INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
*Int. J. Robust Nonlinear Control* **9**, 1117–1141 (1999)

# FAST REFERENCE GOVERNORS FOR SYSTEMS WITH STATE AND CONTROL CONSTRAINTS AND DISTURBANCE INPUTS

ELMER G. GILBERT[1]* AND ILYA KOLMANOVSKY[2]

[1] *Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, U.S.A.*
[2] *P/T Control Systems Department, Ford Research Laboratory, Ford Motor Company, Dearborn, MI 48121-2053, U.S.A.*

## SUMMARY

Reference governors are applied to closed-loop tracking systems that are linear and discrete time and have constraints on state and control variables. Earlier results are extended in significant ways. Disturbance inputs, whose values belong to a specified set, are allowed and a general class of reference governors is introduced. Each governor in the class guarantees constraint satisfaction for all reference and disturbance inputs. Moreover, if the reference input is ultimately confined to a neighbourhood of a constraint-admissible constant input, the eventual action of the reference governor reduces to a unit delay. By appropriately selecting reference governors from the allowed class it is possible to simplify significantly their implementation. The increase in on-line speed of operation overcomes prior limits on the practical application of reference governors. Algorithmic procedures are described which facilitate design of the reference governors. Several examples are presented. They illustrate the design process and the excellence of response to large inputs. Copyright © 1999 John Wiley & Sons, Ltd.

Key words: reference governors; constrained control; disturbance inputs; nonlinear control; discrete time

## 1. INTRODUCTION

This paper addresses the implementation of nonlinear controllers for tracking systems in which the equations of motion are linear but there are pointwise-in-time constraints on control and/or state variables. The purpose of the feedback control is good tracking performance, i.e. a small closed-loop error for a large class of reference commands $r(t)$, and disturbance inputs, $w(t)$. Powerful design techniques are available for synthesizing linear controllers which, ideally, achieve this objective. Suppose now that linear system design objectives have been met and the output of the controller, $y(t)$, is subject to saturation, $u(t) = \text{sat}\, y(t)$, where $\text{sat}\, y = y$ if and only if $y \in Y$, the saturation constraint set. If the reference command is large enough to cause saturation, the effect of the saturation nonlinearity is often unacceptable. The error may diverge or become disproportionately large and persist for a long time.

*This paper was submitted for publication by editor J. Stoustrup*

* Correspondence to Elmer G. Gilbert, Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, U.S.A. E-mail: elmerg@umich.edu

*Received 20 February 1998*
CCC 1049-8923/99/151117–25$17.50 *Revised 28 January 1999*
 *Accepted 30 June 1999*

The traditional way of treating such problems is through design compromises: reducing the performance of the linear controller, putting limits on the reference commands, increasing the size of saturation sets by using more powerful actuators. More recently, the greater effectiveness of nonlinear control techniques has become widely recognized. This led to a large and diverse literature. Major research directions include: state regulators with large domains of attraction, moving horizon optimal control, model predictive control, anti-windup schemes, applications of positively invariant sets. See Reference 1 for an extensive bibliography on systems with saturation.

In this paper we pursue the promising approach to tracking systems originated by Kapasouris *et al.*[2,3]. The key idea is to attenuate the reference command when the possibility of present and future constraint violation occurs. The basic arrangement reduced to its simplest form, is shown in Figure 1. The required attenuation is produced by the reference governor, a nonlinear filter of variable band-width which guarantees $y(t) \in Y$. If the controlled process is the closed-loop system, saturation is avoided and the desirable response properties of the closed-loop linear design are fully maintained. Usually the reference governor is active only for short intervals of time. Thus, for most of the time, the input to the controlled process $v(t)$ closely approximates $r(t)$. In effect, the reference governor is an add-on device whose purpose is to protect the linear system from large excursions of the reference command. A similar device, the error governor,[4,5] which is applied inside the closed-loop system, will not be considered here. While error governors are easier to implement than reference governors, they have serious limitations: constraints on plant state variable cannot be treated, the plant must be stable, and it is difficult to ascertain the response characteristics and stability of the resulting nonlinear closed-loop system.

The tool for predicting the possibility of present and future constraint violation is an appropriate output admissible set, a set of initial states such that $y(t) \in Y$ for all $t \geqslant 0$. In prior implementations of reference governors, it is assumed that these sets are positively invariant. There is a large literature on such sets; see, for example, References 6–17 and references therein. Most of this literature concerns applications of the sets to the implementation of state regulators, an entirely different function than the one used in the implementation of reference governors. The regulation approach can be extended in limited ways to tracking systems. The most common idea is to restrict the reference commands to the class of functions generated by the solution of linear dynamical systems. See References 18 and 16 which consider constant inputs, and Reference 19, which considers general exponential inputs. Reference governors operate effectively without such special assumptions on the class of inputs.

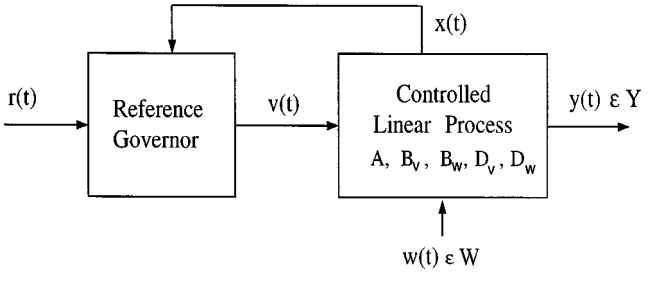


Figure 1. The reference governor and controlled process with constrained output.

 

Practical applications of the governors in References 2, 3 and 5 are limited because the required on-line characterization of the output admissible sets is extremely time consuming. This difficulty is a consequence of their continuous-time nature. Gilbert and Tan[14,20] develop computational tools for discrete-time maximal output admissible sets and use them in the implementation of discrete-time error governors. For the aircraft control problem considered in Reference 5 they obtain computational times which are several orders of magnitude shorter. Subsequently, Gilbert *et al.*[4,21] investigated discrete-time reference governors. Again, they show that practical implementations are possible for systems of moderately high order. Moreover, the theoretical response results are much stronger than those reported in Reference 2.

It is also possible to develop a class of discrete-time reference governors based on a different underlying strategy. See the papers by Bemporad and Mosca.[22–26] They are influenced by ideas from model predictive control where it is assumed that the reference input has a known functional form. Their command generator processes the reference command and produces a parametrically expressed *virtual* reference input. The command generator continuously updates the parameters in such a way that constraint violation is avoided and certain response properties are optimized.

This paper, which has its origins in Reference 27, considers discrete-time reference governors and builds on results of Reference 4 in three significant ways. First, we allow unknown disturbance inputs. The extension depends on the theory of maximal output admissible sets for systems with disturbance inputs. Second, we demonstrate that desirable response properties of reference governors can be maintained when their implementation is based on a family of output admissible sets that are appropriately defined subsets of the maximal output admissible set. The design flexibility afforded by the corresponding *family* of the reference governors leads to simplifications which greatly reduce the number of on-line computations required to implement reference governor operation. Third, we describe algorithmic design procedures which facilitate the implementation of the reference governor.

To be more specific, consider again the system in Figure 1. We assume the controlled process is asymptotically stable and has the form

$$x(t+1) = Ax(t) + B_v v(t) + B_w w(t) \tag{1}$$

$$y(t) = Cx(t) + D_v v(t) + D_w w(t) \tag{2}$$

where $x(t) \in \mathbb{R}^n$, $v(t) \in \mathbb{R}^m$, $w(t) \in \mathbb{R}^l$, $y(t) \in \mathbb{R}^p$. Let $\mathscr{W}$, $\mathscr{Y}$ and $\mathscr{R}$ denote, respectively, the sets of all sequences $w: Z^+ \to W \subset \mathbb{R}^l$, $y: Z^+ \to Y \subset \mathbb{R}^p$ and $r: Z^+ \to \mathbb{R}^m$, where $Z^+$ is the set of non-negative integers. The desired constraints are imposed by requiring $y \in \mathscr{Y}$. Expression (2) is not restricted to controller or actuator limits; general constraints on process state variables can be included. The purpose of the reference governor is to enforce the constraint $y \in \mathscr{Y}$ for all $r \in \mathscr{R}$ and $w \in \mathscr{W}$. Admittedly, the model for the disturbance input is simplistic since it involves only the specification of the set $W$. Although the model leads to more conservative designs when it is contrasted with the stochastic models, it does guarantee protection against the worst possible action of the disturbance. In this sense the model is more typical of those used in game theoretic control problems.[28,29] For the present we avoid mathematical technicalities and concentrate on an informal discussion of the reference governor.

As in Reference 4, the reference governor is a multivariable low pass filter with unit static gain and an adjustable bandwidth parameter $K$, which is a nonlinear function of $r(t)$, its own state and the state of the controlled process:

$$v(t+1) = v(t) + K(r(t), v(t), x(t))(r(t) - v(t)) \tag{3}$$

 

It is required that $K(r(t), v(t), x(t) \in [0, 1]$. This implies that $v(t+1)$ is an interprolation between $r(t)$ and $v(t)$ in the sense that it belongs to the line segment joining $r(t)$ and $v(t)$. Ideally, $K(r(t), v(t), x(t)) = 1$. Then $v(t+1) = r(t)$ and the reference command suffers only a unit delay. When $r(t)$ is large and the possibility of future constraint violation occurs, $K(r(t), v(t), x(t))$ is reduced so that $v(t+1)$ is closer to $v(t)$, which has been chosen on the previous time step to be constraint admissible. If $m > 1$ an alternative scheme, not considered in detail in Reference 4, is to have a separate bandwidth parameter for each component of (3). Then $K(r(t), v(t), x(t))$ becomes a diagonal matrix of $m$ parameters which, subject to the avoidance of future constraint violations, are chosen to minimize a norm of $v(t+1) - r(t)$. Note that $v(t+1)$ is determined entirely by data at time $t$. This is a crucial feature of (3); it allows one sample period for the computation of $K(r(t), v(t), x(t))$.

To put what follows in a clearer perspective, let us sketch the main ideas for the case where $K$ is a scalar and future constraint violation is predicted by a *maximal* output admissible set, $O_\infty$. In words, $O_\infty$ is the set of all initial states which, with $K = 0$, guarantee output constraint satisfaction in (1)–(3). More precisely.

$$O_\infty \triangleq \{(v(0), x(0): y(t) = Cx(t) + D_v v(0) + D_w w(t) \in Y \quad \forall t \in Z^+ \text{ and } \forall w \in \mathscr{W}\} \tag{4}$$

This set is positively disturbance invariant for the system defined by (1)–(3) with $K = 0$, i.e. $(v(t), x(t)) \in O_\infty$ implies $(v(t+1), x(t+1)) \in O_\infty$ for all $w(t) \in W$. The function $K$ is evaluated by

$$K(r, v, x) = \max\{\kappa \in [0, 1]: \begin{bmatrix} v \\ \hline Ax + B_v v + B_w w \end{bmatrix} + \kappa \begin{bmatrix} r - v \\ \hline 0 \end{bmatrix} \in O_\infty, \quad \forall w \in W\} \tag{5}$$

Since the set $O_\infty$ is positively invariant for system (1)–(3) with $K(r(t), v(t), x(t)) \equiv 0$, it follows that $(v, x) \in O_\infty$ implies that the inclusion in (5) is satisfied when $\kappa = 0$. Thus for all $(v(t), x(t)) \in O_\infty$, $K(r(t), v(t), x(t))$ is defined by (5). Further, when $K(r, v, x)$ is used in (1)–(3), it guarantees that $(v(t+1), x(t+1)) \in O_\infty$. Consequently, $(v(0), x(0)) \in O_\infty$, implies $(v(t), x(t)) \in O_\infty$ for all $t \in Z^+$ which in turn implies $y \in \mathscr{Y}$. Since $K(r, v, x)$ is the maximum value of $\kappa$ in (5), $v(t+1)$ is moved as far as possible toward $r(t)$ along the line segment connecting $v(t)$ and $r(t)$. In fact, it will be proved that under reasonable conditions, less stringent than those in Reference 4, that there exists a $\tilde{t} \in Z^+$ such that $K(r(t), v(t), x(t)) = 1$ for all $t \geqslant \tilde{t}$. Thus the ultimate response of the overall system is that of the *linear* controlled process with an added unit delay.

There are two major difficulties in the implementation of the reference governor, both connected with the computation of $K(r, v, x)$. The first problem is addressed in Reference 16. While $O_\infty$ does not, in general, have a finite characterization, it does have a disturbance invariant approximation, $\tilde{O}_\infty$, which is finitely determined and is readily computed. The second problem concerns the complexity of $\tilde{O}_\infty$. It often arises when the reference governor is applied to sampled-data control of a continuous-time system and the sample rate is high: as the sampling rate is increased, $\tilde{O}_\infty$ becomes increasingly complex and the time available for computing $K(r, v, x)$ is decreased. This is where the family of reference governors comes in. By modifying expression (5) in a simple way, it is possible to replace $O_\infty$ by any set $P$ having the property $\tilde{P} \subset P \subset \tilde{O}_\infty$, where $\tilde{P}$ is a set defined by the data that describe the controlled process. Usually, the restriction imposed by $\tilde{P} \subset P \subset \tilde{O}_\infty$ is not severe; in particular, $P$ does not have to be disturbance invariant and it may have a structure that is relatively simple compared to the complex structure of $\tilde{O}_\infty$.

 

We are now in position to summarize the arrangement and contributions of the present paper. Section 2 states overriding assumptions on problem data and introduces a special coordinate system for (1)–(3) which simplifies the subsequent treatment of the disturbance invariant sets. Known properties of $\bar{O}_\infty$, the set in the special co-ordinate system that corresponds to $O_\infty$, and its finitely determined approximation, $\tilde{O}_\infty$, are reviewed in Section 3. In these matters we rely heavily on Reference 16 which summarizes and extends results from the prior literature[8–10,30] on positively disturbance invariant sets. Section 4 contains the main theoretical results. It defines precisely the family of reference governors and states their response properties. In brief $y \in \mathscr{Y}$ for all $r \in \mathscr{R}$ and $w \in \mathscr{W}$ and, if $r(t)$ ultimately enters an appropriate neighbourhood of a constraint-admissible constant input, $K(r(t), v(t), x(t)) = 1$ for all sufficiently large $t$. It is not necessary, as in Reference 4, for $r(t)$ to converge to a constraint admissible value. Section 5 considers, for polyhedral $Y$, computational procedures which lead to functions $K$ that can be implemented economically on-line. The procedures include numerical tests for the inclusions $\tilde{P} \subset P \subset \tilde{O}_\infty$ and algorithmic methods for reducing complexity of $P$. Section 6 contains two examples which illustrate the feasibility of the proposed implementations. Section 7 summarizes the main contributions and compares them with results from the prior literature.

Notations are standard. The $n \times n$ identity matrix is $I_n$; the $n$ by $m$ zero matrix if $O_{n\times m}$; the superscript T denotes vector or matrix transpose. Let $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$, $L \in \mathbb{R}^{m\times n}$, and $X$, $Y \subset \mathbb{R}^n$. Then: $x_i$ is the $i$th component of $x$, $\mathscr{B}^n = \{x\colon \|x\|_\infty \leqslant 1\}$ is the closed unit hypercube, $LX$ is the image of $X$ under $L$, $\alpha X = \{\alpha x\colon x \in X\}$ is the scalar multiple of $X$, $X + Y = \{x + y\colon x \in X, y \in Y\}$ is the Minkowski sum of $X$ and $Y$, $\{x\}$ is the set consisting of the single point $x$, $\{x\} + \alpha\mathscr{B}^n$ is the infinity-norm closed ball of radius $\alpha$ centred at $x$, $coX$ is the convex hull of $X$, int $X$ is the interior of $X$. Suppose $X$ is compact. Then the support function of $X$, $h_X\colon \mathbb{R}^n \to \mathbb{R}$, is defined by $h_X(\eta) = \max_{x\in X} \eta^{\mathrm{T}} x$.

## 2. ASSUMPTIONS AND PRELIMINARIES

We assume that system (1)–(2) and its constraints satisfy the following assumptions: (A1) $A$ is asymptotically stable, (A2) $C$, $A$ is observable, (A3) $W$ and $Y$ are compact and (A4) $0 \in Y$, $0 \in W$. The first assumption is reasonable as (1)–(2) typically models closed-loop systems. The second assumption is not really restrictive since the constraint $y(t) \in Y$, $t \in Z^+$, acts only on the observable co-ordinates of (1)–(2). The third and the fourth assumptions are quite realistic and while not essential everywhere ease the details of the exposition.

The co-ordinate change mentioned in the introduction is achieved by letting

$$x = \bar{B}_v v + \bar{x}, \quad \bar{B}_v = (I_n - A)^{-1} B_v, \quad H_0 = C(I_n - A)^{-1} B_v + D_v \tag{6}$$

Then, (1)–(3) becomes

$$\bar{x}(t+1) = A\bar{x}(t) - \bar{B}_v \bar{K}(r(t), v(t), \bar{x}(t))(r(t) - v(t)) + B_w w(t) \tag{7}$$

$$y(t) = H_0 v(t) + C\bar{x}(t) + D_w w(t) \tag{8}$$

$$v(t+1) = v(t) + \bar{K}(r(t), v(t), \bar{x}(t))(r(t) - v(t)) \tag{9}$$

where $\bar{K}(r, v, \bar{x})$ is a scalar or a diagonal matrix that determines the bandwidth of the reference governor. Hereafter, we focus our attention on (7)–(9), realizing that it is always possible to return to the original co-ordinate system and obtain $K$ from $K(r, v, x) = \bar{K}(r, v, -\bar{B}_v v + x)$. Note that $H_0$ is the static gain from $v$ to $y$ for system (1)–(2).

 

## 3. THE SETS $\bar{O}_\infty$ AND $\tilde{O}_\infty$

The determination of $\bar{K}$ is based on the maximal output admissible set, $\bar{O}_\infty$, associated with the system

$$v(t+1) = v(t) \tag{10}$$

$$\bar{x}(t+1) = A\bar{x}(t) + B_w w(t) \tag{11}$$

$$y(t) = H_0 v(t) + C\bar{x}(t) + D_w w(t) \tag{12}$$

which is obtained by setting $\bar{K}(r(t), v(t), \bar{x}(t)) = 0$ in (7)–(9). The advantage of the special co-ordinate system is now apparent: unlike (1), equation (11) does not contain $v(t)$. The simplifies the treatment of

$$\bar{O}_\infty \triangleq \{(v(0), \bar{x}(0)) \in \mathbb{R}^{m+1}: y \in \mathscr{Y}\ \forall w \in \mathscr{W}\} \tag{13}$$

It is evident that $\bar{O}_\infty$ is positively disturbance invariant, i.e. $(v(0), \bar{x}(0)) \in \bar{O}_\infty$ implies $(v(t), \bar{x}(t)) \in \bar{O}_\infty$ for all $t \in Z^+$ and $w \in \mathscr{W}$.

The remainder of this section reviews known results concerning $\bar{O}_\infty$ and its finitely determined approximation $\bar{O}_\infty$. The presentation follows the pattern of Reference 16 which includes the treatment of systems such as (10)–(12) that are Lyapunov stable but not asymptotically stable. Since complete proofs can be found in this reference, none are given here. To obtain explicit characterization of $\bar{O}_\infty$ and to state concisely the needed results it is necessary to introduce some additional notations and intermediate results.

We begin with a set operation of P-subtraction,[16] or Minkowsky subtraction, as it is sometimes called.[31] The prefix P acknowledges Pontryagin who appears to be the first who introduced the difference in a control application.[32] It has appeared elsewhere as well, often in an implicit rather than explicit from. See, e.g. References 19, 31 and 33–36. Let $U, V \subset \mathbb{R}^p$. Their P-difference is

$$U \sim V \triangleq \{z \in \mathbb{R}^p: z + v \in U\ \forall v \in V\} \tag{14}$$

It is possible that $U \sim V = \emptyset$. Generally, $(U \sim V) + V \neq U$, i.e. P-subtraction is not an additive inverse of the Minkowsky summation. Properties of P-subtraction will be introduced as needed in the sequel. To give key formulas of this section a pragmatic flavour, it is worthwhile to note the following result. Suppose $U$ is the polyhedron

$$U = \{u \in \mathbb{R}^p: \gamma_i^{\mathrm{T}} u \leqslant \beta_i,\ i = 1, \ldots, N\} \tag{15}$$

where $\gamma_i \in \mathbb{R}^p$, $\gamma_i \neq 0$ and $\beta_i \in \mathbb{R}$, $i = 1, \ldots, N$, and $V$ is compact. Then, $U \sim V$ is a polyhedron defined by

$$U \sim V = \{z \in \mathbb{R}^p: \gamma_i^{\mathrm{T}} z \leqslant \beta_i - h_V(\gamma_i),\ i = 1, \ldots, N\} \tag{16}$$

Next, consider $F_t$, the set of states $\bar{x}(t)$ in (11) that are reachable at time $t$, starting at $\bar{x}(0) = 0$. This set is given by

$$F_0 \triangleq \{0\}$$

$$F_t \triangleq \{x(t): x(t) = \sum_{k=0}^{t-1} A^{(t-k-1)} B_w w(k),\ w \in \mathscr{W}\}$$

$$= B_w W + AB_w W + \cdots + A^{t-1} B_w W, \quad t \geqslant 1 \tag{17}$$

 

By (A3) and (A4), $F_t$ is compact and $F_t \subset F_{t+1}$. In fact, the sequence of sets $\{F_t\colon t \in Z^+\}$ has a Hausdorff[28,31] limit.

*Theorem 3.1*

There exists a unique compact set $F$ with the following properties: (i) $F_t \subset F$ for all $t \in Z^+$. (ii) For every $\varepsilon > 0$ there exists $\hat{t} \in Z^+$ such that $F \subset F_t + \varepsilon \mathscr{B}^n$ for all $t > \hat{t}$.

Generally, since it is defined by the Minkowsky sum of infinitely may terms, $F$ does not have an explicit characterization. However, it is possible by various procedures, to obtain super sets of $F$ which approximate $F$. As will be seen in Section 5, such sets can be exploited in numerical computations that involve $F$.

P-subtraction and $F_t$ lead to a concise, recursive characterization of $\bar{O}_\infty$. Define

$$\bar{O}_t \triangleq \{(v(0), x(0)) \in \mathbb{R}^{m+n}\colon y(\tau) \in Y,\ \tau = 0, \ldots, t\ \forall w \in \mathscr{W}\} \tag{18}$$

Then, it is clear that

$$\bar{O}_\infty = \bigcap_{t \in Z+} O_t \tag{19}$$

Moreover, $\bar{O}_t$ has the characterization

$$\begin{aligned} \bar{O}_0 &= \{(v, \bar{x}) \in \mathbb{R}^{m+n}\colon H_o v + C\bar{x} \in Y_0\} \\ \bar{O}_{t+1} &= \bar{O}_t \bigcap \{(v, \bar{x}) \in \mathbb{R}^{m+n}\colon H_0 v + CA^{t+1}\bar{x} \in Y_{t+1}\} \end{aligned} \tag{20}$$

where

$$\begin{aligned} Y_0 &\triangleq Y \sim DW \\ Y_{t+1} &\triangleq Y_t \sim CA^t B_w W \end{aligned} \tag{21}$$

To state the remaining results, two additional sets must be defined:

$$Y_\infty \triangleq \bigcap_{t \in Z+} Y_t, \quad V \triangleq \{v \in \mathbb{R}^m\colon H_0 v \in Y_\infty\} \tag{22}$$

If $Y_\infty \neq \emptyset$, $Y_\infty$ is compact and $V$ is closed. There is a simple, constant input interpretation of these sets. In the disturbance free situation constrained equilibria of (1) are defined by $v(t) = \hat{v}$, $x(0) = \bar{B}_v \hat{v}$ (or $\bar{x}(0) = 0$) and $y(t) \equiv \hat{y} = H_0 \hat{v} \in Y$. If $y \in \mathscr{Y}$ is to be satisfied for all $w \in \mathscr{W}$, and the same constant input and $x(0)$, then $\hat{v}$ must be more tightly constrained; specifically, $V$ is the set of all $\hat{v}$ and $Y_\infty$ is the set of all $\hat{y}$.

*Theorem 3.2*

(i) $\bar{O}_\infty \neq \emptyset$ if and only if $0 \in Y_\infty$. (ii) If $\bar{O}_\infty \neq \emptyset$ it satisfies the inclusions

$$V \times F \subset \bar{O}_\infty \subset V \times \mathbb{R}^n \tag{23}$$

Figure 2 depicts, schematically, the geometric situation described by Theorem 3.2. Shown are the cylinder set $V \times \mathbb{R}^n$ and the set $V \times F$ where $F$ is the set given in Theorem 3.1. Neglect, for the present, the shaded sets; clearly, $\bar{O}_\infty$ satisfies (23) and it is obvious that $\bar{O}_\infty$ must contain points in the boundary of $V \times F$. In the disturbance-free case, the picture is simpler since $V \times F = V \times \{0\}$.

 

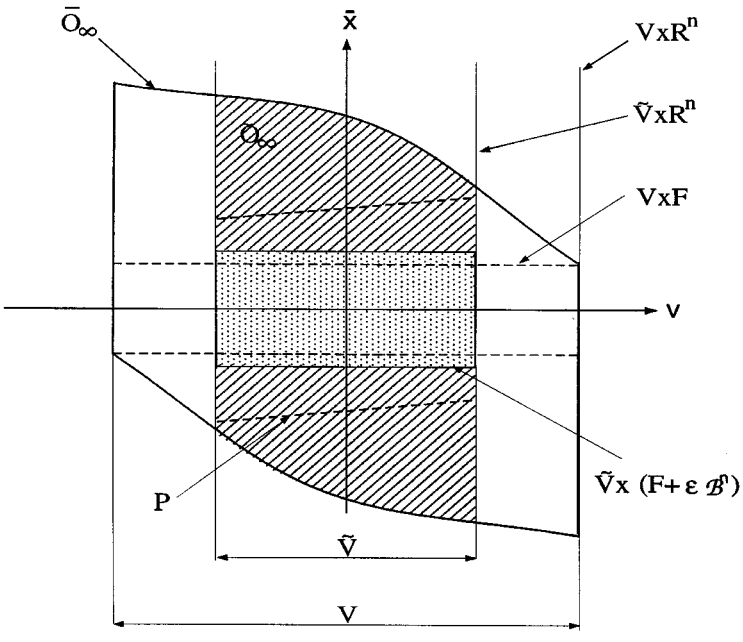


Figure 2. The sets described in Theorems 3.2 and 3.3

With respect to our needs in the next section, $\bar{O}_\infty$ has two deficiencies: the left inclusion in (23) is too weak; the recursive determination of $\bar{O}_\infty$ is generally intractable from a computational point of view. By weak we mean that $V \times (F + \varepsilon\mathscr{B}^n) \subset \bar{O}_\infty$ is not satisfied for some $\varepsilon > 0$. By intractable we mean it is generally true that there exists no $t^* \in Z^+$ such that $\bar{O}_\infty = \bar{O}_{t^*}$. As will be seen in the following two theorems we overcome these deficiences by replacing $V$ and $\bar{O}_\infty$ by smaller sets, $\tilde{V}$ and $\tilde{O}_\infty$, which, ideally, approximate $V$ and $\bar{O}_\infty$. Suppose $0 \in \text{int } Y_\infty$ and let $\tilde{Y} \in \mathbb{R}^m$ satisfy the following conditions:

$$\tilde{Y} \subset \text{int } Y_\infty, \quad 0 \in \text{int } \tilde{Y}, \quad \tilde{Y} \text{ is compact} \tag{24}$$

Define

$$\tilde{V} \triangleq \{v \in \mathbb{R}^m \colon H_0 v \in \tilde{Y}\} \tag{25}$$

and

$$\begin{aligned} \tilde{O}_\infty &\triangleq \bar{O}_\infty \bigcap (\tilde{V} \times \mathbb{R}^n) \\ &= \{(v, \bar{x}) \in \mathbb{R}^{m+n} \colon H_0 v \in \tilde{Y},\ H_0 v + CA^t \bar{x} \in Y_t\ \forall t \in Z^+\} \end{aligned} \tag{26}$$

Note that $\tilde{O}_\infty$ is the maximal output admissible set for (10)–(12) with the original constraint $y(t) \in Y$ and an *added* constraint $H_0 v(t) \in \tilde{Y}$. Hence, $\tilde{O}_\infty$ is d-invariant.

*Theorem 3.3*

Suppose (24) is satisfied. Then there exists $\varepsilon > 0$ such that

$$\tilde{P}(\varepsilon) \triangleq \tilde{V} \times (F + \varepsilon\mathscr{B}^n) \subset \tilde{O}_\infty \subset \tilde{V} \times \mathbb{R}^n \tag{27}$$

 

The modifications produced by the introduction of $\tilde{V}$ are shown by the shaded sets in Figure 2. By choosing $\tilde{Y} \approx Y_\infty$, it follows that $\tilde{O}_\infty \approx \bar{O}_\infty$. Since $0 \in \operatorname{int} P(\varepsilon)$, $0 \in \operatorname{int} \tilde{O}_\infty$. It is obvious that both $\bar{O}_\infty$ and $\tilde{O}_\infty$ are unbounded if $\operatorname{rank} H_0 < m$; if $\operatorname{rank} H_0 = m$ they are both compact.

*Theorem 3.4*

The set $\tilde{O}_\infty$ is finitely determined, i.e. there exists $t^* \in Z^+$ such that $\tilde{O}_\infty = \tilde{O}_{t^*}$.

It is shown in Reference 16 that $\tilde{O}_{t+1} = \tilde{O}_t$ implies $t^* = t$. Hence, $\tilde{O}_\infty$ may be obtained by an algorithmic recursion of a form similar to (20) and (21).

*Algorithm 3.1.* Define

$$Q_0 = \{(\bar{y}, \bar{x}) \in \mathbb{R}^{p+n}: \bar{y} \in \tilde{Y}, \bar{y} + C\bar{x} \in Y_0\}$$
$$Q_{t+1} = Q_t \bigcap \{(\bar{y}, \bar{x}) \in \mathbb{R}^{p+n}: \bar{y} + CA^{t+1}\bar{x} \in Y_{t+1}\} \tag{28}$$

Apply (21) and (28) for $t = 0, 1, \ldots$, until $Q_{t+1} = Q_t$. Set $t^* = t$ and

$$\tilde{O}_\infty = \{(v, \bar{x}) \in \mathbb{R}^{m+n}: (H_0 v, \bar{x}) \in Q_{t^*}\} \tag{29}$$

There are advantages in doing the recursions on $Q_t$ rather than $O_t$. The computations are simpler and the set $Q_{t^*}$ is compact. Additional computational issues, including those associated with $\tilde{Y}$, are discussed in Section 5.

## 4. MAIN RESULTS

We are now ready to give precise definitions of $\bar{K}$ and determine response properties of the resulting reference governors.

Let $\varepsilon$ be any positive number such that $\tilde{P}(\varepsilon) \subset \tilde{O}_\infty$. By Theorem 3.3, $\varepsilon$ exists. The definition of $\bar{K}$ is based on any closed set $P$ satisfying

$$\tilde{P}(\varepsilon) \subset P \subset \tilde{O}_\infty \tag{30}$$

Let

$$P_W = P \sim \begin{bmatrix} 0 \\ \hline \bar{B}_w \end{bmatrix} W \tag{31}$$

Since (30) implies $V \times F \subset P$, it follows from (17) and the definition of P-subtraction that $P_W$ is non-empty. Also, by P-subtraction it follows[16] that $P_W$ is closed. Consider the inclusion

$$\begin{bmatrix} v \\ \hline A\bar{x} \end{bmatrix} + \begin{bmatrix} I_m \\ \hline -\bar{B}_v \end{bmatrix} \kappa(r - v) \in P_W \tag{32}$$

The motivation for this condition comes from the system equations (7) and (9). If it holds, $r(t) = r, v(t) = v, \bar{x}(t) = \bar{x}$ and $\bar{K}(r(t), v(t), \bar{x}(t)) = \kappa$, then $(v(t+1), \bar{x}(t+1)) \in P \subset \tilde{O}_\infty$ for all $w(t) \in W$.

We define the scalar bandwidth parameter, $\bar{K}: \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \to [0, 1]$, by

$$\bar{K}(r, v, \bar{x}) = \begin{cases} 0 & \text{if (32) has no solution for } \kappa \in [0, 1] \\ & \text{otherwise} \\ \max\{\kappa: \text{(32) is satisfied, } \kappa \in [0, 1]\} \end{cases} \tag{33}$$

 

Since $P_W$ is closed it follows that the maximum exists. The d-invariance of $\tilde{O}_\infty$ and the possibility that $\kappa = 0$ guarantee that $\bar{K}(r(t), v(t), \bar{x}(t))$ is chosen so that $(v(t), \bar{x}(t)) \in \tilde{O}_\infty$ implies that $(v(t+1), \bar{x}(t+1)) \in \tilde{O}_\infty$. Hence, $(v(0), \bar{x}(0)) \in \tilde{O}_\infty$ implies that $(v(t), \bar{x}(t)) \in \tilde{O}_\infty$ for all $t \in Z^+$. From (13) nd (26) it follows that $y \in \mathscr{Y}$. By maximizing $\kappa$, $v(t+1)$ is made as close as possible to $r(t)$.

The alternative, multiparameter definition of $\bar{K}$, mentioned in Section 1, is implemented a follows. The reasoning is essentially the same as in the preceding paragraph. Let $\kappa$ be a diagonal matrix of $m$ scalar parameters:

$$\kappa = \text{diag}[\kappa_i] \in \mathbb{R}^{m \times m}, \qquad \kappa_i \in [0, 1], \quad i = 1, \ldots, m \tag{34}$$

Define $\bar{K} : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{m \times m}$ by

$$\bar{K}(r, v, \bar{x}) = \begin{cases} 0 \quad \text{if (32), (34) have no solution} \\ \quad \text{otherwise} \\ \text{argmin}_\kappa \{\| v + \kappa(r - v) - r \| : \text{(32), (34) are satisfied}\} \end{cases} \tag{35}$$

Here $\| \cdot \|$ is an arbitrary norm on $\mathbb{R}^m$ and the purpose of (35) is to make $v(t+1)$ as close as possible to $r(t)$. Again, the optimization problem has a solution. Note that the $\kappa_i$, $i = 1, \ldots, m$, which solve it may not be uniquely determined. In such rather rare cases we suppose the ambiguity in (35) is resolved in a consistent manner; for any such resolution our subsequent results remain valid. While (35) offers more freedom than (33) in the determination of $v(t+1)$, it has the disadvantage that the resulting reference governor is more complex to implement and to analyse.

Through the flexibility in the choice of $P$ provided by (30) both definitions of $\bar{K}$ lead to a family of reference governors. The disturbance-free reference governor considered in Reference 4 corresponds to the choice $P = \tilde{O}_\infty$. There are, at first glance, compelling reasons for making this specific choice in both the disturbance-free and disturbance cases. For $(v, \bar{x}) \in \tilde{O}_\infty$, (32) always has a solution, thus eliminating the first alternatives in (33) and (35). Moreover, (32) then allows maximum freedom in the determination of $\bar{K}$. Thus $\bar{K}$ is as large as possible which causes the reference governor to respond as quickly as possible. The principal disadvantage of $P = \tilde{O}_\infty$ has been presented in Section 1. Computational effort may become prohibitive. As will be seen in Section 6, choosing $P$ to be a relatively simple, proper subset of $\tilde{O}_\infty$ can greatly simplify the computation of $\bar{K}$. If, at the same time, $P \approx \tilde{O}_\infty$ the corresponding decreases in $\bar{K}(r(t), v(t), \bar{x}(t))$ are small and the reference governor performs almost as well as with $P = \tilde{O}_\infty$.

We are now ready to state the response properties of the reference governors. The first theorem is an obvious consequence of our preceding discussion.

*Theorem 4.1*

Assume: $\tilde{Y}$ satisfies (24), $P$ is closed and satisfies (30) for some $\varepsilon > 0$, $\bar{K}$ is defined by (33) or (35), and $(v(0), \bar{x}(0)) \in \tilde{O}_\infty$. Then the following results hold for all $r \in \mathscr{R}$ and $w \in \mathscr{W}$: (i) $(v(t), \bar{x}(t)) \in \tilde{O}_\infty$ for all $t \in Z^+$; (ii) $y \in \mathscr{Y}$.

The theorem says nothing about the performance of the reference governor, i.e. the behavior of $\bar{K}(r(t), v(t), \bar{x}(t))$. This is the objective of the next theorem. To state the results a positive number $G$, which depends on $A$ and $\bar{B}_v$, is needed. Let $\tilde{x}(t), t \in Z^+$, be defined by $\tilde{x}(t+1) =$

 

$A\tilde{x}(t) - B_v v(t)$, $\tilde{x}(0) = 0$. The number $G$ is any upper bound on the induced $l_\infty$ norm of the operator which carries $v$ into $\tilde{x}$, i.e. $\tilde{x}(t) \in G\mathscr{B}^n$ for all $t \in Z^+$ and for all $v$: $Z^+ \to \mathscr{B}^m$.

*Theorem 4.2*

Suppose the assumptions of Theorem 4.1 are satisfied. Assume: $\tilde{Y}$ is convex, $\{r_0\} + \varepsilon(4G)^{-1}\mathscr{B}^m \subset \tilde{V}$ and $\|r(t) - r_0\|_\infty \leqslant \varepsilon(4G)^{-1}$ for all $t \geqslant \hat{t}$. Then there exists $\tilde{t} \geqslant \hat{t}$ such that $v(t+1) = r(t)$ for all $t \geqslant \tilde{t}$.

The proof of the theorem is given in the appendix.

*Remark 4.1*

The assumptions for the theorem in Reference 4, stated in the notations of this paper, are stronger: $\bar{K}$ is defined only by (33), $W = \{0\}$, $P = \tilde{O}_\infty$ and $r(t) \to r_0 \in \tilde{V}$ as $t \to \infty$.

*Remark 4.2*

The theorem places no restrictions on the initial behaviour of $r(t)$ and $\hat{t}$ may be as large as desired. However, the finite settling time, $\tilde{t}$, depends on $\hat{t}$ and on the behaviour of $r(t)$ for $t \leqslant \hat{t}$.

*Remark 4.3*

A natural choice for the initial condition of the reference governor is $v(0) = 0$. Then the requirement $(v(0), \bar{x}(0)) \in \tilde{O}_\infty$ is satisfied if and only if $\bar{x}(0)$ belongs to the maximal output admissible set for the controlled process (1)–(2) with $v(t) \equiv 0$. However, there is more flexibility if $v(0)$ is chosen once $\bar{x}(0)$ is known. Then it suffices to have $\bar{x}(0)$ belong to the projection of $\tilde{O}_\infty$ on the $\bar{x}$ co-ordinates: $\bar{x}(0) \in \{\bar{x} \in \mathbb{R}^n : \exists\, v \in \mathbb{R}^m, (v, \bar{x}) \in \tilde{O}_\infty\}$.

*Remark 4.4*

The key assumption is the ultimate entry of $r(t)$ into an $l_\infty$ ball contained in $V$. Since the size of the ball is proportional to $\varepsilon$ and the upper bound on $\varepsilon$ is determined by (30), the result becomes strongest when $P = \tilde{O}_\infty$. In practice, the requirement on $r(t)$ is very conservative; often, the settling time is finite for inputs which remain in $\tilde{V}$ but go well outside $\{r_0\} + \varepsilon(4G)^{-1}\mathscr{B}^m$.

*Remark 4.5*

Convexity of $\tilde{Y}$ is a crucial assumption. Simple examples show that the theorem may fail when $Y_\infty$ and $\tilde{Y}$ are both non-convex. The failures have a simple explanation: if $(v(0), \bar{x}(0))$ is in a component of $\tilde{O}_\infty$ which does not include $r_0$, $v(t)$ can get trapped in the corresponding component of $\tilde{V}$. Fortunately, in most applications $Y$ is convex. Then $Y_\infty$ is convex and a natural convex approximation of $Y_\infty$ exists: $\tilde{Y} = (1 - \sigma)Y_\infty$, $0 < \sigma \ll 1$. If $Y_\infty$ is not convex, reasonable convex choices for $\tilde{Y}$, which satisfy (24), may still exist. In such cases, $\tilde{V}$ and $\tilde{O}_\infty$ may not approximate $V$ and $\bar{O}_\infty$ closely.

*Remark 4.6*

Recall that the set of all constant, constraint admissible inputs to the controlled process is $V$. When $\tilde{V} \subset V$ approximates $V$ the reference governor does not reduce greatly the allowed set of constant inputs, $r(t) \equiv r_0$, for which $v(t)$ ultimately equals $r_0$.

 

*Remark 4.7*

It is also possible to prove negative results. For example, $v(t+1) = r(t)$ must fail for large $t$ when there exists a $\tilde{t}$ such that $r(t) \notin \tilde{V}$ for all $t \geqslant \tilde{t}$.

## 5. COMPUTATIONAL ISSUES

Implementation of the reference governor requires a concrete representation for $P$ and on-line scheme for generating $\bar{K}$. We now address computational aspects of these requirements. It is assumed hereafter that $Y$ is a polytope (bounded convex polyhedron), $Y = \{y \in \mathbb{R}^p: \gamma_i^{\mathrm{T}} y \leqslant 1, i = 1, \ldots, M\}$. This allows us to exploit computational advantages of (16). More general $Y$ can be handled by replacing them by polytope approximations.[16] The following, easily verified, formulas will be used. Suppose $\eta \in \mathbb{R}^n$. Then: $h_{\mathscr{B}^n}(\eta) = \|\eta\|_1$; if $X = \{x \in \mathbb{R}^n: x^{\mathrm{T}} Q^{-1} x \leqslant 1\}$, $Q = Q^{\mathrm{T}} > 0$ then $h_X(\eta) = \sqrt{\eta^{\mathrm{T}} Q \eta}$; $h_{X_1 + X_2}(\eta) = h_{X_1}(\eta) + h_{X_2}(\eta)$; for $X \subset \mathbb{R}^s$ and $L \subset \mathbb{R}^{n \times s}$, $h_{LX}(\eta) = h_X(L^{\mathrm{T}}\eta)$.

We begin with the computation of $\tilde{Y}$ and $\tilde{O}_\infty$. Because of (16) it is natural to represent $\tilde{Y}$ as a polytope of the form, $\tilde{Y} = \{y \in \mathbb{R}^p: \gamma_i^{\mathrm{T}} y \leqslant \sigma_i, i = 1, \ldots, M\}$ where $0 < \sigma_i < 1$, $i = 1, \ldots, M$. This, in turn, causes $\tilde{O}_\infty$ to be a polyhedron $\tilde{O}_\infty = \{z \in \mathbb{R}^{m+n}: \Psi_i^{\mathrm{T}} z \leqslant 1, i = 1, \ldots, N\}$. The remainder of the section concerns computational procedures for: testing $P$ so that it satisfies (30), obtaining $P$ as a simplified polyhedral approximation of $\tilde{O}_\infty$, determining $\bar{K}$ from $P$.

In the disturbance free case things are straightforward: $Y_\infty = Y_t = Y$ and a suitable choice for $\tilde{Y}$ is $\tilde{Y} = \sigma Y$, where $0 < \sigma < 1$ and, ideally, $\sigma \approx 1$. Then $\sigma_i = \sigma$, $i = 1, \ldots, M$. The computation of $\tilde{O}_\infty$ follows the plan of Algorithm 3.1, where the required testing $Q_{t+1} = Q_t$ is done numerically by solving a set of linear programming problems.[14] Generally, each iteration of Algorithm 3.1 generates redundant inequalities that can be eliminated from the description of $Q_t$ without changing $Q_t$. The efficiency of the iterative process is generally enhanced by eliminating these redundant inequalities at each iteration. Further, $\tilde{O}_\infty$ is simplified in that $N$ is then minimized.

Details for the disturbance case are more complex because $Y_\infty$ is the limit of the sequence $\{Y_t: t \in Z^+\}$ and must be approximated. The set $Y_t$ can be computed recursively by (16) and (21) and it is not difficult to obtain constants $0 < \lambda < 1$ and $\mu > 0$, such that $CA^tBW + CA^{t+1}BW + \cdots \subset \mu\lambda^t \mathscr{B}^m$ for all $t \in Z^+$. Thus, $Y_t \sim \mu\lambda^t \mathscr{B}^m \subset Y_t \sim CA^tBW \sim CA^{t+1}BW \sim \cdots = Y_\infty$ and the approximation $Y_\infty \approx Y_t \sim \mu\lambda^t \mathscr{B}^m$ becomes arbitrarily good as $t \to \infty$. Thus, the obvious choice for $\tilde{Y}$ is $\tilde{Y} = Y_t \sim \mu\lambda^t \mathscr{B}^m$, where the $\sigma_i$ are computed recursively using (16) and (21) and the required support functions are evaluated by $h_{CA^tBW}(\gamma_i) = h_W(B^{\mathrm{T}}(A^t)^{\mathrm{T}} C^{\mathrm{T}} \eta)$ and $h_{\mu\lambda^t \mathscr{B}^m} = \mu\lambda^t \|\gamma_i\|_1$. Because the bound provided by $\mu\lambda^t \mathscr{B}^m$ is conservative, $\tilde{Y} \subset \operatorname{int} Y_\infty$. Moreover, if $0 \in int\ Y_\infty$ and $t$ is sufficiently large, $\sigma_i > 0$. Hence, (24) holds if $0 < \sigma_i$, $i = 1, \ldots, M$, a condition that can be tested numerically as $t$ increases. The accuracy of the approximation $\tilde{Y} \approx Y_\infty$ can be judged by the convergent behavior of $\sigma_i$ as $t$ increases. Once $\tilde{Y}$ is defined, Algorithm 3.1 is applied to obtain $\tilde{O}_\infty$, taking care to eliminate redundant inequalities.

To test the inclusions of (30) we suppose that $P = \{z \in \mathbb{R}^{m+n}: \theta_i^{\mathrm{T}} z \leqslant 1, i = 1, \ldots, N_P\}$. This assumption is consistent with the polyhedral representation of $\tilde{V}$ and $\tilde{O}_\infty$ and it makes simple the computation of $\bar{K}$. Non-polyhedral $P$, such as ellipsoids, can be handled too, but the testing of (30) is more complex.

The right inclusion of (30) is satisfied trivially by setting $P = \tilde{O}_\infty$. More generally, it is satisfied if and only if $h_P(\Psi_i) \leqslant 1$ for $i = 1, \ldots, N$. Since $P$ is polyhedral, $h_P(\Psi_i) = \max\{\Psi_i^{\mathrm{T}} z: z \in P\}$ can be evaluated numerically by linear programming. Similarly the left hand side of (30) is satisfied for some $\varepsilon > 0$ if and only if $h_{\tilde{V} \times F}(\theta_i) < 1$ for $i = 1, \ldots, N_P$. These inequalities can be expressed more

 

directly in terms of the support functions of $\tilde{V}$ and $F$. Let $\theta_i^{\mathrm{T}} = [(\theta_i^v)^{\mathrm{T}} \ (\theta_i^x)^{\mathrm{T}}]$ where $\theta_i^v \in \mathbb{R}^m$ and $\theta_i^x \in \mathbb{R}^n$; then $h_{\tilde{V}\times F}(\theta_i) = h_{\tilde{V}}(\theta_i^v) + h_F(\theta_i^x)$. Thus, the left inclusion of (30) is satisfied if and only if

$$h_F(\theta_i^x) < 1 - h_{\tilde{V}}(\theta_i^v) = 1 - \max\{(\theta_j^v)^{\mathrm{T}} v\colon \gamma_j^{\mathrm{T}} H_0 v \leqslant \sigma_j, \quad j = 1, \ldots, M\},$$
$$i = 1, \ldots, N_P \tag{36}$$

While the right-hand side of (36) can be evaluated by linear programming, the left-hand side is a problem because $F$ does not generally have an explicit representation. Moreover, computing $F_t$ as a Minkowski sum is computationally very expensive.[37,38]

We consider two relatively efficient approaches to completing the verification of (36). In the first approach we exploit (17), leting $t \to \infty$, and properties of support functions to obtain: $h_F(\theta_i^x) = \sum_{t=0}^{\infty} h_W((A^t B_w)^{\mathrm{T}} \theta_i^x)$. Each of the terms in this sum is non-negative and can be bounded. Suppose $\rho_W > 0$ is chosen so that $W \subset \rho_W \mathscr{B}^m$; then $h_W((A^t B_w)^{\mathrm{T}} \theta_i^x) \leqslant \rho_W \| (A^t B_w) \theta_i^x \|_1$. Hence, there exist $\rho > 0$ and $0 < \lambda < 1$ such that $h_W((A^t B_w)^{\mathrm{T}} \theta_i^x) \leqslant \rho \lambda^t \| \theta_i^x \|_1$ for all $t \in Z^+$ and $i = 1, \ldots, N_P$. Combining these results shows that (36) is satisfied if

$$\sum_{t=0}^{\tau} h_W((A^t B_w)^{\mathrm{T}} \theta_i^x) < 1 - h_{\tilde{V}}(\theta_i^v) - \rho \lambda^{\tau+1} (1-\lambda)^{-1} \| \theta_i^x \|_1, \quad i = 1, \ldots, N_P \tag{37}$$

The finite sum on the left-hand side is easy to compute while the condition imposed by (37) approaches (36) a $\tau \to \infty$.

In the second approach $F$ is bounded by an ellipsoid of the form $Z = \{x \in \mathbb{R}^n\colon x^{\mathrm{T}} Q^{-1} x \leqslant 1\}$ where $Q = Q^{\mathrm{T}} > 0$. The matrix $Q$ is obtained[16] by solving the Lyapunov-like equation

$$Q - \gamma^{-1} A Q A^{\mathrm{T}} = (1-\gamma)^{-1} B R B^{\mathrm{T}} + S \tag{38}$$

where $S = S^{\mathrm{T}} > 0$, $R = R^{\mathrm{T}} > 0$ is a matrix such that $W \subset \{w \in \mathbb{R}^w\colon w^{\mathrm{T}} R^{-1} w \leqslant 1\}$, and $\gamma$ is an number such that $\mu^2 < \gamma < 1$ where $\mu$ is the spectral radius of $A$. The bound $F \subset \operatorname{int} Z$ becomes sharper as $R$ and $S$ are made smaller. Since $h_F(\eta) \leqslant h_Z(\eta)$ for all $\eta \in \mathbb{R}^n$, it follows that (36) is satisfied if

$$\sqrt{(\theta_i^x)^{\mathrm{T}} Q^{-} \theta_i^x} \leqslant 1 - h_{\tilde{V}}(\theta_i^v), \quad i = 1, \ldots, N_P \tag{39}$$

These inequalities are generally a weaker sufficient condition than (37), particularly if $\tau$ is large. However, they are less expensive computationally than (37), even when several $Q$'s, corresponding to different values of $\gamma$ are used in (38). Certainly, (39) is an effective test when $\tilde{V} \times F$ is well inside $\tilde{O}_\infty$ and $P \approx \tilde{O}_\infty$.

In summary, computational tools for testing (30) are available. They have simple numerical realizations when $Y$ is a polytope, $P$ is a polyhedron, and it is possible to compute the support function of $W$. Of course, testing (30) is not needed if $\sigma_1, \ldots, \sigma_M > 0$ and $P = \tilde{O}_\infty$. This is a good reason for choosing $P = \tilde{O}_\infty$. If a simpler $P$ is desired there are various ways to proceed including heuristic design procedures. The approach which follows is based on the idea that $P$ should approximate $\tilde{O}_\infty$. It can be automated and appears to be quite effective.

It begins with an outer approximation of $\tilde{O}_\infty$, which is obtained by deleting some of the half spaces whose intersection form $\tilde{O}_\infty$. To achieve good accuracy a numerical criterion for evaluating the relative importance of the various half spaces is needed. Suppose $\tilde{O}_\infty = O(\mathscr{I}) = \{(v, \bar{x}) \in \mathbb{R}^{m+n}\colon H_0 v \in \tilde{Y}, \psi_i^{\mathrm{T}}(v, \bar{x}) \leqslant 1, \ i \in \mathscr{I}\}$, where $\psi_i \in \mathbb{R}^{m+n}$ and $i \in \mathscr{I} \subset Z^+$ defines the linear inequalities which arise from $H_0 v + C A^t \bar{x} \in Y_t, \quad t = 0, \ldots, t^*$, in (26). Define $\varepsilon_i(\mathscr{I}) = \max\{\psi_i^{\mathrm{T}}(v, \bar{x})\colon H_0 v \in \tilde{Y}, \psi_j^{\mathrm{T}}(v, \bar{x}) \leqslant 1, j \in \mathscr{I}, j \neq i\}$ — for $i \in \mathscr{I}$. Because $\tilde{Y}$ is a polytope, the

 

$\varepsilon_i(\mathscr{I})$ can be computed by solving linear programming problems. Clearly, $\varepsilon_i(\mathscr{I}) \geqslant 0$ and its magnitude is a measure of the important of the $i$th half space. Half spaces corresponding to small $\varepsilon_i(\mathscr{I})$ are successively eliminated to generate approximations $O(\mathscr{I}_k)$ where $\mathscr{I} \supset \mathscr{I}_1 \supset \mathscr{I}_2 \supset \cdots$. Note that half space corresponding to the linear inequalities in $H_0 v \in \tilde{Y}$ are not eliminated; we do this so that $O(\mathscr{I}_k) \subset \tilde{V} \times \mathbb{R}^n$, an inclusion which is needed to enforce the condition $P \subset \tilde{V} \times \mathbb{R}^n$, contained in (30). There are many possible ways of doing the elimination. For example, $\mathscr{I}_{k+1}$ can be generated by eliminating from $\mathscr{I}_k$ an index corresponding to the smallest value of $\varepsilon_i(\mathscr{I}_k)$. Obviously, this approach is numerically expensive because of the large number of linear programs that must be solved at each stage of the elimination process. A less refined but more efficient approach is to eliminate from $\mathscr{I}_k$ the several indices corresponding to the smallest several values of $\varepsilon_i(\mathscr{I}_k)$. The process of successive elimination is continued until a reasonable compromise between complexity (cardinality of $\mathscr{I}_k$) and accuracy of the approximation is obtained.

Let $O(\hat{\mathscr{I}})$ be the approximation of $\tilde{O}_\infty$. Since $\tilde{O}_\infty \subset O(\hat{\mathscr{I}})$, $O(\hat{\mathscr{I}})$ cannot be used for $P$. It is necessary to shrink $O(\hat{\mathscr{I}})$. The shrinking is done only on the $\bar{x}$ coordinates so that the '$v$ width' of $P$ is maintained at $\tilde{V}$; see Fig. 2. Let the shrinkage factor be $\alpha^{-1}$ where $\alpha > 1$. Then, $P = \{(v, \bar{x}) \in \mathbb{R}^{m+n}: H_0 v \in \tilde{Y},\ \psi_j^{\mathrm{T}}(v, \alpha\bar{x}) \leqslant 1,\ j \in \hat{\mathscr{I}}\}$. The inclusions of (30) are tested by a search procedure on $\alpha > 1$ until they hold for an $\alpha$ which approximates the smallest possible value for $\alpha$. While in the disturbance-free case, there always exists a value of $\alpha > 1$ such that (30) holds, in the disturbance case an admissible value of $\alpha$ may fail to exist because of the need to satisfy the right inclusion of (30). However, the likelihood of success is high if $O(\hat{\mathscr{I}})$ is a good approximation of $\tilde{O}_\infty$. Also a good approximation is likely to produce a value of $\alpha$ close to 1, so that $P$ closely approximates $\tilde{O}_\infty$.

Once $P$ is determined the on-line evaluation of $\bar{K}$ by (33) is relatively simple. Since, $P$ is a polytope, so is $P_W$ and (33) is implemented by examining a set of inequalities which are dependent on $\kappa$. Suppose $P_W = \{(v, \bar{x}) \in \mathbb{R}^{m+n}: \zeta_j^{\mathrm{T}}(v, \bar{x}) \leqslant 1,\ j = 1, \ldots, N_P\}$. Let $\Theta_j \in \mathbb{R}^m$, $\Phi_j \in \mathbb{R}^{m+n}$ be defined as

$$\Theta_j = \left[\begin{array}{c} I_m \\ \hline -B_v \end{array}\right]^{\mathrm{T}} \zeta_j, \quad \Phi_j = \left[\begin{array}{c|c} I_m & 0 \\ \hline 0 & A \end{array}\right]^{\mathrm{T}} \zeta_j \tag{40}$$

Given $r$ and $v$, compute

$$\theta_j = \Theta_j^{\mathrm{T}}(r - v) \in \mathbb{R}, \qquad \phi_j = 1 - \Phi_j^{\mathrm{T}} \left[\begin{array}{c} v \\ \hline \bar{x} \end{array}\right] \in \mathbb{R} \tag{41}$$

Then inclusion (32) holds for some $\kappa \in [0, 1]$ if and only if $\kappa$ satisfies the inequalities: $0 \leqslant \kappa \leqslant 1$, $\kappa\theta_j \leqslant \phi_j$, $j = 1, \ldots, N_P$. These inequalities are expressed equivalently by $\kappa_j^{\mathrm{L}} \leqslant \kappa \leqslant \kappa_j^{\mathrm{U}}$, $j = 1, \ldots, N_P$ where

$$\begin{aligned}
&\phi_j \geqslant 0, \quad \theta_j \leqslant \phi_j \Rightarrow \kappa_j^{\mathrm{L}} = 0, \quad \kappa_j^{\mathrm{U}} = 1 \\
&\phi_j \geqslant 0, \quad \theta_j > \phi_j \Rightarrow \kappa_j^{\mathrm{L}} = 0, \quad \kappa_j^{\mathrm{U}} = \phi_j/\theta_j \\
&\phi_j < 0, \quad \theta_j \leqslant \phi_j \Rightarrow \kappa_j^{\mathrm{L}} = \phi_j/\theta_j, \quad \kappa_j^{\mathrm{U}} = 1 \\
&\phi_j < 0, \quad \theta_j > \phi_j \Rightarrow \kappa_j^{\mathrm{L}} = 1, \quad \kappa_j^{\mathrm{U}} = 0
\end{aligned} \tag{42}$$

 

The implication on the last line is a convenient artifice for starting that (32) has no solution for some $\kappa \in [0, 1]$. By defining

$$\kappa^{\mathrm{L}} = \max\{\kappa_j^{\mathrm{L}}: i = 1, \ldots, N_P\}, \kappa^{\mathrm{U}} = \min\{\kappa_j^{\mathrm{L}}, j = 1, \ldots, N_P\}, \tag{43}$$

it is seen that

$$\begin{aligned} \tilde{K}(r, v, \bar{x}) &= 0, \kappa^{\mathrm{L}} > \kappa^{\mathrm{U}} \\ &= \kappa^{\mathrm{U}}, \kappa^{\mathrm{L}} \leqslant \kappa^{\mathrm{U}} \end{aligned} \tag{44}$$

The computational effort associated with the on-line computations (41)–(44) is easily estimated. The number of flops (multiplies, adds or comparisons) is: $(3m - 1 + 2m + 2n)N_P$ for (41), $3N_P$ for (42), $2N_P$ for (43). This gives a total of $(5m + 2n + 4)N_P$ flops. If the inequalities in the definition of $P_W$ come in symmetric pairs, $\zeta_{j+1} = -\zeta_j$, as is the case when $Y$ is a hypercube, the operations in (41) need not be fully repeated. Only $(5m/2 + n + 4)N_P$ flops are needed for (41)–(43).

When $\bar{K}$ is defined by (35), conditions (32) and (34) are expressed by $\Theta_j^{\mathrm{T}}(\mathrm{diag}[\kappa_i])(r - v) \leqslant 1 - \phi_j$, $j = 1, \ldots, N_P$, and $0 \leqslant \kappa_i \leqslant 1$, $i = 1, \ldots, m$. Thus, (32) and (34) are expressed by $N_P + 2m$ linear inequalities on the $\kappa_i$. If the linear inequalities are inconsistent $\bar{K}(r, v, \bar{x}) = 0$. Otherwise, $\bar{K}(r, v, \bar{x})$ is defined by minimizing $\|v + (\mathrm{diag}[\kappa_i])(r - v) - r\|$ subject to the linear inequalities. The on-line solution of this problem becomes computationally expensive if either $m$ is large or $N_P$ is large.

## 6. EXAMPLES

Simplicity of the first example makes it easier to illustrate the preceding ideas and the roles played by $\tilde{O}_\infty$ and $P$. The continuous-time plant is an inverted pendulum with a dc motor acting at its base. In scaled variables the equation of motion is $\ddot{\theta} - \theta = u$, where $\theta$ is the deflection of the pendulum from the vertical. The motor current, $u$, saturates so that $|u| \leqslant 0{\cdot}1$. The sampled-data controller, which has a sample period of $T = 0{\cdot}1$, uses a zero order hold on $u$ and linear state feedback: $u(t) = -6x_1(t) - 2x_2(t) - 6w(t) + 5v(t)$. Here, $x_1(t) = \theta(tT)$, $x_2(t) = \dot{\theta}(tT)$, $w(t)$ is measurement noise (e.g., quantizing error) and $v$(t) is the reference command. Going through the usual conversion of a continuous-time system to a discrete-time system and substituting the feedback law, yields the description of the controlled process (1) where $y(\mathrm{t}) = u(t)$ and $Y = [-0{\cdot}1, +0{\cdot}1]$. The eigenvalues are 1·1052 and 0·9048 for the open-loop discrete-time plant, and $0{\cdot}8898 \pm 0{\cdot}1947\mathrm{i}$ for the controlled process. In all what follows it is assumed that $x(0) = 0$.

Figure 3 shows step responses of the controlled process when $w(\mathrm{t}) \equiv 0$ and saturation is allowed to occur. To make the responses easier to see, the samples of $x_1(t)$ are joined by straight-line segments. In all cases the amplitude of the step is sufficient to cause saturation, although for $v(t) \equiv r = 0{\cdot}03$ saturation is brief so that the response is close to the response of the linear system without saturation. Significantly larger steps cause instability or a sluggish response with large overshoot.

We consider two implementations of the disturbance-free reference governor, one for $P = \tilde{O}_\infty$ and one for a simplified $P \subset \tilde{O}_\infty$. Since $H_0 = -1$, $V = Y$. We set $\tilde{V} = \tilde{Y} = [-0{\cdot}095, +0{\cdot}095]$ and apply Algorithm 3.1. The result is $\tilde{O}_\infty = \{(v, \bar{x}) \in \mathbb{R}^3: |-v| \leqslant 0{\cdot}095, |-v + CA^t\bar{x}| \leqslant 0{\cdot}1, t = 0, 1, \ldots, 21\}$. Thus, $\tilde{O}_\infty$ is the intersection of 46 half spaces and $t^* = 21$. See the cross-sections in Figure 4. A simplified $P$ is obtained by following the ideas in Section 5. Figure 4 suggests that a fairly accurate outer approximation of $\tilde{O}_\infty$ can be obtained by using only the inequalities for

 

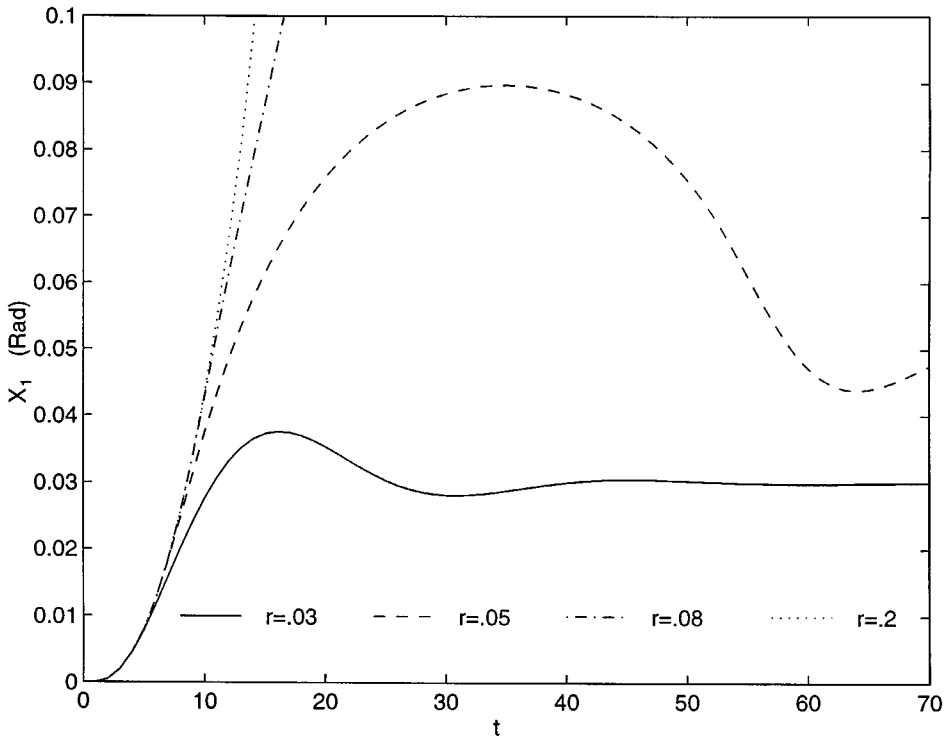


Figure 3. Inverted pendulum example: responses with saturation and without reference governor

$t = 0, 3, 6, \ldots, 21$. This gives a somewhat larger set, consisting of the intersection of 18 half-spaces. We obtain $P$, which is inside $\tilde{O}_\infty$, by shrinking this larger set in its $\bar{x}$ co-ordinates: $P = \{(v, \bar{x}) \in \mathbb{R}^3 : |-v| \leqslant 0{\cdot}095,\ 1 - v + \alpha CA^t\bar{x}| \leqslant 0{\cdot}1,\ t = 0, 3, 6, \ldots, 21\},\ \alpha > 1$. Using linear programming it is possible to determine the smallest value of $\alpha$ which causes $P \subset \tilde{O}_\infty$. This gives $\alpha = 1{\cdot}06$. A somewhat larger value, $\alpha = 1{\cdot}15$, is used in our implementation of (33); this allows us to see that good results can be obtained even if $P$ does not closely approximate $\tilde{O}_\infty$. The number of flops required for evaluation of $K$ is reduced from 391 for $P = \tilde{O}_\infty$ to 153 for the simplified $P$. Note that for all the sets in Figure 4, $\tilde{Y} = \tilde{V}$ remains unchanged so that the extent of the sets in the $v$ coordinate is the same.

Figure 5 shows the responses of the two reference governors for $v(0) = 0$ and the same reference commands which are used in Figure 3. Unlike Figure 3, they are all well behaved and the loss in performance from using the simplified $P$ is small. When $\alpha = 1{\cdot}06$ the difference in responses is even less, with the response curves almost on top of one another. For $r \in \tilde{V}$, the reference governor conforms to the predictions of Theorem 4.2. For instance, with $r = 0{\cdot}08$, $v(t + 1) = r(t)$ for all $t \geqslant 27$. For $r = 0{\cdot}2$, $r \notin \tilde{V}$ and Remark 4.7 applies.

Next, consider the effect of a disturbance where $W = [-0{\cdot}002, +0{\cdot}002]$. Significantly larger disturbance sets cannot be used, because then $\bar{O}_\infty = \emptyset$. For instance, this happens if $W = [-0{\cdot}005, +0{\cdot}005]$. Computations of $Y_t$ for large $t$ show $Y_\infty = (0{\cdot}0595 \ldots)\,[-1, +1]$; our choice for $\tilde{Y} = \tilde{V}$ is $[-0{\cdot}059, +0{\cdot}059]$. The number of half spaces required to express $\tilde{O}_\infty$ is 54 and $t^* = 25$. Without the reference governor, the disturbance can cause significant saturation to

 

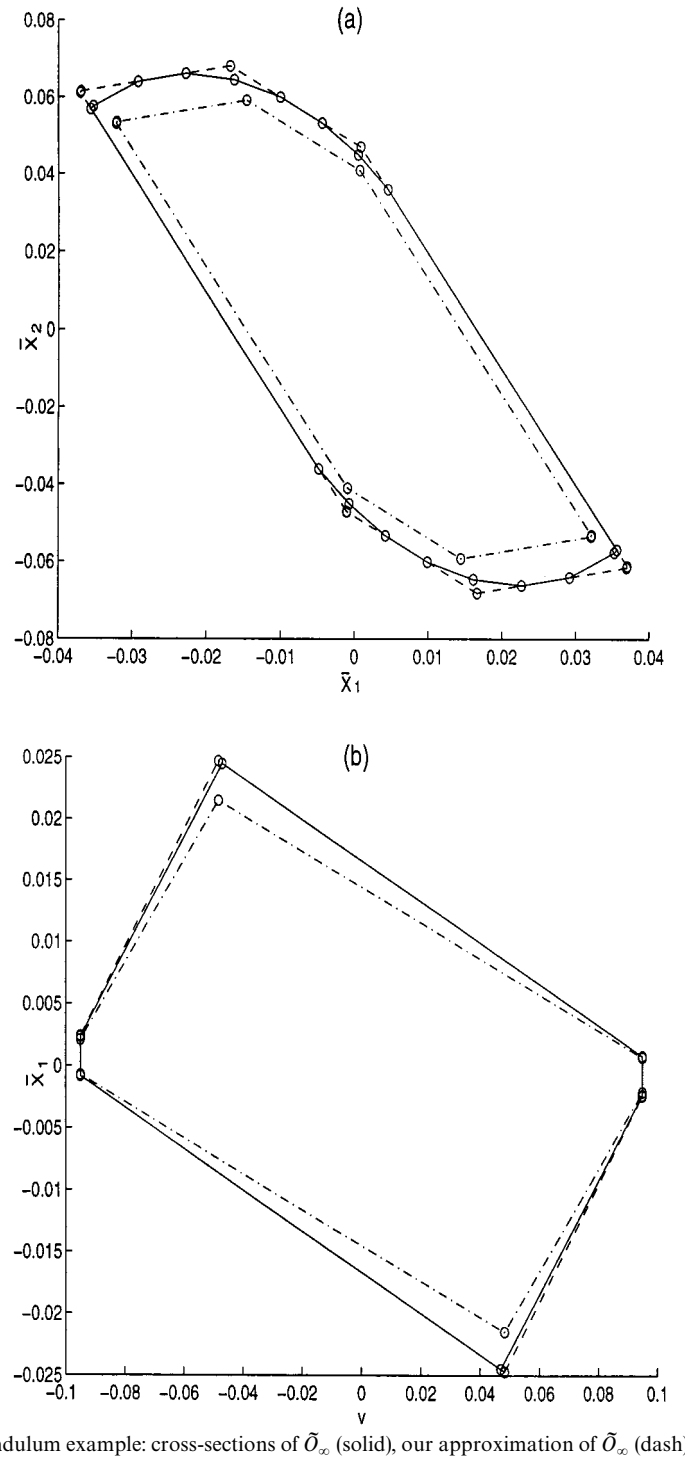


Figure 4. Inverted pendulum example: cross-sections of $\tilde{O}_\infty$ (solid), our approximation of $\tilde{O}_\infty$ (dash) and $P$ (dash–dot) by: (a) $v = 0$; (*b*) $\bar{x}_2 = 0$

 

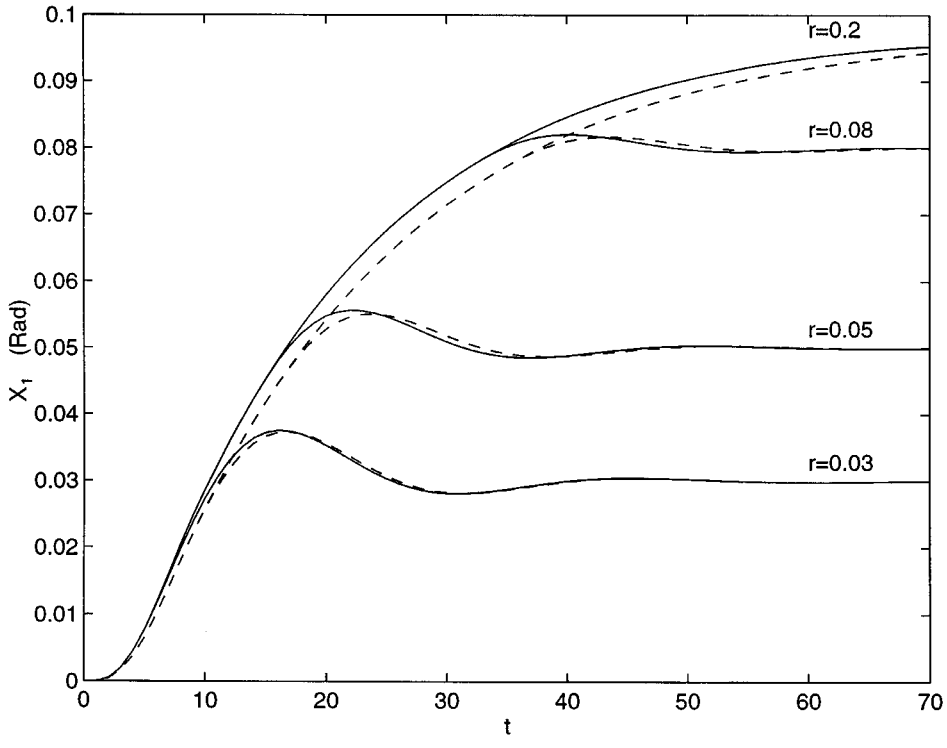


Figure 5. Inverted pendulum example: responses with the reference governor in the disturbance free case for $P = \tilde{O}_\infty$ (solid) and for $P \subset \tilde{O}_\infty$ (dash)

take place when fairly small step inputs are applied. For example, we have performed simulations which show that the system can go unstable with a step $r = 0{\cdot}04$.

In Figure 6 the input is the step $r(t) \equiv r = 0{\cdot}04$ and the disturbance is given by $w(t) = -0{\cdot}002\,\mathrm{sign}(\sin 2t)$. The disturbance is high in frequency so its effect on the response of $x_1(t)$ is not obvious. However, from the plot of $u(t)$, it is clear that the possibility of saturation increases considerably with the disturbance. Note that the response is slower than would be expected for the same input in the disturbance-free case. This is explained easily. The set $P = \tilde{O}_\infty$ is considerably smaller than the set $P$ used in the disturbance-free implementations; hence $K(r(t), v(t), x(t))$ is smaller.

In some applications very small sample periods $T$ are desired, so that the controlled system approaches continuous-time behaviour. It has been observed in Section 1 that decreases in $T$ are troublesome when $P = \bar{O}_\infty$. Figure 4 suggests the nature of the problem. As $T$ decreases from $T = 0{\cdot}1$, more half spaces are required to represent the finer boundary structure of $\tilde{O}_\infty$, but $\tilde{O}_\infty$ does not change very much since dynamics of the control process are represented quite well with $T = 0{\cdot}1$. Thus, $P$, which is well inside $\tilde{O}_\infty$ for $T = 0{\cdot}1$ remains inside for smaller $T$. Simulations show what happens when $T$ is reduced by a factor of 100 to $T = 0{\cdot}001$. The responses are almost the same as with $T = 0{\cdot}1$. The on-line computational requirements for $T = 0{\cdot}001$ and $P = \tilde{O}_\infty$ are vastly different. Although the step responses are again almost the same, the number of half spaces that define $P$ increase from 18 to 4000.

 

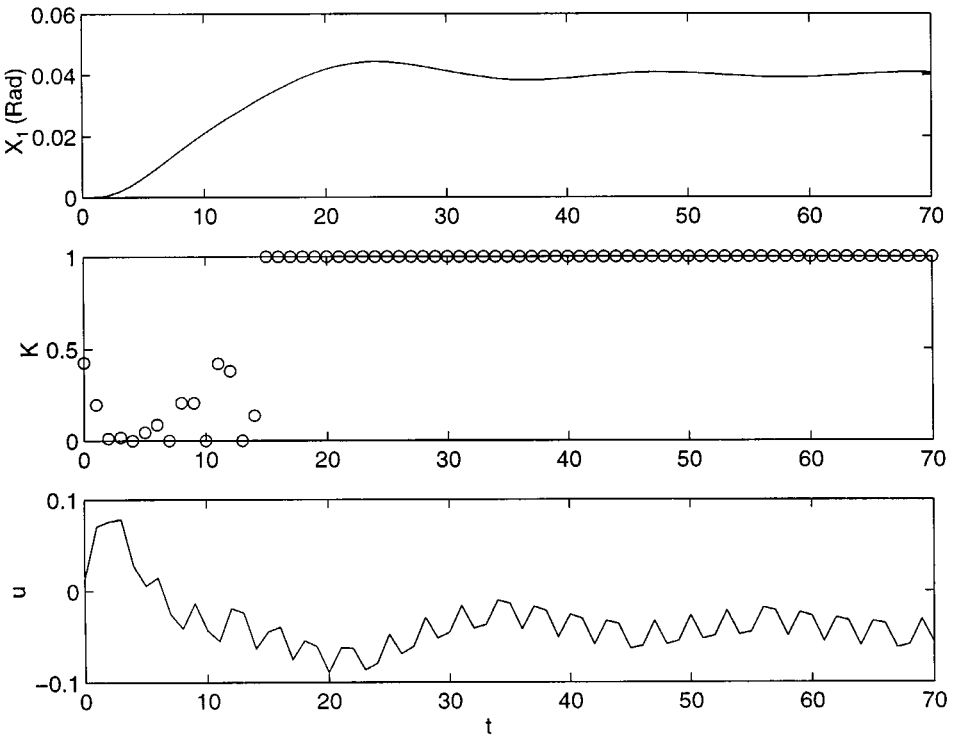


Figure 6. Inverted pendulum example: responses with the reference governor, disturbance, and $r = 0{\cdot}04$

The second example is a pitch acceleration control system for a bank-to-turn missile. It is described by Blakelock,[39] pp. 287–289, and has been used with modifications in the controller gains as an example by Rodriguez and Wang.[40] To allow comparisons with the reference governor results in Reference 40, we choose the parameters of Reference 40. It is assumed that there is no disturbance input.

The missile input, elevator deflection $\delta_e$, saturates at $\pm 4{\cdot}2^\circ$. Available outputs are normal acceleration, $a_z$, and pitch rate $\dot{\theta}$. Laplace transforms of these variables are related by

$$A_z(s) = \frac{-0{\cdot}6368(s - 47{\cdot}318)(s + 47{\cdot}329)}{(s - 10{\cdot}77)(s + 13{\cdot}827)}\Delta_e(s), \qquad \dot{\Theta}(s) = \frac{-701(s + 3{\cdot}194)}{(s - 10{\cdot}77)(s + 13{\cdot}827)}\Delta_e(s) \quad (45)$$

Note that the missile is open-loop unstable and that there is a non-minimum phase zero in the first transfer function. The control objective is to have the normal acceleration follow a reference command, $a_{\text{ref}}$. The two-degree-of-freedom controller uses pitch rate damping in an inner loop and integral plus proportional control in the main feedforward loop:

$$\Delta_e(s) = \frac{4(s + 2)}{s(s + 35)}(A_{\text{ref}}(s) - A_z(s)) + \frac{2{\cdot}2}{s + 35}\dot{\Theta}(s) \quad (46)$$

Eigenvalues of the closed-loop system are at $-15{\cdot}87 \pm 33{\cdot}40\text{i}$ and $-1{\cdot}885 \pm 2{\cdot}189\text{i}$.

 

The discrete-time, controlled process is obtained by generating $\delta_e$ through a zero-order hold driven by a discrete-time controller which is a zero-order hold approximation[41] of the continuous-time controller, (46). To obtain near continuous-time system performance, a small sample period, $T = 0{\cdot}005$ s, is selected. The input to the controlled process, $v(t)$, corresponds to samples of $a_{\text{ref}}$. The output of the discrete-time controller, $y(t) = \delta_e(tT)$, generates the input of the zero-order hold. It is assumed that the system is disturbance free. Thus, (1) is characterized by $m = 1$, $n = 4$, $W = \{0\}$ and the saturation constraint results in $Y = [-4{\cdot}2, +4{\cdot}2]$. Eigenvalues of $A$ are at $0{\cdot}9295 \pm 0{\cdot}1609$i and $0{\cdot}9904 \pm 0{\cdot}0109$i. Without the reference governor saturation does not occur for step commands which have an amplitude less than 31·2. Slightly larger step inputs cause the system response to diverge.

Our reference governor is obtained by setting $\tilde{Y} = [-4{\cdot}0, +4{\cdot}0]$, a reduction in $Y$ of about 5%. Since $T$ is very small, the eigenvalues of $A$ are close to the unit circle. This leads to a complex representation for the polytope $\tilde{O}_\infty$: $t^* = 379$ and there are 762 half spaces. The complexity is not surprising. It reflects the inherent complexity which has been observed in the implementation of continuous-time error governors and reference governors.[2,3,9,40,42]

Figure 7 shows step responses of the reference governor system when $P = \tilde{O}_\infty$. For $r(t) \equiv 15$, the reference governor takes no corrective action and, except for the delay of one sample period, the response is the response of the linear controlled process. For the larger step inputs, $K(r(t), v(t), x(t)) < 1$ for some values of $t$, slowing the response slightly. As expected, instability is avoided. Since, $H_0 = -0{\cdot}1044$, $\tilde{V} = [-38{\cdot}31, 38{\cdot}31]$. Expected results are verified: for steps of

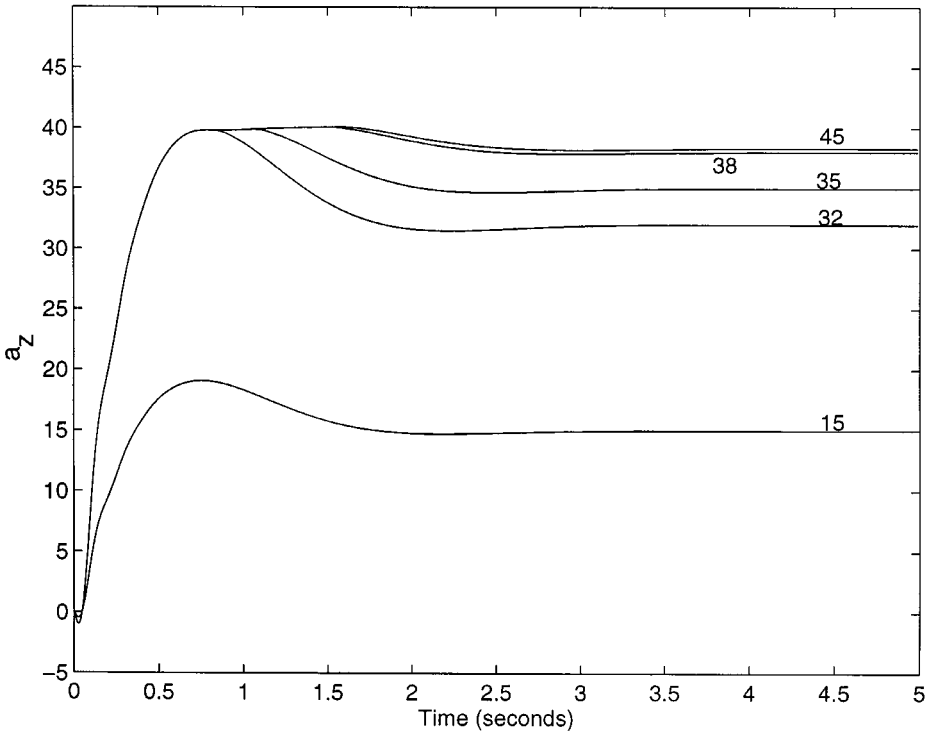


Figure 7. Missile example: time history of $a_z$ with the reference governor and $P = \tilde{O}_\infty$

 

amplitude less than 38·31, $K(r(t), v(t), x(t)) = 1$ for $t$ sufficiently large; for larger steps, $K(r(t), v(t), x(t)) = 0$ for $t$ sufficiently large. The responses in Figure 7 agree closely with those in [40]. The principal difference is that for $r(t) = 45$, the response in Reference 40 approaches $40{\cdot}23 = (0{\cdot}1044)^{-1}$ 4·2 instead of the 38·31 imposed by our replacement of $Y$ by $\tilde{Y}$. We recall the importance of the replacement: it guarantees the finite determination of $\tilde{O}_\infty$ and the finite settling time of $K(r(t), v(t), x(t))$. These issues are not considered in Reference 40.

The on-line computational load for $K$ is very large. Using the symmetry of the constraints, it is approximately 8000 flops. Thus, it is of interest to simplify $P$. Since $\tilde{O}_\infty \subset \mathbb{R}^5$ and is quite complex, the search for an acceptable $P$ cannot be guided by graphical results as it was in our previous example. Instead we take advantage of the numerical ideas described in Section 5. The $\varepsilon_i$ were evaluated for $i = 1, \ldots, 760$. Because of the symmetry of the half spaces it was only necessary to solve half as many linear programming problems. Since the $\varepsilon_i$ are relatively large for $i = 1, \ldots, 58$, the corresponding half spaces were all kept in the initial approximation of $\tilde{O}_\infty$. It was also decided, somewhat arbitrarily, that 2 out of every 16 subsequent half spaces should be kept. This yielded an outer approximation of $\tilde{O}_\infty$, based on the 148 half spaces corresponding to $t \in \mathscr{I}_1$. The set $\mathscr{I}_1$ was further reduced by the stagewise process of Section 5. Using the symmetry of the half spaces, two indices were removed from $\mathscr{I}_k$ at each stage. The process of elimination was continued until the $\varepsilon_i$ which had to be eliminated had increased noticeably. At this point the resulting index set, $\hat{\mathscr{I}}$, contained only 68 indices. The shrinking of $O(\hat{\mathscr{I}})$ about the $\bar{x}$ coordinates to obtain $P$ was attempted for several values of $\alpha$. It was found that the shrinking failed for $\alpha = 1{\cdot}2$ but succeeded for $\alpha = 1{\cdot}3$. The latter value was used in the implementation of (44). The resulting $P$ has 70 inequalities. The number of required flops is about 600, less than one tenth the number for $P = \tilde{O}_\infty$.

The responses, shown in Figure 8, are slightly different in form but only slightly slower.

## 7. CONCLUSION

The reference governor is a supplementary, nonlinear prefilter which enforces point-wise in-time constraints in linear control systems by adjusting the bandwidth parameter $K$. For sufficiently small deviations of the command input it acts only as a unit delay and thus preserves desirable input response and disturbance-rejection properties of the linear system. For large deviations of the command input its output is temporarily slowed, in an optimal way, to insure the absence of present and future constraint violations. This action leads to well-behaved, large-input response properties, a conclusion which is well supported by both the theory and the numerical experiments.

Our results extend prior research[2–4,21–25,27,40,42] on reference governors. The most important new results are in four general areas: (i) finite settling time for $K$ in the presence of persistently varying inputs, (ii) inclusion of disturbance inputs, (iii) the family of functions $K$ generated by choices of $P$ satisfying (30), (iv) specific procedures for implementing off-line design computations associated with (30). As far as we know, the contributions in areas (i) and (iii) are entirely original. With respect to (ii), our developments follow and extend the path originated in Reference 27. See also Reference 26, which treats both disturbance inputs and robustness issues by exploiting mathematical ideas similar to those first employed in Reference 27. Some of the procedures in area (iv) are motivated by ideas discussed in Reference 16.

We believe that the simplifications of $P$, allowed by strict right-side inclusions in (30), are especially important in applications. The complexity problem associated with high sampling rates

 

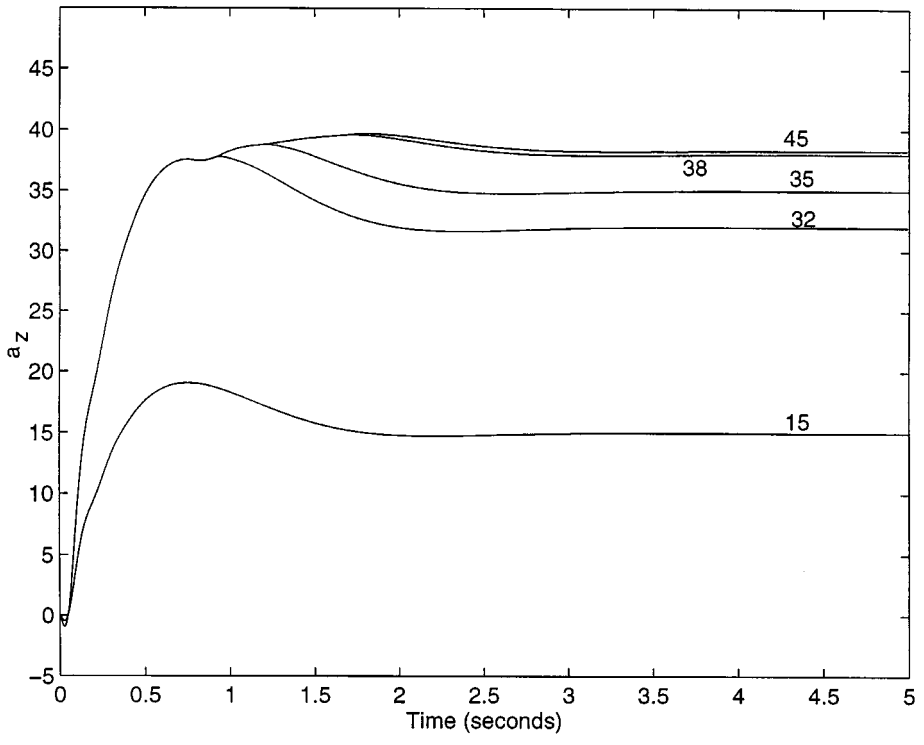


Figure 8. Missile example: time history of $a_z$ with the reference governor and simplified $P$

is an important example. High rates occur commonly when digitally controlled systems are required to mimic closely the response properties of continuous-time systems. Examples in Section 6 illustrate dramatically the simplifications that are possible. There are may other situations in which $\tilde{O}_\infty$ is naturally complex. For instance, $Y$ may be a complex polytope that for computational reasons serves as an approximation to a non-polyhedral constraint set, such as an ellipsoid.

Finally, flexibility in the choice of $P$ can be exploited to guarantee proper reference governor operation in presence of system uncertainty. In principle, the idea is simple. Let $S$ be the set of uncertain systems. Determine a $P$ that satisfies (30) for all $(A, B_v, B_w, C, D_v, D_w, W, Y) \in \mathscr{S}$. Then, for an appropriately modified set of initial conditions, $(v(0), x(0))$, the results of Theorems 4.1 and 4.2 hold for every system in $\mathscr{S}$. In this simplified description we have neglected a minor algebraic complication. Each system in $\mathscr{S}$ generates a distinct change of variables (6). Thus, the actual determination of $P$ must be done in the common coordinate system $(v, x)$, transforming for each system in $\mathscr{S}$, inclusions (30) from their $(v, \bar{x})$ co-ordinates to the $(v, x)$ co-ordinates. The computational procedures for testing the transformed inclusions and computing $K$ are straightforward modifications of the ones discussed in Section 5. This approach to reference governor robustness in quite different than the one describe in Reference 26 and has several advantages: it is conceptually simpler, it can be applied to more general sets $\mathscr{S}$ and it may give stronger results.

 

## APPENDIX A. PROOF OF THEOREM 4.2

The proof of Theorem 4.2 rests on the following lemma.

*Lemma*

Consider the system

$$v(t+1) = v(t) + \kappa(t)(r(t) - v(t)), \quad v(t) \in \mathbb{R}^m, \quad r(t) \in \mathbb{R}^m, \quad t \in Z^+ \tag{47}$$

where $\kappa: Z^+ \to [0, 1]$ and $v(0)$ are arbitrary. Suppose there exist $\hat{t}$, $\delta > 0$ and $r_0$ such that $\|r(t) - r_0\|_\infty \leqslant \delta$ for all $t \geqslant \hat{t}$. Then there exists $t' \geqslant \hat{t}$ such that $\kappa(t)\|v(t) - r(t)\|_\infty \leqslant 2\delta$ for all $t \geqslant t'$.

*Proof.* There are two possibilities for each $i \in \{1, \ldots, m\}$: either there exists $t_i \geqslant \hat{t}$ such that $|v_i(t_i) - r_{0i}| \leqslant \delta$ or $|v_i(t) - r_{0i}| > \delta$ for all $t \geqslant \hat{t}$. The first possibility, (47) and $\kappa(t) \in [0, 1]$ imply that $|v_i(t) - r_{0i}| \leqslant \delta$ and $\kappa(t)|v_i(t) - r_i(t)| \leqslant 2\delta$ for all $t \geqslant t_i$. Suppose the second possibility occurs with $v_i(\hat{t}) > r_{0i} + \delta$. Then it is clear that $r_{0i} + \delta < v_i(t+1) \leqslant v_i(t)$ for all $t \geqslant \hat{t}$. Since $v_i(t)$, $t > \hat{t}$, is non-increasing and bounded from below, it has a limit. Similarly, it has a limit if $v_i(\hat{t}) < r_{0i} - \delta$. Therefore, by (47), $\kappa(t)|v_i(t) - r_i(t)| \to 0$ and there exists $t_i \geqslant \hat{t}$ such that $\kappa(t)|v_i(t) - r_i(t)| \leqslant 2\delta$ for all $t \geqslant t_i$. Thus, with $t' = \max\{t_i: t = 1, \ldots, m\}$, the proof is complete. □

*Proof of Theorem 4.2.* The first step in the proof is to show that the assumption on $r(t)$ implies the existence of $\tilde{t} \geqslant \hat{t}$ such that

$$v(t) \in \tilde{V} \text{ and } \bar{x}(t) \in F + \tfrac{3}{4}\varepsilon\mathscr{B}^n \quad \text{for all } t \geqslant \tilde{t} \tag{48}$$

For notational simplicity let $\bar{\kappa}(t) = \bar{K}(r(t), v(t), \bar{x}(t))$. Suppose $\bar{K}$ is defined by (33). Then by the Lemma there exists $t' \geqslant \hat{t}$ such that

$$\|\bar{\kappa}(t)(r(t) - v(t))\|_\infty \leqslant \frac{\varepsilon}{2G} \quad \text{for all } t \geqslant t' \tag{49}$$

The same inequality is satisfied if $\bar{K}$ is defined by (35): apply the Lemma with $m = 1$ to each component of (9) to obtain $\bar{\kappa}_i(t)|r_i(t) - v_i(t)| \leqslant (\varepsilon/2G)$ for all $t \geqslant t'_i$, $i = 1, \ldots, m$; then (49) holds with $t' = \max\{t'_1, \ldots, t'_m\}$. For $t \geqslant t'$ the solution of (7) can be written as $\bar{x}(t) = A^{t-t'}\bar{x}(t') + \bar{x}_{\mathrm{RG}}(t) + \bar{x}_w(t)$, where $\bar{x}_{\mathrm{RG}}(t)$ and $\bar{x}_w(t)$ are zero-state responses of (7) which are, respectively, due to each of the input terms on the right-hand side of (7). By (49) and the definition of $G$, $\bar{x}_{\mathrm{RG}}(t) \in (\varepsilon/2)\mathscr{B}^n$; by the definitions of $F_t$ and $F$, $\bar{x}_w(t) \in F_{t-t'} \subset F$. Moreover, by the asymptotic stability of $A$ there exists $\tilde{t} \geqslant t'$ such that $A^{t-t'}\bar{x}(t') \in (\varepsilon/4)\mathscr{B}^n$ for all $t \geqslant \tilde{t}$. Combining these facts proves $\bar{x}(t) \in F + (3\varepsilon/4)\mathscr{B}^n$. Part (i) of Theorem 4.1 and (26) complete the verification of (48).

Apply (48) to $v(t+1)$ and $\bar{x}(t+1)$ using (9) and (7). Since the resulting inclusions must hold for all $w(t) \in W$, it follows that

$$v(t) + \bar{\kappa}(t)(r(t) - v(t)) \in \tilde{V} \quad \text{for all } t \geqslant \tilde{t} \tag{50}$$

$$A\bar{x}(t) - \bar{B}_v\bar{\kappa}(t)(r(t) - v(t)) \in (F + \tfrac{3}{4}\varepsilon\mathscr{B}^n) \sim B_w W \quad \text{for all } t \geqslant \tilde{t} \tag{51}$$

These inclusions, together with (30) and (31), confirm that $\kappa = \bar{\kappa}(t)$ satisfies (32) with $r = r(t)$, $v = v(t)$ and $\bar{x} = \bar{x}(t)$. In fact, (51) shows that the last $n$ components of (32) are *inactive* constraints on $\bar{\kappa}(t)$, i.e. for all $\kappa$ in a sufficiently small neighborhood of $\bar{\kappa}(t)$ the last $n$ components of (32) are satisfied automatically. Thus, for $t \geqslant \tilde{t}$ active constraints in the optimization problems of (33) and (35) must come from (50).

Arguing by contradiction, this result leads to the proof. Suppose $\bar{K}$ is defined by (33) and $\bar{\kappa}(t) < 1$ for some $t \geqslant \tilde{t}$. Since $r(t)$ and $v(t)$ both belong to the convex set $\tilde{V}$, it is possible to find $\kappa > \bar{\kappa}(t)$, $\kappa - \bar{\kappa}(t)$ sufficiently small, which causes (32) to remain valid. Thus, $\bar{\kappa}(t) = \bar{K}(r(t), v(t), \bar{x}(t))$ fails to satisfy the maximum condition of (33), which is a contradiction. Hence, $\bar{\kappa}(t) = 1$ and $v(t+1) = r(t)$.

When $\bar{K}$ is defined by (35) the corresponding argument is more involved. Let $\mu = (\kappa_1, \ldots, \kappa_m)$, $\bar{\mu}(t) = (\bar{\kappa}_1(t), \ldots, \bar{\kappa}_m(t)) \in \mathbb{R}^m$ and $\mathscr{N}(r, v) = \{\mu \in \mathbb{R}^m: \mu + \kappa(r - v) \in \tilde{V}, \kappa \in \mathbb{R}^{m \times m} \text{ satisfies (34)}\}$. Then (50) is equivalent to $\bar{\mu}(t) \in \mathscr{N}(r(t), v(t))$. Clearly, $v + \kappa(r - v)$ is an affine vector function of $\mu$. From this and the convexity of $\tilde{V}$, $\mathscr{N}(r(t), v(t))$ is convex. Since $\|\cdot\|$ is a convex function, it also follows that

 

$f(\mu) = \| v(t) + \kappa(r(t) - v(t)) - r(t) \|$ is a convex function of $\mu$. Let $e \in \mathbb{R}^m$ denote the vector whose components are all equal to 1. Obviously, $e \in \mathcal{N}(r(t), v(t))$ and $f(e) = 0$. Now suppose, contrary to the Theorem, that $v(t+1) \neq r(t)$ for some $t \geqslant \tilde{t}$. Then, $f(\bar{\mu}(t)) > 0$. Define $\mu = \bar{\mu}(t) + \lambda(e - \bar{\mu}(t))$. By our convexity results and $f(e) = 0$, it follows that $\mu \in \mathcal{N}(r(t), v(t))$ and $f(\mu) \leqslant (1 - \lambda) f(\bar{\mu}(t))$ for all $\lambda \in [0, 1]$. Thus, for $\lambda > 0$ sufficiently small, (32) is satisfied and $f(\mu) < f(\bar{\mu}(t))$. Hence, the minimization condition in (35) is violated and the proof is complete. □

## REFERENCES


1. Bernstein, D. S. and A. N. Michel, 'A chronological bibliography on saturating actuators', *Int. J. Robust Nonlinear Control*, **5**, 375–380 (1995).
2. Kapasouris, P., M. Athans and G. Stein, 'Design of feedback control systems for unstable plants with saturating actuators', in *Proc. IFAC Symp. on Nonlinear Control System Design*, A. Isodori (Ed.), Pergamon Press, 1990, pp. 302–307.
3. Kapasouris, P., 'Design for performance enhancement in feedback control systems with multiple saturating nonlinearities', Ph.D. dissertation, MIT 1988.
4. Gilbert, E. G., I. Kolmanovsky and K. T. Tan, 'Discrete-time reference governors and the nonlinear control of systems with state and control constraints', *Int. J. Robust Nonlinear Control*, **5**, 487–504 (1995).
5. Kapasouris, P., M. Athans and G. Stein, 'Design of feedback control systems for stable plants with saturating actuators', *Proc. IEEE Conf. on Decision and Control*, Austin, Texas, 1988, pp. 469–475.
6. Benzaouia, A. and C. Burgat, 'Regulator problem for linear discrete-time systems with nonsymmetrical constrained control', *Int. J. Control*, **48**, 2441–2451 (1988).
7. Bitsoris, G., 'Positively invariant polyhedral sets of discrete-time linear systems', *Int. J. Control*, **47**, 1713–1726 (1988).
8. Blanchini, F., 'Constrained control for perturbed linear systems', *Proc. 29th IEEE Conf. on Decision and Control*, Honolulu, Hawaii, 1990, pp. 3464–3467.
9. Blanchini, F., 'Constrained control for systems with unknown disturbances', in *Control and Dynamic Systems*, C. T. Leondes (Ed.), Vol. 51, Academic Press, New York, 1992 130–182.
10. Blanchini, F., 'Control synthesis for discrete-time linear systems with control and state bounds in the presence of disturbances', *J. Opt. Theory Appl.*, **65**, 29–40 (1990).
11. Blanchini, F., 'Feedback control for linear time-invariant systems with state and control bounds in the presence of disturbances', *IEEE Trans. Automat. Control*, **AC-35**, 1231–1234 (1990).
12. Blanchini, F., 'Ultimate boundness control for uncertain systems via set-induced Lyapunov functions', *IEEE Trans. Automat. Control*, **39**, 428–433 (1994).
13. Blanchini, F. and M. Sznaier, 'Persistent disturbance rejection via static state feedback', *IEEE Trans. Automat. Control*, **40**, 1127–1131 (1995).
14. Gilbert, E. G. and K. T. Tan, 'Linear systems with state and control constraints: the theory and applications of maximal output admissible sets', *IEEE Trans. Automat. Control*, **36**, 1008–1020 (1991).
15. Kolmanovsky, I. and E. G. Gilbert, 'Multimode regulator for systems with state and control constraints and disturbance inputs', in *Control Using Logic-Based Switching*, A. S. Morse (Ed.), Spring, Berlin, 1996, pp. 118–127.
16. Kolmanovsky, I. and E. G. Gilbert, 'Theory and computation of disturbance invariant sets for discreter-time linear systems', *Math. Meth. Engng.: Theory, Meth. Appl.*, **4**, 317–367 (1998).
17. Vassilak, M., J. C. Hennet and G. Bitsoris, 'Feedback control of linear systems under state and control constraints', *Int. J. Control*, **47**, 1727–1735 (1988).
18. Blanchini, F. and S. Miani, 'Constant signal tracking for state constrained dynamic systems', *Proc. 36th Conf. on Decision and Control*, San Diego, California, 1997, pp. 2870–2875.
19. Mayne, D. Q. and W. R. Schroeder, 'Robust time-optimal control of constrained linear systems', *Automatica*, **33**, 2103–2118 (1997).
20. Tan, K. T., 'Maximal output admissible sets and the nonlinear control of linear discrete-time systems with state and control constraints', Ph.D. dissertation, Department of Aerospace Engineering, the University of Michigan, 1991.
21. Gilbert, E. G., I. Kolmanovsky and K. T. Tan, 'Nonlinear control of discrete-time linear systems with state and control constraints: a reference governor with global convergence properties', *Proc. 33rd IEEE Conf. on Decision and Control*, Orlando, Florida, 1994, pp. 144–149.
22. Bemporad, A. and E. Mosca, 'Constraint fulfillment in control systems via predictive reference management', *Proc. 33rd IEEE Conf. Decision and Control*, Orlando, Florida, 1994, pp. 3017–3022.
23. Bemporad, A. and E. Mosca, 'Constraint fulfillment in feedback control via predictive reference management', *Proc. 3rd IEEE Conf. on Control Applications*, Glasgow, U.K., 1994, pp. 1909–1914.
24. Bemporad, A. and E. Mosca, 'Nonlinear predictive reference governor for constrained control system', *Proc. 34th IEEE Conf. on Decision and Control*, New Orleans, Louisiana, 1995, pp. 1205–1210.


 

25. Bemporad, A., A. Cosavola, and E. Mosca, 'Nonlinear control of constrained linear systems via predictive reference management', *IEEE Trans. Automat. Control*, **42**, 340–349 (1996).
26. Casavola, A. and E. Mosca, 'Reference governor for constrained linear systems subject to bounded input disturbances', *Proc. 35th Conf. on Decision and Control*, Kobo, Japan, 1996, pp. 3531–3536.
27. Gilbert, E. G. and I. Kolmanovsky, 'Reference governors for systems with state and control constraints and disturbance inputs', *Proc. 34th IEEE Conf. on Decision and Control*, New Orleans, Louisiana, 1995, pp. 1189–1194.
28. Aubin, J. P., *Viability Theory*, Birkhauser, Basel, 1991.
29. Krasovskii, N. N. and A. I. Subbotin, *Game-Theoretical Control Problems*, Springer, Berlin, 1988.
30. Kolmanovsky, I. and E. G. Gilbert, 'Maximal output admissible sets for discrete-time systems with disturbance inputs', *Proc. American Control Conf.* Seattle, Washington, 1995, pp. 1995–2000.
31. Schneider, R., *Convex Bodies: The Brunn–Minkowski Theory*, Cambridge University Press, Cambridge, 1993.
32. Pontryagin, L. S., 'Linear differential games', *Sov. Math. Dokl*, **8**(3), 769–771 (1967).
33. Bertsekas, D. P. and I. B. Rhodes, 'On the minimax reachability of target sets and target tubes', *Automatica*, **7**, 233–247 (1971).
34. Delfour, M. C. and S. K. Mitter, 'Reachability of perturbed systems and min sup problems', *SIAM J. Control*, **7**, 521–533 (1969).
35. Demyanov, V. F. and A. M. Rubinov, *Foundations of Nonsmooth Analysis and Quasidifferential Calculus*, Nauka, Moscow, 1990, pp. 124–125 [in Russian].
36. Glover, J. D. and F. C. Schweppe, 'Control of linear dynamical systems with set constrained disturbances', *IEEE Trans. Automat. Control*, **AC-16**, 411–423 (1971).
37. Keerthi, S. S. and K. Sridharan, 'Solution of parameterized linear inequalities by Fourier elimination and its applications', *J. Optim. Theory Appl.*, **65**, 161–169 (1990).
38. Keerthi, S. S. and E. G. Gilbert, 'Computation of minimum time feedback control laws for discrete-time linear systems with state-control constraints', *IEEE Trans. Automat. Control*, **AC-32**, 432–435 (1987).
39. Blakelock, J. H., *Automatic Control of Aircraft and Missiles*, 2nd edn., Wiley, New York, 1991.
40. Rodriguez, A. A. and Y. Wang, 'Performance enhancement methods for unstable bank-to-turn (BTT) missiles with saturating actuators', *Int. J. Control*, **63**(4), 641–678 (1996).
41. Franklin, G. F., J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems*, Addison-Wesley Publishing Company, Reading, MA, 1990.
42. Rodriguez, A. A. and J. R. Cloutier, 'Control of a back-to-turn (BTT) missile with saturating actuators', *Proc.* 1994 *American Control Conf.*, Baltimore, Maryland, 1994, pp. 1660–1664.
43. Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, 1972.

 