq \mathcal{B}_y(\mathcal{V}_0^*)$. If continuation payoffs come from $\mathcal{V}_0^*$, then the strategy of myopic best responses until the first state transition must deliver at least $\underline{v}^i$ before and after the state transition, hence $\mathcal{B}_y(\mathcal{V}_0^*) \subseteq \mathcal{V}_0^*$. Since $\mathcal{S}_y$ is monotone, it follows that $\widetilde{\mathcal{W}}^n$ is monotonically decreasing in the set-inclusion sense with $\widetilde{\mathcal{W}}_y^n \supseteq \mathcal{E}_r(y)$. Thus, $(\widetilde{\mathcal{W}}^n)_{n \geq 0}$ converges to some limit $\widetilde{\mathcal{W}}^\infty$ that satisfies $\widetilde{\mathcal{W}}_y^\infty \supseteq \mathcal{E}_r(y)$ for each n. It remains to show that $\widetilde{\mathcal{W}}_y^\infty \subseteq \operatorname{conv} \mathcal{S}_y(\widetilde{\mathcal{W}}^\infty)$ so that $\widetilde{\mathcal{W}}^\infty$ is mutually stationary by Lemma 6.6, hence equal to $\mathcal{E}_r$ by Lemma 6.7.

Fix an arbitrary state y and $w \in \widetilde{\mathcal{W}}_y^\infty$. Since $w \in \widetilde{\mathcal{W}}_y^n = \operatorname{conv} \mathcal{S}_y(\widetilde{\mathcal{W}}^{n-1})$ for each n, Carathéodory's theorem implies that there exist three (potentially identical) payoff pairs $w_{n,1}$, $w_{n,2}$, and $w_{n,3}$ such that w lies in their convex hull, and each $w_{n,k}$ is decomposed by some $(\alpha_{n,k}, \delta_{n,k})$. Since $\mathcal{A}(y)$ is finite, by passing to a subsequence we may assume that $\mathcal{A}_\diamond := \operatorname{supp}(\alpha_{n,k})$ is constant along the entire sequence. Let $\mathcal{X}(\mathcal{W})$ denote the set of all $(\alpha, \delta) \in \Upsilon_{\mathcal{A}_\diamond}$ that satisfy $g(y, \alpha) + \delta\lambda(y, \alpha) + r\delta(y') \in \mathcal{W}_{y'}$ for each $y' \in \mathcal{Y}$, and observe that $\mathcal{X}(\mathcal{W})$ is compact. The decomposition of each $w_{n,k}$ implies that $(\alpha_{n,k}, \delta_{n,k})$ lies in $\mathcal{X}(\widetilde{\mathcal{W}}^{n-1})$. Because the entire sequence is contained in the compact set $\mathcal{X}(\mathcal{V}_0^*)$, it must converge along a subsequence $(n_\ell)_{\ell \geq 0}$ to some limit $(\alpha_k, \delta_k) \in \mathcal{X}(\mathcal{V}_0^*)$. Because the sets $\widetilde{\mathcal{W}}^m$ are nested, for any m there exists $\ell(m)$ such that the subsequence remains in the compact set $\mathcal{X}(\widetilde{\mathcal{W}}^m)$ for

all $n_{\ell(m)} > m$. Therefore, the limit must lie in $\mathcal{X}(\widetilde{\mathcal{W}}^m)$ for each m, hence also in $\mathcal{X}(\widetilde{\mathcal{W}}^\infty)$. In particular $w_k := g(y, \alpha_k) + \delta_k \lambda(y, \alpha_k)$ lies in $\mathcal{S}_y(\widetilde{\mathcal{W}}^\infty)$. Continuity of g and λ implies that w_k must be the limit of $w_{n,k} = g(y, \alpha_{n,k}) + \delta_{n,k}\lambda(y, \alpha_{n,k})$, Thus, w must lie in the convex hull of w_1, w_2 and w_3 as, otherwise, w would not be in the convex hull of $w_{n,1}$, $w_{n,2}$, and $w_{n,3}$ for n sufficiently large. This shows that $w \in \operatorname{conv} \mathcal{S}_y(\widetilde{\mathcal{W}}^\infty)$, and completes the proof. □

The proof of Proposition 4.9 proceeds along similar lines, except that establishing $\mathcal{W}_y^\infty \subseteq \mathcal{B}_y(\mathcal{W}^\infty)$ is more challenging. However, since that step requires only minor adaptations from the corresponding proof in Bernard (2024), it is deferred to Online Appendix H.3.

C Enforceability

Proof of Lemma 4.5. Suppose first that Condition (i) holds. Let $M_y^i(a^i, \alpha^{-i})$, $\Lambda_y^i(a^i, \alpha^{-i})$, and $G_y^i(a^i, \alpha^{-i})$ denote the linear extensions of $M_y^i(a)$, $\Lambda_y^i(a)$, and $G_y^i(a)$ defined in Definition 4.4 and in Section 7.1, respectively. Assumption 1.(ii.a) implies that $M_y^i(a^i, \alpha^{-i}) = M^i(a^i)$ is independent of α^{-i}.[12] Since $M_y^i(a^i)$ has full rank, it is invertible, hence for any α^{-i} and any δ^i, there exists β^i that satisfies $\beta^i M_y^i(a^i) = -G_y^i(a^i, \alpha^{-i}) - \delta^i \Lambda_y^i(a^i, \alpha^{-i})$. If all enforceability constraints hold with equality, player i is indifferent among all pure actions, hence they are willing to play any mixed action against α^{-i}. To show that Assumption 1.(ii.c) holds, fix arbitrary α^{-i} and δ, and let a_*^i denote a maximizer of $g^i(y, a^i, \alpha^{-i}) + \delta^i \lambda(y, a^i, \alpha^{-i})$. Then δ provides sufficient incentives to player i with $\beta^i = 0$, and incentives for player $-i$ can be provided by β^{-i} that solves the enforceability constraints with equality.

Suppose next that Condition (ii) holds. Because μ and λ are affine, a mixed action α^i affects the distribution of state transitions and the public signal only through its expected value $\bar{\alpha}^i := \sum_{a^i} \alpha^i(a^i) a^i$. Thus, α is enforceable only if for both players i, α^i maximizes $g^i(y, \tilde{\alpha}^i, \alpha^{-i})$ among all actions $\tilde{\alpha}^i$ with identical expected value. Otherwise, there exists a deviation with a higher flow payoff that is indistinguishable from α^i. Phrased differently, α is enforceable only if for both players i, the expected flow payoff $g^i(y, \alpha)$ lies on the concave envelope

$$g^i_{\text{cav}}(y, x, \alpha^{-i}) = \max_{\alpha^i : \bar{\alpha}^i = x} g^i(y, \alpha)$$

[12] There exists a change of basis matrix Q_y with $Q_y \tilde{M}_y^i(a) = M_y^i(a)$ such that a^1 affects only the first m dimensions of $\tilde{M}_y^i(a)$ and a^2 affects only the remaining ones. The argument that $\tilde{M}_y^i(a)$ is independent of a^{-i} is then identical to the argument for state processes with a product structure in the proof of Lemma 5.5.

of $g^i(y, \cdot, \alpha^{-i})$ when plotted against the expected action taken by player i. Fix an action profile α, for which α^i attains this maximum for $i = 1, 2$, and fix any δ. Since μ is affine without being constant, there exist vectors c^i_y and $b^i_y \neq 0$ such that $\mu(y, \tilde{a}^i, \alpha^{-i}) - \mu(y, \alpha) = c^i_y + b^i_y(\tilde{a}^i - \bar{\alpha}^i)$ for any deviation $\tilde{a}^i$. Thus, action profile α is enforced by (β, δ) in state y for any β with $-\beta^i b^i_y$ between the left- and right-derivatives of $g^i_{\text{cav}}(y, \bar{\alpha}^i, \alpha^{-i}) + \delta^i \lambda(y, \bar{\alpha}^i, \alpha^{-i})$ at $\bar{\alpha}^i$. Such β^i exists because $b^i_y \neq 0$ and $g^i_{\text{cav}}(y, \bar{\alpha}^i, \alpha^{-i}) + \delta^i \lambda(y, \bar{\alpha}^i, \alpha^{-i})$ is concave in $\bar{\alpha}^i$ for any δ^i since λ is affine. In particular, α is enforceable if and only if $g^i(y, \alpha) = g^i_{\text{cav}}(y, \alpha)$ for both players i.

To show that Assumption 1.(ii.c) holds, fix an arbitrary enforceable action profile α, which necessitates $g^{-i}(y, \alpha) = g^{-i}_{\text{cav}}(y, \alpha)$. Together with strict concavity of g^{-i}, this implies that α^{-i} is either a pure action or a mixture between two adjacent pure actions. In particular, (a^i, α^{-i}) is enforceable for any a^i. Fix now arbitrary δ and let a^i_* denote a maximizer of $g^i(y, a^i, \alpha^{-i}) + \delta^i \lambda(y, a^i, \alpha^{-i})$. Because (a^i_*, α^{-i}) is enforceable, we can find suitable β^{-i} as above. □

Proof of Lemma 5.5. The statement is trivial if $\mathcal{A}_\diamond$ is a singleton because $\Psi_{\mathcal{A}_\diamond}$ is defined by the polyhedral constraints (10). Suppose next that Assumption 2.(i) is satisfied. Then $\lambda_{y,y'}(\tilde{a}^i, \alpha^{-i}) - \lambda_{y,y'}(a^i, \alpha^{-i})$ does not depend on α^{-i} for any state y': player $-i$ cannot affect the intensity of states $y' \in \mathcal{Y}_i$ and the difference is 0 for states $y' \in \mathcal{Y}_{-i}$ for any α^{-i}. It follows that the constraints (10) are affine in α^{-i} and δ^i. In particular, the set $\Upsilon_{\mathcal{A}_\diamond}$ of all pairs (α, δ) that satisfy (10) is a closed convex polyhedron. Therefore, so is its projection $\Psi_{\mathcal{A}_\diamond}$ onto the coordinates that correspond to δ. Finally, suppose that Assumption 2.(ii) is satisfied. Let $\Psi^i_{\mathcal{A}^i_\diamond}(\alpha^{-i})$ denote the set of all solutions to (10) for a single player i. Since incentives are provided along a single dimension, $\Psi^i_{\mathcal{A}^i_\diamond}(\alpha^{-i})$ is a closed interval for any α^{-i}. We show in Lemma F.7 that $\Psi^i_{\mathcal{A}^i_\diamond}(\alpha^{-i})$ is continuous and monotone in α^{-i}, hence the union $\Psi^i_{\mathcal{A}_\diamond}$ of all such sets over $\text{supp}(\alpha^{-i}) \subseteq \mathcal{A}^{-i}_\diamond$ is a closed interval. Therefore, $\Psi_{\mathcal{A}_\diamond} = \Psi^1_{\mathcal{A}_\diamond} \times \Psi^2_{\mathcal{A}_\diamond}$ is a rectangle. □

Proof of Lemma 7.6. Under these conditions, player i's enforceability constraints do not depend on α^{-i}. Thus, mixing only reduces the set of incentives via $\Psi_{y,\alpha} = \bigcap_{a^i \in \text{supp}(\alpha^i)} \Psi_{y,a^i,\alpha^{-i}}$ and similarly for $\Xi_{y,\alpha}(w, N, \mathcal{W})$. In particular, each optimality equation is maximized in a pure action profile, and $\mathcal{S}_{y,\mathcal{A}_\diamond}(\mathcal{W}) \subseteq \bigcap_{a \in \mathcal{A}_\diamond} \mathcal{S}_{y,a}(\mathcal{W})$ and $\mathcal{K}_{y,\mathcal{A}_\diamond}(\mathcal{W}) \subseteq \bigcap_{a \in \mathcal{A}_\diamond} \mathcal{K}_{y,a}(\mathcal{W})$. □