JOTA manuscript No.
(will be inserted by the editor)

Robust Sub-Optimality of Linear-Saturated Control via Quadratic Zero-Sum Differential Games

Dario Bauso · Rosario Maggistro ·
Raffaele Pesenti

Communicated by Dean A. Carlson

Received: date / Accepted: date

Abstract In this paper, we determine the approximation ratio of a linear-saturated control policy of a typical robust-stabilization problem. We consider a system, whose state integrates the discrepancy between the unknown but bounded disturbance and control. The control aims at keeping the state within

Dario Bauso
University of Groningen
Groningen, Netherland
d.bauso@rug.nl

Rosario Maggistro, Corresponding author
Università Ca' Foscari Venezia
Venezia, Italy
rosario.maggistro@unive.it

Raffaele Pesenti
Università Ca' Foscari Venezia
Venezia, Italy
pesenti@unive.it

a target set, whereas, the disturbance aims at pushing the state outside of the target set by opposing the control action. The literature often solves this kind of problems via a linear-saturated control policy. We show how this policy is an approximation for the optimal control policy by reframing the problem in the context of quadratic zero-sum differential games. We prove that the considered approximation ratio is asymptotically bounded by 2, and it is upper bounded by 2 in the case of 1-dimensional system. In this last case, we also discuss how the approximation ratio may apparently change, when the system's demand is subject to uncertainty. In conclusion, we compare the approximation ratio of the linear-saturated policy with the one of a family of control policies which generalize the bang-bang one.

Keywords Robust optimization · Bounded disturbances · Differential games · Linear-saturated control

Mathematics Subject Classification (2000) 93D21 · 49N70 · 91A05

1 Introduction

Linear-saturated controls (LSC) are inexpensive to design, to implement and to manage. For this reason, a decision maker may wonder whether they are worth implementing even when they are sub-optimal solution for the problem at hand. The aim of this paper is to support the decision makers by comparing LSC policies with respect to optimal ones in the case of robustly ε-stabilizing a linear system with quadratic costs over a finite horizon.

The stability analysis and the stabilization of linear systems with saturating controls are two main topics considered by the control research community (see, e.g., [1–3]). Several approaches were also proposed, e.g., Lyapunov and LMI based approaches [4–8], and Riccati based approaches in [9, 10]. In particular, we refer the reader to, e.g., [11–13] for systems of integrators as the ones considered in this paper and their generalization, even to unstable plants.

A LSC policy can be also viewed as the simplest form of piece-wise linear control policy [14]; when it assumes values on all the state space of the considered problem, it degenerates into the linear control policy, which is robust optimal for quadratic zero-sum games [15].

The contribution of this paper consists in assessing the approximation ratio of the LSC policy, i.e., the ratio between the payoff induced by the LSC policy and the one induced by the optimal control policy for a linear system in presence of quadratic costs on the state and control (while we disregard a possible rectangular term on the same variables). We prove that this approximation ratio tends to 2 when the horizon length tends to infinity.

This paper, in spirit with [16, 17], can be framed in the literature on production-distribution systems, that have to face an unknown but bounded demand. The inventory problem, dealt with in [16], reduces to our problem (1). In [16], a graph describes the topology of a production-distribution system, where the system manager controls the flows of materials over arcs in order to meet the demand at the nodes. The node states, i.e., the deviations from inventory safety stocks, integrates the deviation between the demand and the flow arriving at and departing from the nodes. The objective of the system manager is to keep the inventory fluctuations bounded around the safety stocks. We finally remark that the modeling of the demand as unknown but bounded variable has a long history in control as well as in robust optimization [18].

The rest of the paper is organized as follows. In Section 2, we describe the problem from a game-theoretic point of view. In Section 3, we prove that the approximation ratio of the LSC policy is asymptotically bounded by 2. In Section 4, we focus on the 1-dimensional version of the differential game and we prove that the associated approximation ratio is upper bounded by 2. In addition, we discuss how uncertainty may affect players' assessment of the value of the approximation ratio. In Section 5, we compare the approximation ratio of the LSC policy with the one associated with a family of control policies that generalize the bang-bang one. Finally, in Section 6, we draw some conclusions.

2 Problem Setup

Consider the following ε-*stabilizability* problem [16] over a horizon of length T of a linear systems with quadratic costs on the state and the control but without rectangular terms:

$$J(z_0, u(.), \omega(.)) = \frac{1}{2}\int_0^T (z(t)'z(t) + u(t)'u(t))\, dt + \frac{1}{2} z(T)'z(T), \tag{1a}$$

$$\dot{z}(t) = u(t) - D\omega(t),\ z(0) = z_0, \qquad 0 \le t \le T, \tag{1b}$$

$$u(t) \in \mathcal{U},\ \omega(t) \in \mathcal{W},\ z(t) \in \mathcal{S}, \qquad 0 \le t \le T, \tag{1c}$$

where $z(t) \in \mathbb{R}^m$ is the continuous time state of the system; $u(t) \in \mathbb{R}^m$ is a bounded control; $\omega(t) \in \mathbb{R}^n$ is an unknown but bounded exogenous disturbance with $n \leq m$; $D \in \mathbb{R}^{m \times n}$ is a matrix describing the physical structure of the system; $\mathcal{S} = \{z \in \mathbb{R}^m : -\varepsilon \leq z \leq \varepsilon\}$, $\mathcal{U} = \{u \in \mathbb{R}^m : -\hat{u} \leq u \leq \hat{u}\}$ and $\mathcal{W} = \{\omega \in \mathbb{R}^n : -\hat{\omega} \leq \omega \leq \hat{\omega}\}$ are three hyper-boxes, where $\varepsilon, \hat{u}, \hat{\omega} > 0$ are a-priori chosen. Finally, $u(.) = \{u(t), 0 \leq t \leq T\} \in U$ and $\omega(.) = \{\omega(t), 0 \leq t \leq T\} \in \Omega$ are the realizations of controls and disturbances respectively, being U and Ω the sets of nonanticipative control in the sense of Elliot-Kalton [19].

We address the problem from a game-theoretic standpoint. We consider two players: player 1 (the minimizer), who plays $u(.)$ aiming at minimizing payoff (1a) while keeping the state within the *target set* $\mathcal{S}$; player 2 (the maximizer), who plays $\omega(.)$ aiming at pushing the state out of $\mathcal{S}$ by maximizing payoff (1a). We say that a control policy $u(.)$ is *robust* for player 1 when it can counteract player 2 *worst disturbance* $\omega^*(.) = \arg\max_{\omega(.)} J(z_0, u(.), \omega(.))$, i.e., player 2's best-response to $u(.)$, so that $z(t) \in \mathcal{S}$ for all $0 \leq t \leq T$ and $J(z_0, u(.), \omega^*(.)) < \infty$. A control policy $u^*(.)$ is *robust optimal* for player 1 when it is robust and it is a best-response to player 2's worst disturbance, i.e., $u^*(.) = \arg\min_{u(.)}\{\max_{\omega(.)} J(z_0, u(.), \omega(.))\}$. Finally, policies $(u^*(.), \omega^*(.))$ are *saddle-point policies* of the game when they satisfy the "Isaacs condition" (see, e.g.,[15, p.353]) that defines the *value of the game* $J^*(z_0)$ as:

$$J^*(z_0) = J(z_0, u^*(.), \omega^*(.)) = \min_{u(.)} \max_{\omega(.)} J(z_0, u(.), \omega(.)) = \max_{\omega(.)} \min_{u(.)} J(z_0, u(.), \omega(.)).$$

In [16], it is shown that player 1 robust policies and, in particular, saddle-point polices exist if and only if

$$int(\mathcal{U}) \supset D\mathcal{W}, \tag{2}$$

where $int(\mathcal{U})$ denotes the set of the internal points of set $\mathcal{U}$, and $D\mathcal{W} = \{y \in \mathbb{R}^m : \exists \omega \in \mathcal{W}\ s.t.\ y = D\omega\}$. Condition (2) allows player 1's controls to dominate any possible realization of the disturbance at each time instant. Formally, for each $\omega \in \mathcal{W}$, there exists $\xi > 0$ such that $u = -(1+\bar{\xi})D\omega \in \mathcal{U}$, for all $0 \leq \bar{\xi} \leq \xi$.

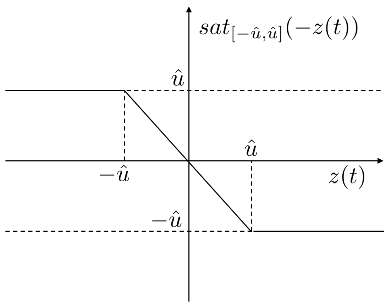

Fig. 1 Linear-saturated control.

In particular, the *linear-saturated control policy*

$$u(t) = \text{sat}_{[-\hat{u},\hat{u}]}(-z(t)) := \big(\text{sat}_{[-\hat{u}_1,\hat{u}_1]}(-z_1(t)), \ldots, \text{sat}_{[-\hat{u}_m,\hat{u}_m]}(-z_m(t))\big) \tag{3}$$

where
$$\text{sat}_{[-\hat{u}_i,\hat{u}_i]}(-z_i) = \begin{cases} -\hat{u}_i, & \text{if } \ z_i > \hat{u}_i, \\ -z_i, & \text{if } \ -\hat{u}_i \leq z_i \leq \hat{u}_i, \\ \hat{u}_i, & \text{if } \ z_i < -\hat{u}_i, \end{cases}$$

(see Fig.1), can keep the state z within $\mathcal{S}$ only if $\varepsilon \geq \max_{\omega \in \mathcal{W}}\{D\omega\}$ and $\hat{u} > \max_{\omega \in \mathcal{W}}\{D\omega\}$, as it can be seen by integrating (1b) and imposing the constraint $z(t) \in \mathcal{S}$. In the following, we assume that condition (2) holds.

Let us denote $J^{LSC}(z_0) = \max_{\omega(.)} J(z_0, \text{sat}_{[-\hat{u},\hat{u}]}(z(.)), \omega(.))$ as the value of (1a) when player 1 is constrained to use the LSC policy (3) and player 2 responds with its best-response. In this paper, we interested in determining an upper bound of the approximation ratio:

$$r = \sup_{z_0} \frac{J^{LSC}(z_0)}{J^*(z_0)}. \tag{4}$$

The ratio r allows to understand how accurately a LSC policy of player 1 can approximate the optimal solution.

3 Asymptotic bound for r

In this section, we prove that an LSC policy is 2-approximating for the differential game (1), at least asymptotically, i.e., we show that 2 is a bound of the approximation ratio (4) when $T \to \infty$.

Initially, let us observe that both the LSC policy $u(z(.)) = \text{sat}_{[-\hat{u},\hat{u}]}(-z(.))$ for player 1 and a possible constant policy $\omega(t) = const$ for player 2 are nonanticipative. Let us also introduce the following notations: $D_{i\cdot}$ denotes the ith row of matrix D; $O_\alpha : \mathbb{R} \to \mathbb{R}^+$ is, for each $\alpha \in \mathbb{N}$, a function such that $\lim_{T\to\infty} \frac{O_\alpha(T)}{T} = 0$; sign:$\mathbb{R} \to \{1, -1\}$ is a function such that $\text{sign}(x) = 1$ if $x \geq 0$ and $\text{sign}(x) = -1$ if $x < 0$.

Finally, let us anticipate some results from next section that are necessary for the proof of the next Theorem 3.1. Lemma 4.1 and Lemma 4.2 in Section 4 consider the 1-dimensional differential game (i.e., $m = n = 1$). They prove that the best-response $\omega^*(.)$ for player 2 is constantly equal to either $\hat{\omega}$ or $-\hat{\omega}$

for $0 \leq t \leq T$ and it is chosen so that $\text{sign}(-z_0 D\omega^*(t)) = 1$, independently from the fact that player 1 implements its best-response $u^*(.)$ (Lemma 4.1) or the LSC policy (Lemma 4.2). The values of the payoff are respectively $J^*(z_0) = \frac{1}{2}T(D\hat{\omega})^2 + O_1(T)$ and $J^{LSC}(z_0) = T(D\hat{\omega})^2 + O_2(T)$.

Theorem 3.1 *Given the differential game* (1)*, the approximation ratio of the LSC policy* (3) *is asymptotically bounded by* 2*, namely,* $r \leq 2$*, when* $T \to \infty$*.*

Proof We prove the result by first determining an upper bound for $J^{LSC}(z_0)$, a lower bound for $J^*(z_0)$, and then determining the ratio of the two bounds. To this end, consider the generic ith component of the state dynamic (1b)

$$\dot{z}_i(t) = u_i(t) - D_{i\cdot}\omega(t), \qquad z(0) = z_{0i}, \qquad i = 1, \ldots, m. \tag{5}$$

To determine an upper bound for $J^{LSC}(z_0)$, we assume that, for each $i = 1, \ldots, m$, player 1 plays the LSC policy, whereas player 2 is allowed to not only fix the value of $\omega(t)$ but also switch its sign, that is, player 2 is free to decide whether to consider dynamics $\dot{z}_i(t) = u_i(t) - D_{i\cdot}\omega(t)$ or $\dot{z}_i(t) = u_i(t) - D_{i\cdot}(-\omega(t))$. The results of Lemma 4.2, and their trivial generalizations when player 2 considers the second dynamics, imply that player 2's best-response $\omega^*(t)$ has components $\omega_j^*(t)$, $j = 1, \ldots, n$, constantly equal to either $\hat{\omega}_j$ or $-\hat{\omega}_j$ for $0 \leq t \leq T$ which are chosen such that $\text{sign}(-z_{0i}D_{i\cdot}\omega^*(t)) = 1$ when player 2 considers the first dynamics and $\text{sign}(-z_{0i}D_{i\cdot}(-\omega^*(t))) = 1$ when player 2 considers the second dynamics, for each $i = 1, \ldots m$. In other words, we can consider the differential game as composed by m different 1-dimensional differential games each one with a payoff $T(D_{i\cdot}\omega)^2 + O_{2i}(T)$, for some $\omega \in \mathcal{W}$

common to all the 1-dimensional differential games. Thus, we can conclude that

$$J^{LSC}(z_0) \leq \max_{\omega\in\mathcal{W}} \sum_{i=1}^{m} T(D_{i\cdot}\omega)^2 + \sum_{i=1}^{m} O_{2i}(T).$$

Now, let $\tilde{\omega} \in \mathbb{R}^n$ be the vector $\tilde{\omega} = \arg\max_{\omega\in\mathcal{W}} \sum_{i=1}^{m}(D_{i\cdot}\omega)^2$, then we can obtain a lower bound for $J^*(z_0)$ by assuming that the set of controls for player 2 is reduced to the set $\tilde{\mathcal{W}} = \{\tilde{\omega}\} \subset \mathcal{W}$. Then player 2 can only play $\omega(t) = \tilde{\omega}$ for $0 \leq t \leq T$, while player 1 plays its best-response. Consider the set I of the indexes $i = 1, \dots, m$ such that $\text{sign}(-z_{0i}D_{i\cdot}\tilde{\omega}) = 1$. For each $i \in I$, the results of Lemma 4.1 allow to consider a 1-dimensional differential game with a payoff $\frac{1}{2}T(D_{i\cdot}\tilde{\omega})^2 + O_{1i}(T)$.

Consider now the set $\tilde{I}$ of the indexes $i = 1, \dots, m$ such that $\text{sign}(-z_{0i}D_{i\cdot}\tilde{\omega}) = -1$. For each $i \in \tilde{I}$, the disturbance $\tilde{\omega}$ initially pushes the state z_i toward 0, while player 1 has no interest in counteracting this action given the payoff (1a) structure. Hence, there exists a time instant $0 \leq \tilde{t}_i \leq \frac{z_{0i}}{D_{i\cdot}\tilde{\omega}}$ such that $z_i(\tilde{t}_i) = 0$. Then, for every $t \geq \tilde{t}_i$ we can consider the 1-dimensional differential game whose initial state is $z_i(\tilde{t}_i)$ and to which we can apply the results of Lemma 4.1 to obtain a payoff $\frac{1}{2}(T - \tilde{t}_i)(D_{i\cdot}\tilde{\omega})^2 + O_{1i}(T - \tilde{t}_i)$.

We set $\tilde{t}_i = 0$ for $i \in I$ to define $\tilde{t} = \max_{i\in I\cup\tilde{I}}\{\tilde{t}_i\}$. Thus, if $T \geq \tilde{t}$, we have

$$J^*(z_0) \geq \frac{1}{2}\sum_{i=1}^{m}(T - \tilde{t}_i)(D_{i\cdot}\tilde{\omega})^2 + \sum_{i=1}^{m} O_{1i}(T - \tilde{t}_i) \geq$$
$$\geq \frac{1}{2}\max_{\omega\in\mathcal{W}}\sum_{i=1}^{m}(T - \tilde{t})(D_{i\cdot}\omega)^2 + \sum_{i=1}^{m} O_{1i}(T - \tilde{t}_i).$$

Finally,

$$r = \frac{J^{LSC}(z_0)}{J^*(z_0)} \leq \frac{\max_{\omega\in\mathcal{W}}\sum_{i=1}^{m} T(D_{i\cdot}\omega)^2 + \sum_{i=1}^{m} O_{2i}(T)}{\frac{1}{2}\max_{\omega\in\mathcal{W}}\sum_{i=1}^{m}(T - \tilde{t})(D_{i\cdot}\omega)^2 + \sum_{i=1}^{m} O_{1i}(T - \tilde{t}_i)} \xrightarrow[T\to\infty]{} 2,$$

then the theorem is proven. □

4 1-Dimensional Differential Game

In this section, we consider the 1-dimensional version of the differential game (1) and we prove that the approximation ratio of LSC policy is upper bounded by 2 for all $T \geq 0$. Hereinafter, without loss of generality, we assume $D = 1$, and hence $\hat{u} \geq \varepsilon > \hat{\omega}$ (as we can always redefine $\hat{\omega}$ by multiplying its original value by an opportune parameter), and $\hat{\omega} = \hat{u} - \delta$, being δ a positive parameter that makes constraints on u and ω equivalent to condition (2).

We determine the players' best-responses using the Isaacs-Hamilton-Jacobi equation (see, e.g., [15] when $z_0 \geq 0$. A symmetric reasoning applies to the case $z_0 < 0$.

We remark that, if $z_0 \geq 0$, the players' policies we introduce next in Lemma 4.1 and in Lemma 4.2 keep the system state positive for all $t > 0$.

Lemma 4.1 *Given the 1-dimensional differential game* (1)*, if* $z_0 \geq 0$*, a saddle-point policy is*

$$(u^*(.), \omega^*(.)) = \left(sat_{[-\hat{u},\hat{u}]}(-z(t) - \hat{\omega}(1 - e^{t-T})), -\hat{\omega}\right), \quad 0 \leq t \leq T. \tag{6}$$

Furthermore, the exact expression of the value of the game is:

$$J^*(z_0) = \begin{cases} \frac{T}{2}\left(\frac{\delta^2 T^2}{3} + \delta T(\delta - z_0) + (\hat{u}^2 - 2\delta z_0 + z_0^2)\right) + \frac{z_0^2}{2}, & \text{if } z_0 > \hat{u} + \delta T, \\ \frac{T}{2}\hat{\omega}^2 + \frac{(\hat{t}^2\delta^2(2\hat{t}-3)+6\delta z_0 \hat{t}(1-\hat{t})+6\hat{t}z_0^2)-3(\hat{u}-z_0)^2+6(\delta\hat{u}+z_0\hat{\omega})}{12}, & \\ & \text{if } \delta + \hat{\omega}e^{-T} \leq z_0 \leq \hat{u} + \delta T, \\ \frac{T}{2}\hat{\omega}^2 + \frac{\hat{\omega}^2(e^{-2T}-1)+4z_0\hat{\omega}(1-e^{-T})+2z_0^2}{4}, & \text{if } 0 \leq z_0 < \delta + \hat{\omega}e^{-T}, \end{cases}$$

where $\hat{t}$ is the solution of the following equation $z_0 - \delta\hat{t} = \delta + \hat{\omega}e^{-T+\hat{t}}$.

The value of the game in the average converges to $\hat{\omega}^2/2$, namely

$$\lim_{T \longrightarrow +\infty} \frac{J^*(z_0)}{T} = \frac{1}{2}\hat{\omega}^2. \tag{7}$$

Proof See Appendix. □

Note that that for $t \to T$ the argument of $u^*(t)$ in (6) tends to $-z(t)$ that is $u^*(t)$ converges to the LSC policy obtained for the static game. On the other hand, for $T \to +\infty$ we find that $u^*(t) = -z(t) - \hat{\omega}$, namely player 1 still plays $-z(t)$ with an additional term $-\hat{\omega}$ to compensate $\omega^*(t)$.

Let us now assume that player 1 is constrained to use the LSC policy (3). We show that player 2's best-response is a bang-bang control.

Lemma 4.2 *Given the 1-dimensional differential game* (1)*, if $z_0 \geq 0$ and player 1 plays the LSC policy* (3)*, then player 2's best-response is the bang-bang control $\omega^*(z(t)) = -sign(z(t))\hat{\omega}$, for $0 \leq t \leq T$.*

Furthermore, the exact expression of the payoff is:

$$J^{LSC}(z_0) = \begin{cases} \frac{T}{2}\Big(\frac{\delta^2 T^2}{3} + \delta T(\delta - z_0) + (\hat{u}^2 - 2\delta z_0 + z_0^2)\Big) + \frac{z_0^2}{2}, & \text{if } z_0 > \hat{u} + \delta T, \\ T\hat{\omega}^2 + \frac{1}{6\delta}\Big(- 6\hat{\omega}\delta^2 e^{\frac{-\hat{u}-\delta T + z_0}{\delta}} + (2\hat{u}^3 - 9\delta\hat{u}^2 + 12\delta^2\hat{u} - \\ \quad -6\delta^3 - 3\hat{u}^2 z_0 + 12\delta\hat{u}z_0 - 6\delta^2 z_0 + z_0^3)\Big), & \text{if } \hat{u} \leq z_0 \leq \hat{u} + \delta T, \\ T\hat{\omega}^2 + \hat{\omega}(\hat{\omega} - z_0)(e^{-T} - 1) + \frac{z_0^2}{2}, & \text{if } 0 < z_0 < \hat{u}. \end{cases}$$

The value of the payoff in the average converges to $\hat{\omega}^2$, namely

$$\lim_{T \longrightarrow +\infty} \frac{J^{LSC}(z_0)}{T} = \hat{\omega}^2. \tag{8}$$

Proof See Appendix. □

We are now ready to determine the approximation ratio (4).

Theorem 4.1 *Given the 1-dimensional differential game* (1)*, if* $z_0 \geq 0$*, the approximation ratio* (4) *is upper bounded by* 2*, namely,* $r \leq 2$*.*

Proof Throughout this proof, we exploit that $r \leq 2$ if and only if $\Delta(z_0, T) := J^{LSC}(z_0) - 2J^*(z_0) \leq 0$ and that the Lambert W function is nonnegative and increasing for $x \geq 0$, and such that $W(ye^y) = y$ for $y \geq 0$.

Next we consider the following four cases:

i) $z_0 \geq \hat{u} + \delta T$,

ii) $\hat{u} < z_0 < \hat{u} + \delta T$,

iii) $\delta + \hat{\omega} e^{-T} \leq z_0 \leq \hat{u}$,

iv) $0 \leq z_0 < \delta + \hat{\omega} e^{-T}$.

i) If $z_0 \geq \hat{u} + \delta T$, we have $r = 1$ as $J^*(z_0) = J^{LSC}(z_0)$.

ii) If $\hat{u} < z_0 < \hat{u} + \delta T$, we have

$$\Delta(z_0, T) = -\hat{\omega}\delta e^{\frac{-\hat{u}-\delta T+z_0}{\delta}} + \frac{1}{6\delta}\left(2\hat{u}^3 - \delta^3 + 6\delta^2\hat{u} - 6\delta\hat{u}^2 - 3\hat{u}^2 z_0 - z_0^3\right)$$
$$+\frac{\delta^2}{6}\left(12W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) + 9W^2\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) + 2W^3\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)\right).$$

Hence, we first show that $\Delta(z_0, T) \leq 0$ if $\Delta(z_0, \frac{z_0-\hat{u}}{\delta}) \leq 0$, then that $\Delta(z_0, \frac{z_0-\hat{u}}{\delta}) \leq 0$ actually holds. To this end, we observe that

$$\frac{\partial \Delta(z_0, T)}{\partial T} = \delta\left(\hat{\omega} e^{\frac{-\hat{u}-\delta T+z_0}{\delta}} - 2\delta W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) - \delta W^2\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)\right).$$

As the function $W(.)$ has nonnegative arguments, we have

$$\frac{\partial \Delta(z_0, T)}{\partial T} \leq \delta\left(\hat{\omega} e^{\frac{-\hat{u}-\delta T+z_0}{\delta}} - \delta W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)\right).$$

Then, we can prove that $\frac{\partial \Delta(z_0,T)}{\partial T} \leq 0$ by showing that $\frac{\hat{\omega}}{\delta} e^{\frac{-\hat{u}-\delta\ T\ +z_0}{\delta}} \leq W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)$. As function $W(.)$ is increasing when its argument is positive, by applying the inverse function we obtain that the inequality holds if and only if

$$\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\hat{u}-\delta T+z_0}{\delta}}\right)e^{\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\hat{u}-\delta T+z_0}{\delta}}\right)} \leq \frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}} \Leftrightarrow$$

$$e^{\frac{\hat{\omega}}{\delta}e^{\frac{-\hat{u}-\delta T+z_0}{\delta}}} \leq e^{\frac{\hat{\omega}}{\delta}} \Leftrightarrow z_0 \leq \hat{u}+\delta T.$$

As, for $\hat{u} \leq z_0 \leq \hat{u}+\delta T$, the difference $\Delta(z_0,T)$ is decreasing in T. Hence, we have that $\Delta(z_0, \frac{z_0-\hat{u}}{\delta}) \leq 0$ implies $\Delta(z_0,T) \leq 0$ for $T \geq \frac{z_0-\hat{u}}{\delta}$. We also observe that $\frac{\partial\Delta(z_0,\frac{z_0-\hat{u}}{\delta})}{\partial z_0} = -\frac{\hat{u}+z_0^2}{2\delta} \leq 0$. Then the maximum value for $\Delta(z_0, \frac{z_0-\hat{u}}{\delta})$ is attained in $z_0 = \hat{u}$. Since $\Delta(\hat{u},0) = -\frac{\hat{u}^2}{2}$ the statement is proved.

iii) If $\delta + \hat{\omega}e^{-T} \leq z_0 \leq \hat{u}$, we have

$$\Delta(z_0,T) = \hat{\omega}(\hat{\omega}-z_0)e^{-T} + \tfrac{1}{6\delta}\left(6\delta^2\hat{u} - \delta^3 - 2z_0^3 - 3\delta\hat{u}^2 + 3\delta z_0^2 - 6\delta\hat{u}z_0\right) +$$

$$+\tfrac{\delta^2}{6}\left(12W\left(\tfrac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) + 9W^2\left(\tfrac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) + 2W^3\left(\tfrac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)\right),$$

and that

$$\frac{\partial\Delta(z_0,T)}{\partial T} = -\hat{\omega}(\hat{\omega}-z_0)e^{-T} - 2\delta^2 W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right) - \delta^2 W^2\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right).$$

Here again, we show that $\Delta(z_0,T)$ is decreasing in T. As the function $W(.)$ has nonnegative arguments, $\frac{\partial\Delta(z_0,T)}{\partial T}$ is trivially nonpositive if $\delta + \hat{\omega}e^{-T} \leq z_0 \leq \hat{u}$. If $\hat{\omega} \leq z_0 \leq \hat{u}$, we have

$$\frac{\partial\Delta(z_0,T)}{\partial T} \leq \hat{\omega}(z_0-\hat{\omega})e^{-T} - \delta^2 W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right).$$

Then, we prove that $\frac{\partial\Delta(z_0,T)}{\partial T} \leq 0$ by showing that

$$\frac{\hat{\omega}(z_0-\hat{\omega})}{\delta^2}e^{-T} \leq W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right).$$

By applying the $W^{-1}(.)$ we obtain that the inequality holds, if and only if

$$\left(\frac{\hat{\omega}(z_0-\hat{\omega})}{\delta^2}e^{-T}\right)e^{\frac{\hat{\omega}(z_0-\hat{\omega})}{\delta^2}e^{-T}} \leq \frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}} \Leftrightarrow$$
$$\Leftrightarrow \frac{z_0-\hat{\omega}}{\delta}e^{\frac{\hat{\omega}(z_0-\hat{\omega})}{\delta^2}e^{-T}} \leq e^{\frac{z_0-\delta}{\delta}}.$$

In the worst case, i.e., $T=0$, it must hold $\frac{z_0-\hat{\omega}}{\delta}e^{\frac{\hat{\omega}(z_0-\hat{\omega})}{\delta^2}} \leq e^{\frac{z_0-\delta}{\delta}}$ or, equivalently, $\delta^2\ln\left(\frac{z_0-\hat{\omega}}{\delta}\right)+\hat{\omega}(z_0-\hat{\omega}) \leq z_0\delta-\delta^2$. As $\ln(x) \leq x-1$ for $x \geq 0$, we consider the stronger condition $\delta(z_0-\hat{\omega})-\delta^2+\hat{\omega}(z_0-\hat{\omega}) \leq z_0\delta-\delta^2$, that in turn holds if and only if $z_0 \leq \hat{u}$.

As, for $\delta+\hat{\omega}e^{-T} \leq z_0 \leq \hat{u}$, the difference $\Delta(z_0,T)$ is decreasing in T. Hence, we have that $\Delta(z_0,\ln\left(\frac{\hat{\omega}}{z_0-\delta}\right)) \leq 0$ implies $\Delta(z_0,T) \leq 0$ for $T \geq \ln\left(\frac{\hat{\omega}}{z_0-\delta}\right) \geq 0$. Since $\Delta(z_0,\ln\left(\frac{\hat{\omega}}{z_0-\delta}\right)) = -\frac{\hat{u}^2}{2}$ for any z_0 in the considered interval, the statement is proved.

iv) If $0 \leq z_0 < \delta+\hat{\omega}e^{-T}$, we have $\Delta(z_0,T) = -\frac{1}{2}\left((1-e^{-T})\hat{\omega}+z_0\right)^2 \leq 0$, and then the statement is proven. □

We conclude this section observing that 2 is a strict upper bound for the approximation ratio, as from (7) and (8) we derive that $\lim_{T\to+\infty}\frac{J^{LSC}(z_0)}{J^*(z_0)} = 2$.

4.1 The Effect of Uncertainty on the Perceived Efficiency

In this subsection, we briefly discuss the ways in which uncertainty may affect the ways in which players assess the value of the approximation ratio. Specifically, we show that the approximation ratio (4) may appear to change when the player-2 best-response is subject to uncertainty.

We assume that the actual response value of player-2 is perturbed in modulus by at most $\sigma > 0$ with respect to the response that both players predict for the opponent play. Considering the 1-dimensional differential game (1), we note that σ has to satisfy the following condition $0 \leq \delta - \sigma < \delta + \sigma < \hat{u}$ and $\varepsilon \geq \hat{u} - \delta + \sigma$ so that the conditions (2) and (1c) hold.

We consider the opposite cases where the perturbation is equal to σ or $-\sigma$. From Lemma 4.2 we know that player 2's best-response coincides with the bang-bang control respectively. Then, we can assume that player 2 implements the following policy $\omega^* + (z(t)) = -\text{sign}(z(t))(\hat{u} - \delta) - \sigma$ for $0 \leq t \leq T$, and we can denote by $J_\sigma^{LSC}(z_0)$ the payoff of (1), where the subscript indicates the presence of uncertainty σ. We say that the perceived efficiency of the LSC policy is given by the ratio between the actual payoff $J_\sigma^{LSC}(z_0)$, given the presence of the uncertainty, and the value that the players estimate as the optimal payoff $J^*(z_0)$ given that they do not know that the player 2's response is affected by uncertainty σ, that is: $r_\sigma = \sup_{z_0} \dfrac{J_\sigma^{LSC}(z_0)}{J^*(z_0)}$.

Using the same arguments of Lemma 4.2 we get that a strict upper bound of the considered ratio is given by $r_\sigma \leq 2 + \zeta(\sigma)$ with $\zeta(\sigma) = \dfrac{2(2\hat{\omega}\sigma + \sigma^2)}{\hat{\omega}} > 0$. Similarly, when the perturbation is equal to $-\sigma$ we get $r_{-\sigma} \leq 2 + \zeta(-\sigma)$ with $\zeta(-\sigma) = \dfrac{2(2\hat{\omega}(-\sigma) + \sigma^2)}{\hat{\omega}} < 0$.

5 Comparison between different approximation ratios

In this section, we compare the approximation ratio of LSC policy with the one of policies parametrized in the threshold value $k > 0$ and subject to (a state based) switching rule of delay-relay type (RTC) depicted in Fig. 2.

In the following, we consider a 1-dimensional system (1).

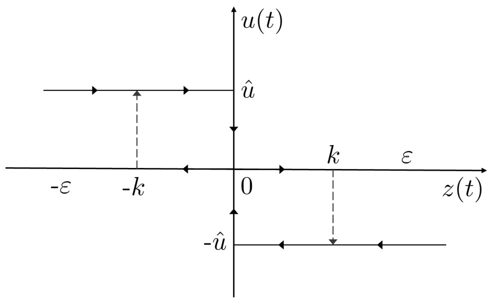

Fig. 2 Switching rule of delay-relay type.

RTC policies are formally described in [20]. They generalize the bang-bang one by allowing hysteresis and they can be described as follows. Assume $0 \leq z_0 \leq k$ and that player 2's policy $\omega^*(.)$ pushes the state to increase its value. Player 1's control $u(.)$ starts to counteract $\omega^*(.)$ jumping from 0 to $-\hat{u}$ when the state goes above the threshold k; $u(.)$ keeps its value equal to $-\hat{u}$ until the state goes below 0 and then jumps to 0 so that the cycle iterates. When $z_0 > k$, $u(.)$ takes on value $-\hat{u}$ from time 0 until the state goes below 0 and then both state and control values cycle as before. Symmetric argument holds if $z_0 \leq 0$ and policy $\omega^*(.)$ pushes the state to decrease its value.
We remark that an RTC policy corresponds to the standard bang-bang policy when $k \to 0$.

Let us now compare the different player 1 policies presented in this paper for a 1-dimensional system (1), with $z_0 = 0$, $\hat{u} = 1$ and $D = 1$. The player 2's best-response is always $\omega^*(t) = -\hat{\omega}$ independently of the policy used by player 1. In particular, if player 1 implements an RTC policy we have that the system dynamics (1b) evolves as follows:

i) $\dot{z}(t) = \hat{\omega}$, which implies $z(t) = \hat{\omega}t$ for $0 \leq t \leq t_1$, where $t_1 = k/\hat{\omega}$ is the switching time when the state goes above the threshold k for the first time,

ii) $\dot{z}(t) = -\hat{u} + \hat{\omega}$, which implies $z(t) = (-\hat{u} + \hat{\omega})t + k\hat{u}/\hat{\omega}$ for $t_1 < t \leq t_2$, where $t_2 = k\hat{u}/\hat{\omega}(\hat{u} - \hat{\omega})$ is the switching time when the state goes below 0 for the first time after t_1,

iii) the system state keeps switching between the two dynamics $\dot{z}(t) = \hat{\omega}$ and $\dot{z}(t) = -\hat{u} + \hat{\omega}$ with frequency $1/t_2 = \hat{\omega}(\hat{u} - \hat{\omega})/k\hat{u}$.

If $\beta t_2 = T$, $\beta \in \mathbb{N}$, and the exit cost is null, the payoff (1a) turns out to be:

$$J^{RTC}(z_0) = \beta \frac{1}{2} \left(\int_0^{t_1} (\hat{\omega}t)^2 \, dt + \int_{t_1}^{t_2} \left((-\hat{u} + \hat{\omega})t + \frac{k\hat{u}}{\hat{\omega}} \right)^2 + \hat{u}^2 \, dt \right) = T \frac{(k^2 + 3\hat{u}\hat{w})}{6}$$

where β is the number of switchings of control $u(.)$ between value 0 and $-\hat{u}$.

Table 1 allows to compare the approximation ratios of LSC and RTC policies for different values of $\hat{\omega}$, k and β. It highlights that the approximation ratios of the LSC policy tends to be better than the one of the RTC policies when $\hat{\omega}$ and β decrease and k increases. The approximation ratios of an RTC policy improves when β increases and $k \to 0$, that is, when player 1 uses a bang-bang control that chatters with a frequency $1/t_2 \to \infty$. Unfortunately, this kind of controls are often very difficult to implement in the practice. For example, the considered 1-dimensional system could be the model of a pump (player 1) in charge of keeping the level $z(.)$ of a water reservoir around a reference 0 level in presence of a demand $\omega(.)$ (player 2). A policy switching continuously the pump on and off would probably quickly damage it.

	$\hat{w}=0.2, k=0.1$		$\hat{w}=0.8, k=0.1$		$\hat{w}=0.2, k=0.9$		$\hat{w}=0.8, k=0.9$	
β	r^{LSC}	r^{RTC}	r^{LSC}	r^{RTC}	r^{LSC}	r^{RTC}	r^{LSC}	r^{RTC}
1	1.19	11.84	1.19	2.92	1.81	12.90	1.81	1.83
10	1.83	5.53	1.83	1.36	1.98	11.86	1.98	1.69
100	1.98	5.12	1.98	1.27	2.00	11.76	2.00	1.67
1000	2.00	5.09	2.00	1.26	2.00	11.75	2.00	1.67

Table 1 Value of the approximation ratios of LSC and RTC policies for different values of $\hat{\omega}$, k and β.

6 Conclusions

In this paper, we proved that the approximation ratio of LSC policy is asymptotically bounded by 2 and compared it with the one of control policies governed by a switching rule. These results provide straightforward performance indicators of the above policies if compared with the optimal ones.

Appendix

Proof of Lemma 4.1

In this proof we show that $(u^*(.), \omega^*(.)) = \big(\text{sat}_{[-\hat{u},\hat{u}]}(-z(t) - \hat{\omega}(1-e^{t-T})), -\hat{\omega}\big)$, for $0 \leq t \leq T$, is a saddle-point policy and we determine the state trajectory $z(t)$.

The values of payoff $J(z_0, u(.), \omega(.))$ (1a) and of limit (7) can be trivially determined by substituting the values of $z(t)$ and $u^*(t)$ in the respective formulas. In particular, we consider the Hamiltonian (9) associated with problem (1) taking into account the constraints on $z(t)$ and $u(t)$ to determine $z(t)$, $u^*(t)$ and $\omega^*(t)$. We have:

$$H(z(t), u(t), \omega(t)) = \frac{1}{2}(z^2(t) + u^2(t)) + p(t)(u(t) - \omega(t)) - \\ - \nu_+(t)(\varepsilon - z(t)) - \nu_-(t)(\varepsilon + z(t)) - \lambda_+(t)(\hat{u} - u(t)) - \lambda_-(t)(\hat{u} + u(t)), \tag{9}$$

where $p(t)$ is the costate and $\nu_+(t), \nu_-(t), \lambda_+(t), \lambda_-(t)$ are four nonnegative functions for $0 \leq t \leq T$, such that $\nu_+(t)(\varepsilon - z(t)) = 0$, $\nu_-(t)(\varepsilon + z(t)) = 0$, $\lambda_+(t)(\hat{u} - u(t)) = 0$, $\lambda_-(t)(\hat{u}+u(t)) = 0$. These last conditions impose that, at each time t, either $\nu_+(t)$ or $\nu_-(t)$, respectively, either $\lambda_+(t)$ or $\lambda_-(t)$ must be equal to 0.

From (9) we obtain that best-responses of the two players are:

$$u^*(t) = \arg\min_{u(t)} H(z(t), u(t), \omega(t)) = -p(t) - \lambda_+(t) + \lambda_-(t) \tag{10a}$$

$$\omega^*(t) = \arg\max_{\omega(t)} H(z(t), u(t), \omega(t)) = -\mathrm{sign}(p(t))\hat{\omega}. \tag{10b}$$

The latter condition implies $p(0) \geq 0$, as we assume $z_0 > 0$, and hence $\omega^*(0)$.

The dynamics on $z(t)$ and $p(t)$ are then

$$\dot{z}(t) = u(t) - \omega(t) = -p(t) - \lambda_+(t) + \lambda_-(t) + \mathrm{sign}(p(t))\hat{\omega}, \text{ with } z(0) = z_0,$$

$$\dot{p}(t) = -z(t) - \nu_+(t) + \nu_-(t), \text{ with } p(T) = z(T).$$

We now prove that there is a saddle-point where $p(t) > 0$ for $0 \leq t \leq T$. To this end, we consider the following two different cases for $p(t)$ and $z(t)$: i) $p(t) \geq \hat{u}$, for $0 \leq t \leq T$, ii) $0 < p(t) < \hat{u}$, for $0 \leq t \leq T$, iii) $p(t) > 0$, for $0 \leq t \leq T$. Specifically, we use the results of the first two cases to address the third general case.

i) If $p(t) \geq \hat{u}$, condition (10a) imposes $\lambda_-(t) = p(t) + u^*(t) + \lambda_+(t) \geq 0$. Then, when $\lambda_-(t) > 0$, we have $u^*(t) = -\hat{u}$ that in turn implies $\lambda_+(t) = 0$. On the other hand, as $\lambda_+(t) \geq 0$ must hold, $\lambda_-(t) = 0$ implies $p(t) = \hat{u}$, $u^*(t) = -\hat{u}$, and $\lambda_+(t) = 0$. As a consequence, straightforward computations indicate the following functions as unique candidate optimal solutions of the differential two–point boundary value problem defined by (10) and (11), for $0 \leq t \leq T$:

$$u^*(t) = -\hat{u}, \qquad \omega^*(t) = -\hat{\omega} \tag{12a}$$

$$z(t) = z_0 - (\hat{u} - \hat{\omega})t = z_0 - \delta t, \qquad p(t) = \frac{1}{2}\delta(t^2 - T(2+T)) + (1 - t + T)z_0. \tag{12b}$$

Note that, given $u^*(t)$ and $\omega^*(t)$, $z(t)$ is decreasing in t, then $z_0 \geq 0$ in $\mathcal{S}$ implies $z(t) \in int\{\mathcal{S}\}$, as $p(t) \geq \hat{u}$ implies $z_0 \geq \hat{u} + \delta T$ and hence $z(t) > 0$ as we show next.

The candidate saddle-point policies (12a) do not contradict the assumptions $p(t) \geq \hat{u}$ for all $0 \leq t \leq T$ if and only if $z_0 \geq \hat{u} + \delta T$. Indeed, we note that $p(T) = z_0 - \delta T \geq \hat{u}$

only if $z_0 \geq \hat{u} + \delta T$ and that function $p(t)$ attains its minimum value in $t^* = \frac{z_0}{\delta}$ which is greater than T for $z_0 \geq \hat{u} + \delta T$.

ii) If $0 \leq p(t) < \hat{u}$, condition (10a) imposes $\lambda_-(t) = 0$, as $\lambda_-(t) > 0$ would imply both $u^*(t) > -\hat{u}$, due to the current assumption on $p(t)$, and $u^*(t) = -\hat{u}$, due to the slackness complementarity condition. Similarly, (10a) imposes also $\lambda_+(t) = 0$. Hence $-\hat{u} < u(t) < 0$ and $\lambda_+(t) = \lambda_-(t) = 0$ for all $0 \leq t \leq T$, then we have

$$u^*(t) = -\hat{\omega}(1 - e^{-T}\cosh(t)) - z_0 e^{-t}, \qquad \omega^*(t) = -\hat{\omega}, \tag{13a}$$

$$z(t) = \hat{\omega}\sinh(t)e^{-T} + z_0 e^{-t}, \qquad p(t) = \hat{\omega}(1 - e^{-T}\cosh(t)) + z_0 e^{-t}. \tag{13b}$$

The candidate saddle-point policies (13a) do not contradict $z(t) \in \mathcal{S}$ and the assumptions $0 \leq p(t) < \hat{u}$ for all $0 \leq t \leq T$ when $0 \leq z_0 < \delta + \hat{\omega}e^{-T}$ as it can be directly verified. In particular, we observe that $z(t)$ is a nonnegative convex function for $t > 0$ if $z_0 \geq 0$. As a consequence, $z(t) \in \mathcal{S}$ if and only if z_0 and $z(T) \in \mathcal{S}$, which in turn requires $\varepsilon \geq \hat{\omega}\frac{1+e^{-T}}{2}$ and which is implied by the fact that $\varepsilon > \hat{\omega}$ holds due to (2).

iii) If $p(t) \geq 0$, we have a more general situation not included in the previous two cases when $\delta + \hat{\omega}e^{-T} \leq z_0 \leq \hat{u} + \delta T$.

We preliminary observe that we can partition the payoff (1) as follows:

$$J(z_0, u(.), \omega(.)) = \underbrace{\int_0^{\hat{t}} \frac{1}{2}(z^2(t) + u^2(t))dt}_{K(z_0, u(.), \omega(.), \hat{t})} + \underbrace{\int_{\hat{t}}^T \frac{1}{2}(z^2(t) + u^2(t))dt + \frac{1}{2}z(T)^2}_{L(z(\hat{t}), u(.), \omega(.), \hat{t})}.$$

for any $0 \leq \hat{t} \leq T$, Then, as equation (1b) describes a first order system, we can determine the value of the payoff (1a) in two steps. First, we compute the value of

$$L^*(z(\hat{t}), \hat{t}) = \min_{u(.)} \max_{\omega(.)} L(z(\hat{t}), u(.), \omega(.), \hat{t})$$

as a function of $z(\hat{t})$. Second, we solve the optimization problem

$$\min_{u(.)} \max_{\omega(.)} \{K(z_0, u(.), \omega(.), \hat{t}) + L^*(z(\hat{t}), \hat{t})\},$$

where $L^*(z(\hat{t}), \hat{t})$ is seen as a final penalty term. In particular, we denote by $\hat{t}$, the first instant, if exists, such that $z_{\hat{t}} := z(\hat{t}) = \delta + \hat{\omega}e^{-T}$.

We obtain the optimal value $L^*(z_{\hat{t}})$ from conditions (13) by translating the time origin in $\hat{t}$, and assuming a horizon length $T-\hat{t}$:

$$L^*(z_{\hat{t}}) = \frac{T-\hat{t}}{2}\hat{\omega}^2 + \frac{(e^{-2(T-\hat{t})}-1)\hat{\omega}^2 + 4z_{\hat{t}}\hat{\omega}(1-e^{-(T-\hat{t})}) + 2z_{\hat{t}}^2}{4}.$$

Next, we solve the optimization problem with respect to $K(z_0, u(t), \omega(t), \hat{t}) + L^*(z_{\hat{t}})$. We observe that the boundary condition is

$$p(\hat{t}) = \left.\frac{\partial L^*(z_{\hat{t}})}{\partial z_{\hat{t}}}\right|_{z_{\hat{t}}=\delta+\hat{\omega}e^{-(T-\hat{t})}} = \hat{\omega} + z_{\hat{t}} - \hat{\omega}e^{-(T-\hat{t})}$$

Now, by contradiction, we show that $u^* = -\hat{u}$ must hold for $t \leq \hat{t}$. Indeed, assume that there exists an instant $\bar{t} \leq \hat{t}$ such that $u^*(\bar{t}) > -\hat{u}$. Then, the latter implies $p(\bar{t}) < \hat{u}$ and being $\dot{p}(t) = -z(t) < 0$ (as $\nu_+(t) = \nu_-(t) = 0$) also $p(\hat{t}) < p(\bar{t}) < \hat{u}$ which contradicts the assumption $p(\hat{t}) = \hat{u}$ and, hence, $\bar{t}$ does not exist. In summary, the saddle-point policies and the associated dynamics on $z(t)$ are:

$$\begin{aligned}
u^*(t) &= \begin{cases} -\hat{u}, & \text{if } t < \hat{t}, \\ -\hat{\omega}(1 - e^{-(T-\hat{t})}\cosh(t-\hat{t})) - z_0 e^{-(t-\hat{t})}, & \text{if } t \geq \hat{t}, \end{cases} \\
\omega^*(t) &= -\hat{\omega}, \\
z(t) &= \begin{cases} z_0 - \delta t, & \text{if } t < \hat{t}, \\ \hat{\omega}\sinh(t-\hat{t})e^{-(T-\hat{t})} + z_0 e^{-(t-\hat{t})}, & \text{if } t \geq \hat{t}. \end{cases}
\end{aligned}$$

Note that even in this case $\varepsilon \geq \hat{\omega}$ implies $z(t) \leq \varepsilon$ for $0 \leq t \leq T$.

As a consequence $\hat{t}$, if exists, is the solution of the following equation $z_0 - \delta\hat{t} = \delta + \hat{\omega}e^{-T+\hat{t}}$ hence $\hat{t} = -1 + \frac{z_0}{\delta} - W\left(\frac{\hat{\omega}}{\delta}e^{\frac{-\delta-\delta T+z_0}{\delta}}\right)$, where $W(.)$ is the Lambert W-function.

It is left to show that $\hat{t}$ always exists (at most we have $\hat{t} = T$). This is straightforward since if $\hat{t}$ did not exist, we would have the lower bound condition $z(t) > \delta + \hat{\omega}e^{-T+t}$ for all $t \leq T$. But the latter is not possible since it must also hold the upper bound condition $z(t) < \hat{u} + \delta(T-t)$, for all $0 \leq t \leq T$, and for $t = T$ we have that both bounds are equal to $\hat{u}$, i.e., $\delta + \hat{\omega}e^{-T+t} = \hat{u} + \delta(T-t) = \hat{u}$. □

Proof of Lemma 4.2

In this proof we determine player 2's best-response, under the assumption that player 1 plays the LSC policy and we determine the state trajectory $z(t)$.

The values of payoff $J(z_0, u(.), \omega(.))$ (1a) and of limit (8) can be trivially determined by substituting the values of $z(t)$ and $u(t)$ in the respective formulas. In particular, we apply the Pontryagin conditions associated with (1) to determine $z(t)$ and $\omega^*(t)$. We consider separately the two cases $0 \leq z_0 \leq \hat{u} \leq \varepsilon$ and $0 \leq z_0 \leq \varepsilon \leq \hat{u}$.

If $0 \leq z_0 \leq \hat{u} \leq \varepsilon$, we have:

$$\begin{aligned}
&\omega^*(t) = -\text{sign}(p(t))\hat{\omega},\\
&\dot{z}(t) = u(z(t)) + \text{sign}(p(t))\hat{\omega},\\
&\dot{p}(t) = -z(t) - u(z(t))\frac{\partial u(z(t))}{\partial z(t)} - p(t)\frac{\partial u(z(t))}{\partial z(t)},\\
&p(T) = z(T).
\end{aligned} \tag{14}$$

To prove that $\omega^*(t) = -\text{sign}(z(t))\hat{\omega}$, initially, consider $0 \leq z_0 \leq \hat{\omega}$, thus $u(z_0) = -z_0$ and observe that possible $\omega^*(.)$ and $z(.)$ components of the solutions of (14) are:

$$\omega^*(t) = -\hat{\omega}, \qquad z(t) = \hat{\omega}(1 - e^{-t}) + z_0 e^{-t} \ > \ 0 \ \ (\text{and} \leq \hat{u}). \tag{15}$$

To prove that $\omega^*(t)$ is actually the solution of (14), we show that any other policy would lead to a worse payoff for player 2. Indeed, suppose that player 2 uses a constant policy $\omega(t) = \hat{\omega}$. We obtain $z(t) = -\hat{\omega}(1-e^{-t})+z_0 e^{-t}$, thus $u(z(t)) = z(t)$ and $z(T) = -\hat{\omega}(1-e^{-T})+z_0 e^{-T}$. Direct computation of the respective values of the payoff shows that policy $\omega(t) = \hat{\omega}$ is worse, respectively not better if $z_0 = 0$, for player 2 than the one of the policy $\omega^*(.)$ in (15). Observe also that, player 2 does not benefit from using a time-varying policy. In this case, we obtain $-\hat{\omega}(1 - e^{-t}) + z_0 e^{-t} \leq z(t) \leq -\hat{\omega}(1 - e^{-t}) + z_0 e^{-t}$, thus $u(z(t)) = z(t)$ and $-\hat{\omega}(1 - e^{-T}) + z_0 e^{-T} \leq z(T) \leq -\hat{\omega}(1 - e^{-T}) + z_0 e^{-T}$. Even in this case the payoff of player 2, due to its quadratic structure, takes on a worse value than the value returned by the policy $\omega^*(.)$ in (15).

It is apparent that if player 2's best-response is $\omega^*(t) = -\hat{\omega}$ when $0 \leq z_0 \leq \hat{\omega}$, then the same control policy is the best-response when z_0 has a greater values. For any other policies, both $u(z(t))$ and $z(t)$ would take on smaller absolute values than in case of $\omega^*(t) = -\hat{\omega}$.

Given the above arguments, state trajectory $z(t)$ associated with $u(.)$ and $\omega^*(.)$ is the defined as $\dot{z}(t) = \tilde{u}(t) + \mathrm{sign}(z(t))\hat{\omega}$. Then,

$$z(t) = \begin{cases} z_0 - \delta t, & \text{if } z_0 \geq \hat{u} + \delta T, \\ (z_0 - \delta t)1\{t \leq \frac{z_0 - \hat{u}}{\delta}\} + (\hat{\omega} + \delta e^{\frac{z_0 - \hat{u}}{\delta} - t}))1\{\frac{z_0 - \hat{u}}{\delta} < t \leq T\}, & \text{if } \hat{u} < z_0 < \hat{u} + \delta T, \\ \hat{\omega}(1 - e^{-t}) + z_0 e^{-t}, & \text{if } 0 \leq z_0 \leq \hat{u}, \end{cases} \cdot$$

Similar arguments hold for $0 < z_0 \leq \varepsilon < \hat{u}$. Even in this case, player 2's best-response is $\tilde{\omega}(t) = -\hat{\omega}$. Also, as $z_0 \leq \varepsilon$, then the state trajectory $z(t)$ associated with $\tilde{u}(.)$ and $\tilde{\omega}(.)$ is $z(t) = \hat{\omega}(1 - e^{-t}) + z_0 e^{-t}$, for $0 \leq t \leq T$. In particular, $\varepsilon \geq \hat{\omega}$ implies $z(t) \leq \varepsilon$. □

References

1. Saberi, A., Lin, Z., Teel, A.: Control of linear systems with saturating actuators. IEEE Trans. Autom. Control 41(3), 368-377 (1996)
2. Tarbouriech, S., Garcia, G., da Silva, J. G., Queinnec, I.: Stability and Stabilization of Linear Systems with Saturating Actuators, London, Springer-Verlag (2011)
3. Benzaouia, A., Mesquine, F., Benhayoun, M.: Saturated Control of Linear Systems. In: Studies in Systems, Decision and Control, Springer, Cham (2018)
4. Hu, T., Teel, A.R., Zaccarian, L.: Stability and performance for saturated systems via quadratic and nonquadratic Lyapunov functions. IEEE Trans. Autom. Control 51(11), 1770-1786 (2006)
5. Henrion, D., Tarbouriech,S.: LMI relaxations for robust stability of linear systems with saturating controls. Automatica 35, 1599-1604 (1999)
6. Benzaouia, A., Tadeo, F., Mesquine, F.: The regulator problem for linear systems with saturation on the control and its increments or rate: an LMI approach. IEEE Trans. Circuits Syst. I Fundam. Theory Appl. 53(12), 2681-2691 (2006)
7. Castelan, E.B., Tarbouriech, S., da Silva Jr., J.M. G., Queinnec, I.: L2-stabilization of continuous-time systems with saturating actuators. Int. J. Robust Nonlinear Control 16, 935-944 (2006)
8. Fang, H., Lin, Z., Hu, T.: Analysis of linear systems in the presence of actuator saturation and L2-disturbances. Automatica 40(7), 1229-1238 (2004)

9. Teel, A.R.: Semi-global stabilizability of linear null controllable systems with input nonlinearities. IEEE Trans. Autom. Control 40(1), 96-100 (1995)
10. Lin, Z., Mantri, R., Saberi, A.: Semi-global output regulation for linear systems subject to input saturation—a low and high gain design. Control Theory Adv. Technol 10(4), 2209-2232 (1995)
11. Teel, A.R.: Global stabilization and restricted tracking for multiple integrators with bounded controls. Syst. Control Lett. 18, 165-171 (1992)
12. da Silva Jr, J. M. G., Tarbouriech, S.: Anti-windup design with guaranteed regions of stability: An LMI-based approach. IEEE Trans. Autom. Control 50, 106-111 (2005)
13. Cao, Y.Y., Lin, Z., Ward, D.G.: An antiwindup approach to enlarging domain of attraction for linear systems subject to actuator saturation. IEEE Trans. Autom. Control 47(1), 140-145 (2002)
14. Bemporad, A., Morari, M., Dua, V., Pistikopoulos, E. N.: The explicit linear quadratic regulator for constrained systems. Automatica 38(1), 3-20 (2002)
15. Basar, T., Olsder, G. J.: Dynamic Noncooperative Game Theory (2nd ed.). Academic Press, London (1995)
16. Bauso, D., Blanchini, F., Pesenti, R.: Robust control policies for multi-inventory systems with average flow constraints. Automatica, 42(8), 1255–1266 (2006)
17. Bagagiolo, F., Bauso, D.: Objective function design for robust optimality of linear control under state-constraints and uncertainty. ESAIM Control Optim. Calc. Var. 17(1), 155-177 (2011)
18. Bertsimas,D., Thiele, A.: A Robust Optimization Approach to Inventory Theory. Oper. Res. 54(1), 150–168 (2006)
19. Elliot, R.J., Kalton, N.J.: The existence of value in differential games. Mem. Amer. Math. Soc. 126. AMS, Providence, USA (1972)
20. Visintin, A.: Differential Models of Hysteresis. Springer-Verlag, Berlin (1994)