Set-Valued Observers and Optimal Disturbance Rejection

Jeff S. Shamma, *Member, IEEE*, and Kuang-Yang Tu, *Member, IEEE*

***Abstract*—A set-valued observer (also called guaranteed state estimator) produces a set of possible states based on output measurements and models of exogenous signals. In this paper, we consider the guaranteed state estimation problem for linear time-varying systems with *a priori* magnitude bounds on exogenous signals. We provide an algorithm to propagate the set of possible states based on output measurements and show that the centers of these sets provide optimal estimates in an ℓ^∞-induced norm sense. We then consider the utility of set-valued observers for disturbance rejection with output feedback and derive the following general separation structure. An optimal controller can consist of a set-valued observer followed by a static nonlinear function on the observed *set* of possible states. A general construction of this function is provided in the scalar control case. Furthermore, in the special case of full-control, i.e., the number of control inputs equals the number of states, optimal output feedback controllers can take the form of an optimal estimate of the full-state feedback controller.**

***Index Terms*—Disturbance rejection, state estimation, observers.**

I. Introduction

STOCHASTIC state estimation provides optimal state estimates based on probabilistic models of exogenous signals. An alternative is to model exogenous signals as deterministic unknown but bounded quantities. The problem is then to construct a *set* of possible state values based on measured outputs. Such an approach has received considerable attention in the controls literature. References [12] and [24] present an overview of work in this area, and [22] contains a collection of related conference papers.

Related to the deterministic setting is induced-norm optimal state estimation. This framework provides optimal state estimates which minimize the induced-norm from exogenous signals to estimation errors. Reference [26] considers the case where exogenous signals and estimation errors are measured using the ℓ^2-norm, or signal energy, which leads to an $\mathcal{H}^\infty$ optimal estimation problem. Reference [33] measures exogenous signals and estimation errors by the ℓ^∞ norm, or signal magnitude, which leads to an ℓ^1 optimal estimation problem.

In this paper, we consider guaranteed state estimation for linear time-varying systems. Under an assumed *a priori* bound on exogenous signals, we present a construction of the set of possible state values. We then relate the centers of these sets to the ℓ^1 optimal estimation problem considered in [33]. In particular, we show that the centers are also optimal in an ℓ^∞ induced-norm sense.

We then investigate the utility of set-valued observers for ℓ^∞-induced norm optimal disturbance rejection. References [30] and [31] considered this disturbance rejection problem in the special case of noise-free state feedback and showed that optimal controllers can be static nonlinear functions of the state. This is in contrast to [15] which showed that optimal *linear* controllers may be dynamic and of arbitrarily high order. In this paper, we consider noisy output feedback. We show that optimal controllers can take the following separation-like structure: 1) a set-valued observer plus 2) a static nonlinear function on the *set* of possible states. A general construction of this function is provided in the scalar control case. Furthermore, in the special case of full-control, i.e., the number of control inputs equals the number of states, optimal output feedback controllers can take the form of an optimal estimate of the full-state feedback controller.

The remainder of this paper is organized as follows. Section II contains preliminary definitions and notation. Section III presents an algorithm which propagates the set-valued estimates based on output measurements and derives the ℓ^∞ induced-norm optimality of the centers of these sets. Section IV discusses applications to disturbance rejection. Finally, Section V contains a simulation example, and Section VI has concluding remarks.

II. Mathematical Preliminaries

A. Basic Notation

For $x \in \mathcal{R}^n$, let x_i denote the ith component of x and define $|x| = \max_i |x_i|$.

Let $\mathcal{Z}^+$ denote the set of nonnegative integers. Let ℓ^∞ denote the set of bounded one-sided sequences in $\mathcal{R}^n$. For $f = \{f(0), f(1), f(2), \cdots\} \in \ell^\infty$, define

$$\|f\| = \sup_{k \in \mathcal{Z}^+} |f(k)|.$$

The dimension n is suppressed in ℓ^∞ for notational convenience. The unit balls in $\mathcal{R}^n$ and ℓ^∞ are denoted $B_{\mathcal{R}^n}$ and B_{ℓ^∞}, respectively.

Define **1** and **0** to be vectors of 1's or 0's, respectively, of appropriate length.

Manuscript received October 19, 1995; revised February 25, 1998. Recommended by Associate Editor, M. A. Dahleh. This work was supported by the NSF under Grant ECS–92258005, EPRI under Grant #8030–23, and Ford Motor Co.

The authors are with the Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712 USA.

Publisher Item Identifier S 0018-9286(99)01298-2.

A set-valued map, denoted $F: X \rightsquigarrow Y$, is a mapping from points $x \in X$ to subsets $F(x) \subset Y$.

B. Projections of Convex Sets

For $M \in \mathcal{R}^{\ell \times n}$ and $m \in \mathcal{R}^{\ell}$, let $\mathrm{Set}(M, m)$ denote the subset of $\mathcal{R}^n$ associated with (M, m) defined by the constraints

$$\mathrm{Set}(M, m) = \{x: Mx \le m\}.$$

For $M_1 \in \mathcal{R}^{\ell \times n}$, $M_2 \in \mathcal{R}^{\ell}$, and $m \in \mathcal{R}^{\ell}$ consider the subset, S, of $\mathcal{R}^n$ defined by

$$S = \{x: M_1 x + M_2 w \le m \text{ for some } w \in \mathcal{R}\}.$$

Define

$$\begin{aligned}\mathrm{Rack}[(M_1 \quad M_2), m] \\ = \{(\tilde{M}, \tilde{m}) \in \mathcal{R}^{\tilde{\ell} \times n} \times \mathcal{R}^{\tilde{\ell}}: S = \mathrm{Set}(\tilde{M}, \tilde{m})\}\end{aligned}$$

i.e., $\mathrm{Rack}[(M_1\ M_2, m]$ is the set of matrix pairs which give a direct characterization of S. While the set S is unique, its matrix representation is not. Hence, $\mathrm{Rack}[M_1\ M_2, m]$ represents a *set* of possible matrix representations. The construction of an element of $\mathrm{Rack}[M_1\ M_2, m]$ may be achieved through the Fourier–Motzkin algorithm which is described in [21].

Now define

$$\mathrm{Rack}^2[M, m] = \mathrm{Rack}[\mathrm{Rack}[M, m]]$$

and recursively define

$$\mathrm{Rack}^k[M, m] = \mathrm{Rack}[\mathrm{Rack}^{k-1}[M, m]].$$

The notation $\mathrm{Rack}^k[M, m]$ is simply a multivariable form of $\mathrm{Rack}[M, m]$. For $M_1 \in \mathcal{R}^{\ell \times n}$ and $M_2 \in \mathcal{R}^{\ell \times k}$, redefine the subset $S \subset \mathcal{R}^n$ as

$$S = \{x: M_1 x + M_2 w \le m \text{ for some } w \in \mathcal{R}^k\}.$$

Then $\mathrm{Rack}^k[(M_1\ \ M_2), m]$ is the set of matrix pairs which give a direct characterization of S.

III. Set-Valued Estimation

A. Set Propagation

This section considers the time-varying discrete-time linear system

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)d(k), \qquad x(0) = x_o \\ y(k) &= C(k)x(k) + n(k) \end{aligned} \tag{1}$$

where $x(k) \in \mathcal{R}^{n_x}$ is the state-vector, $y(k) \in \mathcal{R}^{n_y}$ is the measured output, $d(k) \in \mathcal{R}^{n_d}$ is a process disturbance, and $n(k) \in \mathcal{R}^{n_y}$ is a measurement noise.

Define $w = \binom{d}{n}$. In input–output form, system (1) takes the form

$$y = T_{yw} w + T_{yx_o} x_o$$

where T_{yw} denotes the mapping from w to y with the initial condition $x_o = 0$, and T_{yx_o} denotes the mapping from x_o to y with the input $w = 0$. Similarly define T_{xw} and T_{xx_o}.

The following assumption reflects the (deterministic) *a priori* model of the exogenous signals and initial condition.

Assumption 3.1: The exogenous inputs satisfy $w \in B_{\ell^\infty}$ and the initial condition satisfies $x(0) \in B_{\mathcal{R}^{n_x}}$.

We are interested in constructing an estimate of the state vector based on output measurements. Toward this end, define the set-valued map $W: \ell^\infty \times \mathcal{Z}^+ \rightsquigarrow \mathcal{R}^{n_x}$ as

$$\begin{aligned} W(y,k) = \{(w, x_o) \in B_{\ell^\infty} \times B_{\mathcal{R}^{n_x}}: \\ y(j) = (T_{yw} w + T_{yx_o} x_o)(j),\ j = 0, \cdots, k\}. \end{aligned}$$

In other words, $W(y,k)$ denotes the set of admissible exogenous signals and initial conditions consistent with measured data up to time k. Similarly, define the set-valued $X: \ell^\infty \times \mathcal{Z}^+ \rightsquigarrow \mathcal{R}^{n_x}$ given by

$$\begin{aligned} X(y,k) = \{x \in \mathcal{R}^{n_x}: x = (T_{xw} w + T_{xx_o} x_o)(k) \\ \text{for some } (w, x_o) \in W(y,k)\} \end{aligned}$$

i.e., $X(y,k)$ denotes the set of possible state-vectors at time k consistent with the measured data up to time k. Finally, define the set-valued $\tilde{X}: \mathcal{R} \rightsquigarrow \mathcal{R}^{n_x}$ by

$$\tilde{X}(y) = \{x \in \mathcal{R}^{n_x}: y = Cx + n \text{ for some } |n| \le 1\}.$$

The set $\tilde{X}(y)$ represents the set of possible states based on a single measurement.

The following algorithm (see also [12, Sec. 20]) propagates the set of possible states.

Algorithm 3.1: Let $y \in \ell^\infty$ be a prescribed measurement trajectory.

Initialization:

$$X(y, 0) = \tilde{X}(y(0)) \cap B_{\mathcal{R}^{n_x}}.$$

Propagation:

$$\begin{aligned} X(y,k) = \{x: x = A(k-1)\tilde{x} + B(k-1)d \text{ for some} \\ \tilde{x} \in X(y, k-1), |d| \le 1\} \cap \tilde{X}(y(k)). \end{aligned}$$

Note that all sets are constructed with a *causal* dependence on the measurement trajectory, y.

The following theorem describes a computational implementation of Algorithm 3.1.

Theorem 3.1: In the framework of Algorithm 3.1

$$\tilde{X}(y(k)) = \mathrm{Set}(\tilde{M}(k), \tilde{m}(k))$$

where

$$\tilde{M}(k) = \begin{pmatrix} C(k) \\ -C(k) \end{pmatrix}, \qquad \tilde{m}(k) = \begin{pmatrix} \mathbf{1} + y(k) \\ \mathbf{1} - y(k) \end{pmatrix}$$

and

$$X(y,k) = \mathrm{Set}(M(k), m(k))$$

where we have the equation, shown at the bottom of the page, in case $A(k-1)$ is invertible. If not, then

$$(M(k), m(k)) \in \mathrm{Rack}^{n_x+n_d}\left[\begin{pmatrix} I & -A(k-1) & -B(k-1) \\ -I & A(k-1) & B(k-1) \\ 0 & 0 & I \\ 0 & 0 & -I \\ \tilde{M}(k) & 0 & 0 \\ 0 & M(k-1) & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \\ \tilde{m}(k) \\ m(k-1) \end{pmatrix}\right].$$

Proof: The condition $x(k) \in \tilde{X}(y(k))$ is equivalent to

$$|y(k) - C(k)x(k)| \le 1$$

which is equivalent to

$$\begin{pmatrix} C(k) \\ -C(k) \end{pmatrix} x(k) \le \begin{pmatrix} \mathbf{1} + y(k) \\ \mathbf{1} - y(k) \end{pmatrix}$$

which is the matrix description of $\tilde{X}(k)$.

Now according to Algorithm 3.1, the condition $x(k) \in X(y,k)$ is equivalently described by the two conditions $x(k) \in \tilde{X}(y(k))$ and

$$x(k) = A(k-1)x(k-1) + B(k-1)d \tag{2}$$

for some $x(k-1) \in X(y, k-1)$ and $|d| \le 1$.

In case $A(k-1)$ is invertible, condition (2) is equivalent to

$$A^{-1}(k-1)x(k) - A^{-1}(k-1)B(k-1)d \in X(y, k-1)$$

for some $|d| \le 1$. Using that $X(y, k-1) = \mathrm{Set}(M(k-1), m(k-1))$ leads to the equivalent statement

$$\begin{pmatrix} M(k-1)A^{-1}(k-1) & -M(k-1)A^{-1}(k-1)B(k-1) \\ 0 & I \\ 0 & -I \\ \tilde{M}(k) & 0 \end{pmatrix} \begin{pmatrix} x(k) \\ d \end{pmatrix} \le \begin{pmatrix} m(k-1) \\ \mathbf{1} \\ \mathbf{1} \\ \tilde{m}(k) \end{pmatrix}$$

for some d. An application of the $\mathrm{Rack}^{n_d}[\cdot]$ operator leads to the desired result.

In case $A(k-1)$ is not invertible, the requirements on $x(k)$ become

$$\begin{pmatrix} I & -A(k-1) & -B(k-1) \\ -I & A(k-1) & B(k-1) \\ 0 & 0 & I \\ 0 & 0 & -I \\ \tilde{M}(k) & 0 & 0 \\ 0 & M(k-1) & 0 \end{pmatrix} \begin{pmatrix} x(k) \\ x(k-1) \\ d \end{pmatrix} \le \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \\ \tilde{m}(k) \\ m(k-1) \end{pmatrix}$$

for some d. An application of the $\mathrm{Rack}^{n_x+n_d}[\cdot]$ operator leads to the desired result. ■

The matrices $M(0)$ and $m(0)$ are initialized as

$$M(0) = \begin{pmatrix} I \\ -I \end{pmatrix}, \qquad m(0) = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix}$$

to reflect the *priori* assumption $x_o \in B_{\mathcal{R}^{n_x}}$.

We see that the set of possible states forms a polytope described by a collection of inequalities. The computational burden of a real-time implementation amounts to the computation of the $\mathrm{Rack}[\cdot]$ operator, which essentially requires the solution of several small linear programs to remove redundant constraints. Since these sets may be described by several inequalities, the real-time applicability of these methods is questionable. This consideration has led to the construction of approximate simplified descriptions of $X(y,k)$, in particular through bounding ellipsoids. See [12], [24], and references contained therein for further discussion on these topics.

Note that we have made no statements regarding the observability of the original system. The above characterization holds regardless of observability or detectability assumptions. However, it is straightforward to show that an appropriate notion of detectability implies that the sets $X(y,k)$ are bounded uniformly.

Finally, we note that the above algorithms easily may be modified to accommodate a known input (such as a control) into the state dynamics or a different set of initial conditions. Such changes will be needed in the forthcoming section on disturbance rejection.

B. ℓ^∞ Induced-Norm Optimal Estimation

In this section, we show that the set-valued observer in Section III-A can be used to provide optimal estimates in an induced-norm sense.

Define the scalar variable

$$z(k) = H(k)x(k).$$

$$(M(k), m(k)) \in \mathrm{Rack}^{n_d}\left[\begin{pmatrix} M(k-1)A^{-1}(k-1) & -M(k-1)A^{-1}(k-1)B(k-1) \\ 0 & I \\ 0 & -I \\ \tilde{M}(k) & 0 \end{pmatrix}, \begin{pmatrix} m(k-1) \\ \mathbf{1} \\ \mathbf{1} \\ \tilde{m}(k) \end{pmatrix}\right]$$

In case z is vector-valued, an optimal estimate can be obtained from optimal estimates of the individual components. As in Section III-A, define T_{zw} and T_{zx_o} as the mappings from exogenous signals and initial conditions, respectively, to z.

We now define our optimal estimation problem.

Definition 3.1: An *estimator* is any causal (possibly nonlinear) mapping $\Phi\colon \ell^\infty \to \ell^\infty$.

Definition 3.2: The estimator Φ^* is *pointwise optimal* if for any other estimator, Φ

$$\sup_{(w,x_o)\in W(y,k)} \frac{|(T_{zw}w + T_{zx_o})(k) - (\Phi^* y)(k)|}{\|(w,x_o)\|} \le \sup_{(w,x_o)\in W(y,k)} \frac{|(T_{zw}w + T_{zx_o})(k) - (\Phi y)(k)|}{\|(w,x_o)\|}, \quad \forall k \in \mathcal{Z}^+$$

for all possible measurement trajectories.

The estimator Φ^* is *uniformly optimal* if for any other estimator Φ

$$\sup_{\substack{(w,x_o)\in B_{\ell^\infty}\times B_{\mathcal{R}^{n_x}} \\ y=T_{yw}w+T_{yx_o}x_o}} \frac{|(T_{zw}w + T_{zx_o}x_o)(k) - (\Phi^* y)(k)|}{\|(w,x_o)\|} \le \sup_{\substack{(w,x_o)\in B_{\ell^\infty}\times B_{\mathcal{R}^{n_x}} \\ y=T_{yw}w+T_{yx_o}x_o}} \frac{|(T_{zw}w + T_{zx_o}x_o)(k) - (\Phi y)(k)|}{\|(w,x_o)\|}, \quad \forall k \in \mathcal{Z}^+.$$

Pointwise optimality is a stronger property than uniform optimality. Pointwise optimality assures that the current estimation error is the smallest possible for the *current* measurement trajectory, whereas uniform optimality assures that the current estimation error is smaller than the smallest worst case estimation error over all trajectories. Thus if the measurement trajectory is benign in some sense, the pointwise optimal estimation error can be *less* than the uniformly optimal estimation error. However, there exists a worst case trajectory for which both errors coincide.

The above measures of estimation performance take the form of induced-norms over bounded sets. Another estimation performance measure is simply direct estimation error, i.e.,

$$\sup_{(w,x_o)\in W(y,k)} |(T_{zw}w + T_{zx_o})(k) - (\Phi y)(k)|. \tag{3}$$

Here, the error is not normalized by the size of the exogenous signals and initial condition which produced the error. In the case of linear system dynamics and linear observers, the two notions coincide. Such an unnormalized measure of estimation performance was considered in [24]. Unnormalized measures of estimation performance are natural in the present case of bounded exogenous signals and initial conditions. However, a benefit of induced-norm optimality is that it assures that "overbounding" the exogenous signals and initial conditions does not deteriorate the estimation performance. For example, while the *a priori* assumptions assure $\|w\| \le 1$, the actual exogenous signals might satisfy $\|w\| \ll 1$. Induced-norm optimality assures that the resulting estimation errors are not affected by the conservative bound. Furthermore, induced-norm optimality can be useful when establishing robustness properties.

Reference [33] considers the uniformly optimal estimation problem. In the case of zero-initial conditions and time-invariant dynamics, the uniformly optimal estimation problem can be solved as a standard ℓ_1 model-matching problem (cf., [14]). For nonzero initial conditions, the model matching problem is time-varying, and the optimal estimate at time k requires storage of all measurements $\{y(0), \cdots, y(k)\}$. Reference [33] goes on to provide an approximately optimal estimator which is recursive after a fixed number of time-steps.

The following proposition summarizes the results of [33] needed here.

Proposition 3.1 [33]: There exists a uniformly optimal linear (time-varying) estimator Q. Furthermore, the associated worst case estimation error $\gamma(k)$, defined by

$$\gamma(k) = \sup_{\substack{(w,x_o)\in B_{\ell^\infty}\times B_{\mathcal{R}^{n_x}} \\ y=T_{yw}w+T_{yx_o}x_o}} \cdot \frac{\|(T_{zw}w + T_{zx_o}x_o)(k) - (Qy)(k)\|}{\|(w,x_o)\|}$$

satisfies

$$\gamma(k) = \sup_{\substack{(w,x_o)\in B_{\ell^\infty}\times B_{\mathcal{R}^{n_x}} \\ 0=T_{yw}w+T_{yx_o}x_o}} \frac{\|(T_{zw}w + T_{zx_o}x_o)(k)\|}{\|(w,x_o)\|}.$$

Proposition 3.1 states that the cost of the uniformly optimal estimator (at any fixed k) is given by the worst case estimation error incurred for the measurement trajectory $y = 0$.

The present estimation problem considers nonzero initial conditions *and* time-varying dynamics. We will show that the set-valued observer in Section III-A defines a pointwise optimal estimator.

Definition 3.3: Consider the set-valued observer of Algorithm 3.1. Define

$$\underline{z}(y,k) = \min\{z\colon z \in Z(y,k)\}$$
$$\overline{z}(y,k) = \max\{z\colon z \in Z(y,k)\}$$
$$z_c(y,k) = \frac{\overline{z}(y,k) + \underline{z}(y,k)}{2}$$

where $Z(y,k) \overset{\text{def}}{=} H(k)X(y,k)$. The *central estimator*, $\Phi_c\colon \ell^\infty \to \ell^\infty$, is defined as

$$(\Phi_c y)(k) = z_c(y,k).$$

Our main result of this section is the following.

Theorem 3.2: The central estimator Φ_c is pointwise optimal.

Note that the central estimate is obviously the optimal for the unnormalized estimation error (3) (cf., [24]).

The remainder of this section is devoted to the proof of Theorem 3.2.

Since we are interested in pointwise optimality, we will consider a single "experiment," i.e., a fixed measurement trajectory y and estimation time k. This will simplify the presentation a great deal by dropping notational dependence on

y and k throughout. Thus, for this *fixed* measurement trajectory y and estimation time k we will use the following shorthand notation:

- $\underline{z}, \overline{z}$, and z_c—rather than $\underline{z}(y,k)$, $\overline{z}(y,k)$, and $z_c(y,k)$;
- z_Q—rather than $(Qy)(k)$;
- γ—rather than $\gamma(k)$;

where Q is the uniformly optimal estimator as in Proposition 3.1 and $\gamma(k)$ is the associated cost at time k.

Define r: $[\underline{z}, \overline{z}] \to \mathcal{R}^+$ by

$$r(\hat{z}) = \min\{\|(w, x_o)\|: (w, x_o) \in W(y,k) \text{ and } \hat{z} = (T_{zw}w + T_{zx_o}x_o)(k).\}.$$

In other words, $r(\hat{z})$ is the size of the smallest exogenous signal/initial condition pair which can produce the measured output as well as the value $\hat{z}$.

Similarly, define ϕ: $[\underline{z}, \overline{z}] \to \mathcal{R}$ by

$$\phi(\hat{z}) = \frac{\hat{z} - z_c}{r(\hat{z})}.$$

Then the estimation error associated with z_c can be expressed alternatively as

$$\sup_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi(\hat{z})|$$

(compare to Definition 3.2).

Note that

$$r(\underline{z}) = r(\overline{z}) = 1.$$

This is a result of the underlying linear dynamics. More precisely, the exogenous signals/initial condition which produce either $\overline{z}$ or $\underline{z}$ are the result of an appropriate linear program. Thus the exogenous signals/initial conditions which achieve the extreme values $\overline{z}$ and $\underline{z}$ are actively constrained by $\|(w, x_o)\| \le 1$. As a result

$$\phi(\underline{z}) = -\phi(\overline{z}).$$

However, $\phi(\cdot)$ need not be a symmetric (odd) function. Furthermore, we see r and ϕ can be derived from appropriate minimum distance problems and are both *continuous* functions.

Claim 3.1: The following inequality holds:

$$\frac{\overline{z} - \underline{z}}{2} \le \gamma.$$

In case of equality, $z_c = z_Q$.

Proof: The uniformly optimal estimator satisfies

$$|\overline{z} - z_Q| \le \gamma r(\overline{z})$$

and

$$|\underline{z} - z_Q| \le \gamma r(\underline{z}).$$

Since $r(\overline{z}) = r(\underline{z}) = 1$, this leads to

$$\overline{z} - \underline{z} \le 2\gamma.$$

In the case of equality, $z_Q = z_c$ is necessary. For example if $z_Q < z_c$

$$\overline{z} - z_Q > \gamma r(\overline{z})$$

which is a contradiction. ∎

Claim 3.2: Suppose $z_c \le z_1 \le \overline{z}$. Suppose $\phi(z_1) \le \gamma$. Then $\phi(z_2) \ge \phi(z_1)$ for all $z_2 \ge z_1$.

Proof: Let (w_1, x_{o1}) produce z_1 with minimum norm, i.e., $(w_1, x_{o1}) \in W(y,k)$

$$\|(w_1, x_{o1})\| = r(z_1)$$

and

$$z_1 = (T_{zw}w_1 + T_{zx_o}x_{o1})(k).$$

Let $(w_*, x_{o^*}) \in B_{\ell^\infty} \times B_{\mathcal{R}^{n_x}}$ correspond to the worst case exogenous signal/initial condition pair for the uniformly optimal observer as in Proposition 3.1. That is

$$\gamma = \frac{|(T_{zw}w_* + T_{zx_o}x_{o^*})(k)|}{\|(w_*, x_{o^*})\|}$$

and

$$0 = (T_{yw}w_* + T_{yx_o}x_{o^*})(k).$$

Without loss of generality, assume that

$$z_* = (T_{zw}w_* + T_{zx_o}x_{o^*})(k) > 0.$$

Choose some $z_2 \ge z_1$. One way to produce z_2 is through

$$(w_2, x_{o2}) = (w_1, x_{o1}) + h(w_*, x_{o^*})$$

where h is appropriately scaled so that

$$hz_* = z_2 - z_1.$$

By construction, (w_2, x_{o2}) is consistent with the measured data. However, it may be that $\|(w_2, x_{o2})\| > 1$.

We now compare $\phi(z_2)$ and $\phi(z_1)$. First

$$\phi(z_2) = \frac{z_2 - z_c}{r(z_2)} \ge \frac{z_2 - z_c}{\|(w_2, x_{o2})\|} \ge \frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\|(w_*, x_{o^*})\|}.$$

Thus proving the claim can be achieved by testing whether

$$\frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\|(w_*, x_{o^*})\|} \ge \frac{z_1 - z_c}{\|(w_1, x_{o1})\|} = \phi(z_1).$$

Toward this end, we see that

$$\frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\|(w_*, x_{o^*})\|} \ge \frac{z_1 - z_c}{\|(w_1, x_{o1})\|}$$
$$\Leftrightarrow$$
$$(z_2 - z_1) + (z_1 - z_c))\|(w_1, x_{o1})\| \ge (z_1 - z_c)\left(\|(w_1, x_{o1})\| + h\|(w_*, x_{o^*})\|\right)$$
$$\Leftrightarrow$$
$$(z_2 - z_1)\|(w_1, x_{o1})\| \ge h(z_1 - z_c)\|(w_*, x_{o^*})\|$$
$$\Leftrightarrow$$
$$(z_2 - z_1)\|(w_1, x_{o1})\| \ge \frac{z_2 - z_1}{z_*}(z_1 - z_c)\|(w_*, x_{o^*})\|$$
$$\Leftrightarrow$$
$$\gamma = \frac{z_*}{\|(w_*, x_{o^*})\|} \ge \frac{z_1 - z_c}{\|(w_1, x_{o1})\|} = \phi(z_1).$$

Using the hypothesis

$$\phi(z_1) \le \gamma$$

completes the proof. ∎

Claim 3.3: Suppose $z_c \geq z_1 \geq \underline{z}$. Suppose $-\phi(z_1) \leq \gamma$. Then $\phi(z_2) \leq \phi(z_1)$ for all $z_2 \leq z_1$.

Proof: The proof is similar to the proof of Claim 3.2. ■

Claim 3.4: The function ϕ is monotonically nondecreasing over the interval $[\underline{z}, \overline{z}]$.

Proof: Claims 3.2 and 3.3 imply that ϕ is monotonic for all $\hat{z}$ such that $|\phi(\hat{z})| \leq \gamma$.

Note that $\phi(z_c) = 0$. Thus by continuity, ϕ is monotonic until $\phi(\hat{z}') = \gamma$ for some $\hat{z}' \in (\underline{z}, \overline{z})$. Assume that such a $\hat{z}'$ satisfies $\hat{z}' > z_c$. Similar arguments hold in case $\hat{z}' < z_c$. Since $\overline{z} \geq \hat{z}'$, Claim 3.2 implies $\phi(\overline{z}) \geq \gamma$, and hence $\phi(\underline{z}) \leq -\gamma$. Claim 3.1 then implies that actually

$$\phi(\overline{z}) = \gamma = -\phi(\underline{z})$$

and $z_c = z_Q$. Since $z_c = z_Q$

$$\phi(\hat{z}) = \frac{\hat{z} - z_Q}{r(\hat{z})} \leq \gamma.$$

Thus, if ever $\phi(\hat{z}') = \gamma$, then $\phi(\hat{z}) \leq \gamma$ for all $\hat{z} \in [\underline{z}, \overline{z}]$, which completes the proof. ■

The proof of Claim 3.4 shows that the function ϕ saturates at $\pm\gamma$ if it ever achieves these values. In this case, $z_c = z_Q$. Furthermore, monoticity implies that ϕ always achieves its extreme values at $\underline{z}$ and $\overline{z}$.

We can now show that z_c is the pointwise optimal estimate. The cost of an alternative estimate, z', may be expressed

$$\max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi'(\hat{z})|$$

where

$$\phi'(\hat{z}) = \frac{\hat{z} - z'}{r(\hat{z})}.$$

In case $z' < z_c$, then

$$\phi'(\overline{z}) > \phi(\overline{z}).$$

In case $z' > z_c$, then

$$\phi'(\underline{z}) < \phi(\underline{z}).$$

In either case

$$\max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi'(\hat{z})| > \max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi(\hat{z})|$$

which completes the proof of Theorem 3.2.

IV. Application to Disturbance Rejection

A. Controlled Invariance with Output Feedback

We will consider discrete-time systems of the form

$$\begin{aligned} x(k+1) &= Ax(k) + B_1 d(k) + B_2 u(k) \\ z(k) &= C_1 x(k) \\ y(k) &= C_2 x(k) + D_{21} n(k) \end{aligned} \tag{4}$$

with the additional dimensions being $u(k) \in \mathcal{R}^{n_u}$ and $z(k) \in \mathcal{R}^{n_z}$.

Let $w = \binom{d}{n}$. The following assumptions hold throughout Section IV. Additional special assumptions will be stated as needed.

Assumption 4.1:

1) The exogenous inputs satisfy $w \in B_{\ell^\infty}$.
2) The matrices B_1 and C_1 have rank n_x.
3) The pair $[A, C_2]$ is detectable.

The objective is to design a controller which maintains $\|z\| \leq \gamma$ in the presence of all $\|w\| \leq 1$ using only output feedback, y. This objective is related to ℓ^1 optimal control for linear systems [14].

This objective is stated more precisely as follows. We will say that a controller is any operator which maps a vector, $\hat{x}_o \in \mathcal{R}^{n_x}$, and output sequence, $\{y(0), y(1), y(2), \cdots\}$, into a control sequence, $\{u(0), u(1), u(2), \cdots\}$ in a causal manner. This relationship is denoted $u = \mathcal{K}[\hat{x}_o]y$. The vector $\hat{x}_o$ is used to initialize the controller and can be viewed as an approximate initial condition for (4).

We now state precisely our performance objective.

Definition 4.1: Let S and E be compact convex sets in $\mathcal{R}^{n_x}$ with $0 \in E \subset S$. A controller $\mathcal{K}$ achieves a performance of γ over the sets (S, E) if for any $\hat{x}_o \in S$ and any initial condition $x(0) \in (\hat{x}_o + E) \cap S$, all solutions to (4) satisfy

$$\|z\| \leq \gamma.$$

The set S represents a class of admissible initial conditions, while the set E represents uncertainty in the controller's knowledge of the initial condition.

In the following, we present a theoretical determination of whether *any* controller can achieve a performance of γ over sets (S, E) which are yet to be specified. The presentation here and in [30] and [31] follows the language of viability theory [1] for differential inclusions. However, similar methods have been used in a variety of different contexts including viability theory and differential inclusions [1], [2], [17], [27], [28], dynamic programming [3], [4], systems with control constraints [5], [6], [13], [18]–[21], construction of reachable sets [10], [11], and time-varying system analysis [8], [8], [29], as well as optimal disturbance rejection [7], [9], [16], [23].

For $\gamma > 0$ define K_γ as

$$K_\gamma = \{x: |C_1 x| \leq \gamma\}.$$

Assumption 4.1 assures that K_γ is bounded. Clearly for a controller $\mathcal{K}$ to achieve a performance of γ, it must assure that $x(k) \in K_\gamma$ always. However, this is only a necessary condition. Also required is that there *always* exists a control value $u(k)$ which assures $x(k+1) \in K_\gamma$ as well. Define the set-valued regulation map $R_{K_\gamma}: \mathcal{R}^{n_x} \rightsquigarrow \mathcal{R}^{n_u}$ as

$$R_{K_\gamma}(x) = \{u: Ax + B_1 d + B_2 u \in K_\gamma, \ \forall |d| \leq 1\}.$$

In words, the regulation map determines the set of control values which assure $x(k+1) \in K_\gamma$. In terms of the regulation map, achieving a performance of γ requires 1) $x(k) \in K_\gamma$; 2) $R_{K_\gamma}(x(k))$ is nonempty; and 3) there exists a $u(k)$ such that $x(k+1)$ has properties 1) and 2).

We see that achieving a performance of γ is essentially equivalent to maintaining controlled invariance within the set of states having the above properties 1) and 2). Reference [31] exploited these notions in the noise-free state feedback case to construct controllers which achieve a performance of

γ whenever possible. Briefly, the state equation portion of (4) was written as the difference inclusion

$$x(k+1) \in G(x(t), u(t)) \quad (5)$$

where $G: \mathcal{R}^{n_x} \times \mathcal{R}^{n_m} \rightsquigarrow \mathcal{R}^{n_x}$ is the set-valued map defined by

$$(x, u) \mapsto \{Ax + B_1 d + B_2 u: |d| \leq 1\}.$$

It was shown that a performance of γ is achievable if and only if $\text{CINV}(K_\gamma)$ is nonempty, where CINV is the controlled invariance kernel defined in Appendix A.

Now consider the case of noisy output feedback. Let the *set* of possible state values at time k be denoted $X(k)$, where the explicit dependence on the output measurements (as in Section III-A) and control inputs is suppressed. More demanding than the state feedback case, we now must find a single control value which "works" for all $x(k) \in X(k)$. In terms of the regulation map, $\cap_{x \in X(k)} R_{K_\gamma}(x)$ must be nonempty. Again, this is only a necessary condition. Similarly to the state-feedback case, we must assure that: 1) $X(k) \subset K_\gamma$; 2) $\cap_{x \in X(k)} R_{K_\gamma}(x)$ is nonempty; and 3) there exists a $u(k)$ such that $X(k+1)$ has properties 1) and 2).

This discussion reveals that achieving a performance of γ in the output feedback case also is equivalent to maintaining controlled invariance. But the invariance is now referring to all possible *sets* of states. The similarities between output feedback and state feedback become more apparent if we express the evolution of the *set* of possible state values as a controlled difference inclusion.

Toward this end, let Σ denote the complete metric space of all nonempty compact subsets of R^{n_x} equipped with the Hausdorff metric [25, p. 279]. Define the set-valued map $F: \Sigma \times \mathcal{R}^{n_u} \rightsquigarrow \Sigma$ as follows. Suppose the current set of possible state values is $X(k)$. Based on $X(k)$, let $Y(k+1)$ be the set of possible output measurements at time $k+1$. This set depends on the specific control input $u(k)$ and all possible disturbances and noises. Thus

$$\begin{aligned} Y(k+1) = \{y: y = C_2(Ax + B_1 d + B_2 u(k)) \\ + n \text{ for some } x \in X(k), |d| \leq 1, |n| \leq 1\}. \end{aligned}$$

As in Algorithm 3.1, the set of state values at time $k+1$ is given by

$$\begin{aligned} X(k+1) = \{x: x = A\tilde{x} + B_1 d + B_2 u(k) \text{ for some} \\ \tilde{x} \in X(k), |d| \leq 1\} \cap \tilde{X}(y(k+1)). \end{aligned}$$

Define F to be the set-valued mapping

$$\begin{aligned} &F(X(k), u(k)) \\ &\quad = \{\{x: x = A\tilde{x} + B_1 d + B_2 u(k) \text{ for some} \\ &\qquad \tilde{x} \in X(k)|d| \leq 1\} \cap \tilde{X}(y(k+1)): \\ &\qquad y(k+1) \in Y(k+1)\}. \end{aligned}$$

In words, $F(X(k), u(k))$ represents the set of candidates for $X(k+1)$ based on $X(k)$ and $u(k)$. Thus an *element* of $F(X(k), u(k))$ is a *set* of possible states.

With this definition, we now may describe the system under output feedback by the controlled difference inclusion

$$X(k+1) \in F(X(k), u(k)). \quad (6)$$

Now let $\tilde{\Sigma}_\gamma \subset \Sigma$ denote the subsets of $\mathcal{R}^{n_x}$ which satisfy the following conditions. A set X belongs to $\tilde{\Sigma}_\gamma$ if

1) $X \subset K_\gamma$;
2) $\cap_{x \in X} R_{K_\gamma}(x)$ is nonempty.

We see that in order to achieve a performance of γ, a controller must assure that $X(k) \in \tilde{\Sigma}_\gamma$ always. Thus the original problem of controlled invariance for the state dynamics is transformed to a problem of controlled invariance for the difference inclusion in (6).

The following separation structure is an immediate consequence of this alternative interpretation of disturbance rejection. Let the term *separation structure controller* refer to a controller such that

$$u(k) = g(X(k))$$

where $g: \Sigma \to \mathcal{R}^m$ is a static nonlinear function on the current *set* of possible states $X(k)$.

Theorem 4.1: If any controller achieves a performance of γ over specified sets (S, E), then there exists a separation structure controller which achieves a performance of γ over the sets (S, E).

Proof: Let $\mathcal{K}$ be any controller which achieves a performance of γ, and suppose we constructed a set-valued observer for the system (4) under the *a priori* assumptions of: 1) known bounds on w; 2) known initial condition set $(\hat{x}_o + E) \cap S$; and 3) known control trajectory, u. (Note that the set-valued observer algorithms of Section III-A can easily be modified to incorporate alternate initial condition sets and known inputs.) Then each exogenous input trajectory leads to a trajectory of observed sets of possible states. Let this relation be denoted by $T_{Xw}(\mathcal{K}[\hat{x}_o], x(0))$.

Now let $\Sigma^* \subset \Sigma$ denote the set of reachable sets of states starting from any $\hat{x}_o \in S$ and $x(0) \in (\hat{x}_o + E) \cap S$. Then $X^* \in \Sigma^*$ if and only if

$$X^* = (T_{Xw}(\mathcal{K}[\hat{x}_o], x(0))w)(k)$$

for some $w \in B_{\ell^\infty}$, $k \in \mathcal{Z}^+$, $\hat{x}_o \in S$, and $x(0) \in (\hat{x}_o + E) \cap S$. Clearly Σ^* is a controlled invariant set for the difference inclusion (6). Furthermore, since $\mathcal{K}$ achieves a performance of γ over (S, E), we have that: 1) $\Sigma^* \subset \tilde{\Sigma}_\gamma$ and 2) $(\hat{x}_o + E) \cap S \in \tilde{\Sigma}_\gamma$. Thus for any $X^* \in \Sigma^*$: 1) $X^* \subset K_\gamma$ and 2) $\cap_{x \in X^*} R_{K_\gamma}(x)$ is nonempty. Furthermore, by controlled invariance, there exists a u such that $F(X^*, u) \subset \Sigma^* \subset \tilde{\Sigma}_\gamma$. We may then define the following regulation map $R^*: \Sigma \rightsquigarrow \mathcal{R}^{n_u}$:

$$R^*(X) = \{u: F(X, u) \subset \Sigma^*\}.$$

This leads to a family of separation structure controllers which achieve the desired performance. The only requirement is that $u(k) \in R^*(X(k))$, e.g.,

$$g(X(k)) = \min\{|u|: u \in R^*(X(k))\}.$$

The existence of a minimum is assured since $X(k)$ is always a compact convex set. The definition of g for sets not in Σ^* is not important because of controlled invariance. ■

We do not attempt to derive any regularity properties, such as continuity, of the separation structure controller. Theorem 4.1 is a direct consequence of the *interpretation* of disturbance rejection with output feedback as controlled invariance for the difference inclusion (6), and hence is primarily of conceptual value.

The controlled invariance kernel algorithm of Appendix B can be used to construct theoretically an invariant set if one exists. However, it is believed that the set-valued mapping F is not lower semicontinuous (since matrix intersection is not lower semicontinuous), and hence this procedure may not lead to a closed invariant-set.

B. Special Cases

Theorem 4.1 does not provide a constructive solution to deriving a separation structure controller. However, there are two special cases for which an explicit construction is possible.

1) Full Control: In this section, we make the following restrictive assumptions.

Assumption 4.2:

1) B_2 is invertible.
2) $C_1 = I$.

The situation in which B_2 is invertible is referred to as "full control" since the number of controls equals the number of states.

We will need to define the sets

$$X_{uo}(k) = \{x(k)\colon w \in B_{\ell^\infty} \text{ with } x(0) = 0 \quad u(j) = y(j) = 0, j = 0, \cdots, k\}$$
$$X_{uo} = \overline{\bigcup_{k \in \mathcal{Z}^+} X_{uo}(k)}.$$

The set $X_{uo}(k)$ is the set of states which are reachable at time k from zero initial conditions while maintaining $w \in B_{\ell^\infty}$, $u = 0$, and $y = 0$. The set X_{uo} is the closure of the union of such sets. As seen previously (cf., Proposition 3.1), this set plays an important role in ℓ^∞-optimal estimation. In some sense, X_{uo} represents a set of unobservable states in the sense that the disturbance and noise may drive the state to anywhere in X_{uo} without providing the controller with any additional information. The detectability assumption assures that X_{uo} is bounded.

Let $v_c(k)$ denote the component-by-component central estimates of the vector $v(k)$. Thus, the set of possible state vectors has a central estimate of $x_c(k)$. This set, in turn, leads to a set of possible values for $Ax(k)$ with a central estimate of $(Ax)_c(k)$. Note that $(Ax)_c(k)$ generally *does not* equal $A(x_c(k))$. At any time k these sets of possible values are defined by: 1) the measurements up to time k; 2) the control inputs up to time k; and 3) the *a priori* disturbance and initial condition assumptions.

Theorem 4.2: Let

$$\gamma = \max_{1 \le i \le n_x} (\max_{|d| \le 1} |e_i^T B_1 d| + \max_{x \in X_{uo}} |e_i^T Ax|)$$

where e_i denotes the standard ith basis vector in $\mathcal{R}^{n_x}$. The controller

$$u(k) = -B_2^{-1}(Ax)_c(k)$$

achieves a performance of γ over the sets (X_{uo}, E), where E is arbitrary.

It is easy to see that the stated γ in Theorem 4.2 is the smallest possible performance level under output feedback. Therefore, the given controller is, in fact, optimal. This controller resembles an optimal estimate of the optimal state feedback control, $u(k) = -B_2^{-1}Ax(k)$. However, an optimal estimate is required for $(Ax)(k)$, rather than $x(k)$. This is not surprising since it is the value of $(Ax)(k)$ which actually determines the current state's effects on future trajectories.

We close this section with a proof of Theorem 4.2. The state dynamics with the above controller take the form

$$x(k+1) = (Ax(k) - (Ax)_c(k)) + B_1 d(k).$$

The desired performance is achieved if for any admissible trajectory, the state satisfies

$$|e_i^T(Ax(k) - (Ax)_c(k))| \le \max_{x \in X_{uo}} |e_i^T X_{uo} x|$$

for all $k \in \mathcal{Z}^+$. A slight modification of Proposition 3.1 to accommodate known inputs assures that the above bound is satisfied.

2) Scalar Control: As opposed to full control, we now consider the other extreme of a scalar control variable. In particular, we will state conditions which assure that the regulation map intersection over the set of possible states is always nonempty. In terms of Section IV-A, we can then explicitly construct a separation structure controller.

We start with the following special assumptions.

Assumption 3:

1) The control signal is scalar-valued.
2) There exists a compact $K \subset K_\gamma$ which is controlled invariant under full-state feedback.
3) The regulation map

$$R_K(x) = \{u\colon Ax + B_1 d + B_2 u \in K,\ \forall |d| \le 1\}$$

admits the representation

$$R_K(x) = \{u\colon \max_i -\beta_i + \alpha_i^T x \le u \le \min_j \beta_j + \alpha_j^T x\}. \tag{7}$$

for appropriate vectors $\alpha_i \in \mathcal{R}^{n_x}$ and scalars $\beta_i > 0$.

Condition 4.3-2 is clearly necessary for the existence of an output feedback controller which achieves the desired performance. Reference [31] shows that regulation maps generally take the above form.

The following theorem is derived in [32].

Theorem 4.3: Define the scalar parameters

$$s_i = \max\{\alpha_i^T x\colon x \in X_{uo}\} \tag{8}$$
$$a_i = \max\{\alpha_i^T x\colon x \in K\}. \tag{9}$$

There exists an output feedback controller which achieves a performance of $\gamma = 1$ over (K, X_{uo}) if and only if for all i, j and all $x \in K$

$$\beta_i - s_i \ge 0 \tag{10}$$
$$(\alpha_i^T - \alpha_j^T)x \le \max\{\beta_i - s_i, \beta_i - (a_i - \alpha_i^T x)\} + \max\{\beta_j - s_j, \beta_j - (a_j + \alpha_j^T x)\}. \tag{11}$$

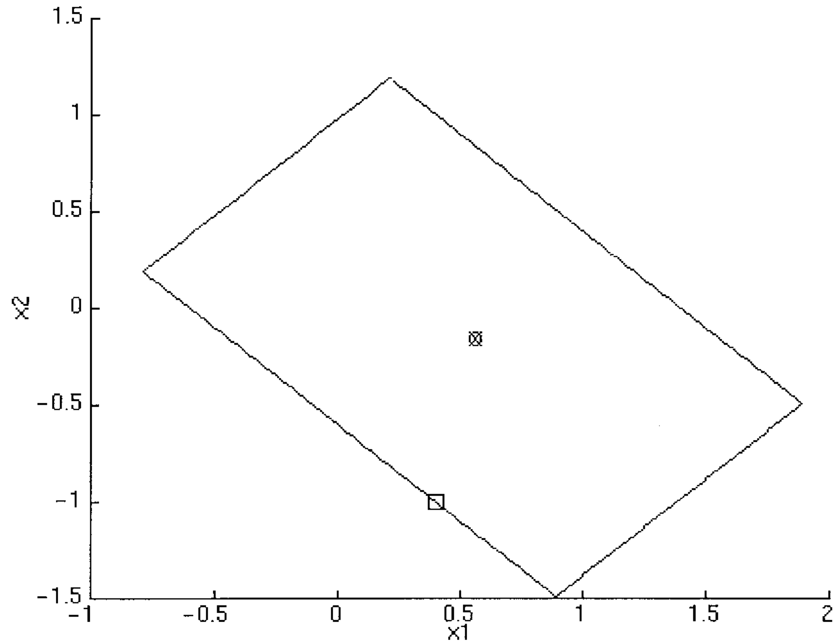


Fig. 1. Set of possible states at time $k = 2$: $* :=$ SVO, $\circ := \ell_1, \square := x_1$.

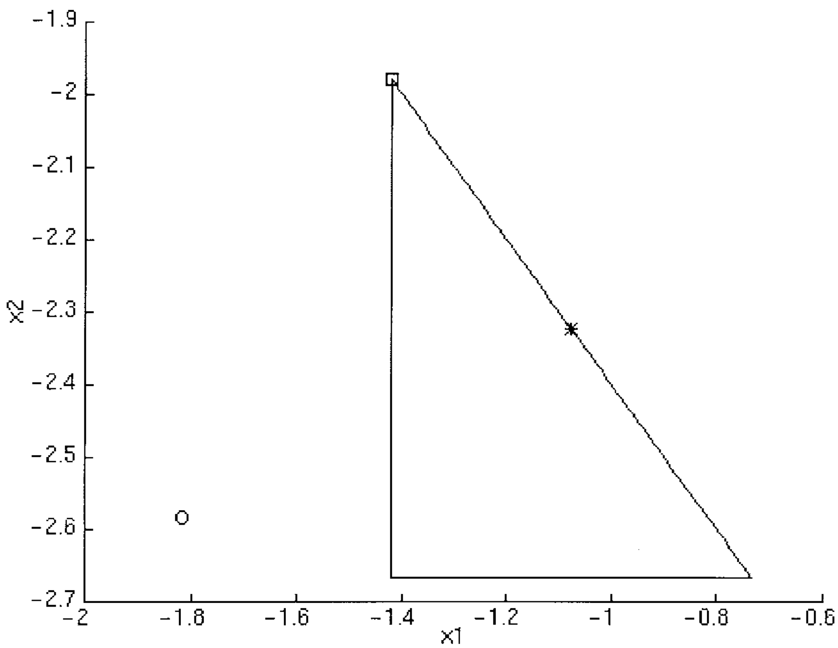


Fig. 2. Set of possible states at time $k = 3$: $* :=$ SVO, $\circ := \ell_1, \square := x_1$.

In terms of the discussion of Section IV-A, Theorem 4.3 provides conditions under which the intersection

$$\bigcap_{x \in X(k)} R_K(x)$$

is never empty. Therefore, a separation structure controller can achieve the desired performance with the static mapping $g(X(k))$ being any selection strategy from the above intersection. The conditions of Theorem 4.3 can be tested *a priori* by solving appropriate linear programs.

V. A NUMERICAL EXAMPLE

This section provides an illustrative numerical example of the set-valued observer. Let

$$x(k+1) = \begin{pmatrix} 0.7 & 0.7 \\ -0.7 & 0.7 \end{pmatrix} x(k) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} d(k)$$
$$y(k) = (1 \quad 1)x(k) + n(k).$$

We are interested in estimating the state x. An optimal estimate of the state amounts to optimal estimates of the individual components x_1 and x_2.

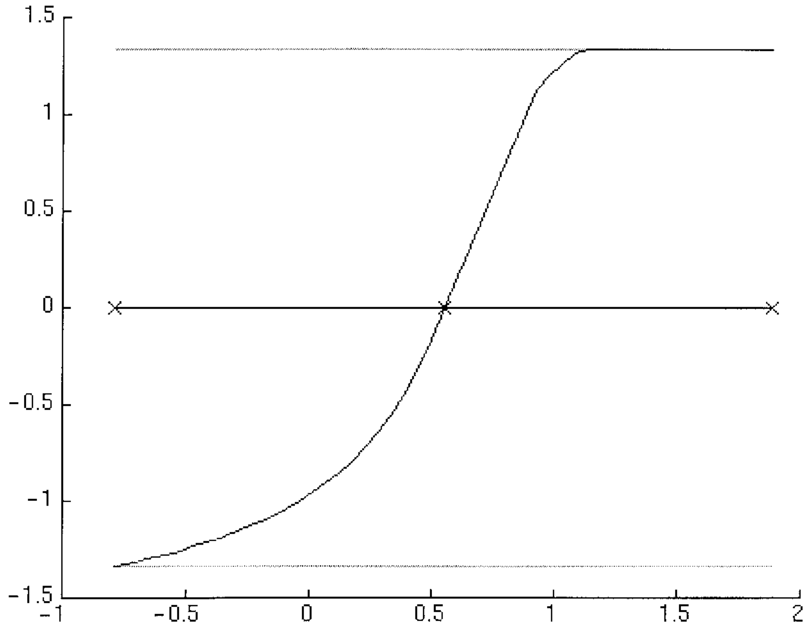


Fig. 3. Estimation error function ϕ for $z = x_1$ at time $k = 2$.

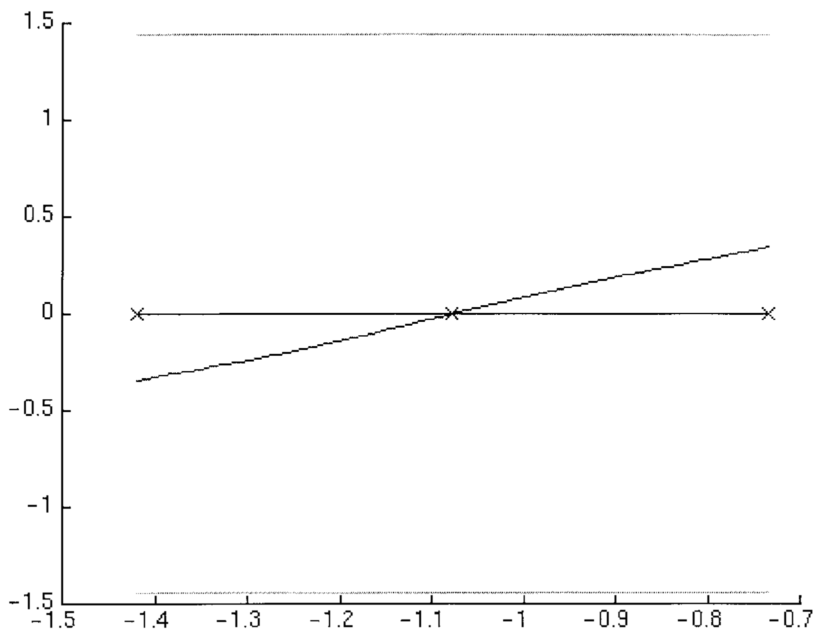


Fig. 4. Estimation error function, ϕ, for $z = x_1$ at time $k = 3$.

This simulation horizon was $k = 0, 1, 2, 3$. The disturbance and noise histories used in the simulation were

$$d(0,1,2,3) = (1,-1,-1,1)$$
$$n(0,1,2,3) = (1,-1,1,-1).$$

The true initial condition was set to $x(0) = 0$.

Figs. 1 and 2 show the set of admissible states at time $k = 2, 3$, respectively. Also shown are the true state, the central estimate, and the uniformly optimal estimate. Note that at time $k = 3$, the uniformly optimal estimate does not lie within the set of admissible states. This illustrates the pointwise optimality of the central estimate. Figs. 3 and 4 plot $\phi(\hat{z})$ at times $k = 2, 3$, respectively, for the estimate $z = x_1$. Note that ϕ is not symmetric, but is monotone as expected. Furthermore, ϕ for time $k = 2$ saturates at $\gamma(2)$ which implies that the central estimate equals the uniformly optimal estimate.

VI. Concluding Remarks

We have considered the guaranteed state estimation problem for discrete-time linear time-varying systems. Based on an *a priori* model of initial conditions and exogenous signals, a set-valued observer was constructed which computes the set of possible state vectors consistent with measured output data. It was shown that the centers of these sets correspond to the optimal state estimate which minimizes the induced norm from exogenous signals/initial conditions to estimation error. The algorithms easily can be modified in the case of known initial conditions and known inputs simply by changing the *a priori* assumptions.

We also considered the utility of set-valued observers for disturbance rejection with output feedback and derived a general, but conceptual, separation structure. An explicit construction is possible in the scalar control case. In the special case of full control, optimal output feedback controllers can resemble an optimal estimate of the full-state feedback controller.

While set-valued observers are of theoretical importance, their real-time applicability to systems with fast dynamics is questionable because of the considerable computational burden in constructing the set-valued estimates. An important research direction toward alleviating this burden is the derivation of fixed-complexity suboptimal set-valued estimates (cf., [24]).

APPENDIX
CONTROLLED INVARIANCE AND DIFFERENCE INCLUSIONS

In this Appendix, we present some material of independent interest regarding controlled difference inclusions and controlled invariance. The material essentially follows [31, Sec. IV], but with somewhat greater generality. The present discussion employs the language of viability theory. However, as mentioned in the main text, similar methods have been used in a variety of different contexts.

Let Σ be a complete metric space. Let $F: \Sigma \times \mathcal{R}^m \rightsquigarrow \Sigma$ be a set-valued mapping whose domain is the entire $\Sigma \times \mathcal{R}^m$. In this section, we consider the controlled difference inclusion

$$s(k+1) \in F(s(k), u(k)).$$

Definition A.1: A subset $\tilde{\Sigma} \subset \Sigma$ is *controlled invariant* if for every $s \in \tilde{\Sigma}$, there exists a $u \in \mathcal{R}^m$ such that $F(s,u) \subset \tilde{\Sigma}$.

Definition A.2: The largest closed subset of $\tilde{\Sigma} \subset \Sigma$ which is controlled invariant is the *controlled invariance kernel* of $\tilde{\Sigma}$ and is denoted $\text{C}\textsc{inv}(\tilde{\Sigma})$.

Define $\text{dom}(F) = \{x \in X: F(x) \text{ is nonempty}.$

Definition A.3 [1, p. 56]: Let X and Y be metric spaces. A set-valued map $F: X \rightsquigarrow Y$ is called *lower semicontinuous* if for any $x \in dom(F)$, $y \in F(x)$, and sequence $x_n \in dom(F)$ converging to x, there exists a sequence of elements $y_n \in F(x_n)$ converging to y.

Proposition A.1 (Controlled Invariance Kernel Algorithm): Suppose the set-valued mapping F satisfies the following.

1) F is lower-semicontinuous.
2) The set

$$\bigcup_n F(s_n, u_n)$$

is bounded if and only if the sequences $\{u_n\} \subset U$ and $\{s_n\} \subset \Sigma$ are bounded.

Let $\tilde{\Sigma} \subset \Sigma$ be compact, and define recursively the subsets $\tilde{\Sigma}_j$ by

$$\tilde{\Sigma}_0 = \tilde{\Sigma}$$
$$\tilde{\Sigma}_{j+1} = \{s \in \tilde{\Sigma}_j: F(s,u) \subset \tilde{\Sigma}_j \text{ for some } u \in \mathcal{R}^m\}.$$

Then

$$\text{C}\textsc{inv}(\tilde{\Sigma}) = \bigcap_{j=0}^{\infty} \tilde{\Sigma}_j.$$

Proof: We first show that if $\tilde{\Sigma}_j$ is closed, the $\tilde{\Sigma}_{j+1}$ is closed. Let $\{s_n\}$ be a sequence in $\tilde{\Sigma}_{j+1}$. Since $\tilde{\Sigma}_{j+1}$ is bounded, we may assume $\{s_n\}$ converges to some s_o. Let u_n be such that

$$F(s_n, u_n) \subset \tilde{\Sigma}_j.$$

The stated assumptions assure the sequence $\{u_n\}$ must be bounded. Therefore, we may assume that the sequence converges to some u_o. By lower semicontinuity, for any $r_o \in F(s_o, u_o)$, there exist $r_n \in F(s_n, u_n) \subset \tilde{\Sigma}_j$ such that the $\{r_n\}$ converge to r_o. Thus $r_o \in \tilde{\Sigma}_j$ since $\tilde{\Sigma}_j$ is closed. This implies $F(s_o, u_o) \subset \tilde{\Sigma}_j$, and hence $s_o \in \tilde{\Sigma}_{j+1}$.

Clearly $\text{C}\textsc{inv}(\tilde{\Sigma})$, if it exists, is contained in $\cap_{j=0}^{\infty} \tilde{\Sigma}_j$. Since the $\tilde{\Sigma}_j$ are nested compact sets, $\cap_{j=0}^{\infty} \tilde{\Sigma}_j$ is empty if and only if $\tilde{\Sigma}_j$ is empty for some j. In this case the proposition holds trivially.

In case $\cap_{j=0}^{\infty} \tilde{\Sigma}_j$ is nonempty, we will show it is controlled invariant. Define the set-valued regulation maps $R_j: \tilde{\Sigma}_j \rightsquigarrow \mathcal{R}^m$ by

$$R_j(s) = \{u: F(s,u) \subset \tilde{\Sigma}_{j-1}\}.$$

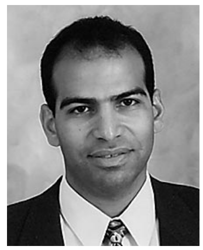
Similar arguments as above show that for any $s \in \cap_{j=0}^{\infty} \tilde{\Sigma}_j$, the $R_j(s)$ are nested compact sets. Therefore $\cap_{j=0}^{\infty} R_j(s)$ is nonempty for every $s \in \cap_{j=0}^{\infty} \tilde{\Sigma}_j$. Thus for any $s \in \cap_{j=0}^{\infty} \tilde{\Sigma}_j$, there exists a $u \in \cap_{j=0}^{\infty} R_j(s)$ such that $F(s,u) \subset \cap_{j=0}^{\infty} \tilde{\Sigma}_j$ which implies the desired controlled invariance. ■

In case F is not lower semicontinuous, the above algorithm still produces the largest invariant set. However, a largest *closed* invariant set may not exist.

REFERENCES

[1] J. P. Aubin, *Viability Theory*. Boston, MA: Birkhäuser, 1991.
[2] J. P. Aubin and A. Cellina, *Differential Inclusions*. New York: Springer-Verlag, 1984.
[3] D. P. Bertsekas, "Linear convex stochastic control problems over an infinite horizon," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 314–315, 1973.
[4] D. P. Bertsekas and I. B. Rhodes, "On the minimax reachability of target sets and target tubes," *Automatica*, vol. 7, pp. 233–247, 1971.
[5] G. Bitsoris and E. Gravalou, "Comparison principle, positive invariance and constrained regulation of nonlinear systems," *Automatica*, vol. 31, pp. 217–222, 1995.
[6] G. Bitsoris and M. Vassilaki, "Constrained regulation of linear systems," *Automatica*, vol. 31, pp. 223–229, 1995.
[7] F. Blanchini, "Feedback control for linear time-invariant systems with state and control bounds in the presence of disturbances," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1231–1234, Nov. 1990.
[8] ——, "Ultimate boundedness control for uncertain discrete-time systems via set-induced Lyapunov functions," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 428–433, Feb. 1994.
[9] F. Blanchini and M. F. Sznaier, "Persistent disturbance rejection via static-state state feedback," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1127–1131, June 1995.
[10] R. K. Brayton and C. H. Tong, "Stability of dynamical systems: A constructive approach," *IEEE Trans. Circuits and Syst.*, vol. CAS-26, no. 4, pp. 224–234, 1979.
[11] ——, "Constructive stability and asymptotic stability of dynamical systems," *IEEE Trans. Circuits and Syst.*, vol. CAS-27, no. 11, pp. 1121–1130, 1980.

[12] F. L. Chernousko, *State Estimation for Dynamic Systems*. Boca Raton, FL: CRC, 1994.
[13] M. Cwikel and P.-O. Gutman, "Convergence of an algorithm to find maximal state constraint sets for discrete-time linear dynamical systems with bounded controls and states," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 457–459, June 1986.
[14] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[15] I. J. Diaz-Bobillo and M. A. Dahleh, "State feedback ℓ^1-optimal controllers can be dynamic," *Syst. Contr. Lett.*, vol. 19, Feb. 1992.
[16] I. J. Fialho and T. T. Georgiou, "ℓ_1 state-feedback control with a prescribed rate of exponential convergence," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1476–1481, Oct. 1997.
[17] H. Frankowska and M. Quincampoix, "Viability kernels of differential inclusions with constraints: Algorithm and applications," *J. Math. Syst., Estimation, and Contr.*, vol. 1, no. 3, pp. 371–388, 1991.
[18] E. G. Gilbert and K. T. Tan, "Linear systems with state and control constraints: The theory and application of maximal output admissible sets," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1008–1020, Sept. 1991.
[19] P.-0. Gutman and M. Cwikel, "Admissible sets and feedback control for discrete-time linear dynamical systems with bounded controls and states," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 373–376, Apr. 1986.
[20] ——, "An algorithm to find maximal state constraint sets for discrete-time linear dynamical systems with bounded controls and states," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 251–254, Mar. 1987.
[21] S. S. Keerthi and E. G. Gilbert, "Computation of minimum-time feedback control laws for discrete-time systems with state-control constraints," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 432–435, May 1987.
[22] A. B. Kurzhanski and V. M. Veliov, Eds., *Modeling Techniques for Uncertain Systems*. Boston, MA: Birkhäuser, 1993.
[23] W.-M. Lu and J. C. Doyle, "Attenuation of persistent L_∞-bounded disturbances for nonlinear systems," preprint, 1995.
[24] M. Milanese and V. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, pp. 997–1009, 1991.
[25] J. R. Munkres, *Topology: A First Course*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.
[26] K. M. Nagpal and P. P. Khargonekar, "Filtering and smoothing in an $\mathcal{H}^\infty$ setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 152–166, 1991.
[27] M. Quincampoix, "An algorithm for invariance kernels of differential inclusions," in *Set-Valued Analysis and Differential Inclusions*, A. B. Kurzhanski and V. M. Veliov, Eds. Boston, MA: Birkhäuser, 1993, pp. 171–183.
[28] M. Quincampoix and P. Saint-Pierre, "An algorithm for viability kernels in Holderian case: Approximation by discrete dynamical systems," *J. Math. Syst., Estimation, and Contr.*, vol. 5, no. 1, pp. 1–13, 1995.
[29] E. D. Santis, "On positively invariant sets for discrete-time linear systems with disturbance: An application of maximal disturbance sets," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 245–249, Jan. 1994.
[30] J. S. Shamma, "Robust stability with time-varying structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 714–724, Apr. 1994.
[31] ——, "Optimization of the ℓ^∞-induced norm under full state feedback," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 533–544, Apr. 1996.
[32] J. S. Shamma and K.-Y. Tu, "Output feedback control for systems with constraints and saturations: Scalar control case," *Syst. Contr. Lett.*, 1998.
[33] P. Voulgaris, "On optimal ℓ_∞ to ℓ_∞ filtering," *Automatica*, vol. 31, no. 3, pp. 489–495, 1995.



Jeff S. Shamma (S'85–M'88) was born in New York, NY, in November 1963, and raised in Pensacola, FL. He received the Ph.D. degree in 1988 from the Department of Mechanical Engineering, Massachusetts Institute of Technology.

After one year of postdoctoral research, he joined the University of Minnesota, where he was an Assistant Professor of Electrical Engineering from 1989 to 1992. He then joined the University of Texas, Austin, where he is currently an Associate Professor of Aerospace Engineering. His research interests include robust control for linear parameter varying and nonlinear systems.

Dr. Shamma is a recipient of a 1992 NSF Young Investigator Award and the 1996 Donald P. Eckman Award of the American Automatic Control Council and was a Plenary Speaker at the 1996 American Control Conference. He has served on the editorial boards of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and *Systems & Control Letters*.

Kuang-Yang Tu (S'95–M'97) was born in Taipei, Taiwan, in December 1967. He received the B.S. degree in industrial engineering from National Chiao Tung University, Hsingchu, Taiwan, in 1989 and the M.S. and Ph.D. degrees in aerospace engineering from the University of Texas, Austin, in 1993 and 1997, respectively.

He is currently a Postdoctoral Researcher with the Flight Dynamics and Control Laboratory at the University of California, Irvine. His research interests include robust control, estimation, guidance control, and optimization.