Repeated games over networks with vector payoffs: the notion of attainability

Ehud Lehrer, Eilon Solan, Dario Bauso

Repeated games over networks with vector payoffs: the notion of attainability

(Invited Paper)

Ehud Lehrer
The School of Mathematical Sciences,
Tel Aviv University,
Tel Aviv 69978, Israel
E-mail: lehrer@post.tau.ac.il

Eilon Solan
The School of Mathematical Sciences,
Tel Aviv University,
Tel Aviv 69978, Israel.
e-mail: eilons@post.tau.ac.il

Dario Bauso
Dipartimento di Ingegneria Chimica,
Gestionale, Informatica e Meccanica
Università di Palermo, Italy
Email: dario.bauso@unipa.it

***Abstract*—We introduce the concept of strongly attainable sets of payoffs in two-player repeated games with vector payoffs in continuous time. A set of payoffs is called *strongly attainable* if player 1 has a strategy guaranteeing, even in the worst case, that the distance between the set and the cumulative payoff shrinks with time to zero. We characterize when any vector is strongly attainable and illustrate the motivation of our study on a multi-inventory application.**

I. Introduction

In the literature on multi-inventory control [1], [3], [4], the supplier tries to meet the demand, which is multi-dimensional and given in continuous time. The goal of the supplier is to ensure that the difference between the *total* demand up to time T and the *total* supply up to time T converges to 0, as T goes to infinity. The literature on Approachability theory, which was initiated by [2], answers a similar question (see, e.g., [5]), when one considers the average demand and average supply instead of total demand and total supply.

Motivated by the problem of multi-inventory control, we study in this paper two-player repeated games in continuous time with vector payoffs. We define a new concept, the concept of *strongly attainable sets*, which is close in spirit to the concept of approachable sets: a set A in the payoff space is *strongly attainable* if player 1 has a strategy such that, for every strategy of player 2, the total payoff up to time T converges to the set A.

The main contribution of this paper are conditions that ensure that all payoff vectors are strongly attainable.

II. Problem Setup and Background

A. System model

We study a two-player repeated game with vector payoffs in continuous time. The set of players is $N = \{1, 2\}$, and the finite set of actions of each player i is A_i. The instantaneous payoff is given by a function $u : A_1 \times A_2 \to \mathbb{R}^m$. We assume w.l.o.g. that payoffs are bounded by 1, so that in fact $u : A_1 \times A_2 \to [-1, 1]^m$. We extend u to $\Delta(A_1) \times \Delta(A_2)$ in a multi-linear fashion. We denote the one-shot vector-payoff game (A_1, A_2, u) by G.

We study non-anticipating behavior strategies with delay. Denote by a_i^t the mixed action chosen by player i in time t.

Definition 1: A *non-anticipating behavior strategy with delay* or simply *behavior strategy* for player i is a process $(a_i^t)_{t\in\mathbb{R}_+}$ with values in $\Delta(A_i)$ such that there exists an increasing sequence of stopping times $(\tau_i^k)_{k\in\mathbb{N}}$ satisfying:

1) τ_i^{k+1} is measurable w.r.t. the information at time τ_i^k (it also depends on the play of the other player up to that time);
2) a_i^t is measurable w.r.t. the information at time τ_i^k, where $\tau_i^k \le t < \tau_i^{k+1}$.

In the sequel we will refer to the stopping times $(\tau_i^k)_{k\in\mathbb{N}}$ in Definition 1 as *the stopping times in the definition of* a_i or as *the stopping times related to* σ_i.

Every pair of non-anticipating strategies with delay $\sigma = (\sigma_i)_{i\in N}$ uniquely determines the play path $(a^t(\sigma))_{t\in\mathbb{R}_+}$. The payoff vector up to time T associated with the pair of strategies σ is given by

$$\gamma^T(\sigma) = \int_0^T u(a^t(\sigma))dt.$$

Since payoffs are bounded by 1, the integral, which is the cumulative payoff up to time T, is well-defined.

For every set $A \subseteq \mathbb{R}^m$ we denote by $B(A, \varepsilon)$ the set of all points whose distance from at least one point in A is less than ε:

$$B(A, \varepsilon) := \{x \in \mathbb{R}^m : d(x, A) < \varepsilon\}.$$

When A is a single point x, we write $B(x, \varepsilon)$ instead of $B(\{x\}, \varepsilon)$.

In this paper we present two new concepts: attainable sets and strongly attainable sets.

Definition 2: (i) The set $A \subseteq \mathbb{R}^m$ is *strongly attainable* by player 1 if there is a strategy σ_1 for player 1 such that for every strategy σ_2 of player 2,

$$\lim_{T\to\infty} d(\gamma^T(\sigma), A) = 0. \tag{1}$$

(ii) The set A is *attainable* by player 1, if the set $B(A, \varepsilon)$ is strongly attainable for every $\varepsilon > 0$.

The definition of an attainable set looks similar to that of approachable set in games played over discrete set of times (see Blackwell, 1956). There is, however, a significant difference between the two.

Notation 1: For every mixed action $p \in \Delta(A_1)$ denote by

$$D_1(p) = \{u(p,q) \colon q \in \Delta(A_2)\},$$

the set of all payoffs that can occur when player 1 plays the mixed action p.

Our goal is to provide a geometric characterization to attainable sets. To this end we define a condition related to auxiliary zero-sum scalar-payoff games. Let $\lambda \in \mathbb{R}^m$. Denote by $\langle \lambda, G\rangle$ the zero-sum game whose set of players and their action sets are as in the original game G. Player 1's payoff[1] is $\langle \lambda, u(a_1,a_2)\rangle$ for every $(a_1,a_2) \in A_1 \times A_2$. As a zero-sum game, the game $\langle \lambda, G\rangle$ has a value, denoted v_λ. The inequality $v_\lambda > 0$ means that there is a mixed action $p \in \Delta(A_1)$ such that $D_1(p)$ is a subset of the open half space $\{x \in \mathbb{R}^m \colon \langle \lambda, x\rangle > 0\}$.

Definition 3: (i) Let $B \subseteq \mathbb{R}^m$ be a subspace and $S \subseteq \Delta(A_1)$ be a subset of player 1's mixed strategies. We say that the payoff function u satisfies *condition* $\boldsymbol{C}(B,S)$ if for every[2] open half space H of B, there is a mixed action $p \in S$ such that $D_1(p) \subseteq H$.

(ii) If $B = \mathbb{R}^m$ and $S = \Delta(A_1)$ we say that u satisfies *condition* $\boldsymbol{C}$.

Note that condition **C** is satisfied if and only if $v_\lambda > 0$ for every $\lambda \in \mathbb{R}^m$.

Remark 1: Standard continuity and compactness arguments imply that if u satisfies condition **C**, then there is $\delta_1 > 0$ such that for every half space F there is $p \in \Delta(A_1)$ satisfying $d(D_1(p), F)) \geq \delta_1$. Stated differently, there is $\delta_2 > 0$ such that for every vector λ whose ℓ_1-norm is 1, $\langle u(p,q), \lambda\rangle < -\delta_2$ for every $q \in \Delta(A_2)$.

The following theorem establishes that condition **C** is equivalent to all vectors being attainable.

Theorem 1: The following statements are equivalent

C1 Every vector $x \in \mathbb{R}^m$ is strongly attainable;
C2 Every vector $x \in \mathbb{R}^m$ is attainable;
C3 The payoff function u satisfies condition **C**.

III. Proof of Theorem 1

We show first that **C3** implies **C1**. Assume that condition **C** is satisfied. Fix a vector x. We show that x is strongly attainable.

We define a strategy σ_1. Fix a constant $\eta = 1$ and let τ_1^k be defined inductively as

$$\tau_1^k = \begin{cases} 0 & k = 1, \\ \tau_1^{k-1} + \frac{\eta}{k} & k > 0. \end{cases} \tag{2}$$

Denote the payoff up-to time τ_1^k by S_k. $\sigma_1(t)$ is defined to be an optimal strategy of player 1 in the game $\langle x - S_k, G\rangle$, for $\tau_1^k \leq t < \tau_1^{k+1}$. That is, σ_1 is constant in the interval $[\tau_1^k, \tau_1^{k+1})$. In case $x - S_k \neq 0$, $\sigma_1(t)$ in this interval is equal to a mixed action that guarantees that the payoff and $x - S_k$ lie on different sides of the hyperplane perpendicular to $x - S_k$. Moreover, in light of Remark 1, there is $\delta_2 > 0$ such that when σ_2 is the strategy played by player 2, $\langle \frac{x-S_k}{\|x-S_k\|_1}, u(\sigma_1(t), \sigma_2(t))\rangle < -\delta_2$.

Denote by a_k the payoff accumulated in the interval $[\tau_1^k, \tau_1^{k+1})$. Thus,

$$\begin{aligned} \|x - S_k\|^2 &= \|x - S_{k-1} - \frac{1}{k} a_k\|^2 = \|x - S_{k-1}\|^2 \\ &+ \frac{1}{k^2}\|a_k\|^2 + 2\frac{1}{k}\|x - S_{k-1}\|_1 \langle \frac{x - S_{k-1}}{\|x - S_{k-1}\|_1}, a_k\rangle. \end{aligned} \tag{3}$$

Thus, $\|x - S_k\|$ is smaller than $\|x - S_{k-1}\|$ as long as

$$\frac{1}{k^2}\|a_k\|^2 + 2\frac{1}{k}\|x - S_{k-1}\|_1 \langle \frac{x - S_{k-1}}{\|x - S_{k-1}\|_1}, a_k\rangle < 0. \tag{4}$$

Recalling that $\|a_k\|^2 \leq 1$, we infer that Eq. (4) occurs when

$$\frac{1}{2k\delta_2} < \|x - S_{k-1}\|_1. \tag{5}$$

We claim that $\liminf \|x - S_k\| = 0$. Otherwise, from a certain k_0 on $\|x - S_k\| > \varepsilon > 0$, meaning that from a certain k_0 on Eq. (5) is satisfied. But then for every $K > k_0$ we obtain from Eq. (3) that

$$\begin{aligned} \|x - S_k\|^2 \leq &\sum_{k=k_0}^{K} \frac{1}{k^2}\|a_k\|^2 \\ &+ 2\frac{1}{k}\|x - S_{k-1}\|_1 \langle \frac{x - S_{k-1}}{\|x - S_{k-1}\|_1}, a_k\rangle \\ \leq \quad &\sum_{k=k_0}^{K} \frac{1}{k^2}\|a_k\|^2 + 2\frac{1}{k}\varepsilon(-\delta_2). \end{aligned} \tag{6}$$

Since the RHS converges to $-\infty$, it implies that the LHS is negative for K large enough, which is a contradiction. Thus indeed, $\liminf \|x - S_k\| = 0$.

We show now that $\limsup \|x - S_k\| = 0$. Fix an $\varepsilon > 0$. By the previous claim there are infinitely many k's for which $\|x - S_k\|^2 < \varepsilon$. If $\|x - S_k\|^2 < \varepsilon$, then $\|x - S_{k+1}\|^2 < \varepsilon + \frac{1}{k^2}\|a_k\|^2 + 2\frac{1}{k}\|x - S_{k-1}\|_1 \langle \frac{x-S_{k-1}}{\|x-S_{k-1}\|_1}, a_k\rangle < 2\varepsilon$ for k sufficiently large. But when k is large enough $\frac{1}{2k\delta_2} < \varepsilon$, and then Eq. (5) is satisfied. In this case $\|x - S_{k+2}\|^2 < \|x - S_k + 1\|^2$. In other words, for k large enough, if $\|x - S_k\|^2 < \varepsilon$, then the next time distance squared, $\|x - S_k + 1\|^2$, cannot jump beyond 2ε. And if this figure jumps above ε, the distance $\|x - S_\ell\|$ is then starting to go down. Thus, for k large enough once $\|x - S_k\|^2$ is smaller than ε, it will remain smaller than 2ε for ever. We conclude that $\limsup \|x - S_k\|^2 < 2\varepsilon$ and since ε is arbitrary, it shows **C1**.

It is clear that **C1** implies **C2** and it remains to show that **C2** implies **C3**. Assume that the payoff function u does not satisfy condition **C**. Then there exists a half space H such that $D_1(p)$ is not a proper subset of H, for every $p \in \Delta(A_1)$. From the minimax theorem this implies that there is $q \in \Delta(A_2)$ such that $D_2(q)$ is disjoint of H. But then any vector $x \in H$ is not attainable by player 1, simply because the strategy

[1] Recall that $\langle x, y\rangle = \sum_{i=1}^m x_i y_i$ for every $x, y \in \mathbb{R}^m$.

[2] H is an open half space if it is of the form $\{x;\ \langle x, \lambda\rangle > 0\}$ for some $\lambda \in \mathbb{R}^m$.

σ_2 that constantly plays q generates a cumulative payoff that, regardless of σ_1 being player by player 1, is always out of H, which shows that **C2** is not satisfied.

IV. Case study

In this section we carry out a numerical analysis aimed at simulating the play path and integral payoff of a given game. The game under consideration is displayed below

$$\left(\begin{array}{cccc} (6,7) & (1,7) & (6,2) & (1,2) \\ (6,-4) & (1,-4) & (6,-9) & (1,-9) \\ (-3,-1) & (-8,-1) & (-3,-6) & (-8,-6) \\ (-3,10) & (-8,10) & (-3,5) & (-8,5) \end{array} \right). \qquad (7)$$

The game is obtained starting from the following multi-inventory case study.

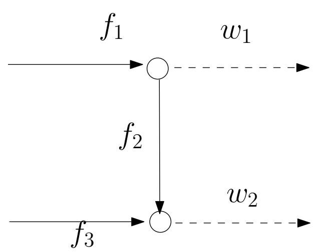

Fig. 1. Network system.

Consider the system depicted in Fig. 1 representing two warehouses, three controlled flows and two uncontrolled flows. A unit of flow f_1 produces one unit of product γ_1 per time unit. Similarly, flow f_2 uses one unit of γ_1 to produce one unit of γ_2 per time unit. A unit of flow f_3 produces one unit of product γ_2 per time unit. Uncontrolled flows w_1 and w_2 represent the exogenous demand of resource γ_1 and γ_2 respectively. The associated dynamics reads then:

$$\left[\begin{array}{c} \dot{\gamma}_1^t \\ \dot{\gamma}_2^t \end{array} \right] = u((a_1^t, a_2^t)) = \underbrace{\left[\begin{array}{ccc} 1 & -1 & 0 \\ 0 & 1 & 1 \end{array} \right]}_{F} \underbrace{\left[\begin{array}{c} f_1^t \\ f_2^t \\ f_3^t \end{array} \right]}_{a_1^t} - \underbrace{\left[\begin{array}{c} w_1^t \\ w_2^t \end{array} \right]}_{a_2^t}.$$

Now, suppose that flows can be processed only in batches and therefore take for instance $f_i \in \{-5,-2,1,6\}$, and $w_i \in \{-3,2\}$.

Let us enumerate all the actions of player 1 and 2, so that we have $A_1 = \{a_{11}, \ldots, a_{1r}\}$ and $A_2 = \{a_{21}, \ldots, a_{2q}\}$ with $r = 4^3$ and $q = 2^2$, where a_{ij} denotes the jth action of player i.

The complete matrix of vector payoffs is then obtained from the following table, where each entry represents a possible vector payoff $u((a_1, a_2))$:

a_1/a_2	a_{21}	$\ldots$	a_{2q}
a_{11}	$Fa_{11} - a_{21}$	$\ldots$	$Fa_{11} - a_{2q}$
$\vdots$	$\vdots$		$\vdots$
a_{1r}			

ϕ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
v_λ	5	3.7	2.4	1.7	1.6	1.5	1.4	1.3	1.2	1.1	1.0

TABLE I
Values of the game $\langle \lambda, G \rangle$ where $\lambda = (\phi, 1-\phi) \in \mathbb{R}^2$ for different ϕ's.

As it will be clearer later on, to our purposes we can simply extract from the above table the rows associated to the following four actions of player 1:

$$a_{11} = (1,-2,6), \quad a_{12} = (1,-2,-5),$$
$$a_{13} = (-5,1,-5), \quad a_{14} = (-5,1,6).$$

This procedure reduces the vector payoff matrix to the 4×4 matrix displayed in (7).

In fact, the above game satisfies condition **C** as $v_\lambda > 0$ for every $\lambda \in \mathbb{R}^m$. Thus, the theoretical results developed in the previous part of the work (see in particular Theorem 1) establish that any vector $x \in \mathbb{R}^2$ is strongly attainable for this game.

To see that $v_\lambda > 0$ for every $\lambda \in \mathbb{R}^m$, we study the one shot game $\langle \lambda, G \rangle$ where $\lambda = (\phi, 1-\phi) \in \mathbb{R}^2$ for $\phi = 0, 0.1, \ldots 1$. More specifically, let $\mathcal{U}_\lambda$ be the payoff matrix and $\mathcal{U}'_\lambda$ its transposed, and consider the optimization variables $\tilde{p} = \frac{p}{v_\lambda}$. Then for each λ we solve the linear program below using Matlab in-built function `linprog`:

$$(LP) \quad \min\{\|\tilde{p}\|_1 | \mathcal{U}'_\lambda \tilde{p} \geq 1, \, \tilde{p} \geq 0\}.$$

The value of the game is then obtained as $v_\lambda = \frac{1}{\|\tilde{p}\|_1}$. In Table I we display the values of the game $\langle \lambda, G \rangle$ for different ϕ's. Note that the minimum $v_\lambda = 1$ in correspondence to $\phi = 1$.

We are now in the position to start with the Monte Carlo simulations. We perform 20 sample paths, and in each one the attainable vector x is randomly chosen with uniform distribution from the interval $[-1, 1]$. At the end of all the 20 sample paths, we compute the sampled average of the variables of interest (integral payoff $\gamma^T(\sigma)$, distance $d(x, \gamma^T(\sigma))$ a.s.o), namely, we average over all the 20 trajectories for varying time T.

To avoid numerical issues, we set the initial payoff $\gamma^0(\sigma) = (0.1, 0.1)$. Each sample path has horizon length τ_1^k with $k = 1, \ldots, 200$, and each interval $[\tau_1^k, \tau_1^{k+1}]$ is subdivided into 10 steps, i.e, the basic length of the interval is $\frac{\tau_1^{k+1} - \tau_1^k}{10}$. In spirit with the proof of Theorem 1, for each period $k = 1, \ldots, 200$, we denote the payoff up-to time τ_1^k by S_k. Then, the strategy of player 1 $\sigma_1(t)$ is the optimal strategy in the game $\langle x - S_k, G \rangle$, for $\tau_1^k \leq t < \tau_1^{k+1}$. More specifically, σ_1 is constant-wise over the intervals $[\tau_1^k, \tau_1^{k+1})$, $k = 1, \ldots, 200$. From a computational standpoint, σ_1 is obtained solving the aforementioned linear program (LP) with $\lambda = x - S_k$.

Without loss of generality, we also take $\sigma_2(t) = (1/4, 1/4, 1/4, 1/4)$ namely we assume that player 2 simply randomizes over $a_2 = 1, \ldots, 4$ with uniform distribution.

All simulations are carried out with MATLAB on an Intel(R) Core(TM)2 Duo CPU P8400 at 2.27 GHz and a 3GB of RAM. The run time of the only Monte Carlo simulations is about 60 seconds. The results of the simulations are summarized in Figs. 2 to 5.

Figure 2 shows the time plot of sampled average of $x - \gamma^T(\sigma)$. In accordance to Theorem 1, any x is strongly attainable and therefore we do expect the trajectories converge to zero as in fact depicted in the figure.

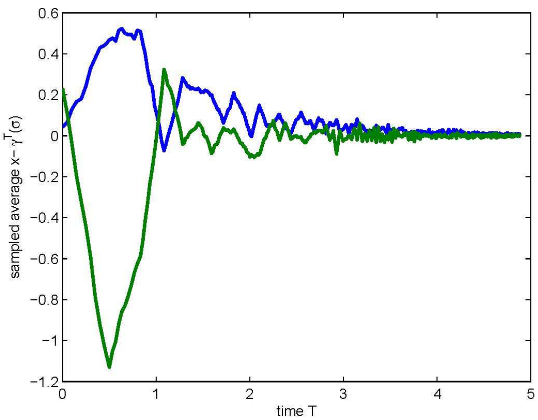

Fig. 2. Sampled average of $x - \gamma^T(\sigma)$ vs. time T.

For the same simulations, Fig. 3 shows the time plot of sampled average of the distance $d(\gamma^T(\sigma), x)$. Again, from Theorem 1, strongly attainability of x implies that the distance tend to zero as illustrated in the plot.

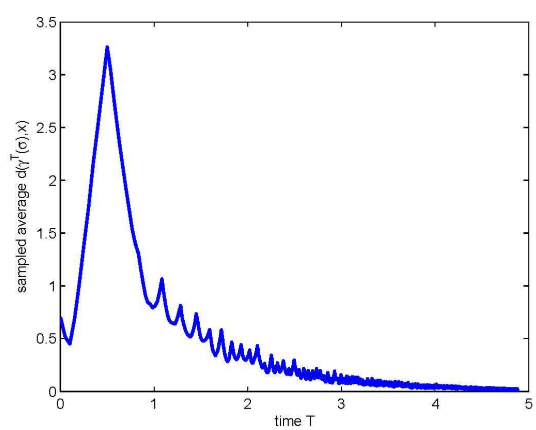

Fig. 3. Sampled average of distance $d(\gamma^T(\sigma), x)$ vs. time T.

Looking at a single path out of the 20 simulations, we plot in Fig. 4 the trajectory $x - \gamma^T(\sigma)$ for varying time T where we have set as attainable vector $x = (0.1, 0.8)$. The plot shows that the trajectory of the integral payoff $\gamma^T(\sigma)$ converges to the prescribed attainable vector x.

Finally, Fig. 5 illustrates the distance $d(\gamma^T(\sigma), x)$ for varying steps N (rather than time T) in the same single simulation mentioned above. The plot shows the converging nature of the trajectory as a function of $\frac{1}{\sqrt{N}}$.

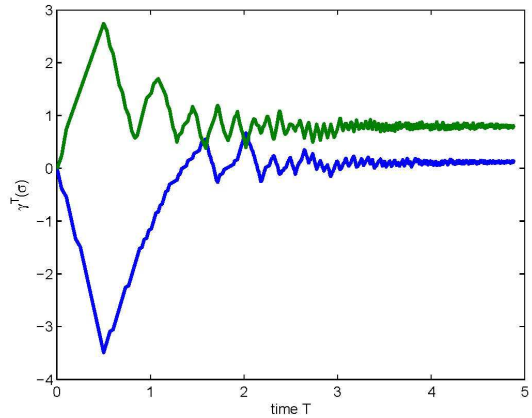

Fig. 4. $\gamma^T(\sigma)$ vs. time T with $x = (0.1, 0.8)$.

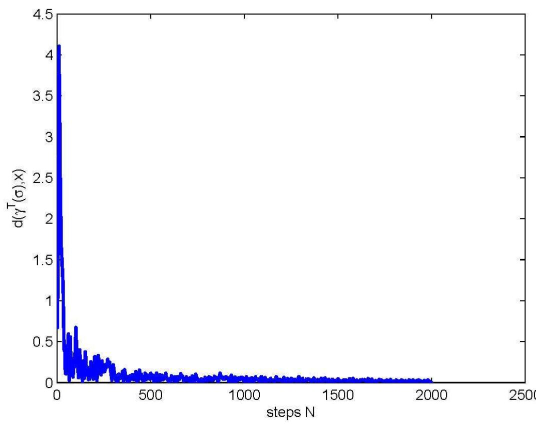

Fig. 5. Distance $d(\gamma^T(\sigma), x)$ vs. steps N with $x = (0.1, 0.8)$.

V. Conclusions

We have introduced the concept of strongly attainable sets of payoffs in two-player repeated games with vector payoffs in continuous time. In particular, a set of payoffs is called *strongly attainable* if player 1 has a strategy that guarantees, even in the worst case, that the distance between the set and the cumulative payoff tends to zero.

As main result we study conditions under which any vector is strongly attainable. Results are illustrated in the context of multi-inventory applications.

Acknowledgment

This research was supported in part by the Google Inter-university center for Electronic Markets and Auctions. Lehrer acknowledges the support of the Israel Science Foundation, Grant #762/045. Solan acknowledges the support of the Israel Science Foundation, Grant #212/09.

References

[1] D. Bauso, F. Blanchini, and R. Pesenti, "Optimization of long-run average-flow cost in networks with time-varying unknown demand," *IEEE Trans. on Automatic Control*, vol. 55, no. 1, pp. 20–31, 2010.

[2] D. Blackwell, “An analog of the minimax theorem for vector payoffs,” *Pacific J. Math.*, vol. 6, no. 1, pp. 1–8, 1956.

[3] F. Blanchini, S. Miani, and W. Ukovich, “Control of production-distribution systems with unknown inputs and system failures,” *IEEE Trans. on Automatic Control*, vol. 45, no. 6, pp. 1072–1081, 2000.

[4] F. Blanchini, F. Rinaldi, and W. Ukovich, “Least inventory control of multi-storage systems with non-stochastic unknown input,” *IEEE Trans. on Robotics and Automation*, vol. 13, no. 5, pp. 633–645, 1997.

[5] E. Lehrer, “Approachability in infinite dimensional spaces,” *International Journal of game Theory*, vol. 31, no. 2, pp. 253–268, 2002.