CONTINUOUS-TIME GAMES WITH IMPERFECT
AND ABRUPT INFORMATION*

BENJAMIN BERNARD†

This version: November 26, 2017

This paper studies two-player games in continuous time with imperfect public monitoring, where information may arrive both continuously, governed by a Brownian motion, and discontinuously, according to Poisson processes. For this general class of two-player games, we characterize the equilibrium payoff set via a differential equation of its boundary. The equilibrium payoff set is obtained from an iterative procedure similar to that known in discrete time. In our setting, however, the resulting payoff set in each step of the iteration is characterized explicitly. Our analysis reveals the drastic influence of abrupt information on the equilibrium payoff set: because of the additional possibility for value burning, the equilibrium payoff set extends much closer to the efficient frontier and its boundary may contain corners and straight line segments.

KEYWORDS: Repeated games, continuous time, imperfect observability, equilibrium characterization, abrupt information.

1 Introduction

In continuous-time games with imperfect monitoring, information may arrive both continuously through a noisy signal and intermittently as occurrences of infrequent but informative events. In many applications, there is a clear distinction between the two types of information. Consider a climate agreement that obligates each signatory to reduce its greenhouse gas emissions. Countries cannot measure each other’s emissions directly, hence they must rely on imperfect information to enforce the agreement. When one country violates the terms of the agreement and emits more than

*I would like to thank Christoph Frei, Tilman Klumpp, participants of the 11th World Congress of the Econometric Society, the 5th World Congress of the Game Theory Society, the Asian Meeting of the Econometric Society 2016, the 2017 Midwest Economic Theory Conference, and seminar participants at the University of Bern for valuable comments and interesting discussions. Financial support by the Natural Sciences and Engineering Research Council of Canada through grant RGPIN/402585-2011 is gratefully acknowledged.

†Department of Economics, UCLA, Los Angeles, CA 90095; benbernard@ucla.edu.
its agreed-upon share of greenhouse gases, other countries may observe an increase in industrial production or an increase in atmospheric greenhouse gas concentration — information that is suggestive but not conclusive proof that the country has violated the agreement. These observations or measurements are possible at any point in time and are best modeled using a continuous but noisy signal. In addition to these continuous processes, countries also observe infrequent but informative political and economic events such as the passing of an environmental bill or the commissioning of a coal power plant. These are events that are inherently discrete and are better modeled using discontinuous processes that jump when one of these events is observed. In another example, consider a partnership between two firms, where each firm chooses hidden effort levels and observes only the total revenue of the partnership. The total revenue moves continuously due to day-to-day fluctuations in supply and demand conditions and it is subject to demand shocks when one of the firms receives bad press. As illustrated in Figure 1, a decomposition of the information leads to two separate signals that are both indicative of the partner’s effort level: the continuous increase in total revenue without the impact of demand shocks and the frequency at which scandals occur. The goal of this paper is to characterize the set of all equilibrium payoffs in two-player games, where information may arrive both continuously and abruptly. Compared to the existing literature, the more general information structure allows for a wider class of incentives that can be provided to players, thereby fundamentally changing the shape of the equilibrium payoff set.

Repeated games provide a very tractable framework to model sustained interaction between strategic decision makers. A key feature of repeated games that pertains to their tractability is the time homogeneity of the decision problem. Because players
face the same decision tomorrow as they do today, subgame perfect equilibria can be constructed in a recursive way as demonstrated in Abreu, Pearce and Stacchetti [1]. While these recursive techniques have been extremely fruitful in understanding incentives in repeated interactions, the continuous-time methods make it possible to obtain explicit results that are not available in discrete time. In his introduction of continuous-time repeated games, Sannikov [16] shows that continuous-time techniques make it possible to relate incentives to the curvature of the equilibrium payoff set. For a class of two-player games with Brownian information, he characterizes the set of equilibrium payoffs $\mathcal{E}(r)$ for any discount rate $r > 0$ by describing its boundary via a differential equation. Such a characterization of the equilibrium payoff set for boundedly patient players is not known in discrete time, where results on equilibrium payoffs are often restricted to folk theorems; see Fudenberg, Levine and Maskin [7].

The assumption that the observed information is Brownian in [16] is not without loss of generality. Aside from the modeling perspective that certain events are inherently discrete, the assumption on Brownian information also restricts the types of incentives that can be provided to players. As argued in Sannikov and Skrzypacz [18], it is too costly to attach mutual punishments to undesirable outcomes of Brownian information. Incentives at the boundary of the equilibrium payoff set in [16] are thus restricted to tangential transfers of value between players. These are precisely the types of incentives used in [7] to prove the folk theorem: since it is impossible to attain asymptotic efficiency if value is destroyed by punishing all players, the only asymptotically efficient means of incentive provision is thorough transfers of value. For impatient players, however, the destruction of value upon the arrival of an undesirable signal can be an efficient way of providing incentives as seen in Green and Porter [8]. Because one of the appeals of continuous-time models are the availability of explicit results for boundedly patient players, we want to allow an information structure, in which players are not inherently restricted to incentives that are efficient in the limit as players become arbitrarily patient. We achieve this by complementing the Brownian information structure with the observation of infrequent events, whose arrival times are governed by Poisson processes. Players’ actions affect the intensities at which these infrequent events occur. While the events are observable by all players, the intensities are not.

This model of mixed monitoring has first been proposed in Sannikov and Skrzypacz [18] in their treatment of discrete-time games with frequent actions. Combining the recursive techniques in [1] with the restrictions on how continuous and abrupt

---

1In the theory of discrete-time repeated games with public monitoring, barely any distributional assumption have to be made on the public signal. The only restriction is that the signals in different periods are independent and conditionally identically distributed, given the played action profile, to ensure homogeneity over time periods. Public signals that satisfy a similar time homogeneity property in a continuous-time setting are Lévy processes—a class of signals that includes not only Brownian motion, but also discontinuous processes with independent increments.
Figure 2: Because Brownian information is used to transfer value between players along tangents, only local information about the geometry of the equilibrium payoff set $\mathcal{E}(r)$ is needed to structure incentives at a payoff $w$ on the boundary. In contrast, abrupt information may induce jumps in the continuation value that are required to remain in $\mathcal{E}(r)$—requiring global information about $\mathcal{E}(r)$.

Information can be used to provide incentives, they establish a payoff bound for discrete-time games when the length of the time period approaches zero. By working directly in a continuous-time setting, we are able to relate these informational restrictions to the boundary of the equilibrium payoff set $\mathcal{E}(r)$ and characterize the set for any discount rate $r > 0$. Our main result is thus a generalization of Theorem 2 in Sannikov [16] to a more general information structure that includes abrupt information, as well as an improvement from a payoff bound in Sannikov and Skrzypacz [18] to an exact description of the equilibrium payoff set. In terms of provided incentives, this paper differs from [18] by the fact that only bounded amounts of value can be transferred or destroyed upon the arrival of an infrequent event. This is not a difference in the underlying model, but rather in the result that we prove.

Cooperation between players requires larger punishments or rewards when players are impatient. To enforce equilibrium behavior, the continuation payoffs after the punishments/rewards have to remain in the bounded equilibrium payoff set. As players get more and more impatient, fewer incentives can be provided through the observation of infrequent events. When player get too impatient, the provided incentives could be insufficient to support equilibrium behavior outside the set of static Nash profiles and the equilibrium payoff set could collapse to the set of static Nash payoffs. This is a feature not observed in [18] as their payoff bound relies on incentives that can be provided when players get arbitrarily patient.

Our proof is based on the fact that the equilibrium payoff set is the largest bounded self-generating set. The continuation value of a perfect public equilibrium (PPE) thus has to remain within $\mathcal{E}(r)$ at all times, which restricts incentives that can be used at the boundary; see Figure 2. Sannikov [16] shows that when information is Brownian, only local information about the boundary of $\mathcal{E}(r)$ is needed to characterize the set of equilibrium payoffs, in the same way that any ODE uses only local information about a function to encode its global properties. Crucial in this respect is that Brownian information is used to transfer value between players tangentially. In contrast, when information arrives according to Poisson processes, such signals can
be used also to provide incentives via value burning, which incurs jumps in players’ continuation value. Because $\mathcal{E}(r)$ is self-generating, the continuation value after such a jump must remain in the equilibrium payoff set. This restriction on incentives, however, requires global information about the set of PPE payoffs, which creates a non-trivial fixed-point problem. We get around this issue with an iterative procedure over the arrival times of infrequent events: at every step of the algorithm, we calculate the largest self-generating set under the restriction that continuation values must jump into the self-generating set from the previous step. Because the restriction on incentives through the abrupt information is fixed in each step, the resulting sets are characterized explicitly by a differential equation. This algorithm is similar in spirit to the discrete-time algorithm in Abreu, Pearce and Stacchetti [1]. However, contrary to its discrete-time counterpart, the sets in every step of the continuous-time algorithm can be computed efficiently as the numerical solution to an explicit ODE. This paper thus also contributes to the literature on computing equilibrium payoffs, providing an alternative to Judd, Yeltekin and Conklin [10] for two-player games, and an extension of Abreu and Sannikov [2] to imperfect monitoring.

Our result generalizes Theorem 2 in Sannikov not only by extending it to a more general information structure, but also by requiring less stringent assumptions on the game primitives. Most notably, we do not require that action profiles are pairwise identifiable, that is, deviations of two players may not be statistically distinguishable. Our result is thus applicable also when the signal is continuous but one-dimensional. This extension contains the important applications of a Cournot duopoly in a single homogeneous good and partnership games, where only the total revenue is observed. When action profiles fail to be pairwise identifiable, they may not be enforceable on any regular tangent hyperplane (cf. Fudenberg, Levine ans Maskin [7]). This means that players may not be willing to transfer future payoffs at any rate and instead have limiting rate at which they are willing to transfer value. While this may result in a collapse of the equilibrium payoff set to the set of static Nash payoff in some games, in other games the players locally keep transferring value at these limiting rates. In these games, the equilibrium payoff set may be flat, i.e., the boundary may have straight line segments.

The remainder of the paper is organized as follows. We introduce the continuous-time model with the general information structure in Section 2. We provide a detailed example of a partnership game with a preview of our results in Section 3. Section 4 contains the important concepts of enforceability and self-generation in our setting. In Section 5, we present our main results, an algorithm that converges to the set of equilibrium payoffs and the characterization of the resulting sets in each step. We discuss how our result relates to the main results in Sannikov [16] and Sannikov and Skrzypacz [18] in Section 6. A description of how to implement the numerical solution of our main result is presented in Section 7 and we conclude in Section 8. The vast majority of the proofs are contained in Appendices A–E.
2 THE SETTING

Consider a game where two players \(i = 1, 2\) continuously choose actions from the finite sets \(A^i\) at each point in time \(t \in [0, \infty)\). The set of all pure action profiles \(a = (a^1, a^2)\) is denoted by \(A = A^1 \times A^2\). Rather than directly observing each other’s actions, players see only the impact of the chosen actions on the distribution of a random signal. The public signal contains continuous, but noisy information modeled by a \(d\)-dimensional Brownian motion \(Z\) and informative, but infrequent observations of events of type \(y \in Y\). We assume that there are finitely many (possibly zero) different types of events in \(Y = \{y_1, \ldots, y_m\}\). Events arrive according to Poisson processes \((J^y)_{y \in Y}\) that are independent from each other and independent of the Brownian motion \(Z\). An event of type \(y\) leads to a jump in the public signal of size \(h(y)\) so that the public signal is given by \(X = Z + \sum_{y \in Y} h(y) \cdot J^y\).

The public information at time \(t\) is a \(\sigma\)-algebra \(\mathcal{F}_t\) that contains the history of the processes \(Z, (J^y)_{y \in Y}\) up to time \(t\), as well as orthogonal information that players may use as a public randomization device. Events of different types are thus observable but their underlying intensities are not. Because we study perfect public equilibria, a player’s choice of action at time \(t\) must be based solely on information in \(\mathcal{F}_t\), which is formalized in the following definition.

**Definition 2.1.** A (public) pure strategy \(A^i\) for player \(i\) is an \((\mathcal{F}_t)_{t \geq 0}\)-predictable stochastic process with values in \(A^i\).

The game primitives \(\mu : A \to \mathbb{R}^d\) and \(\lambda(y | \cdot) : A \to (0, \infty)\) determine the impact of a chosen action profile on the drift rate of the public signal and the intensity of events of type \(y \in Y\), respectively. Let \(\lambda(a) := (\lambda(y_1 | a), \ldots, \lambda(y_m | a))^\top\) denote the vector of all intensities. We assume that events of any type \(y\) are possible after any history, that is, it is a game of full support public monitoring.

**Assumption 1** (Full support). \(\lambda(y | a) > 0\) for all \(a \in A\) and all \(y \in Y\).

Because at any time \(t\), the chosen strategy profile affects the future distribution of the public signal, play of a strategy profile \(A = (A^1, A^2)\) induces a family of probability measures \(Q^A = (Q^A_t)_{t \geq 0}\), under which players observe the public signal.\(^2\)

On \([0, T]\) for any \(T > 0\), the public signal signal takes the form

\[
X_t = \int_0^t \mu(A_s) \, ds + Z^A_t + \sum_{y \in Y} h(y) \cdot J^y_t,
\]

\(^2\)Formally, \(Z\) and \((J^y)_{y \in Y}\) are defined on a probability space \((\Omega, \mathcal{F}, P)\) for a preliminary probability measure \(P\). Under \(P\), \(Z\) is a standard Brownian motion and \((J^y)_{y \in Y}\) all have intensity 1. The family \(Q^A = (Q^A_t)_{t \geq 0}\) is defined via its density process relative to \(P\), given by

\[
\exp \left( \int_0^t \mu(A_s) \, dZ_s - \int_0^t \left( \frac{1}{2} |\mu(A_s)|^2 + \sum_{y \in Y} \lambda(y | A_{s-}) - 1 \right) \, ds \right) \prod_{0 \leq s \leq t} (1 + (\lambda(y | A_{s-}) - 1) \Delta J^y_s).
\]
under $Q^A_T$, where $Z^A = Z - \int \mu(A_s) \, ds$ is a $Q^A_T$-Brownian motion describing noise in the continuous component and $J^y$ has instantaneous intensity $\lambda(y \mid A)$ under $Q^A_T$.

Remark 2.1. Note that it is possible to consider signals of the slightly more general form $X = \sigma Z + \sum_{y \in Y} h(y) J^y$ for a $k$-dimensional Brownian motion $Z$ and covariance matrix $\sigma \in \mathbb{R}^{d \times k}$ with rank $d$. Then $\sigma$ has right-inverse $\sigma^\top (\sigma \sigma^\top)^{-1}$ and the game is equivalent to the game with public signal

$$\tilde{X}_t = \int_0^t \sigma^\top (\sigma \sigma^\top)^{-1} \mu(A_s) \, ds + Z^A_t + \sum_{y \in Y} \sigma^\top (\sigma \sigma^\top)^{-1} h(y) J^y_t.$$ 

Indeed, the information carried by $\tilde{X}$ is identical to the information in $X = \sigma \tilde{X}$.

Anderson [3] and Simon and Stinchcombe [19] demonstrate that seemingly simple strategies need not necessarily lead to unique outcomes in continuous time. This is not a problem in our model because actions taken by agents do not immediately generate information. Indeed, this class of games are games of full support public monitoring: Assumption [1] in conjunction with the unbounded support of the normal distribution implies that any outcome is possible after play of any strategy profile. In public monitoring games one can identify the probability space with the path space of all publicly observable processes, and hence, a realized path $\omega \in \Omega$ leads to the unique outcome $A(\omega)$. This is analogous to discrete-time repeated games with full-support public monitoring; see Mailath and Samuelson [13] for a thorough exposition of discrete-time games.

Each player $i$ receives an expected flow payoff $g^i : \mathcal{A} \to \mathbb{R}$. In a game of imperfect information, players’ expected payoffs depend on their opponents’ actions only through their effect on the distribution of the public signal. That is, player $i$’s expected flow payoff is of the form $g^i(a) = f^i(a^i, \mu(a), \lambda(a))$.

Definition 2.2. Let $r > 0$ be a discount rate common to both players.

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

$$W^i_t(A) = \int_t^\infty e^{-r(s-t)} \mathbb{E}_{Q^A_s} [g^i(A_s) \mid \mathcal{F}_t] \, ds. \quad (1)$$

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

$$W^i_t(A) \geq W^i_t(\bar{A}^i, A^{\bar{i}}) \quad \text{a.s.}$$

holds at all times $t$, where $A^{\bar{i}}$ denotes the strategy of player $i$’s opponent.
(iii) We denote the set of all payoffs achievable by PPE by

\[ \mathcal{E}(r) := \{ w \in \mathbb{R}^2 \mid \text{there exists a PPE } A \text{ with } W_0(A) = w \text{ a.s.} \} \].

Because the weights \( re^{-r(s-t)} \) in (1) integrate up to one, the continuation value of a strategy profile is a convex combination of stage game payoffs. The set of feasible payoffs is thus given by the convex hull of pure action payoffs \( V := \text{conv } \{ g(a) \mid a \in A \} \).

By deviating to her strategy of myopic best responses \( \arg \max g^i(\cdot, A^{-i}) \), player \( i \) can ensure that her payoff in equilibrium dominates her minmax payoff \( v^i = \min_{a^{-i} \in A^{-i}} \max_{a^i \in A^i} g^i(a^i, a^{-i}) \).

The set of equilibrium payoffs is thus contained in the set of all feasible and individually rational payoffs \( V^* := \{ w \in V \mid w^i \geq \bar{y}^i \text{ for all } i \} \). Let \( A^N \subseteq A \) denote the set of static Nash equilibria and denote by \( V^N := \text{conv } \{ g(a) \mid a \in A^N \} \) the corresponding payoffs. Clearly, the inclusions \( V^N \subseteq \mathcal{E}(r) \subseteq V^* \subseteq V \) hold. Observe that \( \mathcal{E}(r) \) is convex because players are allowed to use public randomization. Indeed, for any two PPE \( A \) and \( A' \) with expected payoffs \( W_0(A) \) and \( W_0(A') \), respectively, any payoff \( \nu W_0(A) + (1 - \nu) W_0(A') \) for \( \nu \in (0, 1) \) can be attained by selecting either \( A \) or \( A' \) according to the outcome of a public randomization device at time 0.

### 3 Example of a partnership game

Consider a simple partnership game among two players, where each player continuously chooses an effort level from the set \( A^i = \{0, 1\} \) at every point in time \( t \). Players cannot observe each other’s effort levels and instead see only the realized revenue \( 2X_t \), which is subject to stochastic demand and hence reflects the chosen effort levels only imperfectly. We suppose that players share the revenue equally and are subject to a cost of effort of \( 5a^i \) so that player \( i \) receives a payoff stream of \( dX_t - 5A_t^i \, dt \), where the revenue per player satisfies

\[ dX_t = \mu(A_t) \, dt + dZ_t^A - 0.05J_t^A. \]

The continuous component \( \mu(A_t) \, dt + dZ_t^A \) of the public signal is rising on average with increasing effort levels but it is very noisy due to intraday fluctuations in demand. The abrupt information represents demand shocks, which incur a cost of 0.05 payoff units to each player. In our model, the frequency of the demand shocks is affected by players’ actions, but not the size of the shocks.

We compare two monitoring structures: (i) when players observe the continuous stream of revenue only and players’ actions do not affect the frequency of demand shocks and (ii) when players observe both the continuous stream of revenue and the demand shocks are informative of players’ behavior. For the continuous monitoring structure, let \( \mu_c(a) = 4a^1 + 4a^2 - a^1a^2 \) denote the expected
change of the continuous signal. Each player $i$’s expected flow payoff is then equal to $g^i(a) = 4(a^1 + a^2) - a^1 a^2 - 5a^i$. In the mixed monitoring structure, we suppose that players’ actions affect the intensity of demand shocks via $\lambda_m(a) = 21 - 4(a^1 + a^2) - 12a^1 a^2$. To keep the expected flow payoff unchanged, we reduce the informativeness of the continuous component of the public slightly to $\mu_m(a) = 1.05 + 3.8(a^1 + a^2) - 1.6a^1 a^2$. For both monitoring structures, the strategy profile of eternal mutual shirking is a perfect public equilibrium attaining the static Nash payoff pair $(0, 0)$.

Figure 3 compares the equilibrium payoff sets $\mathcal{E}_c(r)$ and $\mathcal{E}_m(r)$ in the continuous and mixed monitoring game, respectively. Observe that the characterization is new even in the continuous monitoring game: since the public signal is one-dimensional, action profiles are not pairwise identifiable and hence the result in Sannikov [16] does not apply. When actions are not pairwise identifiable, the boundary of the equilibrium payoff set may involve straight line segments which are not in the set of static Nash payoffs. The slopes of these line segments correspond to the limiting rates of value transfers between players or the limiting rates of exposure to the public signal that are consistent with equilibrium behavior. To support the immediate benefit of free-riding in the action profile $a = (0, 1)$, player 1 is willing to have a higher exposure to the public signal than player 2 has. If day-to-day fluctuations move unfavorably, player 1 is willing to absorb a larger share of the losses. In equilibrium, he is not willing to have an exposure that exceeds player 2’s exposure by a factor of 2.

When monitoring includes abrupt information, players have access to value burning. The action profile $(1, 1)$ of mutual effort becomes enforceable by burning 0.125
payoff units of each player upon the arrival of a demand shock. As a result, the set \( \mathcal{E}_m(r) \) has a corner outside the set of static Nash payoffs at the payoff pair \( w \) that maximizes the sum of payoffs among all equilibrium payoffs. When the continuation value is equal to \( w \), both players exert effort and continue to do so without monitoring the continuous revenue. Players trust each other to exert effort until a demand shock occurs. Because we consider games of full support public monitoring, a demand shock will occur eventually, after which play becomes dynamic again.

Figure 3 shows the drastic effect that abrupt information can have on the equilibrium payoff set. By burning only a small amount of value, the equilibrium payoff set extends much closer to the efficient frontier. The payoff bound in Sannikov and Skrzypacz [18] already shows that \( \mathcal{E}_c(r) \) remains below \( L \) for any discount rate because the action profile of mutual effort is not enforceable on the negative diagonal by observing the total revenue only. In our model, we can observe that the same payoff bound holds when abrupt information is observed but players are sufficiently impatient: the mutual punishment of burning 0.125 payoff units upon the arrival of a demand shock is consistent with equilibrium behavior only if \( r \leq 15 \). Since jumps in the continuation value are proportional to the discount rate, for \( r > 15 \) the punishment would involve a continuation payoff outside the set \( \mathcal{V}^* \) of feasible and individually rational payoffs. This shows that \( \mathcal{E}(r) \) is not lower hemi-continuous in the discount rate when players observe abrupt information.

4 Enforceability and self-generation

In games of imperfect monitoring, players’ incentives are necessarily tied to the public signal and thus, we first need to understand the dependence of the continuation value on the public signal. The following stochastic differential representation is the extension of Proposition 1 in Sannikov [16] to games with abrupt information.

**Lemma 4.1.** For a two-dimensional process \( W \) and a pure strategy profile \( A \), the following are equivalent:

(i) \( W \) is the discounted expected payoff under \( A \).

(ii) \( W \) is a bounded semimartingale which satisfies for \( i = 1, 2 \) that

\[
dW_i^t = r \left( W_i^t - g_i(A_t) \right) dt + r \beta_i^t \left( dZ_t - \mu(A_t) dt \right) + r \sum_{y \in Y} \delta_i^t(y) \left( dJ_y^t - \lambda(y | A_t) dt \right) + dM_i^t
\]

for a martingale \( M_i^t \) (strongly) orthogonal to \( Z \) and all \( J_y^t \) with \( M_0^i = 0 \), predictable processes \( \beta_i^t \) and \( \delta_i^t(y) \), \( y \in Y \), satisfying \( \mathbb{E}_{\mathbb{Q}} \left[ \int_0^T |\beta_i^t|^2 dt \right] < \infty \) and \( \mathbb{E}_{\mathbb{Q}} \left[ \int_0^T |\delta_i^t(y)|^2 \lambda(y | A_t) dt \right] < \infty \) for any \( T \geq 0 \).
The process $r^i$ is the sensitivity of player $i$’s continuation value to the continuous component of the public signal, and the processes $r^i(y)$ are the impacts on player $i$’s continuation value when an event of type $y \in Y$ occurs. To keep the notation succinct, we use $\delta^i$ to refer to the row vector $(\delta^i(y_1), \ldots, \delta^i(y_m))$ containing the impacts on $i$’s continuation payoffs for all types of events and $\lambda(a)$ to refer to the column vector $(\lambda(y_1|a), \ldots, \lambda(y_m|a))$ containing the intensities of all events.

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, Pearce and Stacchetti [1]. The representation in (2) shows that in continuous-time games, the continuation value is linear in the public signal and hence, so is the continuation promise. Similarly to Sannikov [16] and Sannikov and Skrzypacz [18], the incentive compatibility conditions take the following form.

**Definition 4.2.** An action profile $a \in A$ is **enforceable** if there exists a continuation promise $(\beta, \delta)$ with $\beta = (\beta^1, \beta^2)^\top \in \mathbb{R}^{2 \times d}$ and $\delta = (\delta^1, \delta^2)^\top \in \mathbb{R}^{2 \times m}$ such that for every player $i$, the sum of expected instantaneous payoff rate $g^i(a)$ and promised continuation rate $\beta^i \mu(a) + \delta^i \lambda(a)$ is maximized in $a^i$. That is, for every $\tilde{a}^i \not\in A^i \backslash \{a^i\}$,

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

We say such a pair $(\beta, \delta)$ enforces $a$. A continuation promise $(\beta, \delta)$ strictly enforces $a$ if (3) holds with strict inequality for both players $i$ and every deviation $\tilde{a}^i \in A^i \backslash \{a^i\}$. A strategy profile is **enforceable** if and only if it takes values in enforceable action profiles almost everywhere.

If players keep their promises, then the continuation promises used to enforce strategy profile $A$ are, in fact, the sensitivities of its continuation value to the public signal. Therefore, no player has an incentive to deviate from the strategy profile $A$ as formalized in the following lemma, which is a generalization of Proposition 2 in Sannikov [16] to our setting.

**Lemma 4.3.** A strategy profile is a PPE if and only if $(\beta^1, \beta^2)$ and $(\delta^1(y), \delta^2(y))_{y \in Y}$ related to $W(A)$ by (2) enforce $A$.

Lemmas 4.1 and 4.3 motivate how we construct equilibrium profiles in continuous time—as the solution to (2) subject to the enforceability constraint in (3). Because there is no notion of a terminal value, the theory of backward stochastic differential equations cannot be applied to infinitely repeated games. Instead, we use time-homogeneity of repeated games to construct forward solutions similarly to discrete time: because equilibria are subgame perfect, only continuation promises that lead to a continuation value in $E(r)$ can be used to provide incentives. The set of equilibrium payoffs is thus self-generating. The following is the definition of a self-generating payoff set in a continuous-time setting.
Definition 4.4. A payoff set $\mathcal{W} \subset \mathbb{R}^2$ is called self-generating if for every $w \in \mathcal{W}$, there exists an enforceable strategy profile $A$, enforced by $(\beta, \delta)$ related to $W(A)$ by (2) such that $W_0(A) = w$ a.s. and $W_\tau(A) \in \mathcal{W}$ a.s. for every stopping time $\tau$.

The following lemma is the analogue of Theorem 1 in Abreu, Pearce and Stacchetti [1]. The proof of Lemma 4.5 is easily derived from the proof of Lemma 2 in Bernard and Frei [6], as that proof works more generally for signals given by any Lévy process.

Lemma 4.5. The set $\mathcal{E}(r)$ is the largest bounded self-generating set.

5 Characterization of equilibrium payoffs

Because the continuation value of a PPE must remain in the equilibrium payoff set at all times, the law of motion (2) places certain restrictions on the continuation promises $(\beta, \delta)$ that can be used to provide incentives. Figure 4 illustrates these restrictions at the boundary of a self-generating set: Upon the arrival of an event $y \in Y$, each player $i$ is punished/rewarded with an amount $r\delta_i(y)$, which induces a jump in the players’ continuation values. For punishments/rewards consistent with equilibrium behavior, the continuation value after the event is still in $\mathcal{E}(r)$. Because abrupt information cannot be anticipated, it is necessary that $W_t + r\delta_t(y)$ holds at all times.

The drift term $r\left( W_t - g(A_t) - \delta_t \lambda(A_t) \right) dt$ in (2) describes the expected movement of the continuation payoff, conditional on the fact that no event occurs. Because the continuation value of a PPE remains in the equilibrium payoff set, the drift has to point towards the interior of the set, i.e., $g(A_t) + \delta_t \lambda(A_t)$ is separated from $\mathcal{E}(r)$ by tangent hyperplanes. This is similar to the equilibrium construction in Fudenberg, Levine and Maskin [7], where payoffs are decomposed as convex combinations of current-period payoffs $g(a)$ outside of $\mathcal{E}(r)$ and expected continuation payoffs in $\mathcal{E}(r)$. In games with abrupt information, the drift points away from $g(A_t) + \delta_t \lambda(A_t)$ instead of $g(A_t)$ because players anticipate the expected impact of rewards/punishments on their continuation payoff. Finally, the diffusion term $r\beta_t \left( dZ_t - \mu(A_t) \right) dt$ in (2) related to the continuous component of the public signal has to be tangential to the boundary. Otherwise the continuation value would immediately escape $\mathcal{E}(r)$ since Brownian motion has unbounded variation.

To formalize these informational restrictions, we introduce the following notation to describe the boundary locally. For a convex set $\mathcal{W}$ and any payoff pair $w \in \partial \mathcal{W}$, denote by $\mathcal{N}_w(\mathcal{W}) := \{ N \in S^1 \mid N^\top(w - v) \geq 0 \text{ for all } v \in \mathcal{W} \}$ the set of all outer-pointing normal vectors to $\partial \mathcal{W}$ at $w$, where the unit circle $S^1$ is the set of all directions. If the boundary is continuously differentiable at $w$, the normal vector is unique and we denote it by $N_w$. The restrictions on the continuation promise $(\beta, \delta)$ used to provide incentives at the boundary of a self-generating set $\mathcal{W}$ are the following:
Figure 4: Because $E(r)$ is self-generating, the continuation value $W$ of a PPE can never escape the set $E(r)$. At the boundary $\partial E(r)$, the drift rate $W_t - g(A_t) - \delta_t \lambda(A_t)$ thus has to point towards the interior of the set. Moreover, the diffusion $r \beta_t (dZ_t - \mu(A_t) dt)$ has to be tangential to $\partial E(r)$ as the continuation value would escape $E(r)$ otherwise due to the unbounded variation of Brownian motion. Finally, an event of type $y \in Y$ incurs a jump in the continuation value of size $r \delta_t(y)$. Since $W$ cannot jump outside of $E(r)$, it is necessary that $W_t + r \delta_t(y) \in E(r)$ for every $y \in Y$.

(i) Jumps within the set: $w + r \delta(y) \in \mathcal{W}$ for every $y \in Y$,

(ii) Inward-pointing drift: $N^T(g(a) + \delta \lambda(a) - w) \geq 0$ for any $N \in \mathcal{N}_w(\mathcal{W})$,

(iii) Tangential volatility: $N^T \beta = 0$ for any $N \in \mathcal{N}_w(\mathcal{W})$.

Sannikov [16] shows that when information arrives continuously, only local information about the boundary is necessary to describe the equilibrium payoff set in the same way that an ordinary differential equation (ODE) uses only local information to describe a function. Crucial in this respect is that Brownian information arrives continuously. Therefore, only the informational restrictions (ii) and (iii) have to be satisfied at the boundary, which are local restrictions on the use of information that depend on the geometry of the equilibrium payoff set only through the normal vector $N_w$ at $w$. This gives rise to an explicit description of the boundary through an ODE in the state $(w, N_w)$. When information arrives discontinuously as well, such a local description is no longer possible as the first restriction is a global restriction that depends on the precise shape of $E(r)$. A local description of the equilibrium payoff set using restrictions (i)–(iii) thus involves $E(r)$, creating a non-trivial fixed-point problem. We solve it with an iteration over the arrival times of infrequent events, where restriction (iii) is relaxed to jumps landing in a fixed payoff set $\mathcal{W}$.

### 5.1 Algorithm

In this section, we present a continuous-time analogue of the algorithm in Abreu, Pearce and Stacchetti [1] and show how it can be used to solve the aforementioned fixed-point problem. We begin with the definition of relaxed self-generation.
Definition 5.1.

(i) Let $\sigma_n$ denote the occurrence of the $n^{th}$ infrequent event and let $W \subseteq \mathbb{R}^2$ be a convex and compact set. We say that a payoff set $X$ is $W$-relaxed self-generating if for every $w \in X$, there exists an enforceable strategy profile $A$, enforced by $(\beta, \delta)$ related to $W(A)$ by (2) such that $W_0(A) = w$ a.s., $W_{\tau}(A) \in X$ a.s. for every stopping time $\tau < \sigma_1$, and $W_{\sigma_1}(A) \in W$ a.s.

(ii) For a convex and compact set $W \subseteq \mathbb{R}^2$, let $B_r(W)$ denote the largest $W$-relaxed self-generating set. Observe that this is well defined since the convex hull of two $W$-relaxed self-generating sets is again $W$-relaxed self-generating.

The operator $B_r$ is closely related to the standard set operator in discrete time introduced in Abreu, Pearce and Stacchetti [1]. Payoffs in $B_r(W)$ can be attained by an enforceable strategy profile with continuation promise at time $\sigma_1$ that lies in $W$. If $W \subseteq B_r(W)$, then the payoff $W_{\sigma_1}$ can be attained by an enforceable strategy profile until the arrival of the next event and so on. The following lemma characterizes the relation between the operator $B_r$ and self-generation.

Lemma 5.2. Let $W \subseteq V$. If $W$ is self-generating, then $W \subseteq B_r(W)$. If $W \subseteq B_r(W)$, then $B_r(W)$ is self-generating.

An $n$-fold application of $B$ to a set $W$ thus ensures that the continuation value after the first $n$ events is in $W$. Because Poisson processes have only countably many jumps, taking the limit as $n$ goes to infinity covers all events. We thus obtain the following algorithm to compute $E(r)$ iteratively.

Proposition 5.3. Let $W_0 = V^*$ and $W_n = B(W_{n-1})$ for $n \geq 1$. Then $(W_n)_{n \geq 0}$ is decreasing in the set-inclusion sense with $\bigcap_{n \geq 0} W_n = E(r)$.

This algorithm is similar to Abreu, Pearce and Stacchetti [1]. However, unlike its discrete-time counterpart, the boundary of the resulting set at each step of the iteration is characterized by an ODE. Since the condition on the incentives provided through jumps is fixed, the characterization of $B_r(W)$ is explicit, which makes the computation efficient. The resulting algorithm thus provides an alternative to discrete-time implementation of Judd, Yeltekin, and Conklin [10].

The arrival of infrequent events carries many similarities to the public signal in discrete time. There exists an embedding into a discrete-time game with periods of random length equal to $\sigma_n - \sigma_{n-1}$, where $\sigma_0 = 0$. At time $\sigma_n$, the signal $y$ is observed if and only if an event $y$ happens at that time. Because $Y$ is finite in our setting, the discrete information does not satisfy a bang-bang result and the jump may go into the interior of $W$. However, since two or more jumps of independent Poisson processes happen at the same time with probability 0, interior points can be attained by a public randomization device with values on the boundary before the next event occurs. It is thus sufficient that the condition on jumps landing in $W$ is verified on the boundary. This makes a characterization of $B_r(W)$ via $\partial B_r(W)$ feasible.
5.2 Stationary payoffs

The simplest $\mathcal{W}$-relaxed self-generating payoff sets consist of only a single payoff pair $w$. These payoff pairs thus have to be attainable by a strategy profile whose continuation value remains in $w$ until an infrequent event occurs. Equation (2) implies that there has to exist an action profile $a$ that can be enforced using only incentives $\delta_0$ relying on the abrupt information with $w = g(a) + \delta_0\lambda(a)$. The drift rate of the constant strategy profile $A \equiv a$ is thus zero and $W(A)$ remains in $w$ until the first event occurs. If, in addition, $w + r\delta_0(y) \in \mathcal{W}$ is satisfied for every event $y \in Y$, then $\{w\}$ is $\mathcal{W}$-relaxed self-generating. We call a payoff pair $\mathcal{W}$-stationary if it admits such a decomposition. We denote by $\mathcal{S}_r(\mathcal{W})$ the set of all $\mathcal{W}$-stationary payoffs. Because $\mathcal{B}_r(\mathcal{W})$ is the largest self-generating set, the set of stationary payoffs is contained in $\mathcal{B}_r(\mathcal{W})$ as formalized by the following lemma.

**Lemma 5.4.** $\mathcal{S}_r(\mathcal{W}) \subseteq \mathcal{B}_r(\mathcal{W})$.

**Proof.** Let the stationary payoff $w$ be attained by $g(a) + \delta_0\lambda(a)$ such that $(0, \delta_0)$ enforces $a$ and $w + r\delta_0(y) \in \mathcal{W}$ for every $y \in Y$. The constant strategy profile $A \equiv a$ is thus enforced by the continuation promise $(\beta, \delta)$ with $\beta \equiv 0$ and $\delta \equiv \delta_0$. A solution $W_y$ to (2) starting in $w$ with $M \equiv 0$ and $A, \beta, \delta$ as given thus has neither drift nor diffusion term. It follows that $W_y$ remains in $w$ until the arrival time $\sigma_1$ of the first event $y$, at which point $W_y$ jumps to $\mathcal{W}$ since $W_{\sigma_1} = W_{\sigma_1} - r\delta_{\sigma_1}(y) = w + r\delta_0(y) \in \mathcal{W}$. \qed

The set of stationary payoffs can be computed easily with the following procedure. The computation involves only finite unions and intersections as well as projections, direct sums, and direct differences of convex sets. Observe first that the set $\Psi_a := \{\delta \in \mathbb{R}^{2 \times m} \mid (0, \delta) \text{ enforces } a\}$ is the intersection of the finitely many half-spaces defined in (3), hence $\Psi_a$ is convex. Since $w = g(a) + \delta\lambda(a)$, the condition on jump $y$ going into $\mathcal{W}$ can be expressed as $g(a) + \delta(\lambda(a) + e_y) \in \mathcal{W}$, where $e_y \in \mathbb{R}^m$ is the basis vector corresponding to event $y \in Y$. The set of all $\delta \in \Psi_a$, for which the reward/punishment after an occurrence of event $y \in Y$ lies in $\mathcal{W}$, is computed as

$$\Psi_y(a) := \Psi_a(\lambda(a) + re_y) \cap (\mathcal{W} - g(a)).$$

Using $w = g(a) + \delta\lambda(a)$ again, we obtain

$$\mathcal{S}_r(\mathcal{W}) = \bigcup_{a \in A} \left( g(a) + \left( \cap_{y \in Y} \Psi_y(a) \right) \lambda(a) \right).$$

\footnote{For two sets $U, V \subset \mathbb{R}^k$, we denote the direct sum and direct difference between the two sets by $U \pm V := \{x \in \mathbb{R}^k \mid \exists u \in U, v \in V \text{ with } x = u \pm v\}$, respectively. Similarly, for two sets of matrices $U \in \mathbb{R}^{k \times j}, V \in \mathbb{R}^{l \times k}$ we denote by $UV$ the set of pointwise multiplication. If $U = \{u\}$ is a singleton, we also write $u \pm V$ and $uV$ for $\{u\} \pm V$ and $\{u\}V$, respectively.}
5.3 Characterization of $\partial B_r(\mathcal{W})$

As we have argued in the introductory paragraph of this section, at the boundary of $B_r(\mathcal{W})$, players must play an action profile such that the continuation value has inward-pointing drift, tangential volatility, and jumps within $\mathcal{W}$. We call such an action profile $a \in A$ restricted-enforceable and denote by

$$\Xi_a(w, N, r, \mathcal{W}) := \left\{ (\beta, \delta) \middle| (\beta, \delta) \text{ enforces } a, N^\top (g(a) + \delta \lambda(a) - w) \geq 0, N^\top \beta = 0, \text{ and } w + r \delta(y) \in \mathcal{W} \text{ for every } y \in Y \right\}$$

the set of all continuation promises $(\beta, \delta)$ that restricted-enforce $a$ for $(w, N, r, \mathcal{W})$.

We characterize $\partial B_r(\mathcal{W})$ in two steps. Similarly to Sannikov [16], the diffusion term makes the boundary of $\partial B_r(\mathcal{W})$ smooth where incentives from the continuous component of the public signal are required. The amount of tangential transfers needed to structure incentives is inversely related to the curvature of $\partial B_r(\mathcal{W})$, giving rise to an ODE. However, in a setting with abrupt information, the continuous component of the public signal may not be necessary to provide incentives. Let

$$\mathcal{G}_r(\mathcal{W}) := \{ w \in \partial B_r(\mathcal{W}) \mid \exists a \in A \text{ with } (0, \delta) \in \Xi_a(w, N, r, \mathcal{W}) \forall N \in \mathcal{N}_w(B_r(\mathcal{W})) \}$$

denote the set of all points where incentives can be provided using the abrupt information only. We also say that such a pair $(a, \delta)$ decomposes $w$ with continuations in $\mathcal{W}$ or that $w$ is decomposable by $a$. These payoff pairs may be attained with strategy profiles, whose continuation value have no diffusion term and hence $B_r(\mathcal{W})$ may have corners in $\mathcal{G}_r(\mathcal{W})$. Stationary payoffs are examples of such points if $S_r(\mathcal{W})$ overlaps with $\partial B_r(\mathcal{W})$. However, $\mathcal{G}_r(\mathcal{W})$ may contain also non-stationary payoff pairs that are attained by strategy profiles whose continuation values have a non-zero drift term.

For any action profile $a \in A$, let $\Psi_a^i$ denote the set of $\delta \in \mathbb{R}^m$, for which continuation promise $(0, \delta)$ provides sufficient incentives to player $i$ to support play of $a$, i.e., $(0, \delta)$ is a solution to (3) for player $i$. We make the following non-empty interiority assumption, stating that if the discontinuous component of the public signal is sufficient to incentivize player $i$ to play $a^i$, given $a^{-i}$, then player $i$ can be strictly incentivized to play $a^i$, given $a^{-i}$, using only the discontinuous information. Observe that the assumption is satisfied for generic $g$ and $\lambda$.

**Assumption 2.** Suppose $\Psi_a^i$ has non-empty interior for any $a \in A$.

**Remark 5.1.** Assumption 2 is a generalization of Assumption 2.(i) in Sannikov [16] as it reduces to a unique best response assumption when $Y = \emptyset$. Note however, that we do not require action profiles to be pairwise identifiable, which is required in Sannikov [13]. We are able to relax this condition despite our general framework by analyzing continuity properties of incentives that are “optimal” on the boundary; see Appendix B and also Appendix E.2 to a lesser extent.
Without pairwise identifiability, action profiles may be restricted-enforceable in some directions but not in others. We show in the appendix that if an action profile \(a\) is restricted-enforceable along a convergent sequence of directions \((N_n)_{n \geq 0}\), then \(a\) is also restricted-enforceable along the limiting direction \(\lim_{n \to \infty} N_n\) unless the limiting direction is a coordinate direction in \(\{\pm e_1, \pm e_2\}\). Assumption 2 guarantees that this holds true also in coordinate directions for the relevant action profiles. This allows us to construct equilibria with the help of the following lemma.

**Lemma 5.5.** Suppose that Assumptions 1 and 2 hold. Let \(C\) be a continuously differentiable curve, oriented by the Gauss map \(w \mapsto N_w\) such that:

(i) There exist measurable selectors \(a^*, \delta^*, \beta^*\) on \(C\) such that the selections satisfy \((\beta^*(w), \delta^*(w)) \in \Xi_{a^*}(w, N_w, r, \mathcal{W})\) for any \(w \in C\) and the curvature at any point \(w \in C\) is given by

\[
\kappa(w) = \frac{2N_w^\top(g(a^*(w)) + \delta^*(w)\lambda(a^*(w)) - w)}{r\|\beta^*(w)\|^2}.
\]

(ii) \(C\) is a closed curve or both of its endpoints are contained in \(\mathcal{B}_r(W)\).

Then \(C \subseteq \mathcal{B}_r(W)\) and the solution \(W\) to (2) with \(A = a^*(W), \delta = \delta^*(W), \beta = \beta^*(W)\), and \(M \equiv 0\) remains on \(C\) until an endpoint of \(C\) is reached or an event occurs.

The above lemma enables us to construct enforceable strategy profiles that remain on a curve with curvature (5). While the presence of abrupt information creates many technical challenges that we address in Appendices B and C, the intuition behind Lemma 5.5 is similar to Sannikov [16] because the diffusion term is the main driver behind Lemma 5.5. The strategy profile constructed in Lemma 5.5 and its continuation promises are Markovian in the continuation value. At any point \(w\) on the curve, incentives provided through the continuous component of the public signal are parallel to the curve. Because of the infinitesimally fast oscillation of Brownian motion, the tangential volatility leads to an outward-pointing drift; see Figure 5. It follows from Itô’s formula that the outward drift is proportional to the curvature and the square of the tangential incentives. For the constructed strategy profile, this outward drift is precisely counteracted by the inward-pointing drift so that

\[
\frac{r^2}{2}\kappa(w)\|\beta_t^2\| = N_w^\top(w - g(A_t) - \delta_t\lambda(A_t))
\]

and the continuation value remains on the curve \(C\). Because \(\delta\) is chosen such that \(W_t + r\delta_t(y) \in \mathcal{W}\) at all times, it is guaranteed that the continuation value jumps to \(\mathcal{W}\) after the occurrence of any event \(y \in Y\). This construction of enforceable strategy profiles will be used several times in the characterization of \(\partial \mathcal{B}_r(W)\).

17
Figure 5: Incentives related to the continuous component of the public signal lead to tangential volatility of the continuation value. Because Brownian motion has unbounded variation, the tangential volatility leads to a second-order outward drift. At points where the curvature is larger, the tangential volatility leads to a stronger outward drift. For continuation promise \((\beta,\delta)\) in Lemma 5.5, this outward drift is precisely counteracted by the inward drift 
\[
r\beta_t(dZ_t - \mu(A_t) \, dt)
\]
so that the continuation value remains on \(C\) until an end point is reached or an infrequent event occurs.

We first characterize the set of payoffs \(G_r(W)\) on the boundary of \(B_r(W)\), where incentives can be provided through the discontinuous information only. The defining characteristics of points in \(G_r(W)\) are that incentives are provided through the abrupt information only such that the drift rate goes towards the interior of \(B_r(W)\) and the continuation payoff after the occurrence of an event is in \(W\). The following lemma establishes that at most one of these defining conditions holds strictly.

**Lemma 5.6.** Suppose that Assumptions 1 and 2 are satisfied and that \(W\) has non-empty interior. For any \(w \in G_r(W)\), it is impossible that there exists \((a, \delta)\) such that two of the following conditions hold simultaneously:

(i) \((0, \delta)\) strictly enforces \(a\),
(ii) \(N^T(g(a) + \delta \lambda(a) - w) > 0\) for some outward normal \(N \in N_w(B_r(W))\),
(iii) \(w + r \delta(y) \in \text{int} W\) for every \(y \in Y\).

**Proof.** Suppose that there exists such a pair \((a, \delta)\). Since \(W\) and \(\Psi_a\) have non-empty interior, there exists \(\delta'\) sufficiently close to \(\delta\) such that all three conditions (i)–(iii) hold simultaneously. Since the conditions are strict, they all hold for \(v\) and \(N'\) sufficiently close to \(w\) and \(N\). Because of condition (i), there exists \(\phi \in \mathbb{R}^d\) sufficiently small such that \((T_\phi, \delta')\) enforces \(a\) for any direction \(T \in S^1\). Let \(C_{w_0,\phi}\) be a solution to (5) with initial value \((w_0, N)\) and selectors \(a^*(w) = a, \delta^*(w) = \delta', \beta^* = T_\nu \phi\), where \(T_\nu\) is the tangent vector to \(C_{w_0,\phi}\) at a point \(v\). Because of conditions (ii) and (iii), choosing \(w_0 \notin \text{cl} B_r(W)\) sufficiently close to \(w\) and choosing \(||\phi||\) sufficiently small guarantees that \(C\) enters the interior of \(B_r(W)\) on both sides of \(w_0\) as illustrated in the left panel of Figure 6. Lemma 5.5 thus shows that \(w_0 \in B_r(W)\), a contradiction. 

A consequence to Lemma 5.6 is that corners are either stationary or that incentives are binding for at least one player \(i\) such that \(w + r \delta(y) \in \partial W\) for some event \(y \in Y\).
verify whether points in the boundary of \( K \) are perfect monitoring. They find that extremal points in the discrete-time analogue of \( \Pi \) are contained in the set \( \partial B_r(W) \) outside of \( G_r(W) \) as depicted in the right panel.

Indeed, if condition (ii) of Lemma 5.6 is violated and \( N^T(g(a) + \delta \lambda(a) - w) = 0 \) for all outward normal vectors \( N \) to \( \partial B_r(W) \) at a corner \( w \), then \( w = g(a) + \delta \lambda(a) \) and hence \( w \) is stationary. If condition (ii) is satisfied, incentives have to be binding and \( w + r \delta(y) \in \partial W \) for at least one \( y \in Y \). Therefore, there must exist an action profile \( a \) such that the corner is on the boundary of the set

\[
K_{r,a}(W) := \{ w \mid \exists \delta \in \Psi_a \text{ with } w + r \delta(y) \in W \text{ for every } y \in Y \}.
\]

Corners of \( G_r(W) \) are thus contained in the set \( K_r(W) = \partial \bigcup_{a \in A} K_{r,a}(W) \).\footnote{Observe that \( \partial K_{r,a}(W) \) is reasonably nice since \( K_{r,a} \) is convex: for any \( w_0, w_1 \in K_{r,a}(W) \), there exist \( (\delta_0, \delta_1) \) with \( w_k + r \delta_k(y) \in W \) for every \( y \in Y \) and \( k = 1, 2 \). Define \( w_\gamma := \gamma w_1 + (1 - \gamma) w_0 \) and \( \delta_\gamma := \gamma \delta_1 + (1 - \gamma) \delta_0 \) and observe that for every \( \gamma \in [0,1] \), \( (0, \delta_\gamma) \) enforces \( a \) and \( w_\gamma + r \delta_\gamma(y) \in W \) for every \( y \in Y \) by convexity of \( \Psi_a \) and \( W \).} A perturbation argument in Appendix C shows that this is, in fact, true for all extremal points of \( G_r(W) \).

**Proposition 5.7.** Suppose that Assumptions 1 and 2 hold. Then any corners of \( G_r(W) \) and any extremal points of continuously differentiable line segments in \( G_r(W) \) are contained in \( S_r(W) \cup K_r(W) \). Moreover, if such a point \( w \) outside of \( S_r(W) \) is decomposable by \( a \in A \), then \( N_w(B_r(W)) \subseteq N_w(K_{r,a}(W)) \).

In contrast to stationary points, which are all contained in \( B_r(W) \), points in \( K_r(W) \) are merely candidates for points in \( G_r(W) \). We show in Section 6 how to verify whether points in \( K_r(W) \) are indeed contained in \( G_r(W) \) and how to compute the boundary of \( B_r(W) \). The key ingredient is the restriction on outward normal vectors. In the next section, we provide a refinement of these candidates specifically for the equilibrium payoff set.

The fact that corners are either stationary points or that incentives have to be binding at corners is similar in spirit to Abreu and Sannikov [2], who find an improvement on the algorithm in Judd, Yeltekin, and Conklin [10] for two-player games with perfect monitoring. They find that extremal points in the discrete-time analogue of \( B(W) \) are attainable either by a stationary strategy profile, or by a current-period
action profile with binding incentives. This is the same in our setting as incentives are binding where the boundary is continuously differentiable also outside of $\mathcal{G}_r(\mathcal{W})$. The following result is a complete characterization of $\mathcal{B}_r(\mathcal{W})$.

**Theorem 5.8.** Suppose that Assumptions 1 and 2 hold and that $\mathcal{W} \subseteq \mathcal{V}^*$ is compact and convex with non-empty interior. Then $\mathcal{B}_r(\mathcal{W})$ is the largest closed convex subset of $\mathcal{V}^*$ such that $\text{ext} \ G_r(\mathcal{W}) \subseteq \mathcal{S}_r(\mathcal{W}) \cup \mathcal{K}_r(\mathcal{W})$ and $\partial \mathcal{B}_r(\mathcal{W}) \setminus \mathcal{G}_r(\mathcal{W})$ is continuously differentiable with curvature at almost every point $w$ given by

$$\kappa(w) = \max_{a \in \mathcal{A}} \max_{(\beta, \delta) \in \Xi(a, w, r, N_w, \mathcal{W})} \frac{2N_w^\top(g(a) + \delta \lambda(a) - w)}{r \|\beta\|^2},$$

(6)

where we set $\kappa(w) = 0$ if the maxima are taken over empty sets.

Outside of $\mathcal{G}_r(\mathcal{W})$, the curvature of the boundary is of the same form as in Lemma 5.5, where the maximum is taken over all restricted-enforceable action profile and all restricted-enforcing continuation promises. The intuition behind this is the same as in Sannikov [16]: if the curvature of $\partial \mathcal{B}_r(\mathcal{W})$ was smaller than the maximal curvature $\kappa^*$ in (5), there would exist an enforceable strategy profile $A^*$ whose continuation value remains on a curve $\mathcal{C}$ with a curvature larger than $\partial \mathcal{B}_r(\mathcal{W})$. Starting at a payoff pair $v$ slightly outside of $\mathcal{B}_r(\mathcal{W})$, the curve would intersect $\partial \mathcal{B}_r(\mathcal{W})$ and reach payoffs in the interior of $\mathcal{B}_r(\mathcal{W})$; see the right panel in Figure 6. Since the continuation value of $A^*$ remains on $\mathcal{C}$ until the arrival of the first even by Lemma 5.5, the continuation value either jumps to $\mathcal{W}$ or reaches an end point of $\mathcal{C}$, where it can be attained by the continuation value of a restricted-enforceable strategy profile. This shows that $\mathcal{C} \cup \mathcal{B}_r(\mathcal{W})$ is $\mathcal{W}$-relaxed self-generating, contradicting maximality of $\mathcal{B}_r(\mathcal{W})$. Note that this argument requires continuity of (5) in initial conditions, which we establish in Appendix 3. Similarly, the curvature of $\partial \mathcal{B}_r(\mathcal{W})$ cannot be larger than the maximum curvature in (5) because otherwise, there would exist no restricted enforceable strategy profile that remains in $\mathcal{B}_r(\mathcal{W})$, contradicting the fact that $\mathcal{B}_r(\mathcal{W})$ is $\mathcal{W}$-relaxed self-generating. There are many details needed to make this argument rigorous, which can be found in Appendices C and E.2.

Even though the curvature is characterized only at almost every point on the boundary, a solution is unique with the additional requirement that it be continuously differentiable. This implies that $\partial \mathcal{B}_r(\mathcal{W})$ is twice continuously differentiable almost everywhere, which is important for the numerical solution of (6) as numerical procedures rely on discretizations. We will elaborate on the numerical implementation in Section 7.

### 5.4 Characterization of $\mathcal{E}(r)$

Since $\mathcal{B}_r$ preserves compactness due to Theorem 5.8, it follows from Proposition 5.3 that $\mathcal{E}(r)$ is compact. An application of Theorem 5.8 for $\mathcal{W} = \mathcal{E}(r)$ thus provides a fixed-point characterization of $\mathcal{B}(\mathcal{E}(r)) = \mathcal{E}(r)$.  

20
In this section, we provide a sharper characterization of $\mathcal{G}_r(\mathcal{E}(r))$. For general sets $\mathcal{W}$, the set $\mathcal{G}_r(\mathcal{W})$ may include points, at which incentives provided through the abrupt information necessarily include outward jumps. In the limit as $\mathcal{B}_r$ is applied infinitely often, these incentives cannot be efficient to support equilibrium behavior: since $\mathcal{E}(r)$ is convex, jumps at the boundary have to satisfy $N^T \delta(y) \leq 0$ for every $y \in Y$ and every $N \in \mathcal{N}_w(\mathcal{E}(r))$. We begin by defining the set of payoffs that can be decomposed using inward jumps regardless of target set $\mathcal{W}$. For any action profile $a \in \mathcal{A}$ and any direction $N \in S^1$, let

$$H_a(N) := \left\{ w \in \mathbb{R}^2 \left| \exists \delta \in \Psi_a \text{ with } N^T \delta(y) \leq 0 \text{ for every } y \in Y \right. \right. \left. \text{ and } N^T(g(a) + \delta \lambda(a) - w) \geq 0 \right\}.$$  

be the half-space of payoffs that can be decomposed with inward jumps with respect to the direction $N$. Observe that $H_a(N)$ cannot extend past $g(a)$, i.e., $g(a)$ is either on the boundary or outside of $H_a(N)$. Let $D_a := \{ N \in S^1 \mid H_a(N) \neq \emptyset \}$ denote the set of directions, with respect to which $a$ can be decomposed using inward jumps only. The set

$$Q_a := \bigcap_{N \in D_a} H_a(N).$$

is the set of all payoffs that can be decomposed with inward-pointing jumps. Since static Nash profiles can be decomposed without any jumps at all, it follows that $Q_a = \{ g(a) \}$ for any static Nash profile $a$. The following lemma follows immediately from the definition of $Q_a$, stating that any payoff $w \in \mathcal{G}_r(\mathcal{W})$ that is decomposed by $a$ is either contained in $Q_a$ or requires outward jumps to be enforced.

**Lemma 5.9.** Let $w \in \mathcal{G}_r(\mathcal{W})$ be decomposed by $a \in \mathcal{A}$. Then either $w \in Q_a$ or any $\delta \in \Psi_a$ with $w + r\delta(y) \in \mathcal{W}$ for every $y \in Y$ and $N^T(g(a) + \delta \lambda(a) - w) \geq 0$ for every $y \in Y$ and $N \in \mathcal{N}_w(\mathcal{B}_r(\mathcal{W}))$ satisfies $N^T \delta(y) > 0$ for some $y \in Y$ and some $N$ outward normal $N$.

Since outward jumps cannot support equilibrium behavior, we obtain the following fixed-point characterization of $\mathcal{E}(r)$ as a consequence to Proposition 5.3, Theorem 5.8, and Lemma 5.9.

**Corollary 5.10.** Under Assumptions 1 and 2, $\mathcal{E}(r)$ is the largest closed subset of $\mathcal{W}^*$ such that $\mathcal{G}_r(\mathcal{E}(r)) \subseteq \bigcup_{a \in \mathcal{A}} (\mathcal{S}_r(\mathcal{W}) \cup \partial \mathcal{K}_{r,a}(\mathcal{E}(r))) \cap Q_a$ and $\partial \mathcal{E}(r) \setminus \mathcal{G}_r(\mathcal{E}(r))$ is continuously differentiable with curvature at almost every point $w$ given by

$$\kappa(w) = \max_{a \in \mathcal{A}} \max_{(\beta, \delta) \in \Xi(a(w, r, N_w, \mathcal{E}(r)))} \frac{2N_w^T(g(a) + \delta \lambda(a) - w)}{r \| \beta \|^2},$$

where we set $\kappa(w) = 0$ if the maxima are taken over empty sets. Straight line segments on $\partial \mathcal{E}(r)$ are bounded on at least one side by a payoff in $\mathcal{S}_r(\mathcal{E}(r))$. 

21
At a first glance, it may seem that the characterization has become more difficult with the exclusion of outward jumps. This is only a notational difficulty. The refinement allows us to exclude many points that would have to be considered as potential corners and extremal points in $G_r(\mathcal{E}(r))$. The following algorithm clarifies that we can exclude all of these points straight from the beginning, leading to a faster computation of $\mathcal{E}(r)$ through a refinement of the operator $B_r$.

**Definition 5.11.** For a compact and convex set $\mathcal{W} \subseteq \mathcal{V}^*$ with non-empty interior, let $\tilde{B}_r(\mathcal{W})$ denote the largest closed subset of $\mathcal{V}^*$ such that

(i) its boundary is a solution to (6) at all points that are not decomposable, and

(ii) all points on the boundary that are decomposable by some $a \in \mathcal{A}$ are contained in the set $(S_r(\mathcal{W}) \cup \partial \mathcal{K}_{r,a}(\mathcal{W})) \cap Q_a$. We refer to these points by $\tilde{G}_r(\mathcal{W})$.

**Proposition 5.12.** Let $\mathcal{W}_0 = \mathcal{V}^*$ and $\mathcal{W}_n := \tilde{B}_r(\mathcal{W}_{n-1})$ for $n \geq 1$. Then $(\mathcal{W}_n)_{n \geq 0}$ is decreasing in the set-inclusion sense with $\bigcap_{n \geq 0} \mathcal{W}_n = \mathcal{E}(r)$.

### 6 Discussion

Figure 3 in Section 3 shows that the presence of abrupt information can drastically increase efficiency for impatient players. Even though abrupt information may be used only to destroy value in equilibrium in the sense that $\partial \mathcal{E}(r)$ moves further away from the stage-game payoff $g(a)$, a small amount of value burning may greatly reduce the amount of tangential transfers necessary. The optimality equation (7) captures this tradeoff. To better understand the different impacts of the two types of information, we perform a comparative statics exercise in the example of the partnership game of Section 3. For any $\gamma \in [0, 1]$, consider the games where information arrives according to

$$\mu_\gamma(a) = 4\gamma(a^1 + a^2) - \gamma a^1 a^2$$

and

$$\lambda_\gamma(a) = (1 - \gamma)(21 - 4(a^1 + a^2) - 12a^1 a^2).$$

Suppose that the expected flow payoff $g'(a) = 4(a^1 + a^2) - a^1 a^2 - 5a^1$ is fixed for all monitoring structures so that the equilibrium payoff sets are comparable on the same scale. For low values of $\gamma$, the continuous information is at its most informative, whereas the abrupt information gets more informative as $\gamma$ increases. Figure 7 shows the equilibrium payoff sets for $r = 5$ and $\gamma \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$.

As the figure shows, neither type of information is strictly better than the other, but the two types of information serve different purposes. Abrupt information may enlarge the set of enforceable action profiles by attaching large punishments/rewards to infrequent events. Information through these events, however, arrives too sparsely for players to continuously adjust incentives: outside the set of stationary payoffs, incentives through the continuous signal is necessary. The more informative the continuous component of the signal is, the more significant transfers between players are possible and the equilibrium payoff set gets wider. The SDE (2) shows that
the punishments/rewards required to enforce action profiles are proportional to the discount rate. Since the continuation value has to remain in $E(r)$ after a jump, fewer incentives can be provided through infrequent events as the discount rate $r$ increases: the observed information arrives too infrequently relative to players’ patience to be relevant to provide incentives. In this partnership game, the demand shocks are not informative enough to enforce mutual effort if and only if $\gamma < (1 - r/15)$.

When the continuous changes in intraday revenue are completely uninformative, the equilibrium payoff set is contained on the positive diagonal between the static Nash payoff $(0, 0)$ and the highest symmetric payoff $w = (1.875, 1.875)$. Since the only restricted-enforceable action profiles are mutual effort and mutual shirking, all equilibrium profiles are versions of (forgiving) grim-trigger strategies: If the expected continuation payoff pair is $w$, both players exert effort until a demand shock occurs. To incentivize the exertion of effort, both players need to burn $\delta^i(y) = [r/8, 1.875]$ payoff units upon the arrival of a demand shock. In the most severe equilibrium punishment, both players burn their entire continuation payoff and the continuation value is absorbed in $(0, 0)$ forever. This corresponds to an unforgiving grim trigger strategy profile where players revert to the static Nash profile forever after bad news is observed. If the punishment is less severe, the continuation value jumps to the relative interior of $E(r)$ and players enter a temporary punishment phase, where they play the static Nash profile for a limited amount of time. During the punishment phase, the continuation value gradually increases as the punishment phase approaches its end. This is seen in Figure 2: while players shirk, the continuation value moves away from the static Nash payoff deterministically on a straight line, i.e., along $E(r)$ until
the continuation value reaches \( w \) again. When the expected discounted payoff is equal to \( w \), players resume exerting effort until the next demand shock occurs, that is, they play a forgiving grim-trigger strategy profile. Note that the length of the punishment phase is deterministic since continuation value moves deterministically during the punishment phase as seen in (2). When \( r = 15 \), the unforgiving grim-trigger strategy profile and eternal shirking are the only equilibrium profiles; for \( r > 15 \), the static Nash profile is the unique PPE.

Let us discuss how our result relates to the payoff bound in Sannikov and Skrzypacz [18]. Since \( \mathcal{E}(r) \) is closed by Corollary 5.10 at any payoff pair \( w \) on the boundary, incentives have to be provided with a continuation promise \((\beta, \delta)\) such that players use the continuous information to transfers value tangentially, the punishments/rewards after any event \( y \in Y \) satisfy \( w + r\delta(y) \in \mathcal{E}(r) \), and \( N^\top(g(a) + \delta\lambda(a) - w) \geq 0 \). Because the equilibrium payoff set is convex, it follows that \( N_w^\top\delta(y) \leq 0 \) for any event \( y \in Y \), i.e., incentives through abrupt information burn value relative to the normal direction \( N_w \) at \( w \). Incentives at the boundary thus satisfy the informational assumptions in Sannikov and Skrzypacz [18], hence their payoff bound applies also to our setting. Note, however, that the algorithms in Propositions 5.3 and 5.12 cannot be started with their payoff bound \( M \) as it is unknown whether or not \( B(M) \) is contained in \( M \) and whether the induced sequence is decreasing. Nevertheless, knowing the payoff bound can help with the implementation.

7 Computation

In this section, we illustrate how to implement Theorem 5.8 and the algorithms in Proposition 5.3 and 5.12 numerically. We also compare the speed of convergence of the two algorithms in the partnership game of Section 3. To compute the set \( B_r(W) \), on calculates first the set of stationary payoffs with (4) and then find the
largest solution to the optimality equation (6) that includes all stationary payoffs. In principle, this is achieved similarly as in Sannikov [16], but we additionally need to account for corners and straight line segments.

7.1 Computing $B_r(W)$ for arbitrary sets $W$

Since $B_r(W)$ is convex, we parametrize the boundary via tangential angle $\theta$. Let $w(\theta)$ denote the set of payoffs in $B_r(W)$ with normal vector $N(\theta) = (\cos(\theta), \sin(\theta))^\top$. Note that $w(\theta)$ is unique where the curvature of $\partial B_r(W)$ is strictly positive, hence one can solve

$$\frac{dw(\theta)}{d\theta} = \frac{T(\theta)}{\kappa(\theta)}$$

numerically, where $T(\theta) = (-\sin(\theta), \cos(\theta))^\top$ and $\kappa(\theta) = \kappa(w(\theta))$ is given by the optimality equation (6). If the maximization in (6) is unbounded for some action profile $a$, we check whether $w(\theta) \in \partial K_{r,a}(W)$ and $N(\theta) \in N_{w(\theta)}(K_{r,a}(W))$. If this is the case, it is possible that $B_r(W)$ has a corner at $w(\theta)$. If the maximization is taken over empty sets, it is possible that $\partial B_r(W)$ has a straight line segment. Because $B_r(W)$ is the largest solution to (6), we search for the maximal angle at corners and the longest straight line segments, for which a closed solution to (6) exists.

We illustrate this procedure by computing $B_r(V^*)$ in the partnership example of Section 3. We first compute the set of stationary payoffs as illustrated in the left panel of Figure 9. Since $S_r(V^*)$ overlaps with $\partial V^*$, it is necessary that $S_r(V^*) \cap \partial V^*$ lies on the boundary of $B_r(V^*)$. Points $A$ and $B$ in Figure 9 thus serve as potential starting points for solving (6). To handle potential corners of $B_r(V^*)$, we next compute the sets $K_{r,a}(V^*)$ for all action profiles $a \in A$. In the partnership example of Section 3, there is only one type of infrequent events, hence $K_{r,a}(V^*)$ can be computed as $V^* - r \Psi_a$. 

Figure 9: The left panel shows the set of stationary payoffs in the mixed-monitoring game for $r = 10$. The right panel shows the candidates for payoff pairs in $G_{10}(V^*)$. 

25
The right panel of Figure 9 shows the intersection $\mathcal{K}_{r,a}(V^*) \cap V^*$ for the different stage game action profiles $a \in \mathcal{A}$. However, none of the points on $\mathcal{K}_r(V^*)$ could possibly be on $\partial \mathcal{B}_r(V^*)$ because the normal vectors point into the wrong direction: from Proposition 5.7, we know that the a normal vector to $\mathcal{B}_r(V^*)$ has to be a normal vector to $\mathcal{K}_{r,a}(V^*)$ as well for action profile $a$ that decomposes $w$. This is impossible since $A$ and $B$ are on $\partial \mathcal{B}_r(V^*)$. We conclude that $\mathcal{B}_r(V^*)$ has corners only at stationary points. We thus search for the maximal angle at $A$, for which a solution connects to either $B$ or the upper right segment of $\mathcal{S}_r(V^*)$. If a solution connects to the other segment of $\mathcal{S}_r(V^*)$, the solution from there to $C$ is given by symmetry.

If $\mathcal{S}_r(W) \cap \partial V^* = \emptyset$, starting points of symmetric games may be found by searching over the positive diagonal with initial angles $\pi/4$ or $5\pi/4$. If the game is not symmetric, one may use an iterative procedure as in Section 8 of Sannikov [16].

7.2 Computing the equilibrium payoff set

The equilibrium payoff set can be computed with the algorithms in Propositions 5.3 and 5.12. Let first $(\mathcal{W}_n)_{n \geq 0}$ denote the sequence in Proposition 5.3 where $\mathcal{W}_n = \mathcal{B}_r(\mathcal{W}_{n-1})$ for any $n$. We iterate the computation as delineated in the previous section until the difference between sets $\mathcal{W}_n$ and $\mathcal{W}_{n-1}$ is sufficiently small. The left panel of Figure 10 illustrates the convergence to $\mathcal{E}(10)$. One can see that in every step of the iteration, we obtain stationary payoffs on $\partial V^* \setminus V^N$ that cannot be contained in the equilibrium payoff set because they necessarily involve outward jumps. The speed of convergence can be improved by implementing the algorithm of Proposition 5.12 instead, where we take the intersection of the stationary payoff set with $\mathcal{Q}_a$ for every action profile $a \in \mathcal{A}$. Since $\mathcal{Q}_a = \{g(a)\}$ for static Nash profiles, this improves the speed of convergence as illustrated in the right panel of Figure 10.
This paper studies a class of continuous-time two-player games with imperfect public observation, where information may arrive both continuously through the observation of a noisy signal, and discontinuously as the occurrences of infrequent but informative events. For this class of games, we characterize the equilibrium payoff set and show how to compute it efficiently. In the presence of abrupt information, this involves an algorithm that relies on a continuous-time analogue to the standard set-operator in Abreu, Pearce and Stacchetti [1]. The notion of relaxed self-generation is a technique that may useful in subsequent research on continuous-time games that involve discontinuous information, such as stochastic games with discrete states.

The application of these games are numerous, as in many situations there is a very intuitive decomposition of the available information into continuous and discontinuous information. The characterization of the equilibrium payoffs set is new even when the signal is continuous but fails the widely assumed pairwise identifiability condition. Our methods thus allow the computation of the equilibrium payoff set in games where the signal is one-dimensional such as a partnership game or a duopoly in a single homogeneous good. Because of the quantitative nature of the result, the impact of information on equilibrium payoffs can be measured precisely, paving the way for future research on information revelation: a company may choose to publicly disclose certain information (make it continuously observable) or keep the information from the public until the media reports on it (abrupt information). Because continuous and discontinuous information have fundamentally different impacts on equilibrium payoffs, a strategic company may prefer one over the other and act accordingly.

A Proofs of auxiliary results in the main text

A.1 Dynamics of the continuation value and continuation promises

The proof is similar to the proof of Lemma 1 in Bernard and Frei [6], with the following additional arguments for the jumps. Because \((J^y)_{y \in Y}\) are pairwise orthogonal and orthogonal to \(Z\), the stable subspace generated by \(Z\) and \((J^y)_{y \in Y}\) is the space of all stochastic integrals with respect to these processes (Theorem IV.36 in Protter [15]). Therefore, we obtain the unique martingale representation property for a square-integrable martingale by Corollary 1 to Theorem IV.37 in [15]. That is, for a bounded \(\mathcal{F}_T\)-measurable random variable \(w^i_T\), there exists an \(\mathcal{F}_0\)-measurable \(c^i_T\), predictable processes \((\beta^i_{t,T})_{0 \leq t \leq T}\), \((\delta^i_{t,T}(y))_{0 \leq t \leq T}\) for all \(y \in Y\) with

\[
\mathbb{E}_{Q^A_t} \left[ \int_0^T \beta^i_{t,T} \, dt \right] < \infty \quad \text{and} \quad \mathbb{E}_{Q^A_t} \left[ \int_0^T |\delta^i_{t,T}(y)|^2 \lambda(y \mid A_t) \, dt \right] < \infty \quad \text{and a } Q^A_t\text{-martingale}
\]
$M^i$ orthogonal to $Z_i$, $(J^y)_{y \in Y}$, with $M^i_0 = 0$ such that

$$w^i_T = c^i_T + \int_0^T r_\gamma^i_t (dZ_t - \mu(A_t) dt) + \sum_{y \in Y} \int_0^T r_\delta^i_t(y) (dJ^y_t - \lambda(y | A_t) dt) + M^i_{T,T}.$$  

The remainder of the equivalence (i) $\iff$ (ii) works analogously to the arguments in Bernard and Frei [6], with the following additional arguments for the jumps: Since $\int r_\delta^i(y) (dJ^y_t - dt)$ has bounded jumps by construction for any $y \in Y$, it follows that $\int_t^\infty e^{-r(s-t)} \delta^i_s(y) (dJ^y_s - \lambda(y | A_s) ds)$ is a BMO-martingale under $Q^A_u$ up to any time $u \in (t, \infty)$. Assumption 1 implies that the jumps of $\lambda(y | A_u) - 1) \Delta J^y_t$ in Footnote [2] are bounded from below by $-1 + \varepsilon$ for any $y \in Y$. Therefore, Remark 3.3 and Theorem 3.6 in Kazamaki [12] imply that $\int_t^\infty e^{-r(s-t)} \delta^i_s(y) (dJ^y_s - \lambda(y | A_s) ds)$ is a BMO-martingale under $Q^A_u$. \(\Box\)

**Proof of Lemma 4.3.** This works analogously to the second statement of Lemma 1 in Bernard and Frei [6]. \(\Box\)

### A.2 Convergence of the algorithm

In this appendix we prove the convergence of the algorithm in Proposition 5.3 to $\mathcal{E}(r)$. We first prove the straightforward lemma that $\mathcal{B}_r(\mathcal{W})$ is monotone in $\mathcal{W}$.

**Lemma A.1.** Let $\mathcal{W} \subseteq \mathcal{W}'$. Then $\mathcal{B}_r(\mathcal{W}) \subseteq \mathcal{B}_r(\mathcal{W}')$.

**Proof.** Any payoff $w \in \mathcal{B}_r(\mathcal{W})$ can be attained by an enforceable strategy profile with continuation value $W$ such that $W \in \mathcal{B}_r(\mathcal{W})$ up to the first jump time $\sigma_1$ with $W_{\sigma_1} \in \mathcal{W}$ a.s. Since $\mathcal{W} \subseteq \mathcal{W}'$, it follows that $\mathcal{B}_r(\mathcal{W})$ is $\mathcal{W}'$-relaxed self-generating. The statement now follows form maximality of $\mathcal{B}_r(\mathcal{W}')$. \(\Box\)

We are now ready to prove Lemma 5.2 and Propositions 5.3 and 5.12.

**Proof of Lemma 5.2.** We first show that $\mathcal{W} \subseteq \mathcal{B}_r(\mathcal{W})$ implies that $\mathcal{B}_r(\mathcal{W})$ is self-generating. To that end, fix a payoff pair $w \in \mathcal{B}_r(\mathcal{W})$ arbitrary and let it be attained by an enforceable strategy profile $A$ such that $W$ remains in $\mathcal{B}_r(\mathcal{W})$ on $[0, \sigma_1)$ and $W_{\sigma_1} \in \mathcal{W}$ a.s. This implies that $W_{\sigma_1}(A) \in \mathcal{B}_r(\mathcal{W})$, hence there exists an enforceable strategy profile $\tilde{A}$ attaining $W_{\sigma_1}(A)$ such that $W$ remains in $\mathcal{B}_r(\mathcal{W})$ up to the second jump time $\sigma_2$ with $W_{\sigma_2} \in \mathcal{W}$ a.s. Therefore, the concatenation $\tilde{A} := A_1\mathbb{1}_{[0, \sigma_1]} + \tilde{A}_{-\sigma_1}\mathbb{1}_{[\sigma_1, \infty]}$ is enforceable and remains in $\mathcal{B}_r(\mathcal{W})$ up to $\sigma_2$. Because Poisson processes have only finitely many jumps on any finite time interval and $Y$ is finite, a countable iteration of this procedure will lead to an enforceable strategy profile, whose continuation value remains in $\mathcal{B}_r(\mathcal{W})$ forever. This shows that $\mathcal{B}_r(\mathcal{W})$ is self-generating.

For the converse, observe that self-generation implies that for any $w \in \mathcal{W}$, there exists an enforceable strategy profile with continuations that remain in $\mathcal{W}$. In particular, $W_{\sigma} \in \mathcal{W}$ a.s. It follows that $\mathcal{W} \subseteq \mathcal{B}_r(\mathcal{W})$ by maximality of $\mathcal{B}_r(\mathcal{W})$. \(\Box\)
Proof of Proposition 5.12. Let \( W \subseteq W' \). Since \( S_r \) and \( K_{r,a} \) are monotone operators, it follows that \( \bar{S}_r(W) \subseteq \bar{S}_r(W') \). Since \( W \subseteq W' \), the ODE (6) is solved over a larger set of controls in \( B_r(W) \) than in \( B_r(W') \) and thus \( B_r(W) \subseteq B_r(W') \). Since \( \bar{B}_r(W) \subseteq \mathcal{V}_r \) by definition, this shows that \( W_n \) is decreasing in the set-inclusion sense. Lemma 5.9 asserts that \( \bar{B}_r(\mathcal{E}(r)) = \mathcal{E}(r) \). Monotonicity of \( \bar{B}_r \) thus shows that \( \mathcal{E}(r) \subseteq W_n \) for any \( n \). The sequence \( (W_n)_{n \geq 0} \) thus converges to a limit \( W_\infty \) that contains \( \mathcal{E}(r) \). It remains to show that \( W_\infty \) is not larger than \( \mathcal{E}(r) \). Let \( (W'_n)_{n \geq 0} \) denote the sequence of iterated applications of \( B_r \) to \( \mathcal{V}_r \). Since \( \bar{B}_r(W) \subseteq B_r(W) \) for any set \( W \), it follows that \( W_1 \subseteq W'_1 \) and hence \( W_{n+1} \subseteq \bar{B}_r(W_n') \subseteq B_r(W_n') = W'_{n+1} \) by monotonicity. Thus, \( \mathcal{E}(r) \subseteq W_n \subseteq W'_n \) for any \( n \) and hence \( W_n \rightarrow \mathcal{E}(r) \) as \( n \rightarrow \infty \). \( \square \)

B Regularity of the optimality equation

The purpose of this appendix is to prove that the optimality equation is locally Lipschitz continuous at almost every point, so that locally, it admits a unique solution. We show in Lemma C.4 that \( \partial B_r(W) \setminus G_r(W) \) is \( C^1 \), hence \( \partial B_r(W) \setminus G_r(W) \) is the unique \( C^1 \) solution to the optimality equation. For any fixed \( r > 0 \), \( a \in \mathcal{A} \), and closed and convex \( W \subseteq \mathcal{V} \), consider the optimality equation in the following form:

\[
\kappa_a(w, N) = \max_{(\phi, \delta) \in \mathcal{I}_a(w, N, r, \mathcal{V})} \frac{2N^T(g(a) + \delta \lambda(a) - w)}{r \|\phi\|^2}.
\]

We start by reducing the two-variable optimization problem to a one-variable optimization by expressing the control \( \phi \) in terms of \( \delta \). For player \( i = 1, 2 \), define

\[
\mathcal{I}_a^i(N, \delta^i) := \{ \phi \in \mathbb{R}^d \mid (T^i \phi, \delta^i) \text{ satisfies (3) for player } i \}
\]

for any direction \( N \in S^1 \) and \( \delta^i \in \mathbb{R}^{\mathcal{V}_i} \). Because \( \mathcal{I}_a^i(N, \delta^i) \) is the intersection of closed half-spaces, it is a (possibly unbounded or empty) closed convex polytope. Therefore, so is \( \Phi_a(N, \delta) := \mathcal{I}_a^1(N, \delta^1) \cap \mathcal{I}_a^2(N, \delta^2) \), the set of all vectors \( \phi \in \mathbb{R}^d \) such that \( (T \phi, \delta) \) enforces \( a \). Let \( \phi(a, N, \delta) \) denote the vector of smallest length in \( \Phi_a(N, \delta) \).
Lemma B.1. Fix $a \in \mathcal{A}$. Then $(N, \delta) \mapsto \phi(a, N, \delta)$ is locally Lipschitz continuous where $\Phi_a(N, \delta) \neq \emptyset$ and $N$ is different from a coordinate direction.

In an intermediate step, we will show that the set-valued map $(N, \delta) \mapsto \Phi_a(N, \delta)$ is locally Lipschitz continuous for $N$ different from coordinate directions. We refer to Aubin and Frankowska \cite{Aubin2009} for a detailed overview of set-valued maps and their properties and state here only the most central property.

Definition B.2. A set-valued map $G : x \mapsto G(x)$ is said to be Lipschitz continuous if $G(x) \subseteq G(\tilde{x}) + K \|x - \tilde{x}\| B_1(0)$ for some constant $K$.

Proof of Lemma B.1. Let $\mathcal{I}_i(\delta^i) := \{ \beta \in \mathbb{R}^d \mid (\beta, \delta^i) \text{satisfies } (3) \}$ for player $i$ be the solution set to (3) for player $i$ and observe that it is a closed convex polytope. Its hyperfaces have normal vectors $\Delta \mu_{ij} := \mu(a) - \mu(a^j_i, a^{-i})$, where $a_i^1, \ldots, a_i^{m_i}$ is an enumeration of $A_i \setminus \{a_i^i\}$. The parameter $\delta^i$ determines the location of these hyperfaces. Observe that a change from $\delta^i$ to $\tilde{\delta}^i$ shifts face $j_i$ by $(\tilde{\delta}^i - \delta^i) \Delta \lambda_{ji}$, where $\Delta \lambda_{ji} := \lambda(a) - \lambda(a^j_i, a^{-i})$. Therefore, the triangle inequality implies that

$$\mathcal{I}_a(\delta) \subseteq \mathcal{I}_a(\tilde{\delta}) + B_1(0) \sum_{j_i = 1, \ldots, m_i} \|\Delta \lambda_{ji}\| \|\tilde{\delta}^i - \delta^i\|,$$

i.e., $\mathcal{I}_a(\delta^i)$ is Lipschitz continuous in $\delta^i$. It is clear that $\mathcal{I}_a(N, \delta^1) = \frac{1}{T} \mathcal{I}_a(\delta^i)$ for $i = 1, 2$ is locally Lipschitz continuous in $(N, \delta)$ for $N$ different from coordinate directions. Since the stretching does not affect the direction of the normal vectors, the normal vectors of $\mathcal{I}_a(N, \delta^i)$ are constant, hence $(N, \delta) \mapsto \Phi_a(N, \delta) = \mathcal{I}_a(N, \delta^1) \cap \mathcal{I}_a(N, \delta^2)$ is locally Lipschitz continuous by Lemma E.2. The statement now follows from the following Lemma.

Lemma B.3. Let $f(x, y)$ be a single-valued Lipschitz continuous map and let $G(x)$ be a set-valued (locally) Lipschitz continuous map. Then $h(x) = \max_{y \in G(x)} f(x, y)$ is (locally) Lipschitz continuous.

Proof. For any $x$, let $U$ be a neighbourhood of $x$ such that $G$ is Lipschitz continuous on $U$ with Lipschitz constant $K_G$. Let $x_1, x_2 \in U$ and suppose without loss of generality that $h(x_1) \geq h(x_2)$. Let $K_f$ be the Lipschitz constant of $f$. Then $f(x_1, y) \leq f(x_2, y) + K_f \|x_2 - x_1\|$ for any $y$, hence

$$h(x_1) - h(x_2) \leq K_f \|x_2 - x_1\| + \max_{y \in G(x_1)} f(x_2, y) - \max_{y \in G(x_2)} f(x_2, y)$$

$$\leq K_f \|x_2 - x_1\| + \max_{y \in G(x_2)} f(x_2, y) - \max_{y \in G(x_2)} f(x_2, y)$$

$$\leq K_f \|x_2 - x_1\| + K_f K_G \|x_2 - x_1\|.$$

\qed
Lemma B.1 significantly simplifies the constraints in the maximization in (9) because we are left with a maximization over δ only. We will prove regularity of a slightly more general form of the optimality equation suitable for the proofs in Appendix C. Instead of requiring \( w + r \delta(y) \in \mathcal{W} \) for every \( y \in Y \) and some fixed set \( \mathcal{W} \), we will require that \( \delta \in \mathcal{D}(w) \) for an affine, compact- and convex-valued correspondence \( w \mapsto \mathcal{D}(w) \subseteq \mathbb{R}^{2 \times m} \). We study the optimality equation of the form

\[
\kappa_a(w, N) = \max_{\delta \in \Psi_a(w,N,r,D)} \frac{2N^T(g(a) + \delta \lambda(a) - w)}{r \|\phi(a, N, \delta)\|^2},
\]

(10)

where \( \Psi_a(w, N, r, \mathcal{D}) := \{ \delta \in \mathcal{D}(w) \mid \Phi_a(N, \delta) \neq \emptyset \text{ and } N^T(g(a) + \delta \lambda(a) - w) \geq 0 \} \).

Observe that (10) is equal to (9) for \( \mathcal{D}(w) = (\mathcal{W} - w)^m/r \), for which \( \delta \in \mathcal{D}(w) \) is equivalent to \( w + r \delta(y) \in \mathcal{W} \) for every \( y \in Y \).

**Lemma B.4.** Let \( w \mapsto \mathcal{D}(w) \) be affine, compact- and convex-valued. Then for any \( a \in \mathcal{A} \), the map \( (w, N) \mapsto \Psi_a(w, N, r, \mathcal{D}) \) is compact- and convex-valued. Moreover, it is locally Lipschitz continuous for \( N \) different from coordinate directions.

**Proof.** Identify \( \mathbb{R}^{2 \times |Y|} \) with \( \mathbb{R}^{2|Y|} \) by setting \( \delta \approx (\delta^1, \delta^2) \). For any subset \( \mathcal{W} \) of \( \mathbb{R}^2 \), define \( \mathcal{W}^{|Y|} := \{ (\delta^1, \delta^2) \in \mathbb{R}^{2|Y|} \mid (\delta^1(y), \delta^2(y))^T \in \mathcal{W} \forall y \in Y \} \). Let \( \Psi_a(w, N) \) and \( \mathcal{J}_a(N) \) denote the set of all \( \delta \), for which \( N^T(g(a) + \delta \lambda(a) - w) \geq 0 \) and \( \Phi_a(N, \delta) \neq \emptyset \), respectively, are satisfied. We begin by showing that \( \mathcal{J}_a(N) \) is closed and convex, hence so is \( \Psi_a(w, N, r, \mathcal{D}(w)) = \mathcal{J}_a(N) \cap \Psi_a(w, N) \cap \mathcal{D}(w) \) as intersection of such sets. Indeed, let \( \delta_1, \delta_2 \in \mathcal{J}_a(N) \). Then there exist \( \phi_1, \phi_2 \) such that \( (\delta_j, T\phi_j) \) for \( j = 1, 2 \) satisfy (3) for every \( \tilde{a}_i \in \mathcal{A} \setminus \{a_i\} \) and \( i = 1, 2 \). By linearity of (3), so does \( (\delta_\nu, T\phi_\nu) \) for \( \nu \in [0, 1] \), where we set \( \delta_\nu := \nu \delta_1 + (1 - \nu) \delta_2 \) and \( \phi_\nu := \nu \phi_1 + (1 - \nu) \phi_2 \). This shows that \( \delta_\nu \in \mathcal{J}_a(N) \), i.e., \( \mathcal{J}_a(N) \) is convex. Let \( (\delta_n)_{n \geq 0} \) be a sequence in \( \mathcal{J}_a(N) \). Then there exists \( (\phi_n)_{n \geq 0} \) such that \( (\delta_n, T\phi_n) \) satisfies (3). Since the inequalities in (3) are not strict, \( (\lim_{n \to \infty} \delta_n, T\lim_{n \to \infty} \phi_n) \) satisfies (3), hence \( \lim_{n \to \infty} \delta_n \in \mathcal{J}_a(N) \) and \( \mathcal{J}_a(N) \) is closed. Compactness of \( \Psi_a(w, N, r, \mathcal{D}(w)) \) now follows because \( \mathcal{D}(w) \) is compact.

For \( \phi \in \mathbb{R}^d \), introduce the auxiliary sets \( \mathcal{J}_a(N, \phi) \) of those \( \delta \in \mathbb{R}^{2|Y|} \) for which \( (T\phi, \delta) \) enforces \( a \). For \( i = 1, 2 \), let \( a_{i}^{1}, \ldots, a_{i}^{m_i} \) be an enumeration of \( \mathcal{A}^i \setminus \{a_i^i\} \) and abbreviate \( \Delta \mu_{j_i}^i := \mu(a) - \mu(a_{j_i}^i, a^{-i}) \) and \( \Delta \lambda_{j_i}^i := \lambda(a) - \lambda(a_{j_i}^i, a^{-i}) \) as in the proof of Lemma B.1 Then \( \mathcal{J}_a(N, \phi) \) is a closed convex polytope, whose hyperfaces have normal vectors

\[
\begin{pmatrix}
\Delta \lambda_{j_1}^1 \\
0
\end{pmatrix}, \quad j_1 = 1, \ldots, m_1, \quad \begin{pmatrix}
0 \\
\Delta \lambda_{j_2}^2
\end{pmatrix}, \quad j_2 = 1, \ldots, m_2
\]

(11)

\footnote{For a set \( \mathcal{X} \subseteq \mathbb{R}^2 \), we denote by \( \mathcal{X}^m \) the \( m \)-fold product of that set, i.e., \( \mathcal{X}^m := \left\{ x \in \mathbb{R}^{2m} \mid (x_k, x_k) \in \mathcal{X} \text{ for every } k = 1, \ldots, m \right\} \).}
and $N$ only determines the position of these hyperfaces. Thus, similarly as in the proof of Lemma B.1, $N \mapsto J_a(N, \phi)$ is Lipschitz continuous with a Lipschitz constant that depends only on $\Delta \mu_{i}$.

In particular, the Lipschitz constant of $N \mapsto J_a(N, \phi)$ is uniformly bounded in $\phi$.

Observe that $w \mapsto \Psi_a(w, N)$ and $w \mapsto J_a(N)$ are affine functions. Lipschitz continuity in $w$ thus follows from Lemma E.1. To obtain Lipschitz continuity in $N$, we verify the conditions of Lemma E.2. Let $W$ be a bounded Lipschitz constants is Lipschitz again. From the fact that the arbitrary union of Lipschitz continuous maps with uniformly bounded Lipschitz constants is Lipschitz continuous, it follows that for any $\phi \in \mathbb{R}^d$, $N \mapsto \Psi_a(w, N) \cap J_a(N, \phi) \cap D(w)$ is Lipschitz continuous in $N$ for non-coordinate directions $\ell$. Since $D(w) \subseteq W$ is constant in $N$ and the intersection of a Lipschitz continuous map with a convex and compact set is Lipschitz continuous, it follows that for any $\phi \in \mathbb{R}^d$, $N \mapsto \Psi_a(w, N) \cap J_a(N, \phi) \cap D(w)$ is Lipschitz continuous. Local Lipschitz continuity of $N \mapsto \Psi_a(w, N, r, D)$ now follows from the fact that the arbitrary union of Lipschitz continuous maps with uniformly bounded Lipschitz constants is Lipschitz again. $\square$

So far we have shown that (10) is locally Lipschitz continuous for almost every direction $N$, where $\phi(a, N, \delta)$ is well defined and bounded away from 0. Define

$$E_a(r, D) := \{(w, N) \in \mathbb{R}^2 \times S^1 | \Psi_a(w, N, r, D) \neq \emptyset\}$$

$$\Gamma_a(r, D) := \{(w, N) \in \mathbb{R}^2 \times S^1 | \exists \delta \in \Psi_a(w, N, r, D) \text{ with } \phi(a, N, \delta) = 0\}$$

and $\Gamma(r, D) := \bigcup_{a \in A} \Gamma_a(r, D)$. Denote by $P := \mathbb{R}^2 \times \{\pm e_1, \pm e_2\}$ the set of points $(w, N) \in \mathbb{R}^2 \times S^1$ with a coordinate normal vector $N$.

Lemma B.5. Let $D$ be an affine, compact- and convex-valued correspondence. If a sequence $(w_n, N_n)_{n \geq 0}$ converges to $(w, N) \notin P$ such that $\Psi_a(w_n, N_n, r, D) \neq \emptyset$ for all $n \geq 0$, then $\Psi_a(w, N, r, D) \neq \emptyset$.

Proof. Let $\delta_n \in \Psi_a(w_n, N_n, r, D)$. Because $D(w_n)$ is uniformly bounded by $D(V)$, the sequence $(\delta_n)_{n \geq 0}$ is uniformly bounded as well. Therefore, $(\delta_n)_{n \geq 0}$ converges along a subsequence $(n_k)_{k \geq 0}$ to some finite limit $\delta$ with $N^T (g(a) + \delta \lambda(a) - w) \geq 0$. Since $D$ is closed-valued and Lipschitz continuous, $\delta(y) \in D(w)$ for every $y \in Y$. It remains to show that $\Phi_a(N, \delta) \neq \emptyset$. Suppose towards a contradiction that the converse is true. Then closedness of $T_\delta(N, \delta)$ for $i = 1, 2$ implies that $T_\delta(N, \delta_1)$ and $T_\delta(N, \delta_2)$ are strictly separated. By continuity, $T_{\delta_1}(N_{\ell_k}, \delta_{1\ell_k})$ and $T_{\delta_2}(N_{\ell_k}, \delta_{2\ell_k})$ are separated as well for $k$ sufficiently large, a contradiction. $\square$
Corollary B.6. For any $a \in A$ and $\varepsilon \geq 0$, $E_a(r, D) \cup P$ and $\Gamma_a(r, D)$ are closed. Therefore, so is $\Gamma(r, D)$.

Proof. Indeed, $\Gamma_a(r, D)$ is closed since $0 \in \Phi_a(N, \delta)$ for some $N \in S^1$ if and only if $0 \in \Phi_a(N, \delta)$ for all $N \in S^1$. □

Proposition B.7. Suppose that $W$ has non-empty interior and that Assumption 3 is satisfied. For any affine, compact- and convex-valued correspondence $D$,

$$\kappa(w, N) = \max_{a \in A} \max_{\delta \in \Psi_a(w, N, r, D)} \frac{2N^T(g(a) + \delta \lambda(a) - w)}{r \|\phi(a, N, \delta)\|^2} \tag{12}$$

is locally Lipschitz continuous outside of $\Gamma(r, D)$, except where $(w, N)$ leaves or enters $E_a(r, D)$ of the maximizing action profile $a$. Here, we interpret $\kappa(w, N) = 0$ on $\bigcap_{a \in A} E_a(r, D)^c$, i.e., where the maxima are taken over empty sets.

When we refer to a solution to (12), we will always mention explicitly with respect to which map $D$ (12) is being solved.

Proof. Suppose first that $N$ is not a coordinate direction, that is, $(w, N) \in E_a(r, D) \setminus (\Gamma_a(r, D) \cup P)$. We first show local Lipschitz continuity of $\kappa_a$ in (10) for fixed $a \in A$. Since $\Gamma_a(r, D)$ is closed by Corollary B.6 there exists an open neighbourhood $U$ of $(w, N)$ bounded away from $\Gamma_a(r, D) \cup P$. Therefore, $\inf_{N, \delta} \|\phi(a, N, \delta)\| \geq c$ and hence the function that is maximized in the right hand side of (10) is Lipschitz continuous on $U$ by Lemma B.4. It follows that $\kappa_a$ is Lipschitz continuous by Lemmas B.3 and B.4. Because (12) is the maximum over finitely many functions $\kappa_a$, it is Lipschitz continuous except where $(w, N)$ leaves the domain of the maximal function $\kappa_a$.

Suppose now that $N$ is a coordinate direction. Let $a$ denote the maximizing action profile at $(w, N)$. Because we show Lipschitz continuity only where maximizers in (12) do not change, we may assume that $a$ maximizes $\kappa_a$ in a neighborhood of $(w, N)$. If $\kappa_a$ is identically zero in a neighborhood, the statement is trivial, hence suppose that a solution is strictly curved. Let $\mathcal{I}_a^i(N, \delta)$ be defined as in the proof of Lemma B.1 and let $\phi^i(a, N, \delta)$ denote the shortest vector in $\mathcal{I}_a^i(N, \delta)$. Suppose that $N$ is coordinate for player $i$, i.e., $N \in \{\pm e_i\}$. Since $W$ and $\Psi_a^i$ have non-empty interior by Assumption 2, then $\mathcal{I}_a^{-i}(N, \delta)$ and $\phi^{-i}(a, N, \delta)$ are locally Lipschitz continuous. Since $\Psi_a^i$ have non-empty interior by Assumption 2 player $i$ can be strictly incentivized to play $a^i$. Thus, there exists $\delta$ such that $(\phi^{-i}(a, N, \delta), \delta)$ enforces $a$. Since $W$ has non-empty interior, there exists such $\delta$ with $w + r\delta(y) \in Y$ for every $y \in Y$. Let

$$\tilde{\kappa}_a(w, N) = \max_\delta \frac{2N^T(g(a) + \delta \lambda(a) - w)}{r \|\phi^i(a, N, \delta)\|^2},$$

where we maximize over all such $\delta$ and observe that $\tilde{\kappa}$ is locally Lipschitz continuous. Since $\tilde{\kappa}_a \leq \kappa_a$ in a neighborhood and $\tilde{\kappa}_a(w, N) = \kappa_a(w, N)$, $\kappa_a$ is locally Lipschitz continuous. □
C CHARACTERIZATION OF $\partial \mathcal{B}_r(\mathcal{W})$

We start by showing that $\partial \mathcal{B}_r(\mathcal{W}) \setminus \mathcal{G}_r(\mathcal{W})$ is given by the optimality equation. Because the continuous part of the signal is what creates the curvature, these steps are similar in ideas to Sannikov [16]. Some technical bounds on the provision of incentives and proximity of solutions to (12) for different choices of $\mathcal{D}$ are deferred to Appendix E.2. We begin with the proof of Lemma 5.5.

Proof of Lemma 5.5. We prove the slightly more general result where instead of $w + r\delta(y) \in \mathcal{W}$ for every $y \in Y$, we require $\delta \in \mathcal{D}(w)$ for a suitable map $\mathcal{D}$. Fix $w$ in the relative interior of $\mathcal{C}$ and choose $\eta > 0$ small enough such that $N_w^+N_v > 0$ for all $v \in \mathcal{C} \cap B_\eta(w)$, where $B_\eta(w)$ denotes the closed ball around $w$ with radius $\eta$. On $B_\eta(w)$, $\partial\mathcal{W}$ 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, where $T_w$ is the vector obtained by rotating $N_w$ by $90^\circ$ in clockwise direction. Denote by $\pi(v) = (\hat{v}, f(\hat{v}))$ the projection of $v \in B_\eta(w)$ onto $\partial\mathcal{W}$ in the direction $N_w$.

Let $(W, A, \beta, \delta, Z, (J^y)_{y \in Y}, M)$ be a weak solution to (2) with initial condition $W_0 = w$ such that $M \equiv 0$ and for all $t \geq 0$, $A_t = a^*(\pi(W_t))$, $\delta_t = \delta^*(\pi(W_t))$, and $\beta_t = T_t\phi^*(\pi(W_t))$ on $[0, \tau)$, where we abbreviated $N_t = N_{\pi(W_t)}$ and $T_t = T_{\pi(W_t)}$, and define $\tau := \sigma_1 \land \inf\{t \geq 0 : W_t \notin B_\eta(w)\}$, where $\sigma_1$ indicates the first time any infrequent event occurs. Since $\delta \in \Psi_A(\pi(W), N, r, \mathcal{D})$ a.e. by construction, it follows that the solution satisfies (b) in Lemma 4.1 up to time $\tau$. Since the maximizer of a measurable function is measurable and $\pi$ is measurable, $A$, $\beta$ and $\delta$ are all predictable. Moreover, because $\delta^*$ is bounded and $\phi$ is a Lipschitz-continuous function of $\delta^*$, they are both square-integrable.

We measure the distance of $W$ to $\mathcal{C}$ by $D_t = N^\top W_t - f(\hat{W}_t)$. 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 second order differentiable at almost every point by Alexandrov’s theorem. In particular, $f'$ has Radon-Nikodym derivative $f''(\hat{W}_t) = -\kappa(\pi(W_t))\ell_t^3$. It follows from the Meyer-Itô formula (see Theorem 19.5 in Kallenberg [11]) that

$$
\begin{align*}
\text{d}D_t &= r\ell_t N_t^\top(W_t - g(A_t) - \delta_t \lambda(A_t)) \, \text{d}t + r\ell_t N_t^\top T_t \phi_t (\text{d}Z_t - \mu(A_t) \, \text{d}t) \\
&\quad + r\ell_t \sum_{y \in Y} N_{t}^\top \delta^*(\pi(W_t); y) \, \text{d}J^y_t - \frac{1}{2} f''(\hat{W}_t) \, \text{d}t,
\end{align*}
$$

The volatility term is zero because $N^\top T = 0$. Note that on $[0, \sigma_1]$, $\Delta J^y \equiv 0$ for any $y \in Y$ implies that $[\hat{W}] = \hat{W}$. Using [5] and the fact that $N_w^+ N_t = T_w^\top T_t = \ell_t^{-1}$, we obtain that on $[0, \tau]$,

$$
\begin{align*}
\text{d}D_t &= r\ell_t N_t^\top(W_t - g(A_t) - \delta_t \lambda(A_t)) \, \text{d}t + r^2 \kappa(\pi(W_t)) \ell_t^3 \left|T_w^\top T_t\right|^2 |\phi_t|^2 \, \text{d}t = r D_t \, \text{d}t,
\end{align*}
$$

34
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 \( D_t = D_0 e^{rt} \), which is identically zero because \( D_0 = 0 \). On \( \{\tau < \sigma_1\} \) we can repeat this procedure and concatenate the solutions to obtain a solution to (12) that remains on \( C \) until either an event \( y \) occurs or an end point of \( C \) is reached. Let \( \rho \) denote the hitting time of an end point of \( C \). Then \( D_0 = 0 \) on \([0, \rho \wedge \sigma_1]\) implies that \( \pi(W) = W \) and hence \( \delta \in \Psi_A(W, N_W, r, D) \).

**Corollary C.1.** For any affine, compact- and convex-valued \( D \), let \( C \) be a \( C^1 \) solution to (12) with positive curvature throughout. Then any payoff in the relative interior is attainable by a strategy profile \( A \), enforced by \( (\beta, \delta) \) with \( \delta \in D((W(A)) \) such that \( W(A) \) remains on \( C \) until either an endpoint of \( C \) is reached or an event occurs.

**Proof.** For any \( w \in C \), let \( a^*(w) \) and \( \delta^*(w) \) denote the maximizers in (12). Since \( C \) is assumed to have positive curvature throughout, the maximization in (12) is not taken over empty sets. By Corollary B.6 the maximizers are attained.

The following two lemmas establish that locally, \( \partial B_r(W) \) coincides with a solution to (6) at almost every point outside \( \mathcal{G}_r(W) \). Lemma C.2 states that it is impossible for a solution to (6) to cut through \( B_r(W) \). For a curve \( C \) with positive curvature throughout, let \( N_C := \{(w, N) \in C \times S^1 \mid N^\top (w - v) \geq 0 \ \forall \ v \in C\} \) denote its outward normal bundle.

**Lemma C.2.** Let \( w \in \partial B_r(W) \) with outward normal \( N' \). Define the projection \( \pi : U \to \partial B_r(W) \) of a suitably small neighborhood \( U \) of \( w \) onto \( \partial B_r(W) \) in the direction of \( N' \) and set

\[
D(w) := \{ \delta \in \mathbb{R}^2 \mid \exists \gamma \in [0, 1] \text{ such that } \gamma w + (1 - \gamma) \pi(w) + r \delta \in W \}. \tag{13}
\]

It is impossible for a \( C^1 \) solution \( C \) to (12) for \( D \) oriented by \( v \mapsto N_v \) with end points \( v_L, v_R \in U \) to simultaneously satisfy

(i) \( v_L + \varepsilon N' \not\in B_r(W) \) and \( v_L + \varepsilon N' \not\in B_r(W) \) for any \( \varepsilon > 0 \),

(ii) there exists \( v_0 \in C \) such that \( v_0 + \eta N' \in B_r(W) \) for some \( \eta > 0 \),

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

(iv) \( N_C \cap (\Gamma(r, D) \cup \mathcal{P}) = \emptyset \),

(v) for any \( a \in A \), \( N_C \cap \partial E_a(r, D) = \emptyset \).

**Proof.** Suppose towards a contradiction that there exists such a curve \( C \). Since \( D \) is affine, compact- and convex-valued, it follows from Conditions (iv) and (v) as well as Proposition B.7 that \( C \) is \( C^2 \) at almost every point. By Condition (iii), there exists a local parametrization \( f \) of \( C \) in the direction \( N' \). Define the orthogonal projection \( \hat{v} = T' \hat{v} \) onto the tangent for any \( v \in U \), where \( T' \) is the counterclockwise rotation.
of $N'$ by $90^\circ$. Denote by $\hat{\pi}(v) = (\hat{v}, f(\hat{v}))$ the projection of $v \in U$ onto $C$ in the direction $N'$. By definition of $B_s(W)$, there exists a solution $(W, A, \beta, \delta, M, Z, (J_y^w)_{y \in Y})$ to (2) with $W_0 = v_0 + \eta N'$ such that on $[0, \sigma_1)$, $(\beta, \delta)$ enforces $A$ with $\delta \in D(W)$. Define the stopping time $\tau_1 := \inf\{t \geq 0 \mid W_t \notin U\}$.

Suppose first that $N_{C} \subseteq E_0(r, D)$ for some $a \in A$, i.e., $C$ is a non-trivial solution to (12). Let $N_t := N_{\hat{W}(t)}$ and $T_t := T_{\hat{W}(t)}$ and observe that these projections are well defined on $[0, \tau_1)$. We measure the distance of $W$ to $C$ by $D_t = N_t^T W_t - f(\hat{W}_t)$. Denote $\ell_t := 1/T_t^T T'$ and $\gamma_t := \ell_t N_t^T T'$ for the sake of brevity and observe that $\gamma := \sup_{w \in \mathcal{C}} N_w^T T'/T_w^T T' < \infty$ by Condition 5. Then, similarly as in Footnote 3 of Hashimoto [9], it follows from Itô’s formula that

$$D_t \geq D_0 + \int_0^t \xi_s \, ds + \int_0^t \xi_s (dZ_s - \mu(A_s) \, ds) + \sum_{y \in Y} \int_0^t \rho_s(y) \, dJ_y^w + \tilde{M}_t,$$

where

$$\xi_t = r \ell_t \left( N_t^T (W_t - g(A_t) - \delta_t \lambda(A_t)) + \frac{r}{2} \kappa(\hat{\pi}(W_t)) \| T_t^T \beta_t + \gamma_t N_t^T \beta_t \|^2 \right)$$

$$= r D_t + r \ell_t \left( N_t^T (\hat{\pi}(W_t) - g(A_t) - \delta_t \lambda(A_t)) + \frac{r}{2} \kappa(\hat{\pi}(W_t)) \| T_t^T \beta_t + \gamma_t N_t^T \beta_t \|^2 \right),$$

$$\xi_t = r \ell_t N_t^T \beta_t, \rho_t(y) = r \ell_t \delta_t(y) \text{ and } \tilde{M}_t = \int_0^t r \ell_t \delta_t \, dM_t.$$ Define the stopping time $\tau_2 := \inf\{t \geq 0 \mid D_t \leq 0\}$ and observe that $\tau_2 \leq \tau_1$ a.s. by Condition (i). We will show that there exists an equivalent probability measure $R$ such that the drift rate of $D_t$ is bounded from below by $r D_t$. Then, $D_t$ becomes arbitrarily large with positive $R$-probability, and hence positive $Q^A$-probability. Because it may take arbitrarily long until an accident arrives, this leads to a contradiction because $\mathcal{V}$ is bounded.

Let $\Xi_1$ denote the set where $N^T (\hat{\pi}(W) - g(A) - \delta \lambda(A)) \geq 0$. On $\Xi_1$, $\zeta_t \geq r D_t$, hence there is no need to change the probability measure. It follows from Condition (iv) that $\beta \neq 0$ on $\Xi_1$. Let $\Xi_2 \subseteq \Xi_1$ be the set where $N_{\hat{C}} \subseteq E_0(r, D)$, i.e., $\Psi_A(\hat{\pi}(W), N, r, D) \neq \emptyset$. Set

$$\hat{\delta} = \arg\min_{x \in \Psi_A(\hat{\pi}(W), N, r, D)} \| x - \delta^1 \| + \| x - \delta^2 \|,$$

then (12) implies that

$$\zeta \geq r D - r \ell N^T (\delta - \hat{\delta}) \lambda(A) - r \ell N^T (g(A) + \hat{\delta} \lambda(A) - \hat{\pi}(W)) \left( 1 - \frac{\| T^T \beta \|^2 - \gamma \| N^T \beta \|^2}{\| \phi(a, N, \delta) \|^2} \right).$$

Denote $\Lambda := \max_{a \in A} \sum_{y \in Y} \lambda(y|a)$ and observe that $N^T (g(A) + \hat{\delta} \lambda(A) - \hat{\pi}(W))$ is uniformly bounded above by the constant $K_1 := \text{diam } \mathcal{V} + \sup(W - \mathcal{V}) \Lambda < \infty$. The
condition that $W + r\delta(y) \in \mathcal{W}$ implies that $\delta(y) \in \mathcal{D}(W)$ on $[0, \tau_2)$ for every $y \in Y$. Due to Lemma E.3 there exist constants $K_2, \Psi$ such that

$$
\zeta \geq rD - r\ell\Lambda K_2 \| N^T\beta \| - r\ell K_1 \frac{2K_2 + 2\gamma}{\Psi} \| N^T\beta \| =: rD_t - K_3 \| \zeta \|.
$$

On the set $\Xi_1 \cap \Xi_2$, condition (v) implies that $\mathcal{N}_c$ is bounded away from $E_A(r, \mathcal{D}) \cup \mathcal{P}$ by virtue of Corollary B.6. Lemma E.4 thus implies that $\| N^T\beta \| \geq K_4$ for some constant $K_4$ and hence

$$
\zeta_t \geq rD_t - r\ell_t K_1 \geq rD_t - \frac{K_1}{K_4} \| \zeta_t \|.
$$

Let $T := \min\{ t \geq 0 \mid D_0(1 + rt)/2 \geq \sup_{w \in \mathcal{Y}} N^T w - f(\hat{w}) \}$ and observe that $T$ is deterministic. We define a density process $L$ on $[0, T]$ by setting

$$
\frac{dL_t}{L_t} = \psi_t \, dZ_t + \sum_{y \in Y} \left( \frac{1}{\lambda(y|A_{t-})} - 1 \right) \, dJ^y_t,
$$

where

$$
\psi_t = K_3 \frac{\zeta_t}{\| \zeta_t \|} 1_{\Xi_1} + \frac{K_1}{K_4} \frac{\zeta_t}{\| \zeta_t \|^2} 1_{\Xi_1 \cap \Xi_2}.
$$

Because $\int_0^T \| \psi_t \|^2 \, dt < \infty$ $Q^A_t$-a.s., it follows from Girsanov’s theorem that $L$ defines a probability measure $R$ equivalent to $Q^A_t$ on $\mathcal{F}_T$ such that $dZ'_t = dZ_t - \psi_t \, dt$ is an $R$-Brownian motion on $[0, T]$, for every $y \in Y$, $J^y$ has intensity 1 and $\tilde{M}_t$ is an $R$-martingale because it is orthogonal to $L$. Then

$$
D_t \geq D_0 + \int_0^t rD_s \, ds + \int_0^t \xi_s^\top \, dZ'_s + \tilde{M}_t + \sum_{y \in Y} \int_0^t \rho_s(y) \, dJ^y_s. \tag{14}
$$

Since $W$ is bounded, $\int_0^t \xi_s^\top \, dZ_s$ is a $BMO(Q^A_t)$-martingale. Therefore, $\int_0^t \xi_s^\top \, dZ'_s$ is a $BMO(R)$-martingale by Theorem 3.6 in Kazamaki [12]. Define the stopping time $\tau_3 := \inf\{ t \geq 0 \mid D_t \leq D_0(1 + rt)/2 \}$ and observe that $\tau_3 \leq \tau_2 \wedge T$. It follows from (14) that

$$
D_{\tau_3} - \frac{D_0}{2} (1 + r\tau_3) \geq \frac{D_0}{2} + F_{\tau_3} + \sum_{y \in Y} \int_0^{\tau_3} \rho_s(y) \, dJ^y_s,
$$

where $F_t = \int_0^t \xi_s \, dZ'_s + \tilde{M}_t$ is an $R$-martingale starting at 0. Define the $R$-martingale $G_t := e^{\mathcal{Y} t} 1_{\{ t < \sigma \}}$ and observe that $G$ is orthogonal to $F$. Because $\tau_3 \leq T$ a.s.,

$$
0 \geq \mathbb{E}_R \left[ \left( D_{\tau_3} - \frac{D_0}{2} (1 + r\tau_3) \right) 1_{\{ T < \sigma \}} \right] \geq \mathbb{E}_R \left[ \frac{D_0}{2} 1_{\{ T < \sigma \}} + F_{\tau_3} 1_{\{ T < \sigma \}} \right] = \frac{D_0}{2} R(T < \sigma) + e^{-\mathcal{Y} T} \mathbb{E}_R [F_{\tau_3} G_T] > 0,
$$

37
Figure 11: Construction of a curve $C'$ that cuts through $B_r(W)$.

where the last inequality follows from the optional stopping theorem and because $R$ is equivalent to $Q_T$. This is a contradiction.

Suppose now that $N \subseteq \bigcap_{a \in A} E_a(r, D)^c$, i.e., $C$ is a straight line segment. Let $D$ denote the distance of $W$ to $C$ in the direction of the normal vector $N'$ of $C$. Condition (iv) makes it possible to apply Lemma E.4, hence any $(\beta, \delta)$ enforcing $A$ it follows that $N^\top \beta \geq K$ for some constant $K$. Similarly as before, the drift of $D_t$ is thus bounded from below by $rD_t - K_1/Kr\ell_t\|N^\top \beta_t\|$. Therefore, there exists an equivalent probability measure under which $D$ grows arbitrarily large with positive probability, a contradiction.

Lemma C.3. Fix $w \in B_r(W) \setminus G_r(W)$ with outward normal $N$, where (i) is locally Lipschitz continuous. Then $\partial B_r(W)$ coincides with a solution to (6) in a neighbourhood of $(w, N)$.

Proof. We first show that a solution to (12) with $D$ given in (13) coincides with $\partial B_r(W)$, which implies that also a solution to (6) stays on $\partial B_r(W)$. In a sufficiently small neighbourhood of $(w, N)$, (12) admits a unique $C^2$ solution that is continuous in initial values. Let $C$ be solution with initial value $(w, N)$ and suppose towards a contradiction that $C$ escapes $\text{cl} B_r(W)$ in a neighbourhood of $w$. Then we can change initial conditions slightly to obtain a curve $C'$ that cuts through $B_r(W)$. Specifically:

- If $\partial B_r(W)$ is not $C^1$ at $w$, we obtain $C'$ as a solution to (6) with initial conditions $(w - \eta N, N)$ for $\eta > 0$ sufficiently small.

- If $\partial B_r(W)$ is $C^1$ at $w$, we obtain $C'$ for initial conditions $(w, N')$, where $N'$ is a slight rotation of $N$ as illustrated in the left panel of Figure 11.

Because the set where (12) fails to be locally Lipschitz continuous is closed by Corollary B.6 and Proposition B.7, a small enough perturbation satisfies $N_C \cap \Gamma_a(r, D) = \emptyset$ for every $a \in A$ and either $N_C \subseteq E_a(r, D)$ or $N_C \cap (E_a(r, D) \cup \mathcal{P}) = \emptyset$ for any $a \in A$, that is, $C'$ satisfies conditions (iv) and (v) of Lemma C.2. By choosing $\eta$ or $N'$ suitably, we can get conditions (i)–(iii) to hold as well, hence $C'$ is impossible due to Lemma C.2. We conclude that $\partial B_r(W)$ is $C^1$ where (12) is locally Lipschitz continuous and that a solution to (12) cannot escape $\text{cl} B_r(W)$.

38
Suppose towards a contradiction that \( C \) falls into the interior of \( B_{r}(W) \) in a neighbourhood of \((w, N)\), that is, there exists \( v \in C \cap \text{int} B_{r}(W) \) arbitrarily close to \( w \).

By convexity of \( B_{r}(W) \), this is not possible if \( C \) is a trivial solution to \((12)\), hence \( C \) is a solution with positive curvature. We may assume without loss of generality that this happens to the right of \( w \) as illustrated in Figure 12. Let \( v \) be close enough to \( w \) such that \((12)\) with \( D \) is Lipschitz continuous on an open neighbourhood of \( N_{C} := \{ (\tilde{w}, N_{\tilde{w}}) \mid \tilde{w} \in C \text{ between } w \text{ and } v \} \). Let \( \delta > 0 \) such that the closed ball \( B_{\delta}(v) \) is contained in the interior of \( B_{r}(W) \). For \( \zeta > 0 \) to be chosen later, let \( W_{\zeta} := \{ w \in V \mid d(w, W) \leq \zeta \} \), where \( d(w, W) \) denotes the minimal distance of \( d \) from \( W \). Set

\[
D_{\zeta}(w) := \{ \delta \in \mathbb{R}^{2} \mid \exists \kappa \in [0, 1] \text{ such that } \kappa w + (1 - \kappa)\pi(w) + r\delta \in W_{\zeta} \},
\]

where \( \pi \) is the projection onto \( \partial B_{r}(W) \) in the direction \( N \). Observe that for \( \zeta \) sufficiently small, \((12)\) with \( D_{\zeta} \) is Lipschitz continuous in a neighbourhood of \( N_{C} \), hence it admits a unique solution \( C_{\zeta} \). Choose now \( \zeta \) small enough such that Lemma E.5 asserts the existence of \( v' \in C_{\zeta} \cap B_{\delta}(v) \).

Because \( C_{\zeta} \) is continuous in initial conditions, a solution \( C'_{\zeta} \) to \((12)\) with \( D_{\zeta} \) for a slight rotation \( N' \) of \( N \) reaches a neighbourhood of \( v' \) in \( B_{r}(W) \). As illustrated in Figure 12, \( C'_{\zeta} \) will escape \( \text{cl} B_{r}(W) \) to the right of \( w \) and enter \( B_{r}(W) \) to the left of \( w \). Thus, for \( N' \) close enough to \( N \), there exist \( v_{L}, v_{R} \in C'_{\zeta} \cap B_{r}(W) \), such that \( \| \tilde{w} - \pi(\tilde{w}) \| \leq \zeta \) for all \( \tilde{w} \in C'_{\zeta} \). By Corollary C.1, for any \( w' \in C'_{\zeta} \) there exists a solution to \((2)\) with \( W_{0} = w' \) such that \( \delta \in \Psi_{A}(W, N, r, \zeta) \) on \([0, \sigma_{1}])\) and \( W \in C'_{\zeta} \) until it reaches an end point of \( C'_{\zeta} \) or an event occurs. Let \( \tau := \inf \{ t \geq 0 \mid W_{t} \in \{ v_{L}, v_{R} \} \} \) and observe that \( W_{\tau} \in B_{r}(W) \) on \( \{ \tau < \sigma_{1} \} \). The condition that \( \delta \in D_{\zeta}(W) \) a.e. for every \( y \in Y \) implies that \( x + r\delta_{0}(y) \in W_{\zeta} \) for some \( x \) between \( W_{t} \) and \( \pi(W_{t}) \). On \([0, \tau \wedge \sigma_{1}])\) it holds that \( \| W_{t} - x \| \leq \zeta \), and hence \( \delta \in \Psi_{A}(W, N_{W}, r, D) \). Because \( W_{\tau} \in B_{r}(W) \) on \( \{ \tau < \sigma_{1} \} \), by definition of \( B_{r}(W) \) there exists a solution \((\tilde{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta})\) with \( \tilde{W}_{0} = W_{\tau} \) such that on \([\sigma_{1}, \sigma_{2}]\), \( (\tilde{\beta}, \tilde{\delta}) \) enforces \( \tilde{A}, \tilde{W} + r\delta_{0}(y) \in D(\tilde{W}) \) a.e. Therefore, a concatenation of \((W, A, \beta, \delta)\) with \((\tilde{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta})\) satisfies the same properties, which shows that \( C'_{\zeta} \subseteq B_{r}(W) \), which is a contradiction to \( w' \notin \text{cl} B_{r}(W) \). \( \square \)

Finally, because \((6)\) is Lipschitz continuous almost everywhere, we need to show that \( \partial B_{r}(W) \) is \( C^{1} \) to grant uniqueness of the solution. By convexity, \( B_{r}(W) \) cannot have inward corners, and it will follow with another escaping argument that it cannot have outward corners outside of \( G_{r}(W) \) either.

**Lemma C.4.** \( \partial B_{r}(W) \setminus G_{r}(W) \) is \( C^{1} \) where \((6)\) fails to be Lipschitz continuous. Moreover, outside of \( P \), the set of all points in \( \partial B_{r}(W) \setminus G_{r}(W) \), where \((6)\) fails to be Lipschitz continuous, has relative measure 0.

**Proof.** We already know from Lemma C.3 that \( \partial B_{r}(W) \setminus G_{r}(W) \) is \( C^{1} \) where it is locally Lipschitz continuous. Suppose, therefore, that \( \partial B_{r}(W) \) has a corner at \( w \in \partial B_{r}(W) \setminus G_{r}(W) \). Therefore, a concatenation of \( (\tilde{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta}) \) with \( \tilde{W}_{0} = W_{\tau} \) such that on \([\sigma_{1}, \sigma_{2}]\), \( (\tilde{\beta}, \tilde{\delta}) \) enforces \( \tilde{A}, \tilde{W} + r\delta_{0}(y) \in D(\tilde{W}) \) a.e. Therefore, a concatenation of \((W, A, \beta, \delta)\) with \((\tilde{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta})\) satisfies the same properties, which shows that \( C'_{\zeta} \subseteq B_{r}(W) \), which is a contradiction to \( w' \notin \text{cl} B_{r}(W) \). \( \square \)
positive length. By shortening the line segment we may assume that
by Corollary B.6, a contradiction.

This implies that

with initial conditions \((v, N')\) and a slight reduction over controls \(\delta \in D_\zeta(w)\) such that \(C'_\zeta\) escapes \(\mathcal{B}_r(W)\). For
small \(\zeta\) and \(N'\) close to \(N\), there exists an enforceable strategy profile attaining \(w' \notin \mathcal{B}_r(W)\) which
reaches \(\mathcal{B}_r(W)\) with certainty. This leads to a contradiction.

Proposition B.7
implies that \((w, N_L)\) and \((w, N_R)\) are contained in \(\partial E_a(r, W)\) and \(\partial E_a'(r, W)\), respectively, where \(N_L\) and \(N_R\) are the extremal normal vectors to \(\partial \mathcal{B}_r(W)\) at \(w\). Since
\(E_a(r, W) \cup \mathcal{P}\) is closed, there exists an open neighborhood \(U\) of \(w\) and a set of normal vectors \(T \subseteq N_w(\mathcal{B}_r(W))\) such that \(U \cap \mathcal{B}_r(W) \times T\) is either contained in or has empty intersection with \(E_a\). Since \(w \notin \mathcal{G}\), it follows that \(\{w\} \times N_w(\mathcal{B}_r(W)) \cap \Gamma(r, W) = \emptyset\). Since \(\Gamma(r, W)\) is closed, \(U \cap \mathcal{B}_r(W) \times T \cap \Gamma(r, W) = \emptyset\) for \(U\) sufficiently small. We can thus construct a solution to \(\mathcal{C}\) to (12) for \(\mathcal{D}\) given in (13) with
initial conditions \((v, N) \in U \cap \mathcal{B}_r(W) \times T\) that cuts through \(\mathcal{B}_r(W)\) with
\(\mathcal{N}_C \subseteq U \cap \mathcal{B}_r(W) \times T\). By choice of \(U\) and \(T\), the conditions in Lemma C.2 are satisfied, which is a contradiction.

For the second statement, suppose that there exists \(\mathcal{C} \subseteq \partial \mathcal{B}_r(W) \setminus \mathcal{G}_r(W)\) of
positive length. By shortening the line segment we may assume that \(\mathcal{N}_C \subseteq \mathcal{P}\) or
\(\mathcal{N}_{\text{int}\mathcal{C}} \cap \mathcal{P} = \emptyset\). Suppose towards a contradiction that \(\mathcal{N}_{\text{int}\mathcal{C}} \cap \mathcal{P} = \emptyset\). Then Proposition B.6
shows that \((w, N)\) enters and leaves \(E_a(r, \mathcal{D})\) of the maximizing action profile
\(a\) at almost every \((w, N) \in \mathcal{N}_C\). Because \(\mathcal{A}\) is finite we may assume that this is the
same action profile. This implies that \(\mathcal{N}_C \subseteq \partial E_a(r, \mathcal{D})\) and hence \(\mathcal{N}_{\text{int}\mathcal{C}} \subseteq E_a(r, \mathcal{D})\)
by Corollary B.6 a contradiction. □

Next, we prove Proposition B.7, showing that extremal points of any \(C^1\) segment in \(\mathcal{G}_r(W)\) and any corners of \(\mathcal{G}_r(W)\) must lie in \(\mathcal{K}_r(W)\).

Proof of Proposition B.7. Suppose first that \(w\) is a corner of \(\mathcal{B}_r(W)\) in \(\mathcal{G}_r(W) \setminus \mathcal{S}_r(W)\). By definition of \(\mathcal{G}_r(W)\), there exists \((a, \delta_0)\) such that \(\delta_0 \in \Psi_a\), \(w + r\delta_0(y) \in W\) for
every \(y \in Y\), and \(N^\top(g(a) + \delta_0\lambda(a) - w) \geq 0\) for all \(N \in N_w(\mathcal{B}_r(W))\). Since \(w\) is not a
stationary point, \(w \neq g(a) + \delta_0\lambda(a)\) and hence \(N^\top(g(a) + \delta_0\lambda(a) - w) > 0\) for almost all normal vectors in \(N_w(\mathcal{B}_r(W))\). Lemma B.5 thus readily implies that \(w \in \partial \mathcal{K}_{r,a}(W)\).
It remains to show that the normal vectors to the sets \(\mathcal{B}_r(W)\) and \(\mathcal{K}_{r,a}(W)\) satisfy the
desired inclusion property. Suppose towards a contradiction that the converse holds, that is, \(N_w(\mathcal{B}_r(W)) \not\subseteq N_w(\mathcal{K}_{r,a}(W))\). Since both \(N_w(\mathcal{B}_r(W))\) and \(N_w(\mathcal{K}_{r,a}(W))\) are closed, there exists \(N \in N_w(\mathcal{K}_{r,a}(W)) \setminus N_w(\mathcal{B}_r(W))\) with \(N^\top(g(a) + \delta_0\lambda(a) - w) > 0\). 

Figure 12: If \(C\) falls into the interior of \(\mathcal{B}_r(W)\), there exists a solution \(C'_\zeta\) to (3) with initial
conditions \((w, N')\) and a slight reduction over controls \(\delta \in D_\zeta(w)\) such that \(C'_\zeta\) escapes \(\mathcal{B}_r(W)\). For
small \(\zeta\) and \(N'\) close to \(N\), there exists an enforceable strategy profile attaining \(w' \notin \mathcal{B}_r(W)\) which
reaches \(\mathcal{B}_r(W)\) with certainty. This leads to a contradiction.
By definition of the normal vector, \( w_\varepsilon := w + \varepsilon T \in \text{int} K_{r,a} \setminus \text{cl} \mathcal{B}_r(W) \) for \( \varepsilon > 0 \) sufficiently small, where \( T \) is orthogonal to \( N \) with \( T^\top N' < 0 \) for all \( N' \in N_w(\mathcal{B}_r(W)) \).

Fix such an \( \varepsilon \) sufficiently small and define \( w_{\varepsilon,\gamma} := \gamma w_{+}(1 - \gamma)w_\varepsilon \) for \( \gamma \in [0,1] \). Since \( w_\varepsilon \) is in the interior of \( K_{r,a}(W) \), there exists \( \delta' \in \text{int} \Psi_a \) with \( w_\varepsilon + r\delta'(y) \in \text{int} W \) for every \( y \in Y \). Let \( \delta_\gamma := \gamma \delta_0 + (1 - \gamma)\delta' \). Since \( W \) and \( \Psi_a \) are convex, it follows that for any \( \gamma > 0 \), \( \delta_\gamma \in \text{int} \Psi_a \) and \( w_{\varepsilon,\gamma} + r\delta_\gamma(y) \in \text{int} W \) for every \( y \in Y \). Moreover for \( \gamma \) sufficiently small, \( N^\top (g(a) + \delta_\gamma \lambda(a) - w_{\varepsilon,\gamma}) > 0 \). A contradiction can thus be obtained in the same way as in the proof of Lemma 5.6.

Suppose next that \( \mathcal{G}_r(W) \) contains a \( C^1 \) line segment \( \mathcal{C} \) of positive length and positive curvature. Let \( w \) be in the relative interior of \( \mathcal{C} \) and denote by \( N_w \) the unique normal vector to \( \partial \mathcal{B}_r(W) \) at \( w \). If there exists \( (a, \delta_w) \in \Psi_a \cap (\mathcal{W} - w)^m/r \) with \( N_w \top (g(a) + \delta_w \lambda(a) - w) > 0 \), then the argument works in the same way as before: conditions (i) and (iii) of Lemma 5.6 have to be violated, showing that \( w \in \partial K_{r,a}(W) \) and if \( N_w \notin N_w(K_{r,a}(W)) \), we can enlarge the set in the same way as before. Suppose, therefore, that every \( w \in \mathcal{C} \) is decomposable only by \( (a_w, \delta_w) \) with \( N_w \top (g(a_w) + \delta_w \lambda(a_w) - w) = 0 \). Since \( w \notin \mathcal{S}_r(W) \), the drift is parallel to \( \partial \mathcal{C} \). Thus, there exists \( v \) outside of \( \mathcal{B}_r(W) \) arbitrarily close to \( w \) with \( N_w \top (g(a_w) + \delta_w \lambda(a_w) - w) > 0 \). If any such \( v \) is in \( K_{r,a} \), we obtain a contradiction in the same way as before. Therefore, \( w \in \partial K_{r,a}(W) \). Finally, if \( N_w \notin N_w(K_{r,a}(W)) \), then there exist payoffs \( v \in \mathcal{C} \) arbitrarily close to \( v \) that are in the interior of \( K_{r,a}(W) \). Since there are only finitely many action profiles, this is a contradiction.

\[ \square \]

**D  Closedness of \( \mathcal{B}_r(W) \)**

This appendix shows that the set \( \mathcal{B}_r(W) \) is closed. It also contains the proof of Theorem 5.8 based on the auxiliary results in Appendix C and this appendix.

**Lemma D.1.** Suppose that \( w \in \mathcal{G}_r(W) \) is either a corner or part of a continuously differentiable line segment in \( \mathcal{G}_r(W) \). Then \( w \in \mathcal{B}_r(W) \).

**Proof.** By Lemma 5.4 any payoff \( w \in \mathcal{S}_r(W) \) is contained in \( \mathcal{B}_r(W) \). Suppose, therefore, that \( w \notin \mathcal{S}_r(W) \) and consider first the case where \( w \) is a corner of \( \mathcal{B}_r(W) \).

Suppose first that there exist an action profile \( a \in \mathcal{A} \), a measurable selection \( \delta^* : K_{r,a} \to \Psi_a \) and a time \( t_0 > 0 \) such that the solution \( (W, A, \beta, \delta, M) \) to (2) with \( W_0 = w \), \( A \equiv a \), \( \beta \equiv 0 \), \( \delta = \delta^*(W) \), and \( M \equiv 0 \) remains in \( \mathcal{B}_r(W) \cap K_{r,a} \) on \( [0, t_0 \wedge \sigma_1] \) and \( W_{t_0} \) is in the interior of \( \mathcal{B}_r(W) \) on the set \( \{ t_0 < \sigma_1 \} \). Observe that such a solution \( W \) to (2) travels on a deterministic path before the arrival of an infrequent event \( y \in Y \). In particular, \( \delta \) is predictable. Since \( W_{t_0} \in \mathcal{B}_r(W) \) on the event \( \{ t_0 < \sigma_1 \} \), there exists a solution \( (\bar{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta}, \bar{M}) \) attaining \( W_{t_0} \) with \( \tilde{\beta}, \tilde{\delta} \) enforcing \( \bar{A} \) such that \( \bar{W} \in \mathcal{B}_r(W) \) up until the arrival of the first infrequent event, at which point \( W \) jumps to \( \bar{W} \). The concatenation \( (\bar{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta}, \bar{M})1_{[0, t_0]} + (\bar{W}, \tilde{A}, \tilde{\beta}, \tilde{\delta}, \bar{M})1_{(t_0, \infty)} \) thus satisfies the same conditions on \( [0, \infty) \), showing that \( w \in \mathcal{B}_r(W) \).

41
Suppose now that no such $a, \delta^*, t_0$ exist. Then there exists no strategy profile attaining $w$ without using the Brownian information to structure incentives. We can thus obtain a contradiction using an escaping argument similarly to Lemma C.2. Let $C_\epsilon$ be a solution to (6) starting in $w - \epsilon N$ for an outward normal vector $N$ and $\epsilon > 0$. For $\epsilon$ sufficiently small, $C_\epsilon$ intersects $\partial B_r(W)$ and satisfies the conditions of Lemma C.2 outside a set of measure 0. Such a curve is impossible by Lemma C.2.

Suppose next that $G_r(W)$ is locally a $C^1$ line segment $C$. Since there are only finitely many action profiles, there exists a subsegment $C'$ of positive length that can be decomposed by some action profile $a \in A$. If $C'$ has positive curvature, then $C' \subseteq \partial K_{r,a}(W)$ by Proposition 5.7. Since $K_{r,a}(W)$ is convex, a measurable selector $\delta^*$ as in the previous case exists and $C' \in G_r(W)$. If $C'$ is a straight line segment, then any payoffs in the relative interior can be attained by public randomization and its endpoints can be attained as before.

**Lemma D.2.** $B_r(W)$ is closed.

**Proof.** By public randomization, a straight line segment is contained in $B_r(W)$ if both of its end points are contained in $B_r(W)$. Similarly, Lemma 5.5 shows that curved parts of $\partial B_r(W) \setminus G_r(W)$ are contained in $B_r(W)$ if its end points are. Since Lemma D.1 shows that $G_r(W) \subseteq G_r(W)$, the closure of $B_r(W)$ is $W$-relaxed generating, hence it is contained in $B_r(W)$ by maximality.

**Proof of Theorem 5.8.** Lemmas C.3 and C.4 imply that $\partial B_r(W) \setminus G_r(W)$ is a $C^1$ solution to (6). It follows from Proposition 5.7 that $G_r(W)$ has the desired properties. Finally, $B_r(W)$ is closed by Lemma D.2.

**E Auxiliary results related to the optimality equation**

**E.1 Lipschitz continuity of set-valued maps**

Consider an arbitrary family $(F_i)_{i \in I}$ of Lipschitz continuous set-valued maps. If their Lipschitz constants $(K_i)_{i \in I}$ are uniformly bounded, then the union $x \mapsto \bigcup_{i \in I} F_i(x)$ is Lipschitz continuous again. However, the intersection of two Lipschitz continuous maps may fail to be Lipschitz continuous in general. In this appendix, we show Lipschitz continuity of the intersection for two special cases that are relevant in our setting.

**Lemma E.1.** The intersection of two convex-valued affine maps is Lipschitz continuous.

**Proof.** Let $F$ and $G$ be two convex-valued affine functions. It is sufficient to show that $F \cap G$ is continuous as it is then Lipschitz continuous since both $F$ and $G$ are affine. Suppose towards a contradiction that $F \cap G$ fails to be continuous at some $x_0$,
that is, there exists $v \in F(x_0) \cap G(x_0)$ such that $B_\varepsilon(v) \cap F(x_0) \cap G(x_0) = \emptyset$ for $\varepsilon > 0$ arbitrarily small and $x \in \text{supp} F \cap G$ arbitrarily close to $x_0$. Since $F$ and $G$ are affine, this is only possible if $N_F = -N_G$, where $N_F$ and $N_G$ denote the normal vectors to $\partial F(x_0)$ and $\partial G(x_0)$, respectively, at $v$. It follows from convexity that this is possible only if $F(x) \cap G(x) = \emptyset$ for $x$ arbitrarily close to $x_0$, contradicting the fact that $x \in \text{supp} F \cap G$.

**Lemma E.2.** Let $F$ and $G$ be Lipschitz continuous maps with bounded support taking values in closed convex polytopes. Denote the outward normal vectors to their hyperfaces by $\pi_i^F(x)$, $i \in I_F$ and $\pi_i^G(x)$, $i \in I_G$, respectively. If for any $J_F \subseteq I_F$, $J_G \subseteq I_G$, the matrix $[(\pi_i^F(x))_{j \in J_F}, (\pi_i^G(x))_{j \in J_G}]$ has constant column rank in a neighbourhood of $x$, then $F \cap G$ is locally Lipschitz continuous at $x$. If the ranks of above matrices are constant on the entire support of $F \cap G$, then $F \cap G$ is Lipschitz continuous.

**Proof.** Fix $x$ in the support of $F \cap G$ and let $K$ be the maximum of the Lipschitz constants of $F$ and $G$. Then Lipschitz continuity of the individual maps implies that

$$F(\bar{x}) \cap G(\bar{x}) \subseteq (F(x) + \|\bar{x} - x\|B_K(0)) \cap (G(x) + \|\bar{x} - x\|B_K(0)).$$

Observe, however, that the right hand side is larger than $F(x) \cap G(x) + \|\bar{x} - x\|B_K(0)$. Let $H(z) := \partial(\{F(x) + B_z(0)\} \cap \{G(x) + B_z(0)\})$ be the level sets of $\partial(F(x) \cap G(x))$. Let $p_1$ denote the point in $H(1)$ with maximal distance from $\partial(F(x) \cap G(x))$ and let $p_0$ be the point in $\partial(F(x) \cap G(x))$ with minimal distance from $p_1$ as illustrated in Figure 13. Let $\{\pi_1, \ldots, \pi_n\}$ be a minimal subset of normal vectors to the hyperfaces of $F(x) \cap G(x)$ that intersect at $p_0$ such that $p_1$ is the unique point in $H(1)$, which is related to $p_0$ by

$$\pi_j^\top (p_1 - p_0) = 1 \quad \text{for } j = 1, \ldots, m \quad \text{and} \quad p_1 - p_0 \in \text{span} \pi_j. \quad (15)$$

By linearity of (15), it follows that $p_{K\|\bar{x} - x\|} := p_0 + K\|\bar{x} - x\|(p_1 - p_0)$ is a point in $H(K\|\bar{x} - x\|)$ with maximal distance of $F(x) \cap G(x)$. Its distance from $p_0$ equals $K\|p_1 - p_0\|\|\bar{x} - x\|$. The statement thus follows once we show that $\|p_1(x) - p_0(x)\|$ is uniformly bounded in $x$. 

---

Figure 13: Level sets $H(z)$ of $\partial(F(x) \cap G(x))$ containing points $p_z$ with maximal distance from $p_0 \in \partial(F(x) \cap G(x))$. Clearly, $\|p_z - p_0\| = z\|p_1 - p_0\|$. The proof proceeds as above.
By minimality of \( \{\pi_1, \ldots, \pi_n\} \), the vectors \( \pi_1, \ldots, \pi_n \) are linearly independent. Thus by assumption, \( \pi_1(\tilde{x}), \ldots, \pi_n(\tilde{x}) \) are linearly independent also for \( \tilde{x} \) in a neighbourhood of \( x \). Since \( F \) and \( G \) are continuous, the norm of the solution is continuous, hence by making the neighbourhood smaller and compact, its maximum is bounded. Because \( F \) and \( G \) have finitely many hyperfaces, the finite maximum over all possible combinations of normal vectors \( \pi_1, \ldots, \pi_n \) yields a bound for \( ||p_1 - p_0|| \) on a sufficiently small neighbourhood of \( x \). Finally, if the rank is constant on \( \text{supp} F \cap G \), then \( ||p_1 - p_0|| \) is uniformly bounded since \( \text{supp} F \cap G \) is compact.

\[ \text{Proof.} \]
We begin the proof by extending \( F \) hence by making the neighbourhood smaller and compact, its maximum is bounded.

\[ \hat{\phi} \]

\[ \Phi_a(N, \delta) := \{ \beta | (\beta, \delta) \text{ enforces } a \text{ and } N^\top \beta = 0 \} \]

the space of value transfers normal to \( N \) that enforce \( a \) given value burning \( \delta \). Let

\[ \Psi_a(w, N, r, \mathcal{W}) := \left\{ \delta \in \mathbb{R}^{2 \times |Y|} \mid \begin{array}{c}
\Phi_a(\delta, N) \neq \emptyset, N^\top (g(a) + \delta \lambda(a) - w) \geq 0, \\
R \in \mathcal{W} \text{ for every } y \in Y
\end{array} \right\}. \]

Relevant is the optimality equation

\[ \kappa_a(w) = \max_{\delta \in \Psi_a(w, N, r, \mathcal{W})} \frac{2N_w^\top (g(a) + \delta \lambda(a) - w)}{r \| \phi(\alpha, N, \delta) \|^2}, \quad (16) \]

where \( \phi(a, N, \delta) \) is the shortest vector in \( \Phi_a(N, \delta) \).

**Lemma E.3.** Let \( \mathcal{C} \) be a \( C^1 \) solution to \([16]\) for fixed \( a, r, \text{ and } \mathcal{W} \) with endpoints \( v_L, v_R \) such that the normal vector \( N_w \) to \( \mathcal{C} \) at \( w \) is contained in \( E_a^\mathcal{g}(r, \mathcal{W}) \setminus (\Gamma_a^\mathcal{g}(r, \mathcal{W}) \cup \mathcal{P}) \) for every \( w \in \mathcal{C} \). Then there exists a constant \( K \) such that for any \( \alpha \geq 0 \), for any \( w \in \mathcal{C} \), any \( (T_a \phi + N_w \chi, \delta) \) enforcing \( a \) with \( N_w^\top (g(a) + \delta \lambda(a) - w) \geq 0 \) and \( w + r \delta(y) \in \mathcal{W} \) satisfies

\[ K \| \chi \| \geq \| \hat{\delta}^1 - \delta^1 \| + \| \hat{\delta}^2 - \delta^2 \|, \quad \frac{2K + 2\alpha}{\Psi} \| \chi \| \geq 1 - \frac{(\| \phi - \alpha \| \| \chi \|)^2}{\| \phi(\alpha, N, \hat{\delta}) \|^2}, \quad (17) \]

where \( \hat{\delta} \) is the element of \( \Psi_a(w, N_w, r, \mathcal{D}) \) that minimizes \( \| \hat{\delta}^1 - \delta^1 \| + \| \hat{\delta}^2 - \delta^2 \| \) and \( \Psi := \inf_{w \in \mathcal{C}} \min_{\delta \in \Psi_a(w, N_w, r, \mathcal{W})} \| \phi(a, N_w, \delta') \|^2 > 0 \).

**Proof.** We begin the proof by extending \( \phi(a, N, \delta) \) to \( \hat{\delta} \) with \( \Phi_a(N, \delta) = \emptyset \) in a Lipschitz continuous way. We achieve this by introducing the function \( \chi(a, N, \delta) \) characterizing the minimal normal component \( N^\top \beta \) necessary to enforce \( a \) given \( \delta \), and setting \( \phi(a, N, \delta) \) equal to the minimal tangential component \( \phi \) necessary such that \( (T \phi + N \chi(a, N, \delta), \hat{\delta}) \) enforces \( a \). Observe that this is indeed an extension of \( \phi(a, N, \delta) \) as \( \chi(a, N, \delta) \equiv 0 \) for \( \delta \) with \( \Phi_a(N, \delta) \neq \emptyset \).
By assumption. Define the function $D(w, \delta) := \min_{x \in \Psi_a(w, N_w, r, W)} \|x^1 - \delta^1\| + \|x^2 - \delta^2\|$.
and observe that it is well-defined for any \( w \in C \) since \( \Psi_\alpha(w, N_w, \rho, \mathcal{W}) \) is non-empty because \( (w, N_w) \in E_\alpha(r, \mathcal{W}) \) by assumption. Observe also that \( D \) is continuous—as it minimizes a convex function over a convex set—and that it is equal to 0 for \( \delta \in \Psi_\alpha(w, N_w, \rho, \mathcal{W}) \), which is precisely the case when \( T^1_\alpha(N_w, \delta^1) \) and \( T^2_\alpha(N_w, \delta^2) \) overlap. Since a change in \( \delta \) corresponds to a linear shift of the hyperfaces of \( T^1_\alpha(N_w, \delta^1) \), \( D(w, \delta) \) is piecewise linear in \( \delta \). For fixed \( w \), the rate at which \( D(w, \delta) \) changes depends on the angle between the closest hyperfaces of \( T^1_\alpha(N_w, \delta^1) \) and \( T^2_\alpha(N_w, \delta^2) \). Since changes in \( \delta \) do not affect these angles and because there are finitely many hyperfaces, there exists \( K_1(w) \) with \( |dD(\delta)/d\delta| \geq K_1(w) > 0 \) outside of \( \Psi_\alpha(w, N_w, \rho, \mathcal{W}) \). Changing \( N_w \) corresponds to rescaling the sets \( T^1_\alpha(N_w, \delta^1) = \frac{1}{N^1_w} T^1_\alpha(\delta) \), but since \( \frac{1}{N^1_w} \geq 1 \), it follows that \( K_1 := \min_{w \in C} K_1(w) > 0 \). Therefore, \( K_1 D(w, \delta) \leq d(T^1_\alpha(N_w, \delta^1), T^2_\alpha(N_w, \delta^2)) \), which implies by virtue of \( \text{(18)} \) that

\[
K_1 \|\delta^1 - \delta^2\| + K_1 \|\delta^1 - \delta^2\| \leq \sup_{w \in C} \left( \frac{N^1_w}{T^1_w} + \frac{N^2_w}{T^2_w} \right) \|\chi(a, N, \delta)\|.
\]

Because \( N_w \) is bounded away from coordinate directions \( ((w, N_w) \not\in \mathcal{P}) \), the supremum on the right-hand side is finite. Since \( \|\chi(a, N, \delta)\| \leq \|\chi\| \), this proves the first inequality of \( \text{(17)} \) for constant \( K \)

\[
K = \frac{1}{K_1} \sup_{w \in C} \left( \frac{N^1_w}{T^1_w} + \frac{N^2_w}{T^2_w} \right).
\]

The second inequality follows once we show that there exists \( K_3 > 0 \) with

\[
\|\phi(a, N, \hat{\delta})\| - \|\phi\| \leq K_3 D(\delta) + K_3 \|\chi\|.
\]

Indeed, the right hand side of \( \text{(19)} \) is bounded by \( K_3(K + 1) \|\chi\| \) due to the already established inequality. Thus,

\[
1 - \frac{\|\phi\| - \alpha \|\chi\|}{\|\phi(a, N, \hat{\delta})\|} \leq \frac{\|\phi(a, N, \hat{\delta})\| - \|\phi\| + \alpha \|\chi\|}{\tilde{\Psi}} \leq \frac{K_3(K + 1) + \alpha}{\tilde{\Psi}} \|\chi\|.
\]

Observe that \( \tilde{\Psi} > 0 \) since \( \mathcal{N}_C \) is bounded away from \( \Gamma(r, D) \cup \mathcal{P} \) by closedness. The second inequality in \( \text{(17)} \) then follows from \( \text{(20)} \) in conjunction with \( 1-x \geq \frac{1}{2} (1-x^2) \).

It remains to show \( \text{(19)} \). Suppose first that \( \|\phi\| \geq \|\phi(a, N, \delta)\| \). Then Lipschitz continuity of \( \phi \) implies that \( \|\phi\| \geq \|\phi(a, N, \hat{\delta})\| - K_\phi \|\hat{\delta} - \delta\| \), which readily implies \( \text{(19)} \). Suppose therefore that \( \|\phi\| < \|\phi(a, N, \delta)\| \). Let \( d_i \) denote the distance of \( \phi \) from \( T^i_\alpha(N, \delta) \) for \( i = 1, 2 \) and observe that \( d_i \leq N^i/T^i \|\chi\| \) as illustrated in the left panel of Figure \( \text{14} \). Define the auxiliary sets \( \tilde{T}^i_\alpha(N, \delta) := T^i_\alpha(N, \delta^i) - N^i/T^i \chi(a, N, \delta) \) so that \( \phi(a, N, \delta) \in \tilde{T}^1_\alpha(N, \delta^1) \cap \tilde{T}^2_\alpha(N, \delta^2) \) as shown in the right panel of Figure \( \text{14} \). Let \( d_i \) denote the distance of \( \phi \) from \( \tilde{T}^i_\alpha(N, \delta) \) and observe that \( d_i \leq d_i \). Let \( q_i \) for \( i = 1, 2 \) denote the point in \( \partial \tilde{T}^i_\alpha(N, \delta) \) closest to \( \phi \) and let \( \phi' \) be the projection of \( \phi \)
onto the plane through \( \phi(a, N, \delta), q_1 \) and \( q_2 \). Let \( j \in \{1, 2\} \) be the index \( i \) for which the angle \( \theta_i \) between \( \phi(a, N, \delta) - \phi' \) and \( \phi(a, N, \delta) - q_i \) is maximal. Then \( \theta_j \geq \gamma/2 \), where \( \gamma \) is the angle between \( \phi(a, N, \delta) - q_1 \) and \( \phi(a, N, \delta) - q_2 \). Let \( \alpha_i \) be the angles between \( \phi(a, N, \delta) - \phi \) and \( \phi(a, N, \delta) - q_i \) and observe that \( \alpha_i \geq \theta_i \). Then

\[
\|\phi(a, N, \delta) - \phi\| = d_j \left(1 + \frac{1}{\tan(\alpha_j)}\right) \leq \left(1 + \frac{1}{\tan(\gamma/2)}\right) \frac{N_j}{T_j} \|\chi\|
\]

as illustrated in Figure 15. Observe that it is impossible for \( \gamma \) to be 0 by the definition of \( \phi(a, N, \delta) \). Since changes in \( N \) and \( \delta \) do not change the direction of the hyperplanes bounding \( \tilde{T}_a(N, \delta) \), a uniform lower bound \( \gamma \) for \( \gamma \) is given by taking the minimum over all strictly positive angles between the finitely many hyperfaces of \( \partial \tilde{T}_a(N, \delta) \) and \( \partial \tilde{T}_a(N, \delta) \). Therefore, \( \|\phi(a, N, \delta) - \phi\| \leq K_4 \|\chi\| \) for

\[
K_4 = \left(1 + \frac{1}{\tan(\gamma/2)}\right) \sup_{w \in C} \left(\frac{N_w^1}{T_w^1} + \frac{N_w^2}{T_w^2}\right).
\]

[19] now follows from the triangle inequality

\[
\|\phi(a, N, \delta) - \phi\| \leq \|\phi(a, N, \delta) - \phi(a, N, \delta)\| + \|\phi(a, N, \delta) - \phi\| \leq K_3 \|\delta\| + K_4 \|\chi\|. \quad \square
\]

**Lemma E.4.** Let \( C \) be a \( C^1 \) solution to [3] for fixed \( r, \mathcal{W} \) oriented by \( w \mapsto N_w \) such that \( \mathcal{N}_c \cap (\Gamma_a(r, \mathcal{W}) \cup E_a(r, \mathcal{W}) \cup \mathcal{P}) = \emptyset \) for some \( a \in \mathcal{A} \). Then there exists \( K > 0 \) such that for any \( w \in C \), any pair \( (T_w \phi + N_w \chi, \delta) \) that enforces \( a \) with \( w + r \delta(y) \in \mathcal{W} \) for every \( y \in Y \) and \( N_w^+(g(a) + \delta \lambda(a) - w) \geq 0 \) satisfies \( K \leq \|\chi\| \).

**Proof.** Let \( (T_w \phi + N_w \chi, \delta) \) enforce \( a \) with \( w + r \delta(y) \in \mathcal{W} \) for every \( y \in Y \) and \( N_w^+(g(a) + \delta \lambda(a) - w) \geq 0 \). The condition \( (C, \mathcal{N}_c) \cap \Gamma(r, \mathcal{W}) = \emptyset \) implies that \( T_w \phi + N_w \chi \neq 0 \). The condition \( (C, \mathcal{N}_c) \cap E_a(r, \mathcal{W}) = \emptyset \) implies that \( \chi \neq 0 \) because \( \Psi_a(w, N_w, r, \mathcal{W}) = 0 \) for any \( w \in C \). This is equivalent to \( T_a^1(N_w, \delta) \cap T_a^2(N_w, \delta) = \emptyset \), implying that the two sets \( T_a^1(N_w, \delta), T_a^2(N_w, \delta) \) are strictly separated since they are closed. Let \( d(w, \delta) \) denote the minimal distance between the two sets. Because \( N_w \) is bounded away from coordinate directions, the map \( (N_w, \delta) \mapsto T_a^1(N_w, \delta) \) is
continuous for \(i = 1, 2\), hence so is \(d(w, \delta)\). Let \(\mathcal{J}_a(w)\) denote the set of all \(\delta\), for which there exists \(\beta\) such that \((\beta, \delta)\) enforces \(a\) with \(w + r(\delta)(y) \in W\) for every \(y \in Y\) and \(N_w^{-}(g(a) + \delta \lambda(a) - w) \geq 0\). Since \(C\) is \(C^1\), \(\mathcal{J}_a\) is continuous in \(w \in C\). The minimum of \(\min_{\delta \in \mathcal{J}_a(w)} d(w, \delta)\) over the compact set \(C\) is attained, hence positive. This implies the statement by virtue of \((18)\).

**Lemma E.5.** Let \(\varepsilon_1 \leq \varepsilon_2\) and let \(\mathcal{D}_1(w)\) and \(\mathcal{D}_2(w)\) be two maps of class \(B\) such that there exists \(\varepsilon > 0\) with \(\mathcal{D}_2(w) \cap H_{\varepsilon_2}(N) \subseteq \mathcal{D}_1(w) \cap H_{\varepsilon_1}(N) \subseteq \mathcal{D}_2(w) \cap H_{\varepsilon_1}(N) + B_{\varepsilon}(0)\) for every \(w \in \mathcal{V}\) and every \(N \in S_1\). Let \((w, N)\) such that \((\varepsilon_1, \mathcal{D}_1(w))\) and \((\varepsilon_2, \mathcal{D}_2(w))\) is Lipschitz continuous in a neighbourhood \(U\) of \((w, N)\). Let \(\mathcal{C}_1\) and \(\mathcal{C}_2\) be two solutions to \((12)\) with \((\varepsilon_1, \mathcal{D}_1(w))\) and \((\varepsilon_2, \mathcal{D}_2(w))\), respectively, with initial value \((w, N)\) such that \(\mathcal{N}_{\mathcal{C}_1}, \mathcal{N}_{\mathcal{C}_2} \subseteq U\). Then there exist constants \(K_1, K_2, K_3\) such that for any \(v \in \mathcal{C}_1\), there exists \(v' \in \mathcal{C}_2\) with

\[
\|v - v'\| \leq K_1\varepsilon(\|v - w\|^2 + K_2e^{K_3\|v - w\|}).
\]

**Proof.** Let \(f\) and \(h\) be parametrizations of \(\mathcal{C}_1\) and \(\mathcal{C}_2\), respectively, in the direction of \(N\). Let \(w\) be the origin. Then \(f\) and \(h\) are solutions to

\[
f''(x) = F(x, f(x), f'(x)), \quad h''(x) = H(x, h(x), h'(x)) \tag{21}
\]

with \(f(0) = h(0) = 0\) and \(f'(0) = h'(0) = 0\) for Lipschitz continuous \(F\) and \(H\) with Lipschitz constants \(K_F\) and \(K_H\), respectively. By Lemma B.1, the right hand side of \((12)\) is Lipschitz continuous in \(\delta\) for \((v, N_v) \in U\) with some Lipschitz constant \(K\). The condition that \(\mathcal{D}_2(w) \cap H_{\varepsilon_2}(N) \subseteq \mathcal{D}_1(w) \cap H_{\varepsilon_1}(N) \subseteq \mathcal{D}_2(w) \cap H_{\varepsilon_1}(N) + B_{\varepsilon}(0)\) implies \(0 \leq F(x, d, v) - H(x, d, v) \leq K\sqrt{|Y|}\varepsilon\), hence integrating \((21)\) yields

\[
f'(x) - h'(x) = \int_0^x F(t, f(t), f'(t)) - H(t, f(t), f'(t)) + H(t, f(t), f'(t)) - H(t, h(t), h'(t))\ dt \\
\leq K\sqrt{|Y|}\varepsilon x + K_H\int_0^x |f(t) - h(t)| + |f'(t) - h'(t)|\ dt
\]

Since \(H(x, d, v) \leq F(x, d, v)\) in a neighbourhood of 0, we may assume \(f'(x) > h'(x)\) and \(f(x) > h(x)\) by choosing \(U\) small enough. Therefore, \(f - h\) satisfies the conditions of Theorem 1.8.1 in Pachpatte [14], which implies that

\[
f'(x) - h'(x) \leq K\sqrt{|Y|}\varepsilon \left(x + K_F\int_0^x t + \frac{t^2}{2} + \frac{1}{8K_F^2}e^{2K_F t}\ dt\right).
\]

Let \(c_1 = 2K_F \vee 1\) and \(c_2 = c_1(K_F + 1/(8K_F^2))\). Using the inequality \(t + t^2/2 \leq e^t\), we obtain \(f'(x) - h'(x) \leq K\sqrt{|Y|}\varepsilon(x + c_2e^{c_1 x})\). Integrating once yields

\[
f(x) - h(x) \leq K\sqrt{|Y|}\varepsilon\left(\frac{x^2}{2} + \frac{c_2}{c_1}e^{c_1 x}\right).
\]

For \(v = (x, h(x))\), let \(v' = (x, f(x))\), hence the result follows from \(x \leq \|v - w\|\).  \(\square\)
References


