Games and Economic Behavior 122 (2020) 83–104

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Solving two-state Markov games with incomplete information on one side ☆

Galit Ashkenazi-Golan [a], Catherine Rainer [b], Eilon Solan [a,*]

[a] The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel
[b] Université de Bretagne Occidentale, 6, avenue Victor-le-Gorgeu, B.P. 809, 29285 Brest cedex, France

ARTICLE INFO

Article history:
Received 5 August 2019
Available online 18 April 2020

JEL classification:
C72
C73
C63

Keywords:
Repeated games with incomplete information on one side
Markov games
Value
Optimal strategy
Algorithm

ABSTRACT

We study the optimal use of information in Markov games with incomplete information on one side and two states. We provide a finite-stage algorithm for calculating the limit value as the gap between stages goes to 0, and an optimal strategy for the informed player in the limiting game in continuous time. This limiting strategy induces an ϵ-optimal strategy for the informed player, provided the gap between stages is small. Our results demonstrate when the informed player should use her information and how.

1. Introduction

In most strategic interactions, the players are not fully informed of the game's parameters, like their opponents' action sets and payoff functions, and sometimes even their own payoff function and the identity of the opponents. This observation motivates the study of games with incomplete information, which was incepted in the fifties. Harsanyi (1967) introduced the model of Bayesian games, which are one-stage games with incomplete information. Aumann and Maschler (1968, 1995) studied repeated games with incomplete information on one side, provided an elegant characterization to the value, and described optimal strategies for the players. This characterization has been extended to continuous-time games by Cardaliaguet (2006) (see also Cardaliaguet and Rainer (2009a,b), Grün (2012a), and Oliu-Barton (2015)), and to repeated games with incomplete information on both sides (see, e.g., Aumann et al. (1968) and Mertens and Zamir (1971)). For recent surveys on the topic, see Aumann and Heifetz (2002); Mertens et al. (2016).

In repeated games with incomplete information, the parameters of the game remain fixed throughout the interaction. Sometimes these parameters change along the play, in a way that is independent of the players' actions. When the state

☆ The authors thank two anonymous referees for their useful comments, and acknowledge the support of the COST Action 16228, the European Network of Game Theory. Ashkenazi-Golan acknowledges the support of the Israel Science Foundation, grant #520/16. Solan acknowledges the support of the Israel Science Foundation, grant #217/17.

* Corresponding author.
E-mail addresses: galit.ashkenazi@gmail.com (G. Ashkenazi-Golan), Catherine.Rainer@univ-brest.fr (C. Rainer), eilons@post.tau.ac.il (E. Solan).

https://doi.org/10.1016/j.geb.2020.04.004
0899-8256/

changes along the play according to a Markov process that is independent of the players behavior, the game is called a *Markov game*. For example, changes in global markets affect local consumers and producers, each of which has a negligible effect on the global market.

One illustrative example, adapted from Hernández and Neeman (2018), involves the allocation of enforcement resources across different locations with the goal of deterring unwanted behavior, like illegal parking. The compliance rate of the agents depends on their information about the amount of enforcement resources. When the amount of enforcement resources is private information of the enforcing agency, and it varies over time, according to, say, the budget allocated for enforcement or the number of other tasks that the enforcing agency should perform, the situation can be modeled as a two-player zero-sum Markov game between the agency and the representative agent.

Two-player zero-sum Markov games with incomplete information on one side have been first studied by Renault (2006), who proved the existence of the uniform value; see also Neyman (2008) and Hörner et al. (2010). Since the state changes over time, the optimal strategy typically involves a repeated revelation of information. Recently Cardaliaguet et al. (2016) studied discounted two-player zero-sum Markov games with incomplete information on one side, where the time duration between stages goes to 0, and characterized the limit value function and the limit optimal strategy of the informed player; see also Gensbittel (2016, 2019). Gensbittel and Renault (2015) studied Markov games when both players have partial information on the state variable.

This paper is part of a project whose goal is to study the optimal use of information in dynamic situations of incomplete information, and to provide an easy to use algorithm for calculating the value and optimal strategies. We study Markov games with incomplete information on one side and two states. Player 1 knows when the state changes, while Player 2 does not know it. The players observe each other's actions and have perfect recall, and thus Player 2 may use past actions of Player 1 to deduce the identity of the current state.

We study the limit value of this game as the gap between stages goes to 0. Consequently, the discount factor as well as the transition probabilities from one state to the other depend on the gap between stages. We will provide an algorithm for calculating the limit value of this game, when the gap between stages and the transition probabilities between the states go to 0, as well as an ϵ-optimal strategy for the informed player, when the gap between stages is small.

In addition to the literature on Markov games, our paper is related to two strands of literature. The literature on dynamic information provision (or dynamic persuasion), as studied, for example, by Renault et al. (2017) and Ely (2017), studies two-player nonzero-sum Markov games where an informed player tries to induce an uninformed player to act in a certain way that is desirable to the informed player. The optimization problem faced by the informed player in this model has similarities to the one in our setup.

Our paper is also related to the literature on numerical schemes for differential (resp. stochastic differential) games with asymmetric information, see, e.g., Cardaliaguet (2009) and Grün (2012b). Instead of constructing the value function on a discrete time and space grid, our construction permits to determine the areas corresponding to the different regimes of the revelation process. It can therefore be linked with the geometric construction of the value function for Dynkin games with asymmetric information in Gensbittel and Grün (2019).

To get some intuition to the problem, we contrast it with the case of repeated games with incomplete information on one side, in which the state does not change along the play, as studied by Aumann and Maschler (1995). Denote by π_n the belief of the uninformed player on the state at stage n. The process $(\pi_n)_{n\in\mathbb{N}}$ is a martingale that is controlled by the informed player. By the Martingale Convergence Theorem the process $(\pi_n)_{n\in\mathbb{N}}$ converges to a limit π_∞, which implies that as the game evolves information stops being revealed. In particular, under the optimal strategies of the players, the stage payoff will converge to $u(\pi_\infty)$, the value of the one-stage game in which the state is chosen according to the probability distribution π_∞ and no player is informed of the chosen state. It can then be proven that the value of the game is the concavification of the function u, that is, the smallest concave function that is larger than or equal to u. This, in turn, implies that the informed player has an optimal strategy in which information is revealed only at the first stage of the play.

Since the state changes along the play, the process $(\pi_n)_{n\in\mathbb{N}}$ is no longer a martingale; indeed, in addition to its dependence on the informed player's actions, the belief has a drift towards the stationary distribution of the associated Markov chain,[1] denoted p^*. We study the discounted game, and are interested in the optimal way of information revelation. As in the case of repeated games with incomplete information on one side, there are two ways in which the informed player can use her information at stage n:

A.1 The informed player may elect not to reveal any information at that stage. The optimal stage payoff is then $u(\pi_n)$, and the belief changes because of the drift towards p^*.

A.2 The informed player may elect to split the belief between two other beliefs: for some $p', p'' \in [0,1]$ and some $q \in [0,1]$, the informed player plays in such a way that $\pi_{n+1} = p'$ with probability q and $\pi_{n+1} = p''$ with probability $1-q$.

Since we study the limit value as the gap between stages goes to 0, it will be more convenient to consider the process in continuous time. One can expect that in continuous time, the interval $[0,1]$ of beliefs will be divided into subintervals,

[1] To be precise, p^* is the probability of state s_1 under the stationary distribution (and $1-p^*$ is the probability of state s_2 under the stationary distribution).

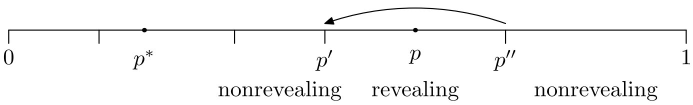

Fig. 1. Possible information revelation.

State s_1	L	R
T	1	0
B	0	0

State s_2	L	R
T	0	0
B	0	1

Fig. 2. The payoff matrices in Example 1.

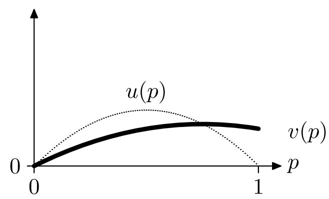

Fig. 3. The value function in Example 1.

as depicted in Fig. 1. When the belief is in some subintervals, the informed player will reveal no information, and due to the transition, the belief of the uninformed player will *slide* towards the invariant distribution p^*. When the belief is in the other subintervals, the informed player will reveal information. In the latter case, the stage payoff if no information is revealed is low, hence the informed player will avoid such beliefs.

This gives rise to two types of information revelations on the side of the informed player:

- If the current belief is within a subinterval $I = [p', p'']$ which the informed player wants to avoid, she will *split* the belief of the uninformed player between the two endpoints of the interval, namely, p' and p''.
- If the current belief is the upper end of the subinterval I and if $p^* < p'$, since the belief drifts towards the invariant distribution p^*, the informed player will be able to reveal information in such a way that $\pi_{n+1} \in \{p', p''\}$. This implies that the belief will remain p'' until it *jumps* to p' at a random time.

If the current belief is the lower end of the subinterval I, if $p^* < p'$, and if the behavior of the informed player alternates in adjacent intervals, then p' is the upper end of a subinterval I' in which the informed player reveals no information, and since the belief drifts towards the invariant distribution p^* it will not get into the subinterval $[p', p'']$. If the interval lies below the invariant distribution p^*, the behavior of the informed player is mirrored.

The above description of the general structure of the optimal information revelation strategy is conjectural. To prove that this description is correct, and to provide an algorithm that calculates the value function and the limit optimal strategy of the informed player, we will write down the equations that a value function that is derived from this description must satisfy, and use the characterization of the value function as given by Gensbittel (2019) to show that this intuition is correct.

We illustrate the optimal revelation strategies by three examples from the seminal work of Aumann and Maschler (1995), adapted to our model. We denote the two states by s_1 and s_2, and denote by p the probability that the uninformed player assigns to state s_1. For expositional ease, in these examples state s_2 is absorbing: once the play reaches it, it remains there forever. In particular, the invariant distribution is $p^* = 0$.

Example 1 *(Nonrevealing optimal strategy).* Consider the Markov game with the payoff matrices that appear in Fig. 2. For every $p \in [0, 1]$ the value of the one-stage game is $u(p) = p(1 - p)$ (see Aumann and Maschler (1995), Section I.2 and the dotted line in Fig. 3). The value function is concave, and the informed Player 1 has no incentive to reveal information to Player 2. Consequently, the limit optimal strategy consists of playing the myopic optimal strategy. The belief, which starts at the initial belief, slides towards the invariant distribution $p^* = 0$, and the limit value function is given by the discounted integral of the function u (see the dark line in Fig. 3).

Example 2 *(Revealing optimal strategy).* Consider the game with the payoff matrices that appear in Fig. 4. For every $p \in [0, 1]$ the value of the one-stage game is $u(p) = -p(1 - p)$ (see Aumann and Maschler (1995), Section I.3, and the dotted line in Fig. 5). The value function is convex, and the informed Player 1 has incentive to reveal her information to Player 2. Consequently, in the optimal strategy Player 1 reveals her information at every stage, and the limit value function is identically 0 (see the dark line in Fig. 5).

We now exhibit a nontrivial and challenging case, where the function u is neither convex nor concave.

State s_1	L	R
T	-1	0
B	0	0

State s_2	L	R
T	0	0
B	0	-1

Fig. 4. The payoff matrices in Example 2.

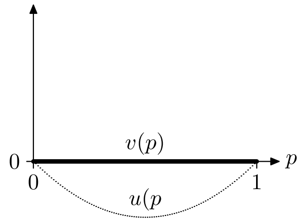

Fig. 5. The value function in Example 2.

State s_1	L	R
T	1	0
B	0	2

State s_2	L	R
T	-2	0
B	0	-1

Fig. 6. The payoff matrices in Example 3.

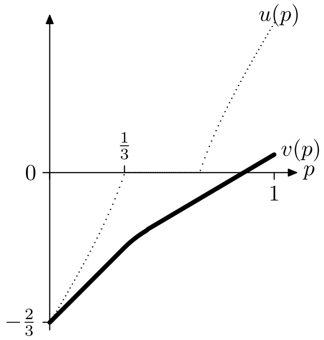

Fig. 7. The value function in Example 3.

Example 3 *(Partial revelation of information).* Consider the Markov game with the payoff matrices that appear in Fig. 6. For every $p \in [0, 1]$ the value of the one-stage game is given by (see Aumann and Maschler (1995), Section I.4 and the dotted line in Fig. 7)

$$u(p) = \begin{cases} \frac{9p^2-9p+2}{6p-3}, & \text{if } 0 \le p \le \frac{1}{3}, \\ 0 & \text{if } \frac{1}{3} < p \le \frac{2}{3}, \\ \frac{9p^2-9p+2}{6p-3} & \text{if } \frac{2}{3} < p \le 1. \end{cases} \tag{1}$$

For $p \le \frac{1}{3}$ the function u is convex, hence it is optimal for Player 1 to reveal some of her information. She should therefore pick some $p_0 \ge \frac{1}{3}$ and split the belief of Player 2 between $p = 0$ and $p = p_0$. Do we have $p_0 = \frac{1}{3}$ or $p_0 > \frac{1}{3}$?

Consider next the case that the initial belief is $p = 1$. If Player 1 reveals no information, at every stage k in which the belief is p_k she obtains the payoff $u(p_k)$, and the belief drifts towards 0. Consequently, the payoff slides down the graph of u. Since in the interval $\frac{1}{3} \le p \le 1$ the graph of u lies below the line segment that connects the points $(\frac{1}{3}, u(\frac{1}{3}))$ and $(1, u(1))$, it is not optimal for Player 1 to hide her information throughout: when the belief reaches some point $p_1 \in [\frac{1}{3}, 1]$ she should start revealing information. What is this point p_1? How much information does Player 1 reveal? We will answer these questions and provide an algorithm that describes the limit strategy in the general case.

The paper is organized as follows. The model as well as known results appear in Section 2. Section 3 details the algorithm and discusses its relation to ϵ-optimal strategies of the informed player, Section 4 demonstrates the algorithm on few examples. Section 5 discusses possible extensions and open problems, and Section 6 proves the correctness of the algorithm.

2. The model

In this paper we study two-player zero-sum Markov games, which were first studied in Renault (2006). A two-player zero-sum *Markov game* G is a vector $(S, A, B, g, \delta, x^1, x^2, p)$ where

- $S = \{s_1, s_2\}$ is the set of states.
- A and B are finite action sets for the two players.
- $g : S \times A \times B \to \mathbb{R}$ is a payoff function.
- δ is the discount rate.
- x^1 and x^2 are the rates of transition.
- p is the prior probability that the initial state is s_1.

The game is played as follows. The initial state is s_1 with probability p, and s_2 with probability $1 - p$. At every stage $k \geq 1$ the players choose independently and simultaneously actions a^k and b^k in their action sets. If $s^k = s_1$, then the new state s^{k+1} is equal to s_1 with probability $1 - x^1$ and to s_2 with probability x^1. If $s^k = s_2$, then the new state s^{k+1} is equal to s_2 with probability $1 - x^2$ and to s_1 with probability x^2.

For every finite set Y, denote by $\Delta(Y)$ the set of probability distributions over Y. We assume that information is asymmetric: Player 1 knows the current state while Player 2 does not. In addition, we assume perfect recall. Consequently, a strategy of Player 1 is a sequence $\sigma = (\sigma_k)_{k\geq 1}$, where $\sigma_k : S^k \times (A \times B)^{k-1} \to \Delta(A)$ for every $k \geq 1$. A strategy for Player 2 is a sequence $\tau = (\tau_k)_{k\geq 1}$, where $\tau_k : (A \times B)^{k-1} \to \Delta(B)$ for every $k \geq 1$. The sets of strategies of Player 1 and Player 2 are denoted by $\mathcal{S}$ and $\mathcal{T}$, respectively. Every pair of strategies $(\sigma, \tau) \in \mathcal{S} \times \mathcal{T}$, together with the prior belief p, induces a probability distribution on the space $(S \times A \times B)^{\mathbb{N}}$ of *plays*, where $\mathbb{N}$ is the set of positive integers, and the payoff is given by

$$g(p, \sigma, \tau) := E_{p,\sigma,\tau}\left[\sum_{k\geq 1} \delta(1-\delta)^{k-1} g(s^k, a^k, b^k)\right].$$

Player 1 is the maximizer and Player 2 is the minimizer. Hence, the value of the game G is given by

$$v := \max_{\sigma\in\mathcal{S}} \min_{\tau\in\mathcal{T}} g(p, \sigma, \tau) = \min_{\tau\in\mathcal{T}} \max_{\sigma\in\mathcal{S}} g(p, \sigma, \tau). \tag{2}$$

The value exists because the payoff is discounted and the strategy spaces of the players are compact in the product topology. A strategy σ (resp. τ) of Player 1 (resp. Player 2) that achieves the maximum (resp. minimum) in the second (resp. third) term in Eq. (2) is called *optimal*.

We will be interested in the value of the game and in the optimal strategy of Player 1 when the duration between stages is small. Consequently, we will parameterize the game with a parameter $n > 0$, that will capture the duration between stages. Thus, given three positive real numbers r, λ_1, and λ_2, we denote by $G^{(n)}(p)$ the Markov game $(S, A, B, g, 1 - e^{r/n}, 1 - e^{\lambda_1/n}, 1 - e^{\lambda_2/n}, p)$. We denote by $v^{(n)}(p) = v^{(n)}(p, r, \lambda_1, \lambda_2)$ the value of the game $G^{(n)}(p)$. It follows that the rates of switching states are roughly λ_1/n and λ_2/n, and therefore the limit invariant distribution as n goes to infinity is

$$p^* := \frac{\lambda_2}{\lambda_1 + \lambda_2}.$$

Denote also

$$\mu := \frac{r}{\lambda_1 + \lambda_2}.$$

By Cardaliaguet et al. (2016) the limit $v := \lim_{n\to\infty} v^{(n)}$ exists and the limit as n goes to infinity of the optimal strategy of Player 1 can be characterized as the solution of a certain optimization problem. We now describe this result. Extend the domain of the payoff function g to $S \times \Delta(A) \times \Delta(B)$ in a bilinear fashion:

$$g(s, x, y) = \sum_{a\in A}\sum_{b\in B} g(s, a, b)x(a)y(b), \quad \forall (x, y) \in \Delta(A) \times \Delta(B), s \in S.$$

For $p \in [0, 1]$, the value of the one-stage game given that the two states s_1 and s_2 are observed by none of the players and s_1 is the current state with probability p (and s_2 with probability $1 - p$) is

$$u(p) := \max_{x\in\Delta(A)} \min_{y\in\Delta(B)} \big(pg(s_1, x, y) + (1 - p)g(s_2, x, y)\big).$$

Let $p \in [0, 1]$ be given and let $(\Omega, \mathcal{F}, P)$ be a sufficiently large probability space. Let $\mathcal{S}(p)$ be the set of all càdlàg, $[0, 1]$-valued processes $(\pi_t)_{t\geq 0}$ defined over $(\Omega, \mathcal{F}, P)$ that satisfy $E[\pi_0] = p$ and $E[\pi_t|\mathcal{F}^{\pi_\cdot}_s] = p^* + (\pi_s - p^*)e^{-(\lambda_1+\lambda_2)(t-s)}$ for every $0 \leq s \leq t$, where $\mathcal{F}^{\pi_\cdot}_t$ is the σ-algebra generated by $(\pi_s)_{s\leq t}$.

Theorem 2.1 *(Cardaliaguet et al., 2016, Theorem 1. P1). The sequence of functions $p \mapsto v^{(n)}(p)$ converges uniformly to a function $v : [0, 1] \to \mathbb{R}$ that satisfies*

$$v(p) = \max_{(\pi_t)_{t\geq 0}\in\mathcal{S}(p)} E\left[\int_0^\infty re^{-rt}u(\pi_t)dt\right], \quad \forall p \in [0,1]. \tag{3}$$

The function v is termed the *limit value function*. The processes $(\pi_t)_{t\geq 0} \in \mathcal{S}(p)$ in Eq. (3) represent the possible revelation mechanisms induced by the actions of the informed player. In particular, the process that realizes the maximum in Eq. (3) represents the optimal revelation process for the continuous-time game, and is termed the *limit optimal strategy* for the informed player. The characterization of v provided by Cardaliaguet et al. (2016) is via a differential equation, as summarized by the next result.

Theorem 2.2 *(Cardaliaguet et al., 2016, Theorem 1. P2). The limit value function v is the unique viscosity solution of the equation*

$$\min\{rv(p) - \langle {}^tRp, Dv(p)\rangle - ru(p); -\lambda_{max}v(p, D^2v(p))\} = 0, \ \forall p \in \Delta(2),$$

where $R = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$ is the generator of the Markov chain, $\Delta(2)$ is the simplex in $\mathbb{R}^2$, and $\lambda_{max}v(p, D^2v(p))$ is the maximal eigenvalue of the restriction of $D^2v(p)$ to the tangent space at p to $\Delta(2)$.

Since in the sequel we will not need the notion of viscosity solution, we do not provide their definition, and refer to Cardaliaguet et al. (2016) for the definition used in the above theorem. In Cardaliaguet et al. (2016) it is also shown how the optimal solution $(p_t)_{t\geq 0}$ in Eq. (3) can be used to identify ϵ-optimal strategies for the informed player in the discrete-time game $G^{(n)}(p)$, provided n is sufficiently large.

Gensbittel (2019) reformulates Theorem 2.2 in terms of directional derivatives. Using the fact that in the two-state case the resulting equations are one-dimensional, we can prove that the limit value function v is differentiable on $[0,1] \setminus \{p^*\}$. This leads to the following simple characterization of v that involves only an ordinary differential equation.

Recall that the *hypograph* of a function $f : [0,1] \to \mathbb{R}$ is the set of all points that are on or below the graph of the function. When f is concave, its hypograph is a convex set, and its set of extreme points coincides with the set of points on the graph of f where f is not affine, plus the corner points $(0, f(0))$ and $(1, f(1))$.

Theorem 2.3. *The function v is the unique continuous, concave function $v : [0,1] \to \mathbb{R}$ which is differentiable on $[0,1]$ except, possibly, at p^*, and that satisfies the following conditions:*

G.1 *$v(p^*) \geq u(p^*)$, with an equality if $(p^*, v(p^*))$ is an extreme point of the hypograph of v.*
G.2 *For every $p \in [0,1] \setminus \{p^*\}$ we have $v'(p)(p - p^*) + \mu(v(p) - u(p)) \geq 0$.*
G.3 *For every extreme point $(p, v(p))$ of the hypograph of v such that $p \neq p^*$ we have*

$$v'(p)(p - p^*) + \mu(v(p) - u(p)) = 0, \tag{4}$$

where for $p = 0$ (resp. $p = 1$), $v'(p)$ stands for the right (resp. left) derivative of v at p.

Proof. By Theorem 2.12 in Gensbittel (2019), the limit value function is the unique concave, Lipschitz function that satisfies

$$r(v(p) - u(p)) - \overrightarrow{D}V(\mathbf{p}, {}^tR\mathbf{p}) \geq 0, \quad \forall p \in [0,1], \mathbf{p} = (p, 1-p),$$

and, if $(p, v(p))$ is an extreme point of the hypograph of v,

$$r(v(p) - u(p)) - \overrightarrow{D}V(\mathbf{p}, {}^tR\mathbf{p}) \leq 0, \tag{5}$$

where $\overrightarrow{D}V(\mathbf{p}, \cdot)$ is the directional derivative of $V : \Delta(2) \ni (p, 1-p) \mapsto V(p, 1-p) := v(p)$, and $R = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$. It follows that

$$\overrightarrow{D}V(\mathbf{p}, {}^tR\mathbf{p}) = \begin{cases} -v'_+(p)\frac{r}{\mu}(p - p^*) & \text{if } p < p^*, \\ -v'_-(p)\frac{r}{\mu}(p - p^*) & \text{if } p > p^*, \\ \overrightarrow{D}V(\mathbf{p}, 0) = 0 & \text{for } \mathbf{p} = (p^*, (1-p^*)). \end{cases} \tag{6}$$

Statement G.1 follows from the above discussion.

Statements G.2 and G.3 will follow once we show that $v'_-(p) = v'_+(p)$ for every $p \in (0,1) \setminus \{p^*\}$.
Suppose first that $p > p^*$. This implies that for every $q > p$ it holds that $q > p^*$, hence $\frac{r}{\mu}(q - p^*) > 0$. Since v is concave, it also implies that $v'_-(q) \leq v'_+(p) \leq v'_-(p)$. From Eqs. (5) and (6) it follows that

$$0 \le \mu(v(q) - u(q)) + v'_-(q)(q - p^*) \le \mu(v(q) - u(q)) + v'_+(p)(q - p^*).$$

From the continuity of u and v we deduce that

$$0 \le \mu(v(p) - u(p)) + (p - p^*)v'_+(p) \le \mu(v(p) - u(p)) + (p - p^*)v'_-(p).$$

If $(p, v(p))$ is an extreme point of the hypograph of v, then

$$\mu(v(p) - u(p)) + (p - p^*)v'_-(p) = 0,$$

and it follows that $v'_-(p) = v'_+(p)$. If $(p, v(p))$ is not an extreme point of the hypograph of v, then there exist $p_1, p_2 \in [0, 1]$ and $\alpha \in (0, 1)$ such that $p = \alpha p_1 + (1 - \alpha)p_2$ and $v(p) = \alpha v(p_1) + (1 - \alpha)v(p_2)$. Since v is concave, it follows that v is affine on the interval $[p_1, p_2]$, and therefore differentiable on its interior. In particular, $v'_-(p) = v'_+(p)$ in this case as well.

Suppose now that $p < p^*$. In this case for $q < p$ we have $\frac{r}{\mu}(q - p^*) < 0$ and $v'_+(p) \le v'_-(p) \le v'_+(q)$ for every $q < p$, and an analogous argument to the one provided above leads to the same result: $v'_-(p) = v'_+(p)$. We conclude that statements G.2 and G.3 hold as well. □

Remark 2.4. The arguments of the proof of G.2 and G.3 cannot be used for $p = p^*$, because ${}^tRp^* = 0$. In fact, the function v may not be differentiable at p^*, see variation b of Example 3 below.

3. An algorithm to calculate the value function and the optimal revelation process

In this section we present a finite stage recursive algorithm for calculating the limit value function and the limit optimal strategy for the informed player. We start by explaining the intuition behind the algorithm.

3.1. Intuition

We shall see that the limit value function at the invariant distribution p^* can be explicitly calculated. The interval $[0, 1]$ will be divided into finitely many subintervals: in some intervals the informed player reveals no information, and she plays in such a way that the belief never enters the remaining intervals.

The algorithm will assume that the limit value was already calculated in a certain closed interval that contains p^*, and will calculate it for a larger interval. The calculation for beliefs smaller than p^* will be analogous to the calculation for beliefs larger than p^*, hence we will concentrate on the latter.

In this section we provide the equations that the limit value function must satisfy under the two types of information revelation that were discussed in the introduction.

3.1.1. No revelation of information

We first provide the equation that the limit value function satisfies in an interval in which no information is revealed by the informed player. Recall that $v^{(n)}(p)$ is the value of the game with time step $1/n$ and initial distribution $p \in [0, 1]$. Let $p^* < p' < p'' \le 1$ and suppose that for every belief $p \in [p', p'']$ of the uninformed player, the optimal strategy of the informed player is not to reveal her information. The continuation payoff is given by $v^{(n)}(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n}))$. Therefore the value function satisfies the relation

$$v^{(n)}(p) = (1 - e^{-r/n})u(p) + e^{-r/n}v^{(n)}(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n})). \tag{7}$$

Simple algebraic manipulations yield that

$$\frac{v^{(n)}(p) - v^{(n)}\left(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n})\right)}{p - \left(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n})\right)}$$
$$= \frac{(1 - e^{-r/n})u(p) - (1 - e^{-r/n})v^{(n)}\left(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n})\right)}{p - \left(pe^{-\lambda_1/n} + (1 - p)(1 - e^{-\lambda_2/n})\right)}.$$

Taking the limit as n goes to ∞ and recalling that $\mu = \frac{r}{\lambda_1+\lambda_2}$ and $p^* = \frac{\lambda_2}{\lambda_1+\lambda_2}$, we obtain that the limit value function $v = \lim_{n\to\infty} v^{(n)}$ is the solution of the following differential equation:

$$v'(p) = \frac{\mu(u(p) - v(p))}{p - p^*}. \tag{8}$$

3.1.2. Some revelation of information

Suppose that the informed player wants to avoid beliefs in some open interval (p', p''). In this case, whenever the belief is in the interval (p', p''), the informed player will reveal information in such a way that the belief changes to either p' or p''. In continuous time this implies that after time 0 the belief will never lie in the open interval (p', p''). Therefore, this kind of information revelation will occur at most once, at the first stage of the game. It follows that for every $n \in \mathbb{N}$ the value function $v^{(n)}$ is affine on the interval $[p', p'']$, and therefore so is the limit value function v, that is,

$$v'_-(p'') = \frac{v(p'') - v(p')}{p'' - p'}. \tag{9}$$

Suppose now that at some stage the belief is p''. Two cases can occur. Suppose first that $p^* < p' < p''$ and that moreover in some interval (p'', p''') the informed player reveals no information. Since p'' is an endpoint of an interval in which no information is revealed, by Eq. (8)

$$v'_+(p'') = \frac{\mu(u(p'') - v(p''))}{p'' - p^*}. \tag{10}$$

Since v is smooth at p'', we have $v'_+(p'') = v'_-(p'')$, and therefore by Eqs. (9) and (10) we have

$$v(p'')\left(\frac{1}{p'' - p'} + \frac{\mu}{p'' - p^*}\right) = \frac{v(p')}{p'' - p'} + \frac{\mu u(p'')}{p'' - p^*},$$

or, equivalently,

$$v(p'') = \frac{v(p')(p'' - p^*) + \mu(p'' - p')u(p'')}{p'' - p^* + \mu(p'' - p')}. \tag{11}$$

Substituting $v(p'')$ from Eq. (11) in Eq. (10) we obtain that

$$v'(p'') = \frac{\mu(u(p'') - v(p'))}{p'' - p^* + \mu(p'' - p')}. \tag{12}$$

Since v is affine on (p', p''), we deduce that for every $p \in [p', p'']$ we have

$$v(p) = v(p') + (p - p')\frac{\mu(u(p'') - v(p'))}{p'' - p^* + \mu(p'' - p')}. \tag{13}$$

Suppose now that $p' < p^* < p''$. When the belief is p'' (resp. p'), it remains p'' (resp. p') until it jumps at a random time to p' (resp. p''). Applying Eq. (11) to the jumps from p' to p'' and from p'' to p', we obtain two affine equations in $v(p')$ and $v(p'')$. If for every $p \in (p', p'')$ the informed player splits the belief to p' and p'', we obtain a strategy for the informed player that guarantees a payoff of

$$v(p) = u(p')\frac{(\mu+1)p'' - p^*}{(p'' - p')(\mu+1)} + u(p'')\frac{p^* - p'(\mu+1)}{(p'' - p')(\mu+1)} + p\mu \cdot \frac{u(p'') - u(p')}{(p'' - p')(\mu+1)}, \quad p \in [p', p'']. \tag{14}$$

Remark 3.1. Substituting $p = p^*$ in Eq. (14) we obtain:

$$v(p^*) = \frac{p'' - p^*}{p'' - p'}u(p') + \frac{p^* - p'}{p'' - p'}u(p''). \tag{15}$$

3.1.3. Conclusion

The intuition we presented describes the conjectured behavior of the belief of Player 2 under the optimal strategy of Player 1: in the first stage the belief may split, and thereafter the behavior alternates between sliding continuously towards the invariant distribution p^* and jumping at a random time to a belief closer to or behind p^*.

To find the points where the behavior of the belief changes, we will begin from “the end”, that is, from $p = p^*$, and work our way towards $p = 1$ (and then towards $p = 0$). Supposing that the limit value function was already calculated for every belief p in some interval $[p^*, p_0]$, we compare the incremental value of the two strategies described in Sections 3.1.1 and 3.1.2, find the maximal interval $[p_0, p_1]$ for which the better strategy yields a higher increment, and accordingly extend the definition of the limit value function to the interval $[p^*, p_1]$. We then conduct the analogous procedure for p's smaller than p^*.

3.2. The algorithm to compute the limit value function

In this section we present the algorithm that calculates the limit value function. We will start with some notations. Given an interval $I \subset [0,1)$, we define the function $a : I \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ by

$$a(p,x) := \sup_{p' \in (p,1]} \frac{\mu(u(p') - x)}{p' - p^* + \mu(p' - p)}, \quad \forall p \in I, x \in \mathbb{R}. \tag{16}$$

Analogously, given some interval $I \subset (0,1]$, we define the function $\widetilde{a} : I \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ by

$$\widetilde{a}(p,x) := \inf_{p' \in [0,p)} \frac{\mu(u(p') - x)}{p' - p^* + \mu(p' - p)}.$$

Since the function u is continuous, for $p \neq p^*$ we have

$$a(p,x) = \max_{p' \in [p,1]} \frac{\mu(u(p') - x)}{p' - p^* + \mu(p' - p)} \text{ and } \widetilde{a}(p,x) = \min_{p' \in [0,p]} \frac{\mu(u(p') - x)}{p' - p^* + \mu(p' - p)}.$$

For a fixed $x \in \mathbb{R}$, the function $a(\cdot, x)$ (resp. $\widetilde{a}(\cdot, x)$) is continuous on $I \setminus \{p^*\}$, and, if $x = u(p^*)$ and $I = [p^*, \widehat{p}]$ (resp. $I = [\widehat{p}, p^*]$) for some $\widehat{p}$, then it is continuous on I.

Define also

$$\begin{cases} \rho(p,x) := \sup\left\{p' \in (p,1] : a(p,x) = \frac{\mu(u(p')-x)}{p'-p^*+\mu(p'-p)}\right\} & \text{if } p > p^*, \\ \widetilde{\rho}(p,x) := \inf\left\{p' \in [0,p) : \widetilde{a}(p,x) = \frac{\mu(u(p')-x)}{p'-p^*+\mu(p'-p)}\right\} & \text{if } p < p^*, \end{cases} \tag{17}$$

with $\inf \emptyset = 1$ and $\sup \emptyset = 0$.

To see the motivation for these definitions, recall the discussion in Section 3.1.2. When Player 1 wants to make the belief of Player 2 jump from some $p' > p$ to p, the value function on the interval $[p, p']$ is affine and given by Eq. (11). In particular, the slope of the value function to the left of p' is given by $\frac{\mu(u(p')-v(p))}{p'-p^*+\mu(p'-p)}$, see Eq. (12). To maximize the payoff in a small neighborhood to the right of p, Player 1 will jump to p from some belief p' that attains the maximum in Eq. (16). The quantity $a(p, v(p))$ is defined to be the slope at such optimal belief p', and $\rho(p, v(p))$ is the largest optimal belief. The quantities $\widetilde{a}(p, v(p))$ and $\widetilde{\rho}(p, v(p))$ have analogous interpretations when $p < p^*$.

We now present the algorithm, which defines in steps a function $w : [0,1] \to \mathbb{R}$ that is later shown to be the limit value function. The initial step of the algorithm identifies a closed interval $[\widetilde{p}_0, p_0]$ that includes the stationary distribution p^*, on which the calculation of the limit value function is simple.

The algorithm then defines iteratively an increasing sequence $(p_k)_{k \geq 0}$ of points in the interval $[p_0, 1]$; at the k'th iteration of the algorithm we define the point p_{k+1} and extend the definition of w to include $(p_k, p_{k+1}]$. This part of the algorithm terminates when $p_k = 1$. Finally, the algorithm defines iteratively a decreasing sequence $(\widetilde{p}_k)_{k \geq 0}$ of points in the interval $[0, \widetilde{p}_0]$ and extends the definition of w to include $[\widetilde{p}_{k+1}, \widetilde{p}_k)$. This part of the algorithm terminates when $\widetilde{p}_k = 0$.

Initialization:

Let $p_0 = \inf\{p > p^*, (\text{cav } u)(p) = u(p)\}$ and $\widetilde{p}_0 = \sup\{p < p^*, (\text{cav } u)(p) = u(p)\}$. Define a function $w : [\widetilde{p}_0, p_0] \to \mathbb{R}$ as follows:

- If $\widetilde{p}_0 = p^* = p_0$, then set $w(p^*) = u(p^*)$.
- If $\widetilde{p}_0 < p_0$, then w is defined by (compare this expression with Eq. (14))

$$w(p) := u(\widetilde{p}_0) \frac{p_0(\mu+1) - p^*}{(p_0 - \widetilde{p}_0)(\mu+1)} + u(p_0) \frac{p^* - \widetilde{p}_0(\mu+1)}{(p_0 - \widetilde{p}_0)(\mu+1)} + p\mu \cdot \frac{u(p_0) - u(\widetilde{p}_0)}{(p_0 - \widetilde{p}_0)(\mu+1)}, \quad p \in [\widetilde{p}_0, p_0]. \tag{18}$$

Increasing part of the algorithm:

I.1. Let $k \geq 0$ and suppose that the function w is already defined on the interval $[p_0, p_k]$.
I.2. If $p_k = 1$, the first part of the algorithm terminates; go to Step D.1.
I.3. If $p_k < 1$, let $\varphi_k : [p_k, 1] \to \mathbb{R}$ be the solution of the following differential equation:

$$\begin{cases} \varphi_k(p_k) = w(p_k), \\ \varphi_k'(p) = \frac{\mu(u(p) - \varphi_k(p))}{p - p^*}, \quad p \in (p_k, 1], \end{cases} \tag{19}$$

and set

$$\psi_k(p) := w(p_k) + (p - p_k) a(p_k, w(p_k)), \quad \forall p \in (p_k, 1]. \tag{20}$$

I.4. If $\rho(p_k, w(p_k)) > p_k$, define

$$p_{k+1} := \rho(p_k, w(p_k)). \tag{21}$$

Extend the domain of w to include $(p_k, p_{k+1}]$ by

$$w(p) := \psi_k(p), \quad \forall p \in (p_k, p_{k+1}]. \tag{22}$$

I.5. Otherwise, $\rho(p_k, w(p_k)) = p_k$. Define

$$p_{k+1} := \inf\{p > p_k : \rho(p, \varphi_k(p)) > p\}, \tag{23}$$

with $\inf\emptyset = 1$. Extend the domain of w to include $(p_k, p_{k+1}]$ by

$$w(p) := \varphi_k(p), \quad \forall p \in (p_k, p_{k+1}].$$

I.6. Increase k by 1 and go to Step I.2.

Decreasing part of the algorithm:

D.1. Let $k \geq 0$ and suppose that the function w is already defined on the interval $[\widetilde{p}_k, \widetilde{p}_0]$.
D.2. If $\widetilde{p}_k = 0$, the algorithm terminates.
D.3. If $\widetilde{p}_k > 0$, let $\varphi_k : [0, \widetilde{p}_k] \to \mathbb{R}$ be the solution of the following differential equation:

$$\begin{cases} \varphi_k(\widetilde{p}_k) = w(\widetilde{p}_k), \\ \varphi_k'(p) = \frac{\mu(u(p)-\varphi_k(p))}{p-p^*}, \quad p \in [0, \widetilde{p}_k). \end{cases} \tag{24}$$

Define

$$\widetilde{\psi}_k(p) := w(\widetilde{p}_k) + (p - \widetilde{p}_k)\widetilde{a}(\widetilde{p}_k, w(\widetilde{p}_k)), \quad \forall p \in [0, \widetilde{p}_k]. \tag{25}$$

D.4. If $\rho(\widetilde{p}_k, w(\widetilde{p}_k)) < \widetilde{p}_k$, define

$$\widetilde{p}_{k+1} := \rho(\widetilde{p}_k, w(\widetilde{p}_k)).$$

Extend the domain of w to include $[\widetilde{p}_{k+1}, \widetilde{p}_k)$ by

$$w(p) := \widetilde{\psi}_k(p), \quad \forall p \in [\widetilde{p}_{k+1}, \widetilde{p}_k).$$

D.5. Otherwise, $\rho(\widetilde{p}_k, w(\widetilde{p}_k)) = \widetilde{p}_k$. Define $\widetilde{p}_{k+1} := \sup\{p < \widetilde{p}_k : \rho(p, \varphi_k(p)) < p\}$, with $\inf\emptyset = 0$. Extend the domain of w to include $[\widetilde{p}_{k+1}, \widetilde{p}_k)$ by

$$w(p) := \varphi_k(p), \quad \forall p \in [\widetilde{p}_{k+1}, \widetilde{p}_k).$$

D.6. Increase k by 1 and go to Step D.2.

The idea is that after the initialization, the algorithm determines for each point p_k whether for beliefs slightly above p_k it is optimal for Player 1 to reveal information or to reveal nothing until the belief reaches p_k. The decision is based on comparison of derivatives: the derivative of φ_k, the nonrevealing payoff, is compared to $a(p_k, w(p_k))$, the highest possible derivative when splitting. The strategy that gives the highest derivative is the one that is played, for as long as it's derivative is indeed the higher one. The changes from a revealing strategy to nonrevealing strategy and vice versa occur at the points $(p_k)_{k\geq 0}$ and $(\widetilde{p}_k)_{k\geq 0}$, where the former lower derivative becomes the higher one. Since the derivative from the right is equal to the derivative from the left in points where the behavior of the informed player changes, the corresponding payoff function, and consequently the limit value function, turn out to be differentiable.

On intervals $(p_k, p_{k+1}]$ (resp. $[\widetilde{p}_{k+1}, \widetilde{p}_k)$) where the function w is defined by Step I.4 (resp. D.4), w is linear, while on intervals $(p_k, p_{k+1}]$ (resp. $[\widetilde{p}_{k+1}, \widetilde{p}_k)$) where the function w is defined by Step I.5 (resp. D.5), w is nonlinear. We therefore call intervals on which w is defined by Steps I.4 and D.4 (resp. I.5 and D.5) *linear intervals* (resp. *nonlinear intervals*).

Remarks 3.2.

1. In the initialization step, under the optimal strategy of Player 1, the belief jumps at random times from $\widetilde{p}_0$ to p_0 and back. When p^* is an extreme point of this interval, say $p^* = \widetilde{p}_0$, substituting $p = p^*$ in Eq. (18) yields $w(p^*) = u(p^*)$. Consequently, in this case at the belief p^* there is no revelation of information.

2. On the interval $[\widetilde{p}_0, p_0]$ the function w coincides with the value function v. Indeed, by Lemma 2 in Cardaliaguet et al. (2016), for every $p \in [\widetilde{p}_0, p_0]$ we have

$$v(p) = \int_0^\infty e^{-rt}(\text{cav } u)(p^* + (p - p^*)e^{-(\lambda_1+\lambda_2)t})dt,$$

with $(\text{cav } u)(p) = u(\widetilde{p}_0) + \frac{u(p_0)-u(\widetilde{p}_0)}{p_0-\widetilde{p}_0}(p - \widetilde{p}_0)$. This integral can be calculated explicitly and it coincides with the expression of w in Eq. (18).
3. In general Eq. (19) does not have an explicit solution. In the special case that $p^* = 0$ and $\mu = 1$, this equation has an explicit solution, given by

$$\varphi_k(p) = \frac{p_k}{p}\varphi(p_k) + \frac{1}{p}\int_{p_k}^{p} u(t)dt.$$

4. Calculating the limit of the term on the right-hand side of Eq. (16) as p' converges to p, we deduce that for every $x \in \mathbb{R}$ we have $a(p, x) \geq \mu \cdot \frac{u(p)-x}{p-p^*}$, provided $p \neq p^*$. In particular, substituting $x = \varphi_k(p)$, the solution of Eq. (19), this gives

$$a(p, \varphi_k(p)) \geq \varphi_k'(p).$$

On a nonlinear interval $(p_k, p_{k+1}]$ we can be even more precise: for every p such that $\rho(p, \varphi_k(p)) = p$, it follows from the definition of $a(p, \varphi_k(p))$ that

$$a(p, \varphi_k(p)) = \mu \cdot \frac{u(p) - \varphi_k(p)}{p - p^*} = \varphi_k'(p). \tag{26}$$

In particular, given that $p_{k+1} = \inf\{p > p_k : \rho(p, \varphi_k(p)) > p\}$, Eq. (26) holds for every $p \in (p_k, p_{k+1})$ as well as for the right (resp. left) derivative of φ_k for $p = p_k$ (resp. $p = p_{k+1}$).

We now state the main theorem of the paper.

Theorem 3.3.

1. *For every $k \geq 0$ such that $p_k < 1$ we have $p_k < p_{k+1}$.*
2. *For every $k \geq 0$ such that $\widetilde{p}_k > 0$ we have $\widetilde{p}_{k+1} < \widetilde{p}_k$.*
3. *The algorithm terminates after a finite number of iterations; that is, there is $k \geq 0$ such that $p_k = 1$ and there is $k \geq 0$ such that $\widetilde{p}_k = 0$.*
4. *The function w generated by the algorithm is the limit value function of the game, i.e., $w = v$.*

The proof of Theorem 3.3 is relegated to Section 6, after the algorithm is demonstrated on some examples.

3.3. On the optimal use of information

Theorem 3.3 states that the algorithm computes the value function v. It also allows to construct an ϵ-optimal strategy for the informed player in the game $G^{(n)}(p)$, provided n is sufficiently large. Indeed, as exhibited by Cardaliaguet et al. (2016), an ϵ-optimal strategy for the informed player in the game $G^{(n)}(p)$ can be constructed using the process $(\pi_t)_{t\geq 0}$ that attains the maximum in Eq. (3). According to this ϵ-optimal strategy, at each stage l of the discrete-time game, the informed player plays an optimal strategy in the two-player zero-sum strategic-form game with payoff function $g(\pi_{l/n}, \cdot, \cdot)$. Thus, $\pi_{l/n}$ serves as a fictitious belief of the uninformed player at stage l. Since the gap between stages is small, the uninformed player can approximate the belief process from the realized actions of the informed player. In this sense, the informed player reveals information about the state of nature along the play. We now argue that the algorithm provides this process.

- As we have seen, the informed player plays in such a way that the belief never lies in the interior of a linear interval. Consequently, the process at time $t = 0$ is defined as follows.
  Let $p \in \Delta(S)$ be the initial distribution. If p is in the interior of a linear interval (p_k, p_{k+1}) for some $k \geq 0$, then the process $(\pi_t)_{t\geq 0}$ has a jump at time $t = 0$: π_0 is equal to p_k or p_{k+1}, where the probability to attain each value is determined so as $\mathbb{E}[\pi_0] = p$, that is, $p_k P[\pi_0 = p_k] + p_{k+1} P[\pi_0 = p_{k+1}] = p$. The analogous statement holds if p lies in a linear interval $(\widetilde{p}_{k+1}, \widetilde{p}_k)$ for some $k \geq 0$, and if $p \in (\widetilde{p}_0, p_0)$. Otherwise we set $\pi_0 = p$.
- When the belief is in a nonlinear interval, the informed player reveals no information. Consequently, in nonlinear intervals the process $(\pi_t)_{t\geq 0}$ is defined as follows.

Let $t_0 \geq 0$, and suppose that π_{t_0} lies in a nonlinear interval $(p_k, p_{k+1}]$ for some $k \geq 1$. The process $(\pi_t)_{t\geq 0}$ evolves continuously: $\pi_{t_0+h} = p^* + (\pi_{t_0} - p^*)e^{-(\lambda_1+\lambda_2)h}$ for every $h > 0$ such that $\pi_{t_0+h} \in [p_k, p_{k+1})$.
If π_{t_0} lies in a nonlinear interval $[\widetilde{p}_{k+1}, \widetilde{p}_k)$ for some $k \geq 1$, the process $(\pi_t)_{t\geq 0}$ is defined analogously.

- When the belief is the upper end of a linear interval, the belief remains p_{k+1} for some time, until it changes to p_k; that is, the informed player reveals information at a random time.
Let $t_0 \geq 0$, and suppose that $\pi_{t_0} = p_{k+1}$, where $(p_k, p_{k+1}]$ is a linear interval. For $t > t_0$ the value of π_t remains p_{k+1}, until it jumps to p_k at the rate $\alpha_k := (\lambda_1 + \lambda_2)\frac{p_{k+1}-p^*}{p_{k+1}-p_k}$. Indeed, the fact that $(\pi_{t-t_0})_{t\geq 0}$ belongs to $\mathcal{S}(p_{k+1})$ implies that $E[\pi_t|\pi_{t_0} = p_{k+1}] = p^* + (p_{k+1} - p^*)e^{-(\lambda_1+\lambda_2)(t-t_0)}$. Writing on the other hand
$$E[\pi_t \mid \pi_{t_0} = p_{k+1}] = p_{k+1} P[T > t \mid \pi_{t_0} = p_{k+1}] + \int_{t_0}^t (p^* + (p_k - p^*)e^{-(\lambda_1+\lambda_2)(t-s)}) P[T \in ds \mid \pi_{t_0} = p_{k+1}],$$
with $T := \inf\{s \geq t_0, \pi_s \neq p_{k+1}\}$, it follows by elementary computations that $P[T > t \mid \pi_{t_0} = p_{k+1}] = e^{-(\lambda_1+\lambda_2)\frac{p_{k+1}-p^*}{p_{k+1}-p_k}(t-t_0)}$.
If $\pi_{t_0} = \widetilde{p}_{k+1}$, where $[\widetilde{p}_{k+1}, \widetilde{p}_k)$ is a linear interval, then the process $(\pi_t)_{t\geq 0}$ is defined analogously for $t > t_0$.
- When the belief is p_0 (resp. $\widetilde{p}_0$), with $p_0 \neq \widetilde{p}_0$, the uninformed player makes the belief jump to $\widetilde{p}_0$ (resp. p_0) at a random time.
Let $t_0 \geq 0$, and suppose that $\pi_{t_0} = p_0$. As in the previous point, for $t > t_0$ the value of π_t remains p_0, until it switches to $\widetilde{p}_0$ with jump rate equal to $(\lambda_1 + \lambda_2)\frac{p_0-p^*}{p_0-\widetilde{p}_0}$. If $\pi_{t_0} = \widetilde{p}_0$, the belief process jumps back to p_0 with jump rate $(\lambda_1 + \lambda_2)\frac{p^*-\widetilde{p}_0}{p_0-\widetilde{p}_0}$.

It follows that the set of possible fictitious beliefs of the uninformed player is divided into finitely many disjoint regions; some regions should be avoided, and the informed player reveals information in such a way that the fictitious belief does not lie in these regions, while in the remaining regions the informed player reveals no information. It turns out that the types of the regions are alternating: each region that should be avoided lies between two nonrevealing regions, and vice versa. This phenomenon already occurred in the setup of Aumann and Maschler (1995), where the state remains constant throughout the play. Yet while in the setup of Aumann and Maschler (1995) the nonrevealing regions are singletons, in our setup the regions are always intervals with non-empty interiors (except possibly $\{p^*\}$).

4. Examples

In this section we illustrate the algorithm on the three examples provided in the Introduction. Recall that in these examples $\mu = r = 1$ and the state s_2 is absorbing, so that $p^* = 0$. We will also analyze two variants of the third example, where state s_2 is not absorbing; the first is a nontrivial example where the algorithm has both an increasing and a decreasing part, and the second will show that the limit value function may be nondifferentiable at p^*.

Example 1, continued. In this example the function u is given by $u(p) = p(1-p)$ for every $p \in [0, 1]$. The function u is concave, and therefore $\widetilde{p}_0 = p_0 = 0$, and $w(0) = 0$. We next compute the solution of Eq. (19) with initial condition $\varphi(0) = 0$. For $p \in [0, 1]$, the solution is (see Remark 3.2.3)
$$\varphi(p) = \frac{0}{p} + \frac{1}{p}\int_0^p t(1-t)dt = \frac{p}{2} - \frac{p^2}{3}.$$
It follows that
$$a(p, \varphi(p)) = \sup_{p' \in (p,1]} \frac{p'(1-p') - \frac{p}{2} + \frac{p^2}{3}}{2p' - p}.$$
For every $p \in [0, 1]$ the supremum is obtained only at $p' = p$, that is $\rho(p, \varphi) = p$ for every $p \in [0, 1]$. This implies that the condition of Step I.5 holds, $p_1 = 1$, and the first part of the algorithm terminates. Since $\widetilde{p}_0 = 0$, the second part of the algorithm is vacuous. In conclusion, the limit value function is given by
$$v(p) = \frac{p}{2} - \frac{p^2}{3}, \ \forall p \in [0, 1],$$
and the optimal strategy of Player 1 is never to reveal her information.

Example 2, continued. Recall that in this example the function u is given by $u(p) = -p(1-p)$ for every $p \in [0, 1]$. Since $(\text{cav } u)(p^*) = 0 = \alpha u(0) + (1-\alpha)u(1), \ \forall \alpha \in [0, 1]$, we have $\widetilde{p}_0 = 0$ and $p_0 = 1$. From Eq. (18) we obtain that $v(p) = 0$ for every $p \in [0, 1]$. Consequently, the optimal strategy of Player 1 is to always reveal her information.

Example 3, continued. In this example the function $u(p)$ is given by Eq. (1) and is represented by the dotted line in Fig. 7. For this example the algorithm runs as follows. Since $p^* = 0$ we have $\widetilde{p}_0 = 0$. Simple calculations show that $p_0 = \frac{1}{3}$, and from Eq. (18) we have $w(p) = -\frac{2}{3} + p$ for every $p \in [0, \frac{1}{3}]$. On $[\frac{1}{3}, \frac{2}{3}]$ the solution of Eq. (19) is $\varphi_0(p) = \frac{1}{3p} v(\frac{1}{3}) + \frac{1}{p}\int_{\frac{1}{3}}^{p} 0\, dx = -\frac{1}{9p}$. It follows that, for $p \in [\frac{1}{3}, \frac{2}{3}]$,

$$\begin{aligned} a(p,\varphi_0(p)) =& \ \sup_{p' \in (p,1]} \frac{u(p') - \varphi_0(p)}{2p' - p} \\ =& \ \max\left\{ \sup_{p' \in (p,\frac{2}{3})} \frac{-\varphi_0(p)}{2p'-p}; \sup_{p' \in [\frac{2}{3},1]} \frac{\frac{9p'^2 - 9p' + 2}{3(2p'-1)} + \frac{1}{9p}}{2p'-p} \right\} \\ =& \ \max\left\{ \frac{1}{9p^2}, \frac{6p+1}{9p(2-p)} \right\}, \end{aligned}$$

where the suprema are respectively attained at p and 1. Solving $\frac{1}{9p^2} = \frac{6p+1}{9p(2-p)}$, we obtain

$$a(p,\varphi_0(p)) = \begin{cases} \frac{1}{9p^2} & \text{with } \rho(p,\varphi_0(p)) = p, \text{ for } p < \overline{p} := \frac{-1+\sqrt{13}}{6}, \\ \frac{6p+1}{9p(2-p)} & \text{with } \rho(p,\varphi_0(p)) = 1, \text{ for } p \geq \overline{p}. \end{cases}$$

Therefore (using Step I.5) $p_1 = \inf\{p > p_0, \rho(p,\varphi_0(p)) > p\} = \overline{p}$ and $w(p) = \varphi_0(p) = -\frac{1}{9p^2}$ for $p \in [\frac{1}{3}, \overline{p}]$. Finally (Step I.4) $p_2 = \rho(\overline{p}, w(\overline{p})) = 1$ and $w(p) = -\frac{1}{9p} + \frac{1}{9p^2}(p - \overline{p})$ on $[\overline{p}, 1]$, and the algorithm terminates. In conclusion, the limit value function is given by

$$v(p) = \begin{cases} p - \frac{2}{3}, & \text{if } 0 \leq p < \frac{1}{3}, \\ -\frac{1}{9p}, & \text{if } \frac{1}{3} \leq p < \overline{p}, \text{ with } \overline{p} = \frac{\sqrt{13}-1}{6} (\simeq 0,434), \\ -\frac{1}{9\overline{p}} + \frac{1}{9\overline{p}^2}(p - \overline{p}), & \text{if } \overline{p} \leq p \leq 1. \end{cases} \tag{27}$$

In particular

$$v(1) = \frac{1}{9\overline{p}^2}(2\overline{p} - 1).$$

The evolution of the optimal belief process in this example is as follows. Suppose that the initial belief $\pi_0 = 1$. The belief process jumps to $\overline{p}$ with jump rate equal to $\frac{1}{1-\overline{p}}$, and then slides towards $\frac{1}{3}$ with speed $\pi'_t = -\overline{p}e^{-(t-T)}$, where T is the jump time. Once it arrives to $\frac{1}{3}$, it jumps to 0 with jump rate 1, where it remains forever. If $\pi_0 \in (\overline{p}, 1)$ or $\pi_0 \in (0, \frac{1}{3})$, the process first splits to the extreme points of the interval, and then continues as described above. As described in Section 3.3, the optimal belief process determines one ϵ-optimal strategy for the informed player when the gap between stages is small.

Example 3, variation a.

Here we assume that $\lambda_1 = 3$ and $\lambda_2 = r = 1$, and therefore state s_2 is no longer absorbing. It follows that $\mu = \frac{1}{4}$ and $p^* = \frac{1}{4} \neq 0$. The algorithm starts at $p^* = \frac{1}{4}$ and runs in both directions. Here we write down the main steps of the algorithms; the detailed calculations appear in Appendix 7.1. The function u is convex on $[0, \frac{1}{3}]$ (see Fig. 8), one can show that $\widetilde{p}_0 = 0$ and $p_0 = \frac{1}{3}$, and therefore the decreasing part of the algorithm is vacuous. For the increasing part of the algorithm, Eqs. (19) and (21) for $k = 0$ can be solved explicitly and give $p_1 \simeq 0.3858$. It is then not difficult to show that $p_2 = 1$. It turns out that the limit value function is

$$v(p) = \begin{cases} \frac{2}{15}(3p-2), & \text{if } 0 \leq p < \frac{1}{3}, \\ -\frac{2}{15} \cdot 3^{-1/4} \cdot (4p-1)^{-1/4}, & \text{if } \frac{1}{3} \leq p < p_1, \\ ap + b & \text{if } p \in (p_1, 1], \end{cases} \tag{28}$$

where $a = \frac{\frac{2}{3} - w(p_1)}{4 - p_1} \simeq 0.21709$ and $b = \frac{2}{3} - 4a \simeq -0.20177$ (see Fig. 8). The optimal revelation process has the following structure. If $\pi_0 = 0$, the process jumps to $p^* = \frac{1}{3}$ with jump rate $\lambda_1 + \lambda_2 = 4$, where it remains forever. If $\pi_0 = 1$, the process jumps to p_1 with jump rate $(\lambda_1 + \lambda_2)\frac{1-p^*}{1-p_1} \simeq 4.342$, then slides towards $\frac{1}{3}$ with speed $\pi'(t) = -4(p_1 - p^*)e^{-4(t-T)} \simeq -0.21e^{-4(t-T)}$, where T is the jump time. For any other starting point, the optimal revelation process can be deduced from these two cases, by adding, if necessary, an initial splitting.

Example 3, variation b.

Suppose now that $\lambda_1 = \frac{4}{3}$, $\lambda_2 = \frac{2}{3}$, which implies that $r = 1$, $p^* = \frac{1}{3}$, and $\mu = \frac{1}{2}$. In this case $(\text{cav } u)(\frac{1}{3}) = u(\frac{1}{3}) = 0$ and $\widetilde{p}_0 = p_0 = \frac{1}{3}$, and in both directions the algorithm starts with an affine part and lasts only one step: $\widetilde{p}_1 = 0$ and $p_1 = 1$. The limit value function turns out to be

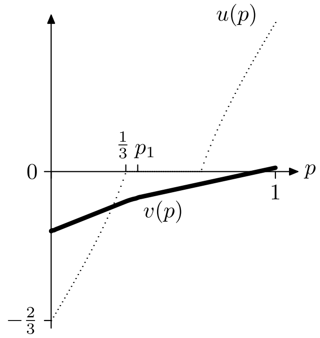

Fig. 8. The value function in Example 3, variation a.

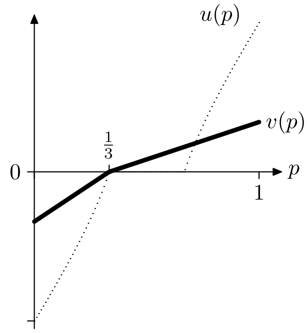

Fig. 9. The value function in Example 3, variation b.

$$v(p)=\begin{cases}\frac{2}{3}p-\frac{2}{9}, & \text{if } 0\le p<\frac{1}{3},\\ \frac{1}{3}p-\frac{1}{9}, & \text{if } \frac{1}{3}\le p<\bar{1}.\end{cases} \tag{29}$$

(See Fig. 9.) Note that v is not differentiable at $p^*=\frac{1}{3}$. For the precise computations, see Appendix 7.2. Here, the set of beliefs is divided into two linear intervals. Consequently, at time 0 the belief is split between 0 and $\frac{1}{3}$ (if the initial belief is smaller than $\frac{1}{3}$) or between $\frac{1}{3}$ and 1 (if the initial belief is larger than $\frac{1}{3}$). Then the optimal revelation process jumps to $p^*=\frac{1}{3}$ with jump rate equal to 2, where it remains forever.

5. Perspectives and open questions

An interesting question concerns the dependence of the value function and the optimal strategy of the informed player on the transition rates λ_1 and λ_2 and on the discount rate r. The invariant measure $p^*=\frac{\lambda_2}{\lambda_1+\lambda_2}$ determines the interval $[\widetilde{p}_0,p_0]$ at the initialization of the algorithm, as well as the rate of information revelation by the informed player when π_t is p_0, $\widetilde{p}_0$, or the extreme point of a linear interval that is farther away from the invariant distribution. The ratio $\mu=\frac{r}{\lambda_1+\lambda_2}$ affects the slope of the nonlinear part (see Eq. (19)) and therefore also the slope of the linear parts of the value function. The larger μ, the more extreme is the slope (larger when the slope is positive, smaller when the slope is negative). Example 3 and its variants show that for different parameters, the value function may have a completely different structure.

The algorithm we presented computes the value function of a zero-sum Markov game, yet it applies to all games where the optimization problem of the informed player can be expressed as an optimization problem with respect to a belief process

$$v(p)=\max_{(p_t)\in\mathcal{S}(p)} E\left[\int_0^\infty re^{-rt}u(p_t)dt\right], \tag{30}$$

whenever the function u is Lipschitz, piecewise twice differentiable, and its second derivative changes signs finitely many times. These last two properties hold when u is semialgebraic, and in particular in repeated games with incomplete information. They also hold for o-minimal structures in general (see, e.g., Coste (1999)). For a study of games with payoffs and transitions that are in some o-minimal structure, see Bolte et al. (2015).

One setup where the optimization problem can be expressed as in Eq. (30) is the model of dynamic information provision, see, e.g., Ely (2017) and Renault et al. (2017). In this setup, the payoff is not zero sum and it depends on the state and on the action of the uninformed player, but not on the action of the informed player. Consequently, the uninformed player is myopic, her action at each stage depends solely on her belief, and it turns out that the limit value function is of type (30), where u is an indicator function, hence discontinuous. By approximating the discontinuous function u with a continuous semialgebraic function $\widehat{u}$, and applying our algorithm to $\widehat{u}$, one can approximate the limit value function of the original problem. This approach has been taken by Ashkenazi-Golan et al. (2020) who study dynamic information provision with two states of nature when the function u is a monotone step function, and provide insights regarding the structure of the optimal strategy of the informed player.

The algorithm is tailored for repeated games with incomplete information and two states. A natural question is whether the algorithm can be generalized to games with more than two states. Unfortunately, the existing literature exhibits the inherent complexity in Markov games with three states. For example, in the model of dynamic information provision, while the two-state case can be easily solved, Renault et al. (2017) provided a well chosen example showing that even in the case where the transition matrix is a homothety, the optimal strategy for the informed player may be complex. A similar phenomenon occurs in related models. In a work in progress, Gensbittel and Rainer (2020) study the continuous time counterpart of Renault et al. (2017) and show that in general, there is no much hope for an explicit construction if the number of states exceeds two. Gensbittel and Grün (2019) studied optimal stopping with asymmetric information in continuous time, and provided an algorithm to compute the optimal strategy of the informed player only in the two-state case. In the case where the cost function is time inhomogeneous but the information does not evolve, that is, the Markov process is constant, Cardaliaguet and Rainer (2009b) provided a three-state example admitting an optimal revelation process which is a continuous martingale, i.e., a stochastic integral with respect to a Brownian motion, living on a hyperspace described by a mean curvature motion. In this example, there exists an optimal revealing process with a simpler structure as well, yet the question if there always exists an optimal revealing process without Brownian part is still open. Another natural extension concerns signals as in Gensbittel (2019), who studied Markov games in continuous time where the uninformed players observe a diffusion depending on the Markov state process. In this model, the value function depends on both the state and the diffusion process. This diffusion process adds a supplementary dimension to the value function and its characterization. Therefore, it is not clear how to adapt the algorithm to this case.

The algorithm provides an optimal strategy for the informed player. An interesting open question is the determination of the optimal strategy for the uninformed player.

A strategy is *uniform ϵ-optimal* if it is ϵ-optimal in all discounted games, provided the discount rate is sufficiently close to 1. Renault (2012) (resp. Gensbittel and Renault (2015)) proved that in zero-sum Markov games with incomplete information on one side (resp. on both sides) both players have uniform ϵ-optimal strategies, for every $\epsilon > 0$. By definition, if uniform ϵ-optimal strategies exist for every $\epsilon > 0$, then the discounted value function converges to a limit when the discount rate goes to 0. To date it is not known whether in continuous-time repeated games with asymmetric information, the discounted value function converges to a limit as the discount rate goes to 0, and in particular, whether uniform ϵ-optimal strategies exist. We hope that our research will help in the solution of this problem.

6. Proof of Theorem 3.3

This section is devoted to the proof of Theorem 3.3. In Section 6.1 we study the sequence (p_k) and show that the algorithm provided in Section 3.2 terminates. In Section 6.2 we show that the function w is concave and differentiable everywhere, except, possibly, at p^*. In Section 6.3 we show that $w = v$.

6.1. On the sequence (p_k)

In this section we study the sequence (p_k). We will show that it is strictly increasing (Lemma 6.2) and that if p_k is defined by Eq. (21) then p_{k+1} is defined by Eq. (23), and vice versa (Lemma 6.3). We will then show that there is $k \in \mathbb{N}$ such that $p_k = 1$. We start with a technical lemma that will determine the value of u and φ on the elements of the sequence (p_k).

Lemma 6.1. *Let $q \in (p^*, 1)$. Suppose that w is defined at q and set $\rho := \rho(q, w(q))$. Suppose that u is twice differentiable on some open interval I that contains ρ.*

1. *If $q < \rho$, then*
   (i) $u'(\rho) = \frac{1+\mu}{\mu} a(q, w(q))$, and
   (ii) $u''(\rho) \leq 0$.
2. *Let $\varphi : [q, 1] \to \mathbb{R}$ be a function satisfying $\varphi(q) = w(q)$ and $\varphi'(p)(p - p^*) = \mu\big(u(p) - \varphi(p)\big)$ on $[q, 1]$. If $q = \rho$, then $\varphi''(q) \leq 0$.*

Proof. We start with the first claim. Set $\Delta(p) := p - p^* + \mu(p - q)$ and $F(p) := \mu \cdot \frac{u(p)-w(q)}{\Delta(p)}$ for every $p \in I$. Since u is differentiable on I, the function F is also differentiable on I, and its derivative is

$$F'(p) = \mu \cdot \frac{u'(p)\Delta(p) - (1+\mu)(u(p) - w(q))}{\Delta^2(p)}. \tag{31}$$

If $q < \rho$, then ρ is a local extremum in I of F, and we have $F'(\rho) = 0$. From Eq. (31) we obtain

$$u'(\rho) = (1+\mu)\frac{u(\rho) - w(q)}{\Delta(\rho)} = (1+\mu)\frac{u(\rho) - w(q)}{\rho - p^* + \mu(\rho - q)}. \tag{32}$$

Item *(i)* follows by the definition of $a(q, w)$.
The second derivative of F at ρ is

$$F''(\rho) = \mu \cdot \frac{u''(\rho)\Delta^2(\rho) - 2(1+\mu)\left(u'(\rho)\Delta(\rho) - (1+\mu)(u(\rho) - w(q))\right)}{\Delta^3(\rho)}. \tag{33}$$

From Eq. (32) the second term in the numerator in Eq. (33) vanishes, hence

$$u''(\rho) = \tfrac{1}{\mu} F''(\rho)\Delta(\rho).$$

Since ρ is a local maximum of F, we have $F''(\rho) \le 0$. Since $\Delta(\rho) > 0$, Item *(ii)* follows.

We turn to the second claim. If $\rho = q$, the maximum of F on $[q, 1]$ is attained at q. Therefore $F'(q) \le 0$. It follows from Eq. (31) that

$$u'(q) \le (1+\mu)\frac{u(q) - \varphi(q)}{q - p^*} = \frac{1+\mu}{\mu}\varphi'(q). \tag{34}$$

Further, from the relation $\varphi'(q)(q - p^*) = \mu(u(q) - \varphi(q))$, we get

$$\varphi''(q)(q - p^*) = \mu u'(q) - (1+\mu)\varphi'(q). \tag{35}$$

Eqs. (34) and (35) imply that $\varphi''(q) \le 0$. □

Lemma 6.2. *For all $k \ge 0$ such that $p^* \le p_k < 1$, we have $p_k < p_{k+1}$.*

Proof. By Step I.4, if $\rho(p_k, w(p_k)) > p_k$, then $(p_k, p_{k+1}]$ is a linear interval and $p_{k+1} = \rho(p_k, w(p_k)) > p_k$, as claimed. If $\rho(p_k, w(p_k)) = p_k$, then by Step I.5, $(p_k, p_{k+1}]$ is a nonlinear interval and

$$p_{k+1} = \inf\{p > p_k, \rho(p, \varphi(p)) > p\}, \tag{36}$$

where φ is the solution of Eq. (19). In this case the result is not trivial. We shall prove it by contradiction.

Suppose to the contrary that $p_{k+1} = p_k$. Then Eq. (36) implies the existence of a sequence $(q^n)_{n\in\mathbb{N}} \subset (p_k, p_k + \epsilon)$ such that $q^n \searrow p_k$ and $\rho(q^n, \varphi(q^n)) > q^n$ for every $n \in \mathbb{N}$. In what follows, we set $\rho^n := \rho(q^n, \varphi(q^n))$.

Let $\overline{\rho}$ be an accumulation point of the sequence $(\rho^n)_{n\in\mathbb{N}}$ and denote still by $(q^n)_{n\in\mathbb{N}}$ a subsequence of $(q^n)_{n\in\mathbb{N}}$ such that $\rho(q^n, \varphi(q^n))$ converges to $\overline{\rho}$. Since $p \mapsto a(p, \varphi(p))$ is continuous,[2] and, by Eq. (16)

$$a(q^n, \varphi(q^n)) = \mu \cdot \frac{u(\rho^n) - \varphi(q^n)}{\rho^n - p^* + \mu(\rho^n - q^n)}, \tag{37}$$

letting n tend to ∞ in Eq. (37) we get

$$a(p_k, \varphi(p_k)) = \mu \cdot \frac{u(\overline{\rho}) - \varphi(p_k)}{\overline{\rho} - p^* + \mu(\overline{\rho} - p_k)}.$$

By assumption, the value $a(p_k, \varphi(p_k))$ is attained only at $p_k = \rho(p_k, \varphi(p_k))$. Thus $\overline{\rho} = p_k$. By taking a subsequence of $(q^n)_{n\in\mathbb{N}}$, still denoted $(q^n)_{n\in\mathbb{N}}$, we can assume that $\rho^{n+1} < q^n < \rho^n$ for every $n \in \mathbb{N}$.

Since the function u is semialgebraic, there exists n_0 such that u is smooth on the interval (p_k, q^{n_0}). Moreover, by Lemma 6.1(1) we have $u''(\rho^n) \le 0$ for every $n \ge n_0$. This implies that u' is nonincreasing and u is concave on (p_k, q^{n_0}). We can strengthen this conclusion: we can choose n_0 such that u' is strictly decreasing and u is strictly concave on (p_k, q^{n_0}). Indeed, suppose that this does not hold. In this case u is linear in a small one-sided neighborhood of p_k: there exist $\widetilde{\epsilon} \le \epsilon$ and $\alpha, \beta \in \mathbb{R}$ such that $u(p) = \alpha p + \beta$ for all $p \in [p_k, p_k + \widetilde{\epsilon}]$. By Lemma 6.1(1), it follows that $\frac{\mu}{1+\mu}\alpha = a(q^n, \varphi(q^n)) = a(p_k, \varphi(p^k))$ for n sufficiently large. By Eq. (37) we therefore have

$$a(p_k, \varphi(p_k)) = \mu \cdot \frac{\frac{1+\mu}{\mu} a(p_k, \varphi(p_k))\rho^n + \beta - \varphi(q^n)}{\rho^n - p^* + \mu(\rho^n - q^n)},$$

or, equivalently, $\beta + \frac{a(p_k, \varphi(p_k))}{\mu}(p^* + \mu q^n) = \varphi(q^n)$. In addition, for every $p > \rho^n$ we have by the definition of ρ^n,

[2] as the maximum of a set of uniformly Lipschitz continuous functions.

$$a(p_k,\varphi(p_k)) = a(q^n,\varphi(p_k)) > \mu \cdot \frac{u(p)-\varphi(q^n)}{p-p^*+\mu(p-q^n)} = \mu \cdot \frac{\frac{1+\mu}{\mu}a(p_k,\varphi(p_k))p+\beta-\varphi(q^n)}{p-p^*+\mu(p-q^n)},$$

or, equivalently, $\beta + \frac{a(p_k,\varphi(p_k))}{\mu}(p^*+\mu q^n) < \varphi(q^n)$, a contradiction. It follows that u' is strictly decreasing in a small one-sided neighborhood of p_k.

Now fix $n > n_0$ and let $\widetilde{\rho} \in (q^n, \rho^n)$. Since u' is strictly decreasing on (p_k, q^{n_0}) and $\rho^{n+1} < q^n$, we have $u'(\rho^n) < u'(\widetilde{\rho}) < u'(\rho^{n+1})$, or, equivalently, using Lemma 6.1(1),

$$\frac{\mu}{1+\mu}u'(\widetilde{\rho}) \in (a(q^n,\varphi(q^n)), a(q^{n+1},\varphi(q^{n+1}))).$$

By the continuity of the function $p \mapsto a(p,\varphi(p))$, there exists $q' \in (q^{n+1}, q^n)$ such that $a(q',\varphi(q')) = \frac{\mu}{1+\mu}u'(\widetilde{\rho})$. It follows that, if ρ' is close enough to ρ^n, it belongs to the set of points where the maximum $a(q',\varphi(q'))$ is attained. Therefore it holds that $\rho' := \rho(q',\varphi(q')) \geq \widetilde{\rho} \geq q^n$. Moreover, since $q' > q^n$, we have

$$a(q^n,\varphi(q^n)) < a(q',\varphi(q')). \tag{38}$$

Consider now the function Ψ on $[q',1]$ defined by $\Psi(p) := \varphi(q') + a(q',\varphi(q'))(p-q')$. The reader can verify that the function Ψ is a solution of

$$\begin{aligned} &\Psi(q') = \varphi(q'),\\ &\Psi'(p)(p-p^*) = \mu(\overline{u}(p)-\Psi(p)), \quad \forall p \in [q',1], \end{aligned}$$

with $\overline{u}(p) := \varphi(q') + \frac{a(q',\varphi(q'))}{\mu}\big(p-p^*+\mu(p-q')\big)$ for every $p \in [q',1]$. It follows that the function $\gamma : [q',1] \to \mathbb{R}$ defined by $\gamma(p) := \Psi(p) - \varphi(p)$ is a solution of

$$\begin{aligned} &\gamma(q') = 0,\\ &\gamma'(p)(p-p^*) = \mu(\widetilde{u}(p)-\gamma(p)), \quad \forall p \in (q',1), \end{aligned} \tag{39}$$

with $\widetilde{u}(p) = \overline{u}(p) - u(p)$, for every $p \in [q',1]$. Eq. (39) can be solved quasi-explicitly:

$$\gamma(p) = c(p)(p-p^*)^{-\mu}, \quad \forall p \in [q',1],$$

with $c(q') = 0$ and $c'(p) = \mu\widetilde{u}(p)(p-p^*)^{\mu-1}$.

By the definition of $a(q',\varphi(q'))$ and $\overline{u}$, the function $\widetilde{u}$ is nonnegative on $[q',1]$. It follows that, for every $p \in [q',1]$ we have $c(p) \geq 0$ and consequently $\gamma(p) \geq 0$, which is equivalent to $\Psi(p) \geq \varphi(p)$. Substituting $p = q^n$, we obtain in particular that $\Psi(q^n) \geq \varphi(q^n)$, or, equivalently,

$$\varphi(q^n) - \varphi(q') \leq a(q',\varphi(q'))(q^n - q'). \tag{40}$$

To derive a contradiction, recall that, by the definition of $a(q',\varphi(q'))$ and $a(q^n,\varphi(q^n))$,

$$u(\rho') = \varphi(q') + \frac{a(q',\varphi(q'))}{\mu}(\rho'-p^*+\mu(\rho'-q'))$$

and

$$u(\rho') \leq \varphi(q^n) + \frac{a(q^n,\varphi(q^n))}{\mu}(\rho'-p^*+\mu(\rho'-q^n)).$$

Combining these two equations with Eq. (40) we obtain

$$a(q',\varphi(q'))(\rho'-p^*+\mu(\rho'-q^n)) \leq a(q^n,\varphi(q^n))(\rho'-p^*+\mu(\rho'-q^n)).$$

Since $\rho' > p^*$ and $\rho' > q^n$ this implies that $a(q',\varphi(q')) \leq a(q^n,\varphi(q^n))$, contradicting Eq. (38). It follows that p_{k+1} that is defined by Eq. (36) satisfies $p_{k+1} > p_k$. $\square$

The following lemma says that linear intervals are followed by nonlinear intervals and vice versa.

Lemma 6.3.

1. *If $p^* < p_0 < 1$, then $\rho(p_0, w) = p_0$ and $(p_0, p_1]$ is a nonlinear interval. (If $p_0 = p^*$, then $(p_0, p_1]$ may be a linear or a nonlinear interval.)*
2. *For every $k \geq 1$ such that $p_k > 1$, if $(p_{k-1}, p_k]$ is a linear interval (resp. a nonlinear interval), then $(p_k, p_{k+1}]$ is a nonlinear interval (resp. a linear interval).*

Proof. By the definition of $\widetilde{p}_0$ and p_0, we have

$$\frac{u(p)-u(p_0)}{p-p_0} < \frac{u(p_0)-u(\widetilde{p}_0)}{p_0-\widetilde{p}_0}, \quad \forall p \in (p_0, 1].$$

Simple (though tedious) algebraic manipulations combining this inequality with Eq. (18) for $p = p_0$ yield

$$\frac{u(p)-w(p_0)}{p-p^*+\mu(p-p_0)} < \frac{u(p_0)-w(p_0)}{p_0-p^*}, \quad \forall p \in (p_0, 1].$$

Claim 1 follows.

We turn to prove Claim 2. Suppose that $(p_{k-1}, p_k]$ is a linear interval. By construction we have

$$a(p_{k-1}, w(p_{k-1})) = \mu \cdot \frac{u(p_k)-w(p_{k-1})}{p_k-p^*+\mu(p_k-p_{k-1})}. \tag{41}$$

Since on the interval $(p_{k-1}, p_k]$ the function w is defined by Eqs. (22) and (20), we have $w(p_{k-1}) = w(p_k) - a(p_{k-1}, w(p_{k-1}))(p_k - p_{k-1})$, and Eq. (41) becomes

$$a(p_{k-1}, w(p_{k-1})) = \mu \cdot \frac{u(p_k)-w(p_k)}{p_k-p^*}. \tag{42}$$

To show that $(p_k, p_{k+1}]$ is a nonlinear interval we will show that $\rho_k := \rho(p_k, w(p_k)) = p_k$. By Eq. (42) and Remark 3.2.4 we have

$$a(p_{k-1}, w(p_{k-1})) = \mu \cdot \frac{u(p_k)-w(p_k)}{p_k-p^*} \le a(p_k, w(p_k)) = \mu \cdot \frac{u(\rho_k)-w(p_k)}{\rho_k-p^*+\mu(\rho_k-p_k)}.$$

Using again the relation $w(p_k) = w(p_{k-1}) + a(p_{k-1}, w(p_{k-1}))(p_k - p_{k-1})$, this last inequality becomes

$$a(p_{k-1}, w(p_{k-1})) \le \mu \cdot \frac{u(\rho_k)-w(p_{k-1})}{\rho_k-p^*+\mu(\rho_k-p_{k-1})}.$$

Since $\rho(p_{k-1}, w(p_{k-1}))$ is the maximal p' that satisfies $a(p_{k-1}, w(p_{k-1})) = \mu \cdot \frac{u(p')-w(p_{k-1})}{p'-p^*+\mu(p'-p_{k-1})}$, this implies that $\rho(p_k, w(p_k)) = \rho(p_{k-1}, w(p_{k-1})) = p_k$, which is what we wanted to prove. For later use we note that in this case we have $a(p_{k-1}, w(p_{k-1})) = a(p_k, w(p_k))$.

Finally assume that $(p_{k-1}, p_k]$ is a nonlinear interval, so that $p_k = \inf\{p > p_{k-1}, \rho(p, w(p)) > p\}$. To prove that $(p_k, p_{k+1}]$ is a linear interval we will show that $\rho(p_k, w(p_k)) > p_k$. Suppose to the contrary that $\rho(p_k, w(p_k)) = p_k$. Then, the algorithm dictates that $p_{k+1} = \inf\{p > p_k, \rho(p, \varphi_k(p)) > p\}$ and $w = \varphi_k$ on $(p_k, p_{k+1}]$. By the definition of p_k, this implies that $p_{k+1} = p_k$, contradicting Lemma 6.2. We conclude that $(p_k, p_{k+1}]$ is a linear interval. □

Lemma 6.4. *The algorithm ends after a finite number of iterations: there exists $k \ge 0$ such that $p_k = 1$ and there exists $\widetilde{k} \ge 1$ such that $\widetilde{p}_{\widetilde{k}} = 0$.*

Proof. We will prove the first claim. The second claim is proven analogously. Assume by contradiction that $p_k < 1$ for every $k \in \mathbb{N}$, and set $p_\infty = \lim_{n\to\infty} p_n$. By Lemma 6.2, $(p_k, p_{k+1}) \neq \emptyset$ for every $k \in \mathbb{N}$. Since u is semialgebraic, there is n_0 sufficiently large such that u is twice differentiable on $[p_{n_0}, p_\infty)$. Let $k \ge n_0$ be such that the interval $(p_k, p_{k+1}]$ is linear. By Eq. (26) and the definition of $a(p_k, w(p_k))$ (Eq. (16)),

- $a(p_k, w(p_k)) = \mu \cdot \frac{u(p_k)-w(p_k)}{p_k-p^*} = \mu \cdot \frac{u(p_{k+1})-w(p_k)}{p_{k+1}-p^*+\mu(p_{k+1}-p_k)}$,
- $a(p_k, w(p_k)) \ge \mu \cdot \frac{u(p)-w(p_k)}{p-p^*+\mu(p-p_k)}$, for every $p \in (p_k, p_{k+1})$.

Equivalently, if we set

$$f(p) = \mu(u(p) - w(p_k)) - a(p_k, w(p_k)) - a(p_k, w(p_k))(p - p^* + \mu(p - p_k)), \tag{43}$$

it holds that $f(p_k) = f(p_{k+1}) = 0$ and $f(p) \le 0$ for every $p \in (p_k, p_{k+1})$.

By Lemma 6.3 there are infinitely many linear intervals. We argue now that, provided k is sufficiently large, if the interval $(p_k, p_{k+1}]$ is linear then there exists $p \in (p_k, p_{k+1})$ with $f(p) < 0$. Indeed, if this is not true, then for every such k sufficiently large, $f(p) = 0$ for every $p \in (p_k, p_{k+1})$. By Eq. (43) this implies that u is affine on (p_k, p_{k+1}). Since u is semialgebraic, it is affine on the whole interval $[p_{n_1}, p_\infty)$, for some large enough n_1. But in this case, for every $k \ge n_1$, if $\rho(p_k, w) > p_k$ then $\rho(p_k, w) = p_\infty$, contradicting the fact that $p_k < p_\infty$ for every k.

We conclude that for every k sufficiently large such that the interval $(p_k, p_{k+1}]$ is linear there is $p \in (p_k, p_{k+1})$ satisfying $f(p) < 0$. In that case we can also find $p' \in (p_k, p_{k+1})$ such that $f''(p') > 0$. Since $f'' = \mu u''$, this implies that $u''(p_k) > 0$. Since u is semialgebraic, this implies that $u''(p) > 0$ for every p sufficiently close to p_∞. However, by Lemma 6.1 (1.ii), $u''(q) \le 0$ for some $q \in (p_k, p_{k+1})$, a contradiction. □

6.2. The function w is differentiable and concave

Lemma 6.5. *The function w is differentiable on $[0, p^*) \cup (p^*, 1]$. If $\widetilde{p}_0 < p^* < p_0$, then w is differentiable everywhere.*

Proof. By its definition, the function w is affine on $[\widetilde{p}_0, p_0]$. Hence w is differentiable on (p^*, p_0), and if $\widetilde{p}_0 < p^* < p_0$ then w is differentiable at p^*.

We next note that w is differentiable on each interval (p_{k-1}, p_k). Indeed, on each of these intervals, w is affine or the solution of a standard first order differential equation.

We now show that w is differentiable at each of the points $(p_k)_{k\geq 1}$, and, if $p^* < p_0$, it is also differentiable at p_0. Denote by $w'_-(p)$ (resp. $w'_+(p)$) the left (resp. right) derivative of w at p.

If $p^* < p_0$, then w is affine on $[p^*, p_0]$ and by Eq. (18) we have $w'_-(p_0) = \frac{\mu(u(p_0)-u(\widetilde{p}_0))}{(p_0-\widetilde{p}_0)(\mu+1)}$. By definition, $w'_+(p_0) = \varphi'_+(p_0) = \mu \cdot \frac{u(p_0)-w(p_0)}{p_0-p^*}$. Substituting $w(p_0)$ by its expression in Eq. (18), we deduce that $w'_+(p_0) = w'_-(p_0)$.

For $k \geq 1$, either $\rho(p_k, w) > p_k$ or $\rho(p_k, w) = p_k$. In the first case, $(p_k, p_{k+1}]$ is a linear interval, and then $w'_-(p_k) = \mu \cdot \frac{u(p_k)-w(p_k)}{p_k-p^*}$ and $w'_+(p_k) = \mu \cdot \frac{u(p_{k+1})-w(p_k)}{p_{k+1}-p^*+\mu(p_{k+1}-p_k)}$. p_k is defined to be $\inf\{p > p_{k-1} : \rho(p, \varphi_{k-1}) > p\}$. Thus, for all $p \in (p_{k-1}, p_k)$ we have

$$\frac{u(p_{k+1}) - w(p)}{p_{k+1} - p^* + \mu(p_{k+1} - p)} \leq \frac{u(p) - w(p)}{p - p^*}.$$

Since p_k is the infimum of a decreasing sequence where the last inequality is reversed, and by the continuity of w and u, we get the equality of the derivatives.

In the second case the interval $(p_{k-1}, p_k]$ is linear. For such k we have $w'_-(p_k) = a(p_{k-1}, w(p_{k-1}))$ and $w'_+(p_k) = \mu \cdot \frac{u(p_k)-w(p_k)}{p_k-p^*}$, and the equality of the derivatives follows from Eq. (42).

Analogous arguments hold for the interval $[0, p^*)$. □

Lemma 6.6. *The function w is concave.*

Proof. On the interval $[\widetilde{p}_0, p_0]$ and on linear intervals the function w is affine. We turn to prove that w is concave on the nonlinear intervals. We will only discuss nonlinear intervals defined by Step I.5.

On nonlinear intervals the function w coincides with the solution φ of Eq. (19). Moreover, $\rho(q, \varphi) = q$ for every q in such an interval. The function u is semialgebraic, hence twice differentiable on [0,1], except possibly at finitely many points. If q is in a nonlinear interval and u is twice differentiable at q, then u is twice differentiable in an open neighborhood of q, and hence, by Lemma 6.1(2), we have $w''(q) = \varphi''(q) \leq 0$.

It follows that the interval $[0, 1]$ can be partitioned into finitely many subintervals such that w' is weakly decreasing on the interior of each of the subintervals. If w is differentiable on $[0, 1]$, we can conclude from this that w' is decreasing everywhere, i.e., w is concave on the whole interval $[0, 1]$. By Lemma 6.5, this is the case when $\widetilde{p}_0 < p^* < p_0$. If $p^* = 0$ (resp. $p^* = 1$) then w is concave on $(0, 1]$ (resp. $[0, 1)$), and hence also on $[0, 1]$.

Suppose then that $p^* \in (0, 1)$, and $p^* = p_0$ or $p^* = \widetilde{p}_0$. In this case we have to examine the behavior of w at p^*, where it may not be differentiable. We will handle the case $p^* = p_0$. The case $p^* = \widetilde{p}_0$ is handled analogously.

Since w is differentiable on $[0, p^*) \cup (p^*, 1]$, both the left and the right derivatives at p^* exist, and it is sufficient to show that $w'_+(p^*) \leq w'_-(p^*)$. We will show that $w'_+(p^*) = a(p^*, w(p^*))$. Indeed, when $p^* = p_0$ we have $w(p^*) = u(p^*)$, and therefore

$$a(p^*, w(p^*)) = \frac{\mu}{1+\mu} \sup_{p\in(p^*,1]} \frac{u(p) - u(p^*)}{p - p^*}. \tag{44}$$

We distinguish between two cases.

- If $\rho(p^*, w) > p^*$, then the interval $[p^*, p_1]$ is linear, that is, $w = \psi_0$, and we have

$$w'_+(p^*) = \psi'_{0+}(p^*) = a(p^*, w(p^*)).$$

- If $\rho(p^*, w) = p^*$, then the interval $[p^*, p_1]$ is nonlinear. In this case, we have

$$\begin{aligned} w'_+(p^*) = \varphi'_{0+}(p^*) = &\ \lim_{p\searrow p^*} \varphi'_0(p) = \lim_{p\searrow p^*} \mu \cdot \tfrac{u(p)-\varphi_0(p)}{p-p^*} \\ = &\ \mu \lim_{p\searrow p^*} \tfrac{u(p)-u(p^*)}{p-p^*} - \mu \lim_{p\searrow p^*} \tfrac{\varphi_0(p)-\varphi_0(p^*)}{p-p^*} \\ = &\ \mu u'_+(p^*) - \mu\varphi'_{0+}(p^*) = \mu u'_+(p^*) - \mu w'_+(p^*). \end{aligned}$$

By Eq. (44) we deduce that

$$w'_+(p^*) = \frac{\mu}{1+\mu} u'_+(p^*) = \frac{\mu}{1+\mu} \sup_{p\in(p^*,1]} \frac{u(p)-u(p^*)}{p-p^*} = a(p^*, w(p^*)).$$

We now calculate $w'_-(p^*)$. If $\widetilde{p}_0 = p_0 = p^*$, a similar argument shows that $w'_-(p^*) = \widetilde{a}(p^*, w(p^*)) = \frac{\mu}{1+\mu} \inf_{p\in[0,p^*)} \frac{u(p)-u(p^*)}{p-p^*}$. However, in this case, at p_0 the function u is equal to its convex hull, and therefore

$$w'_+(p^*) = \frac{\mu}{1+\mu} \sup_{p\in(p^*,1]} \frac{u(p)-u(p^*)}{p-p^*} \le \frac{\mu}{1+\mu} \inf_{p'\in[0,p^*)} \frac{u(p')-u(p^*)}{p'-p^*} = w'_-(p^*),$$

as desired. If $\widetilde{p}_0 < p_0 = p^*$, Eq. (18) yields

$$w'_-(p^*) = \frac{\mu}{1+\mu} \frac{u(p_0)-u(\widetilde{p}_0)}{p_0-\widetilde{p}_0}.$$

From the definition of $\widetilde{p}_0$ and p_0 we deduce that

$$\frac{u(p_0)-u(\widetilde{p}_0)}{p_0-\widetilde{p}_0} \ge \sup_{p\in(p_0,1]} \frac{u(p)-u(p_0)}{p-p_0}, \tag{45}$$

and once again it follows that $w'_+(p^*) \le w'_-(p^*)$. □

6.3. The functions w and v coincide

Proposition 6.7. *For every $p \in [0,1]$ we have $w(p) = v(p)$.*

Proof. To prove the claim we show that the function w satisfies the conditions of Theorem 2.3. Condition G.1 holds by the definition of w on the interval $[\widetilde{p}_0, p_0]$. By Lemmas 6.5 and 6.6, w is concave and differentiable on $[0,1]\setminus\{p^*\}$.

Since w is affine on the interval $[\widetilde{p}_0, p_0]$ and on linear intervals $[p_k, p_{k+1}]$, and since by Lemma 6.3 the two end-points of these intervals lie in nonlinear intervals, it follows that all the extreme points of the hypograph of w lie in nonlinear intervals $[p_k, p_{k+1}]$. On these intervals, from Eq. (26), the relation $w'(p)(p-p^*) + \mu(w(p)-u(p)) = 0$ holds, and therefore Condition G.3 holds. Moreover, Condition G.2 holds on nonlinear intervals $[p_k, p_{k+1}]$.

It remains to show that Condition G.2 holds: $w'(p)(p-p^*) + \mu(w(p)-u(p)) \ge 0$ on the interval $[\widetilde{p}_0, p_0]$ and on linear intervals. On a linear interval $(p_k, p_{k+1}]$ we have $w(p) = w(p_k) + (p-p_k)a(p_k, w(p_k))$ and $w'(p) = a(p_k, w(p_k))$. It then follows by the definition of $a(p_k, w(p_k))$ that on these intervals

$$\begin{aligned}(p-p^*)w'(p) \quad &+ \quad \mu(w(p)-u(p)) \\ &= \quad (p-p^*)a(p_k, w(p_k)) + \mu\big(w(p_k) + (p-p_k)a(p_k, w(p_k)) - u(p)\big) \\ &= \quad \big(p-p^* + \mu(p-p_k)\big)a(p_k, w(p_k)) + \mu\big(w(p_k) - u(p)\big) \\ &\ge \quad 0,\end{aligned}$$

as desired.

On the interval $[\widetilde{p}_0, p_0]$ the function w is affine, thus w' is constant and therefore the function $w(p) + \frac{w'(p)(p-p^*)}{\mu}$ is affine as well. The points $(\widetilde{p}_0, u(\widetilde{p}_0))$ and $(p_0, u(p_0))$ are on the graph of this last function. These points and the interval connecting them are on the graph of the function cav u. It follows that for every $p \in (\widetilde{p}_0, p_0)$ we have $u(p) \le w(p) + \frac{w'(p)(p-p^*)}{\mu}$, which implies that $w'(p)(p-p^*) + \mu(w(p)-u(p)) \ge 0$. □

7. Appendix: the computation for Example 3, variations a. and b.

7.1. Example 3, variation a.

We here analyze the algorithm when $\lambda_1 = 3$ and $\lambda_2 = r = 1$, which implies that $\mu = \frac{1}{4}$ and $p^* = \frac{1}{4}$.

1. Initialization: Simple calculations yield that $\widetilde{p}_0 = 0$ and $p_0 = \frac{1}{3}$. By Eq. (18) we obtain $w(p) = \frac{2}{15}(3p-2)$ for $p \in [0, \frac{1}{3}]$. In particular $w(\frac{1}{3}) = -\frac{2}{15}$.
2. We compute the solution of Eq. (19) with the initial condition $\varphi(\frac{1}{3}) = -\frac{2}{15}$. For $p \in [\frac{1}{3}, \frac{2}{3}]$, the solution is

$$\varphi(p) = -\frac{2}{15} 3^{-1/4}(4p-1)^{-1/4},$$

and we obtain

$$a(p, \varphi(p)) = \max\left\{\frac{u(p)-\varphi(p)}{4p-1}, \frac{u(1)-\varphi(p)}{4-p}\right\} = \max\left\{-\frac{\varphi(p)}{4p-1}, \frac{\frac{2}{3}-\varphi(p)}{4-p}\right\}.$$

Following Step I.5 of the algorithm, p_1 is the last $p \in (\frac{1}{3}, 1]$ that satisfies the relation

$$a(p, \varphi(p)) = \frac{u(p) - \varphi(p)}{4p - 1}.$$

On $[\frac{1}{3}, \frac{2}{3}]$, this relation is equivalent to

$$(4p-1)^{-\frac{5}{4}}(1-p) = 3^{\frac{1}{4}},$$

which yields $p_1 \simeq 0.3858$.

3. The next step is to determine p_2 and the function w on the interval $(p_1, p_2]$. As already noted, we have

$$a(p_1, w(p_1)) = \sup_{p \in [p_1, 1]} \frac{u(p') - w(p_1)}{4p - p_1 - 1} = \frac{\frac{2}{3} - w(p_1)}{4 - p_1},$$

and the supremum is attained at $p' = 1$. Let $\Psi(p) = w(p_1) + (p - p_1)a(p_1, w(p_1))$, for every $p \in [p_1, 1]$. It can be shown that, for every $p \in (p_1, 1)$ we have $a(p_1, w(p_1)) > \frac{u(p) - \Psi(p)}{4p-1}$. Therefore $p_2 = 1$ and $w(p) = \Psi(p)$ for every $p \in [p_1, 1]$. The first part of the algorithm ends.

4. Since $\widetilde{p}_0 = 0$, the second part of the algorithm is vacuous.

Thus, the limit value function is given by

$$v(p) = \begin{cases} \frac{2}{15}(3p - 2), & \text{if } 0 \le p < \frac{1}{3}, \\ -\frac{2}{15} \cdot 3^{-1/4} \cdot (4p-1)^{-1/4}, & \text{if } \frac{1}{3} \le p < p_1, \\ ap + b & \text{if } p \in (p_1, 1], \end{cases} \tag{46}$$

where $a = \frac{\frac{2}{3} - w(p_1)}{4 - p_1} \simeq 0.21709$ and $b = \frac{2}{3} - 4a \simeq -0.20177$.

7.2. Example 3, variation b.

We here analyze the algorithm when $\lambda_1 = \frac{4}{3}$, $\lambda_2 = \frac{2}{3}$, and $r = 1$, which implies that $p^* = \frac{1}{3}$ and $\mu = \frac{1}{2}$.

1. Initialization: Simple calculations show that $(\text{cav } u)(\frac{1}{3}) = u(\frac{1}{3}) = 0$ and $\widetilde{p}_0 = p_0 = \frac{1}{3}$.
2. We turn to the first part of the algorithm. The reader can verify that

$$a(\tfrac{1}{3}, w(\tfrac{1}{3})) = \sup_{p' \in (\frac{1}{3}, 1]} \frac{\mu(u(p') - w(\frac{1}{3}))}{p' - \frac{1}{3} + \mu(p' - \frac{1}{3})} = \max_{p' \in [\frac{2}{3}, 1]} \frac{\frac{9p'^2 - 9p' + 2}{6p' - 3} - 0}{2(p' - \frac{1}{3}) + (p' - \frac{1}{3})} = \max_{p' \in [\frac{2}{3}, 1]} \frac{3p' - 2}{6p' - 3}.$$

This maximum is obtained at $p' = 1$. Therefore, $\rho(p_0, w(p_0)) = 1$, and the condition of Step I.4 holds. We obtain $a(\frac{1}{3}, w(\frac{1}{3})) = \frac{1}{3}$, and the first part of the algorithm ends with $p_1 = 1$.

3. For the second part of the algorithm we compute

$$\begin{aligned} \widetilde{a}(\widetilde{p}_0, w(p_0)) = {} & \widetilde{a}(\tfrac{1}{3}, w(\tfrac{1}{3})) \\ = {} & \inf_{p' \in [0, \frac{1}{3})} \frac{\mu(u(p) - w(\frac{1}{3}))}{p' - \frac{1}{3} + \mu(p' - \frac{1}{3})} = \inf_{p' \in [0, \frac{1}{3})} \frac{\frac{9p'^2 - 9p' + 2}{6p' - 3}}{3p' - 1} = \inf_{p' \in [0, \frac{1}{3})} \frac{3p' - 2}{6p' - 3}. \end{aligned}$$

This infimum is attained at $p' = 0$. Therefore, $\rho(\widetilde{p}_0, w(\widetilde{p}_0)) = 0$, and the condition of Step D.4 holds. We obtain $\widetilde{a}(\frac{1}{3}, w(\frac{1}{3})) = \frac{2}{3}$, and the algorithm ends with $\widetilde{p}_1 = 0$.

In conclusion, the limit value function is given by:

$$v(p) = \begin{cases} \frac{2}{3}p - \frac{2}{9}, & \text{if } 0 \le p < \frac{1}{3}, \\ \frac{1}{3}p - \frac{1}{9}, & \text{if } \frac{1}{3} \le p < \bar{1}. \end{cases} \tag{47}$$

Acknowledgments

We thank the referees for their careful reading and their pertinent remarks which helped a lot to improve the paper.

References

Ashkenazi-Golan, G., Hernández, P., Neeman, Z., Solan, S., 2020. Dynamic Bayesian persuasion with two states. Preprint.
Aumann, R.J., Heifetz, A., 2002. Incomplete information. In: Handbook of Game Theory with Economic Applications, Vol. 3, pp. 1665–1686. Chapter 43.
Aumann, R.J., Maschler, M.B., 1968. Repeated games of incomplete information: the zero-sum extensive case. In: Report of the U.S. Arms Control and Disarmament Agency ST-143. Washington, D.C., pp. 37–116. Chapter III.
Aumann, R.J., Maschler, M.B., 1995. Repeated Games with Incomplete Information. MIT Press.
Aumann, R.J., Maschler, M.B., Stearns, R.E., 1968. Repeated games of incomplete information: an approach to the non-zero sum case. In: Report of the U.S. Arms Control and Disarmament Agency ST-143. Washington, D.C., pp. 117–216. Chapter IV.
Bolte, J., Gaubert, S., Vigeral, G., 2015. Definable zero-sum stochastic games. Math. Oper. Res. 40 (1), 171–191.
Cardaliaguet, P., 2006. Differential games with asymmetric information. SIAM J. Control Optim. 46 (3), 816–838.
Cardaliaguet, P., 2009. Numerical approximation and optimal strategies for differential games with lack of information on one side. In: Advances in Dynamic Games and Their Applications. In: Annals of the International Society for Dynamic Games, vol. 10. Birkhauser Boston, Inc., Boston, MA, pp. 159–176.
Cardaliaguet, P., Rainer, C., 2009a. Stochastic differential games with asymmetric information. Appl. Math. Optim. 59 (1), 1–36.
Cardaliaguet, P., Rainer, C., 2009b. On a continuous-time game with incomplete information. Math. Oper. Res. 34 (4), 769–794.
Cardaliaguet, P., Rainer, C., Rosenberg, D., Vieille, N., 2016. Markov games with frequent actions and incomplete information. Math. Oper. Res. 41 (1), 49–71.
Coste, M., 1999. An introduction to o-minimal geometry. Raag notes. Institut de Recherche Mathématiques de Rennes.
Ely, J., 2017. Beeps. Am. Econ. Rev. 107 (1), 31–53.
Gensbittel, F., 2016. Continuous-time limit of dynamic games with incomplete information and a more informed player. Int. J. Game Theory 45 (1-2), 321–352.
Gensbittel, F., 2019. Continuous-time Markov games with asymmetric information. Dyn. Games Appl. 9 (3), 671–699.
Gensbittel, F., Grün, C., 2019. Zero-sum stopping games with asymmetric information. Math. Oper. Res. 44 (1), 277–302.
Gensbittel, F., Rainer, C., 2020. Optimal dynamic provision on continuous time. Working paper.
Gensbittel, F., Renault, J., 2015. The value of Markov chain games with incomplete information on both sides. Math. Oper. Res. 40 (4), 820–841.
Grün, C., 2012a. A BSDE approach to stochastic differential games with incomplete information. Stoch. Process. Appl. 122 (4), 1917–1946.
Grün, C., 2012b. A probabilistic-numerical approximation for an obstacle problem arising in game theory. Appl. Math. Optim. 66 (3), 363–385.
Harsanyi, J.C., 1967. Games with incomplete information played by "Bayesian" players, I–III. Part I. The basic model. Manag. Sci. 14 (3), 159–182.
Hernández, P., Neeman Z., 2018. How Bayesian persuasion can help reduce illegal parking and other socially undesirable behavior. Preprint.
Hörner, J., Rosenberg, D., Solan, E., Vieille, N., 2010. On a Markov game with one-sided information. Oper. Res. 58, 1107–1115.
Mertens, J.-F., Sorin, S., Zamir, S., 2016. Repeated Games. Cambridge University Press.
Mertens, J.F., Zamir, S., 1971. The value of two-person zero-sum repeated games with lack of information on both sides. Int. J. Game Theory 1 (1), 39–64.
Neyman, A., 2008. Existence of optimal strategies in Markov games with incomplete information. Int. J. Game Theory 37 (4), 581–596.
Oliu-Barton, M., 2015. Differential games with asymmetric and correlated information. Dyn. Games Appl. 5 (3), 378–396.
Renault, J., 2006. The value of Markov chain games with lack of information on one side. Math. Oper. Res. 31, 490–512.
Renault, J., 2012. The value of repeated games with an informed controller. Math. Oper. Res. 37 (1), 154–179.
Renault, J., Solan, E., Vieille, N., 2017. Optimal dynamic information provision. Games Econ. Behav. 104, 329–349.