# Continuous-Time Stochastic Games with Imperfect Public Monitoring*

Benjamin Bernard†

This version: August 1, 2023

*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 about past actions. 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, the correspondence of initial state to equilibrium payoffs is preserved, the characterization is possible for any level of discounting, and the characterization is applicable to games that are not irreducible.*

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

## 1 Introduction

A wide range of economic applications can be cast as strategic interactions between different parties, in which the environment changes in response to the parties' behavior. During the global financial crisis, financial institutions were unable to raise sufficient capital to meet their short-term liabilities because investors had lost confidence in the financial sector. The past performance of the financial sector had led credit lines to dry up and created an environment where financial institutions were not able to bridge short-term liquidity gaps in the usual way. The development of new technologies by competing research institutions exhibits a similar history-dependent environment. The successful discovery of a new technology changes the research environment forever: competing researchers cannot patent similar work anymore and any effort put into such a discovery was exerted in vain. It is impossible to

*I am grateful for comments and insightful discussions with Andrew Clausen, Eduardo Faingold, Johannes Hörner, Michihiro Kandori, Yuliy Sannikov, and participants of the 12th World Congress of the Econometric Society, GAMES 2020.1, AMETS, the NZ Economics eSeminar series, the NTU-AS symposium as well as seminars at the University of Edinburgh, University of Tokyo, University of Wisconsin–Madison, Hong Kong Baptist University, and the Toulouse School of Economics. Financial support from the Ministry of Science and Technology in Taiwan through grant 108-2410-H-002-249 and the Center for Research in Econometric Theory and Applications by the Ministry of Education in Taiwan through grant 109L900203 is gratefully acknowledged.

†Department of Economics, University of Wisconsin–Madison, 1180 Observatory Dr, Madison, WI 53706; bbernard3@wisc.edu.

forecast the exact time of a financial crisis or the discovery of a new technology. The occurrence of such a state change is random and the likelihood depends on the involved parties' actions. A game-theoretic model that accounts for these sudden state changes is a stochastic game. No deterministic-time dynamic game can capture these sudden and potentially drastic changes in the environment.[1]

In the literature on discrete-time stochastic games, two classes of games have been studied frequently. The first class are *absorbing games*, in which all but the initial state are absorbing. In this class of games, perfect public equilibria (PPE) can be characterized with repeated-game techniques in the continuation states and an enforceability condition in the initial state.[2] The second class of games are *irreducible games*, in which the underlying state process is an irreducible Markov chain. In those games, the distribution over states in any stationary strategy profile converges to a steady-state distribution of the induced Markov chain. Thus, in the limit as players get arbitrarily patient, the impact of the initial state becomes negligible and the limit PPE payoff set is independent of the initial state. Hörner, Sugaya, Takahashi, and Vieille (2011) show that it is possible to adapt the linear program of Fudenberg and Levine (1994) from repeated games to characterize the limit PPE payoff set in such stochastic games.[3] Key in this endeavor is the fact that in the limit, the continuation value of any PPE with extremal payoffs must lie below the tangent hyperplane. In the limit as the discount factor tends to one, these are the only constraints that matter, giving rise to a linear program. When the PPE payoff sets differ from state to state, such as in many non-irreducible games or when players are impatient, the continuation values need not lie below the tangent hyperplane. Consequently, the techniques from discrete time then do not readily generalize to such cases. We show in this paper that for continuous-time stochastic game with finitely many states, such a characterization is possible.

We study a class of two-player stochastic games with finitely many states and imperfect public monitoring. The state process is observed by the players and the frequency of state transitions depends on the chosen actions. Additionally, players continuously observe a public signal, whose distribution depends on the chosen actions. In such a model, the effects of state changes on continuation payoffs can be cleanly separated from the effects of the public signal: continuous movements in the continuation value due to information from the public signal have to lie "below the tangent hyperplane," whereas discontinuous movements at a time of a state transition may lie elsewhere. This allows us to use the techniques developed in Sannikov (2007) to relate the continuous movement of the continuation value to the boundary of the set of PPE payoffs $\mathcal{E}_y$ with initial state $y$ and obtain a differential characterization of its boundary. In a stochastic game, the continuation value after a state transition to state $y'$ has to come from the set $\mathcal{E}_{y'}$. This implies that (a) there is room for some trade-off between incentives provided to players through state transitions versus the

[1] Stochastic games have been used also to model dynamic competition with inventories (e.g., Kirman and Sobel (1974)), the extraction of a common resource (e.g., Levhari and Mirman (1980)), economic growth (e.g., Bernheim and Ray (1989)), entry and exit dynamics in oligopolistic competition (e.g., Ericson and Pakes (1995)), strategic pricing (e.g., Bergemann and Välimäki (1996)), social uprisings (e.g., Acemoglu and Robinson (2001)), research and development (e.g., Grossman and Hart (1983)), and strategic experimentation (e.g., Keller, Rady, and Cripps (2005)).

[2] See, for example, Kohlberg (1974), Mertens and Neyman (1981), and Mertens, Neyman, and Rosenberg (2009) for results in absorbing zero-sum games.

[3] See Fudenberg and Yamamoto (2011) for a direct proof of the folk theorem in irreducible stochastic games with imperfect public monitoring.

public signal and that (b) the differential characterizations between different states are coupled.

We show that this set of coupled differential characterizations can be decoupled through an iterative procedure over state transitions. They key is to view the stochastic game as a concatenation of a set of auxiliary repeated games. For a fixed family of payoff sets $\mathcal{W} = (\mathcal{W}_y)_y$, consider an auxiliary game, in which stage game $y$ is continuously repeated until the time of the first state transition such that terminal payoffs after a transition to state $y'$ come from a the set $\mathcal{W}_{y'}$. Because the state is fixed throughout the auxiliary game, its termination of the game is simply a discontinuous signal about past actions. The auxiliary games thus fall into a class of repeated games studied in Bernard (2023). With some modifications to the techniques in Bernard (2023), it is thus possible to obtain a differential characterization of the boundary of the PPE payoff set $\mathcal{B}_y(\mathcal{W})$ of the auxiliary game in state $y$. Crucially, the characterization of the sets $\mathcal{B}_y(\mathcal{W})$ and $\mathcal{B}_{y'}(\mathcal{W})$ in two different states $y$ and $y'$ are not coupled because the terminal payoffs at the end of the auxiliary games come from a fixed family of payoff sets. We can recover equilibrium payoffs of the stochastic game if the family $\mathcal{B}(\mathcal{W}) = \big(\mathcal{B}_y(\mathcal{W})\big)_y$ is *self-generating*, that is, if $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for each state $y$. Then, each payoff pair $w \in \mathcal{B}_y(\mathcal{W})$ can be attained with a strategy profile that is incentive-compatible until the transition to the next state $y'$, at which point the continuation value $w'$ comes from $\mathcal{W}_{y'}$. Since $\mathcal{W}_{y'} \subseteq \mathcal{B}_{y'}(\mathcal{W})$, the value $w'$ can be attained by an equilibrium in the auxiliary game of state $y'$ until the next state transition, and so on. A countable concatenation of these equilibria of the auxiliary games then yields a PPE of the stochastic game.

Similarly to the algorithm in Abreu, Pearce, and Stacchetti (1990), an iterated application of the operator $\mathcal{B}$ to the set of feasible payoffs $\mathcal{V}$ converges to the family of PPE payoff sets.[4] At first glance, it may seem that the construction of PPE has been reduced to a discrete-time stochastic game, whose period lengths are determined by the state transitions. The concatenation procedure, however, only deals with incentives that are provided at the time of state transitions. In a continuous-time setting, players are not inert between state transitions and the operator $\mathcal{B}$ also has to ensure that players have no incentive to deviate at any other time. Payoffs of perfect public equilibria, in which players condition only on state transitions are characterized as a fixed point of a linear program similarly to discrete time. The incentive constraints between state transitions give rise to two ordinary differential equations (ODEs) that describe the boundary of each set $\mathcal{B}_y(\mathcal{W})$: the optimal incentives related to the absence of state transitions yields the first-order *abrupt-information optimality equation*, first described in Bernard (2023), whereas the optimal tangential value transfers based on the public signal gives rise to a second-order *optimality equation*, first described in Sannikov (2007).

The algorithm may converge in finitely many applications of $\mathcal{B}$. Because any payoff pair in $\mathcal{B}_y(\mathcal{W})$ can be attained by a strategy profile that is enforceable until the next state transition, the set of PPE payoffs $\mathcal{E}_y$ in an absorbing state $y$ is found by a single application of $\mathcal{B}_y$. Similarly, if it is impossible for the state to return to state $y$ once it leaves, then $\mathcal{E}_y$ can be computed with a single application of $\mathcal{B}_y$—as long as the set of PPE payoffs in successor states has been computed first. If there are no possible cycles among states, one can compute the family of PPE payoffs with finitely many applications of $\mathcal{B}_y$,

[4] An adaptation of the repeated-games algorithm to zeros-um stochastic games is given in Rosenberg and Sorin (2001).

one for each state $y$. Because differential equations can be solved numerically rather easily, an ODE description of equilibrium payoffs also contributes to the literature on computation of equilibrium payoffs in stochastic games. While current algorithms restrict attention to 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 imperfect public monitoring.

It is worth noting that the techniques in this paper work for any discount rate $r$, i.e., they allow us to characterize the family of PPE payoff sets for impatient players. Thus, even for irreducible games, for which we have developed a decent understanding from discrete time, the techniques provide additional insights: since convergence to a steady state distribution is not required, the correspondence $y \mapsto \mathcal{E}_y(r)$ of initial states to equilibrium payoffs is preserved.

In many applications, particular attention is given to stationary Markov or Markov-perfect equilibria because of their simplicity.[5] Because stationary Markov equilibria are time independent, their characterization is very similar to the discrete-time counterpart; see Haller and Lagunoff (2000).[6] In addition to characterizing PPE and Markov-perfect equilibrium payoffs, we characterize payoffs of what we call *state-order dependent* PPE. A state-order dependent PPE depends on the public history only through the order of states that have been attained so far. State-order dependent PPE thus strike a balance between simplicity of implementation and attaining a multitude of equilibrium payoffs. We illustrate the different equilibrium notions in an example of a regime-change game.

The techniques used to solve stochastic games in discrete time often depend crucially on the form of the underlying state process. The analysis of irreducible games often relies on convergence to a steady state distribution (e.g., Hörner et al. (2011) and Pęski and Toikka (2017)), whereas the analysis of absorbing games often hinges on the fact that there is only one non-absorbing state. The techniques in this paper only rely on the fact that there are finitely many states. Beyond this restriction, the results apply to stochastic games with an arbitrary state process. The continuous-time techniques do, however, require that the public signal satisfies a pairwise full rank condition to ensure that the aforementioned ODEs are continuous in initial conditions. This condition is stronger than the pairwise identifiability condition in Sannikov (2007) because the set of admissible incentives related to state transitions may, in general, be discontinuous in the players' continuation value. The techniques in this paper do not readily extend to infinite state spaces or games with perfect monitoring. If the state process follows a diffusion process, techniques to characterize perfect public equilibria have been developed in concurrent work by Faingold and Sannikov (2020). Continuous-time stochastic games with perfect monitoring have been studied in Neyman (2017).

The remainder of this paper is organized as follows. The model is presented in Section 2. Section 3

[5] Existence of stationary Markov equilibria in behavior strategies in our setting with finitely many actions and states has been established in Fink (1964). Haller and Lagunoff (2000) show genericity and Doraszelski and Escobar (2010) show stability and purifiability. In more general settings, stationary Markov equilibria may not exist; see the papers by Levy (2013) and Levy and McLennan (2015) for counterexamples. Finally, see He and Sun (2017) for the most general existence result to date.

[6] This paper restricts attention to pure-strategy equilibria. This is not a severe restriction for PPE because players can mix artificially by playing different pure actions for infinitesimally short periods of time. However, this restriction is more severe for stationary Markov equilibria because those are time independent.

contains a motivating example of a contestable democracy. In Section 4, we introduce the notions of enforceability and self-generation in our setting. The set of Markov-perfet equilibrium payoffs and state-order dependent equilibrium payoffs are characterized in Section 5. Section 6 describes the characterization of the set of all PPE payoffs. Section 7 discusses notions related to the computation of equilibrium payoffs and Section 8 concludes. The proofs are contained in Appendices A–D.

## 2 Model

We consider games between 2 players $i = 1, 2$ with a finite set of states $\mathcal{Y}$. State $y \in \mathcal{Y}$ determines the set $\mathcal{A}^i(y)$ of each player $i$'s available actions. We assume that $\mathcal{A}^i(y)$ is finite for any state $y \in \mathcal{Y}$ and we denote by $\mathcal{A}(y) := \mathcal{A}^1(y) \times \mathcal{A}^2(y)$ the set of available action profiles in state $y$. Denote by $\mathcal{A} = \bigcup_{y\in\mathcal{Y}} \mathcal{A}(y)$ the set of all action profiles. The chosen action profile $a \in \mathcal{A}(y)$ affects the rate $\lambda_{y,y'}(a)$, at which the state process $S = (S_t)_{t\geq 0}$ transitions from state $y$ to state $y'$. We denote by $\mathcal{Z}_y(a)$ the set of states $y'$ that can be reached from state $y$ if $a$ is played, i.e., those with $\lambda_{y,y'}(a) > 0$. Denote by $\mathcal{Z}_y = \bigcup_{a\in\mathcal{A}(y)} \mathcal{Z}_y(a)$ the set of all possible successor states of $y$, and by $\mathcal{Z}$ the set of ordered pairs $(y, y')$ with $y' \in \mathcal{Z}_y$. Players choose actions continuously from $\mathcal{A}(S_t)$ at any time $t \in [0, \infty)$.

In a game of imperfect monitoring, players cannot observe the actions chosen by their opponents, but only the impact of the chosen actions on the distribution of a publicly observable signal $X$ as well as the state process $S$. Specifically, we assume that in state $y$, the chosen action profile $a$ affects the drift rate $\mu(y, a)$ of a Brownian signal given by $dX_t = \mu(y, a)\, \mathrm{d}t + \sigma(y)\, \mathrm{d}Z_t$, where $Z$ is a $d$-dimensional Brownian motion and the volatility matrix $\sigma(y) \in \mathbb{R}^{d(y)\times d}$ in state $y$ is of full rank $d(y) \leq d$. With slight abuse of notation, we denote its right inverse by $\sigma^{-1}(y) = \sigma(y)^\top \big(\sigma(y)\sigma(y)^\top\big)^{-1}$. We do allow for the possibility that $d(y) = 0$ in some states $y$, in which case the only public signal about past play is the state process.[7] The public information $\mathcal{F}_t$ at time $t$ is a $\sigma$-algebra that contains the history of the processes $S$ and $Z$ up to time $t$. It may contain additional information that players can use 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 that takes values in $\mathcal{A}^i(S_-)$ or the limit of a sequence of such processes.[8]

*Remark* 2.1. Stochastic games are tractable despite their generality because they are time homogeneous: the set of available actions, the payoffs, and the conditional distribution over future states and public signal realizations do not depend on the time. In a continuous-time setting, this imposes that, conditional on the current state and the chosen action profile, the increments of the signal and the state process are independent and identically distributed, that is, they are Lévy increments. It

[7] Discrete arrival of information about past play without impact on the state can be accommodated by incorporating duplicate states $y_1, y_2$ with identical sets of available actions, payoffs, drift rate and volatility of the public signal, and conditional transition probabilities to states in $\mathcal{Y} \setminus \{y_1, y_2\}$. The "transition" between duplicate states is then just a signal about past play. This framework thus allows a mixed signal structure containing Brownian and Poisson information as in Sannikov and Skrzypacz (2010) and Bernard (2023).

[8] Because the times of state transitions are totally inaccessible, they cannot be approximated by a sequence of predictable times. Thus, no strategy profile can anticipate state changes. Allowing the limit closes the strategy space by allowing non-predictable public randomization.

follows from Corollary 10.23 and Theorem 11.4 in Kallenberg (2002) that the distribution of the state process and a continuously arriving public signal *have to be* of the chosen form.

Because at any time $t$, the chosen strategy profile affects the future distribution of the state process and the public signal, play of a strategy profile $A = (A^1, A^2)$ induces a family of probability measures $Q^A = (Q_t^A)_{t \geq 0}$, under which players observe the game. On any interval $[0, T]$ for $T > 0$, the public signal takes the form

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

under $Q_T^A$, where $Z^A = Z - \int \sigma^{-1}(S_s)\mu(S_s, A_s)\, \mathrm{d}s$ is a $Q_T^A$-Brownian motion describing noise in the public signal. Moreover, under $Q_T^A$, state transitions to state $y$ at time $t$ occur with instantaneous intensity $\lambda_{S_{t-},y}(A_t)$. For a mathematical foundation of the model and details on the change of probability measures, see Appendix A.

Simon and Stinchcombe (1989) and Neyman (2017) demonstrate that in continuous-time games of perfect monitoring, strategies may not lead to unique outcomes if they depend on the immediate past of the opponent's chosen actions. The restriction to public strategies resolves this issue because actions taken by opponents do not immediately generate unambiguous information. Formally, in public monitoring games, one can identify the probability space $\Omega$ with the path space of all publicly observable processes, hence a realized path $\omega \in \Omega$ leads to the unique outcome $(A(\omega), S(\omega))$. In particular, each strategy in a public monitoring game is admissible in the sense of Neyman (2017).

**Definition 2.2.**

(i) In each state $y$, player $i$ receives an unobservable expected flow payoff $g^i(y, \cdot\,) : \mathcal{A}(y) \to \mathbb{R}$.[9]

(ii) Player $i$'s *discounted expected future payoff* (or *continuation value*) under strategy profile $A$ at any time $t \geq 0$ is given by

$$W_t^i(S_t, A) = \int_t^\infty r\mathrm{e}^{-r(s-t)} \mathbb{E}_{Q_s^A}\Big[g^i(S_s, A_s)\,\Big|\,\mathcal{F}_t\Big]\, \mathrm{d}s, \tag{2}$$

where $r > 0$ is the discount rate of both players and the distribution of $(S_s)_{s \geq t}$ is determined uniquely by $S_t$ and $(A_s)_{s \geq t}$.

(iii) A strategy profile $A$ is a *perfect public equilibrium (PPE)* for discount rate $r$ if for every player $i$ and all possible deviations $\tilde{A}^i$,

$$W^i(S, A) \geq W^i\big(S, (\tilde{A}^i, A^{-i})\big) \text{ a.e.,}^{10} \tag{3}$$

where $A^{-i}$ denotes the strategy of player $i$'s opponent in profile $A$.

(iv) We denote the set of all payoff pairs that are achievable by a perfect public equilibrium when

[9] The expected flow payoff corresponds to the ex-ante stage game payoff in discrete-time games.

[10] Deviations of a PPE are not profitable *almost everywhere* (a.e.), that is, the inequality $W_t^i(y, A; \omega) \geq W_t^i\big(y, \tilde{A}^i, A^{-i}; \omega\big)$ holds for every pair $(\omega, t)$ except on a set of $P \otimes$ *Lebesgue*-measure 0.

the initial state is $y$ and the discount rate is $r$ by

$$\mathcal{E}_y(r) := \Big\{ w \in \mathbb{R}^2 \;\Big|\; \text{there exists a PPE } A \text{ with } \mathbb{E}[W_0(y, A)] = w \text{ a.s.} \Big\}.$$

We denote by $\mathcal{E}(r)$ the family $\left(\mathcal{E}_y(r)\right)_{y \in \mathcal{Y}}$ of equilibrium payoff sets.

The form of the players' continuation value in (2) shows that the players' strategies affect their expected payoffs directly through their expected flow payoff and indirectly, through the impact on the distribution of the public signal and the state process, which is reflected in the change of probability measure. Because the weights in (2) integrate to one, the continuation value of a strategy profile is a convex combination of "stage game" payoffs, hence the set of feasible payoffs is contained within the set $\mathcal{V} := \operatorname{conv}\{g(y, a) \mid a \in \mathcal{A}(y), y \in \mathcal{Y}\}$. By deviating to the strategy of myopic best replies to their opponent's strategy profile, each player $i$ can ensure an equilibrium payoff of at least

$$\underline{v}^i = \min_{y \in \mathcal{Y}} \min_{a^{-i} \in \mathcal{A}^{-i}(y)} \max_{a^i \in \mathcal{A}^i(y)} g^i(y, (a^i, a^{-i})).$$ [11]

It follows that $\mathcal{V}^* := \{w \in \mathcal{V} \mid w^i \geq \underline{v}^i \text{ for every player } i\}$ is an upper bound for the set of equilibrium payoffs $\mathcal{E}_y(r)$ in each state $y \in \mathcal{Y}$. Moreover, each $\mathcal{E}_y(r)$ is convex because players are allowed to use public randomization.

Because the set of public perfect equilibria is typically very large, it is worth studying simple classes of PPE that might be easier to implement and, thus, perhaps more reasonable to expect in reality. The most frequently studied of such equilibria are stationary Markov or Markov-perfect equilibria, which depend on the entire public history only through the current state. Since there are only finitely many Markov-perfect equilibria, those propose a sharp contrast to the multitude of PPE. The techniques in this model also allow us to easily characterize the payoffs of a family of equilibria between Markov perfect equilibria and general PPE. We call those state-order dependent PPE.

**Definition 2.3.**

(i) A PPE $A$ is a *Markov perfect equilibrium* or *stationary Markov equilibrium* for initial state $S_0$ if there exists a map $a^* : \mathcal{Y} \to \mathcal{A}$ with $a_*(y) \in \mathcal{A}(y)$ for every state $y$ such that $A = a^*(S_-)$.

(ii) A PPE is *state-order dependent* for initial state $S_0$ if there exists a map $a_* : \bigcup_{k=1}^{\infty} \mathcal{Y}^k \to \mathcal{A}$ with $a_*(y_1, \ldots, y_k) \in \mathcal{A}(y_k)$ for any sequence of states of any length $k$ such that $A = a_*(\hat{S}_-)$, where $\hat{S}_t$ is the sequence of states visited up to and including time $t$.

(iii) We denote by $\mathcal{E}^M(r)$ and $\mathcal{E}^S(r)$ the families of payoff pairs that are achievable in Markov-perfect equilibria and state-order dependent equilibria, respectively. Both of those are without the use of a public randomization device. We denote by $\mathcal{E}^S_R(r)$ the family of state-order dependent equilibrium payoffs with public randomization. Note that

$$\mathcal{E}^M(r) \subseteq \mathcal{E}^S(r) \subseteq \mathcal{E}^S_R(r) \subseteq \mathcal{E}(r) \subseteq \mathcal{V}^*.$$

[11] Taking the minimum over states does not yield a tight bound, but it will suffice for our purposes.

| | A | I |
|---|---|---|
| P | $5, 3$ | $1, -2$ |
| M | $2, -1$ | $0, 0$ |

$g(y_1, a)$

| | P | M |
|---|---|---|
| A | $3, 5$ | $-1, 2$ |
| I | $-2, 1$ | $0, 0$ |

$g(y_2, a)$

| | A | I |
|---|---|---|
| P | 0.5 | 4.5 |
| M | 0.25 | 2.25 |

$\lambda_{y,y'}(a)$

| | P | M |
|---|---|---|
| A | $7, (-1)^i$ | $1, (-1)^i 3$ |
| I | $2, 0$ | $0, 0$ |

$\mu(y_i, a)$

**Figure 1:** Impact of actions on expected flow payoffs, state transitions, and the drift rate of the public signal.

A state-order dependent PPE depends on the public history only through the order of states visited, but neither on the time of state transitions, nor on how much time has been spent in any of the states. Such strategy profiles depend on significantly less information than the the full state process. A player following a state-order dependent strategy does not react, for example, to a deviation by the opponent that makes all state transitions ten times more frequent because such a deviation leaves the distribution over the order of states unchanged. Note that if the state process is cyclic, the sequence of state transitions is determined uniquely from the initial state. Thus, in a cyclic game, the sequence of action profiles taken—but not their duration—in state-order dependent PPE is known at the beginning of the game. This is the case in the following example of a regime-change game: if one ideological group is removed from office, it is because the other group has taken over.

## 3 Example of a regime-change game

Suppose that proponents of two ideologies compete for power. State $y_i$ for $i = 1, 2$ indicates that group with ideology $i$ is in charge. At every instant, the incumbent can either work on policy reforms or manipulate the media, that is, $\mathcal{A}^i(y_i) = \{P, M\}$. The non-incumbent can either accommodate the incumbent or attempt to instigate a revolution, i.e., $\mathcal{A}^{-i}(y_i) = \{A, I\}$. The intensity of state changes is indicated in the left panel of Figure 1. State changes are much more frequent when the non-incumbent is actively trying to instigate a revolution. The incumbent can somewhat counteract this effect by manipulating the media. We suppose that the cooperative actions $A$ and $P$ have a flow cost of 2 $\mathrm{d}t$ and we normalize the flow cost of the non-cooperative actions $I$ and $M$ to zero.

Suppose that welfare $X$ has two dimensions: $X^1$ measures components of welfare that are enjoyed equally by the two groups, whereas $X^2$ measures welfare of issues, on which the two groups have diametrically opposite views with group 1 enjoying negative changes in $X^2$. In a sense, $(X^1, X^2)$ is an orthogonal decomposition of welfare. This leads to a realized flow payoff of

$$\mathrm{d}V_t^1 := \mathrm{d}X_t^1 - \mathrm{d}X_t^2 - 2 \cdot 1_{\{A^1 \in \{A,P\}\}}\, \mathrm{d}t, \qquad \mathrm{d}V_t^2 := \mathrm{d}X_t^1 + \mathrm{d}X_t^2 - 2 \cdot 1_{\{A^2 \in \{A,P\}\}}\, \mathrm{d}t.$$

The expected flow payoff of group $i$ is $g^i(y, a) = \mu^1(y, a) + (-1)^i \mu^2(y, a) - 2 \cdot 1_{\{A^i \in \{A,P\}\}}$, leading to the stage-game payoff pairs summarized in Figure 1. In addition to being payoff-relevant, welfare changes carry information about past play. The chosen action profiles affect the drift rate of welfare as indicated in Figure 1. The total increase in welfare is highest if the incumbent works on policy reforms and the non-incumbent accommodates, which leads to compromises and to more moderate

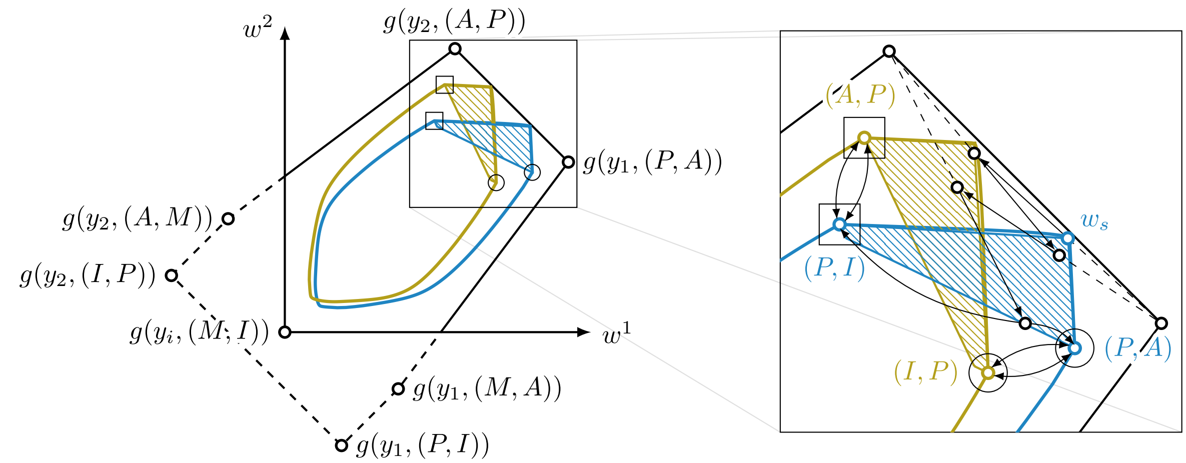



**Figure 2:** Depicted in color is the family $\mathcal{E}(r) = \{\mathcal{E}_{y_1}(r), \mathcal{E}_{y_2}(r)\}$ of PPE payoff sets. The hatched area are payoff pairs that can be attained state-order dependent PPE with public randomization. There are two Markov-perfect equilibria, yielding payoff pairs in $\{\hat{w}_\square(y_1), \hat{w}_\square(y_2)\}$ and $\{\hat{w}_\circ(y_1), \hat{w}_\circ(y_2)\}$ indicated with squares and circles, respectively. The sample path of a state-order dependent PPE attaining $w_s$ is shown.

policies being implemented. If the non-incumbent actively instigates a revolution, the total increase in welfare is lower and more one-sided. The volatility matrix is the identity matrix in both states.

Note that each stage game has a unique Nash equilibrium, in which both parties choose the cooperative action, giving each group its highest-possible payoff of the stage game. Thus, if one group was to rule forever, the efficient PPE would demand that the non-incumbent accommodates at all times to ensure moderate policies are implemented. However, the possibility for state changes prevents permanent collaboration as the opposition is tempted to accelerate a state change by instigating a revolution. Figure 2 illustrates the different sets of equilibrium payoffs in this game for discount rate $r = 0.5$. There are two Markov-perfect equilibria, which involve one group exerting effort at all times and the other group working on policies when they are in charge but instigating a revolution when in opposition. Markov-perfect equilibria are inefficient because stationarity requires that one group instigates a revolution whenever it is not in charge, even if doing so after every second revolution so is not necessary to deter deviations. A higher sum of payoffs can be attained with a state-order dependent PPE. The hatched areas in Figure 2 is the family $\mathcal{E}_R^S(r)$. By allowing players to make their actions dependent on the sequence of states, it is possible to delay instigation of a revolution. Figure 2 shows a state-order dependent PPE that attains payoff pair $w_s$, which supports collaboration in the first four terms. In the fifth term, the opposition randomizes among its actions to select a stationary continuation equilibrium. In this example we have $\mathcal{E}^M(r) = \mathcal{E}^S(r)$, that is, state-order dependence does not allow one to attain a larger set of payoffs than stationarity in the state unless one also allows public randomization.

Outside the set of state-order dependent PPE payoffs, the boundary is given by the solution to two ODEs, called the optimality equations. Along the curved segments, the PPE is generically unique: if the continuation value is equal to some $w$ on the boundary, players play the (generically unique) action profile $a_w$ given by the optimality equations. Figure 3 shows the optimal action profiles at the boundary of $\mathcal{E}_{y_1}(r)$. A PPE attaining extremal payoffs remains on the boundary of $\mathcal{E}_S(r)$. The PPE

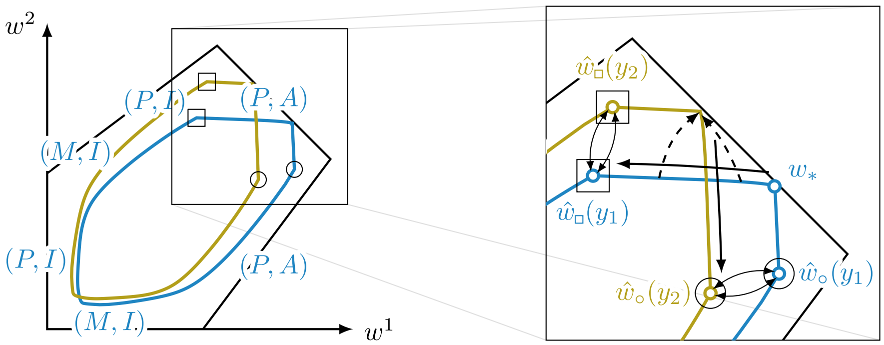



**Figure 3:** In the PPE that attains $w_* \in \mathcal{E}_{y_1}(r)$, both groups cooperate initially. The solid arrows indicate in which direction the continuation value drifts in absence of any revolutions. The dashed arrows show the optimal rewards/punishments provided upon a state transition.

payoff pair with the highest sum of payoffs $w_*$ lies on one such solution; see the zoom-in of Figure 3.[12] Initially, both groups cooperate. Different from the state-order dependent PPE, the continuation value is not locally constant and instead drifts slowly along the boundary towards the Markov-perfect equilibrium 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. While the non-incumbent is willing to cooperate initially, doing so forever is not incentive-compatible, hence group 2 changes action at time $t_0$. If a state transition occurs before $t_0$ is reached, the continuation value jumps to the symmetric boundary of $\mathcal{E}_{y_2}(r)$, where both groups cooperate until $\hat{w}_{\circ}(y_2)$ is reached. In this PPE, the continuous public signal is completely ignored and incentives are provided entirely through state transitions.

At other extremal equilibrium payoffs, the continuous signal is used to provide incentives. The opposition mostly attempts to instigate a revolution. Doing so is temporarily costly, but it improves the likelihood of a state change that will put the opposition in power. On the upper part of $\partial\mathcal{E}_{y_1}(r)$, the incumbent is mostly willing to exert effort since an equilibrium reward for the opposition necessarily also rewards the incumbent. This is not the case on the lower part of the boundary $\partial\mathcal{E}_{y_1}(r)$, where the incumbent exerts effort only when the non-incumbent accommodates.

## 4 Construction of perfect public equilibria

As in any game with imperfect public monitoring, players' incentives are tied to outcomes of the public signal. Thus, we first state the dependence of the continuation value on the public signal and the state process. It will be convenient to denote by

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

the process that counts the number of state transitions to state $y$. Note that the increment $\mathrm{d}J_t^y$ is the increment of a counting process with instantaneous intensity $\lambda_{S_{t-},y}(A_t)$, which means $\mathrm{d}J^y \equiv 0$

[12] The payoff pair $w_*$ lies close to the payoff pair $w_s$ that can be attained by a state-order dependent PPE. While they are hard to distinguish even in the zoom-in, $w_*$ is strictly preferred to $w_s$ by both groups.

if $y$ cannot be reached directly from state $S_{t-}$. We obtain the following stochastic differential representation of the players' 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*

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

*for a martingale $M^i$ orthogonal to $Z$ and $(J^y)_{y\in\mathcal{Y}}$ with $M_0^i = 0$, predictable and locally square-integrable processes $\beta^i$ and $\delta^i(y)$ for $y \in \mathcal{Y}$, and state process $S$ induced by $A$ with $S_0 = y_0$.*

The first term in (5) is a drift term that describes the expected movement of the continuation value. It points away from the expected flow payoff: if $W_t^i < g^i(S_t, A_t)$, then player $i$ extracts an instantaneous payoff rate that exceeds his continuation value and, therefore, doing so has to decrease his future payoff in expectation. The second term is a diffusion term that describes the exposure of the continuation value to the public signal. In a continuous-time setting, rewards or punishments are provided proportionally to how far the observed change $\sigma(S_t)\mathrm{d}Z_t$ in the signal is above or below the expected change $\mu(S_t, A_t)\,\mathrm{d}t$. The term $r\beta_t^i$ is the sensitivity of the exposure. The value $\beta_t^{ik}$ is positive if a high realization of $Z^k$ is good news about past play of player $i$. The third term captures the impact of state changes (or the lack thereof) on the continuation value of player $i$. The term $r\delta_t^i(y)$ is player $i$'s reward if a state transition to state $y$ occurs. If a state transition to state $y$ is good news for player $i$, then the absence of such a state transition 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 of the reward/punishment upon the occurrence of a state transition. The final term is a martingale, which captures the use public randomization. Changes in the continuation value due to public randomization average out in expectation and $M = 0$ if players do not use it.

### 4.1 Enforceability

In discrete-time games, incentives are provided by a *continuation promise* that maps the public signal to a promised continuation payoff for every player; see, for example, Abreu et al. (1990). The representation in (5) shows that in continuous-time games, the continuation value is linear in the public signal, hence so is the continuation promise. To state the incentive-compatibility conditions, it will be convenient to denote by $\lambda(y, a) := \left(\lambda_{y,y_1}(a), \ldots, \lambda_{y,y_{|\mathcal{Y}|}}(a)\right)^\top$ the column vector of all transition intensities to successor states when the state is equal to $y$ and action profile $a$ is played.

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

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

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 $S$.

Action profile $a$ is enforceable in state $y$ if the continuation promise guarantees that the sum of expected instantaneous payoff rate $g^i(y,a)$ and promised continuation rate $\beta^i\mu(y,a)+\delta^i\lambda(y,a)$ for every player $i$ is maximized in $a^i$ over all unilateral deviations. Suppose that players keep their promises and the continuation promise used to enforce $A$ are, in fact, the sensitivities of the continuation value to the public signal and the state process. Then, no player has an incentive to deviate at any point in time and the strategy profile is an equilibrium. This is formalized in the following lemma, which is the continuous-time analogue to the one-shot deviation principle.

**Lemma 4.3.** *A strategy profile $A$ is a PPE for initial state $y_0$ if and only if $(\beta,\delta)$ 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 make the following assumptions on stage-game payoffs and the public signal.

**Definition 4.4.** For any state $y$, any action profile $a \in \mathcal{A}(y)$, and any player $i = 1, 2$, let $M^i_y(a)$ denote the $d \times (|\mathcal{A}^i(y)| - 1)$-dimensional matrix, whose column vectors are given by $\mu(y, \tilde{a}^i, a^{-i}) - \mu(y, a)$ for deviations $\tilde{a}^i \in \mathcal{A}^i(y) \setminus \{a^i\}$. An action profile $a$ is said to have *pairwise full rank* if the matrix $[M^1_y(a), M^2_y(a)]$ has rank $|\mathcal{A}^1(y)| + |\mathcal{A}^2(y)| - 2$.

**Assumption 1.** For each $y \in \mathcal{Y}$, every action profile $a \in \mathcal{A}(y)$ has pairwise full rank.

As in Fudenberg, Levine, and Maskin (1994), an action profile $a$ has pairwise full rank if and only if it satisfies two weaker conditions: $a$ has *individual full rank* if $M^i_y(a)$ has rank $|\mathcal{A}^i(y)| - 1$ and $a$ is *pairwise identifiable* if $\operatorname{span} M^1_y(a) \cap \operatorname{span} M^2_y(a) = \{0\}$. Assumption 1 ensures regularity of the boundaries of the family of PPE payoff sets. The individual full rank condition ensures that the optimality equation is locally Lipschitz continuous in incentives from state transitions. As in Sannikov (2007), pairwise identifiability guarantees that the characterizing ODE is locally Lipschitz continuous in incentives from the public signal in non-coordinate directions. For local Lipschitz continuity in coordinate directions, we additionally need that the game has a *product structure*, i.e., that the players' impact on the distribution of the public signal are orthogonal to each other.

**Assumption 2.** $\operatorname{span} M^1_y(a) \perp \operatorname{span} M^2_y(a)$ for each $y \in \mathcal{Y}$ and each $a \in \mathcal{A}(y)$.

### 4.2 Self-generation

*Remark* 4.1. Since Assumption 2 implies pairwise identifiability, Assumption 2 in conjunction with individual full rank implies Assumption 1. The assumptions are not needed and the entire characterization holds if there is no Brownian signal.

Lemmas 4.1 and 4.3 motivate how we construct equilibrium profiles in continuous time—as solutions to the SDE (5) subject to the enforceability constraint (6). To do so, we use the fact that stochastic games are time homogeneous: since the continuation profile of a PPE after any time $t$ is also an equilibrium of the entire game, the continuation value has to remain within the family of equilibrium payoff sets at all times. This property is known as self generation. In our setting, it is formalized as follows.

**Definition 4.5.** A family of payoff sets $(\mathcal{W}_y)_{y\in\mathcal{Y}} \subseteq \mathbb{R}^2$ is *self-generating* if for every state $y \in \mathcal{Y}$ and every payoff pair $w \in \mathcal{W}_y$, there exists a solution $(W, S, A, \beta, \delta, M)$ to (5) such that $(\beta, \delta)$ enforces $A$, $S_0 = y$ a.s., $W_0 = w$ a.s., and $W_\tau \in \mathcal{W}_{S_\tau}$ a.s. for every stopping time $\tau$.[13]

*Remark* 4.2. The stochastic differential equation (5) does not in general admit strong solutions, that is, it may not be possible to solve (5) for a fixed Brownian motion and a fixed state process. In the proofs that are contained in the appendices, we use weak solutions to (5), in which the Brownian motion, the state process, and the probability space are part of the solution. To keep the notation simple, we do not make this distinction in the main text. See Appendix B for details.

Similarly as in discrete-time games, the family of equilibrium payoff sets is the largest family of bounded self-generating payoff sets.[14]

**Lemma 4.6.** *The family* $\big(\mathcal{E}_y(r)\big)_{y\in\mathcal{Y}}$ *is the largest family of bounded self-generating payoff sets.*

The characterization of $\mathcal{E}(r)$ as the largest family of bounded self-generating payoff sets allows us to construct equilibria using a stochastic control approach. Self-generation implies that the continuation value of each PPE corresponds to an enforceable solution to (5) that locally remains in $\mathcal{E}_y(r)$ while the state is equal to $y$, which jumps to $\mathcal{E}_{y'}(r)$ when the state transitions to state $y'$. For any $w \in \partial\mathcal{E}_y(r)$ and any normal vector $N_w$ to $w \in \partial\mathcal{E}_y(r)$, it is thus necessary that:

(I1) The drift points towards the interior of the set: ${N_w}^\top\big(g(y,a) + \delta\lambda(y,a) - w\big) \geq 0$,

(I2) The volatility is tangential to the set: ${N_w}^\top\beta = 0$,

(I3) State transitions are within $\mathcal{E}(r)$: $w + r\delta(y') \in \mathcal{E}_{y'}(r)$ for every $y' \in \mathcal{Z}_y(a)$.

Observe that restriction (I2) is necessary because of unbounded variation of Brownian motion: if incentives were provided in any non-tangential direction, the continuation value would escape $\mathcal{E}(r)$ immediately; see Figure 4. This means that at extremal equilibrium payoffs, a Brownian signal can be used only to transfer values tangentially to the set. In contrast, incentives from state transitions can be used to transfer value, destroy value, or generate value for both players simultaneously.

### 4.3 Iteration over state transitions

In repeated games with a Brownian public signal, informational restriction (I2) is the crucial condition needed to describe the boundary of the equilibrium payoff set through an explicit ODE. It implies that the space of all admissible incentives depends on the geometry of the equilibrium payoff set

[13] The processes $(J^y)_{y\in\mathcal{Y}}$ appearing in (5) are defined from $S$ as in (4).

[14] One can show that the union $(\mathcal{W}_y \cup \mathcal{V}_y)_{y\in\mathcal{Y}}$ of two families $(\mathcal{W}_y)_{y\in\mathcal{Y}}$, $(\mathcal{V}_y)_{y\in\mathcal{Y}}$ of self-generating payoff sets is again self-generating. Thus, there is, in fact, a largest bounded self-generating family of payoff sets equal to the union of all such families of payoff sets.

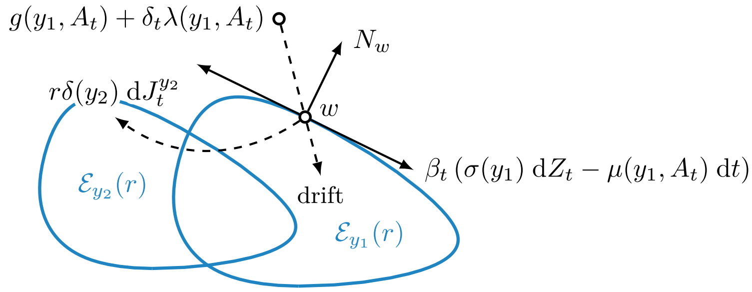



**Figure 4:** While the state is equal to $y_1$, the continuation value of a PPE moves continuously in $\mathcal{E}_{y_1}(r)$. At the boundary, self-generation implies that the continuation value has to move tangentially to the boundary, given through the incentives arising from the public signal. When the state changes to $y_2$, the continuation value changes discontinuously to a value in $\mathcal{E}_{y_2}(r)$.

only through local information, i.e., the direction of the tangent—in the same way that an ODE uses only local information to encode the entire function. In a stochastic game, the space of admissible incentives additionally includes the rewards and punishments after state transitions. Thus, at the boundary of $\mathcal{E}_y(r)$, the set of admissible incentives depends on local information exclusively only if the state process can never return to state $y$. Otherwise, (I3) depends on global information about $\mathcal{E}_y(r)$ through the sequences of states, in which the state process can return to $y$. Then, the ODEs describing the boundaries are no longer explicit. In order to solve them numerically, we approximate the boundary through a sequence of ODEs. Key is the relaxation of restriction (I3) to a condition that requires continuation values after a state change to come from a fixed family of payoff sets.

**Definition 4.7.** Fix a family of payoff sets $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$ such that each $\mathcal{W}_y$ is compact and convex. Let $\rho$ denote the occurrence of the first state change.

(i) A payoff set $\mathcal{X}_y$ is *$\mathcal{W}$-relaxed self-generating* in state $y$ if for each $w \in \mathcal{X}_y$, there exists a solution $(W, S, A, \beta, \delta, M)$ to (5) with $W_0 = w$ a.s., $S_0 = y$ a.s., $W_\rho \in \mathcal{W}_{S_\rho}$ a.s., and for almost every $(\omega, t)$ with $0 \le t < \rho(\omega)$, we have $W_t(\omega) \in \mathcal{X}_y$ and $(\beta_t(\omega), \delta_t(\omega))$ enforces $A_t(\omega)$ in $S_t(\omega)$.

(ii) A family $\mathcal{X} = (\mathcal{X}_y)_{y\in\mathcal{Y}}$ of payoff sets is *$\mathcal{W}$-relaxed self-generating* if each set $\mathcal{X}_y$ is $\mathcal{W}$-relaxed self-generating in state $y$.

(iii) $\mathcal{B}(\mathcal{W}) = \left(\mathcal{B}_y(\mathcal{W})\right)_{y\in\mathcal{Y}}$ is the largest family of $\mathcal{W}$-relaxed self-generating payoff sets.

The operator $\mathcal{B}$ bears many similarities to the standard set operator in Abreu et al. (1990). Payoff pairs in $\mathcal{B}_y(\mathcal{W})$ for any state $y \in \mathcal{Y}$ can be attained by an enforceable strategy profile with continuation promise from $\mathcal{W}_{S_\rho}$ at time $\rho$. The additional requirement due to the continuous-time setting is that before the state transition, the continuation value remain in $\mathcal{B}_y(\mathcal{W})$. In particular, incentives related to the absence of state transitions and the arrival of the public signal have to enforce the strategy profile at all times before the state transition.

It follows immediately from the definition that that $\mathcal{B}_y(\mathcal{E}(r)) = \mathcal{E}_y(r)$ for any state $y$: any payoff pair in $\mathcal{B}_y(\mathcal{E}(r))$ can be attained by a locally enforceable strategy profile with equilibrium continuation values after a state transition. Conversely, any equilibrium payoff in $\mathcal{E}_y(r)$ must be attainable by such a strategy profile as well. We characterize the boundary of $\mathcal{B}_y(\mathcal{W})$ for an arbitrary family of payoff sets $\mathcal{W}$ in Section 6.3. This is sufficient to solve for the family of PPE payoff sets in games, in which no state can be reached twice. For games, in which each state may be visited more than once, the operator $\mathcal{B}$ allows us to develop an approximation result similarly to Abreu et al. (1990). We first observe that $\mathcal{B}$ is closely related to the concept of self-generation.

**Lemma 4.8.** *Let $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$ such that $\mathcal{W}_y \subseteq \mathcal{V}^*$ for every state $y \in \mathcal{Y}$. If $\mathcal{W}$ is self-generating, then $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for every state $y \in \mathcal{Y}$. Conversely, if $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for every state $y \in \mathcal{Y}$, then $\mathcal{B}(\mathcal{W})$ is self-generating.*

Since payoff pairs in $\mathcal{B}_y(\mathcal{W})$ can be attained with a continuation promise in $\mathcal{W}_{S_\rho}$ at time $\rho$, the condition that $\mathcal{W}_{y'} \subseteq \mathcal{B}_{y'}(\mathcal{W})$ for every $y'$ allows us to attain $W_\rho$ with an enforceable strategy profile that remains in $\mathcal{B}_{S_\rho}(\mathcal{W})$ until the next state transition, and so on. Because Poisson processes have only countably many jumps, repeating this concatenation argument countably many times yields solutions that remain in $\mathcal{B}_S(\mathcal{W})$ forever. The proof in Appendix B additionally deals with some measurability issues of the concatenation. We obtain the following algorithm to approximate $\mathcal{E}(r)$.

**Proposition 4.9.** *Let $\mathcal{W}_0 = (\mathcal{W}_{0,y})_{y\in\mathcal{Y}}$ be the family of payoff sets with $\mathcal{W}_{0,y} = \mathcal{V}^*$ 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 $\bigcap_{n\geq 0} \mathcal{W}_{n,y} = \mathcal{E}_y(r)$.*

Contrary to the algorithm in Hörner et al. (2011), this algorithm does not require players to be arbitrarily patient, nor does it impose any conditions on the underlying state process. As a consequence, the algorithm is also applicable to games that violate Assumption A in Hörner et al. (2011), requiring that the limit equilibrium payoff set is independent of the initial state. Their assumption A is typically encountered in the form of irreducibility of the state process. However, there are other sufficient conditions; see Dutta (1995). The trade-off between our algorithm and the algorithm in Hörner et al. (2011) is that, in the general setting, we require rather strong conditions on the distribution of the public signal. Our algorithm is also applicable if there is no Brownian public signal and the only information revealed is through the state process. In that case, no assumptions are required.

## 5 Markov perfect and state-order dependent equilibria

In a state-order dependent equilibrium, players condition their behavior only on state changes. Thus, the strategy profile and the continuation value must be locally constant between consecutive state transitions. Similarly to self-generation, in a state-order dependent PPE the continuation value after a state transition must itself be attainable by a locally constant strategy profile and so on. We introduce the notion of mutual stationarity that allows a recursive construction of stationary PPE similarly to relaxed self-generation for general PPE. If the continuation value is locally constant, the SDE representation (5) implies that $\beta = 0$, i.e., incentives are provided through state changes exclusively. Let $\Psi_{y,a}$ denote the set of all incentives $\delta$, for which $(0, \delta)$ enforces $a$ in state $y$. For a family of payoff sets $\mathcal{W} = (\mathcal{W}_y)_{y\in\mathcal{Y}}$, let $\Psi_{y,a}(w, \mathcal{W})$ denote the set of all $\delta \in \Psi_{y,a}$ with $w + r\delta(y') \in \mathcal{W}_{y'}$ for all successor states $y'$ of $y$, i.e., all $y' \in \mathcal{Z}_y(a)$.

**Definition 5.1.** A family $\mathcal{W} = (\mathcal{W}_y)_{y \in \mathcal{Y}}$ of payoff sets is *mutually stationary* if for every $y \in \mathcal{Y}$ and every $w \in \mathcal{W}_y$, there exist $a \in \mathcal{A}(y)$ and $\delta \in \Psi_{y,a}(w, \mathcal{W})$ with $w = g(y,a) + \delta\lambda(y,a)$.

**Lemma 5.2.** *The family of state-order dependent equilibrium payoffs $\mathcal{E}^S(r)$ is the largest bounded family of mutually stationary payoff sets.*

Mutual stationarity is the analogue to self-generation for state-order dependent PPE. The set $\mathcal{E}^S(r)$ can be approximated with a similar localized construction, where we impose that the continuation value after a state change comes from a fixed family of payoff sets.

**Definition 5.3.** Fix a family $\mathcal{W} = (\mathcal{W}_y)_{y \in \mathcal{Y}}$ of payoff sets.

(i) A payoff set $\mathcal{X}_y$ is $\mathcal{W}$-*stationary* in state $y$ if for each $w \in \mathcal{X}_y$, there exist $a \in \mathcal{A}(y)$ and $\delta \in \Psi_{y,a}(w, \mathcal{W})$ with $w = g(y,a) + \delta\lambda(y,a)$.

(ii) A family $\mathcal{X} = (\mathcal{X}_y)_{y \in \mathcal{Y}}$ of payoff sets is $\mathcal{W}$-*stationary* if each $\mathcal{X}_y$ is $\mathcal{W}$-stationary in state $y$.

(iii) Let $\mathcal{S}(\mathcal{W}) = \big(\mathcal{S}_y(\mathcal{W})\big)_{y \in \mathcal{Y}}$ denote the largest family of $\mathcal{W}$-stationary payoff sets.

Observe that $\mathcal{S}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W})$ for any state $y \in \mathcal{Y}$: any payoff pair $w \in \mathcal{S}_y(\mathcal{W})$ can be attained by an enforceable solution to (5) that remains in $w$ until the first state transition occurs, at which point the continuation value jumps to $\mathcal{W}_{S_\rho}$. Therefore, the singleton $\{w\}$ is $\mathcal{W}$-relaxed self-generating and hence contained in $\mathcal{B}_y(\mathcal{W})$. Similarly to Lemma 4.8 and Proposition 4.9, we obtain the analogue results for state-order dependent PPE.

**Lemma 5.4.** *If $\mathcal{W}_y \subseteq \mathcal{S}_y(\mathcal{W})$ for every state $y$, then $\mathcal{S}(\mathcal{W})$ is mutually stationary.*

**Proposition 5.5.** *Let $\mathcal{W}_0$ be as in Proposition 4.9 and let $\mathcal{W}_n := \mathcal{S}(\mathcal{W}_{n-1})$ for any $n \geq 1$. Then $(\mathcal{W}_{n,y})_{n \geq 0}$ is decreasing for every $y \in \mathcal{Y}$ with $\bigcap_{n \geq 0} \mathcal{W}_{n,y} = \mathcal{E}^S_y(r)$. Moreover, if we set $\mathcal{W}_{n,y} = \operatorname{conv} \mathcal{S}_y(\mathcal{W}_{n-1})$ for each $n \geq 1$, then $\bigcap_{n \geq 0} \mathcal{W}_{n,y} = \mathcal{E}^S_{R,y}(r)$.*

The construction in Proposition 5.5 highlights the fact that state-order dependent PPE are locally constant until the next state transition occurs, at which point the continuation value again has to be stationary. If players are allowed to use a public randomization device, the continuation value after a state transition may come from $\operatorname{conv} \mathcal{S}_y(\mathcal{W})$.

### 5.1 Characterization of $\mathcal{S}(\mathcal{W})$

We begin by illustrating the computation of $\mathcal{S}_y(\mathcal{W})$ if $y$ has at most one successor state. If $y$ has no successor state, then $\mathcal{S}_y(\mathcal{W})$ is simply the set of static Nash equilibrium payoffs of stage game $y$. If $y$ has exactly one successor state $y'$, then $\delta^i$ in (5) is one-dimensional. Let $\mathcal{S}_{y,a}(\mathcal{W})$ denote the set of all stationary payoffs, for which there exists $\delta_0 \in \Psi_{y,a}(w, \mathcal{W})$ with $w = g(y,a) + \delta_0\lambda(y,a)$. The stationarity condition of a payoff $w \in \mathcal{S}_{y,a}(\mathcal{W})$ implies that $w + r\delta_0 = g(y,a) + \delta_0\lambda(y,a) + r\delta_0 \in \mathcal{W}_y$. The set of all payoffs that can be reached from $\mathcal{S}_{y,a}(\mathcal{W})$ after a state transition is thus

$$\big(g(y,a) + \Psi_{y,a}(\lambda(y,a) + r)\big) \cap \mathcal{W}_{y'}. \tag{7}$$

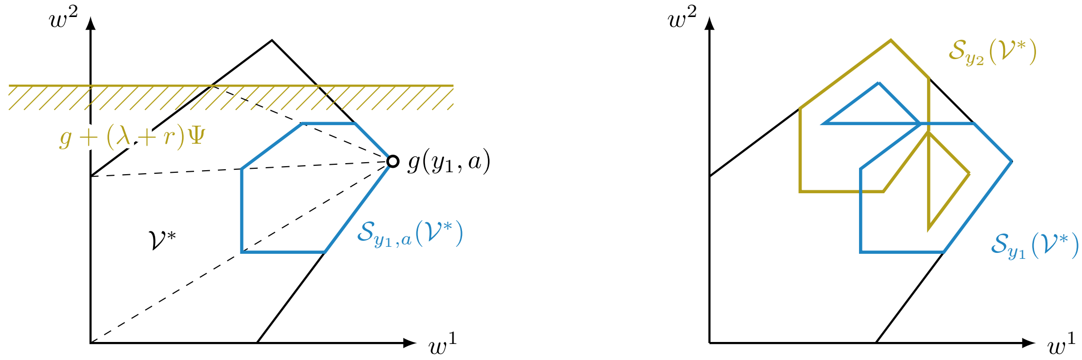



**Figure 5:** In the regime-change example, mutual effort $a = (P, A)$ can be enforced in state $y_1$ by any $\delta$ with $\delta^2 \leq 1.25$ and $\delta^1 \geq -12$. As illustrated in the left panel, the set $\mathcal{S}_{y_1,a}(\mathcal{V}^*)$ is given by shrinking the set of $(g + (\lambda + r)\Psi) \cap \mathcal{V}^*$ towards $g$. The right panel shows the family $\mathcal{S}(\mathcal{V}^*)$.

Note that (7) consists only of scalings, translates, and intersections of closed, convex sets and is thus easily computed. The continuation value $w$ is then the expected value of the current expected flow payoff $g(y, a)$ and the continuation value in (7) after a state transition occurs. Geometrically, $\mathcal{S}_{y,a}(\mathcal{W})$ is obtained by shrinking the set in (7) towards $g(y, a)$ by a factor $\frac{\lambda(y,a)}{\lambda(y,a)+r}$; see Figure 5. Repeating this construction for all action profiles $a \in \mathcal{A}(y)$ yields $\mathcal{S}_y(\mathcal{W})$.

If a state $y$ has more than one successor state, an a analogue construction is carried out in the incentive space rather than in the payoff space. The condition that the continuation value after a transition to state $y'$ comes from the set $\mathcal{W}_{y'}$ can be expressed as $g(y, a) + f_{y,y'}(\delta) \in \mathcal{W}_{y'}$, where $f_{y,y'}(\delta) = \delta(\lambda(y, a) + re_{y'})$ and $e_{y'}$ is the unit vector in the direction of state $y'$ among all successor states of $y$. This eliminates the variable $w$ and allows us to parametrize the set of stationary payoffs via incentives $\delta$. Such incentives have to come from $\Psi_{y,a}$ and they have to satisfy the jump condition for every $y \in Y$, i.e., the set of all such incentives is given by

$$\mathcal{X}_{y,a}(\mathcal{W}) := \Psi_{y,a} \cap \bigcap_{y' \in \mathcal{Z}_y(a)} f^{-1}_{y,y'}(\mathcal{W}_{y'} - g(y, a)),$$

where $f^{-1}_{y,y'}$ denotes the inverse image under $f_{y,y'}$. It is now straightforward that

$$\mathcal{S}_y(\mathcal{W}) = \bigcup_{a \in \mathcal{A}} (g(y, a) + \mathcal{X}_{y,a}(\mathcal{W})\lambda(y, a)), \tag{8}$$

where $\mathcal{X}_{y,a}(\mathcal{W})\lambda(y, a) := \{v \in \mathbb{R}^2 \mid \exists\, \delta \in \mathcal{X}_{y,a}(\mathcal{W}) \text{ with } \delta\lambda(y, a) = v\}$ denotes the projection of $\mathcal{X}_{y,a}(\mathcal{W})$ onto $\mathbb{R}^2$ in the direction $\lambda(y, a)$. Note that $\Psi_{y,a}$ is a closed, convex polytope characterized by the affine inequalities in (6). Consider first the case, where $\mathcal{W}$ is a family of polygons $\mathcal{W}_y$ with extremal points $v_{y,1}, \ldots, v_{y,n_y}$ and corresponding normal vectors $N_{y,1}, \ldots, N_{y,n_y}$. Then the set $\mathcal{X}_{y,a}(\mathcal{W})$ is a convex polytope, characterized by the affine inequalities in (6) and

$$N_{y',j}{}^{\top}(\delta(\lambda(y, a) + re_{y'})) \leq N_{y',j}{}^{\top} x_{y',j}, \qquad j = 1, \ldots, n_{y'},\ y' \in \mathcal{Z}_y(a).$$

If $|\mathcal{Z}_y(a)|$ is sufficiently small so that extremal points of $\mathcal{X}_{y,a}(\mathcal{W})$ can be computed efficiently, then

$\mathcal{X}_{y,a}(\mathcal{W})\lambda(y,a)$ can be computed by projecting the extremal points of $\mathcal{X}_{y,a}(\mathcal{W})$.Alternatively, one can maximize $N^\top \delta\lambda(y,a)$ over $\delta \in \mathcal{X}_{y,a}(\mathcal{W})$ for a sufficiently rich grid of normal vectors $N$. For general $\mathcal{W}$, one can use inner and outer polygon approximations $\underline{\mathcal{W}}_y$ and $\overline{\mathcal{W}}_y$ of $\mathcal{W}_y$, respectively, as in Yeltekin et al. (2017) and then use the steps above.

### 5.2 Markov perfect equilibria

Markov-perfect equilibria are a special class of state-order dependent equilibria that depend on the sequence of observed states only through the current state. For initial state $y_0$, let $\mathcal{Y}(y_0)$ denote the set of all states that can be reached from $y_0$, including $y_0$ itself. The following is a verification result for whether or not a given map from states to action profiles is a Markov-perfect equilibrium.

**Proposition 5.6.** *A map $a_* : \mathcal{Y} \to \mathcal{A}$ with $a_*(y) \in \mathcal{A}(y)$ for each $y$ is a Markov-perfect equilibrium for initial state $y_0$ if and only if for every state $y \in \mathcal{Y}(y_0)$, player $i = 1, 2$, and $a^i \in \mathcal{A}^i(y)$, we have*

$$w_*^i(y) \geq g^i(y, (a^i, a_*^{-i}(y))) + \sum_{y' \in \mathcal{Y}} \delta_*^i(y, y')\lambda_{y,y'}(a^i, a_*^{-i}(y)), \tag{9}$$

*where we denote by $\Lambda_{y_0}(a_*)$ the matrix with entries $\lambda_{y,y'}(a_*(y))$ in row $y'$ and column $y$ for $(y, y') \in \mathcal{Y}(y_0)^2$, by $G_{y_0}(a_*)$ and $w_*$ the matrices with entries $g^i(y, a_*(y))$ and $w_*^i(y)$, respectively, in row $i$ and column corresponding to $y \in \mathcal{Y}(y_0)$, by $\mathbf{1}$ the $|\mathcal{Y}(y_0)|$-dimensional row vector containing all ones, and*

$$w_* := rG_{y_0}(a_*)\big(\operatorname{diag}(r\mathbf{1} + \mathbf{1}\Lambda_{y_0}(a_*)) - \Lambda_{y_0}(a_*)\big)^{-1}, \qquad \delta_*(y, y') := \frac{w_*(y') - w_*(y)}{r}. \tag{10}$$

*Moreover, $W(a_*(S_-)) = w_*(S)$, where $S$ is the state process starting in $y_0$.*

It is worth noting that for a fixed map $a_*$ from states to action profiles, $w_*$ and $\delta_*$ are given explicitly by the respective expressions in (10). A naive algorithm is thus to verify the conditions of all such maps $a_*$ since we restrict attention to pure-strategy equilibria in this paper. While this comes at almost no loss of generality for continuous-time PPE, it is a significant restriction for Markov-perfect equilibria; see also Footnote 6.

## 6 Characterization of $\mathcal{B}(\mathcal{W})$

In Section 4, we motivated the construction of equilibria as enforceable solutions to (5). Due to the iterative procedure in Proposition 4.9, it is sufficient to construct such solutions up until the first state transition if we additionally impose that continuation values after a state transition come from a fixed family of payoff sets. This local construction of strategy profiles can be viewed as equilibria of an auxiliary repeated game that ends at the time of a state transition with terminal payoffs from $\mathcal{W} = (\mathcal{W}_y)_{y \in \mathcal{Y}}$. A "state transition" in the auxiliary game is then simply a discrete signal about past play that ends the game. The fact that this local construction is a continuous-time repeated game with public signal $(X, S)$ means that we can characterize $\mathcal{B}_y(\mathcal{W})$ for a fixed state $y$ with techniques similar to those in Bernard (2023). This section thus follows that paper rather closely.

In the previous section, we established that the family of $\mathcal{W}$-stationary payoff sets is contained in $\mathcal{B}(\mathcal{W})$. In a non-stationary payoff, the public signal as well as the history of the state process may be used to provide intertemporal incentives. Incentives at the boundary of $\mathcal{B}_y(\mathcal{W})$ must satisfy an analogue to informational restrictions (I1)–(I3), formalized as follows.

**Definition 6.1.** For a payoff pair $w \in \mathbb{R}^2$, a direction $N \in S^1$, and a family $\mathcal{W}$ of payoff sets, we say that a continuation promise $(\beta, \delta)$ from the set

$$\Xi_{y,a}(w, N, \mathcal{W}) := \left\{ (\beta, \delta) \;\middle|\; \begin{array}{l} (\beta, \delta) \text{ enforces } a \text{ in state } y, N^\top(g(y,a) + \delta\lambda(y,a) - w) \geq 0, \\ N^\top\beta = 0, \text{and } w + r\delta(y') \in \mathcal{W}_{y'} \text{ for all } y' \in \mathcal{Z}_y(a) \end{array} \right\}$$

*restricted-enforces* $a$ in state $y$. An action profile $a \in \mathcal{A}(y)$ is *restricted-enforceable* for $(w, N, \mathcal{W})$ in state $y$ if the set $\Xi_{y,a}(w, N, \mathcal{W})$ is non-empty.

### 6.1 Incentives provided through state transitions

Stationary payoffs are those payoffs that can be attained when rewards or punishments are provided only after a state transition occurs. In this subsection, we characterize those payoff pairs that can be attained by an enforceable strategy profile if, in addition, players condition on the absence of state transitions. For a state $y$ and an action profile $a \in \mathcal{A}(y)$, denote by

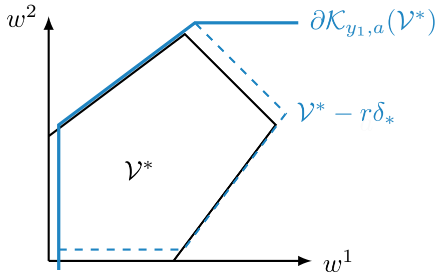
$$\mathcal{K}_{y,a}(\mathcal{W}) = \left\{ w \in \mathbb{R}^2 \;\middle|\; \Psi_{y,a}(w, \mathcal{W}) \neq \emptyset \right\}$$

the set of all payoff pairs, at which incentives can be provided through continuation values in $\mathcal{W}_{y'}$ after a transition to any state $y'$. The set $\mathcal{K}_{y,a}(\mathcal{W})$ has a particularly simple geometric interpretation if $y$ has a unique successor state $y'$. In that case, extremal points of $\mathcal{K}_{y,a}(\mathcal{W})$ are translates of $\partial\mathcal{W}_{y'}$ by the extremal punishments/rewards necessary to enforce $a$. See Figure 6 for this construction in the regime-change game of Section 3. In general, at the boundary of $\mathcal{K}_{y,a}(\mathcal{W})$, there must be at least one player $i$, for whom the minimal or maximal continuation values in $\mathcal{W}$ just provide sufficient incentives not to deviate from $a^i$. We say that incentives at these payoff pairs are *binding* in state $y$ and *extremal* in $\mathcal{W}$ for player $i$. The set of all payoff pairs, at which incentives through the state process are binding and extremal, is denoted by $\mathcal{K}_y(\mathcal{W}) = \bigcup_{a \in \mathcal{A}(y)} \mathcal{K}_{y,a}(\mathcal{W})$.

Whether the state process is sufficient to provide incentives depends not only on the location of the payoff pair, but also on the direction of the outward normal vector due to informational restriction (I1). Incentives from state transitions are sufficient for $(w, N)$ in the set

$$\Gamma_y(\mathcal{W}) := \left\{ (w, N) \in \mathbb{R}^2 \times S^1 \;\middle|\; \begin{array}{l} \text{There exist } a \in \mathcal{A} \text{ and } \delta \in \Psi_{y,a}(w, \mathcal{W}) \\ \text{with } N^\top(g(y,a) + \delta\lambda(y,a) - w) \geq 0 \end{array} \right\}.$$

For any convex set $\mathcal{X}$ and any $w \in \partial\mathcal{X}$, let $\mathcal{N}_w(\mathcal{X})$ denote the set of outward normal vectors to $\mathcal{X}$ at $w$. Denote by $\mathcal{N}_{\mathcal{X}} := \{(w, N) \mid w \in \partial\mathcal{X} \text{ and } N \in \mathcal{N}_w(\mathcal{X})\}$ the *normal bundle* of $\mathcal{X}$. With slight abuse of terminology, we will refer to the boundary of $\mathcal{B}_y(\mathcal{W})$ within $\Gamma_y(\mathcal{W})$ when referring to boundary points $w \in \partial\mathcal{B}_y(\mathcal{W})$, for which there exists an outward normal vector $N$ with $(w, N) \in \Gamma_y(\mathcal{W})$.



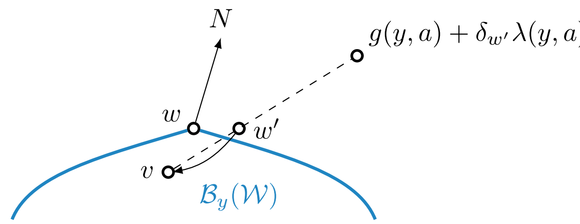




**Figure 6:** The left panel shows $\mathcal{K}_{y_1,a}(\mathcal{V}^*)$ in the regime-change game for $a = (M, I)$. Extremal points in $\mathcal{K}_{y_1,a}(\mathcal{V}^*)$ are offset from extremal points in $\mathcal{V}^*$ by the minimal rewards/punishments $r\delta_*^i$ needed to enforce $(M, I)$. The right panel illustrates the perturbation argument behind the proof of Lemma 6.2: if the lemma was violated, then $w'$ slightly outside of $\mathcal{B}_y(\mathcal{W})$ could be attained by a $\mathcal{W}$-enforceable solution to (5).

Our first result states that at the boundary of $\mathcal{B}_y(\mathcal{W})$, the continuation value of any strategy profile must evolve tangentially to $\mathcal{B}_y(\mathcal{W})$ unless incentives are binding and extremal.

**Lemma 6.2.** *For any state $y$, any payoff pair $w \in \partial\mathcal{B}_y(\mathcal{W}) \setminus \mathcal{K}_y(\mathcal{W})$ and any normal vector $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$, there exist no $a \in \mathcal{A}(y)$ and $\delta \in \Psi_{y,a}(w, \mathcal{W})$ with $N^\top(g(y,a) + \delta\lambda(y,a) - w) > 0$.*

The idea is that if such $(w, N, a, \delta)$ existed, then $w$ could be attained by an enforceable solution to (5) whose continuation value enters strictly into the interior of $\mathcal{B}_y(\mathcal{W})$. Continuity of $w \mapsto \Psi_a(w, \mathcal{W})$ within $\mathcal{K}_{y,a}(\mathcal{W})$ established in Lemma E.3 then implies that a payoff pair $w' \notin \operatorname{cl}\mathcal{B}_y(\mathcal{W})$ sufficiently close to $w$ can be attained by an enforceable solution $W$ to (5) that reaches some payoff pair $v$ in the interior of $\mathcal{B}_y(\mathcal{W})$; see Figure 6. By definition of $\mathcal{B}_y(\mathcal{W})$, payoff pair $v$ can be attained by an enforceable solution $W'$ to (5) that remains in $\mathcal{B}_y(\mathcal{W})$ until the first state transition occurs. Therefore, the concatenation of $W$ and $W'$ shows that $w' \in \mathcal{B}_y(\mathcal{W})$, a contradiction.

Lemma 6.2 has two implications. First, it tells us that all non-stationary corners are contained in $\mathcal{K}_y(\mathcal{W})$. Indeed, the only way for $g(y,a) + \delta\lambda(y,a)$ to be tangential to $w$ in three distinct directions is if $w = g(y,a) + \delta\lambda(y,a)$, i.e., $w \in \mathcal{S}_{y,a}(\mathcal{W})$.[15] Second, Lemma 6.2 implies that at any boundary payoff outside of $\mathcal{S}_y(\mathcal{W}) \cup \mathcal{K}_y(\mathcal{W})$, the boundary is a continuously differentiable solution to the *abrupt-information optimality equation*

$$N_w{}^\top w = \max_{a\in\mathcal{A}} \max_{\delta\in\Psi_{y,a}(w,\mathcal{W})} N_w{}^\top(g(y,a) + \delta\lambda(y,a)), \tag{11}$$

To interpret (11), suppose first that $N_w^i = \frac{1}{\sqrt{2}}$ so that $w$ maximizes the sum of payoffs in $\mathcal{B}_y(\mathcal{W})$. There are two reasons why $N_w{}^\top\delta\lambda(y,a)$ may be non-positive for some action profile $a$. First, it may be necessary to destroy a certain amount of value upon a state transition in order to incentivize $a$, i.e., $\delta^i(y') \leq 0$ is imposed by the enforceability constraint. This may occur if $a$ is a cooperative action profile and a state transition is more likely under a deviation. The abrupt-information optimality equation (11) states that on the efficient frontier, the lowest amount of value is burnt. Second, it is possible that state $y$ is the most "efficient" state in the sense that the sum of payoffs are lower

[15] Formally, the order of quantifiers in Lemma 6.2 allows $\delta$ to depend on $N$. Since $g(y,a) + \Psi_{y,a}(w, \mathcal{W})\lambda(y,a)$ is convex, three different supporting tangents go through $w$ only if $w \in g(y,a) + \Psi_{y,a}(w, \mathcal{W})\lambda(y,a)$.

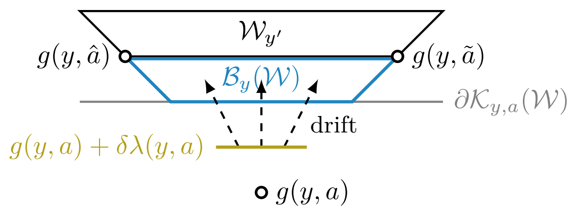



**Figure 7:** Consider a game with two states $y$ and $y'$, two static Nash equilibria $\hat{a}$ and $\tilde{a}$ in state $y$, and an action profile $a$ with $\lambda(y,a)=1$ that is enforced in state $y$ by $\delta^2 \leq 1$ and $\delta^1 \in [-1,1]$ with expected flow payoffs as indicated above. For $r=1$ and $\mathcal{W}_{y'}$ as above, all payoffs on $\partial\mathcal{K}_{y,a}(\mathcal{W}) \cap \mathcal{B}_y(\mathcal{W})$ except the extremal payoffs are strictly decomposed by $a$ and $\delta$ with $\delta^2 = 1$ as indicated by the strictly inward-pointing drift. The remaining sides of $\mathcal{B}_y(\mathcal{W})$ are straight solutions to (11).

in the other states, i.e., $N^\top(v-w) \leq 0$ for any $v \in \mathcal{W}_{y'}$ and any successor state $y'$. In this case, (11) aims to maximize expected flow payoffs while minimizing the likelihood of state transitions. If the expression ${N_w}^\top \delta\lambda(y,a)$ is positive, then state changes are efficient in expectation. Then the maximization in (11) trades off expected flow payoffs and the frequency of state transitions, with the latter becoming more important the higher the payoff difference is across states. The same logic applies for arbitrary $N_w$ if the sum of payoffs is replaced with the weighted sum of payoffs.

It remains to characterize the corners in $\mathcal{K}_y(\mathcal{W})$.

**Definition 6.3.** Consider $w \in \partial\mathcal{B}_y(\mathcal{W})$ and a set of action profiles $\mathcal{A}_w \subseteq \mathcal{A}(y)$.

(i) $\mathcal{A}_w$ *decomposes* $w$ if for any $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$, there exist $a \in \mathcal{A}_w$ and $\delta \in \Psi_{y,a}(w,\mathcal{W})$ with $N^\top(g(y,a)+\delta\lambda(y,a)-w) \geq 0$. Such $w$ is said to be *decomposable.*

(ii) $\mathcal{A}_w$ *strictly decomposes* $w$ if $\mathcal{A}_w$ decomposes $w$ and for each $N \in \operatorname{ext}\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$, there exist $a \in \mathcal{A}_w$ and $\delta \in \Psi_{y,a}(w,\mathcal{W})$ with $N^\top(g(y,a)+\delta\lambda(y,a)-w) > 0$.

(iii) $\mathcal{A}_w$ *minimally decomposes* $w$ if $\mathcal{A}_w$ decomposes $w$ and no proper subset of $\mathcal{A}_w$ decomposes $w$.

The following lemma is a complete characterization of boundary points of $\mathcal{B}_y(\mathcal{W})$ within $\Gamma_y(\mathcal{W})$. If the continuous public signal is uninformative, Lemma 6.4 characterizes $\mathcal{B}_y(\mathcal{W})$.

**Lemma 6.4.** *Consider $(w,N) \in \mathcal{N}_{\mathcal{B}_y(\mathcal{W})} \cap \Gamma_y(\mathcal{W})$ with non-stationary $w$. Then exactly one of the following conditions holds:*

*(i) There exists a set $\mathcal{A}_w \subseteq \mathcal{A}(y)$ that strictly and minimally decomposes $w$ such that $w \in \partial\mathcal{K}_{y,a}(\mathcal{W})$ for each $a \in \mathcal{A}_w$ and*

$$\mathcal{N}_w(\mathcal{B}_y(\mathcal{W})) \subseteq \mathcal{N}_w(\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W})), \tag{12}$$

*where we denote $\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W}) := \bigcap_{a\in\mathcal{A}_w} \mathcal{K}_{y,a}(\mathcal{W})$.*

*(ii) $(w,N)$ satisfies (11) and either of the following conditions hold:*

*(a) $w$ is in the interior of $\mathcal{K}_{y,a_*}(\mathcal{W})$ and $\mathcal{N}_w(\mathcal{B}_y(\mathcal{W})) = \{N\}$,*

*(b) $w$ is decomposed by $a_*$ and $w \in \partial\mathcal{K}_{y,a_*}(\mathcal{W})$,*

*where $a_*$ is the action profile attaining the maximum in (11).*

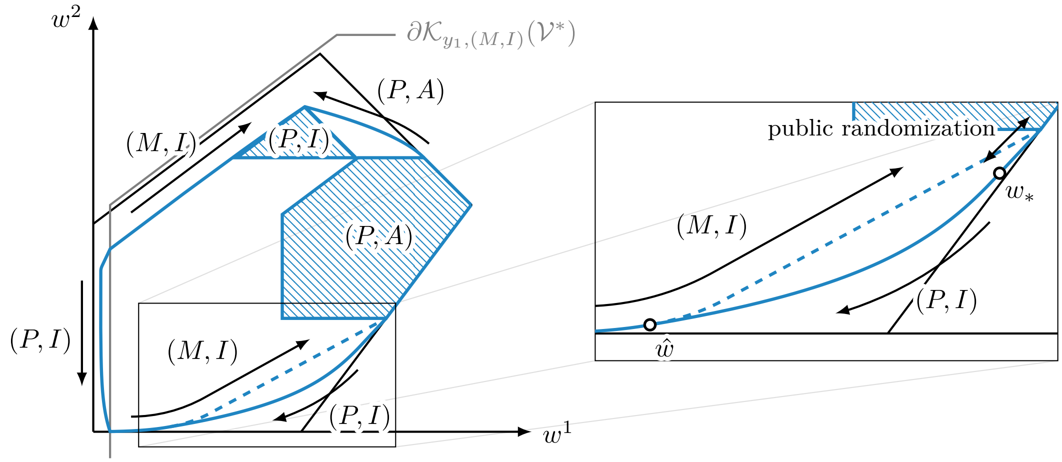



**Figure 8:** The set of all payoff pairs in $\mathcal{B}_y(\mathcal{V}^*)$ that can be attained with incentives related to the state process only is characterized by Lemma 6.4. Its boundary consists of stationary payoffs, solutions to (11), and two short segments on which public randomization is optimal. All corners are either stationary or the starting points on $\partial\mathcal{K}_{y_1,(M,I)}(\mathcal{V}^*)$ of continuously differentiable solutions to (11).

The essence of Lemma 6.4 is that if the state process is sufficient to provide incentives locally, the boundary is a differentiable solution to (11) except at payoff pairs $w \in \mathcal{K}_y(\mathcal{W})$, where incentives are binding and extremal. Statement (ii) characterizes those solutions, where (ii.a) concerns the relative interior of those solutions and (ii.b) describes their end points. Statement (i) describes strictly decomposable boundary payoffs in $\mathcal{K}_y(\mathcal{W})$, where (12) relates the shape of the boundary of $\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W})$ for the minimally decomposing action profiles $\mathcal{A}_w$ to the shape of the boundary of $\mathcal{B}_y(\mathcal{W})$. Specifically, $\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W})$ is locally a subset of $\mathcal{B}_y(\mathcal{W})$. The set of points $\mathcal{K}_y(\mathcal{W})$, where the boundary is not a solution to (11), has measure 0 in $\mathbb{R}^2$ since it is the finite union of one-dimensional objects. In the regime-change game of Section 3, the intersection of $\mathcal{K}_y(\mathcal{V}^*)$ with $\mathcal{V}^*$ is only one straight line segment; see Figure 8. Nevertheless, it is possible that boundary segments of positive length are given by $\partial\mathcal{K}_{y,a}(\mathcal{W})$; see Figure 7. While smooth line segments in $\Gamma_y(\mathcal{W})$ are decomposable by a single action profile, corners may not be. If a corner is minimally decomposed by a non-singleton $\mathcal{A}_w$, statement (i) implies that it must be strictly decomposable. If a corner is not strictly decomposable, it is either stationary or the starting point of a continuously differentiable solution to (11) as in case (ii.b).

Figure 8 shows all payoffs in $\mathcal{B}_y(\mathcal{W})$ in the regime-change example that can be attained with incentives from the state process only. Let us briefly discuss how strategy profiles are related to the solutions of these ODEs in the regime-change game; see Figure 8. At stationary payoffs, players play a locally constant strategy profile with fixed rewards/punishments attached to a state transition. At solutions to (11), players locally play the unique maximizing action profile in (11). In absence of state transitions, the continuation value travels along the boundary as a deterministic solution to (5). Let $t_0$ and $t_1$ denote the deterministic times required for such a solution to reach $\hat{w}$ from $w_*$ when $(P, I)$ is played and from $\hat{w}$ to reach $\mathcal{S}_{y_1}(\mathcal{V}^*)$ if $(M, I)$ is played. Let $\tau$ denote the time of the

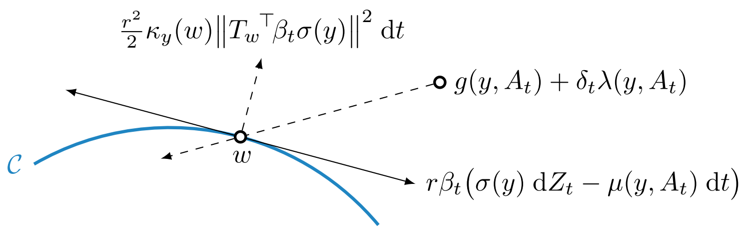



**Figure 9:** The infinitesimally quick oscillation in the continuation value due to unbounded variation of value transfers causes the continuation value to drift away from the curve in orthogonal direction.

first state transition. Then $w_*$ is locally attained by the strategy profile

$$A(t, S_s, s < t) = \begin{cases} (P, I) & \text{if } t < \min\{t_0, \tau\}, \\ (M, I) & \text{if } t_0 \leq t < \min\{t_1, \tau\}, \\ (P, A) & \text{if } t_1 \leq t < \tau. \end{cases} \tag{13}$$

with a suitable continuation value in $\mathcal{V}^*$ if a state transition occurs. Contrary to stationary profiles, the continuation value after a state transition is not fixed but continually adjusted with time to preserve incentives as the continuation profile changes. From (13), we see that the non-incumbent first attempts to instigate a revolution. If unsuccessful up to time $t_0$, there will be a punishment phase of length $t_1 - t_0$, during which the mutually uncooperative $(M, I)$ is played. The punishment is costly for both, hence the continuation value of both players increases as the punishment phase progresses towards its end. At time $t_1$, players start to cooperate and play $(P, A)$ until a state transition occurs.

### 6.2 Incentives provided through the public signal

Let us now add incentives from the public signal to those from state transitions. Since the unbounded variation of Brownian motion drives the argument, the argument is similar to that in Sannikov (2007). Consider, for a moment, that an enforceable solution $W$ to (5) for initial state $y \in \mathcal{Y}$ remains on a continuously differentiable curve $\mathcal{C}$ before any state transitions occur. Then the public signal must be used to provide incentives via tangential value transfers as illustrated in Figure 9. Due to unbounded variation of Brownian motion, players transfer value very rapidly. The infinitesimally quick tangential oscillation created by those transfers causes the continuation value to drift away from the curve in orthogonal direction, similar in spirit to the centrifugal force in physics. The larger the curvature is, the stronger is the outward drift created by the tangential transfers. The magnitude of the outward drift is given by Itō's formula and it is equal to $\frac{r^2}{2}\kappa_y(w)\left\|{T_w}^\top \beta_t \sigma(y)\right\|^2$. For the continuation value to stay on the curve, the outward drift has to be offset precisely by the inward drift, hence

$$\frac{r^2}{2}\kappa_y(w)\left\|{T_w}^\top \beta_t \sigma(y)\right\|^2 = r{N_w}^\top\big(g(y, A_t) + \delta_t \lambda(y, A_t) - w\big).$$

To complete the argument, it is necessary that whenever the continuation value revisits the same point on the curve, the tangential transfers induce the same curvature. This is the case when the

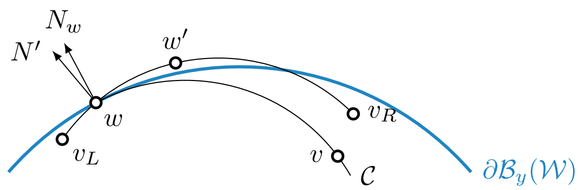



**Figure 10:** If a solution $\mathcal{C}$ to (14) starting at $(w, N_w)$ falls into the interior of $\mathcal{B}_y(\mathcal{W})$, then a solution $\mathcal{C}'$ with initial conditions $(w, N')$ for a slight rotation $N'$ of $N_w$ would leave and re-enter $\mathcal{B}_y(\mathcal{W})$. But then the strategy profile of the maximizers in (14) would attain $w' \in \mathcal{B}_y(\mathcal{W})$, a contradiction.

chosen action profiles and the provided continuation promises are Markovian in the continuation value. The following lemma states that at the boundary of $\mathcal{B}_y(\mathcal{W})$, enforceable action profiles and their continuation promises are generically unique, given by those that maximize the expression for the curvature. In particular, action profiles and continuation promises are Markovian and, hence, the continuation value of a locally enforceable strategy remains on the boundary of $\mathcal{B}_y(\mathcal{W})$.

**Lemma 6.5.** *Suppose that Assumptions 1–2 hold. Outside of $\Gamma_y(\mathcal{W})$, the boundary $\partial\mathcal{B}_y(\mathcal{W})$ is continuously differentiable with curvature at almost every point $w$ given by*

$$\kappa_y(w) = \max_{a \in \mathcal{A}} \max_{(\beta,\delta) \in \Xi_{y,a}(w,N_w,\mathcal{W})} \frac{2{N_w}^\top(g(y,a) + \delta\lambda(y,a) - w)}{r\left\|{T_w}^\top \beta\sigma(y)\right\|^2}. \tag{14}$$

The optimality equation (14) shows the trade-off between incentives provided through state transitions and the public signal. If payoff pair $w$ is efficient in direction $N_w$, then state transitions destroy value. Thus, the frequency of state transitions should be minimized and incentives should be provided mainly through the public signal. If $w$ is inefficient in direction $N_w$, then state transitions increase the weight sum of payoffs, hence the public signal is used only to provide the residual incentives that cannot be provided through state transitions.

Figure 10 illustrates why the curvature in (14) is maximized over all restricted-enforceable action profiles and incentives. If the curvature of the boundary were smaller, then a solution $\mathcal{C}$ to (14) starting at the boundary would fall into the interior of $\mathcal{B}_y(\mathcal{W})$. Thus, a solution $\mathcal{C}'$ to (14) for slightly rotated initial conditions would leave and re-enter $\mathcal{B}_y(\mathcal{W})$. In particular, payoff pairs on $\mathcal{C}'$ outside of $\mathcal{B}_y(\mathcal{W})$ could be attained by the enforceable strategy profile given by the maximizer in (14), contradicting maximality of $\mathcal{B}_y(\mathcal{W})$. This argument requires continuity of solutions to (14) in initial conditions, which we discuss in Section 6.4. On the other hand, a solution $\mathcal{C}$ to (14) cannot escape $\mathcal{B}_y(\mathcal{W})$ either because $\mathcal{C}$ maximizes the curvature over all restricted-enforceable action profiles and their incentives. Details of the arguments are contained in Appendix D.1.

### 6.3 Family of PPE payoff sets

Lemmas 6.4 and 6.5 give rise to the following result, characterizing the family of payoff sets $\mathcal{B}(\mathcal{W})$ for any family $\mathcal{W}$ of convex and compact payoff sets. It also characterizes the family of PPE payoff sets through $\mathcal{E}_y(r) = \mathcal{B}_y(\mathcal{E}(r))$ if there are no possible cycles among states. If the state process can return to the same state, then Theorem 6.6 provides an approximation for $\mathcal{E}(r)$ via Proposition 4.9.

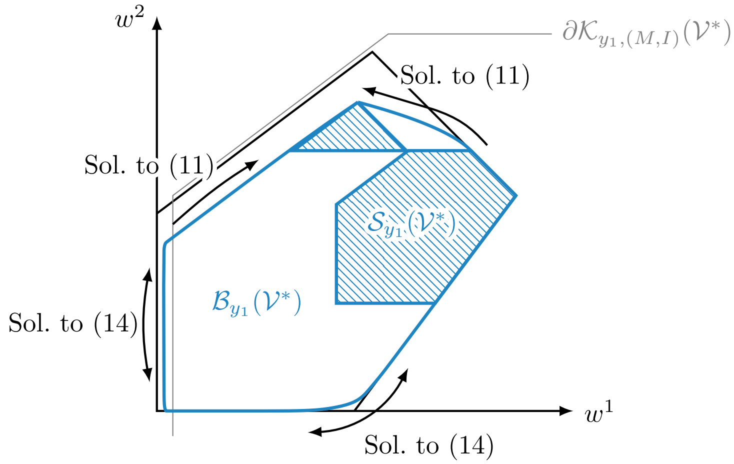



**Figure 11:** The set $\mathcal{B}_{y_1}(\mathcal{V}^*)$ of relaxed-generating payoff sets in the regime-change example after one step of the algorithm in Proposition 4.9. In this example, the boundary is a differentiable solution to (14) or (11) outside of $\mathcal{S}_{y_1}(\mathcal{V}^*)$. In general, it could have corners in $\mathcal{K}_{y_1}(\mathcal{V}^*)$.

**Theorem 6.6.** *Let $\mathcal{W}$ be a family of convex and compact payoff sets. Then $\mathcal{B}_y(\mathcal{W})$ is the largest closed convex subset $\mathcal{X}$ of $\mathcal{V}^*$ that contains $\mathcal{S}_y(\mathcal{W})$ such that:*

*(i) Outside of $\mathcal{S}_y(\mathcal{W}) \cup \mathcal{K}_y(\mathcal{W})$, the boundary $\partial\mathcal{X}$ is continuously differentiable and $(w, N_w) \in \mathcal{N}_\mathcal{X}$ solves (11) within $\Gamma_y(\mathcal{W})$ and (14) outside of $\Gamma_y(\mathcal{W})$,*

*(ii) Every corner $w \in \partial\mathcal{X}$ is either stationary, minimally decomposable by a maximizer $a_*$ of (11) for $(w, N)$ with $N \in \operatorname{ext}\mathcal{N}_w(\mathcal{X})$ and $w \in \partial\mathcal{K}_{y,a_*}(\mathcal{W})$, or minimally and strictly decomposable by some $\mathcal{A}_w \subseteq \mathcal{A}(y)$ with $\mathcal{N}_w(\mathcal{X}) \subseteq \mathcal{N}_w(\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W}))$.*

Since we have discussed the individual components of the result after Lemmas 6.4 and 6.5 already, we move on to the discussion of $\mathcal{B}(\mathcal{V}^*)$ in our regime-change example. Figure 11 shows that the addition of the public signal enlarges $\mathcal{B}_{y_1}(\mathcal{V}^*)$. While it is hard to make out with the naked eye, the solutions to (14) are strictly curved everywhere. All the corners outside the set of stationary payoffs are smoothened out by the use of the public signal: while the state process alone is not sufficient to enforce $(M, I)$ to the left of $\partial\mathcal{K}_{y_1,(M,I)}$, players can compensate the lack of incentives provided after a state transition through tangential value transfers. Strategy profiles attaining payoff pairs on the lower boundary are given by the maximizing action profile in (14). Writing the strategy profiles as an explicit function of $X$ and $S$ as in (13), however, is not so straightforward anymore. Players essentially have to solve the SDE (5) for the promised values $(\beta, \delta)$ and the observed values of $\sigma(y)\,\mathrm{d}Z_t - \mu(y, A_t)\,\mathrm{d}t$.

The operator $\mathcal{B}$ in stochastic games differs from the corresponding operator in repeated games through the fact that continuation values after state transitions come from a family of payoff sets $\mathcal{W}_y$. Aside from a more complex mathematical foundation required for the model, this also complicates the regularity conditions needed for Lipschitz continuity of the optimality equation. The appendices contain all the proofs of the model setup in Section 4, the characterization of Markov-perfect and state-order dependent PPE payoff sets in Section 5, as well as local Lipschitz continuity of (14).

Once local Lipschitz continuity of the optimality equation is established, the proofs of Lemmas 6.2–6.5 are identical to Bernard (2023) and, hence, omitted here. This is because the local construction of equilibria relies on auxiliary repeated games with a signal structure covered in Bernard (2023).

### 6.4 Discussion of assumptions

Continuity of solutions to (14) in initial conditions requires that $\kappa_y$ is locally Lipschitz continuous in $(w, N)$. One can show that for a fixed action profile, the right-hand side of (14) is locally Lipschitz continuous at $(w, N)$ in the interior of the effective domain $\{(w, N) \mid \Xi_{y,a}(w, N, \mathcal{W}) \neq \emptyset\}$ of $\Xi_{y,a}$. The expression for the curvature may thus fail to be Lipschitz continuous only where the maximizing action profile in (14) fails to be restricted-enforceable within a small neighborhood. Assumptions 1 and 2 guarantee that the maximizing action profile is restricted-enforceable for small perturbations.

To understand the additional conditions needed in comparison to repeated games, it will be helpful to represent the enforceability condition (6) in matrix notation. Let $G^i_y(a)$ denote the row vector with entries $g^i(y, \tilde{a}^i, a^{-i}) - g^i(y, a)$ and let $M^i_y(a)$ and $\Lambda^i_y(a)$ denote the matrices with column vectors $\mu(y, \tilde{a}^i, a^{-i}) - \mu(y, a)$ and $\lambda(y, \tilde{a}^i, a^{-i}) - \lambda(y, a)$ for each $\tilde{a}^i \in \mathcal{A}^i(y) \setminus \{a^i\}$, respectively. Then $a$ is enforceable in state $y$ if and only if $G^i_y(a) \leq \beta^i M^i_y(a) + \delta^i \Lambda^i_y(a)$ for $i = 1, 2$ has a solution, where the inequality is understood elementwise. In a repeated game, the set of available incentives $\Xi_a(N)$ depends only on the direction of the tangent. Pairwise identifiability of the public signal guarantees that all enforceable action profiles can be enforced on any tangents that are not parallel to a coordinate direction: if $G^i_y(a) \leq \beta^i M^i_y(a)$ has a solution, so does $G^i_y(a) \leq T^i \phi M^i_y$, where $T = (T^1, T^2)$ is the tangent vector. In a stochastic game, the set of available incentives $\delta$ also depends on $w$ and $\mathcal{W}$ through the condition $w + r\delta(y') \in \mathcal{W}_{y'}$. Thus, whether $G^i_y(a) \leq \beta^i M^i_y(a) + \delta \Lambda^i_y(a)$ has a solution in the first place will generally depend on $w$ and $\mathcal{W}$. Since we have little control over where $w$ and $\mathcal{W}_{y'}$ lie, to ensure that it has a solution for any admissible $\delta$ if it has a solution for $\delta = 0$, we impose the individual full rank condition so that $G^i_y(a) - \delta \Lambda^i_y(a) \leq \beta^i M^i_y(a)$ can be solved with equality. Together with pairwise identifiability, this yields Assumption 1.

To guarantee Lipschitz continuity in coordinate directions, Sannikov (2007) shows that a sufficient condition for repeated games is the unique best response property: if $a^i$ is a best response to $a^{-i}$, then $G^i_y(a) < 0$. This condition ensures that we have to provide incentives only to player $-i$ for tangents $T^i$ close to coordinate directions. Suppose $\beta^{-i} = T^{-i}\phi$ provides sufficient incentives to player $-i$ to not deviate from $a$, then player $i$ has no incentive to deviate for continuation promise $T^i\phi$ either as long as $|T^i|$ is sufficiently small so that $G^i_y(a) < -\varepsilon \leq T^i \phi M^i_y(a)$ for some $\varepsilon > 0$ that exists since $G^i_y(a) < 0$. In a stochastic game, this argument no longer works since $G^i_y(a) + \delta \Lambda^i(a)$ may not be strictly smaller than 0 even if $G^i_y(a)$ is. Instead, we need that the public signal has a product structure so that the enforceability conditions for the two players completely disentangle.

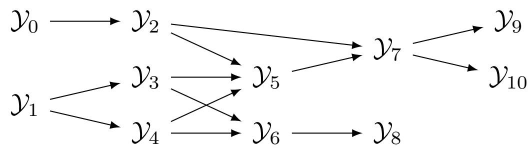



**Figure 12:** Communicating classes of a stochastic game form a directed acyclic graph.

## 7 Computation

In the terminology of Markov processes, 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 (either directly or indirectly). Computation of the family of equilibrium payoff sets proceeds by communicating classes. Communicating classes of a Markov process can be organized in a directed acyclic graph as illustrated in Figure 12.[16] A communicating class $\mathcal{Y}_0$ is a *direct predecessor class* of $\mathcal{Y}_1$, denoted $\mathcal{Y}_0 \prec \mathcal{Y}_1$, if some state in $\mathcal{Y}_1$ can be reached directly from some state in $\mathcal{Y}_0$. If $\mathcal{Y}_0 \prec \mathcal{Y}_1$, we also say that $\mathcal{Y}_1$ is a *direct successor class* of $\mathcal{Y}_0$. Consider a communicating class $\mathcal{Y}_e$ without direct successor class, that is, a class at the end of the directed graph. Since no states outside of $\mathcal{Y}_e$ can ever be reached, the subfamily $\big(\mathcal{E}_y(r)\big)_{y\in\mathcal{Y}_e}$ of equilibrium payoff sets can be computed from the algorithm in Proposition 4.9 without considering states in $\mathcal{Y}_e^c$. This is particularly simple if $\mathcal{Y}_e = \{y_e\}$ is a singleton, that is, $y_e$ is an absorbing state. The continuation game is then just a repeated game and hence $\mathcal{E}_{y_e}(r)$ can be computed with a single computation of $\mathcal{B}_{y_e}(\mathcal{V}^*)$, which reduces to Theorem 2 in Sannikov (2007) for absorbing states. One can then proceed backwards in the directed graph: consider a communicating class $\mathcal{Y}'$, for which all subfamilies of equilibrium payoff sets of direct successor classes $\mathcal{Y}_{e_1}, \dots, \mathcal{Y}_{e_n}$ have been computed already. In the computation of $\big(\mathcal{E}_y(r)\big)_{y\in\mathcal{Y}'}$, incentives from state transitions to states in $\mathcal{Y}_E := \bigcup_{k=1}^n \mathcal{Y}_{e_k}$ do not need to be computed iteratively as in Proposition 4.9, but only incentives via state transitions within the communicating class $\mathcal{Y}'$ have to be accounted for in an iterative fashion. This becomes again particularly simple if $\mathcal{Y}' = \{y'\}$ is a singleton. Then $\mathcal{E}_{y'}(r) = \mathcal{B}_{y'}\big((\mathcal{E}_y(r))_{y\in\mathcal{Y}_E}\big)$ can be solved in a single application of Theorem 6.6 rather than an iterated application. We refer to Section 8 in Bernard (2023) for notes on the implementation of Theorem 6.6.

## 8 Conclusion

Based on recent developments in continuous-time repeated games, this paper provides a unifying framework for the analysis of stochastic games with imperfect public monitoring in a continuous-time setting. The methodology is not limited to irreducible games or absorbing games and it is applicable to any stochastic game, as long as the public signal satisfies Assumptions 1 and 2. The paper characterizes the set of all PPE payoffs, the set of all Markov-perfect payoffs, as well as the payoffs of a notion of state-order dependent PPE that are simple to compute and implement, yet not as limiting as Markov-perfect equilibria.

[16] Indeed, if there was a cycle of communicating classes, then every state within the cycle could be reached from any other state in the cycle, hence the union of all communicating classes in the cycle forms a single communicating class.

Crucially, the analysis is carried out for any discount rate $r > 0$. This preserves the correspondence from initial states to equilibrium payoffs and allows us to draw conclusions on the values players 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. While we do not directly analyze the limit as players get arbitrarily patient, the extremal incentives characterized in this paper may be helpful to tackle such a problem in the future. Such a characterization would be new in stochastic games, in which the limit payoff set is not independent of the initial state.

The analysis in this paper relies on an iterative procedure over state transitions, allowing us to draw from the techniques of continuous-time repeated games in each step of the iteration. This is possible because we consider stochastic games with finitely many states and, hence, the state is constant almost everywhere. The techniques in the present paper cannot be extended in a straightforward manner to stochastic games with a continuum of states. If the state process follows a diffusion process, techniques are developed in Faingold and Sannikov (2020). For other state processes, such an extension provides an interesting direction for future research.

## 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. 2023. "Continuous-Time Games with Imperfect and Abrupt Information." *Accepted for publictation by The Review of Economic Studies.*

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.

Doraszelski, Ulrich, and Juan F. Escobar. 2010. "A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification." *Theoretical Economics*, **5(3)**: 369–402.

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. 2020. "Moral Hazard in Stochastic Differential Games: Beyond Markov Equilibrium." working paper.

Fink, A. M. 1964. "Equilibrium in a stochastic $n$-person game." *Journal of Science of the Hiroshima University, Series A-I (Mathematics)*, **28(1)**: 89–93.

Fudenberg, Drew, and David K. Levine. 1994. "Efficiency and Observability with Long-Run and Short-Run Players." *Journal of Economic Theory*, **62(1)**: 103–135.

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.

Fudenberg, Drew, David K. Levine, and Eric Maskin. 1994. "The Folk Theorem with Imperfect Public Information." *Econometrica*, **62(5)**: 997–1039.

Grossman, Sanford J., and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica*, **51(1)**: 7–45.

Haller, Hans, and Roger Lagunoff. 2000. "Genericity and Markovian Behavior in Stochastic Games." *Econometrica*, **68(5)**: 1231–1248.

He, Wei, and Yeneng Sun. 2017. "Stationary Markov perfect equilibria in discounted stochastic games." *Journal of Economic Theory*, **169**: 35–61.

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.

Jacod, J., and A. Shiryaev. 2002. *Limit Theorems for Stochastic Processes. Grundlehren der mathematischen Wissenschaften.* 2 ed., Springer Berlin Heidelberg.

Kallenberg, Olav. 2002. *Foundations of modern probability. Probability and its Applications.* Second ed., Springer-Verlag, New York.

Kazamaki, N. 2006. *Continuous Exponential Martingales and BMO. Lecture Notes in Mathematics*, Springer Berlin Heidelberg.

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.

Levy, Yehuda. 2013. "Discounted Stochastic Games With No Stationary Nash Equilibrium: Two Examples." *Econometrica*, **81(5)**: 1973–2007.

Levy, Yehuda John, and Andrew McLennan. 2015. "Corrigendum to "Discounted Stochastic Games With No Stationary Nash Equilibrium: Two Examples"." *Econometrica*, **83(3)**: 1237–1252.

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 Juuso Toikka. 2017. "Value of Persistent Information." *Econometrica*, **85(6)**: 1921–1948.

Protter, P.E. 2005. *Stochastic Integration and Differential Equations. Stochastic Modelling and Applied Probability.* 2 ed., Springer Berlin Heidelberg.

Rosenberg, Dinah, and Sylvain Sorin. 2001. "An operator approach to zero-sum repeated games." *Isral Journal of Mathematics*, **121**: 221—246.

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

Sannikov, Yuliy, and Andrzej Skrzypacz. 2010. "The Role of Information in Repeated Games With Frequent Actions." *Econometrica*, **78(3)**: 847–882.

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

Sun, Yeneng. 2006. "The exact law of large numbers via Fubini extension and characterization of insurable risks." *Journal of Economic Theory*, **126(1)**: 31–69.

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

Yen, N. D. 1995. "Lipschitz Continuity of Solutions of Variational Inequalities with a Parametric Polyhedral Constraint." *Mathematics of Operations Research*, **20(3)**: 695–708.

# A Mathematical foundation

In this appendix we provide a mathematical foundation for the model in Section 2 and provide the proofs of the results in Section 4. Recall that $\mathcal{Z}$ is the set of ordered pairs $(y, y')$ of states, where $\lambda_{y,y'}(a) > 0$ for some action profile $a \in \mathcal{A}(y)$. 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 then defined by

$$S_0 = y_0, \qquad S_t := \begin{cases} S_{t-} & \text{if } \Delta J_t^{S_{t-},y} = 0 \text{ for all } y \text{ with } (S_{t-}, y) \in \mathcal{Z}, \\ y & \text{if } \Delta J_t^{S_{t-},y} = 1, \end{cases} \tag{15}$$

i.e., $S$ is a piecewise constant stochastic process that jumps to $y$ at time $t$ if and only if $\Delta J_t^{S_{t-},y} = 1$. When we require existence of a state process $S$ with certain properties in Section 2, the mathematically precise statement is to require existence of independent Poisson processes $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ such that the process $S$ induced by (15) has the desired properties. The family $Q^A = (Q_t^A)_{t\geq 0}$ of probability measures is defined via its density process relative to $P$, given by

$$\frac{\mathrm{d}Q_t^A}{\mathrm{d}P} := \mathcal{E}_t\left(\int_0^\cdot \mu(S_s, A_s)^\top (\sigma^{-1}(S_s))^\top \mathrm{d}Z_s + \sum_{(S_{s-},y)\in\mathcal{Z}(A_{s-})} \int_0^\cdot (\lambda_{S_{s-},y}(A_{s-}) - 1)(\mathrm{d}J_s^{S_{s-},y} - \mathrm{d}s)\right), \tag{16}$$

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 the process $X$. Formally, the filtration $\mathbb{F}$ contains the filtration generated by $Z$ and $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$, and not just the filtration generated by $Z$ and $S$ so that the density process is adapted to $\mathbb{F}$. Note that the instantaneous intensities of the processes $J^{y,y'}$ are equal to 1 under $Q^A$, where $S \neq y$ so that players learn nothing from these processes. A mathematical foundation 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 $P$. It follows from Girsanov's theorem that the public signal under $Q^A$ indeed takes the form (1) and that state transitions occur with instantaneous intensities $\lambda_{S_{t-},y}(A_t)$.

The proofs of Lemmas 4.1 and 4.3 in this appendix are adaptations of the proofs in Bernard (2023) to the setting of stochastic games. Because the public signal and the state process generate the same information as the public signal in Bernard (2023), those adaptations are minor.

*Proof of Lemma 4.1.* Fix a strategy profile $A$ and observe that $W := W(S_0, A)$ is bounded because it takes values in $\mathcal{V}$. Fix a player $i$ and a time $T > 0$, and define the $\mathcal{F}_T$-measurable random variable $w_T^i := W_T^i - r\int_0^T (W_t^i - g^i(S_t, A_t))\,\mathrm{d}t$. Because $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ are pairwise orthogonal and orthogonal to $\int \sigma(S_t)\,\mathrm{d}Z_t$, the stable subspace generated by $\int \sigma(S_t)\,\mathrm{d}Z_t$ and $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ is the space of all stochastic integrals with respect to these processes by Theorem IV.36 in Protter (2005). Therefore, we obtain the unique martingale representation property for square-integrable martingales by Corollary 1 to Theorem IV.37 in Protter (2005). That is, there exist an $\mathcal{F}_0$-measurable $c_T^i$, predictable processes $(\beta_{t,T}^i)_{0\leq t\leq T}$ and $(\delta_{t,T}^i(y, y'))_{0\leq t\leq T}$ for $(y, y') \in \mathcal{Z}$ with $\mathbb{E}_{Q_T^A}\Big[\int_0^T |\beta_{t,T}^i \sigma(S_t)|^2\,\mathrm{d}t\Big] < \infty$ and $\mathbb{E}_{Q_T^A}\Big[\int_0^T |\delta_{t,T}^i(y, y')|^2 \lambda_{y,y'}(A_t)\,\mathrm{d}t\Big] < \infty$ for all $(y, y') \in \mathcal{Z}$ and a $Q_T^A$-martingale $N^i$ orthogonal to $\int \sigma(S_t)\,\mathrm{d}Z_t$ and $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$ with $N_0^i = 0$ such that

$$w_T^i = c_T^i + r\int_0^T \beta_{t,T}^i(\sigma(S_t)\,\mathrm{d}Z_t - \mu(S_t,A_t)\,\mathrm{d}t) + \sum_{(y,y')\in\mathcal{Z}} r\int_0^T \delta_{t,T}^i(y,y')(\mathrm{d}J_t^{y,y'} - \lambda_{y,y'}(A_t)\mathrm{d}t) + N_{T,T}^i.$$

To prove that (5) holds, we need to show that $c_T^i$, $\beta_{t,T}^i$, $\delta_{t,T}^i(y,y')$ and $N_{t,T}^i$ do not depend on $T$. It follows from (2) and Fubini's theorem that

$$w_T^i = W_T^i + r\int_0^T g^i(S_t,A_t)\,\mathrm{d}t - r\int_0^\infty \int_0^{s\wedge T} r\mathrm{e}^{-r(s-t)}\mathbb{E}_{Q_s^A}\Big[g^i(S_s,A_s)\ \Big|\ \mathcal{F}_t\Big]\,\mathrm{d}t\,\mathrm{d}s. \tag{17}$$

Let $\tilde{T} \le T$ and take conditional expectations on $\mathcal{F}_{\tilde{T}}$ under $Q_T^A$ of (17) to deduce that

$$\begin{aligned}\mathbb{E}_{Q_T^A}\Big[w_T^i\,\Big|\,\mathcal{F}_{\tilde{T}}\Big] - w_{\tilde{T}}^i &= \mathbb{E}_{Q_T^A}\Big[W_T^i\,\Big|\,\mathcal{F}_{\tilde{T}}\Big] - W_{\tilde{T}}^i + r\int_{\tilde{T}}^T \mathbb{E}_{Q_t^A}\Big[g^i(S_t,A_t)\,\Big|\,\mathcal{F}_{\tilde{T}}\Big]\,\mathrm{d}t\\ &\quad - r\int_{\tilde{T}}^\infty \int_{\tilde{T}}^{s\wedge T} r\mathrm{e}^{-r(s-t)}\mathbb{E}_{Q_s^A}\Big[g^i(S_s,A_s)\,\Big|\,\mathcal{F}_{\tilde{T}}\Big]\,\mathrm{d}t\,\mathrm{d}s\\ &= \mathbb{E}_{Q_T^A}\Big[W_T^i\,\Big|\,\mathcal{F}_{\tilde{T}}\Big] - W_{\tilde{T}}^i - \int_T^\infty r\mathrm{e}^{-r(s-T)}\mathbb{E}_{Q_s^A}\Big[g^i(S_s,A_s)\,\Big|\,\mathcal{F}_{\tilde{T}}\Big]\,\mathrm{d}s\\ &\quad + \int_{\tilde{T}}^\infty r\mathrm{e}^{-r(s-\tilde{T})}\mathbb{E}_{Q_s^A}\Big[g^i(S_s,A_s)\,\Big|\,\mathcal{F}_{\tilde{T}}\Big]\,\mathrm{d}s\\ &= 0.\end{aligned}$$

Taking $\tilde{T} = 0$, this shows that $c_T^i = W_0^i$ does not depend on $T$. It also implies

$$w_{\tilde{T}}^i = W_0^i + r\int_0^{\tilde{T}} \beta_{t,T}^i(\sigma(S_t)\,\mathrm{d}Z_t - \mu(S_t,A_t)\,\mathrm{d}t) + \sum_{(y,y')\in\mathcal{Z}} r\int_0^{\tilde{T}} \delta_{t,T}^i(y,y')(\mathrm{d}J_t^{y,y'} - \lambda_{y,y'}(A_t)\,\mathrm{d}t) + N_{\tilde{T},T}^i,$$

which yields $\beta_{\cdot,T}^i = \beta_{\cdot,\tilde{T}}^i$ and $\delta_{\cdot,T}^i(y,y') = \delta_{\cdot,\tilde{T}}^i(y,y')$ for every $(y,y') \in \mathcal{Z}$ a.e. on $[0,\tilde{T}]$ and $N_{\tilde{T},T}^i = N_{\tilde{T},\tilde{T}}^i$ a.s. by the uniqueness of the orthogonal decomposition. Taking $\mathcal{F}_t$-conditional expectations, we deduce $N_{t,\tilde{T}}^i = M_{t,T}^i$ a.s. for $t \in [0,\tilde{T}]$, proving that the integral representation is independent of $T$. We thus omit the subscript $T$ and $\tilde{T}$ of the constructed processes $\beta^i$, $\big(\delta^i(y,y')\big)_{(y,y')\in\mathcal{Z}}$, and $N^i$. To arrive at (5), we set

$$M^i = N^i + \sum_{(y,y')\in\mathcal{Z}} r\int_0^{\cdot} \delta_t^i(y,y')1_{\{S_{t-}\neq y\}}(\mathrm{d}J_t^{y,y'} - 1\,\mathrm{d}t)$$

and $\delta^i(y) := \delta^i(S_-,y)1_{\{(S_-,y)\in\mathcal{Z}\}}$ for any $y \in \mathcal{Y}$. Because $N^i$ is orthogonal of $Z$ and $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$, and the processes $J^{y,y'}$ are orthogonal to each other and of $Z$, it follows that $M^i$ is a martingale orthogonal to $\int \sigma(S_t)\,\mathrm{d}Z_t$ and to $J^y = \sum_{0<s\le t}\Delta J_t^{S_{s-},y}$ for every $y \in \mathcal{Y}$. Since the processes $\delta^i(y,y')$ are square-integrable by construction, so is $\delta^i(y)$. Moreover, $\delta^i(y)$ is predictable because $\delta^i(y,y')$

and $S_-$ are. By construction,

$$M^i + \sum_{y \in \mathcal{Y}} r \int_0^{\cdot} \delta^i(y)\big(\mathrm{d}J_t^y - \lambda_{S_{t-},y}(A_t)\,\mathrm{d}t\big) = N^i + \sum_{(y,y') \in \mathcal{Z}} r \int_0^{\tilde{T}} \delta^i_{t,T}(y,y')\big(\mathrm{d}J_t^{y,y'} - \lambda_{y,y'}(A_t)\,\mathrm{d}t\big),$$

hence $W$ satisfies (5) for processes $\beta^i$, $\big(\delta^i(y)\big)_{y \in \mathcal{Y}}$, and $M^i$.

To show the converse, we derive from Itō's formula that

$$\begin{aligned} \mathrm{d}(\mathrm{e}^{-rt} W_t^i) &= -r\mathrm{e}^{-rt} g^i(S_t, A_t)\,\mathrm{d}t + r\mathrm{e}^{-rt}\beta_t^i\left(\sigma(S_t)\,\mathrm{d}Z_t - \mu(S_t,A_t)\,\mathrm{d}t\right) \\ &\quad + r\mathrm{e}^{-rt}\sum_{y\in\mathcal{Y}} \delta_t^i(y)\big(\mathrm{d}J_t^y - \lambda_{S_{t-},y}(A_t)\,\mathrm{d}t\big) + \mathrm{e}^{-rt}\,\mathrm{d}M_t^i. \end{aligned} \tag{18}$$

Since $M^i$ is strongly orthogonal to $\int \sigma(S_t)\,\mathrm{d}Z_t$ and $(J^y)_{y\in\mathcal{Y}}$, it is also strongly orthogonal to the density process given in (16). Therefore, $M^i$ is a martingale also under $Q^A$. Integrating (18) from $t$ to $T$ and taking $Q_T^A$-conditional expectations on $\mathcal{F}_t$ thus yields

$$W_t^i = \int_t^T r\mathrm{e}^{-r(s-t)}\mathbb{E}_{Q_s^A}\Big[g^i(S_s,A_s)\,\Big|\,\mathcal{F}_t\Big]\,\mathrm{d}s + \mathrm{e}^{-r(T-t)}\mathbb{E}_{Q_T^A}\Big[W_T^i\,\Big|\,\mathcal{F}_t\Big].$$

Since $W$ is bounded, the second summand converges to zero a.s. as $T$ tends to $\infty$, hence $W_t^i$ is indeed $i$'s continuation value under strategy profile $A$ in state $S_t$. $\square$

*Proof of Lemma 4.3.* Fix a strategy profile $A$ and let $\tilde{A}$ be a strategy profile involving a unilateral deviation of some player $i$, that is, $\tilde{A}^{-i} = A^{-i}$ a.e. For $(\beta,\delta)$ related to $W(S,A)$ by (5), integrating (18) from $t$ to $u$ yields

$$\begin{aligned} W_t^i(S_t,A) &= -\int_t^u \mathrm{e}^{-r(s-t)}\Big(\beta_s^i\left(\sigma(S_s)\,\mathrm{d}Z_s - \mu(S_s,A_s)\,\mathrm{d}s\right) - g^i(S_s,A_s)\,\mathrm{d}s - \mathrm{d}M_s^i\Big) \\ &\quad - \sum_{y\in\mathcal{Y}}\int_t^u \mathrm{e}^{-r(s-t)}\delta_s^i(y)\left(\mathrm{d}J_s^y - \lambda_{S_{s-},y}(A_s)\,\mathrm{d}s\right) + \mathrm{e}^{-r(u-t)}W_u^i(S_u,A). \end{aligned}$$

Note that the term $\mathrm{e}^{-r(u-t)}W_u^i(S_u,A)$ vanishes as we let $u\to\infty$ because $W(S,A)$ is in the bounded set $\mathcal{V}$. Since $M$ is a martingale up to time $u$ also under $Q_u^{\tilde{A}}$, taking conditional expectations yields

$$\begin{aligned} W_t^i(S_t,\tilde{A}) &= \lim_{u\to\infty}\mathbb{E}_{Q_u^{\tilde{A}}}\left[\int_t^u r\mathrm{e}^{-r(s-t)}g^i(S_s,\tilde{A}_s)\,\mathrm{d}s\,\middle|\,\mathcal{F}_t\right] \\ &= W_t^i(S_t,A) + \lim_{u\to\infty}\mathbb{E}_{Q_u^{\tilde{A}}}\left[\int_t^u r\mathrm{e}^{-r(s-t)}\Big(\big(g^i(S_s,\tilde{A}_s) - g^i(S_s,A_s)\big)\,\mathrm{d}s\right. \\ &\quad \left. + \beta_s^i\left(\sigma(S_s)\,\mathrm{d}Z_s - \mu(S_s,A_s)\,\mathrm{d}s\right) + \sum_{y\in\mathcal{Y}}\delta_s^i(y)\left(\mathrm{d}J_s^y - \lambda_{S_{s-},y}(A_s)\,\mathrm{d}s\right)\Big)\middle|\,\mathcal{F}_t\right] \text{ a.s.} \end{aligned}$$

Note that the state process $S$ is the same process under strategy profile $A$ and $\tilde{A}$ (as a map from $\Omega\times[0,\infty)$ to $\mathcal{Y}$), but it has a different distribution under $Q_u^A$ and $Q_u^{\tilde{A}}$. Because $\beta$ is constructed using a martingale representation result for the bounded random variable $w_T^i$ in (17), the process

$\int_t^\cdot r\mathrm{e}^{-r(s-t)}\beta_s^i\left(\sigma(S_s)\,\mathrm{d}Z_s-\mu(S_s,A_s)\,\mathrm{d}s\right)$ is a bounded mean oscillation (BMO) martingale under the probability measure $Q_u^A$ up to any time $u>t$. It follows from Theorem 3.6 in Kazamaki (2006) that $\int_t^\cdot r\mathrm{e}^{-r(s-t)}\beta_s^i\left(\sigma(S_s)\,\mathrm{d}Z_s-\mu(S_s,\tilde{A}_s)\,\mathrm{d}s\right)$ is a martingale under $Q_u^{\tilde{A}}$. Since $W(A)$ lies in $\mathcal{V}$, each $\delta(y)$ is uniformly bounded $P$-a.s., hence also $Q_u^{\tilde{A}}$-a.s. for any $u>t$. The lemma after Theorem IV.29 in Protter (2005) thus implies that $\int_t^\cdot r\mathrm{e}^{-r(s-t)}\delta_s^i(y)\left(\mathrm{d}J_s^y-\lambda_{S_{s-},y}(\tilde{A}_s)\,\mathrm{d}s\right)$ is a $Q_u^{\tilde{A}}$-martingale up to any time $u>t$. Together with Fubini's theorem, this implies

$$W_t^i(S_t,\tilde{A})-W_t^i(S_t,A)=\int_t^\infty \mathrm{e}^{-r(s-t)}\mathbb{E}_{Q_s^{\tilde{A}}}\Big[g^i(S_s,\tilde{A}_s)-g^i(S_s,A_s)+\beta_s^i\left(\mu(S_s,\tilde{A}_s)\right. \qquad (19)$$

$$\left.-\mu(S_s,A_s)\right)+\delta_s^i\left(\lambda(S_s,\tilde{A}_s)-\lambda(S_s,A_s)\right)\Big|\,\mathcal{F}_t\Big]\,\mathrm{d}s\quad\text{a.s.}$$

If $(\beta,\delta)$ enforces $A$ in state $S$, the above conditional expectation is non-positive, hence $A$ is a PPE. To show the converse, assume towards a contradiction that there exist a player $i$ and a set $\Xi\subseteq\Omega\times[0,\infty)$ with $P\otimes \mathit{Lebesgue}(\Xi)>0$, such that some strategy $\hat{A}^i$ satisfies

$$g^i(S,\hat{A})-g^i(S,A)+\beta^i\left(\mu(S,\hat{A})-\mu(S,A)\right)+\delta^i\left(\lambda(S,\hat{A})-\lambda(S,A)\right)>0$$

on the set $\Xi$, where we denoted $\hat{A}=(\hat{A}^i,A^{-i})$ for the sake of brevity. Because $\beta$ and $\delta$ are predictable, we can and do choose $\Xi$ predictable as well. Thus, $\tilde{A}^i:=\hat{A}^i 1_\Xi+A^i 1_{\Xi^c}$ is predictable and, in particular, a strategy for player $i$. For $\tilde{A}=(\tilde{A}^i,A^{-i})$, the expectation in (19) is strictly positive for $t=0$, which is a contradiction. □

## B Concatenations of solutions and convergence of algorithm

The main idea behind the proof of Lemma 4.8 is the following. Any payoff pair in $\mathcal{B}_y(\mathcal{W})$ can be attained by an en enforceable solution to (5) that remains in $\mathcal{B}_y(\mathcal{W})$ until the time $\rho$ of the first state transition and $W_\rho\in\mathcal{W}_{\mathcal{S}_\rho}$. Since $\mathcal{W}_{y'}\subseteq\mathcal{B}_{y'}(\mathcal{W})$, we would like to attain $W_\rho$ by another enforceable solution $W$ to (5) until the next state transition and then concatenate the two solutions. Such a concatenation, however, is subject to some subtle measurability issues.

Without restrictions on $\beta$ and $\delta$, solutions to (5) are weak solutions, that is, the Brownian motion, the Poisson processes, and the entire probability space are part of the solution. We refer to $(\Omega,\mathcal{F},\mathbb{F},P,Z,(J^{y,y'})_{(y,y')\in\mathcal{Z}})$ as the *stochastic framework* for a solution of (5). Thus, more formally,

$$\mathcal{B}_y(\mathcal{W})=\left\{w\in\mathcal{V}\;\middle|\;\begin{array}{l}\text{There exists a stochastic framework }(\Omega,\mathcal{F},\mathbb{F},P,Z,(J^{y,y'})_{(y,y')\in\mathcal{Z}})\\ \text{containing a solution }(W,S,A,\beta,\delta,M)\text{ to (5) on }[\![0,\rho]\!]\text{ with }W_0=w,\\ S_0=y,\text{ and }W_\rho\in\mathcal{W}_{S_\rho}\ P\text{-a.s., such that on }[\![0,\rho)\!),W\in\mathcal{B}(\mathcal{W})\text{ and}\\ (\beta,\delta)\text{ enforces }A\text{ in }S,\text{where }\rho\text{ is the first jump time of }(J^{y,y'})_{(y,y')\in\mathcal{Z}}.\end{array}\right\}.^{17,18}$$

[17]For two stopping times $\rho$ and $\tau$, the set $[\![\rho,\tau)\!):=\{(\omega,t)\in\Omega\times[0,\infty)\mid\rho(\omega)\le t<\tau(\omega)\}$ is called the (left-closed, right-open) stochastic interval from $\rho$ to $\tau$. Closed, open, and left-open, right-closed stochastic intervals are defined analogously.

[18]We say that a stochastic process $X$ satisfies a certain property on $\Xi\subseteq\Omega\times[0,\infty)$ if $X_t(\omega)$ satisfies that property

When we aim to concatenate solutions at time $\rho$, the continuation solutions may live on separate probability spaces for each realization of $W_\rho$. It is thus not a priori clear that there exists a probability space that contains the entire concatenation. The following result, adapted from Lemma A.1 in Bernard (2023), establishes that this is indeed possible.

**Lemma B.1.** *Let $X$ be a $\mathcal{V}$-valued random variable with distribution $\nu$. The following are equivalent:*

*(a)* $X \in \mathcal{B}_y(\mathcal{W})$ *$\nu$-a.s.,*

*(b) There exists a solution $(W, S, A, \beta, \delta, M)$ to (5) in a framework $(\Omega, \mathcal{F}, \mathbb{F}, P, Z, (J^{y,y'})_{(y,y')\in\mathcal{Z}})$ such that $X$ is $\mathcal{F}_0$-measurable with $\nu = P \circ X^{-1}$, $W_0 = X$ $P$-a.s., $S_0 = y$ $P$-a.s., $W_\rho \in \mathcal{W}_{S_\rho}$ $P$-a.s., and on $[\![0, \rho))$, we have $W \in \mathcal{B}_y(\mathcal{W})$ and $(\beta, \delta)$ enforces $A$ in state $S$, where $\rho$ is the first jump time of any of the processes $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$.*

Crucial in the proof are (a) that the path space of the solutions coincide for different realizations of $X$ and (b) that the path space (with the Skorohod metric) is complete and separable. Then the different probability spaces can be aggregated with a regular conditional probability. Neither of these properties depend on the fact that we are studying stochastic games in this paper, hence the proof of Lemma A.1 in Bernard (2023) goes through.

With the help of Lemma B.1, we can formalize the concatenation procedure.

**Lemma B.2.** *Fix a state $y$, a payoff set $\mathcal{X}_y$, and a family of payoff sets $\mathcal{W} = (\mathcal{W}_{y'})_{y'\in\mathcal{Y}}$ such that for every $w \in \mathcal{X}_y$, there exists a stochastic framework $(\Omega, \mathcal{F}, \mathbb{F}, P, Z, (J^{y,y'})_{(y,y')\in\mathcal{Z}})$ containing a solution $(W, S, A, \beta, \delta, M)$ to (5) for $S$ defined in (15) on the stochastic interval $[\![0, \rho \wedge \tau]\!]$, where $\rho$ is the first jump time of $(J^{y,y'})_{(y,y')\in\mathcal{Z}}$, $\tau$ is an $\mathbb{F}$-stopping time, such that:*

*(i) $W_0 = w$ and $S_0 = y$ $P$-a.s.,*

*(ii) $W \in \mathcal{X}_y$, $(\beta, \delta)$ enforces $A$ in $S$, and $W + r\delta(y') \in \mathcal{W}_{y'}$ for each $(y, y') \in \mathcal{Z}(A)$ on $[\![0, \rho \wedge \tau))$,*

*(iii) On $\{\tau < \rho\}$, we have $W_\tau \in \mathcal{B}_y(\mathcal{W})$.*

*Then for every $w \in \mathcal{X}_y$, there exists a stochastic framework $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P}, \hat{Z}, (\hat{J}^{y,y'})_{y,y'\in\mathcal{Z}})$ containing a solution $(\hat{W}, \hat{S}, \hat{A}, \hat{\beta}, \hat{\delta}, \hat{M})$ to (5) that coincides with $(W, S, A, \beta, \delta, M)$ on $[\![0, \tau \wedge \rho))$, such that $(\beta, \delta)$ enforces $A$ in state $S$ on $[\![0, \hat{\rho}))$, $W \in \mathcal{B}_{S_\tau}(\mathcal{W})$ on $[\![\tau, \hat{\rho}))$ and $W_{\hat{\rho}} \in \mathcal{W}_{S_{\hat{\rho}}}$ $P$-a.s., where $\hat{\rho}$ is the first jump time of any of the processes $(\hat{J}^{y,y'})_{y,y'\in\mathcal{Z}}$. In particular, $\mathcal{X}_y \subseteq \mathcal{B}_y(\mathcal{W})$.*

*Proof.* Fix $w \in \mathcal{X}$, a stochastic framework $(\Omega, \mathcal{F}, \mathbb{F}, P, Z, (J^{y,y'})_{(y,y')\in\mathcal{Z}})$, an $\mathbb{F}$-stopping time $\tau$, and a solution $(W, A, \beta, \delta, M)$ to (5) on $[\![0, \tau]\!]$ with all the stated properties. By Lemma B.1, on $\{\tau < \rho\}$ there exists a stochastic framework $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P}, \tilde{Z}, (\tilde{J}^{y,y'})_{(y,y')\in\mathcal{Z}})$ containing a $\mathcal{W}$-enforceable solution $(\tilde{W}, \tilde{S}, \tilde{A}, \tilde{\beta}, \tilde{\delta}, \tilde{M})$ to (5) such that $\tilde{W}_0$ is distributed under $\tilde{P}$ as $W_\tau$ under $P$. Without loss of generality we can let $\Omega$ and $\tilde{\Omega}$ be the Polish path spaces of the solutions so that there exists a regular conditional probability $\hat{P}$ on $\hat{\Omega} = \Omega \times \tilde{\Omega}$ with marginal $P$ on $\Omega$. Let $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{t\geq 0}$ denote the $\hat{P}$-augmented filtration such that $\hat{\mathcal{F}}_t$ contains $\mathcal{F}_{t\wedge\tau} \vee \tilde{\mathcal{F}}_{t-\tau\vee 0}$. Define the processes $\hat{Z}$ and $\hat{J}^{y,y'}$ by setting

$$\hat{Z}_t = Z_{t\wedge\tau} + \tilde{Z}_{t-\tau} 1_{\{t>\tau\}}, \qquad \hat{J}^{y,y'}_t = J^{y,y'}_{t\wedge\tau} + \tilde{J}^{y,y'}_{t-\tau} 1_{\{t>\tau\}}.$$

for $P \otimes$ *Lebesgue*-almost every $(\omega, t) \in \Xi$.

Because Brownian motion and Poisson processes have independent and identically distributed increments, $\hat{Z}$ is an $\hat{\mathbb{F}}$-Brownian motion and $\hat{J}^{y,y'}$ for any $y \in Y$ is an $\hat{\mathbb{F}}$-Poisson process. Define the concatenated processes $\hat{W}$, $\hat{S}$, $\hat{A}$, $\hat{\beta}$, $\hat{\delta}$, and $\hat{M}$ by setting

$$\hat{W} := (W 1_{[\![0,\tau]\!]} + \tilde{W}_{\cdot-\tau} 1_{(\!(\tau,\infty)\!)}) 1_{\{\tau<\rho\}} + W 1_{\{\tau\geq\rho\}}$$

and similarly for $\hat{A}$, $\hat{S}$, $\hat{\beta}$, $\hat{\delta}$, and $\hat{M}$. By construction, $\hat{S}$ satisfies (15) for $\hat{A}$ and $(\hat{W}, \hat{S}, \hat{A}, \hat{\beta}, \hat{\delta}, \hat{M})$ coincides with $(W, S, A, \beta, \delta, M)$ on $[\![0, \tau \wedge \rho)\!)$. By the properties of a regular conditional probability, $\hat{W}_0 = w$ and $\hat{W}_{\hat{\rho}} \in \mathcal{W}_{S_{\hat{\rho}}}$ $\hat{P}$-a.s. Moreover, $(\hat{W}, \hat{A}, \hat{\beta}, \hat{\delta}, \hat{M})$ has all the desired properties on $[\![0, \hat{\rho})\!)$. □

*Proof of Lemma 4.8.* Suppose first that $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ for each state $y$. Because Poisson processes have only countably many jumps, an iteration of the concatenation procedure in Lemma B.2 yields an enforceable solution to (5) that remains in $\mathcal{B}(\mathcal{W})$ forever, showing that $\mathcal{B}(\mathcal{W})$ is self-generating. If $\mathcal{W}$ is self-generating, then for any state $y \in \mathcal{Y}$ and any $w \in \mathcal{W}_y$, there exists an enforceable solution $W$ to (5) that is in $\mathcal{W}_S$ a.e. In particular, $W_\rho \in \mathcal{W}_{S_\rho}$ a.s. It follows that $\mathcal{W}_y \subseteq \mathcal{B}_y(\mathcal{W})$ by maximality of $\mathcal{B}_y(\mathcal{W})$. Since the state was arbitrary the result follows. □

As a last step in preparation for the proof of Proposition 4.9, we show monotonicity of $\mathcal{B}_y$.

**Lemma B.3.** *Let $\mathcal{W}_y \subseteq \mathcal{W}'_y$ for every $y \in \mathcal{Y}$. Then $\mathcal{B}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W}')$ for every $y \in \mathcal{Y}$.*

*Proof.* Fix any state $y \in \mathcal{Y}$ and any payoff pair $w \in \mathcal{B}_y(\mathcal{W})$. By definition, there exists a solution $(W, A, \beta, \delta, M)$ to (5) for $S$ defined in (15) with initial state $y$ such that $W_0 = w$ a.s., $W_\rho \in \mathcal{W}_{S_\rho}$ a.s., and on $[\![0, \rho)\!)$, $(\beta, \delta)$ enforces $A$ in state $y$ and $W \in \mathcal{B}_y(\mathcal{W})$. Since this implies that also $W_\rho \in \mathcal{W}'_{S_\rho}$ a.s., it follows that $\mathcal{B}(\mathcal{W})$ is $\mathcal{W}'$-relaxed self-generating. In particular, $\mathcal{B}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W}')$ by maximality of $\mathcal{B}(\mathcal{W}')$. □

*Proof of Proposition 4.9.* Lemma 4.8 implies that $\mathcal{E}_y(r) \subseteq \mathcal{B}_y(\mathcal{E}(r))$ for every state $y \in \mathcal{Y}$ and that $\mathcal{B}(\mathcal{E}(r))$ is self-generating. Since $\mathcal{E}(r)$ is the largest family of bounded self-generating sets, it follows that $\mathcal{B}_y(\mathcal{E}(r)) \subseteq \mathcal{E}_y(r)$ for every $y \in \mathcal{Y}$ and hence $\mathcal{B}(\mathcal{E}(r)) = \mathcal{E}(r)$. Since $\mathcal{E}_y(r) \subseteq \mathcal{V} = \mathcal{W}_{0,y}$ for every $y \in \mathcal{Y}$, monotonicity of $\mathcal{B}$ implies that $\mathcal{E}_y(r) \subseteq \mathcal{W}_{1,y} \subseteq \mathcal{W}_{0,y}$. An iterated application of Lemma B.3 thus shows that $(W_{n,y})_{n\geq 0}$ is decreasing in the set-inclusion sense and that it is bounded from below by $\mathcal{E}_y(r)$ for any $y \in \mathcal{Y}$. Those sequences must converge to a limit $\mathcal{W}_\infty$ with $W_{\infty,y} \supseteq \mathcal{E}_y(r)$ for every $y \in \mathcal{Y}$ such that $\mathcal{B}_y(\mathcal{W}_\infty) \subseteq \mathcal{W}_{\infty,y}$. The last step of the proof is to show that $\mathcal{W}_{\infty,y} \subseteq \mathcal{B}(\mathcal{W}_{\infty,y})$ for each state $y$ so that $\mathcal{W}_\infty = \mathcal{B}(\mathcal{W}_\infty)$ is self-generating by Lemma 4.6.

To that end, fix a state $y$ and a payoff pair $w \in \mathcal{W}_{\infty,y}$. Since $w \in \mathcal{W}_{n,y}$ for each $n$, there exist enforceable solutions $(W^n, S^n, A^n, \beta^n, \delta^n, M^n)$ to (5) on some stochastic basis $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, Z^n, J^n, P^n)$ such that $W^n \in \mathcal{W}_{n,y}$ a.e. on $[\![0, \rho)\!)$ and $W^n_\rho \in \mathcal{W}_{n-1,S_\rho}$ a.s. Through the use of public randomization, we may choose $W^n$ that remains on the boundary of $\partial\mathcal{W}_{n,y}$ on $[\![0, \rho)\!)$. Moreover, we may choose such $M^n$ with finite variation. Because $\beta^n$ is given by the maximizer of the Lipschitz continuous optimality equation (14) where $\beta^n \neq 0$, there must exist a constant $K > 0$ such that $\|\beta^n \sigma(y)\| \leq K$, i.e., the quadratic variation of the continuation value is uniformly bounded. We will show that the solutions converge in law along a subsequence to a limit $(W, S, A, \beta, \delta, M)$ that solves (5). Convergence in law then implies that the limit has the remaining desired properties. Indeed, if there exists a set $\Xi \subseteq [\![0, \rho)\!)$ of positive $P \otimes Lebesgue$-measure on which $W \notin \mathcal{W}_{\infty,y}$, then closedness of $\mathcal{W}_{\infty,y}$ implies that there exists $n$ sufficiently large, for which $W^n \notin \mathcal{W}_{n,y}$ with positive measure, a contradiction. Similarly, convergence in law implies $W_\rho \in \mathcal{W}_{\infty,S_\rho}$ a.s., and that $(\beta, \delta)$ enforce $A$ a.e. on $[\![0, \rho)\!)$.

Convergence of $(W^n)_{n\geq 0}$ in law along a subsequence is shown in two steps. First, we show that the sequence is tight by verifying that the associated semimartingale characteristics satisfy the sufficient conditions of Theorem VI.5.17 in Jacod and Shiryaev (2002). Second, Prokhorov's theorem states that any tight sequence converges in law along a subsequence.[19] We show convergence in law of $\tilde{W}^n := (e^{-rt} W^n_t)_{t\geq 0}$ because its semimartingale characteristics or more easily computed. Convergence of $(\tilde{W}^n)_{n\geq 0}$ in law then implies convergence in law of $(W_n)_{n\geq 0}$. Since each $\tilde{W}^n$ is uniformly bounded under $Q_{A^n}$, it is a so-called special semimartingale, admitting a unique decomposition $\tilde{W}^n = W^n_0 + B^n + \tilde{M}^n$ into a predictable process $B^n$ of finite variation and a $Q_{A^n}$-local martingale $\tilde{M}^n$. The semimartingale characteristics of $\tilde{W}^n$ under $Q_{A^n}$ is the triplet $(B^n, C^n, \nu^n)$, where $C^n$ is the predictable quadratic variation of $\tilde{M}^n$ and $\nu^n$ is the compensated jump measure of $\tilde{W}^n$; see Chapter II of Jacod and Shiryaev (2002) for a detailed introduction.[20] It follows from (18) that the first two characteristics of $\tilde{W}^n$ are given by

$$B^n = -\int_0^{\cdot} r\mathrm{e}^{-rt} g(y, A^n_t)\, \mathrm{d}t, \qquad\qquad C^{n,ij} = \int_0^{\cdot} r^2 \mathrm{e}^{-2rt} \sigma(y)^\top \beta^{n,i\top}_t \beta^{n,j}_t \sigma(y)\, \mathrm{d}t.$$

Define the auxiliary processes $G^n := \mathrm{Var}(B^{n,1}) + \mathrm{Var}(B^{n,2}) + C^{n,11} + C^{n,22}$ for each $n$. Since we chose $\beta^n \sigma(y)$ uniformly bounded by $K$, it follows that each $G^n$ is majorized by

$$F := \int_0^{\cdot} r\mathrm{e}^{-rs} \max_{y\in\mathcal{Y}} \max_{a\in\mathcal{A}} \left| g^1(y,a) + g^2(y,a) \right| \mathrm{d}s + 2\int_0^{\cdot} r^2 \mathrm{e}^{-2rs} K^2\, \mathrm{d}s.^{21}$$

Aldous' criterion implies that $(G_n)_{n\geq 0}$ is tight; see Theorem VI.4.5 in Jacod and Shiryaev (2002). Thus, the sequence $(G_n)_{n\geq 0}$ converges in law along a subsequence to some limit process $G$ by Prokhorov's theorem. Since $F$ is deterministic, $F$ majorizes $G$. Together with the fact that $\left(\tilde{W}^n_0\right)_{n\geq 0} = (M^n_0)_{n\geq 0}$ is tight because each $M^n_0$ takes values in the compact set $\mathcal{V}$, this shows that $\left(\tilde{W}^n\right)_{n\geq 0}$ is tight by virtue of Theorem VI.5.17 in Jacod and Shiryaev (2002).

With the same argument and majorizing process $F$, it follows that $(B_n)_{n\geq 0}$ and $\left(\tilde{M}^n\right)_{n\geq 0}$ are tight. Since the path space $\mathcal{A}^{[0,\infty)}$ of any $A^n$ is compact by Tychonov's theorem, the sequence $(A_n)_{n\geq 0}$ is uniformly tight. By Prokhorov's theorem, there exists a subsequence $(n_k)_{k\geq 0}$, along which $(\tilde{W}^n)_{n\geq 0}$, $(A_n)_{n\geq 0}$, $(B_n)_{n\geq 0}$, and $(\tilde{M}^n)_{n\geq 0}$ all converge to limits $\tilde{W}$, $A$, $B$, and $\tilde{M}$, respectively. Convergence in law implies that $-\int_0^{\cdot} r\mathrm{e}^{-rt} g(y, A_t)\, \mathrm{d}t = \tilde{B}$. After a suitable transformation with Girsanov's theorem, the laws of $(Z_n)_{n\geq 0}$ and $(J^{y,y',n})_{n\geq 0}$ are constant, hence they converge trivially

[19]See, for example, Theorem VI.3.5 in Jacod and Shiryaev (2002).

[20]Because each $\tilde{W}^n$ is bounded, we do not need to truncate its jumps to compute the semimartingale characteristics. In the notation of Jacod and Shiryaev (2002), we can choose "truncation function" $h(x) = x$, hence the second modified characteristic coincides with the second characteristic.

[21]A process $F$ majorizes $G$ if $F - G$ is increasing.

to a Brownian motion $Z$ and Poisson processes $J^{y,y'}$ along $(n_k)_{k\geq 0}$. By Proposition IX.1.1 in Jacod and Shiryaev (2002), the limit process $\tilde{M}$ of $(\tilde{M}^n)_{n\geq 0}$ is an $(\mathbb{F}, Q^A)$-martingale. Let $\beta$ and $\delta$ be defined by the martingale representation of $\tilde{M}$. By Theorem VI.6.26 in Jacod and Shiryaev (2002), $[\tilde{M}^{n_k}, \tilde{M}^{n_k}] \to [\tilde{M}, \tilde{M}]$ in law, showing that $\beta^{n_k} \to \beta$ and $\delta^{n_k} \to \delta$ in law as well. This concludes the proof that any payoff pair $w \in \mathcal{W}_{\infty,y}$ can be attained by an enforceable solution to (5) that remains in $\mathcal{W}_{\infty,y}$ up until time $\rho$ with $W_\rho \in \mathcal{W}_{\infty,S_\rho}$. Thus, the family $\mathcal{W}_\infty$ is $\mathcal{W}_\infty$-relaxed self-generating, hence it is contained in the family $\mathcal{B}(\mathcal{W}_\infty)$ by maximality. □

## C Stationary Markov and state-order dependent equilibria

We begin with the proofs of the results characterizing state-order dependent PPE.

*Proof of Lemma 5.2.* 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 exist $a_0 \in \mathcal{A}(y_0)$ and $\delta_0$ such that $(0, \delta_0)$ enforces $a_0$, $w_0 = g(y_0, a_0) + \delta_0\lambda(y_0, a_0)$, and $w_0 + r\delta_0(y) \in \mathcal{W}_y$ for every successor state $y$ of $y_0$. A solution to (5) for $A \equiv a_0$, $\beta \equiv 0$, $\delta \equiv \delta_0$, and $M \equiv 0$ starting at $w_0$ thus remains in $w_0$ until the first time $\rho_1$, when a state transition occurs. It follows from the choice of $\delta_0$ that $W_{\rho_1} \in \mathcal{W}_{S_{\rho_1}}$. Note that $w_0$, $\delta_0$, and $a_0$ can be viewed as a function from the initial state. Therefore, $W_{\rho_1}$ is fully determined by $w_0$, $\delta_0$, and the state $S_\rho$, that is, $W_\rho$ is a function of $(y_0, S_{\rho_1})$. Since $W_{\rho_1} \in \mathcal{W}_{S_{\rho_1}}$, by definition of mutual stationarity, we can decompose $W_{\rho_1}$ by $a_1$ and $\delta_1$, which are functions of $W_{\rho_1}$, hence functions of $(y_0, S_{\rho_1})$. We can thus get an enforceable solution to (5) on $[\![\rho_1, \rho_2)\!)$ that remains in $W_{\rho_1}$ until $\rho_2$ and jumps to $\mathcal{W}_{\rho_2}$ at time $\rho_2$. Similarly to the proof of Lemma B.2, we can concatenate the solutions at state transitions. 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 $a_k$ is a function of $(y_0, S_{\rho_1}, \ldots, S_{\rho_k})$.

Since the union of two families of mutually stationary payoff sets is again mutually stationary, this shows that $\mathcal{E}^S(r)$ contains the largest bounded family of mutually stationary payoff sets. Thus, it remains to show that $\mathcal{E}^S(r)$ is mutually stationary.

Fix an initial state $y_0$ and a payoff pair $w_0 \in \mathcal{E}^S_{y_0}(r)$. Let $A$ be a state-order dependent PPE attaining $w_0$ for initial state $y_0$, corresponding to selector $a_*$. Let $\tau$ be a stopping with $\tau \leq \rho_1$ a.s. It follows along the same lines as in the proof of Proposition 5.6 that $W_\tau(A) = w_0$. Since $\tau$ was arbitrary, $W$ is locally constant on $[\![0, \rho)\!)$. Lemma 4.1 thus implies that $\beta \equiv 0$, $M \equiv 0$, and $\delta \equiv \delta_0$ for some $\delta_0$ such that

$$w_0 = g(y, a_*(y_0)) + \sum_{y' \in \mathcal{Y}(y_0)} \delta_0(y')\lambda_{y_0,y'}(a_*(y_0)).$$

State-order dependence implies that also the continuation profile after the first state transition is state-order dependent and hence $W_{\rho_1} \in \mathcal{E}^S_y(r)$ on the event $\{S_{\rho_1} = y\}$ for all states $y$, for which $\{S_{\rho_1} = y\}$ has positive measure, i.e., for all $y$ such that $(y_0, y) \in \mathcal{Z}$. The SDE representation of the continuation value implies that $W_{\rho_1} - w_0 = r\delta_0(y)$ on $\{S_{\rho_1} = y\}$ for all successor states $y$ of $y_0$ and hence $w_0 + r\delta_0(y) \in \mathcal{E}^S_y(r)$. Finally, since $A$ is a PPE, Lemma 4.3 implies that $(0, \delta_0)$ enforces $a_*(y_0)$. $\square$

*Proof of Lemma 5.4.* Since $\mathcal{S}(\mathcal{W})$ is $\mathcal{W}$-stationary, for each $y \in \mathcal{Y}$ and each $w \in \mathcal{S}_y(\mathcal{W})$, there exist $a \in \mathcal{A}(y)$ and $\delta$ such that $(0, \delta)$ enforces $a$ in state $y$, $w = g(y, a) + \delta\lambda(y, a)$, and $w + r\delta(y') \in \mathcal{W}_{y'}$ for every $(y, y') \in \mathcal{Z}$. If $\mathcal{W}_{y'} \subseteq \mathcal{S}_{y'}(\mathcal{W})$ for every $(y, y') \in \mathcal{Z}$, then $\mathcal{S}(\mathcal{W})$ is also $\mathcal{S}(\mathcal{W})$-stationary and hence mutually stationary. $\square$

*Proof of Proposition 5.5.* This proof mimics the proof of Proposition 4.9. We first note that $\mathcal{S}$ is monotone by definition, that is, $\mathcal{S}_y(\mathcal{W}) \subseteq \mathcal{S}_{y'}(\mathcal{W}')$ for any two families $\mathcal{W}, \mathcal{W}'$ of payoff sets with $\mathcal{W}_y \subseteq \mathcal{W}'_y$ for every state $y$. Note that $\mathcal{S}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W}) \subseteq \mathcal{V}^*$ for any family $\mathcal{W}$ of feasible payoff sets, hence $\mathcal{W}_{1,y} \subseteq \mathcal{W}_{0,y}$. Since $\mathcal{E}^S_y(r) \subseteq \mathcal{V}^*$, monotonicity of $\mathcal{S}_y$ implies that $\mathcal{E}^S_y(r) \subseteq \mathcal{W}_{1,y} \subseteq \mathcal{W}_{0,y} = \mathcal{V}^*$ for every state $y$. An iterated application of $\mathcal{S}$ thus yields a family of payoff sets that is decreasing in the set-inclusion sense. Each member of the family must thus converge to a limit set $\mathcal{W}_{\infty,y} = \bigcap_{n\geq 0} \mathcal{W}_{n,y}$ such that $\mathcal{W}_{\infty,y} \supseteq \mathcal{E}^S_y(r)$ and $\mathcal{S}(\mathcal{W}_\infty) \subseteq \mathcal{W}_\infty$. As in the proof of Proposition 4.9 we deduce that also $\mathcal{W}_\infty \subseteq \mathcal{S}(\mathcal{W}_\infty)$. Thus, Lemma 5.4 implies that $\mathcal{W}_\infty$ is mutually stationary, hence $\mathcal{W}_\infty = \mathcal{E}^S(r)$. $\square$

*Proof of Proposition 5.6.* Suppose first that PPE $A = a_*(S_-)$ is a stationary Markov perfect equilibrium for initial state $y_0$. For any state $y \in \mathcal{Y}(y_0)$, let $S^y$ denote the state process defined in (15) with initial state $y$ and set $w_y := W_0(y, a_*(S^y_-))$. Fix an arbitrary stopping time $\tau$ and define the processes $J^{\tau,y,y'} := J^{y,y'}_{\cdot+\tau} - J^{y,y'}_\tau$ for every $(y, y') \in \mathcal{Z}$. Let $S^\tau$ be defined as in (15) from processes $(J^{\tau,y,y'})_{y,y'\in\mathcal{Z}}$ with initial state $S_\tau$. Since $J^{y,y'}$ is a Lévy process for every $(y, y') \in \mathcal{Z}$, the process $J^{\tau,y,y'}$ 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

$$\tilde{A} := A_{\cdot+\tau-} = a_*(S_{\cdot+\tau-}) = a_*(\tilde{S}_-) = \sum_{y\in\mathcal{Y}(y_0)} a_*(\tilde{S}_-)1_{\{S_\tau=y\}} \overset{\mathrm{d}}{=} \sum_{y\in\mathcal{Y}(y_0)} a_*(S^y_-)1_{\{S_\tau=y\}}$$

and hence

$$W_\tau(A) = W_0(a_*(\tilde{S})) \overset{\mathrm{d}}{=} \sum_{y\in\mathcal{Y}(y_0)} W_0(a_*(S^y_-))1_{\{S_\tau=y\}} = \sum_{y\in\mathcal{Y}(y_0)} w_y 1_{\{S_\tau=y\}}.$$

Since $\tau$ was arbitrary, this shows that $W(A) = w_*(S)$, where $w_*(y) = w_y$ for every $y \in \mathcal{Y}(y_0)$. In particular, $W(A)$ 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

$$w_*(y) = g(y, a_*(y)) + \sum_{y'\in\mathcal{Y}(y_0)} \delta_*(y, y')\lambda_{y,y'}(a_*(y)) \tag{20}$$

for every state $y \in \mathcal{Y}(y_0)$. Moreover, since $W$ has to jump from $w_*(y)$ to $w_*(y')$ when a state transition from $y$ to $y'$ occurs, it follows that $\delta_*(y, y') = \frac{1}{r}(w_*(y') - w_*(y))$ for every $(y, y') \in \mathcal{Y}(y_0)^2$. Substituting the expression for $\delta_*(y, y')$ into (20) and solving for the vector $w_*$ yields (10). Note here that

$$B = \operatorname{diag}(r\mathbf{1} + \mathbf{1}\Lambda_{y_0}(a_*)) - \Lambda_{y_0}(a_*)$$

is indeed invertible since dividing each row $y'$ by $r + \Lambda_{y_0}(a_*)e_{y'}$ turns $B$ into the identity matrix minus a strictly substochastic matrix. Finally, Lemma 4.3 implies (9).

Suppose that the converse holds, i.e., there exists $a_* : \mathcal{Y} \to \mathcal{A}$ with $a_*(y) \in \mathcal{A}(y)$ such that for $w_*$ and $\delta_*$ defined in (10) for states $\mathcal{Y}(y_0)$, inequalities (9) hold. Let $S$ be the state process starting in initial state $y_0$. Let $W$ be a solution to (5) for $A = a_*(S_-)$, $\beta \equiv 0$, $\delta = \delta_*(S_-)$, and $M \equiv 0$, starting in initial state $w_*(y_0)$. Since (10) is equivalent to (20), it follows that $W$ is locally constant unless a state change from $y$ to $y'$ occurs, at which point there is a jump of size $w(y') - w(y)$ in $W$ and hence $W = w_*(S)$. Due to (9), $(\beta, \delta)$ enforces $A$ and hence Lemma 4.3 implies that $A$ is a PPE. $\square$

## D Characterization of $\mathcal{B}(\mathcal{W})$

Since the state is locally constant for enforceable solutions to (5) attaining payoffs in $\mathcal{B}(\mathcal{W})$, the characterization of $\mathcal{B}_y(\mathcal{W})$ for a fixed state $y$ is very similar to the case of repeated games with signal $(S, X)$. There are two differences, both of which affect local Lipschitz continuity of the optimality equations. First, the set $\mathcal{B}_y(\mathcal{W})$ is not necessarily contained within $\operatorname{conv} g(y, \mathcal{A}(y))$ because continuation payoffs partially come from $\operatorname{conv} g(y', \mathcal{A}(y'))$ for successor states $y'$ of $y$. Second, for repeated games, the approximation via the algorithm in Proposition 4.9 leads to a decreasing sequence of payoff sets, each satisfying $\mathcal{B}(\mathcal{W}) \subseteq \mathcal{W}$. In stochastic games, incentives related to state transitions may not come from a superset of $\mathcal{B}_y(\mathcal{W})$ since rewards after a state transition come from the sets $\mathcal{W}_{y'}$ for successor states $y'$ of $y$. We defer the proof of local Lipschitz continuity to Appendix E and elaborate in this appendix how the techniques from repeated games for signal $(S, X)$ can be modified to characterize the boundaries of $\mathcal{B}_y(\mathcal{W})$ in stochastic games.

The first result shows how to construct enforceable solutions to the SDE (5) that locally remain on a curve that that solve the optimality equations. We state the result for general measurable selectors $a_*$, $\beta_*$, and $\delta_*$ so that we can apply it to the maximizers in either (11) or (14).

**Lemma D.1.** *Fix a state $y$. Let $\mathcal{C}$ be a continuously differentiable curve oriented by $w \mapsto N_w$ such that there exist measurable selectors $a_*$, $\beta_*$, $\delta_*$ on $\mathcal{C}$ that for any $w \in \mathcal{C}$,*

*(i)* $(\beta_*(w), \delta_*(w)) \in \Xi_{y,a_*(w)}(w, N_w, \mathcal{W})$,

*(ii) The curvature $\kappa_y(w)$ of $\mathcal{C}$ at $w$ satisfies*

$$\kappa_y(w)\left\|{T_w}^\top \beta_*(w)\sigma(y)\right\|^2 = \frac{2}{r}{N_w}^\top\big(g(y, a_*(w)) + \delta_*(w)\lambda(y, a_*(w)) - w\big).$$

*(iii) If $\beta_*(w) = 0$ for a segment of positive length, then ${T_w}^\top\big(g(y, a_*(w)) + \delta_*(w)\lambda(y, a_*(w)) - w\big)$ is either non-positive or non-negative throughout the segment.*

*Then the solution $(W, S, A, \beta, \delta, M)$ to (5) with $A = a_*(W_-)$, $\beta = \beta_*(W_-)$, $\delta = \delta_*(W_-)$, and $M \equiv 0$ is $\mathcal{W}$-enforceable in state $y$ and it remains on $\mathcal{C}$ until an end point of $\mathcal{C}$ is reached or the state changes.*

*Proof.* Fix $w$ in the relative interior of $\mathcal{C}$ and choose $\eta > 0$ small enough such that ${N_w}^\top N_v > 0$ for all $v \in \mathcal{C} \cap B_\eta(w)$, where $B_\eta(w)$ denotes the closed ball around $w$ with radius $\eta$. On $B_\eta(w)$, $\mathcal{C}$ admits a local parametrization $f$ in the direction $N_w$. For any $v \in B_\varepsilon(w)$, define the orthogonal projection $\hat{v} = {T_w}^\top v$ onto the tangent and denote by $\pi(v) = (\hat{v}, f(\hat{v}))$ the projection of $v \in B_\eta(w)$ onto $\mathcal{C}$ in the direction $N_w$. Let $(W, A, \beta, \delta, Z, (J^{y,y'})_{(y,y') \in \mathcal{Z}}, M)$ be a weak solution to (5) starting at $W_0 = w$ with $A = a_*(\pi(W_-))$, $\beta = \beta_*(\pi(W_-))$, $\delta = \delta_*(\pi(W_-))$, and $M \equiv 0$ on $[\![0, \tau)\!)$, where we set $\tau := \rho \wedge \inf\{t \geq 0 \mid W_t \notin B_\eta(w)\}$, where $\rho$ is the first jump time of any of the processes $(J^{y,y'})_{(y,y') \in \mathcal{Z}}$. By (iii), such a solution exists. Since $\pi$ is measurable, the processes $A$, $\beta$ and $\delta$ are all predictable.

We measure the distance of $W$ to $\mathcal{C}$ by $D_t = {N_w}^\top W_t - f(\hat{W}_t)$. Note that $f$ is differentiable by assumption and $(-f'(\hat{W}_t), 1) = \ell_t N_t$, where $\ell_t := \|(-f'(\hat{W}_t), 1)\|$. Since $f$ is locally convex it is second-order differentiable at almost every point by Alexandrov's Theorem. In particular, $f'$ has Radon-Nikodým derivative $f''(\hat{W}_t) = -\kappa_y(\pi(W_t))\ell_t^3$. It follows from Itō's formula that

$$\begin{aligned} \mathrm{d}D_t &= r\ell_t {N_t}^\top (W_t - g(y, A_t) - \delta_t \lambda(y, A_t))\, \mathrm{d}t + r\ell_t {N_t}^\top \beta_t \left(\sigma(y)\, \mathrm{d}Z_t - \mu(y, A_t)\, \mathrm{d}t\right) \\ &\quad + r\ell_t \sum_{y,y' \in \mathcal{Z}} {N_{t-}}^\top \delta_t(y')\, \mathrm{d}J_t^{y,y'} - \frac{1}{2} f''(\hat{W}_{t-})\, \mathrm{d}[\hat{W}]_t, \end{aligned}$$

where we abbreviated $N_t = N_{\pi(W_t)}$ and $T_t = T_{\pi(W_t)}$. Since $N^\top \beta = 0$, the volatility term vanishes and we can write $\beta = TT^\top \beta$. Before the first state transition, we have $\Delta J^{y,y'} \equiv 0$ for any $y, y' \in \mathcal{Z}$, hence $[\hat{W}] = \langle \hat{W} \rangle = \langle {T_w}^\top W \rangle$. Using (ii) and the fact that ${N_w}^\top N_t = {T_w}^\top T_t = \ell_t^{-1}$, we obtain that on $[\![0, \tau)\!)$,

$$\mathrm{d}D_t = r\ell_t {N_t}^\top (W_t - g(y, A_t) - \delta_t \lambda(y, A_t))\, \mathrm{d}t + \frac{r^2}{2} \kappa_y(\pi(W_t))\, \ell_t^3 \left|{T_w}^\top T_t\right|^2 \left\|{T_t}^\top \beta_t \sigma(y)\right\|^2 \mathrm{d}t = rD_t\, \mathrm{d}t,$$

where we used ${N_t}^\top (W_t - \pi(W_t)) = {N_t}^\top N_w D_t = \ell_t^{-1} D_t$ in the second equality. It follows that $D_t = D_0 \mathrm{e}^{rt} = 0$ since $D_0 = 0$. On $\{\tau < \rho\}$ we can repeat this procedure and concatenate the solutions as in the proof of Lemma B.2 to obtain a solution $(W, S, A, \beta, \delta, M)$ to (5) with $W \in \mathcal{C}$ until time $\rho \wedge \xi$, where $\xi$ is the first time $W$ reaches an end point of $\mathcal{C}$. Condition (i) implies $(\beta, \delta) \in \Xi_{y,A}(\pi(W), N_W, \mathcal{W})$ on $[\![0, \rho \wedge \xi)\!)$, which implies that the solution is $\mathcal{W}$-enforceable since $\pi(W) = W$ on $[\![0, \rho \wedge \xi)\!)$. ☐

## D.1 Proof of Lemma 6.5

The proof of Lemma 6.5 relies on a perturbation argument that shows that solutions to (14) with initial conditions $(w, N) \in \mathcal{N}_{\mathcal{B}_y(\mathcal{W})}$ can neither escape nor fall into the interior of $\mathcal{B}_y(\mathcal{W})$. Since the state is fixed, the perturbation argument is analogous to Bernard (2023). We only sketch the key steps here insofar as they are relevant to establish local Lipschitz continuity.

A key step in the proof is the so-called escaping lemma, stating that it is impossible for a solution $\mathcal{C}$ to the optimality equation in some state $y$ to cut through the interior of $\mathcal{B}_y(\mathcal{W})$; see Figure 13. Because the optimality equation maximizes the curvature over all restricted-enforceable action profiles, any enforceable solution to (5) that attains a payoff pair in $\mathcal{B}_y(\mathcal{W})$ "above the curve" (in the shaded area of Figure 13) must stay above $\mathcal{C}$ with positive probability. If incentives related

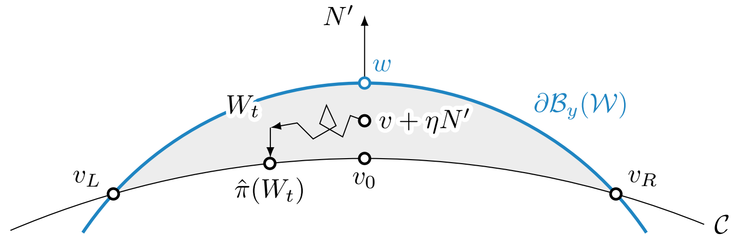



**Figure 13:** A solution $\mathcal{C}$ to the optimality equation that "cuts through" $\mathcal{B}_y(\mathcal{W})$. The proof of the escaping lemma compares the law of motion of $W$ attaining $v + \eta N'$ to the maximal incentives in the optimality equation at its projection $\hat{\pi}(W)$ onto $\mathcal{C}$ to deduce that $W$ must escape $\mathcal{B}_y(\mathcal{W})$ with positive probability.

to the public signal are parallel to $\mathcal{C}$, the continuation value will not get closer to $\mathcal{C}$ because the curvature of $\mathcal{C}$ is maximal. The continuation value will thus escape to either side with positive probability. If incentives related to the public signal involve a normal component, the continuation value is just as likely to escape $\mathcal{B}_y(\mathcal{W})$ as it is to cross $\mathcal{C}$. In either case, the continuation value escapes $\mathcal{B}_y(\mathcal{W})$ with positive probability, contradicting relaxed self-generation.

Incentives related to the state transition differ at any $W_t$ in the shaded area of Figure 13 from the incentives at its projection $\hat{\pi}(W_t)$ onto $\mathcal{C}$. To compare the law of motion in the shaded are of Figure 13 to the curvature of the curve $\mathcal{C}$, we need to consider the following generalization of the optimality equation that enlarges the set of incentives after a state transition in a Lipschitz continuous way.

For a convex set $\mathcal{X}$ and any $h \geq 0$, let $\mathcal{X}_h := \{v \in \mathbb{R}^2 \mid \min_{w \in \mathcal{X}} \|v - w\| \leq h\}$ denote the set of all payoff pairs within distance $h$ from $\mathcal{X}$. Call a family $\mathcal{L} = (\mathcal{L}_y)_{y \in \mathcal{Y}}$ of set-valued maps $\mathcal{L}_y : \mathbb{R}^2 \rightrightarrows \mathbb{R}^{2\times|\mathcal{Y}|}$ a *Lipschitz expansion* of $\mathcal{W} = (\mathcal{W}_y)_{y \in \mathcal{Y}}$ if

$$\mathcal{L}_y(w) = \left\{\delta \in \mathbb{R}^{2\times|\mathcal{Y}|} \;\middle|\; w + r\delta(y') \in \mathcal{W}_{y',\,h(w)} \text{ for all } y' \text{ with } (y, y') \in \mathcal{Z}\right\}$$

for a non-negative Lipschitz-continuous function $h : \mathbb{R}^2 \to \mathbb{R}$. We denote by $\mathcal{L}^0$ the trivial Lipschitz expansion with $h \equiv 0$. Let $\Upsilon_{y,a}(w, N, \mathcal{L})$ denote the set of all $\delta \in \mathcal{L}_y(w)$, for which there exists $\beta$ with $N^\top \beta = 0$ such that $(\beta, \delta)$ enforces $a$. Denote by

$$E_{y,a}(\mathcal{L}) := \left\{(w, N) \in \mathbb{R}^2 \times S^1 \;\middle|\; \Upsilon_{y,a}(w, N, \mathcal{L}) \neq \emptyset\right\}$$

the effective domain of of $(w, N) \mapsto \Upsilon_{y,a}(w, N, \mathcal{L})$. For any $N$ and any $\delta \in \Upsilon_{y,a}(w, N, \delta)$, let $\Phi_{y,a}(N, \delta)$ denote the set of all $\phi \in \mathbb{R}^d$, for which $(T\phi, \delta)$ enforces $a$, where $T$ is the clockwise orthogonal vector to $N$. Let $\phi_y(a, N, \delta)$ denote the shortest vector in $\Phi_{y,a}(N, \delta)$. We can now write the optimality equation in the following form

$$\kappa_{y,\mathcal{L}}(w, N) := \max_{a \in \mathcal{A}(y)} \max_{\delta \in \Upsilon_{y,a}(w,N,\mathcal{L})} \frac{2{N_w}^\top(g(y, a) + \delta\lambda(y, a) - w)}{r\|\phi_y(a, N, \delta)\sigma(y)\|^2} \vee 0. \tag{21}$$

**Lemma D.2.** *Let $w \in \partial\mathcal{B}(\mathcal{W})$ with outward normal $N'$. Let $\pi : U \to \partial\mathcal{B}(\mathcal{W})$ be the projection of a neighborhood $U$ of $w$ onto $\partial\mathcal{B}(\mathcal{W})$ in the direction of $N'$ and set*

$$\mathcal{L}_k(v) := \left\{\delta \in \mathbb{R}^{2\times m} \;\middle|\; v + r\delta(y) \in \mathcal{W}_{k\|\pi(v)-v\|} \textit{ for all } y \in Y\right\} \tag{22}$$

*Let $\mathcal{C}$ be a $C^1$-solution to (21) for $\mathcal{L}_k$ with $k \geq 1$, oriented by $v \mapsto N_v$ with end points $v_L, v_R \in U$. It is impossible that the following properties hold simultaneously:*

*(i)* $v_L + \varepsilon N' \notin \mathcal{B}(\mathcal{W})$ *and* $v_L + \varepsilon N' \notin \mathcal{B}(\mathcal{W})$ *for any* $\varepsilon > 0$,

*(ii)* *there exists* $v_0 \in \mathcal{C}$ *such that* $v_0 + \eta N' \in \mathcal{B}(\mathcal{W})$ *for some* $\eta > 0$,

*(iii)* $\inf_{v \in \mathcal{C}} {N_v}^\top N' > 0$,

*(iv)* $\inf_{v \in \mathcal{C}} |N_v^i| > 0$ *for* $i = 1, 2$,

*(v)* $\mathcal{N}_\mathcal{C} \cap \Gamma(\mathcal{L}_k) = \emptyset$.

*Proof.* The proof is lengthy but entirely analogous to the proof of Lemma C.2 in Bernard (2023). □

*Proof of Lemma 6.5.* Fix a state $y$. First, Lemma D.2 shows that if $w$ is a corner of $\mathcal{B}_y(\mathcal{W})$, then $(w, N) \in \Gamma_y(\mathcal{W})$ for every outward normal vector $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$. If $(w, N) \notin \Gamma_y(\mathcal{W})$, then $(w', N') \notin \Gamma_y(\mathcal{W})$ for $(w', N')$ sufficiently close to $(w, N)$. A solution to (22) with initial conditions $(w - \varepsilon N, N)$ for $\varepsilon > 0$ sufficiently small thus cuts through $\mathcal{B}_y(\mathcal{W})$, which is an impossibility due to Lemma D.2. In particular, the boundary of $\mathcal{B}_y(\mathcal{W})$ is smooth outside of $\Gamma_y(\mathcal{W})$. Next, fix a payoff-direction pair $(w, N) \in \mathcal{N}_{\mathcal{B}_y(\mathcal{W})} \setminus \Gamma_y(\mathcal{W})$ and a solution $\mathcal{C}$ to (22) with initial condition $(w, N)$. Since $\Gamma_y(\mathcal{W})$ is closed, there is a neighborhood of $(w, N)$ that is not contained in $\Gamma_y(\mathcal{W})$ either, on which (22) is continuous in initial conditions by Lemma E.6. Suppose that $\mathcal{C}$ escapes $\mathcal{B}_y(\mathcal{W})$ in a neighborhood of $(w, N)$. Then the solution to (22) for a slightly rotated initial condition would cut through $\mathcal{B}_y(\mathcal{W})$, satisfying all the conditions of Lemma D.2—an impossibility. If $\mathcal{C}$ falls into the interior of $\mathcal{B}_y(\mathcal{W})$, then the solution $\mathcal{C}'$ with initial conditions $(w, N')$ for a slight rotation of $N$ leaves and enters $\mathcal{B}_y(\mathcal{W})$ as in Figure 10. Assumption 1 implies that for any $\delta$ and any non-coordinate tangent vector $T$, there exists a row vector $\phi$ such that $(T\phi, \delta)$ solves the enforceability constraint (6) with equality. In particular, $\Xi_{y,a}(w, N, \mathcal{W})$ is non-empty for non-coordinate $N$. Moreover, for coordinate normal vectors $N = \pm e_i$, individual full rank implies that incentives can be provided to player $-i$ through $\beta^{-i}$ for any $\delta$ that provide sufficient incentives to player $i$. In particular, $\Xi_{y,a}(w, N, \mathcal{W})$ is non-empty also in this case. Thus, we can let $a_*$, $\beta_*$, and $\delta_*$ of Lemma D.1 be the maximizers in (22). By Lemma D.1, the solution $W$ to (5) with $A = a_*(W_-)$, $\beta = \beta_*(W_-)$, $\delta = \delta_*(W_-)$, and $M \equiv 0$ remains on $\mathcal{C}'$ until an end point of $\mathcal{C}'$ is reached or a state transition occurs. By Lemma B.2, $\mathcal{C}' \subseteq \mathcal{B}_y(\mathcal{W})$, a contradiction. □

## D.2 Proof of Lemma 6.4

We begin with two preliminary lemmas.

**Lemma D.3.** *Let $\mathcal{A}_w$ strictly and minimally decomposes $w \notin \mathcal{S}_y(\mathcal{W})$. Then for each $a \in \mathcal{A}_w$, there exists $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$ and $\delta \in \Psi_{y,a}(w, \mathcal{W})$ with $N^\top(g(y, a) + \delta\lambda(y, a) - w) > 0$.*

*Proof.* Fix such $w$ and $\mathcal{A}_w$ and suppose that towards a contradiction that there exists $a \in \mathcal{A}_w$ such that $\mathcal{D}_{y,a} := g(y, a) + \Psi_{y,a}(w, \mathcal{W})\lambda(y, a)$ lies in every lower half-space $H(w, N) = \{v \mid N^\top(v - w) \leq 0\}$ for $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$. Since $w$ is not stationary, $w \notin \mathcal{D}_{y,a}(w)$, hence $N^\top(g(y, a) + \delta\lambda(y, a) - w) < 0$ for all non-extremal $N \in \mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$. By definition of strict decomposability, for any extremal normal

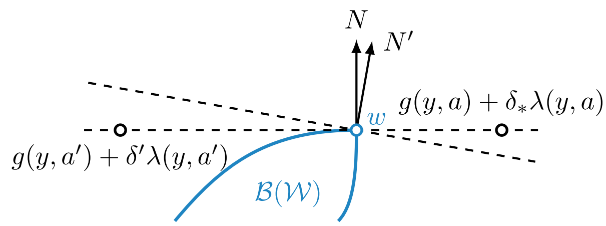



**Figure 14:** Since ${N'}^\top\big(g(y,a)+\delta_*\lambda(y,a)-w\big)\geq 0$ for slight clockwise rotations of $N$, $g(y,a)+\delta_*\lambda(y,a)$ must lie to the right of $w$ and, hence, decompose $w$.

vector $N$ in $\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$, there is $\hat{a}\in\mathcal{A}_w$ and $\hat{\delta}\in\Psi_{y,\hat{a}}(w,\mathcal{W})$ with $N^\top\big(g(y,\hat{a})+\hat{\delta}\lambda(y,\hat{a})-w\big)>0$, hence $\hat{a}\neq a$. It follows that $\mathcal{A}_w\setminus\{a\}$ decomposes $w$, contradicting minimality of $\mathcal{A}_w$. $\square$

**Lemma D.4.** *Let $w\in\partial\mathcal{B}_y(\mathcal{W})\setminus\mathcal{S}_y(\mathcal{W})$ be minimally and strictly decomposed by $\mathcal{A}_w$. Then $w\in\mathcal{B}(\mathcal{W})$, $w\in\partial\mathcal{K}_a(\mathcal{W})$ for each $a\in\mathcal{A}_w$ and* (12) *is satisfied.*

*Proof.* Fix such a payoff pair $w\in\partial\mathcal{B}_y(\mathcal{W})\setminus\mathcal{S}_y(\mathcal{W})$ and a set of action profiles $\mathcal{A}_w$. By Lemma D.3, for each $a\in\mathcal{A}_w$, there exist $N\in\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$ and $\delta\in\Psi_{y,a}(w,\mathcal{W})$ with $N^\top\big(g(y,a)+\delta\lambda(y,a)-w\big)>0$. Thus, Lemma 6.2 implies that $w\in\partial\mathcal{K}_{y,a}(\mathcal{W})$ for each $a\in\mathcal{A}_w$. Strict decomposition implies that there exists a public randomization of action profile in $\mathcal{A}_w$ such that the drift points strictly towards the interior of $\mathcal{B}_y(\mathcal{W})$, hence $w\in\mathcal{B}_y(\mathcal{W})$. For the last statement, suppose towards a contradiction that $\partial\mathcal{B}_y(\mathcal{W})$ enters the interior of $\mathcal{K}_{y,\mathcal{A}_w}(\mathcal{W})$ and, hence, (12) is not satisfied. Since $w\mapsto\Psi_a(w,\mathcal{W})$ is continuous in $\mathcal{K}_{y,a}(\mathcal{W})$ by Lemma E.3, strict decomposition of $w$ implies that there exists $a\in\mathcal{A}_w$ and $\delta\in\Psi_{y,a}(v,\mathcal{W})$ with ${N_v}^\top\big(g(y,a)+\delta\lambda(y,a)-v\big)>0$ for $v\in\partial\mathcal{B}_y(\mathcal{W})$ sufficiently close to $w$, contradicting Lemma D.3. $\square$

*Proof of Lemma 6.4.* Fix $(w,N)\in\mathcal{N}_{\mathcal{B}_y(\mathcal{W})}\cap\Gamma_y(\mathcal{W})$ such that $w\in\mathcal{K}_y(\mathcal{W})\setminus\mathcal{S}_y(\mathcal{W})$. Lemma 6.2 implies that $(w,N')\in\Gamma_y(\mathcal{W})$ for every $N'\in\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$. In particular, the payoff pair $w$ is decomposable. If $w$ is strictly decomposable, the result follows from Lemma D.4. Suppose, therefore, that $w$ is not strictly decomposable, i.e., there exists an extremal normal vector $N$, for which $N^\top\big(g(y,a)+\delta\lambda(y,a)-w\big)\leq 0$ for all $\delta\in\Psi_{y,a}(w,\mathcal{W})$ and every $a\in\mathcal{A}$. Because $(w,N)\in\Gamma_y(\mathcal{W})$, there must exist at least one action profile, for which the expression is non-negative and hence (11) holds. If $\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))=\{N\}$, then the maximizer $a_*$ of (11) minimally decomposes $w$. If $w$ is a corner, then, because $w$ is decomposable, there exists $a\in\mathcal{A}$ and $\delta\in\Psi_{y,a}(w,\mathcal{W})$ with ${N'}^\top\big(g(y,a)+\delta\lambda(y,a)-w\big)\geq 0$ for arbitrarily slight rotations $N'\in\operatorname{int}\mathcal{N}_w(\mathcal{B}_y(\mathcal{W}))$ of the extremal normal vector $N$. Without loss of generality, suppose that $N'$ are clockwise rotations of $N$. Since $\Psi_{y,a}(w,\mathcal{W})$ is closed, there must exist $\delta_*\in\Psi_{y,a}(w,\mathcal{W})$ with ${N'}^\top\big(g(y,a)+\delta_*\lambda(y,a)-w\big)\geq 0$ for $N'=N$ as well as for slight clockwise rotations $N'$ of $N$. Therefore, $a$ maximizes (11) and $g(y,a)+\delta_*\lambda(y,a)$ lies on the tangent to $\mathcal{B}_y(\mathcal{W})$ to the right of $w$ as illustrated in Figure 14. It follows that $a$ decomposes $w$. $\square$

### D.3 Proof of Theorem 6.6

*Proof of Theorem 6.6.* Fix a family of payoff sets $\mathcal{W}$, a state $y$ and consider a payoff set $\mathcal{X}$ that satisfies conditions (i) and (ii) of Theorem 6.6. It follows from Lemmas D.1 and D.4 that any payoff pair on the boundary $\partial\mathcal{X}$ can be attained with a locally enforceable solution to (5) that remains in $\operatorname{cl}\mathcal{X}$ until the first state transition.[22] It follows that $\operatorname{cl}\mathcal{X} \subseteq \mathcal{B}_y(\mathcal{W})$ by maximality of $\mathcal{B}_y(\mathcal{W})$. Next, by Lemmas D.4 and 6.5 the boundary of $\mathcal{B}_y(\mathcal{W})$ satisfies conditions (i) and (ii). This implies that $\operatorname{cl}\mathcal{B}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W})$, hence $\mathcal{B}_y(\mathcal{W})$ is closed. Finally $\mathcal{S}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W})$ since for each $w \in \mathcal{S}_y(\mathcal{W})$, the singleton $\{w\}$ is relaxed self-generating. $\square$

## E Regularity of the optimality equations

To show local Lipschitz continuity of the optimality equations, we show that both ODEs are a maximization of a locally Lipschitz continuous function over a locally Lipschitz continuous set of available incentives. For set-valued maps, Lipschitz continuity is defined as follows.

**Definition E.1.** A set-valued map $G : x \mapsto G(x) \subseteq \mathbb{R}^k$ is said to be *Lipschitz continuous* if there exists a constant $K$ such that $G(x) \subseteq G(\tilde{x}) + K\|x - \tilde{x}\|B_1(0)$ for any $x$ and $\tilde{x}$, where $B_1(0)$ denotes the closed unit ball in $\mathbb{R}^k$ centered at the origin and $+$ is the setwise addition.

### E.1 Abrupt-information optimality equation

For any state $y$ and any $a \in \mathcal{A}(y)$, let $\Psi^i_{y,a}$ denote the set of all $\delta^i$ that solve the enforceability constraint (6) for player $i = 1, 2$. We view $\Psi_{y,a} = \Psi^1_{y,a} \times \Psi^2_{y,a}$ as a subset of $\mathbb{R}^{2|\mathcal{Y}|}$ with typical element $\delta = (\delta^1, \delta^2)$. Observe that $\Psi_{y,a}$ is the intersection of closed half-spaces with normal vectors $\Delta\lambda^1_{y,a}(\tilde{a}^1) := \big(\lambda(y, \tilde{a}^1, a^2) - \lambda(y, a), 0\big)$ for $\tilde{a}^1 \in \mathcal{A}^1(y) \setminus \{a^1\}$ and $\Delta\lambda^2_{y,a}(\tilde{a}^2) := \big(0, \lambda(y, a^1, \tilde{a}^2) - \lambda(y, a)\big)$ for $\tilde{a}^2 \in \mathcal{A}^2(y) \setminus \{a^2\}$. Let $\Theta_{y,a}(w)$ denote the set of all $\delta \in \mathbb{R}^{2|\mathcal{Y}|}$ with $w + r\delta(y') \in \mathcal{W}_{y'}$ for each $y' \in \mathcal{Z}_y(a)$, where we omit the dependence on the family $\mathcal{W}$ since it is usually fixed. Let us abbreviate $\Psi_{y,a}(w) := \Psi_{y,a}(w, \mathcal{W})$ and observe that $\Psi_{y,a}(w) = \Psi_{y,a} \cap \Theta_{y,a}(w)$.

Recall that $\mathcal{K}_{y,a}(\mathcal{W})$ is the set of all payoff pairs $w$, for which $\Psi_{y,a}(w)$ is non-empty, and that $\mathcal{K}_y(\mathcal{W}) = \bigcup_a \partial\mathcal{K}_{y,a}(\mathcal{W})$. If each set $\mathcal{W}_{y'}$ is closed or convex, respectively, then so are $\Theta_{y,a}(w)$, $\Psi_{y,a}(w)$, and $\mathcal{K}_{y,a}(\mathcal{W})$. We begin with the following two observations.

**Lemma E.2.**

*(i) If there exists $w \in \mathcal{K}_{y,a}(\mathcal{W})$, for which $\Psi_{y,a}$ intersect the interior of $\Theta_{y,a}(w)$, then $\Psi_{y,a}$ intersects the interior of $\Theta_{y,a}(v)$ for every $v \in \operatorname{int}\mathcal{K}_{y,a}(\mathcal{W})$.*

*(ii) Let $H^i_{y,a}(\tilde{a}^i)$ denote the hyperface of $\Psi_{y,a}$ with normal vector $\Delta\lambda^i_{y,a}(\tilde{a}^i)$. If $\Psi_{y,a}(w) \subseteq H^i_{y,a}(\tilde{a}^i)$ for some $w \in \operatorname{int}\mathcal{K}_{y,a}(\mathcal{W})$, then $\Psi_{y,a}(v) \subseteq H^i_{y,a}(\tilde{a}^i)$ for all $v \in \mathcal{K}_{y,a}(\mathcal{W})$.*

[22] Where the maximizer in (11) changes, public randomization may be necessary even on the boundary; see Lemma C.5 in Bernard (2023) for details.

*Proof.* For the first statement, fix $v \in \mathcal{K}_{y,a}(\mathcal{W})$ and $\delta_v \in \Psi_{y,a} \cap \operatorname{int} \Theta_{y,a}(v)$. Because $\Psi_{y,a}$ is constant and $\Theta_{y,a}(w)$ is Lipschitz continuous, there exists an open neighborhood $U$ of $v$ such that $\Psi_{y,a}$ intersects the interior of $\Theta_{y,a}(w)$ for any $w \in U$. In particular, $v \in \operatorname{int} \mathcal{K}_{y,a}(\mathcal{W})$. By convexity of $\mathcal{K}_{y,a}(\mathcal{W})$, we can write any other $w \in \operatorname{int} \mathcal{K}_{y,a}(\mathcal{W})$ as a strict convex combination $w = xv + (1-x)u$ for some $u \in \mathcal{K}_{y,a}(\mathcal{W})$. Convexity of $\Psi_{y,a}$ and each $\mathcal{W}_{y'}$ imply that $\delta_* := x\delta_v + (1-x)\delta_u \in \Psi_{y,a}$ and $w + r\delta_*(y') \in \mathcal{W}_{y'}$ for each $y' \in \mathcal{Z}_y(a)$. In particular, $\delta_* \in \Psi_{y,a}(w)$.

For the second statement, fix $w \in \operatorname{int} \mathcal{K}_{y,a}(\mathcal{W})$, a player $i$, and a deviation $\tilde{a}^i$ with $\Psi_{y,a}(w) \subseteq H^i_{y,a}(\tilde{a}^i)$. Suppose towards a contradiction that there exists $v \in \mathcal{K}_{y,a}(\mathcal{W})$ and $\delta_v \in \Psi_{y,a}(v)$ for which deviations to $\tilde{a}^i$ by player $i$ are deterred strictly. Write $w$ as a strict convex combination $w = xv + (1-x)u$ of $v$ and $u \in \mathcal{K}_{y,a}(\mathcal{W})$. Convexity of $\Psi_{y,a}$ and $\mathcal{W}$ imply that $\delta_* := x\delta_v + (1-x)\delta_u \in \Psi_{y,a}(w)$ for arbitrary $\delta_u \in \Psi_{y,a}(u)$. Since deviations to $\tilde{a}^i$ are deterred strictly by $\delta_*$, this is a contradiction. $\square$

The following lemma establishes that in the interior of its domain $\mathcal{K}_{y,a}(\mathcal{W})$, the maximization in the abrupt-information optimality equation (11) is taken over a locally Lipschitz continuous set of parameters.

**Lemma E.3.** *$w \mapsto \Psi_{y,a}(w)$ is continuous on $\mathcal{K}_{y,a}(\mathcal{W})$ and locally Lipschitz continuous in $\operatorname{int} \mathcal{K}_{y,a}(\mathcal{W})$.*

*Proof.* We first show local Lipschitz continuity. If $\mathcal{K}_{y,a}(\mathcal{W})$ contains a single element $v$, for which $\Psi_{y,a}$ intersects the interior of $\Theta_{y,a}(v)$, then $\Psi_{y,a}$ intersects the interior of $\Theta_{y,a}(w)$ for every $w$ in the interior of $\mathcal{K}_{y,a}(\mathcal{W})$. It follows from Lemma B.3 in Bernard (2023) that $w \mapsto \Psi_{y,a}(w)$ is locally Lipschitz continuous in the interior of $\mathcal{K}_{y,a}(\mathcal{W})$. Suppose, therefore, that $\Psi_{y,a}$ does not intersect the interior of $\Theta_{y,a}(w)$ for any $w \in \mathcal{K}_{y,a}(\mathcal{W})$. Then $\Theta_{y,a}(w)$ moves parallel to the hyperfaces of $\Psi_{y,a}$ that contain $\Psi_{y,a}(w)$. Since the set of binding hyperfaces does not change in the interior of $\mathcal{K}_{y,a}(\mathcal{W})$ by Lemma E.2, the maximal change in $\Psi_{y,a}(w)$ is bounded by the maximal change in $\Theta_{y,a}(w)$.

It remains to show that $w \mapsto \Psi_{y,a}(w)$ is continuous at the boundary of $\mathcal{K}_{y,a}(\mathcal{W})$. Because it is the intersection of two upper semicontinuous maps, it is again upper semicontinuous. For lower semicontinuity, fix $w \in \partial\mathcal{K}_{y,a}(\mathcal{W})$ and let $\tilde{\mathcal{A}}^i$ be the set of those deviations to $\tilde{a}^i$ for which incentives bind at $w$ in state $y$. Thus, the hyperfaces of $\Psi_{y,a}$ that intersect $\Theta_{y,a}(w)$ are precisely those with normal vectors $\Delta\lambda^i_{y,a}(\tilde{a}^i)$ for $\tilde{a}^i \in \tilde{\mathcal{A}}^1 \cup \tilde{\mathcal{A}}^2$. Let $(w_n)_{n\geq 0} \subseteq \mathcal{K}_{y,a}(\mathcal{W})$ be any sequence approximating $w$. Since $w_n \in \mathcal{K}_{y,a}(\mathcal{W})$, there must exist at least one face of $\Psi_{y,a}$ that intersects $\Theta_{y,a}(w_n)$. Since $\Psi_{y,a}$ has finitely many hyperfaces, by passing to a subsequence we may assume that the same set of faces intersect $\Theta_{y,a}(w_n)$ for any $n \geq 0$. Moreover, those hyperfaces are a subset of $\tilde{\mathcal{A}}^1 \cup \tilde{\mathcal{A}}^2$. Let $\hat{\Psi}_{y,a}(v)$ denote the intersection of those faces with $\Theta_{y,a}(v)$. Then $\Psi_{y,a}(w) \subseteq \hat{\Psi}_{y,a}(w)$ and $\hat{\Psi}_{y,a}(w_n) \subseteq \Psi_{y,a}(w_n)$ for any $n$.

Observe that any $\delta$ in the relative interior of $\hat{\Psi}_{y,a}(w)$ is in the relative interior of $\hat{\Psi}_{y,a}(w_n)$ for $n$ sufficiently large since $\Psi_{y,a}$ is constant and $\Theta_{y,a}(w_n)$ converges to $\Theta_{y,a}(w)$. In particular, any $\delta_* \in \Psi_{y,a}(w)$ that lies in the relative interior of $\hat{\Psi}_{y,a}(w)$ can be approximated by a constant sequence $\delta_n = \delta_*$ for $n$ sufficiently large. If $\hat{\Psi}_{y,a}(w) = \{\delta_*\}$ is a singleton (hence so is $\Psi_{y,a}(w)$), then an arbitrary sequence in $\hat{\Psi}_{y,a}(w_n)$ approaches $\delta_*$. If $\hat{\Psi}_{y,a}(w)$ is not a singleton, fix arbitrary $\delta_0$ in the relative interior of $\hat{\Psi}_{y,a}(w)$ (hence in the relative interior of $\hat{\Psi}_{y,a}(w_n)$ for $n$ sufficiently large) and parametrize $\delta$ in the relative boundary of $\hat{\Psi}_{y,a}(v)$ through the direction $e_\delta$ of $\delta - \delta_0$. Any $\delta_* \in \Psi_{y,a}(w)$ on the relative boundary of $\hat{\Psi}_{y,a}(w)$ can then be approximated by $\delta_n$ on the relative boundary of $\hat{\Psi}_{y,a}(w_n)$ parametrized by $e_{\delta_*}$. □

To obtain a geometric interpretation of the abrupt-information optimality equation (11), note that we can equivalently write it as $N_w{}^\top w = \max_x N_w{}^\top x$, where the maximization is over $x$ in

$$\mathcal{D}_y(w) := \operatorname{conv}\left( \bigcup_{a \in \mathcal{A}(y)} g(y,a) + \Psi_{y,a}(w)\lambda(y,a) \right).$$

Let us denote by $\bar{\mathcal{S}}_y(\mathcal{W})$ the set of all payoff pairs $w$ for which $w \in \mathcal{D}_y(w)$. At any $w \notin \bar{\mathcal{S}}_y(\mathcal{W})$, a solution to (11) evolves in the direction of a tangent $T_w$ to $\mathcal{D}_y(w)$ through $w$. Because $\mathcal{D}_y(w)$ is convex, there are two such directions at any payoff pair $w \notin \bar{\mathcal{S}}_y(\mathcal{W})$. Call a solution to (11) an *oriented solution* if the normal vector $N_w$ always points towards the same side of the solution. Note that because there are two oriented solutions starting at any $w \notin \bar{\mathcal{S}}_y(\mathcal{W})$, initial conditions of (11) consist of a payoff pair $w$ and the selection of an orientation.

**Proposition E.4.** *Outside the set $\mathcal{K}_y(\mathcal{W}) \cup \bar{\mathcal{S}}_y(\mathcal{W})$, oriented solutions to* (11) *in state $y$ are continuously differentiable and they are unique and continuous in initial conditions.*

*Proof.* Fix a state $y$ and a payoff pair $w \notin \mathcal{K}_y(\mathcal{W}) \cup \bar{\mathcal{S}}_y(\mathcal{W})$. Since $\mathcal{D}_y(w)$ is closed and $w \notin \bar{\mathcal{S}}_y(\mathcal{W})$, the set $\mathcal{D}_y(w)$ is bounded away from $w$. Because $w \notin \mathcal{K}_y(\mathcal{W})$, the map $\mathcal{D}_y$ is locally Lipschitz continuous by Lemma E.3, hence $\mathcal{D}_y(v)$ is bounded away from $v$ close to $w$. Because $\mathcal{D}_y$ is uniformly bounded, the direction of the tangent $T_w$ changes continuously in $w$, i.e., oriented solutions to (11) are continuously differentiable.

Since the maximum of a locally Lipschitz continuous function over a locally Lipschitz continuous set of maximizers is locally Lipschitz continuous by Lemma B.2 of Bernard (2023), uniqueness and continuity in initial conditions follows from Lemma E.3 and the Picard-Lindelöf theorem. □

If $w$ lies in the interior of $\mathcal{D}_y(w)$, then there is no tangent to $\mathcal{D}_y(w)$ through $w$, hence the abrupt-information optimality equation has no solution. Such payoff pairs lie in the interior of $\bar{\mathcal{S}}_y(\mathcal{W})$, hence in the interior of $\mathcal{B}_y(\mathcal{W})$ by the following lemma. Together with Proposition E.4, the following lemma implies that solutions to the abrupt-information optimality equation (11) exist outside of $\operatorname{int} \mathcal{B}_y(\mathcal{W})$ and they are unique and continuous in initial conditions outside of $\mathcal{K}_y(\mathcal{W})$.

**Lemma E.5.** $\bar{\mathcal{S}}_y(\mathcal{W}) \subseteq \mathcal{B}_y(\mathcal{W})$.

*Proof.* Since $w \in \mathcal{D}_y(w)$ for any $w \in \bar{\mathcal{S}}_y(\mathcal{W})$, we can write $w$ as a convex combination of $w_k = g(y,a_k) + \delta_k \lambda(y,a_k)$ with weights $x_k \geq 0$. Let $(A,\delta)$ be an infinitesimally quick public randomization that selects $(a_k,\delta_k)$ with probability $x_k$ at every instant of time. Then an exact law of large numbers implies that $W \equiv w$ is a $\mathcal{W}$-enforceable solution to (5) for $(A,\delta)$ as well as $\beta = M = 0$ up until the first state transition. See Sun (2006) for a sufficiently general statement of the exact law of large numbers and the proof of Lemma C.5 in Bernard (2023) for details on this construction. □

## E.2 Optimality equation

We prove local Lipschitz continuity for the optimality equation in the form (21).

**Lemma E.6.** *Suppose that Assumptions 1 and 2 are satisfied. Then* (21) *is locally Lipschitz continuous outside of* $\Gamma_y(\mathcal{L}) = \bigcup_{a \in A(y)} \Gamma_{y,a}(\mathcal{L})$*, where*

$$\Gamma_{y,a}(\mathcal{L}) := \{(w, N) \mid \textit{there exists } \delta \in \mathcal{L}_y(w) \textit{ with } (0, \delta) \textit{ enforces } a\}.$$

The basic idea behind the proof of Lemma E.6 is to show that for any given $a$, the right-hand side of (21) is locally Lipschitz continuous on $E_{y,a}(\mathcal{L}) \setminus \Gamma_{y,a}(\mathcal{L})$. Then $\kappa_{y,\mathcal{L}}$ is locally Lipschitz continuous except at $\bigcup_{a \in \mathcal{A}(y)} \partial E_{y,a}(\mathcal{L})$, where the set of restricted-enforceable action profiles changes, and on $\Gamma_y(\mathcal{L})$, where the denominator is 0. The following lemma shows that $\partial E_{y,a}(\mathcal{L})$ is contained in $\mathbb{R}^2 \times \{\pm e_1, \pm e_2\}$, hence the proof of Lemma E.6 need only deal with Lipschitz continuity in coordinate directions.

**Lemma E.7.** *Suppose that Assumption 1 is satisfied. For any Lipschitz expansion* $\mathcal{L}$*, any state* $y$*, and any* $a \in \mathcal{A}(y)$*, we have* $\mathbb{R}^2 \times (S^1 \setminus \{\pm e_1, \pm e_2\}) \subseteq E_{y,a}(\mathcal{L})$*. Moreover,* $(w, N) \mapsto \Upsilon_{y,a}(w, N, \mathcal{L})$ *is locally Lipschitz continuous on* $\mathbb{R}^2 \times (S^1 \setminus \{\pm e_1, \pm e_2\})$*.*

*Proof.* Fix any $\mathcal{W}$, $y$, $a$, $w \in \mathbb{R}^2$, and any non-coordinate $N$. Choose any $\delta$ in $\mathcal{L}_y(w)$. Since $M^i(a)$ has individual full rank by Assumption 1, there exists $\beta^i$ that solves (6) for player $i$ with equality. Since $a$ is pairwise identifiable by Assumption 1, it follows from Lemma 2 in Sannikov (2007) applied to payoff function $\tilde{g}(y, a) = g(y, a) + \delta\lambda(y, a)$ that there exists $\beta'$ with $N^\top \beta' = 0$ such that $(\beta', \delta)$ enforces $a$. In particular, $\Upsilon_{y,a}(w, N, \mathcal{L}) \neq \emptyset$ and hence $(w, N) \in E_{y,a}(\mathcal{L})$. Moreover, this construction shows that $\Upsilon_{y,a}(w, N, \mathcal{L}) = \mathcal{L}_y(w)$ for non-coordinate $N$, which is locally Lipschitz continuous. □

*Proof of Lemma E.6.* Let $G^i_y(a)$ denote the row vector with entries $g^i(y, \tilde{a}^i, a^{-i}) - g^i(y, a)$ and let $\Lambda^i_y(a)$ denote the matrix with column vectors $\lambda(y, \tilde{a}^i, a^{-i}) - \lambda(y, a)$. For non-coordinate $N$, $\Phi_{y,a}(N, \delta)$ is the set of all $\phi$ that satisfy $\phi M^i_y(a) \leq -\frac{1}{T^i}\big(G^i_y(a) + \delta^i \Lambda^i_y(a)\big)$ for $i = 1, 2$. Since $M^i_y(a)$ does not depend on $N$ or $\delta$ and the right-hand side is locally Lipschitz continuous in $N, \delta$, the set $\Phi_{y,a}(N, \delta)$ is a constant-rank polyhedron with locally Lipschitz continuous right-hand side. It follows from the main result of Yen (1995) that the projection of 0 onto $\Phi_{y,a}(N, \delta)$ is locally Lipschitz continuous in $(N, \delta)$. In particular, $\phi_y(a, N, \delta)$ is locally Lipschitz continuous in $(N, \delta)$ for non-coordinate $N$.

Consider now a coordinate direction $N = \pm e_i$. Then any $\beta$ with $N^\top \beta$ cannot provide any incentives to player $i$, hence any $\delta \in \Upsilon_{y,a}(w, N, \mathcal{L})$ must satisfy $G^i_y(a) + \delta^i \Lambda^i_y(a) \leq 0$. Denote by $\Phi^{-i}_{y,a}(\delta)$ the set of all $\phi$ that satisfy $\phi M^{-i}_y(a) \leq -(G^i_y(a) + \delta^{-i} \Lambda^{-i}_y(a))$. As above, the shortest vector $\phi^{-i}_y(a, \delta)$ in $\Phi^{-i}_{y,a}(\delta)$ is locally Lipschitz continuous in $\delta$ as the projection of 0 onto $\Phi^{-i}_{y,a}(\delta)$. The shortest vector in $\Phi_{y,a}(N, \delta) = \frac{1}{T^1}\Phi^1_{y,a}(\delta) \cap \frac{1}{T^2}\Phi^2_{y,a}(\delta)$ is thus at least as long as $\frac{1}{T^{-i}}\phi^{-i}_y(a, \delta)$. Since $\operatorname{span} M^i_y(a)$ is orthogonal to $\operatorname{span} M^{-i}_y(a)$ it follows that $\frac{1}{T^{-i}}\phi^{-i}_y(a, \delta) M^i_y(a) = 0$, hence $\phi_y(a, N, \delta) = \frac{1}{T^{-i}}\phi^{-i}_y(a, \delta)$, which is locally Lipschitz continuous in a neighborhood of $(N, \delta)$.

Since $\Gamma_y(\mathcal{W})$ is closed, for any $(w, N) \notin \Gamma_y(\mathcal{W})$, there exists a neighborhood on which $\|\phi_y(a, N, \delta)\|$ is bounded away from 0. For non-coordinate $N$, it follows from Lemma E.7 that

$\kappa_{y,\mathcal{L}}(w, N)$ is locally Lipschitz continuous as the maximum of a locally Lipschitz continuous function over a locally Lipschitz continuous set; see Lemma B.2 in Bernard (2023). For $(w, N)$ with coordinate $N$, fix any action profile $a$ with $\Upsilon_{y,a}(w, N, \mathcal{L}) = \emptyset$. This means that there exists no $\delta \in \mathcal{L}(w)$ with $0 \in \Phi^i_{y,a}(\delta)$. Since $\Phi^i y, a(\delta)$ is closed, this implies that 0 is bounded away from $\Phi^i_{y,a}(\delta)$, hence $\|\phi_y(a, \tilde{N}, \delta)\|$ converges to $\infty$ as $\tilde{N}$ converges to $N$ because it is the shortest vector in $\Phi_{y,a}(\tilde{N}, \delta) = \frac{1}{\tilde{T}^1}\Phi^1_{y,a}(\delta) \cap \frac{1}{\tilde{T}^2}\Phi^2_{y,a}(\delta)$. Let $\mathcal{A}_y(w, N)$ denote the action profiles $a \in \mathcal{A}(y)$, for which

$$\kappa_{y,a,\mathcal{L}}(w, N) := \max_{\delta \in \Upsilon_{y,a}(w,N,\mathcal{L})} \frac{2{N_w}^\top(g(y, a) + \delta\lambda(y, a) - w)}{r\|\phi_y(a, N, \delta)\sigma(y)\|^2} \vee 0$$

is strictly positive. Then $\kappa_{y,a,\mathcal{L}}(w', N') > 0$ for $(w', N')$ in a neighborhood of $(w, N)$ and hence $\kappa_{y,\mathcal{L}}(w', N') = \max_{a \in \mathcal{A}_y(w',N')} \kappa_{y,a,\mathcal{L}}(w, N)$ in a neighborhood of $(w, N)$. In particular, $\kappa_{y,\mathcal{L}}$ is locally Lipschitz continuous at $(w, N)$. $\square$