21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Zero-sum Stochastic Games with Nonsymmetric Partial Observation

Arnab Basu[1] and Lukasz Stettner[2]

***Abstract*— Assume that the state process is a controlled Markov process (X_t), which is partially observed with observation process (Y_t). We consider a game with two players: the maximizer who controls off-line he is choosing for each time a control function (control rule), which is a function of the current state of the process and filtering process plus eventually (in the case of MINMAX game) current value of control of the minimizer (the oponent). The minimizer knows control functions of the maximizer but gets only a partial observation (Y_t) of the state process plus, in the case of MAXMIN game, the current value of control of the maximizer. In the paper upper and lower values of the game are characterized and a simple counterexample for which upper value is strictly greater than the lower value of the game is shown.**

I. Model Formulation

We consider a nonsymmetric partially observed zero-sum (between two players - maximizer and minimizer) discrete-time stochastic game where the maximizer i.e. player I controls off-line choosing control functions, which are functions of the current state of the process plus filtering process and in the case of MINMAX game also control of the minimizer, whereas player II (the minimizer) gets a corrupted information about the state process and knows previous control functions chosen by the maximizer plus additionally in the case of MAXMIN game control value values of maximizer. Such an asymmetrical partially-observed stochastic game (POASG) is determined by eight objects $(E, D, U, V, p, q, \mu, \{r_t\}_{t\in\mathbb{N}_0})$ where E is a finite state space which is observable to player I but not to player II, D is a finite observation space observable to both players, U and V are finite action spaces of players I and II, resp. Denote by $\mathcal{P}(E)$ the set of probability measures on E. Let $\{X_t\}_{t\in\mathbb{N}_0}$ be a controlled (by both players) Markov chain on E with stochastic transition kernel $p : E\times U\times V \mapsto \mathcal{P}(E)$ given by

$$P\left[X_{t+1} = i|\mathcal{F}_t\right] = p(i|X_t, u_t, v_t),\ i \in E, \tag{1}$$

with $\{u_t\}_{t\in\mathbb{N}_0}, \{v_t\}_{t\in\mathbb{N}_0}$ being the control processes of player I and II resp. taking values in U and V resp., and $\mathcal{F}_t \equiv \sigma\{(X_s, Y_s) : 0 \le s \le t\}$ with $\{Y_t\}_{t\in\mathbb{N}_0}$ being the corresponding observation process on D with observation transition kernel $q : E \mapsto \mathcal{P}(D)$ given by

$$P\left[Y_t = d|\mathcal{X}^{(t)} \vee \mathcal{F}_{t-1}\right] = q(d|X_t),\ d \in D, \tag{2}$$

where $\mathcal{X}^{(t)} \equiv \sigma\{X_s : 0 \le s \le t\}$. We denote by $\mu \in \mathcal{P}(X)$ the initial state distribution of $\{X_t\}_{t\in\mathbb{N}_0}$, by $r_t : E\times U\times V \mapsto \mathbb{R}$ the one-stage cost function at time t, the observation σ field $\mathcal{Y}^{(t)} \equiv \sigma\{Y_s : 0 \le s \le t\}$ and the maximizer controls σ field $\mathcal{U}^{(t)} = \sigma\{u_s, s \le t\}$ up to time t. The players depending on the kind of game have or do not have current information concerning the control chosen by their oponent. To be more precise in the MINMAX game the first player (maximizer) at each time moment gets an information about the current control chosen by the minimiser. In the case of MAXMIN game the second player (minimizer) at each time moment gets the control chosen by the maximizer. The strategy u_t of the maximizer at time t in the MINMAX game is of the form of function $\tilde{u}_t$ of X_t, filtering process (defined later) π_t and control of the minimizer v_t taking values in U, while in the case of MAXMIN game it is a function of X_t, filtering process $\tilde{\pi}_t^-$ with observation up to time $t-1$ of the maximizer controls. We shall use the following notation $\mathcal{U}_{MINMAX} \overset{def}{=} \{\tilde{u} : E \times \mathcal{P}(E) \times V \to U\}$ and $\mathcal{U}_{MAXMIN} \overset{def}{=} \{\tilde{u} : E \times \mathcal{P}(E) \to U\}$. The strategy of the minimizer v_t at time t takes values in V and is adapted to the filtration $\mathcal{Y}^{(t)}$ in MINMAX game and to $\mathcal{Y}^{(t)} \vee \mathcal{U}^{(t)}$ for MAXMIN game.

The cost functional is over a finite horizon T and is defined as

$$\mathcal{J}_T(\mu, \{u_t\}, \{v_t\}) \overset{def}{=} E_\mu\left[\sum_{t=0}^{T} r_t(X_t, u_t, v_t)\right],\ i \in E, \tag{3}$$

where, without loss of generality, we assume $r_T(i, u, v) = r_T(i)$ for all $(i, u, v) \in E \times U \times V$, assuming that initial state of the process (X_t) is μ.

There are many papers concerning stochastic games with partial for both players but there are of few papers concerning dynamical models in which the observation is asymmetric. In the class of repeated dynamical models with asymmetric information we have the papers by Sorin [6], [7], [8] and the review paper by Zamir [9]. Particular versions of such games were studied in [4] and [5]. In this paper we are studying general version Markov game with partial observation in the setting typical for stochastic system theory in which we have state and observation processes. To avoid problems with selectors we consider the simplest case with finite state, observation and control parameters spaces. The paper is using various results from discrete time filtering theory (see

The work of A. Basu was supported by Department of Science and Technology, Govt. of India research grant no: SR/FTP/MS-08/2009 and in part by Research Chair grant no: 190 of IIM Bangalore as well as Warsaw Center of Mathematical and Computational Science grant. The work of L. Stettner was supported by NCN grant DEC-2012/07/B/ST1/03298

[1] Arnab Basu is with Quantitative Methods and Information Systems Area, Indian Institute of Management Bangalore, India 560076. arnab.basu@iimb.ernet.in

[2] Lukasz Stettner is with Institute of Mathematics Polish Academy of Sciences, Sniadeckich 8, 00-956 Warsaw, also Vistula University stettner@impan.pl

ISBN: 978-90-367-6321-9

[3]). The results in this note are formulated without proofs. For proofs and further details see [1].

II. Filtering processes

The minimizer does not observe the state process. However based on his available observation he is able to track conditional law of the state process. Let for $i \in E$, $l \in D$, $\eta \in \mathcal{P}(E)$, $\tilde{u} \in \mathcal{U}_{MINMAX}$ and $v \in V$

$$\tilde{\mathfrak{M}}^{\tilde{u},v}(l,\eta)(i) \stackrel{def}{=} \frac{\sum_{j\in E} q(l|i)p(i|j,\tilde{u}(j,\eta,v),v)\eta(j)}{\sum_{k\in E}\sum_{j\in E} q(l|k)p(k|j,\tilde{u}(j,\eta,v),v)\eta(j)} \tag{4}$$

and for any bounded measurable $f : E\times\mathcal{P}(E) \mapsto \mathbb{R}$, $\tilde{u} \in \mathcal{U}_{MINMAX}$, $v \in V$,

$$\tilde{\mathfrak{P}}^{\tilde{u},v} f(j,\eta) \stackrel{def}{=} \sum_{k\in E}\sum_{l\in D} q(l|k)p(k|j,\tilde{u}(j,\eta,v),v)f(k,\tilde{\mathfrak{M}}^{\tilde{u},v}(l,\eta)). \tag{5}$$

Define also for $y \in D$ and $\eta \in \mathcal{P}(E)$

$$M_0(y,\eta)(i) \stackrel{def}{=} \frac{q(Y_0|i)\mu(i)}{\sum_{k\in E} q(Y_0|k)\mu(k)} \tag{6}$$

Given control processes $\{u_t\}_{t\in\mathbb{N}_0}, \{v_t\}_{t\in\mathbb{N}_0}$, where $\{u_t\}_{t\in\mathbb{N}_0}$ is determined by the functions $\tilde{u}_t \in \mathcal{U}_{MINMAX}$ and initial law μ of X_0 define filtering process $\{\pi_t\}_{t\in\mathbb{N}_0}$ as follows for $i \in E$

$$\pi_0(i) \stackrel{def}{=} M_0(Y_0,\mu)(i) \tag{7}$$

and for $t \geq 1$ recursively

$$\pi_t(i) \stackrel{def}{=} \tilde{\mathfrak{M}}^{\tilde{u}_{t-1},v_{t-1}}(Y_t,\pi_{t-1})(i) \tag{8}$$

Following Lemma 1.1 of [3] we have

Theorem 1. *The filtering process (Y_t) has the following interpretation for $i \in E$, P a.e.*

$$\pi_t(i) = P\{X_t = i|\mathcal{Y}^{(t)}\}. \tag{9}$$

Moreover for a bounded measurable $f : E\times\mathcal{P}(E) \mapsto \mathbb{R}$ we have

$$E\{f(X_{t+1},\pi_{t+1})|\mathcal{F}_t\} = \tilde{\mathfrak{P}}^{\tilde{u}_t,v_t} f(X_t,\pi_t) \tag{10}$$

P a.e. for $t \geq 0$.

Let now for $i \in E$, $l \in D$, $\eta \in \mathcal{P}(E)$, $u \in U$ and $v \in V$

$$\mathfrak{M}^{u,v}(l,\eta)(i) \stackrel{def}{=} \frac{\sum_{j\in E} q(l|i)p(i|j,u,v)\eta(j)}{\sum_{k\in E}\sum_{j\in E} q(l|k)p(k|j,u,v)\eta(j)} \tag{11}$$

and also for $\tilde{u}_t \in \mathcal{U}_{MAXMIN}$,

$$K_t(u,\eta)(i) \stackrel{def}{=} \frac{\mathbf{1}_{\tilde{u}_t(i,\eta)=u}\eta(i)}{\sum_{j\in E}\mathbf{1}_{\tilde{u}_t(j,\eta)=u}\eta(j)} \tag{12}$$

for any bounded measurable $f : E\times\mathcal{P}(E) \mapsto \mathbb{R}$, $u \in U$, $v \in V$,

$$\mathfrak{P}^{u,v} f(j,\eta) \stackrel{def}{=} \sum_{k\in E}\sum_{l\in D} q(l|k)p(k|j,u,v)f(k,\mathfrak{M}^{u,v}(l,\eta)). \tag{13}$$

Furthermore let for $i \in E, \eta \in \mathcal{P}(E)$, $u \in U$, $v \in V$, $\tilde{u}_0, \in \mathcal{U}_{MAXMIN}$,

$$\tilde{\pi}_0^-(i) \stackrel{def}{=} \frac{q(Y_0|i)\mu(i)}{\sum_k q(Y_0|k)\mu(k)}, \tag{14}$$

$$\tilde{\pi}_0^u(i) \stackrel{def}{=} K_0(u,\tilde{\pi}_0^-)(i), \tag{15}$$

and

$$\tilde{\pi}_0(i) \stackrel{def}{=} \tilde{\pi}_0^{\tilde{u}_0(X_0,\tilde{\pi}_0^-)}(i). \tag{16}$$

Then we define inductively for $t \geq 1$, $i \in E$ and control processes $\{u_t\}_{t\in\mathbb{N}_0}, \{v_t\}_{t\in\mathbb{N}_0}$, where $\{u_t\}_{t\in\mathbb{N}_0}$ is determined by the functions $\tilde{u}_t \in \mathcal{U}_{MAXMIN}$

$$\tilde{\pi}_t^-(i) \stackrel{def}{=} \mathfrak{M}^{u_{t-1},v_{t-1}}(Y_t,\tilde{\pi}_{t-1})(i), \tag{17}$$

$$\tilde{\pi}_t^u(i) \stackrel{def}{=} K_t(u_t,\tilde{\pi}_t^-)(i), \tag{18}$$

and

$$\tilde{\pi}_t(i) \stackrel{def}{=} \tilde{\pi}_t^{\tilde{u}_t(X_t,\tilde{\pi}_t^-)}(i). \tag{19}$$

Theorem 2. *We have that P a.e. for $i \in E$, $u \in U$ and $t \geq 0$*

$$\tilde{\pi}_t(i) = P\{X_t = i|\mathcal{Y}^{(t)} \vee \mathcal{U}^{(t)}\}, \tag{20}$$

$$\tilde{\pi}_t^u(i) = P\{X_t = i|\left\{\tilde{u}_t(X_t,\pi_t^-) = u\right\} \vee \mathcal{Y}^{(t)} \vee \mathcal{U}^{(t-1)}\}, \tag{21}$$

with $\mathcal{U}^{(t-1)} = \{\emptyset,\Omega\}$, conditional mean with respect to an event and σ field defined in the Appendix and

$$\tilde{\pi}_t^-(i) = P\{X_t = i|\mathcal{Y}^{(t)} \vee \mathcal{U}^{(t-1)}\}. \tag{22}$$

Furthermore for a bounded measurable $f : E\times\mathcal{P}(E) \mapsto \mathbb{R}$ we have

$$E\{f(X_{t+1},\tilde{\pi}_{t+1}^-)|\mathcal{F}_t\} = \mathfrak{P}^{u_t,v_t} f(X_t,\tilde{\pi}_t) \tag{23}$$

P a.e. for $t \geq 0$.

III. MINMAX game

In this game the maximizer has a priority in the sense that he knows the control chosen by the minimizer. We first define by backward induction a system of equations and selectors.

Let $\bar{w}_T(i) = r_T(i)$ for $i \in E$ and

$$\bar{w}_{T-1}^+(i,v) = \sup_{u\in U}\left[r_{T-1}(i,u,v) + P^{uv}\bar{w}_T(i)\right] \tag{24}$$

for

$$P^{uv}\bar{w}_T(i) = \sum_{j\in E}\bar{w}_T(j)p(j|i,u,v) \tag{25}$$

with supremum attained for $\hat{u}_{T-1}(i,v)$ which defines optimal $\hat{\tilde{u}}_{T-1} \in \mathcal{U}_{MINMAX}$ at time $T-1$. Let $\hat{v}_{T-1}$ be defined for $\eta \in \mathcal{P}(E)$ as selector in the equation

$$\inf_{v\in V}\sum_{k\in E}\bar{w}_{T-1}^+(k,v)\eta(k) = \sum_{k\in E}\bar{w}_{T-1}^+(k,\hat{v}_{T-1}(\eta))\eta(k). \tag{26}$$

Define

$$\bar{w}_{T-1}(i,\eta) = \bar{w}_{T-1}^+(i,\hat{v}_{T-1}(\eta)). \tag{27}$$

Further step in backward induction is defined as follows: for $i \in E$, $\eta \in \mathcal{P}(E)$ and $\tilde{u} \in \mathcal{U}_{MINMAX}$, $v \in V$ let for $t \leq T-2$

$$\bar{w}_t^+(i,\eta,\tilde{u},v) = r_t(i,\tilde{u}(i,\eta,v),v) + \tilde{\mathfrak{P}}^{\tilde{u},v}\bar{w}_{t+1}(i,\eta) \tag{28}$$

For a given control function $\tilde{u} \in \mathcal{U}_{MINMAX}$ of the maximizer the minimizer is choosing the control $\hat{v}_t^{\tilde{u}}(\eta)$ such that

$$\inf_{v\in V}\sum_{k\in E}\bar{w}_t^+(k,\eta,\tilde{u},v)\eta(k) = \sum_{k\in E}\bar{w}_t^+(k,\eta_t,\tilde{u},\hat{v}_t^{\tilde{u}}(\eta))\eta(k). \tag{29}$$

The maximizer is choosing his control function $\hat{\tilde{u}}_t$ at time t as the maximizer of

$$\sup_{\tilde{u}\in\mathcal{U}_{MINMAX}} E_\mu\left[\bar{w}_t^+(X_t,\pi_t,\tilde{u},\hat{v}_t^{\tilde{u}}(\pi_t))\right] = E_\mu\left[\bar{w}_t^+(X_t,\pi_t,\hat{\tilde{u}}_t,\hat{v}_t^{\hat{\tilde{u}}_t}(\pi_t))\right] \tag{30}$$

Note that the above supremum is attained since random variables under expected value have only finite number of admissible values therefore in fact we maximize over a finite number of values. Consequently the upper value of the game at time t is equal to

$$\bar{w}_t(i,\eta) = \bar{w}_t^+(i,\eta,\hat{\tilde{u}}_t,\hat{v}_t^{\hat{\tilde{u}}_t}(\eta)) \tag{31}$$

and by backward induction we obtain successively the value $\bar{w}_0(i,\eta)$.

Theorem 3. *The function of μ of the form*

$$\sum_{i\in E}\sum_{y\in D}\bar{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{32}$$

is the upper value of the game with cost functional (3) *in the sense for any sequence of control functions $(\tilde{u}_t)$, where $\tilde{u}_t \in \mathcal{U}_{MINMAX}$ and corresponding sequence of control values u_t we have*

$$\mathcal{J}_T(\mu,\{u_t\},\{\hat{v}_t\}) \leq \sum_{i\in E}\sum_{y\in D}\bar{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{33}$$

for $\hat{v}_t = \hat{v}_t^{u_t}(\pi_t)$ for $t \leq T-2$ and $\hat{v}_{T-1} = \hat{v}_{T-1}(\pi_{T-1})$ with equality whenever $\hat{u}_t = \hat{\tilde{u}}_t(X_t,\pi_t,\hat{v}_t^{\hat{\tilde{u}}_t})$ for $t \leq T-2$ and $\hat{u}_{T-1} = \hat{\tilde{u}}_{T-1}(X_{T-1},\hat{v}_{T-1})$. Moreover for any control (v_t) of the minimizer we have

$$\mathcal{J}_T(\mu,\{\hat{u}_t\},\{v_t\}) \geq \sum_{i\in E}\sum_{y\in D}\bar{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{34}$$

for $\hat{u}_t = \hat{\tilde{u}}_t(X_t,\pi_t,v_t)$ for $t \leq T-2$ and $\hat{u}_{T-1} = \hat{\tilde{u}}_{T-1}(X_{T-1},v_{T-1})$ with equality for $\hat{v}_t = \hat{v}_t^{\hat{\tilde{u}}_t}(\pi_t)$ for $t \leq T-2$ and $\hat{v}_{T-1} = \hat{v}_{T-1}(\pi_{T-1})$.

IV. MAXMIN GAME

We now assume that the minimizer knows at each time the current value of the control of the maximizer. By backward induction again we have $\underline{w}_T(i) = r_T(i)$ and

$$\underline{w}_{T-1}^+(i,u,v) \stackrel{def}{=} r_{T-1}(i,u,v) + P^{uv}\underline{w}_T(i). \tag{35}$$

Let for given $u \in U$, $i \in E$, $\eta \in \mathcal{P}(E)$, the value $\hat{v}_{T-1}(\eta,u)$ be defined as minimizer of

$$\inf_{v\in V}\sum_{j\in E}\underline{w}_{T-1}^+(j,u,v)\eta(j) = \sum_{j\in E}\underline{w}_{T-1}^+(j,u,\hat{v}_{T-1}(\eta,u))\eta(j). \tag{36}$$

Then $\hat{\tilde{u}}_{T-1} \in \mathcal{U}_{MAXMIN}$ is defined as maximizer in

$$\begin{aligned}\sup_{\tilde{u}\in\mathcal{U}_{MAXMIN}} & E_\mu\Big[\underline{w}_{T-1}^+(X_{T-1}, \tilde{u}(X_{T-1},\tilde{\pi}_{T-1}^-),\hat{v}_{T-1}(\tilde{\pi}_{T-1}^{\tilde{u}(X_{T-1},\tilde{\pi}_{T-1}^-)}, \tilde{u}(X_{T-1},\tilde{\pi}_{T-1}^-)))\Big] = \\ & E_\mu\Big[\underline{w}_{T-1}^+(X_{T-1},\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-), \hat{v}_{T-1}(\tilde{\pi}_{T-1}^{\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)}, \hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)))\Big]\end{aligned} \tag{37}$$

and we define for $i \in E$ and $\eta \in \mathcal{P}(E)$

$$\underline{w}_{T-1}(i,\eta) \stackrel{def}{=} \underline{w}_{T-1}^+(i,\hat{\tilde{u}}_{T-1}(i,\eta), \hat{v}_{T-1}(K_{T-1}(\hat{\tilde{u}}_{T-1}(i,\eta),\eta),\hat{\tilde{u}}_{T-1}(i,\eta))). \tag{38}$$

Inductively we define for $t \leq T-2$ given $\underline{w}_{t+1}$

$$\underline{w}_t^+(i,\eta,u,v) \stackrel{def}{=} r_t(i,u,v) + \mathfrak{P}^{uv}\underline{w}_{t+1}(i,\eta). \tag{39}$$

and $\hat{v}_t(\eta,u)$ is defined as minimizer in

$$\inf_{v\in V}\sum_{j\in E}\underline{w}_t^+(j,\eta,u,v)\eta(j) = \sum_{j\in E}\underline{w}_t^+(j,\eta,u,\hat{v}_t(\eta,u))\eta(j). \tag{40}$$

The optimal control function $\hat{\tilde{u}}_t$ of the maximizer is defined by the formula

$$\begin{aligned}\sup_{\tilde{u}\in\mathcal{U}_{MAXMIN}} & E_\mu\Big[\underline{w}_t^+(X_t,\tilde{\pi}_t^{\tilde{u}(X_t,\tilde{\pi}_t^-)},\tilde{u}(X_t,\tilde{\pi}_t^-), \hat{v}_t(\tilde{\pi}_t^{\tilde{u}(X_t,\tilde{\pi}_t^-)},\tilde{u}(X_t,\tilde{\pi}_t^-)))\Big] = \\ & E_\mu\Big[\underline{w}_t^+(X_t,\tilde{\pi}_t^{\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)}, \hat{\tilde{u}}_t(X_t,\pi_t^-),\hat{v}_t(\tilde{\pi}_t^{\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)},\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)))\Big]\end{aligned} \tag{41}$$

and finally we define lower value of the game at time t

$$\underline{w}_t(i,\eta) \stackrel{def}{=} \underline{w}_t^+(i,K_t(\hat{\tilde{u}}_t(i,\eta),\eta), \hat{\tilde{u}}_t(i,\eta),\hat{v}_t(K_t(\hat{\tilde{u}}_t(i,\eta),\eta),\hat{\tilde{u}}_t(i,\eta))). \tag{42}$$

Theorem 4. *The function of μ of the form*

$$\sum_{i\in E}\sum_{y\in D}\underline{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{43}$$

is the lower value of the game with cost functional (3) *in the sense for any sequence of control functions $(\tilde{u}_t)$ such that $\tilde{u}_t\in\mathcal{U}_{MAXMIN}$ that determines control sequence of the maximizer (u_t) and $\hat{v}_t=\hat{v}_t(\tilde{\pi}_t^{\tilde{u}_t(X_t,\tilde{\pi}_t^-)},\tilde{u}_t(X_t,\tilde{\pi}_t^-)))$, for $t\leq T-2$ and $\hat{v}_{T-1}=\hat{v}_{T-1}(\tilde{\pi}_{T-1}^{\tilde{u}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)},\tilde{u}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)))$ we have*

$$\mathcal{J}_T(\mu,\{u_t\},\{\hat{v}_t\})\leq\sum_{i\in E}\sum_{y\in D}\underline{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{44}$$

with equality for $\hat{u}_t=\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)$ for $t\leq T-2$ and $\hat{u}_{T-1}=\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)$. Furthermore for any strategy of the minimizer (v_t) we have

$$\mathcal{J}_T(\mu,\{\hat{u}_t\},\{v_t\})\geq\sum_{i\in E}\sum_{y\in D}\underline{w}_0(i,M_0(y,\mu))q(y|i)\mu(i) \tag{45}$$

where $\hat{u}_t=\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)$ for $t\leq T-2$ and $\hat{u}_{T-1}=\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)$ with equality for $\hat{v}_t=\hat{v}_t(\tilde{\pi}_t^{\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)},\hat{\tilde{u}}_t(X_t,\tilde{\pi}_t^-)))$ for $t\leq T-2$ and $\hat{v}_{T-1}=\hat{v}_{T-1}(\tilde{\pi}_{T-1}^{\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)},\hat{\tilde{u}}_{T-1}(X_{T-1},\tilde{\pi}_{T-1}^-)))$.

V. COUNTEREXAMPLE

We consider time horizon $T=2$ with $E=D=\{0,1\}=U=V$. Let $r_0=r_2\equiv 0$, $r_1>0$ and $r_1(1,1,1)=0$, $r_1(1,1,0)=r_1$, $r_1(0,1,1)=r_1$, $r_1(0,1,0)=0$, $r_1(1,0,0)=0$, $r_1(1,0,1)=r_1$, $r_1(0,0,1)=0$, $r_1(0,0,0)=r_1$.

Then $\max_{u\in U}r_1(1,u,0)=r_1$ with $\hat{u}_1(1,0)=1$, $\max_{u\in U}r_1(1,u,1)=r_1$ with $\hat{u}_1(1,1)=0$, $\max_{u\in U}r_1(0,u,0)=r_1$ with $\hat{u}_1(0,0)=0$, $\max_{u\in U}r_1(0,u,1)=r_1$ with $\hat{u}_1(0,1)=1$, and consequently $\bar{w}_1\equiv r_1$ with optimal control function $\hat{\tilde{u}}_1(i,\eta,v)=\mathbf{1}_{i\neq v}$ and arbitrary $v_1\in V$.

Now for $\eta\in\mathcal{P}(E)$

$$\inf_{v\in V}\left[r_1(0,1,v)\eta(0)+r_1(1,1,v)\eta(1)\right]=r_1\left[\eta(0)\wedge\eta(1)\right] \tag{46}$$

with optimal control $\hat{v}_1(\eta,1)=\mathbf{1}_{\eta(0)\leq\eta(1)}$, and

$$\inf_{v\in V}\left[r_1(0,0,v)\eta(0)+r_1(1,0,v)\eta(1)\right]=r_1\left[\eta(0)\wedge\eta(1)\right] \tag{47}$$

with $\hat{v}_1(\eta,1)=\mathbf{1}_{\eta(1)<\eta(0)}$ as optimal control.

Therefore for any $u\in U$ we have $r(1,u,\hat{v}_1(u,\eta))=r_1\mathbf{1}_{\eta(1)<\eta(0)}$ and $r(0,u,\hat{v}_1(u,\eta))=r_1\mathbf{1}_{\eta(0)\leq\eta(1)}$. Consequently no matter what control function $\tilde{u}_1\in\mathcal{U}_{MAXMIN}$ we choose we obtain $\underline{\mathrm{w}}_1(1,\eta)=r_1\mathbf{1}_{\eta(1)<\eta(0)}$ and $\underline{\mathrm{w}}_1(0,\eta)=r_1\mathbf{1}_{\eta(0)\leq\eta(1)}$. and we see that depending on the value of the process $\pi_1\equiv\eta$ we may be below the value r_1, which corresponds to the upper value of the game.

APPENDIX

Given two events A and B we define conditional probability

$$P_A(B)\stackrel{def}{=}P\left[A|B\right]=\frac{P\left[A\cap B\right]}{P\left[B\right]}. \tag{48}$$

Moreover for random variables X and Y and an event A such that $P(A)>0$ the conditional law $P\left[X=j|A\vee\sigma(Y)\right]$ is understood as $P_A\left[X=j|\sigma(Y)\right]$. We can even extend this value to the such events A for which $P(A)=0$ letting then $P\left[X=j|A\vee\sigma(Y)\right]:=P\left[X=j|\sigma(Y)\right]$. The following Lemma was used in the analysis of the lower value function i.e. the max-min case.

Lemma 1. *Given three random variables X,Z,Y taking values in a countable set, say, $S\equiv\mathbb{N}$, we have P a.e.*

$$P\left[X=j|\sigma(Z)\vee\sigma(Y)\right]=f(j,Z,Y), \tag{49}$$

where

$$f(j,i,k)=P\left[X=j|Z=i,Y=k\right] \tag{50}$$

provided that $P\left[Z=i,Y=k\right]>0$. Furthermore,

$$P\left[X=j|\{Z=i\}\vee\sigma(Y)\right]\stackrel{def}{=} P_{\{Z=i\}}\left[X=j|\sigma(Y)\right]=f(j,i,Y), \tag{51}$$

P a.e. implying that $P\left[X=j|\{Z=i\}\vee\sigma(Y)\right]$ is $\sigma(Y)$-measurable.

In general case when random variable Z can be arbitrary the conditional value $P\left[X=j|\{Z=z\}\vee\sigma(Y)\right]$ is understood as $f(j,z,Y)$ where $f(\cdot,Z,Y)$ is a regular conditional probability (the existence of which follows e.g. from Theorem 5.3 of [2]) such that

$$P\left[X=j|\sigma(Z)\vee\sigma(Y)\right]=f(j,Z,Y) \tag{52}$$

P a.e..

REFERENCES

[1] A. Basu, L. Stettner, Zero-sum Stochastic Games with Nonsymmetric Partial Observation, in preparation.
[2] O. Kallenberg, Foundations of Modern Probability, Springer 1997,
[3] W. Runggaldier, L. Stettner, Approximations of Discrete Time Partially Observed Control Problems, Applied Mathematics Monographs CNR, Giardini Editori, Pisa 1994.
[4] D. Rosenberg, E. Solan and N. Vieille, The MaxMin value of stochastic games with imperfect monitoring, Internat. J. Game Theory 32(1), Special anniversary issue Part 2, pp. 133-150, 2003.
[5] Rosenberg and N. Vielle, The Maxmin of Recursive Games with Incomplete Information on One Side, Math. Oper. Res. 25(1), pp. 23-35, 2000.
[6] S. Sorin, Match with Lack of Information on One Side, Part 1, Internat. J. Game Theory 13(4), pp. 201-255, 1984.
[7] S. Sorin, Big Match with Lack of Information on One Side, Part 2, Internat. J. Game Theory 14(3), pp. 173-204, 1985.
[8] S. Sorin and S. Zamir, A 2-Person Game with Lack of Information on $1\frac{1}{2}$ Sides, Math. Oper. Res. 10(1), pp. 17-23, 1985.
[9] S. Zamir, Repeated Games of Incomplete Information: Zero-Sum, Handbook of Game Theory with Economic Applications, Elsevier Science Publishers, Amsterdam, Holland, Vol. 1, pp. 109-154, 1992.