Absolute Stability Analysis of Discrete-Time Systems with Composite Quadratic Lyapunov Functions [1]

Tingshu Hu[1] Zongli Lin[2]

[1] Department of Electrical & Computer Engineering, University of Massachusetts Lowell
Lowell, MA 01854 Email: tingshu@gmail.com
[2] Charles L. Brown Department of Electrical & Computer Engineering, University of Virginia, P.O. Box 400743
Charlottesville, VA 22904-4743 Email: zl5y@virginia.edu

Abstract

A generalized sector bounded by piecewise linear functions was introduced in [13] for the purpose of reducing conservatism in absolute stability analysis of systems with nonlinearity and/or uncertainty. This paper will further enhance absolute stability analysis by using the composite quadratic Lyapunov function whose level set is the convex hull of a family of ellipsoids. The absolute stability analysis will be approached by characterizing absolutely contractively invariant (ACI) level sets of the composite quadratic Lyapunov functions. This objective will be achieved through three steps. The first step transforms the problem of absolute stability analysis into one of stability analysis for an array of saturated linear systems. The second step establishes stability conditions for linear difference inclusions and then for saturated linear systems. The third step assembles all the conditions of stability for an array of saturated linear systems into a condition of absolute stability. Based on the conditions for absolute stability, optimization problems are formulated for the estimation of the stability region. Numerical examples demonstrate that stability analysis results based on composite quadratic Lyapunov functions improve significantly on what can be achieved with quadratic Lyapunov functions.

Key words: Absolute stability, piecewise linear sector, composite quadratic function, invariant set, saturation.

[1] Work supported in part by NSF under grant CMS-0324329.

1 Introduction

A discrete-time system with multiple nonlinear components is described as,

$$x^+ = Ax + B\psi(Fx, t), \tag{1}$$

where x and x^+ stand for $x(t)$ and $x(t+1)$, respectively, $A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times m}$ and $F \in \mathbf{R}^{m\times n}$ are constant matrices. And $\psi(\cdot,\cdot) : \mathbf{R}^m \times \mathbf{Z} \to \mathbf{R}^m$ represents the nonlinearities, possibly time-varying and uncertain. In the classical absolute stability theory, a nonlinear/uncertain/time-varying component is described with a conic sector. This description allows the nonlinear system practically accessible with tools originally developed for linear systems such as frequency analysis, robustness analysis and more recently, the LMI optimization technique (see, e.g., [1, 4, 7, 22, 24, 26, 27, 32, 34]). The conic sector takes into account both the nonlinearity and the possible time-varying uncertainty of the component but could be too conservative for a particular component for which more specific properties can be obtained such as an actuator with saturation or dead zone. For this reason, subclasses of the conic sector which impose additional restrictions on the derivative of $\psi(\cdot,\cdot)$ have been considered and less conservative conditions for absolute stability have been derived (see, e.g., [8, 9, 22, 27, 29].)

Also motivated by the objective of reducing the conservatism of stability analysis, we introduced a generalized sector in [13] for more flexible and more specific description of a nonlinear component. In contrast to using two straight lines to bound a conic sector, we use two odd symmetric piecewise linear functions that are convex or concave over $\mathbf{R}_{\geq 0}$ to bound an (uncertain) nonlinear function. To be specific, we will call the "generalized sector" the piecewise linear sector. Some common nonlinearities, such as saturation (or saturation-like) and dead zone functions can be exactly or arbitrarily closely described with a piecewise linear sector. When global absolute stability is out of the question or is unable to be confirmed, a region of absolute stability has to be estimated. In such a situation, a more detailed description of the nonlinearity by a piecewise linear sector would promise a larger estimate of the stability region than that by a conic sector.

In [13], the region of absolute stability is estimated with absolutely contractively invariant (ACI) ellipsoids and their convex hull for continuous-time systems with one nonlinear component. The objectives of maximizing ACI ellipsoids were formulated into LMI optimization problems. It was also established that if we have a group of ACI ellispoids, then their convex hull is also ACI. Because of this, larger estimates of the stability region can be produced. The main results in [13] were developed using quadratic Lyapunov functions, which have been extensively used for absolute stability and robustness analysis due to numerical issues.

While a piecewise linear sector promises a larger estimate of the stability region than that by the conic sector, the conservatism of absolute stability analysis can be further reduced by exploring nonquadratic Lyapunov functions. Apart from using the Lur'e type Lyapunov functions, an earlier attempt was made in [28] by combining several quadratic functions. Recent years have witnessed an extensive search for nonquadratic Lyapunov functions, among which are piecewise quadratic Lyapunov functions ([23, 25, 33]), polyhedral Lyapunov functions ([2, 5]), and homogeneous polynomial Lyapunov functions (HPLFs) ([6, 21]).

In [16], the composite quadratic Lyapunov function was introduced for enlarging the stability region of saturated linear systems. This type of functions were further explored in [10, 20] where the conditions of stability for linear differential inclusions and those for saturated linear systems were significantly improved.

The essential difference between [16] and [20] is that, in [16], the invariance of the convex hull of a family of ellipsoids is concluded from the invariance of each individual ellipsoid, while in [20], the invariance of each individual ellipsoid is no longer required. All the results in [10, 20] were developed for continuous-time systems. Their discrete-time counterparts will be established in this paper as basic tools for absolute stability analysis.

In this paper, we will consider a discrete-time system with multiple nonlinear components, each of which is bounded by a piecewise linear sector. We will use the composite quadratic Lyapunov functions to perform absolute stability analysis for such a system. The first step toward this goal (contained in Section 3) is to transform the absolute stability analysis problem into the stability analysis of an array of saturated linear systems. This is achieved by describing a piecewise linear function with an array of saturation functions. The second step (in Section 4) is to derive stability conditions for a saturated linear system by using composite quadratic functions. This is achieved by first developing a stability condition for linear difference inclusions (LDIs) and then obtaining a regional LDI description for the saturated linear system. The last step (in Section 5) is to put together the conditions of stability for all the saturated linear systems into the condition of absolute stability for the system with a piecewise linear sector condition. In Section 5, we also formulate optimization problems for enlarging the estimate of the stability region and use two examples to demonstrate the advantage of using composite quadratic functions over using quadratic functions. In particular, the estimate of the stability region for each case is significantly enlarged by using composite quadratic functions.

The problems in this paper are formulated under the discrete-time setting. However, all the results can be readily extended to continuous-time systems.

Notation:

- For two integers $k_1, k_2, k_1 < k_2$, we denote $I[k_1, k_2] = \{k_1, k_1 + 1, \cdots, k_2\}$.
- We use $\mathrm{sat}(\cdot)$ to denote the standard saturation function, i.e., for $u \in \mathbf{R}^m$, $[\mathrm{sat}(u)]_i = \mathrm{sgn}(u_i)\min\{1, |u_i|\}$.
- For N vectors $x^i \in \mathbf{R}^n, i \in I[1, N]$, we use $\mathrm{co}\{x^i : \ i \in I[1, N]\}$ to denote the convex hull of these vectors, i.e.,
$$\mathrm{co}\{x^i : \ i \in I[1, N]\} := \left\{ \sum_{i=1}^{N} \gamma_i x^i : \ \sum_{i=1}^{N} \gamma_i = 1, \ \gamma_i \geq 0 \right\}.$$
- For N functions $\psi^i : \mathbf{R}^n \to \mathbf{R}^m$, $i \in I[1, N]$, we use $\mathrm{co}\{\psi^i : \ i \in I[1, N]\}$ to denote the set of functions $\psi : \mathbf{R}^n \to \mathbf{R}^m$ such that
$$\psi(x) \in \mathrm{co}\{\psi^i(x) : \ i \in I[1, N]\} \qquad \forall\, x \in \mathbf{R}^n.$$
- For a matrix $F \in \mathbf{R}^{m \times n}$, denote
$$\mathcal{L}(F) := \left\{ x \in \mathbf{R}^n : \ |Fx|_\infty \leq 1 \right\}.$$
- Let $P = P^{\mathrm{T}} \in \mathbf{R}^{n \times n}$ be a positive-definite matrix. For a positive number ρ, denote
$$\mathcal{E}(P, \rho) := \left\{ x \in \mathbf{R}^n : x^{\mathrm{T}} P x \leq \rho \right\}.$$
For simplicity, we use $\mathcal{E}(P)$ to denote $\mathcal{E}(P, 1)$.

2 The generalized sector and absolute stability

2.1 Concave and convex functions

The generalized sector, as introduced in [13], is defined in terms of two concave/convex functions. Given a scalar function $\phi : \mathbf{R} \to \mathbf{R}$. Assume that

1) $\phi(\cdot)$ is continuous, piecewise differentiable and $\phi(0) = 0$.

2) $\phi(\cdot)$ is odd symmetric, i.e., $\phi(-u) = -\phi(u)$.

A function $\phi(u)$ satisfying the above assumption is said to be *concave* if it is concave for $u \geq 0$ and is said to be *convex* if it is convex for $u \geq 0$. These definitions are made for simplicity. It should be understood by odd symmetry that a concave function is convex for $u \leq 0$ and a convex function is concave for $u \leq 0$.

2.2 The generalized sector and absolute stability: definitions

Consider the system

$$x^+ = Ax + B\psi(Fx, t), \tag{2}$$

where $A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times m}, F \in \mathbf{R}^{m\times n}$ are given matrices, and $\psi(\cdot,\cdot) : \mathbf{R}^m \times \mathbf{Z} \to \mathbf{R}^m$ is a decoupled vector function, i.e.,

$$\psi(u, t) = [\phi_1(u_1, t) \;\; \phi_2(u_2, t) \;\; \cdots \;\; \phi_m(u_m, t)]^{\mathrm{T}},$$

and $\phi_i(\cdot,\cdot) : \mathbf{R} \times \mathbf{Z} \to \mathbf{R}$. Our objective is to estimate the stability region of (2) with invariant level sets of a certain Lyapunov function.

Let $V : \mathbf{R}^n \to \mathbf{R}$ be a positive definite Lyapunov function candidate. Given a positive number ρ, a level set of V is

$$L_V(\rho) := \{x \in \mathbf{R}^n : \;\; V(x) \leq \rho\}.$$

The level set $L_V(\rho)$ is said to be *contractively invariant* for (2) if

$$\Delta V(x, t) = V(Ax + B\psi(Fx, t)) - V(x) < 0 \tag{3}$$

for all $x \in L_V(\rho) \setminus \{0\}$ and $t \in \mathbf{Z}$. Clearly, if $L_V(\rho)$ is contractively invariant, then it is inside the stability region. If $V(x)$ is a quadratic function $x^{\mathrm{T}}Px$, then $L_V(\rho) = \mathcal{E}(P, \rho)$.

In the above definition of contractive invariance, the nonlinear function $\psi(u, t)$ is assumed to be known. For practical reasons, we would like to study the invariance of a level set for a class of nonlinear functions, for example, a class of $\psi(u, t)$, every component of which is bounded by a pair of convex/concave nonlinear functions, i.e.,

$$\phi_i(u_i, t) \in \mathrm{co}\{\breve{\phi}_i(u_i), \bar{\phi}_i(u_i)\} \quad \forall\, u_i \in \mathbf{R},\, t \in \mathbf{Z}, \qquad i \in I[1, m], \tag{4}$$

where $\breve{\phi}_i(\cdot)$ and $\bar{\phi}_i(\cdot)$ are known scalar functions. We use $\mathrm{co}\{\breve{\phi}_i, \bar{\phi}_i\}$ to denote the generalized sector for the i-th component of $\psi(\cdot,\cdot)$. The multivariable sector for $\psi(\cdot,\cdot)$ is the convex hull of 2^m decoupled functions

$$\psi^j(u) = [\phi_1^j(u_1) \;\; \phi_2^j(u_2) \;\; \cdots \;\; \phi_m^j(u_m)]^{\mathrm{T}}, \quad j \in I[1, 2^m],$$

where $\phi_i^j = \bar{\phi}_i$ or $\breve{\phi}_i$. We denote this multivariable sector as $\text{co}\{\psi^j : \ j \in I[1, 2^m]\}$. In general case, we may also consider a sector as the convex hull of N decoupled functions ψ^j, $\text{co}\{\psi^j : \ j \in I[1, N]\}$. We say that $\psi(\cdot, \cdot)$ satisfies a generalized sector condition if

$$\psi(u, t) \in \text{co}\{\psi^j(u) : \ j \in I[1, N]\} \quad \forall\, u \in \mathbf{R}^m,\, t \in \mathbf{Z},$$

and denote it as $\psi \in \text{co}\{\psi^j : \ j \in I[1, N]\}$.

Definition 1 *A level set $L_V(\rho)$ is said to be absolutely contractively invariant (ACI) over the sector $\text{co}\{\psi^j : j \in I[1, N]\}$ if it is contractively invariant for (2) under every possible $\psi(\cdot, \cdot)$ satisfying the generalized sector condition.*

Clearly, if $L_V(\rho)$ is ACI, then every trajectory starting from it will converge to the origin under any $\psi(\cdot, \cdot)$ satisfying the generalized sector condition. Hence $L_V(\rho)$ is a region of absolute stability. We may use more general functions to define the boundary of a sector to capture more details about the nonlinear components. As was explained in [13], the reason that we have chosen concave/convex functions is for the simplicity and completeness of the results they lead to. If we use other more general functions as the boundaries, the condition for ACI may be hard to describe or numerically non-tractable. We may also choose asymmetric functions or even symmetric functions. We have settled on odd symmetric functions since we will be focusing on level sets which are symmetric about the origin. Let us next state a simple result.

Lemma 1 *Assume that the Lyapunov function $V(x)$ is convex. Given a level set $L_V(\rho_0)$ and a sector $\text{co}\{\psi^j : j \in I[1, N]\}$. $L_V(\rho_0)$ is ACI over the sector if and only if it is contractively invariant for*

$$x^+ = Ax + B\psi^j(Fx), \tag{5}$$

for every $j \in I[1, N]$.

Proof. The "only if" part is obvious. Now we assume that $L_V(\rho_0)$ is contractively invariant for (5) for every $j \in I[1, N]$. Consider any $\psi \in \text{co}\{\psi^j : \ j \in I[1, N]\}$. We need to show that $L_V(\rho_0)$ is contractively invariant for

$$x^+ = Ax + B\psi(Fx, t). \tag{6}$$

Let x be such that $V(x) = \rho \in (0, \rho_0]$. Since $L_V(\rho_0)$ is contractively invariant for (5), we have

$$V(Ax + B\psi^j(Fx)) < V(x) = \rho \qquad \forall j \in I[1, N].$$

Since V is a convex function and $\psi(Fx, t) \in \text{co}\{\psi^j(Fx) : \ j \in I[1, N]\}$, we have

$$V(Ax + B\psi(Fx, t)) \le \max\{V(Ax + B\psi^j(Fx)) : \ \ j \in I[1, N]\} < V(x),$$

which implies the contractive invariance of $L_V(\rho_0)$ for (6). □

A similar result exists for continuous-time systems, where the differentiability of V is required instead of the convexity in Lemma 1.

We will restrict our attention to the level sets of convex Lyapunov functions. By Lemma 1, the absolute contractive invariance of a level set is equivalent to its contractive invariance under every vertex function ψ^j. Since a nonlinear function can be well approximated with a piecewise linear function, we will focus our attention on the case where the components of ψ^j are piecewise linear convex/concave functions. This will make the stability analysis problems numerically approachable. In this case, we call the generalized sector a piecewise linear sector.

2.3 A class of piecewise linear functions

Consider the class of piecewise linear functions:

$$\phi(u) = \begin{cases} k_0 u, & \text{if } u \in [0, b_1], \\ k_1 u + c_1, & \text{if } u \in (b_1, b_2], \\ \quad \vdots & \\ k_N u + c_N, & \text{if } u \in (b_N, \infty), \end{cases} \tag{7}$$

where $0 < b_1 < b_2 < \cdots < b_N$. The values of $\phi(u)$ for $u < 0$ can be determined by odd symmetry. It is easy to see that if $\phi(u)$ is concave, then $k_0 > k_1 > k_2 > \cdots > k_N > -\infty$ and $0 < c_1 < c_2 < \cdots < c_N$. In the case that $k_0 > 0$ and $k_N = 0$, $\phi(u)$ is a saturation like function with a saturation bound c_N. If $\phi(u)$ is convex, then $k_0 < k_1 < k_2 < \cdots < k_N < \infty$ and $0 > c_1 > c_2 > \cdots > c_N$. We note that $b_1, b_2, \cdots, b_N$ can be determined from $c_1, c_2, \cdots, c_N$ by the continuity of the function,

$$b_i = -\frac{c_i - c_{i-1}}{k_i - k_{i-1}}, \quad (c_0 = 0),$$

and vise versa. Fig. 1 plots a piecewise linear concave function with four bends.

3 Describing the system under sector condition with saturated linear systems

Lemma 1 transforms the problem of verifying the absolute contractive invariance of a level set into one of verifying the contractive invariance of the level set under individual vertex functions. This simplifies the problem to some degree but there is still no solution even if each component of the vertex functions is piecewise linear. The only situation that we are able to address is where each component of the vertex functions is a piecewise linear function with only one bend over $[0, \infty)$. In this case, every vertex function can be decomposed into a linear term and a standard saturation function and the corresponding vertex system is a saturated linear system, for which a set of analysis tools have been recently developed (see, e.g., [16, 18, 19, 20]). In this section, we use an array of saturation functions to describe a piecewise linear sector. By doing this, we transform the absolute stability problem into the stability analysis of an array of systems with saturation nonlinearities.

Let us first examine a scalar concave/convex piecewise linear function. The following lemma establishes a connection between a piecewise linear function and an array of saturation functions.

Lemma 2 *Consider a piecewise linear concave/convex function*

$$\phi(u) = \begin{cases} k_0 u, & \text{if } u \in [0, b_1], \\ k_1 u + c_1, & \text{if } u \in (b_1, b_2], \\ \quad \vdots & \\ k_N u + c_N, & \text{if } u \in (b_N, \infty). \end{cases} \tag{8}$$

For $j \in I[1, N]$, define

$$\phi^j(u) := k_j u + c_j \,\mathrm{sat}\left(\frac{k_0 - k_j}{c_j} u\right). \tag{9}$$

Then

$$\phi^j(u) \in \mathrm{co}\{k_0 u, \phi(u)\}, \tag{10}$$

$$\phi(u) \in \mathrm{co}\{\phi^j(u) : \ j \in I[1, N]\}. \tag{11}$$

Moreover, if ϕ is concave, then for $u \geq 0$,

$$\phi(u) \leq \phi^j(u) \leq k_0 u, \quad \phi(u) = \min\{\phi^j(u) : \ j \in I[1, N]\}. \tag{12}$$

If ϕ is convex, then for $u \geq 0$,

$$\phi(u) \geq \phi^j(u) \geq k_0 u, \quad \phi(u) = \max\{\phi^j(u) : \ j \in I[1, N]\}. \tag{13}$$

Proof. We only give the proof for the case where $\phi(u)$ is concave. The case where $\phi(u)$ is convex is similar. The proof is illustrated in Fig. 1, where the piecewise linear function in solid line is $\phi(u)$. Also plotted in the figure are $\phi^2(u)$ and the straight line $k_0 u$, both in dash-dotted lines.

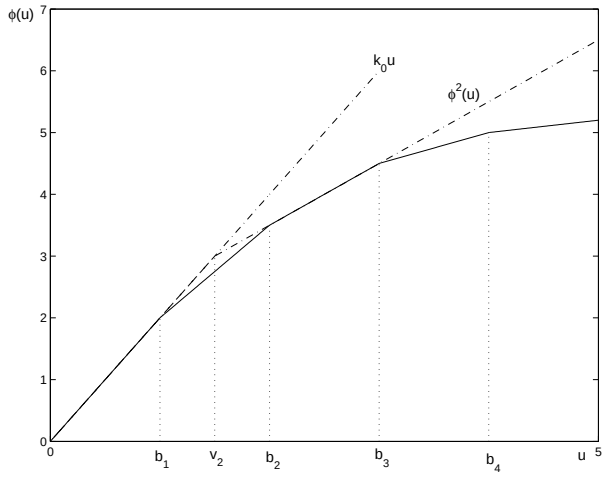

Figure 1: Illustration for the proof of Lemma 2

Let $v_j = \frac{c_j}{k_0 - k_j}$. Then $(v_j, k_0 v_j)$ is the intersection of the straight line $v = k_0 u$ with the other straight line obtained by extending the $(j+1)$th section of $\phi(u)$ to the left (see Fig. 1). It is clear that $0 < v_j \leq b_j$ ($v_1 = b_1$) and

$$\phi^j(u) = \begin{cases} k_0 u, & \text{if } u \in [0, v_j], \\ k_j u + c_j, & \text{if } u \in (v_j, \infty). \end{cases}$$

It follows that

$$\phi(u) = \begin{cases} \phi^1(u), & \text{if } u \in [0, b_2], \\ \quad \vdots & \\ \phi^j(u), & \text{if } u \in (b_j, b_{j+1}], \\ \quad \vdots & \\ \phi^N(u), & \text{if } u \in (b_N, \infty). \end{cases}$$

Hence,

$$\phi(u) \in \text{co}\{\phi^j(u) : \ j \in I[1, N]\}, \quad \forall\, u \in \mathbf{R}. \tag{14}$$

Since ϕ is concave, we have (see Fig. 1),

$$k_j u + c_j \geq \phi(u), \quad k_0 u \geq \phi(u) \qquad \forall u \geq 0$$

It follows that,

$$\phi(u) \leq \phi^j(u) \leq k_0 u, \qquad \forall u \geq 0, j \in I[1, N], \tag{15}$$

which implies (10). From (14), we have

$$\phi(u) \geq \min\{\phi^j(u) : \ j \in I[1, N]\}.$$

From (15), we have

$$\phi(u) \leq \min\{\phi^j(u) : \ j \in I[1, N]\}.$$

Thus (12) is verified. □

Here we note that relations equivalent to (10) and (11) were contained in the proof of Theorem 2 in [13]. Applying Lemma 2, we can use a new set of vertex functions to replace the original piecewise linear vertex functions. The new set of vertex functions have components of the form ϕ^j as defined in (9). Let us first consider the scalar case where $\phi(\cdot, \cdot) : \mathbf{R} \times \mathbf{Z} \to \mathbf{R}$ belongs to the sector $\text{co}\{\bar{\phi}, \breve{\phi}\}$ with $\bar{\phi}$ and $\breve{\phi}$ being piecewise linear concave/convex functions

$$\bar{\phi}(u) = \begin{cases} k_{01} u, & \text{if } u \in [0, b_{11}], \\ k_{11} u + c_{11}, & \text{if } u \in (b_{11}, b_{21}], \\ \quad \vdots & \\ k_{N_1 1} u + c_{N_1 1}, & \text{if } u \in (b_{N_1}, \infty). \end{cases} \tag{16}$$

and

$$\breve{\phi}(u) = \begin{cases} k_{02} u, & \text{if } u \in [0, b_{12}], \\ k_{12} u + c_{12}, & \text{if } u \in (b_{12}, b_{22}], \\ \quad \vdots & \\ k_{N_2 2} u + c_{N_2 2}, & \text{if } u \in (b_{N_2}, \infty). \end{cases} \tag{17}$$

Define

$$\bar{\phi}^\ell(u) := k_{\ell 1} u + c_{\ell 1} \,\text{sat}\left(\frac{k_{01} - k_{\ell 1}}{c_{\ell 1}} u\right), \quad \ell \in I[1, N_1] \tag{18}$$

$$\breve{\phi}^j(u) := k_{j2} u + c_{j2} \,\text{sat}\left(\frac{k_{02} - k_{j2}}{c_{j2}} u\right), \quad j \in I[1, N_2]. \tag{19}$$

Then by (12) and (13) of Lemma 2,

$$\mathrm{co}\{\bar{\phi}, \breve{\phi}\} \subset \mathrm{co}\{\bar{\phi}^\ell, \breve{\phi}^j : \ \ell \in I[1, N_1], j \in I[1, N_2]\}. \tag{20}$$

Note that the sector on the right hand side has vertex functions as the sum of linear functions and saturation functions.

We intend to use the sector on the right hand side of (20) to replace the one on the left hand side so that we can use stability analysis tools available for systems with saturation nonlinearities. However, the difference between the two sectors may introduce conservatism. The following lemma lists several cases where the two sectors are the same.

Lemma 3 *Assume that $\bar{\phi}(u) \geq \breve{\phi}(u)$ for $u \geq 0$. For the following cases,*

$$\mathrm{co}\{\bar{\phi}, \breve{\phi}\} = \mathrm{co}\{\bar{\phi}^\ell, \breve{\phi}^j : \ \ell \in I[1, N_1], j \in I[1, N_2]\}. \tag{21}$$

1. *$\bar{\phi}$ is convex and $\breve{\phi}$ is concave;*
2. *$\bar{\phi}$ or $\breve{\phi}$ is linear;*
3. *Both $\bar{\phi}$ and $\breve{\phi}$ are convex, $\breve{\phi}$ has only one bend and $\breve{\phi} \leq \bar{\phi}^\ell$ for all $\ell \in I[1, N_1]$;*
4. *Both $\bar{\phi}$ and $\breve{\phi}$ are concave, $\bar{\phi}$ has only one bend and $\bar{\phi} \geq \breve{\phi}^j$ for all $j \in I[1, N_2]$.*

The four cases in Lemma 3 are plotted in Fig. 2.

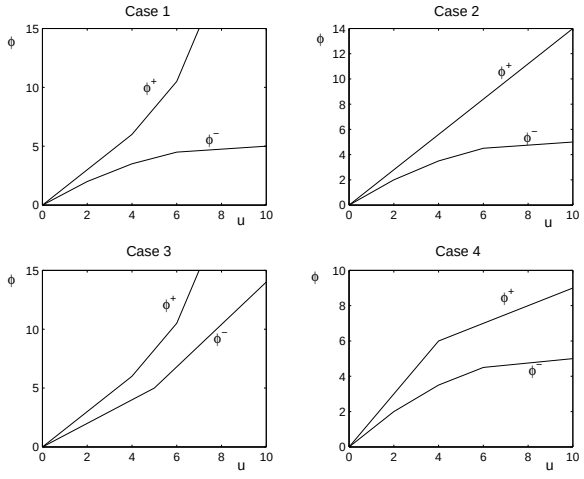

Figure 2: The four cases of Lemma 3: $\bar{\phi} = \phi^+, \breve{\phi} = \phi^-$

Proof. Let $\bar{\phi}$ and $\breve{\phi}$ be in the form of (16) and (17). Then $k_{01} \geq k_{02}$. For Case 1, it follows from (12) and (13) of Lemma 2 that for all $u \geq 0$,

$$\bar{\phi}(u) \geq \bar{\phi}^\ell(u) \geq k_{01}u \geq k_{02}u \geq \breve{\phi}^j(u) \geq \breve{\phi}(u), \quad \ell \in I[1, N_1], j \in I[1, N_2]. \tag{22}$$

Hence we have

$$\bar{\phi}^\ell, \breve{\phi}^j \in \mathrm{co}\{\bar{\phi}, \breve{\phi}\} \qquad \forall\, \ell \in I[1, N_1], j \in I[1, N_2]. \tag{23}$$

and (21) is obtained.

For Case 2, if $\bar{\phi}$ is linear, then $N_1 = 1$ and

$$\bar{\phi}(u) = \bar{\phi}^{\ell}(u) = k_{01}u \geq k_{02}u \geq \breve{\phi}^{j}(u) \geq \breve{\phi}(u), \quad \ell \in I[1, N_1], j \in I[1, N_2], \tag{24}$$

and we also obtain (23) and (21). If $\breve{\phi}$ is linear, the argument is similar.

For Case 3, we have $N_2 = 1$ and

$$\bar{\phi}(u) \geq \bar{\phi}^{\ell}(u) \geq \breve{\phi}^{j}(u) = \breve{\phi}(u), \quad \ell \in I[1, N_1], j \in I[1, N_2]. \tag{25}$$

For Case 4, we have $N_1 = 1$ and

$$\bar{\phi}(u) = \bar{\phi}^{\ell}(u) \geq \breve{\phi}^{j}(u) \geq \breve{\phi}(u), \quad \ell \in I[1, N_1], j \in I[1, N_2]. \tag{26}$$

Both (25) and (26) imply (23) and hence (21). □

Now we consider the case where $m > 1$. Assume that the i-th component $\phi_i(u_i, t)$ belongs to the sector $\mathrm{co}\{\bar{\phi}_i, \breve{\phi}_i\}$. Then the piecewise linear sector for $\psi(u, t)$ is

$$\mathrm{co}\{\psi^j : \ j \in I[1, 2^m]\}$$

where the i-th component of $\psi^j = \bar{\phi}_i$ or $\breve{\phi}_i$. Using Lemma 2 for each i, we can find N_i functions of the form

$$\phi_i^{\ell}(u_i) = \pi_{i\ell}u_i + \omega_{i\ell}\,\mathrm{sat}(\gamma_{i\ell}u_i), \quad \ell \in I[1, N_i], \tag{27}$$

such that

$$\mathrm{co}\{\bar{\phi}_i, \breve{\phi}_i\} \subset \mathrm{co}\{\phi_i^{\ell} : \ \ell \in I[1, N_i]\}. \tag{28}$$

Hence we have

$$\mathrm{co}\{\psi^j : \ j \in [1, 2^m]\} \subset \mathrm{co}\{[\phi_1^{\ell_1} \ \phi_2^{\ell_2} \ \cdots \ \phi_m^{\ell_m}]^{\mathrm{T}} : \ \ell_i \in I[1, N_i], i \in I[1, m]\}, \tag{29}$$

where the sector on the right hand side has $K := N_1 \times N_2 \times \cdots \times N_m$ vertices. If we assign a number $q \in I[1, K]$ to each vertex corresponding to $(\ell_1, \ell_2, \cdots, \ell_m)$, then each of the vertices has the form

$$[\phi_1^{\ell_1}(u_1) \ \phi_2^{\ell_2}(u_2) \ \cdots \ \phi_m^{\ell_m}(u_m)]^{\mathrm{T}} = \Pi_q u + \Omega_q\,\mathrm{sat}(\Gamma_q u), \quad q \in I[1, K], \tag{30}$$

where

$$\begin{aligned}
\Pi_q &= \mathrm{diag}\{\pi_{1\ell_1}, \pi_{2\ell_2}, \cdots, \pi_{m\ell_m}\}, \\
\Omega_q &= \mathrm{diag}\{\omega_{1\ell_1}, \omega_{2\ell_2}, \cdots, \omega_{m\ell_m}\}, \\
\Gamma_q &= \mathrm{diag}\{\gamma_{1\ell_1}, \gamma_{2\ell_2}, \cdots, \gamma_{m\ell_m}\}.
\end{aligned}$$

Replacing $\psi(u, t)$ of system (2) with a vertex function of the form (30), and letting $A_q = A + B\Pi_q F, B_q = B\Omega_q$ and $F_q = \Gamma_q F$, we obtain

$$x^+ = A_q x + B_q\,\mathrm{sat}(F_q x), \quad q \in I[1, K]. \tag{31}$$

It then follows from Lemma 1 that if a convex level set is contractively invariant for each of the systems in (31), then it is absolutely contractively invariant for system (2) under the sector condition $\phi_i \in \mathrm{co}\{\bar{\phi}_i, \breve{\phi}_i\}$ for each $i \in I[1, m]$.

4 Analysis of saturated linear systems with composite quadratic functions

In Sections 2 and 3, we converted the problem of verifying the ACI of a convex set into one of checking its contractive invariance for an array of saturated linear systems (31). In this section, we study the contractive invariance of the convex hull of a family of ellipsoids, which can be described as the level set of the composite quadratic Lyapunov function as introduced in [16]. In [16], such a level set was used to estimate the stability region for saturated continuous-time systems. In [20], the condition for such a level set to be contractively invariant was substantially relaxed through a stability condition for linear differential inclusions recently developed in [10]. In this section, we will derive a stability condition for linear difference inclusions by using the composite quadratic function and then establish the contractive invariance of its level set for saturated discrete-time systems. First we give a brief review of the composite quadratic function and its properties.

4.1 The composite quadratic Lyapunov function

Given a family of positive definite matrices $Q_j \in \mathbf{R}^{n\times n}, Q_j = Q_j^{\mathrm{T}} > 0, j \in I[1,J]$. Let

$$\Gamma = \left\{\gamma \in \mathbf{R}^J : \ \gamma_1 + \gamma_2 + \cdots + \gamma_N = 1, \gamma_j \geq 0\right\}.$$

The *composite quadratic function* is defined as

$$V_c(x) := \min_{\gamma\in\Gamma} x^{\mathrm{T}} \left(\sum_{j=1}^{J} \gamma_j Q_j\right)^{-1} x. \tag{32}$$

For simplicity, we say that V_c is composed from $Q_j, j \in I[1,J]$. When $J = 1$, V_c reduces to a quadratic function $x^{\mathrm{T}}Q_1^{-1}x$. It is evident that V_c is homogeneous of degree 2, i.e., $V_c(\alpha x) = \alpha^2 V_c(x)$. Also established in [10, 16] is that V_c is convex and continuously differentiable.

Through convex analysis, it was also shown in [17] that

$$\begin{aligned} V_c(x) &= \max\{c^{\mathrm{T}}xx^{\mathrm{T}}c : \ c \in \cap_{j=1}^{J}\mathcal{E}(Q_j)\} \\ &= \max\{c^{\mathrm{T}}xx^{\mathrm{T}}c : \ c^{\mathrm{T}}Q_j c \leq 1, j \in I[1,J]\}. \end{aligned} \tag{33}$$

For $\rho > 0$, the level set of V_c is

$$L_{V_c}(\rho) := \left\{x \in \mathbf{R}^n : \ V_c(x) \leq \rho\right\}.$$

It is clear from the definition that

$$V_c(x) \leq x^{\mathrm{T}}Q_j^{-1}x \qquad \forall j \in I[1,J]. \tag{34}$$

Hence $L_{V_c}(\rho) \supset \mathcal{E}(Q_j^{-1},\rho)$ for all $\rho > 0$ and $j \in I[1,J]$. Actually, from [16], we further have

$$L_{V_c}(\rho) = \mathrm{co}\left\{\mathcal{E}(Q_j^{-1},\rho) : j \in I[1,J]\right\},$$

where the right hand side denote the convex hull of the ellipsoids $\mathcal{E}(Q_j^{-1},\rho) : j \in I[1,J]$, i.e.,

$$\mathrm{co}\left\{\mathcal{E}(Q_j^{-1},\rho), j \in I[1,J]\right\} = \left\{\sum_{j=1}^{J}\gamma_j x_j : \ x_j \in \mathcal{E}(Q_j^{-1},\rho), \gamma \in \Gamma\right\}.$$

For a compact convex set S, a point x on the boundary of S is called an extreme point if it cannot be represented as the convex combination of any other points in S. As a result, an extreme point of $L_{V_c}(\rho)$ must be on the boundary of $\mathcal{E}(Q_j^{-1},\rho)$ for some $j \in I[1,J]$. In other words, if x is an extreme point of $L_{V_c}(\rho)$, we must have $V_c(x) = x^{\mathrm{T}}Q_j^{-1}x = \rho$ for some j.

4.2 Stability condition for linear difference inclusions

Consider the following linear difference inclusion

$$x^+ \in \mathrm{co}\{A_i x : \ i \in I[1,N]\}, \tag{35}$$

where $A_i \in \mathbf{R}^{n\times n}, i \in I[1,N]$, are given. Let V_c be composed from $Q_j = Q_j^T > 0, j \in I[1,J]$. The following is a discrete-time counterpart of the stability condition for linear differential inclusions in [10, 20]. Similarly to [10, 20], a dual result can also be established by using a function conjugate to V_c. For simplicity, we only present the following result.

Theorem 1 *If there exist $\lambda_{ijk} \geq 0, \sum_{k=1}^J \lambda_{ijk} = 1, i \in I[1,N], j,k \in I[1,J]$, such that*

$$A_i^T \left(\sum_{k=1}^J \lambda_{ijk} Q_k\right)^{-1} A_i < Q_j^{-1}, \quad i \in I[1,N], j \in I[1,J], \tag{36}$$

then $V_c(x^+) < V_c(x)$ for all $x \neq 0$. If $J = 2$, then (36) is a necessary condition.

Proof. By homogeneity of the LDI and V_c, $V_c(x^+) < V_c(x)$ for all $x \neq 0$ is equivalent to $V_c(x^+) < V_c(x)$ for all $x \in L_{V_c}(1) \setminus \{0\}$. By the convexity of V_c and linearity of $A_i x$, this is equivalent to $V_c(A_i x) < 1$ for all $x \in L_{V_c}(1) \setminus \{0\}$ and each i. Again by the convexity of V_c and linearity of $A_i x$, this is equivalent to $V_c(A_i x) < 1$ for every extreme point of $L_{V_c}(1)$. If x is an extreme point of $L_{V_c}(1)$, then we have $V_c(x) = x^{\mathrm{T}}Q_j^{-1}x = 1$ for some $j \in I[1,J]$. Therefore, $V_c(x^+) < V_c(x)$ for all $x \neq 0$ if and only if $V_c(A_i x) < 1$ for all $x \in \mathcal{E}(Q_j^{-1}), j \in I[1,J]$.

Under the condition (36) and by the definition of V_c, for $x \in \mathcal{E}(Q_j^{-1}) \setminus \{0\}$ we have

$$V_c(A_i x) \leq x^{\mathrm{T}} A_i^{\mathrm{T}} \left(\sum_{k=1}^J \lambda_{ijk} Q_k\right)^{-1} A_i x < x^{\mathrm{T}} Q_j^{-1} x \leq 1,$$

which confirms the sufficiency of the condition.

Before proving the necessity of the condition for the case $J = 2$, we need to show the following result.

Claim 1 *Given $R = R^T, R_1 = R_1^T, R_2 = R_2^T \in \mathbf{R}^{n\times n}$. We have*

$$c^T R c < \max\{c^T R_1 c, c^T R_2 c\} \quad \forall\, c \in \mathbf{R}^n \setminus \{0\}, \tag{37}$$

if and only if there exists $\alpha \in [0,1]$ such that $R < \alpha R_1 + (1-\alpha)R_2$.

The "if" part is obvious. We show "only if". We can break (37) into two inequalities,

$$c^{\mathrm{T}} R c < c^{\mathrm{T}} R_1 c, \ \text{ if } \ c^{\mathrm{T}}(R_2 - R_1)c \leq 0, \tag{38}$$

$$c^{\mathrm{T}} R c < c^{\mathrm{T}} R_2 c, \ \text{ if } \ c^{\mathrm{T}}(R_1 - R_2)c \leq 0. \tag{39}$$

By $\mathcal{S}$-procedure (e.g., see [4]), this implies the existence of $\alpha_1, \alpha_2 \geq 0$ such that

$$R < R_1 + \alpha_1(R_2 - R_1), \quad R < R_2 + \alpha_2(R_1 - R_2). \tag{40}$$

If either $\alpha_1 \leq 1$ or $\alpha_2 \leq 1$, then we are done. Now suppose $\alpha_1, \alpha_2 > 1$. Let $\delta \in [\frac{\alpha_2 - 1}{\alpha_1 + \alpha_2 - 1}, \frac{\alpha_2}{\alpha_1 + \alpha_2 - 1}]$. If we multiply the two inequalities in (40) with δ and $1 - \delta$ respectively and add them up, then we obtain $R < \alpha R_1 + (1 - \alpha)R_2$ with $\alpha = \delta(1 - \alpha_1) + (1 - \delta)\alpha_2 \in [0, 1]$. This proves Claim 1.

We now proceed to show the necessity of the condition for the case $J = 2$. It suffices to verify that $V_c(A_i x) < 1$ for all $x \in \mathcal{E}(Q_j^{-1}), j \in I[1, 2]$ leads to (36) for every i and j. By (33), $V_c(A_i x) < 1$ for all $x \in \mathcal{E}(Q_j^{-1})$ implies that

$$\max\{c^{\mathrm{T}} A_i x x^{\mathrm{T}} A_i^{\mathrm{T}} c : c^{\mathrm{T}} Q_1 c \leq 1, c^{\mathrm{T}} Q_2 c \leq 1, x^{\mathrm{T}} Q_j^{-1} x \leq 1\} < 1. \tag{41}$$

Observing that $\max\{c^{\mathrm{T}} A_i x x^{\mathrm{T}} A_i^{\mathrm{T}} c : x^{\mathrm{T}} Q_j^{-1} x \leq 1\} = c^{\mathrm{T}} A_i Q_j A_i^{\mathrm{T}} c$, we have

$$\max\{c^{\mathrm{T}} A_i Q_j A_i^{\mathrm{T}} c : c^{\mathrm{T}} Q_1 c \leq 1, c^{\mathrm{T}} Q_2 c \leq 1\} < 1. \tag{42}$$

It can be verified routinely that (42) is equivalent to

$$c^{\mathrm{T}} A_i Q_j A_i^{\mathrm{T}} c < \max\{c^{\mathrm{T}} Q_1 c, c^{\mathrm{T}} Q_2 c\} \quad \forall\, c \neq 0. \tag{43}$$

By Claim 1, this implies the existence of $\lambda_{ijk} \geq 0, k = 1, 2, \lambda_{ij1} + \lambda_{ij2} = 1$, such that

$$A_i Q_j A_i^{\mathrm{T}} < \sum_{k=1}^{2} \lambda_{ijk} Q_k \tag{44}$$

By Schur complement, the above is equivalent to

$$\begin{bmatrix} \sum_{k=1}^{2} \lambda_{ijk} Q_k & A_i \\ A_i^{\mathrm{T}} & Q_j^{-1} \end{bmatrix} > 0 \tag{45}$$

and then to (36). □

From the proof of Theorem 1, we see that condition (36) basically implies that V_c is a common Lyapunov function for all the vertex systems $x^+ = A_i x, i \in I[1, N]$. To determine if there exists such a V_c for a given J, we need to solve $N \times J$ matrix inequalities in (36) with variables $Q_j, j \in I[1, J]$ and $\lambda_{ijk}, i \in I[1, N], j, k \in I[1, J]$. What we have done is to minimize the number β such that there exist Q_j's and λ_{ijk}'s satisfying

$$A_i^{\mathrm{T}} \left(\sum_{k=1}^{J} \lambda_{ijk} Q_k \right)^{-1} A_i < \beta Q_j^{-1}, \quad i \in I[1, N], j \in I[1, J]. \tag{46}$$

or equivalently,

$$\begin{bmatrix} \sum_{k=1}^{J} \lambda_{ijk} Q_k & A_i Q_j \\ Q_j A_i^{\mathrm{T}} & \beta Q_j \end{bmatrix} > 0, \quad i \in I[1, N], j \in I[1, J]. \tag{47}$$

It fact, it can be shown similarly to the proof of Theorem 1 that if (47) is satisfied, then $V_c(x^+) < \beta V_c(x)$ for all $x \neq 0$. If the optimal $\beta \leq 1$, then the stability of the LDI is confirmed with V_c composed from the solution Q_j's.

Remark 1 *We note that (47) contains bilinear matrix inequalities (BMIs). A straightforward method to solve an optimization problem with BMI constraints is to break it up into a few problems with LMI constraints and solve them iteratively. However, this method does not work well on our particular problem. Some methods to address BMI constraints were developed in [3, 11, 12], etc. In our computation, we adopt the path-following method in [12] and it turned out to be very effective. We first solve the problem of minimizing β by assuming that all the Q_j's are equal to Q. This turns out to be a generalized eigenvalue problem. We then assign this optimal Q to all Q_j's and randomly pick initial values for λ_{ijk} so that $\sum_{k=1}^{J}\lambda_{ijk}=1$. The path-following algorithm is started with these Q_j's and λ_{ijk}'s.*

It is evident that as J is increased, the optimal β will decrease and the stability condition is less conservative. And as expected, the computation effort will be more intense.

The following simple example shows that, when applied to stability analysis of LDIs, the composite quadratic functions with $J=2$ can improve substantially on what can be achieved with quadratic functions.

Example 1 Consider the following LDI

$$x^+ \in \mathrm{co}\{A_1x, A_2(a)x\},$$

where

$$A_1=\begin{bmatrix} 0.4 & -0.4 \\ 0.4 & 0.4 \end{bmatrix}, \quad A_2(a)=\begin{bmatrix} 0.4 & -0.4/a \\ 0.4a & 0.4 \end{bmatrix},\ a\geq 1.$$

The maximal a such that there exists a common quadratic Lyapunov function for A_1 and $A_2(a)$ is $a_1=4.676$. By using Theorem 1 with $J=2$, the maximal a that guarantees the existence of a common Lyapunov function V_c is $a_2=7.546$.

4.3 Invariant level set for saturated linear systems

Consider the saturated system

$$x^+ = Ax + B\,\mathrm{sat}(Fx), \tag{48}$$

where $A\in\mathbf{R}^{n\times n}, B\in\mathbf{R}^{n\times m}$ and $F\in\mathbf{R}^{m\times n}$. We will use LDIs to describe this system within a local region of state space. This description is made possible with a tool from [18, 19].

Consider the set of $m\times m$ diagonal matrices whose diagonal elements are either 1 or 0. There are 2^m such matrices and we label them as D_i, $i\in I[1,2^m]$. Denote $D_i^- = I - D_i$. Given two vectors, $u,v\in\mathbf{R}^m$,

$$\left\{D_iu + D_i^-v:\ i\in[1,2^m]\right\}$$

is the set of vectors obtained by choosing some elements from u and the rest from v.

Lemma 4 *[19] Let $H\in\mathbf{R}^{m\times n}$ be given. Then for all $x\in\mathcal{L}(H)$,*

$$\mathrm{sat}(Fx)\in\mathrm{co}\{(D_iF+D_i^-H)x:\ i\in I[1,2^m]\}.$$

Let h_ℓ be the ℓ-th row of H. Given $Q=Q^T>0$. $\mathcal{E}(Q^{-1})\subset\mathcal{L}(H)$ if and only if $h_\ell Q h_\ell^T\leq 1$ for all $\ell\in I[1,m]$.

Let $Q_j = Q_j^{\mathrm{T}}, j \in I[1,J]$ be positive definite matrices and let V_c be defined as in (32). The following theorem gives a sufficient condition for the contractive invariance of $L_{V_c}(1)$ for (48).

Theorem 2 *If there exist an* $H \in \mathbf{R}^{m\times n}$ *and* $\lambda_{ijk} \geq 0$, $\sum_{k=1}^{J} \lambda_{ijk} = 1$, $i \in I[1,2^m], j,k \in I[1,J]$, *such that*

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \lambda_{ijk} Q_k & (A + B(D_i F + D_i^- H))Q_j \\ Q_j(A + B(D_i F + D_i^- H))^T & Q_j \end{array} \right] > 0, \;\; i \in I[1,2^m],\, j \in I[1,J], \tag{49}$$

$$h_\ell Q_j h_\ell^T \leq 1, \quad \ell \in I[1,m],\; j \in I[1,J], \tag{50}$$

where h_ℓ *is the* ℓ*-th row of* H. *Then* $L_{V_c}(1)$ *is contractively invariant, i.e.,*

$$V_c(x^+) < V_c(x) \quad \forall\, x \in L_{V_c}(1) \setminus \{0\}. \tag{51}$$

Proof. By Lemma 4, (50) implies that $\mathcal{E}(Q_j^{-1}) \subset \mathcal{L}(H)$ for all $j \in I[1,J]$. Since $L_{V_c}(1)$ is the convex hull of $\mathcal{E}(Q_j^{-1}), j \in I[1,J]$, it follows that $L_{V_c}(1) \subset \mathcal{L}(H)$. Also by Lemma 4, we have

$$\mathrm{sat}(Fx) \in \mathrm{co}\{(D_i F + D_i^- H)x :\, i \in I[1,2^m]\} \quad \forall\, x \in L_{V_c}(1).$$

If we let $A_i = A + B(D_i F + D_i^- H)$, then for all $x \in L_{V_c}(1)$,

$$x^+ = Ax + B\,\mathrm{sat}(Fx) \in \mathrm{co}\{A_i x :\; i \in I[1,2^m]\}.$$

By Schur complement, (49) implies

$$A_i^{\mathrm{T}} \left(\sum_{k=1}^{J} \lambda_{ijk} Q_k \right)^{-1} A_i < Q_j^{-1}, \quad i \in I[1,N], j \in I[1,J].$$

Hence by Theorem 1, we have

$$V_c(x^+) < V_c(x) \qquad \forall\, x \in L_{V_c}(1) \setminus \{0\}.$$

□

In Theorem 2, the condition for $L_{V_c}(1)$ to be contractively invariant involves $\mathcal{E}(Q_j^{-1}) \subset \mathcal{L}(H)$ for all $j \in I[1,J]$. For the special case $m = 1$, we can use different H for different $\mathcal{E}(Q_j^{-1})$. By doing this, the condition for stability can be relaxed significantly.

Theorem 3 *Assume that* $m = 1$. *If there exist* $H_j \in \mathbf{R}^{1\times n}$ *and* $\alpha_{jk} \geq 0$, $\beta_{jk} \geq 0$, $\sum_{k=1}^{J} \alpha_{jk} = 1$, $\sum_{k=1}^{J} \beta_{jk} = 1$, $j,k \in I[1,J]$ *such that*

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \alpha_{jk} Q_k & (A + BF)Q_j \\ Q_j(A + BF))^T & Q_j \end{array} \right] > 0,\; j \in I[1,J], \tag{52}$$

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \beta_{jk} Q_k & (A + BH_j)Q_j \\ Q_j(A + BH_j))^T & Q_j \end{array} \right] > 0,\; j \in I[1,J], \tag{53}$$

$$H_j Q_j H_j^T \leq 1, \;\; j \in I[1,J]. \tag{54}$$

Then $L_{V_c}(1)$ *is contractively invariant.*

To prove Theorem 3, we need a result which only holds for systems with one saturation component.

Lemma 5 *Suppose* $x_0 \in \mathrm{co}\{x_j : \ j \in I[1, J_0]\}$ *and* $Fx_j \geq 0$ *for all* $j \in I[1, J_0]$. *Then*

$$Ax_0 + B\mathrm{sat}(Fx_0) \in \mathrm{co}\{Ax_j + BFx_j, Ax_j + B\mathrm{sat}(Fx_j) : \ \ j \in I[1, J_0]\}. \tag{55}$$

Proof. To prove (55), it suffices to show that for any $c \in \mathbf{R}^n$,

$$c^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(Fx_0)) \in [\alpha_{\min}, \alpha_{\max}], \tag{56}$$

where

$$\begin{aligned}
\alpha_{\min} &= \min\{c^{\mathrm{T}}(Ax_j + BFx_j), c^{\mathrm{T}}(Ax_j + B\mathrm{sat}(Fx_j)) : \ \ j \in I[1, J_0]\}, \\
\alpha_{\max} &= \max\{c^{\mathrm{T}}(Ax_j + BFx_j), c^{\mathrm{T}}(Ax_j + B\mathrm{sat}(Fx_j)) : \ \ j \in I[1, J_0]\}.
\end{aligned}$$

If $c^{\mathrm{T}}B \geq 0$, then by the concavity of $\mathrm{sat}(u)$ for $u \geq 0$ and $Fx_0 \geq 0$, we have

$$c^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(Fx_0)) \leq c^{\mathrm{T}}(Ax_0 + BFx_0) \leq \max\{c^{\mathrm{T}}(Ax_j + BFx_j) : \ \ j \in I[1, J_0]\} \leq \alpha_{\max},$$

and

$$c^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(Fx_0)) \geq \min\{c^{\mathrm{T}}(Ax_j + B\mathrm{sat}(Fx_j)) : \ \ j \in I[1, J_0]\} \geq \alpha_{\min}.$$

Hence we obtain (56).

If $c^{\mathrm{T}}B \leq 0$, then by the convexity of $-\mathrm{sat}(u)$ for $u \geq 0$ and $Fx_0 \geq 0$,

$$c^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(Fx_0)) \geq c^{\mathrm{T}}(Ax_0 + BFx_0) \geq \min\{c^{\mathrm{T}}(Ax_j + BFx_j) : \ \ j \in I[1, J_0]\} \geq \alpha_{\min},$$

and

$$c^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(Fx_0)) \leq \max\{c^{\mathrm{T}}(Ax_j + B\mathrm{sat}(Fx_j)) : \ \ j \in I[1, J_0]\} \leq \alpha_{\max}.$$

We also have (56). □

Proof of Theorem 3. By Schur complement, conditions (52), (53) are equivalent to

$$(A + BF)^{\mathrm{T}} \left(\sum_{k=1}^{J} \alpha_{jk} Q_k \right)^{-1} (A + BF) < Q_j^{-1}, \quad j \in I[1, J] \tag{57}$$

$$(A + BH_j)^{\mathrm{T}} \left(\sum_{k=1}^{J} \beta_{jk} Q_k \right)^{-1} (A + BH_j) < Q_j^{-1}, \quad j \in I[1, J]. \tag{58}$$

It follows that for each $\rho > 0, j \in I[1, J]$,

$$V_c(Ax + BFx) \leq x^{\mathrm{T}}(A + BF)^{\mathrm{T}} \left(\sum_{k=1}^{J} \alpha_{jk} Q_k \right)^{-1} (A + BF)x < x^{\mathrm{T}} Q_j^{-1} x \leq \rho \quad \forall x \in \mathcal{E}(Q_j^{-1}, \rho) \setminus \{0\}. \tag{59}$$

and

$$V_c(Ax + BH_j x)) \leq x^{\mathrm{T}}(A + BH_j)^{\mathrm{T}} \left(\sum_{k=1}^{J} \beta_{jk} Q_k \right)^{-1} (A + BH_j)x < x^{\mathrm{T}} Q_j^{-1} x \leq \rho \quad \forall x \in \mathcal{E}(Q_j^{-1}, \rho) \setminus \{0\}. \tag{60}$$

By (54) we have $\mathcal{E}(Q_j^{-1}) \subset \mathcal{L}(H_j)$. Hence by Lemma 4 we have $Ax + \text{sat}(Fx) \in \{(A+BF)x, (A+BH_j)x\}$ for all $x \in \mathcal{E}(Q_j^{-1})$. It follows from (59), (60) and the convexity of V_c that for each $\rho \in (0,1]$, we have

$$V_c(Ax + B\text{sat}(Fx))) < \rho \quad \forall x \in \mathcal{E}(Q_j^{-1}, \rho) \setminus \{0\}. \tag{61}$$

To prove the theorem, it suffices to show that for any $\rho \in (0,1]$ and $x_0 \in \partial L_{V_c}(\rho)$,

$$V_c(Ax_0 + B\text{sat}(Fx_0)) < V_c(x_0) = \rho. \tag{62}$$

Here we only prove the case where $\rho = 1$. The proof for $\rho \in (0,1)$ follows from same arguments. Hence we need to show that for any $x_0 \in \partial L_{V_c}(1)$,

$$V_c(Ax_0 + B\text{sat}(Fx_0)) < 1. \tag{63}$$

Now we consider an arbitrary $x_0 \in \partial L_{V_c}(1)$. For simplicity, assume that $Fx_0 \geq 0$ (the proof for $Fx_0 \leq 0$ is similar). If $x_0 \in \partial\mathcal{E}(Q_j^{-1}) \cap \partial L_{V_c}(1)$ for some $j \in I[1,J]$, then by (61)

$$V_c(Ax_0 + B\text{sat}(Fx_0)) < 1.$$

Now we assume that $x_0 \notin \partial\mathcal{E}(Q_j^{-1})$ for any j. Since $L_{V_c}(1)$ is the convex hull of $\mathcal{E}(Q_j^{-1}, 1), j \in I[1,J]$, all the extreme points of $L_{V_c}(1)$ belong to the union of $\partial\mathcal{E}(Q_j^{-1}, 1), j \in I[1,J]$. Hence there exist an integer $J_0 \leq J$, some vectors $x_j \in \partial\mathcal{E}(Q_j^{-1}), j \in I[1,J_0]$, such that

$$x_0 \in \text{co}\{x_j : \ j \in I[1,J_0]\}.$$

(Here we have assumed for simplicity that x_0 is only related to the first J_0 ellipsoids. Otherwise, the ellipsoids can be reordered to meet this assumption.) From (61), we have

$$V_c(Ax_j + B\text{sat}(Fx_j)) < 1, \quad j \in I[1,J_0]. \tag{64}$$

and by (59)

$$V_c(Ax_j + BFx_j) < 1, \quad j \in I[1,J_0]. \tag{65}$$

We first consider the case where $Fx_j \geq 0$ for all $j \in I[1,J_0]$. By Lemma 5, we have

$$Ax_0 + B\text{sat}(Fx_0) \in \text{co}\{Ax_j + BFx_j, Ax_j + B\text{sat}(Fx_j) : \ j \in I[1,J_0]\}. \tag{66}$$

Then (63) follows from (64), (65), (66) and the convexity of V_c.

Next we consider the case where $Fx_j < 0$ for some $j \in I[1,J_0]$. Since $Fx_j \geq 0$ does not hold for all $j \in I[1,J_0]$, we can get an intersection of the set $\text{co}\{x_1, x_2, \cdots, x_{J_0}\}$ with the half space $Fx \geq 0$. This intersection is also a polygon and can be denoted as $\text{co}\{y_1, y_2, \cdots, y_{J_1}\}$. Since $Fx_0 \geq 0$, we have $x_0 \in \text{co}\{y_1, y_2, \cdots, y_{J_1}\}$ and by Lemma 5, we have

$$Ax_0 + B\text{sat}(Fx_0) \in \text{co}\{Ay_j + BFy_j, Ay_j + B\text{sat}(Fy_j) : \ j \in I[1,J_1]\}. \tag{67}$$

We note that some y_j's belong to $\{x_1, x_2, \cdots, x_{J_0}\}$, others are not. For those $y_j \notin \{x_i : \ i \in I[1, J_0]\}$, we must have $Fy_j = 0$ and $y_j \in \text{co}\{x_i : \ i \in I[1, J_0]\}$. It follows from (65) and the convexity of V_c that for these y_j, we also have

$$V_c(Ay_j + BFy_j) < 1.$$

and

$$V_c(Ay_j + B\text{sat}(Fy_j)) = V_c(Ay_j + BFy_j) < 1.$$

In summary, we have

$$V_c(Ay_j + BFy_j) < 1, \quad V_c(Ay_j + B\text{sat}(Fy_j)) < 1, \quad j \in I[1, J_1]. \tag{68}$$

Combining (67), (68) and the convexity of V_c, we obtain (63). □

We finally consider the special case where both $m = 1$ and $J = 1$. In this case, V_c reduces to a quadratic function $x^{\text{T}}Q^{-1}x$ and the level set $L_{V_c}(1)$ reduces to an ellipsoid $\mathcal{E}(Q^{-1})$. Accordingly, the conditions in Theorem 3 can be transformed into LMIs by introducing the new variable $Y = HQ$. For continuous-time system, it was shown that the corresponding condition is necessary and sufficient for the contractive invariance of an ellipsoid [15]. The counterpart result for discrete-time system is established in Appendix A and summarized as follows.

Theorem 4 *Assume $m = 1$. Given $Q = Q^T > 0$. The ellipsoid $\mathcal{E}(Q^{-1})$ is contractively invariant for (48) if and only if there exists an $H \in \mathbf{R}^{1\times n}$ such that*

$$\begin{bmatrix} Q & (A+BF)Q \\ Q(A+BF)^T & Q \end{bmatrix} > 0, \tag{69}$$

$$\begin{bmatrix} Q & (A+BH)Q \\ Q(A+BH)^T & Q \end{bmatrix} > 0, \tag{70}$$

$$HQH^T \le 1. \tag{71}$$

5 Estimation of stability region with ACI level sets

5.1 Conditions for ACI level sets

We return to the system with a sector condition

$$x^+ = Ax + B\psi(Fx, t), \tag{72}$$

where $\psi_i \in \text{co}\{\breve{\phi}_i, \bar{\phi}_i\}, i \in I[1, m]$, and $\bar{\phi}_i, \breve{\phi}_i$ are given piecewise linear concave/convex functions. We would like to estimate the stability region using ACI level sets of V_c which is composed from J matrices $Q_j = Q_j^{\text{T}} > 0$.

Section 3 transforms the problem of verifying the absolute contractive invariance of a level set into one of verifying its contractive invariance for an array of saturated linear systems:

$$x^+ = A_q x + B_q \text{sat}(F_q x), \quad q \in I[1, K]. \tag{73}$$

Conditions for the contractive invariance of a level set of V_c for each of the above systems are presented in Theorem 2 for the general case. Putting all these conditions together, we obtain the condition of ACI for a level set $L_{V_c}(1)$ as follows.

Theorem 5 *If there exist* $H_q \in \mathbf{R}^{m\times n}$ *and* $\lambda_{qijk} \geq 0$, $\sum_{k=1}^{J} \lambda_{qijk} = 1$, $q \in I[1,K]$, $i \in I[1,2^m], j,k \in I[1,J]$, *such that*

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \lambda_{qijk} Q_k & (A_q + B_q(D_i F_q + D_i^- H_q))Q_j \\ Q_j(A_q + B_q(D_i F_q + D_i^- H_q))^T & Q_j \end{array} \right] > 0,\ q \in I[1,K], i \in I[1,2^m],\ j \in I[1,J], \tag{74}$$

$$h_{q,\ell} Q_j h_{q,\ell}^T \leq 1, \quad q \in I[1,K],\ \ell \in I[1,m],\ j \in I[1,N], \tag{75}$$

where $h_{q,\ell}$ *is the* ℓ*-th row of* H_q. *Then* $L_{V_c}(1)$ *is absolutely contractively invariant.*

In what follows, we consider the case $m = 1$, when B has only one column and ψ is a scalar function. Let the two boundary functions $\bar{\phi}$ and $\breve{\phi}$ be given as in (16) and (17), where $\bar{\phi}$ has N_1 bends and $\breve{\phi}$ has N_2 bends. So we have $K = N_1 + N_2$ saturated linear systems in (73). The matrices of each system is given as follows. For $q \in I[1, N_1]$,

$$A_q = A + k_{q1} BF, \quad B_q = c_{q1} B, \quad F_q = \frac{k_{01} - k_{q1}}{c_{q1}} F,$$

and for $q \in I[N_1 + 1, N_1 + N_2]$,

$$A_q = A + k_{q-N_1,2} BF, \quad B_q = c_{q-N_1,2} B, \quad F_q = \frac{k_{02} - k_{q-N_1,2}}{c_{q-N_1,2}} F.$$

It is easy to verify that $A_q + B_q F_q = A + k_{01} BF$ for $q \in I[1, N_1]$ and $A_q + B_q F_q = A + k_{02} BF$ for $q \in I[N_1 + 1, N_1 + N_2]$. Combining Lemma 1 and Theorem 3, we have the following result.

Theorem 6 *Assume that* $m = 1$. *If there exist* $H_{qj} \in \mathbf{R}^{1\times n}$ *and* $\alpha_{1jk}, \alpha_{2jk}, \beta_{qjk} \geq 0$, $\sum_{k=1}^{J} \alpha_{1jk} = 1$, $\sum_{k=1}^{J} \alpha_{2jk} = 1$, $\sum_{k=1}^{J} \beta_{qjk} = 1$, $q \in I[1,K]$, $j,k \in I[1,J]$ *such that*

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \alpha_{1jk} Q_k & (A + k_{01} BF)Q_j \\ Q_j(A + k_{01} BF)^T & Q_j \end{array} \right] > 0, \quad j \in I[1,J], \tag{76}$$

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \alpha_{2jk} Q_k & (A + k_{02} BF)Q_j \\ Q_j(A + k_{02} BF)^T & Q_j \end{array} \right] > 0, \quad j \in I[1,J], \tag{77}$$

$$\left[\begin{array}{cc} \sum_{k=1}^{J} \beta_{qjk} Q_k & (A_q + B_q H_{qj})Q_j \\ Q_j(A_q + B_q H_{qj})^T & Q_j \end{array} \right] > 0, \quad q \in I[1,K], j \in I[1,J], \tag{78}$$

$$H_{qj} Q_j H_{qj}^T \leq 1,\ \ q \in I[1,K], j \in I[1,J]. \tag{79}$$

Then $L_{V_c}(1)$ *is absolutely contractively invariant.*

We finally consider the special case where $m = 1$ and $J = 1$. For this case, the level sets of V_c reduce to ellipsoids. Using Theorem 4, the necessary and sufficient condition for the ACI of an ellipsoid can be obtained.

Theorem 7 *Assume that* $m = 1$. *Given* $Q = Q^T > 0$. *The ellipsoid* $\mathcal{E}(Q^{-1})$ *is ACI if and only if there exist* $H_q \in \mathbf{R}^{1\times n}$, $q \in I[1,K]$ *such that*

$$\left[\begin{array}{cc} Q & (A + k_{01} BF)Q \\ Q(A + k_{01} BF)^T & Q \end{array} \right] > 0, \tag{80}$$

$$\left[\begin{array}{cc} Q & (A + k_{02} BF)Q \\ Q(A + k_{02} BF)^T & Q \end{array} \right] > 0, \tag{81}$$

$$\left[\begin{array}{cc} Q & (A_q + B_q H_q)Q \\ Q(A_q + B_q H_q)^T & Q \end{array} \right] > 0, \quad q \in I[1,K], \tag{82}$$

$$H_q Q H_q^T \leq 1,\ \ q \in I[1,K]. \tag{83}$$

Proof. The sufficiency follows directly from applying Theorem 6 with $J = 1$. We prove the necessity. The ACI of the ellipsoid implies that for each $\rho \in (0, 1]$, $\mathcal{E}(Q^{-1}, \rho)$ is contractively invariant for

$$x^+ = Ax + B\bar{\phi}(Fx). \tag{84}$$

When ρ is sufficiently small, $\bar{\phi}(Fx) = k_{01}Fx$ for all $x \in \mathcal{E}(Q^{-1}, \rho)$. Hence $\mathcal{E}(Q^{-1}, \rho)$ must be contractively invariant for

$$x^+ = Ax + k_{01}BFx. \tag{85}$$

And by linearity, $\mathcal{E}(Q^{-1})$ must also be contractively invariant for (85). This proves (80). For $q \in I[1, N_1]$, we have

$$x^+ = A_q x + B_q \mathrm{sat}(F_q x) = Ax + B\bar{\phi}^q(Fx), \tag{86}$$

where

$$\bar{\phi}^q(u) = k_{q1}u + c_{q1}\mathrm{sat}\left(\frac{k_{01} - k_{q1}}{c_{q1}}u\right).$$

By Lemma 2, we have $\bar{\phi}^q(u) \in \mathrm{co}\{k_{01}u, \bar{\phi}(u)\}$. Hence the contractive invariance of $\mathcal{E}(Q^{-1})$ for (86) follows from its contractive invariance for both (84) and (85). Applying Theorem 4 to (86), there exists H_q satisfying (82) and (83). Similar arguments can be applied to $q \in I[N_1 + 1, N_1 + N_2]$ by using the contractive invariance of $\mathcal{E}(Q^{-1})$ for $x^+ = Ax + B\breve{\phi}(Fx)$. □

5.2 Optimization of ACI level sets

We recall that the set $L_{V_c}(1)$ is characterized by the matrices $Q_j, j \in I[1, J]$. By the definition of V_c and using Schur complement, we have

$$L_{V_c}(1) = \left\{x \in \mathbf{R}^n : \exists \gamma \in \Gamma \text{ s.t. } \begin{bmatrix} 1 & x^{\mathrm{T}} \\ x & \sum_{j=1}^J \gamma_j Q_j \end{bmatrix} \geq 0\right\}.$$

To obtain better estimation of the stability region, we would like to determine an ACI level set $L_{V_c}(1)$ as large as possible. Typically, we are given a set of reference points $x_p \in \mathbf{R}^n, p \in I[1, N]$ and would like to determine an ACI set so that it contains $\alpha x_p, p \in I[1, N]$ with α as large as possible. When it specifies to the level sets of V_c, we would like to determine matrices $Q_j, j \in I[1, J]$ satisfying the conditions of Theorem 5, 6, or 7 and further more, it satisfies

$$\begin{bmatrix} 1 & \alpha x_p^{\mathrm{T}} \\ \alpha x_p & \sum_{j=1}^J \gamma_{pj} Q_j \end{bmatrix} \geq 0, \quad \gamma_{pj} \geq 0, \sum_{j=1}^J \gamma_{pj} = 1,$$

with α as large as possible. For the case $m = 1$, this objective can be formulated (by applying Theorem 6) into the following optimization problem:

$$\sup_{Q_j, H_{qj}, \alpha_{1jk}, \alpha_{2jk}, \beta_{qjk}, \gamma_{pj}} \alpha \tag{87}$$

$$\text{s.t.} \quad a) \begin{bmatrix} 1 & \alpha x_p^{\mathrm{T}} \\ \alpha x_p & \sum_{j=1}^{J} \gamma_{pj} Q_j \end{bmatrix} \geq 0, \quad p \in I[1, N],$$

$$b) \begin{bmatrix} \sum_{k=1}^{J} \alpha_{1jk} Q_k & (A + k_{01} BF) Q_j \\ Q_j (A + k_{01} F)^{\mathrm{T}} & Q_j \end{bmatrix} > 0, \quad j \in I[1, J],$$

$$c) \begin{bmatrix} \sum_{k=1}^{J} \alpha_{2jk} Q_k & (A + k_{02} F) Q_j \\ Q_j (A + k_{02} F)^{\mathrm{T}} & Q_j \end{bmatrix} > 0, \quad j \in I[1, J],$$

$$d) \begin{bmatrix} \sum_{k=1}^{J} \beta_{qjk} Q_k & (A_q + B_q H_{qj}) Q_j \\ Q_j (A_q + B_q H_{qj})^{\mathrm{T}} & Q_j \end{bmatrix} > 0, \quad q \in I[1, K], j \in I[1, J],$$

$$e) \ H_{qj} Q_j H_{qj}^{\mathrm{T}} \leq 1, \ \ q \in I[1, K], j \in I[1, J],$$

$$f) \ Q_j = Q_j^{\mathrm{T}} > 0, \alpha_{1jk}, \alpha_{2jk}, \beta_{qjk}, \gamma_{pj} \geq 0, \ p \in I[1, N], q \in I[1, 2^m], j, k \in I[1, J],$$

$$g) \ \sum_{k=1}^{J} \alpha_{1jk} = 1, \ \sum_{k=1}^{J} \alpha_{2jk} = 1, \ \sum_{k=1}^{J} \beta_{qjk} = 1, \ \sum_{j=1}^{J} \gamma_{pj} = 1.$$

Similar optimization problem can be formulated based on Theorem 5 for the case $m > 1$. To simplify computation, we can introduce new parameters $Y_{qj} := H_{qj} Q_j, q \in I[1, K], j \in I[1, J]$. Then items d) and e) in the optimization problem can be replaced with

$$d) \begin{bmatrix} \sum_{k=1}^{J} \beta_{qjk} Q_k & A_q Q_j + B_q Y_{qj} \\ Q_j A_q^{\mathrm{T}} + Y_{qj}^{\mathrm{T}} B_q^{\mathrm{T}} & Q_j \end{bmatrix} > 0, \quad q \in I[1, K], j \in I[1, J],$$

$$e) \begin{bmatrix} 1 & Y_{qj} \\ Y_{qj}^{\mathrm{T}} & Q_j \end{bmatrix} \geq 0, \ \ q \in I[1, K], j \in I[1, J].$$

The resulting optimization problem has BMIs as constraints. In our computation, we again used the path-following method in [12] and it also turned out to be very effective. Before starting the path-following algorithm, we solved (87) under the assumption that all Q_j's are equal to Q. In this case, all the constraints are LMIs. We then assign this optimal Q to each Q_j, pick random α_{1jk}'s, α_{2jk}'s β_{qjk}'s and γ_{pj}'s satisfying f) and g), and then start the path-following algorithm. Based on our computational experience, all the initial values set up this way lead to the same final solutions.

We note that if we let $J = 1$ in (87), then g) is gone and we obtain an optimization problem to maximize the ACI ellipsoids with respect to x_p's. In this case, all the constraints are LMIs. Similar LMI problem can be derived for the case $m > 1$.

5.3 Numerical examples

The following two examples illustrate that using ACI level sets of V_c with $J = 2$ to estimate the stability region can improve significantly on what can be achieved by using ACI ellipsoids.

Example 2 Consider a second order system with one nonlinear component,

$$x^+ = Ax + B \tan^{-1}(Fx),$$

where

$$A = \begin{bmatrix} 0.8 & 0 \\ 0 & 1.2 \end{bmatrix}, \quad B = \begin{bmatrix} 0.6 \\ -0.8 \end{bmatrix}, \quad F = \begin{bmatrix} 0.5 & 1 \end{bmatrix}.$$

As in [13], we use the linear function $\psi_1(u) = u$ and a piecewise linear function ψ_2 to bound $\tan^{-1}(u)$, where ψ_2 is obtained by connecting a finite number of points on $\tan^{-1}(u)$ including the origin. Here we select six points $(u, \tan^{-1}(u)), u = 0, 1, 2, 3, 5, 8$. The resulting piecewise linear function $\psi_2(u)$ has the form of (7) with $N = 5$,

$$(k_0, k_1, k_2, k_3, k_4, k_5) = (0.7845, 0.3218, 0.1419, 0.0622, 0.0243, 0)$$
$$(c_1, c_2, c_3, c_4, c_5) = (0.4636, 0.8234, 1.0625, 1.2517, 1.4464).$$

We choose the reference point as $x_0 = \begin{bmatrix} 0 & 1 \end{bmatrix}^{\mathrm{T}}$.

We first consider quadratic functions ($J = 1$). The maximal α such that αx_0 is inside an ACI ellipsoid is $\alpha_1 = 3.6009$.

Next we consider V_c which is composed from two quadratic functions ($J = 2$). The maximal α such that αx_0 is inside an ACI $L_{V_c}(1)$ is $\alpha_2 = 5.0970$. (The maximal α such that αx_0 is inside the true stability region is 5.693, as detected by simulation). For verification, the two matrices defining the optimal V_c is given as follows,

$$Q_1 = \begin{bmatrix} 61.9296 & -20.4315 \\ -20.4315 & 25.3023 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 26.7952 & 5.2779 \\ 5.2779 & 26.1436 \end{bmatrix}.$$

The resulting invariant level sets by using different Lyapunov functions are compared in the left box of Fig. 3, where the outermost boundary is the optimal $L_{V_c}(1)$ for x_0. It can be seen that $L_{V_c}(1)$ is the convex hull of two ellipsoids (thin solid curve). The ellipsoids plotted in dashed lines are the maximal ACI ellipsoids with respect to x_0 and $x_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathrm{T}}$, respectively. To verify that the resulting level set $L_{V_c}(1)$ is actually invariant, we plot

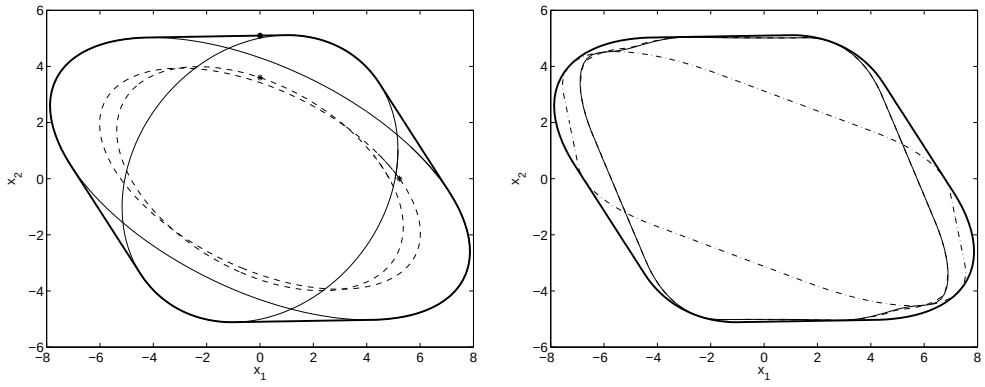

Figure 3: Left: ACI level sets; Right: next-step maps from ∂L_{V_c}

the image of its boundary under the next-step map $x \mapsto Ax + B\tan^{-1}(Fx)$ in the right box of Fig. 3 (see the thin solid lines). As a comparison, we also plot the images of the boundary under the map $x \mapsto Ax + B\psi_1(Fx)$ (dash-dotted) and the map $x \mapsto Ax + B\psi_2(Fx)$ (dashed), respectively. Since ψ_2 is very close to $\tan^{-1}$, the dashed curve and the thin solid curve are very close.

Example 3 Consider a second order system with two nonlinear components,

$$x^+ = Ax + B\psi(Fx),$$

where

$$A = \begin{bmatrix} 1 & -0.04 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, \quad F = \begin{bmatrix} 0.2727 & -0.2242 \\ -0.5097 & -0.1563 \end{bmatrix};$$

and $\psi_1(u_1) \in \text{co}\{\text{sat}(u_1), 2u_1\}$, $\psi_2(u_2) \in \text{co}\{\text{sat}(u_2), 2u_2\}$. The eigenvalues of A are $1 \pm j0.2$ and the eigenvalues of $A + BF$ are $0.65 \pm j0.3841$. If we let B_1 and B_2 be the two columns of B and let F_1, F_2 be the two rows of F, we obtain four vertex systems,

$$x^+ = (A + 2BF)x, \tag{88}$$

$$x^+ = (A + 2B_1F_1)x + B_2\text{sat}(F_2x), \tag{89}$$

$$x^+ = (A + 2B_2F_2)x + B_1\text{sat}(F_1x), \tag{90}$$

$$x^+ = Ax + B\text{sat}(Fx). \tag{91}$$

We take the reference point as $x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We would like to determine an ACI set such that it contains αx_0 with α as large as possible. If we optimize over all the ACI ellipsoids, the maximal α is $\alpha_1 = 10.1473$. If we optimize over ACI level sets of V_c with $J = 2$, the maximal α is $\alpha_2 = 19.0170$. Plotted in the left box of Fig. 4 are the optimized ACI ellipsoid (dashed line) and ACI $L_{V_c}(1)$ (thick solid line). Also plotted in this box in thin solid lines are the two ellipsoids $\mathcal{E}(Q_1^{-1})$ and $\mathcal{E}(Q_2^{-1})$ whose convex hull is $L_{V_c}(1)$. For verification, Q_1 and Q_2 are given as follows,

$$Q_1 = \begin{bmatrix} 47.1064 & -92.8231 \\ -92.8231 & 356.9252 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 40.9330 & -21.3000 \\ -21.3000 & 372.7298 \end{bmatrix}.$$

To demonstrate that this level set $L_{V_c}(1)$ is indeed absolutely contractively invariant, we computed the images of the boundary of $L_{V_c}(1)$ under the four next-step maps corresponding to the four vertex systems (88) - (91). These images are plotted with thin solid curves in the right box of Fig. 4. As can be clearly seen, all these images are within the level set $L_{V_c}(1)$.

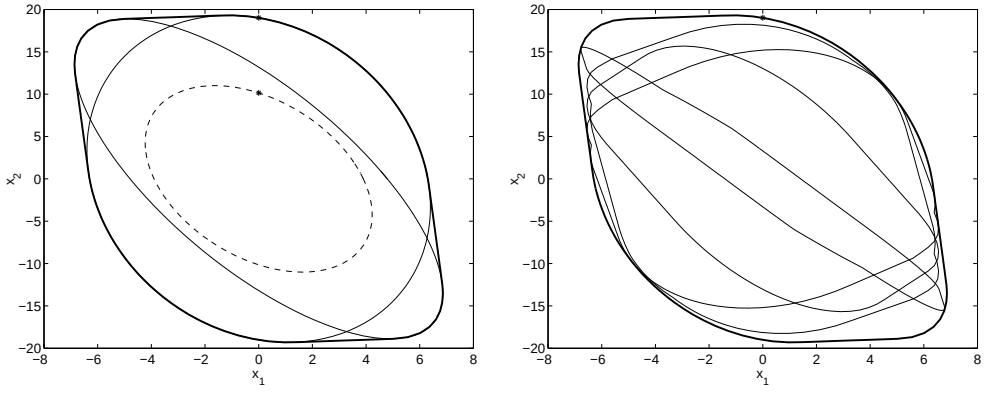

Figure 4: Left: ACI level sets; Right: next-step maps from $\partial L_{V_c}(1)$

As a matter of fact, we can also use V_c to determine a region of instability using numerical simulation. The idea is to generate the "worst switching" among the four vertex systems so that at each point x, x^+ is chosen as the one which maximizes V_c. By doing this, a diverging trajectory can be potentially produced. Fig. 5 plots a diverging trajectory produced this way (dash-dotted line). The initial state is $\begin{bmatrix} 0 & 24.5 \end{bmatrix}^{\text{T}}$ (marked with "*").

A nearly closed trajectory is also produced under the worst switching strategy (see the curve in thin solid line). The region outside this nearly closed curve is deemed unstable. Also plotted in Fig. 5 is the same ACI level set as in Fig. 4. We notice that there is some gap between the estimated stability region (the ACI set) and the deemed region of instability. The true region of instability could include some points inside the nearly closed curve and the real stability region must be larger than the ACI set.

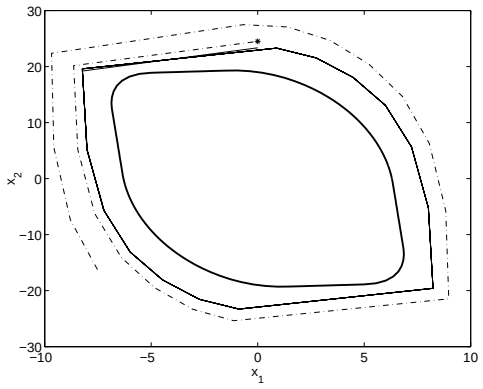

Figure 5: Region of instability and a diverging trajectory

We note that with different switching strategies (possibly the worst case with respect to different Lyapunov functions), different regions of instability can be detected. And the true region of instability for the system with sector condition will be the union of all these.

6 Conclusions

We used composite quadratic Lyapunov function to enhance absolute stability analysis of systems with a piecewise linear sector condition. The composite quadratic Lyapunov function was introduced in our recent work [16] for stability analysis of saturated linear systems. Its properties have been further studied in [10, 17, 20] and great potential has been demonstrated in the stability analysis of linear differential inclusions and saturated linear systems. Inspired by recent development in continuous-time systems, this paper first establishes stability conditions through composite quadratic Lyapunov functions for linear difference inclusions and saturated linear discrete-time systems. These results are used to establish conditions for absolute stability through a connection to saturated linear systems. With these conditions, the problem of estimating the stability region is formulated as optimization problems with bilinear matrix inequalities which can be effectively solved with the path-following method in [12]. As illustrated by numerical examples, the stability region estimated by the composite quadratic Lyapunov function can be significantly larger than those by quadratic functions. The composite quadratic functions can also be used to generate the "worst switching" among a group of vertex systems for the purpose of detecting a potentially diverging trajectory.

References

[1] M. Arcak, M. Larsen and P. Kokotovic, "Circle and Popov criteria as tools for nonlinear feedback design," *Automatica*, Vol. 39, pp. 643-650, 2003.

[2] F. Blanchini, "Nonquadratic Lyapunov functions for robust control," *Automatica*, **31**, pp. 451-461, 1995.

[3] E. Beran, L. Vandenberghe, and S. Boyd, "A global BMI algorithm based on the generalized Benders decomposition," *Proc. of the European Control Conference*, paper no.934, 1997.

[4] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, SIAM Studies in Appl. Mathematics, Philadelphia, 1994.

[5] R. K. Brayton and C. H. Tong, "Stability of dynamical systems: a constructive approach," *IEEE Trans. on Circuits and Systems*, **26**, pp. 224-234, 1979.

[6] G. Chesi, A. Garulli, A. Tesi and A. Vicino, "Homogeneous Lyapunov functions for systems with structured uncertainties," *Automatica*, **39**, pp. 1027-1035, 2003.

[7] Y-S Chou, A. L. Tits and V. Balakrishnan, "Stability multipliers and μ upper bounds: connections and implications for numerical verification of frequency domain conditions," *IEEE Trans. Automat. Contr.*, Vol. 44, No. 5, pp. 906-913, 1999.

[8] A. Dewey and E. Jury, "A stability inequality for a class of nonlinear feedback systems," *IEEE Trans. Automat. Contr.*, Vol. 11, pp. 54-62, 1966.

[9] W. M. Haddad and V. Kapila,"Absolute stability criteria for multiple slope-restricted monotonic nonlinearities," *IEEE Trans. Automat. Contr.*, Vol. 40, No. 2, pp. 361-365, 1995.

[10] R. Goebel, A. R. Teel, T. Hu and Z. Lin, "Dissipativity for dual linear differential inclusions through conjugate storage functions," *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004.

[11] K. C. Goh, M. G. Safonov, and G. P. Papavassilopoulos, "A global optimization approach for the BMI problem," *Proc. of the IEEE Conference on Decision and Control*, pp. 2009-2014, 1994.

[12] A. Hassibi, J. How and S. Boyd, "A path-following method for solving BMI problems in control," *Proc. of American Control Conference*, pp. 1385-1389, 1999.

[13] T. Hu, B. Huang and Z. Lin, "Absolute stability with a generalized sector condition," *IEEE Transactions on Automatic Control*, Vol. 49, No. 4, pp. 535-548, 2004.

[14] T. Hu and Z. Lin. *Control Systems with Actuator Saturation: Analysis and Design*, Birkhäuser, Boston, 2001.

[15] T. Hu and Z. Lin, "Exact characterization of invariant ellipsoids for linear systems with saturating actuators," *IEEE Transactions on Automatic Control*, Vol. 47, No. 1, pp. 164-169, 2002.

[16] T. Hu and Z. Lin, "Composite quadratic Lyapunov functions for constrained control systems," *IEEE Transactions on Automatic Control*, Vol.48, No. 3, pp.440-450, 2003.

[17] T. Hu and Z. Lin, "Properties of composite quadratic Lyapunov functions," *IEEE Transactions on Automatic Control*, Vol.49, No. 7, pp. 1162-1167, 2004.

[18] T. Hu, Z. Lin and B. M. Chen, "An analysis and design method for linear systems subject to actuator saturation and disturbance," *Automatica*, Vol. 38, No. 2, pp. 351-359, 2002.

[19] T. Hu, Z. Lin and B. M. Chen, "Analysis and design for linear discrete-time systems subject to actuator saturation," *Systems & Control Letters*, Vol. 45, No. 2, pp. 97-112, 2002.

[20] T. Hu, Z. Lin, R. Goebel and A. R. Teel, "Stability regions for saturated linear systems via conjugate Lyapunov functions," *Proceedings of the 43rd IEEE Conf. on Dec. and Contr.*, 2004.

[21] Z. Jarvis-Wloszek and A. K. Packard, "An LMI method to demonstrate simultaneous stability using non-quadratic polynomial Lyapunov functions," *Proceedings of the IEEE Conf. on Decision and Control*, pp. 287-292, Las Vegas, NV, 2002.

[22] E. I. Jury and B. W. Lee, "On the stability of a certain class of nonlinear sample-data systems," *IEEE Trans. Automat. Contr.*, Vol. 9, Jan., pp. 51-61, 1964.

[23] M. Johansson and A. Rantzer, "Computation of piecewise quadratic Lyapunov functions for hybrid systems," *IEEE Trans. Automat. Contr.*, **43**, No. 4, pp. 555-559, 1998.

[24] A. Molchanov and D. Liu, "Robust absolute stability of time-varying nonlinear discrete-time systems," *IEEE Transactions on Circuit and Systems I: Fundamental Theory and Appicaltions*, Vol. 49, No. 8, pp. 1129-1137, 2002.

[25] A. P. Molchanov, Y. Pyatnitskiy "Criteria of asymptotic stability of differential and difference inclusions encountered in control theory," *Systems & Control Letters*, **13**, pp. 59-64, 1989.

[26] K.S. Narendra and J. Taylor, *Frequency Domain Methods for Absolute Stability*, Academic Press, New York, 1973.

[27] J. B. Pearson and J. E. Gibson, "On the asymptotic stability of a class of saturating sampled-data systems," *IEEE Trans. on Application and Industry*, No. 71, pp. 81-86, March, 1964.

[28] H. M. Power and A. C. Tsoi, "Improving the predictions of the circle criterion by combining quadratic forms," *IEEE Transactions on Automatic Control*, **28**, pp. 65-67, 1973.

[29] V. Singh, "A stability inequality for nonlinear feedback systems with slope-restricted nonlinearity," *IEEE Trans. Automat. Contr.*, Vol. 29, No. 8, pp. 743-744, 1984.

[30] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.

[31] T. Wada, M. Ikeda, Y. Ohta and D. D. Siljak, "Parametric absolute stability of Lur'e systems," *IEEE Trans. Automatic Control*, AC-43, N0. 11, pp. 1649-1653, 1998.

[32] J. L. Willems, "The computation of finte stability regions by means of open Lyapunov surfaces," *Int. J. Control*, Vol. 10, No. 5, pp. 537-544, 1969.

[33] L. Xie, S. Shishkin and M. Fu, "Piecewise Lyapunov functions for robust stability of linear time-varying systems," *Systems & Control Letters*, **31**, pp.165-171, 1997.

[34] V. A. Yakubovich, "Frequency conditions for the absolute stability of control systems with several nonlinear and linear nonstationary blocks," *Automation and Remote Control*, Vol. 28, pp. 857-880, 1967.

A Four equivalent statements about invariant ellipsoids

Proposition 1 *Given an ellipsoid $\mathcal{E}(P,\rho)$ and a row vector $F \in \mathbf{R}^{1\times n}$. Assume that*

$$(A+BF)^T P(A+BF) - P < 0. \tag{92}$$

The following four statements are equivalent:

a) The ellipsoid is contractively invariant for

$$x^+ = Ax + B\text{sat}(Fx). \tag{93}$$

b) There exists a function $f : \mathbf{R} \to [-1,1]$ such that the ellipsoid is contractively invariant for

$$x^+ = Ax + Bf(x). \tag{94}$$

c) The ellipsoid is contractively invariant for

$$x^+ = Ax + B\text{sat}(F_0 x), \quad F_0 = -(B^T P B)^{-1} B^T P A. \tag{95}$$

d) There exists an $H \in \mathbf{R}^{1\times n}$ such that

$$(A+BH)^T P(A+BH) - P < 0, \tag{96}$$

and $\mathcal{E}(P,\rho) \subset \mathcal{L}(H)$.

We notice that Theorem 4 follows from the equivalence of a) and d). Let $Q = P/\rho$, then by using Schur complement it can be shown that (69) and (70) are equivalent to (92) and (96). Also $HQH^{\text{T}} \le 1$ is equivalent to $\mathcal{E}(P,\rho) \subset \mathcal{L}(H)$.

Statement a) in Proposition 1 is about the invariance of an ellipsoid under a given feedback law. Statement b) is about the existence of a bounded feedback law to make the ellipsoid invariant. In statement c), the feedback law $u = \text{sat}(F_0 x)$ maximizes the convergence rate with respect to the function $V(x) = x^{\text{T}} P x$. Statement d) is about the existence of a feedback law linear inside the ellipsoid to make it invariant. There is no direct way to verify a), b) or c). The significance of Proposition 1 lies in the fact that all of them can be verified through d) which is numerically tractable.

The equivalence of a) and b) was established in [14] (see also [19]) and the equivalence of b) and c) was established in [14] (page 258). It is also clear that d) implies b) and hence a) and c). In what follows, we will show that c) is equivalent to d).

Define

$$\rho_c^* = \sup\{\rho > 0 : \ \ \mathcal{E}(P,\rho) \ \text{ is contractively invariant for (95)}\}$$

and

$$\rho_d^* = \sup_H \rho \tag{97}$$

$$s.t. \ \ (A+BH)^{\text{T}} P(A+BH) - P < 0, \tag{98}$$

$$HP^{-1}H^{\text{T}} \le \frac{1}{\rho}. \tag{99}$$

Note that $HP^{-1}H^{\mathrm{T}} \le \frac{1}{\rho}$ is equivalent to $\mathcal{E}(P,\rho) \subset L(H)$. Since there exists an H satisfying (98) (actually, if we let $H = F$), the "$<$" in (98) can be replaced with "$\le$". Hence the above optimization problem can also be written as

$$\begin{aligned}(\rho_d^*)^{-1} = \min_H\ & HP^{-1}H^{\mathrm{T}} \\ s.t.\ & \left[\begin{array}{cc} P & (A+BH)^{\mathrm{T}} \\ A+BH & P^{-1} \end{array}\right] \ge 0.\end{aligned} \tag{100}$$

To prove Proposition 1, it suffices to show that

Lemma 6 $\rho_c^* = \rho_d^*$.

Proof. It is evident that $\rho_c^* \ge \rho_d^*$ (note that b) and c) are equivalent). If there exists an $x \in \mathbf{R}^n \setminus \{0\}$ such that

$$(Ax + B\mathrm{sat}(F_0x))^{\mathrm{T}}P(Ax + B\mathrm{sat}(F_0x)) - x^{\mathrm{T}}Px = 0,$$

then $\rho_c^* < \infty$ and

$$\begin{aligned}\rho_c^* = \min_{x\in\mathbf{R}^n\setminus\{0\}}\ & x^{\mathrm{T}}Px \\ \text{s.t. }\ & (Ax + B\mathrm{sat}(F_0x))^{\mathrm{T}}P(Ax + B\mathrm{sat}(F_0x)) - x^{\mathrm{T}}Px = 0.\end{aligned} \tag{101}$$

To show $\rho_c^* \le \rho_d^*$, it suffices to consider the case where $\rho_d^* < \infty$ and to construct an $x_0 \in \mathbf{R}^n$ such that $x_0^{\mathrm{T}}Px_0 = \rho_d^*$ and x_0 satisfies (101).

For clarity, we divide the proof into 3 steps.

Step 1: Transformation of the optimization problem and normalization.

Denote

$$W(H) = \left[\begin{array}{cc} P & (A+BH)^{\mathrm{T}} \\ A+BH & P^{-1} \end{array}\right]$$

and define

$$\begin{aligned}S_1(\alpha) &:= \left\{H \in \mathbf{R}^{1\times n} :\ \ W(H) \ge \alpha I\right\} \\ S_2(\beta) &:= \left\{H \in \mathbf{R}^{1\times n} :\ \ HP^{-1}H^{\mathrm{T}} \le \beta\right\}.\end{aligned}$$

Since $S_1(\alpha)$ is convex and $S_2(\beta)$ is strictly convex (any point between two distinct boundary points is not on the boundary), (100) implies that $S_1(0)$ and $S_2((\rho_d^*)^{-1})$ has a unique intersection H_* with $\lambda_{\min}(W(H_*)) = 0$, where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a matrix. Furthermore, for all $H \in S_2((\rho_d^*)^{-1}) \setminus \{H_*\}$, we have $H \notin S_1(0)$, i.e., $\lambda_{\min}(W(H)) < 0$. It follows that

$$\begin{aligned}0 = \max_H\ & \lambda_{\min}(W(H)) \\ \text{s.t. }\ & HP^{-1}H^{\mathrm{T}} \le (\rho_d^*)^{-1},\end{aligned}$$

and the above problem has a unique optimal solution H_*.

Without loss of generality and for simplicity, we assume that $P = I$ and $\rho_d^* = 1$. Otherwise we can make it so with a state transformation $\bar{x} = (P/\rho_d^*)^{\frac{1}{2}}\, x$. Thus we have

$$0 = \max_H \lambda_{\min}(W(H)) \tag{102}$$

$$\text{s.t. } HH^{\mathrm{T}} \leq 1, \tag{103}$$

where

$$W(H) = \begin{bmatrix} I & (A+BH)^{\mathrm{T}} \\ A+BH & I \end{bmatrix}.$$

Step 2: The eigenvector of $W(H_*)$.

Let H_* be the unique optimal solution to (102). Then $\lambda_{\min}(W(H_*)) = 0$. Suppose that the multiplicity of the zero eigenvalue of $W(H_*)$ is p and let $X \in \mathbf{R}^{n\times p}$ span the eigenspace of the zero eigenvalue. Then

$$W(H_*)X = 0.$$

We claim that $p = 1$. Here we need to resort to eigenvalue perturbation theory (see, e.g., [30]) to prove this claim. Let $H = H_* + k\Delta H$, where ΔH represents the direction of perturbation. As k is increased from 0, $\lambda_{\min}(W(H_* + k\Delta H)$ increases with a slope greater than ε if and only if

$$X^{\mathrm{T}} \frac{\partial W(H_* + k\Delta H)}{\partial k} X > \varepsilon I.$$

We note that, for a fixed ΔH,

$$X^{\mathrm{T}} \frac{\partial W(H_* + k\Delta H)}{\partial k} X = X^{\mathrm{T}} \begin{bmatrix} 0 & (B\Delta H)^{\mathrm{T}} \\ B\Delta H & 0 \end{bmatrix} X.$$

Let H_0 be such that $W(H_0) > \eta I$ for some $\eta > 0$ (such H_0 exists by assumption, e.g., $H_0 = F$). Consider $\Delta H = H_0 - H_*$. Then for $k \in (0, 1)$,

$$W(H_* + k\Delta H) = W((1-k)H_* + kH_0) = (1-k)W(H_*) + kW(H_0) > k\eta I.$$

Hence we must have

$$X^{\mathrm{T}} \begin{bmatrix} 0 & (B\Delta H)^{\mathrm{T}} \\ B\Delta H & 0 \end{bmatrix} X > \eta I. \tag{104}$$

Since both B and ΔH are of rank 1, the matrix $\begin{bmatrix} 0 & (B\Delta H)^{\mathrm{T}} \\ B\Delta H & 0 \end{bmatrix}$ has at most one positive eigenvalue. For (104) to be true, X can only have one column, i.e., $p = 1$.

The optimality of the solution H_* means that $\lambda_{\min}(W(H))$ cannot be increased by varying H in a neighborhood of H_* along any direction which keeps it within the constraint $HH^{\mathrm{T}} \leq 1$, i.e.,

1) $\lambda_{\min}(W(H_* + k\Delta H))$ cannot be increased for $k > 0$ or $k < 0$ if ΔH is tangential to the sphere surface $HH^{\mathrm{T}} = 1$ at H_*;

2) $\lambda_{\min}(W(H_* + k\Delta H))$ cannot be increased for $k > 0$ if ΔH points inward of the sphere $HH^{\mathrm{T}} \leq 1$ from H_*.

Partition X as $X = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, where $v_1, v_2 \in \mathbf{R}^n$. By the eigenvalue perturbation theory, items 1) and 2) reduce to

$$X^{\mathrm{T}} \frac{\partial W(H_* + k\Delta H)}{\partial k} X = 2v_2^{\mathrm{T}} B \Delta H v_1 = 0, \qquad \text{if } \Delta H {H_*}^{\mathrm{T}} = 0, \tag{105}$$

$$X^{\mathrm{T}} \frac{\partial W(H_* - kH_*)}{\partial k} X = -2v_2^{\mathrm{T}} B H_* v_1 < 0. \tag{106}$$

For (105), we note that $\lambda_{\min}(W(H_* + k\Delta H))$ cannot be increased for either $k > 0$ or $k < 0$, hence the derivative has to be zero. For (106), we recall that there is one direction $\Delta H = H_0 - H_*$ which makes $\lambda_{\min}(W(H_* + k\Delta H))$ strictly increase. Hence $v_2^{\mathrm{T}} B H_* v_1 \neq 0$.

Combining (105) and (106), we see that $v_2^{\mathrm{T}} B \neq 0$ and v_1 must be proportional to ${H_*}^{\mathrm{T}}$. For simplicity, we take $v_1 = {H_*}^{\mathrm{T}}$. Recall that $X = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is an eigenvector corresponding to the zero eigenvalue of $W(H_*)$, we have,

$$\begin{bmatrix} I & (A + BH_*)^{\mathrm{T}} \\ A + BH_* & I \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0,$$

from which we obtain

$$v_2 = -(A + BH_*){H_*}^{\mathrm{T}}, \tag{107}$$

$${H_*}^{\mathrm{T}} = (A + BH_*)^{\mathrm{T}}(A + BH_*){H_*}^{\mathrm{T}}. \tag{108}$$

Step 3: Construction of x_0.

From (106), (107), $v_1 = {H_*}^{\mathrm{T}}$ and $H_* {H_*}^{\mathrm{T}} = 1$, we have

$$-(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}A{H_*}^{\mathrm{T}} > 1 \Longrightarrow F_0 H_*^{\mathrm{T}} > 1. \tag{109}$$

Let $x_0 = {H_*}^{\mathrm{T}}$, then $H_* x_0 = x_0^{\mathrm{T}} x_0 = 1$ and $\mathrm{sat}(F_0 x_0) = 1 = H_* x_0$. It follows from (108) that

$$\begin{aligned} (Ax_0 + B\mathrm{sat}(F_0 x_0))^{\mathrm{T}}(Ax_0 + B\mathrm{sat}(F_0 x_0)) &= x_0^{\mathrm{T}}(A + BH_*)^{\mathrm{T}}(A + BH_*)x_0 \\ &= H_*(A + BH_*)^{\mathrm{T}}(A + BH_*)H_*^T \\ &= 1 \\ &= x_0^{\mathrm{T}} P x_0. \end{aligned}$$

Recall that we have assumed $P = I$ and $\rho_d^* = 1$. This shows that there exists an x_0 such that $x_0^{\mathrm{T}} P x_0 = 1$ and x_0 satisfies (101). It follows that $\rho_c^* \leq 1 = \rho_d^*$. □

From the proof of Lemma 6, we see that at the optimal solution H_*, the matrix

$$\begin{bmatrix} P & (A + BH_*)^{\mathrm{T}} \\ A + BH_* & P^{-1} \end{bmatrix}$$

has a single eigenvalue at 0. This is because B has only one column.