HAL
open science

Stability and Observer Design for Lur'e Systems with Multivalued, Nonmonotone, Time-Varying Nonlinearities and State Jumps

Aneel Tanwani, Bernard Brogliato, Christophe Prieur

To cite this version:

Aneel Tanwani, Bernard Brogliato, Christophe Prieur. Stability and Observer Design for Lur'e Systems with Multivalued, Nonmonotone, Time-Varying Nonlinearities and State Jumps. SIAM Journal on Control and Optimization, 2014, 52 (6), pp.3639-3672. 10.1137/120902252 . hal-01083629

HAL Id: hal-01083629

https://inria.hal.science/hal-01083629

Submitted on 23 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stability and Observer Design for Lur'e Systems with Multivalued, Non-monotone, Time-Varying Nonlinearities and State Jumps

Aneel Tanwani* Bernard Brogliato* Christophe Prieur†

February 21, 2015

Abstract

This paper deals with the stability and observer design for Lur'e systems with multivalued nonlinearities, which are not necessarily monotone or time-invariant. Such differential inclusions model the motion of state trajectories which are constrained to evolve inside time-varying non-convex sets. Using Lyapunov-based analysis, sufficient conditions are proposed for local stability in such systems, while specifying the basin of attraction. If the sets governing the motion of state trajectories are moving with bounded variation, then the resulting state trajectories are also of bounded variation, and unlike the convex case, the stability conditions depend on the size of jumps allowed in the sets. Based on the stability analysis, a Luenberger-like observer is proposed which is shown to converge asymptotically to the actual state, provided the initial value of the state estimation error is small enough. In addition, a practically convergent state estimator, based on the high-gain approach, is designed to reduce the state estimation error to the desired accuracy in finite time for larger initial values of the state estimation error. The two approaches are then combined to obtain global asymptotically convergent state estimates.

1 Introduction

Lur'e systems comprise an asymptotically stable linear system with an additive nonlinearity in the feedback loop as shown in Figure 1. Due to the practical significance, the problem of stability in such systems has received much attention in the literature. The basic question is: for what kind of nonlinearities, the resulting feedback system is asymptotically stable? For the most part, researchers have assumed that the nonlinearities are monotone and satisfy certain sector condition. However, if the monotonicity assumption is relaxed for the nonlinearities, then one can obtain certain sufficient conditions on linear dynamics which render the system stable. Our paper follows this line of thought, by addressing the stability and observer design for a certain type of Lur'e systems where nonlinearities belong to the normal cones associated with time-varying non-convex sets. Such multivalued nonlinearities don't necessarily satisfy monotone condition and may induce jumps in the state trajectory.

The kind of models considered in this paper could also be classified as the *sweeping processes*, introduced in [35] (see [28] for tutorial exposition). Such models have found several applications in non-smooth dynamics such as impact mechanics [33], electrical circuits [1], and could be used to generalize several existing classes of discontinuous dynamical systems such as linear complementarity systems (LCS) [10]. The perturbed version of Moreau's sweeping process is described by the following differential inclusion:

$$-\dot{x}(t) + f(t, x(t)) \in \mathcal{N}(S(t); x(t)), \quad \text{for almost all } t \in [t_0, \infty), \tag{1a}$$

$$x(t_0) = x_0 \in S(t_0), \tag{1b}$$

*Team BipOp, Institut National de Recherche en Informatique et Automatique (INRIA), Rhône-Alpes, 655 Avenue de l'Europe, 38334 Montbonnot Cedex, France. Email:{aneel.tanwani,bernard.brogliato}@inria.fr.

†Gipsa-lab, Department of Automatic Control, 11 Rue des Mathématiques, BP 46, 38402 Saint Martin d'Hères, France. Email: christophe.prieur@gipsa-lab.fr. This work has been partially supported bytheLabExPERSYVAL-Lab(ANR-11-LABX-0025-01)

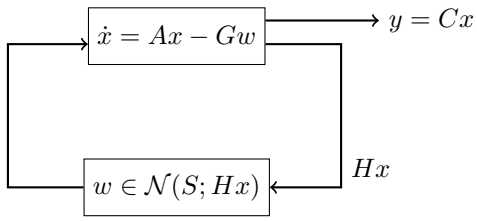


Figure 1: Lur'e systems with multivalued nonlinearities in feedback.

where $S : [t_0, \infty) \rightrightarrows \mathbb{R}^n$ is a set-valued map, and $\mathcal{N}(S(t); x(t))$ denotes the Fréchet normal cone to the set $S(t) \subset \mathbb{R}^n$ at a point $x(t) \in S(t)$. Intuitively speaking, the aforementioned differential inclusion states that the state $x(t)$ moves in the direction given by $f(t, x(t))$ when $x(t)$ is in the interior of the set $S(t)$ since, in that case, $\mathcal{N}(S(t); x(t)) = \{0\}$. However, when $x(t)$ is on the boundary of $S(t)$, the Fréchet normal at $x(t)$ is no longer just zero, and it acquires the value from the normal cone defined at that boundary point so that $x(t)$ stays inside the set $S(t)$, for all $t \geq t_0$.

The important considerations in solving the differential inclusion (1) are the geometry of sets $S(t)$, and the variation (measured using Hausdorff metric) of these sets with time. The case where $S(t)$ is convex renders itself amenable to analysis; see [7, 8, 9] for well-posedness and our preliminary results on observer design for such cases. However, in this paper, we will assume that there exists an $r > 0$ such that $S(t)$ is an *r-prox-regular* set and closed for each $t \geq t_0$. Some simple examples include: the complement of an open ball of radius greater than r in $\mathbb{R}^n$; union of disconnected closed intervals on the real line separated by a distance greater than $2r$; and sub-level sets of any twice differentiable function. Every convex set is r-prox-regular where one can take r to be arbitrarily large. Several characterizations of prox-regular sets appear in [39], which indicate that a set S is r-prox-regular if and only if there exists a unique $\bar{z} := \arg\min_{z \in S} |y - z|$ in S for every $y \in \mathbb{R}^n$ satisfying $\min_{z \in S} |y - z| < r$. An application of differential inclusions with prox-regular sets appears in [32] for modeling the motion of a crowd where the agents in the crowd are required to keep a certain distance among each other, say $2r$. If we lump the positions of N agents in the crowd into a single vector, say x, then this vector must belong to a prox-regular set $\{x \in \mathbb{R}^N : |x_i - x_j| > r, 1 \leq i, j \leq N, i \neq j\}$. One can see the application of prox-regular sets in optimal control problems [16], as they appear to be the first step in studying problems in non-convex setup. Solution concepts for differential inclusions of type (1), where $S(t)$ is assumed to be r-prox-regular, are discussed in [10, 16, 17], and we recall some of them later on.

This paper first addresses the problem of stability for the perturbed Moreau's sweeping process described by (1), where we fix $f(t, x(t)) := Ax(t)$. Depending on how the function $S(\cdot)$ evolves, the state $x(\cdot)$ evolves accordingly. As a simpler case, we first consider the case where $S(\cdot)$ varies in an absolutely continuous manner which allows the resulting state trajectory to be absolutely continuous. The major difference in the stability analysis, compared to the case where $S(\cdot)$ is convex-valued, is that the mapping $x \mapsto \mathcal{N}(S(t); x)$ is *maximal monotone* when $S(t)$ is convex, which is not the case for a prox-regular set $S(t)$, even though the set $\mathcal{N}(S(t); x)$ is convex. Unlike the global stability obtained in the convex case, the region of stability in an r-prox-regular case is determined by r. Nonetheless, we recover the results for convex sets as a specific case of the results presented.

We then address the problem of designing observers for such systems. Observer design using passivity approach for single valued Lur'e systems was considered in [3, 18] and later generalized to Lur'e systems with multivalued nonlinearity by [8, 9, 38]. Similar approach has been adopted for observer design in linear complementarity systems by [22], and several generalizations of these results have appeared recently in [4, 12, 25, 26]. The stability results developed in this paper are utilized to design a state estimator where we show that the dynamics of state estimation error are asymptotically stable under certain conditions. We remark that, compared to the earlier work on observer design [9], the major contribution of this work is to study the observer problem for non-convex sweeping processes within the general solution framework comprising functions of bounded variations (which allow state jumps with Zeno behavior), and to derive

stability criteria for error dynamics that depend on the parameter r which, roughly speaking, measures the lack of convexity.

The paper is organized as follows: In Section 2, we give a formal definition of the prox-regular set along with some of its properties, and recall the result on existence and uniqueness of solution for system (1) when the system admits absolutely continuous state trajectories. For such systems, the problems of stability and observer design are studied in Section 3. The conditions for stability appear in Section 3.1, followed by a two-stage observer design in Section 3.2. In Section 4, we develop parallel results when the trajectories of the system are of bounded variation, and hence possibly discontinuous. Sufficient conditions for stability and the results on observer design appear in Section 4.2 and Section 4.3, respectively. After some concluding remarks, we have collected the proofs of some technical results in the Appendix.

2 Preliminaries

In this section, we introduce the prox-regular sets and derive some properties which will be used in the development of main results later. Before proceeding with the formal treatment, we introduce some standard notation that will be used throughout the paper.

Notations: We denote the Euclidean norm of a vector $x \in \mathbb{R}^n$ by $|x|$ and the induced norm of a matrix A with $\|A\|$. The maximum and minimum eigenvalues of a matrix P are denoted by $\sigma_{\max}(P)$ and $\sigma_{\min}(P)$, respectively. For a matrix H, σ_H^+ denotes its least positive singular value. The preimage of a set $S \subseteq \mathbb{R}^l$ under a function $f : \mathbb{R}^n \to \mathbb{R}^l$ is denoted by $f^{-1}(S)$, that is, $f^{-1}(S) := \{z \in \mathbb{R}^n \mid f(z) \in S\}$, and the range of $f(\cdot)$ is defined as range $f := \{y \in \mathbb{R}^l \mid f(z) = y \text{ for some } z \in \mathbb{R}^n\}$. For an interval $I \subseteq \mathbb{R}$, and a function $f : I \to \mathbb{R}^n$, the variation of $f(\cdot)$ over the interval I is the supremum of $\sum_{i=1}^k |f(s_i) - f(s_{i-1})$ over the set of all finite sets of points $s_0 < s_1 < \cdots < s_k$ (called partitions) of I. When this supremum is finite, the mapping $f(\cdot)$ is said to be of *bounded variation* on I. We say that $f(\cdot)$ is of *locally bounded variation* if it is of bounded variation on each compact subinterval of I. If $f(\cdot)$ is right-continuous and of (locally) bounded variation, we call it (locally) *rcbv*. A function of locally bounded variation on I has at most a countable number of jump discontinuities in I. Moreover, it has right and left limits everywhere. The right and left limits of the function $f(\cdot)$ at $t \in I$ are denoted by $f(t^+) := \lim_{s \searrow t} f(s)$ and $f(t^-) := \lim_{s \nearrow t} f(s)$, respectively, provided they exist. In this notation, right continuity of $f(\cdot)$ in t, means that $f(t^+) = f(t)$.

The distance between a point $x \in \mathbb{R}^n$ and a set $S \subseteq \mathbb{R}^n$ is denoted by $d(x, S) := \inf_{z \in S} |x - z|$ and we let $\text{prox}_S(x) := \arg\min_{z \in S} |x - z|$. For any two sets $S, S' \subset \mathbb{R}^n$, we denote their Hausdorff distance by $d_H(S, S')$ which is defined as:

$$d_H(S, S') := \max\{\sup_{x \in S'} d(x, S), \sup_{x \in S} d(x, S')\}. \tag{2}$$

In this paper, we will consider set-valued maps $S : [t_0, \infty) \rightrightarrows \mathbb{R}^l$, for some fixed $t_0 \in \mathbb{R}$. The variation of $S(\cdot)$ over an interval $[t_0, t]$ denoted by $v_S(t)$, is obtained by replacing $|f(s_i) - f(s_{i-1})|$ with $d_H(S(s_i), S(s_{i-1}))$ in the definition of the variation of $f(\cdot)$, that is,

$$v_S(t) := \sup_{t_0 = s_0 < s_1 < \cdots < s_k = t} \sum_{i=1}^{k} d_H(S(s_i), S(s_{i-1}))$$

where the supremum is taken over the set of all partitions of $[t_0, t]$. We denote by $\mathcal{L}_1(I, \mathbb{R}^n; \nu)$ and $\mathcal{L}_1^{loc}(I, \mathbb{R}^n; \nu)$ the space of integrable and locally integrable functions, respectively, from the interval I to $\mathbb{R}^n$ with respect to the measure ν. If the measure is not specified then the integration is with respect to the Lebesgue measure. An absolutely continuous (AC) function $f : I \to \mathbb{R}^n$ is a function that can be written as $f(t) - f(t_0) = \int_{t_0}^t \dot{f}(s) ds$ for any $t_0, t \in I$, $t_0 \leq t$, and some $\dot{f} \in \mathcal{L}_1(I, \mathbb{R}^n)$, which is considered as its derivative.

2.1 Non-convex Analysis

To formally define the notion of a normal cone for non-convex sets, the definition of normal cone from convex analysis is extended as follows.

Definition 1 (Fréchet Normals [34])**.** For a closed set $S \subset \mathbb{R}^n$, and $x \in S$, the vector $w \in \mathbb{R}^n$ is called a Fréchet normal to the set S at x if, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$\langle w, x' - x \rangle \leq \epsilon |x' - x| \qquad \forall\, x' \in S, \quad |x - x'| < \delta. \tag{3}$$

The set of all Fréchet normals at a point $x \in S$ form a cone denoted by $\mathcal{N}(S; x)$.

Definition 2 (Uniformly Prox-Regular Set [13, 39])**.** A set S is called uniformly prox-regular with constant $1/r$, or simply r-prox-regular, if for each $x \in S$, and each $w \in \mathcal{N}(S; x)$ with $|w| < 1$, it holds that $\text{prox}_S(x + rw) = \{x\}$, that is, x is the unique nearest vector to $x + rw$ in the set S.

Thus, it follows from the definition that S is an r-prox-regular set, if and only if, for each $x, x' \in S$, and each $w \in \mathcal{N}(S; x)$, with $|w| < 1$, we have

$$|rw|^2 = |x + rw - x|^2 < |x + rw - x'|^2 = |x - x'|^2 + 2\,\langle rw, x - x' \rangle + |rw|^2,$$

or equivalently for each $w \in \mathcal{N}(S; x)$,

$$\left\langle \frac{w}{|w|}, x - x' \right\rangle \geq -\frac{1}{2r}|x - x'|^2, \qquad \forall\, x' \in S. \tag{4}$$

In the above inequality, if we let $r \to \infty$, then the expression on the right-hand side becomes zero and we see that w is the normal vector at $x \in S$ in the classical sense of convex analysis. For that reason, we say that the case $r \to \infty$ corresponds to S being convex. In our development, the convex sets will be treated as a particular case of the r-prox-regular sets by taking $r \to \infty$.

Remark 1. For closed sets, one may find various notions of normals such as proximal normals, Clarke normals, and limiting (Mordukhovich) normals. It has been proven that, see for example [5, Theorem 3.2], these notions of normals coincide in case the set is uniformly prox-regular. So we choose to denote the normal cone simply by $\mathcal{N}(S; x)$, implicitly assuming that it may refer to any of the existing notion of a normal cone.

We now recall a fundamental result from non-convex analysis that gives a characterization of the cone normal to the preimage of a set under a constraint qualification. It is used later in rewriting the system dynamics under a coordinate transformation. For proof, see [40, Theorem 6.14].

Lemma 1 (Chain rule)**.** *Consider a nonempty, closed, r-prox-regular set $S \subseteq \mathbb{R}^l$, $r > 0$, and a linear map $H : \mathbb{R}^n \to \mathbb{R}^l$, so that $S \subseteq \operatorname{range} H$. Let $S' := H^{-1}(S)$, and assume that the following constraint qualification holds:*

$$\textit{For each } z \in S' \textit{ and } w \in \mathcal{N}(S; Hz), H^\top w = 0 \textit{ only if } w = 0. \tag{5}$$

Then, for each $z \in \mathbb{R}^n$, and $v = Hz$, it holds that

$$\mathcal{N}(S'; z) := \{H^\top w \,|\, w \in \mathcal{N}(S; v)\} = H^\top \mathcal{N}(S; Hz). \tag{6}$$

Remark 2. If H has full row rank, then the constraint qualification (5) holds automatically. However, (5) does not imply that H has full row rank. An equivalent way of expressing (5) is

$$\ker(H^\top) \cap \mathcal{N}(S; Hz) = \{0\}, \qquad \forall\, z \in H^{-1}(S).$$

Two further properties of the prox-regular sets required for stability analysis appear in Lemma 3 and Lemma 4a. These properties are derived using the metric regularity of a linear surjective map stated in the following lemma:

Lemma 2. *Consider a linear map $H : \mathbb{R}^n \to \mathbb{R}^l$ and a nonempty closed set $S \subseteq H$, then it holds that*

$$d(x, H^{-1}(S)) \leq \frac{1}{\sigma_H^+} d(Hx, S), \qquad \text{for} \quad x \in \mathbb{R}^n, \tag{7}$$

where σ_H^+ denotes the least positive singular value of H.

To arrive at (7), one observes that the regularity modulus of a linear mapping is $\sup\{d(0, H^{-1}y) \,|\, y \in \mathbb{R}^l, |y| = 1\}$, see [15, Example 1.1]. Comparing this characterization to the definition of the singular values, we immediately obtain (7).

The next result states that the preimage of a prox-regular set under a metric regular map is also prox-regular. The proof, which is formally developed in Appendix A, is based on relating the normal cone to a set with the subdifferential of the distance function to that set and using the chain rule for subdifferentials.

Lemma 3. *Consider a nonempty, closed, r-prox-regular set $S \subset \mathbb{R}^l$, $r > 0$, and a linear map $H : \mathbb{R}^n \to \mathbb{R}^l$, so that S is in the range space of H. Then the set $S' := H^{-1}(S)$ is uniformly r'-prox-regular with $r' := r\sigma_H^+/\|H\|^2$.*

The last lemma in this section states how the variation of a set changes under the preimage of a metric regular map, and is also proved in Appendix A.

Lemma 4a. *For a multivalued function $S : [t_0, \infty) \rightrightarrows \operatorname{range} H$, assume that $v_S(\cdot)$ is locally absolutely continuous. Let $S'(t) := H^{-1}(S(t))$, then $v_{S'}(\cdot)$ is also locally absolutely continuous and furthermore, $\dot{v}_{S'}(t) \leq \frac{1}{\sigma_H^+} \dot{v}_S(t)$, for Lebesgue almost all $t \in [t_0, \infty)$.*

2.2 Solution Concepts

To study the solution of system (1), we consider $S : [t_0, \infty) \rightrightarrows \mathbb{R}^n$ and introduce the following hypotheses:

(H1) There exists a constant $r > 0$ such that, for each $t \in [t_0, \infty)$, $S(t)$ is a non-empty, closed and r-prox-regular set.

(H2) The function $v_S(\cdot) : [t_0, \infty) \to \mathbb{R}_+$ is locally absolutely continuous and $|\dot{v}_S(t)|$ is bounded by v for all t except for a set of Lebesgue measure zero, i.e., $\operatorname{ess\,sup}_{t \geq t_0} |\dot{v}_S(t)| = v$.

Let us state a general result on the existence of solutions for perturbed Moreau's sweeping process (1).

Theorem 1 (AC solutions [16, Theorem 1])**.** *Consider system* (1) *over the interval $I := [t_0, \infty)$ and assume that $f(t, x)$ satisfies the following assumptions:*

(A1) *$f(\cdot, x)$ is a Lebesgue-measurable function for each $x \in \mathbb{R}^n$,*

(A2) *There exists a nonnegative function $h_1^M(\cdot) \in \mathcal{L}_1^{loc}(I, \mathbb{R})$ such that for each $t \in I$, and $|x_i| \leq M$, $i = 1, 2$,*

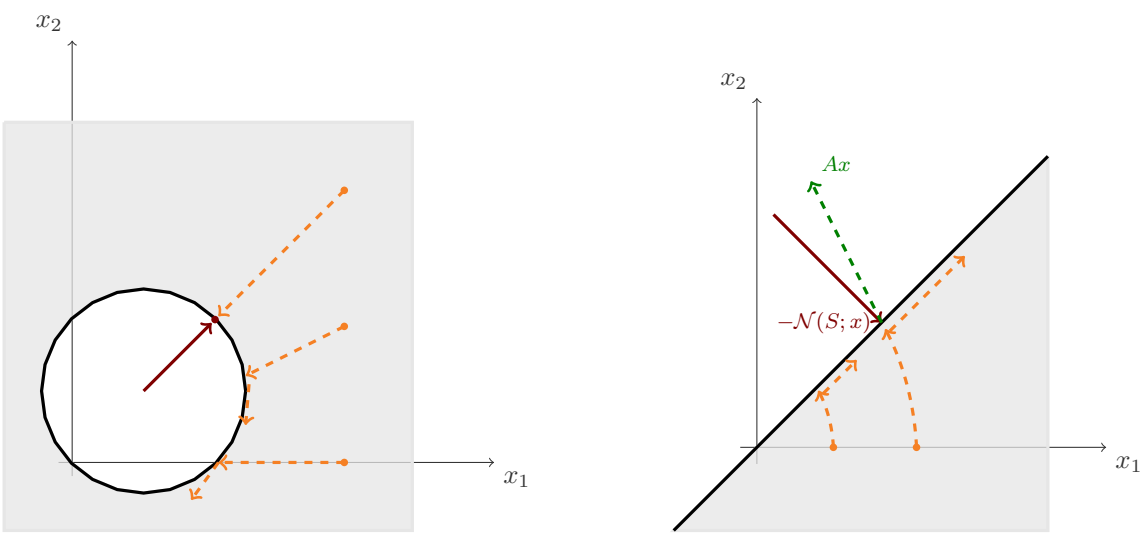
$$|f(t, x_1) - f(t, x_2)| \leq h_1^M(t)|x_1 - x_2|.$$

(A3) *There exists a nonnegative function $h_2(\cdot) \in \mathcal{L}_1^{loc}(I, \mathbb{R})$ such that $|f(t, x)| \leq h_2(t)(1 + |x|)$ for all $x \in \cup_{s \in I} S(s)$.*

Then under the hypotheses **(H1)** *and* **(H2)***, for each $x_0 \in S(t_0)$, there exists a unique locally absolutely continuous solution $x(\cdot)$ that satisfies* (1)*, and $x(t) \in S(t)$ for all $t \in I$. Furthermore, the following bound holds:*

$$|\dot{x}(t) + f(t, x(t))| \leq |f(t, x(t))| + |\dot{v}_S(t)| \quad \text{a.e. } t \in I. \tag{8}$$

Remark 3. In the formulation of Theorem 1 in [16], the bound on right-hand side in (8) was expressed differently in terms of certain system parameters. However, the more compact expression used in (8) appears in [16, page 358, eq. (3.15)], which we find more suitable for results in this paper. Also, the interval I was taken to be compact in [16], but the result can be applied repeatedly over the intervals $[t_0, t_0+1]$, $[t_0+1, t_0+2]$, and so on, to obtain the statement written here.



(a) The system is not globally asymptotically stable when the constraint set is nonconvex and A is Hurwitz.

(b) The trajectories can grow unbounded even with A Hurwitz, and S convex.

Figure 2: The trajectories of system (9) for two different cases.

3 Absolutely Continuous Solutions

For our initial results, we start with the following system class where the state admits absolutely continuous solutions:

$$\dot{x}(t) = Ax(t) - Gw_t \quad \text{for almost all } t \in [t_0, \infty), \tag{9a}$$

$$w_t \in \mathcal{N}(S(t); Hx(t)) \tag{9b}$$

where $x(t) \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times l}$, and $H \in \mathbb{R}^{l\times n}$. The initial state $x(t_0)$ is assumed to satisfy $Hx(t_0) \in S(t_0)$. For such systems, the following is assumed so that the lemmas developed in the previous section could be applied.

Assumption 1. The matrix $H \in \mathbb{R}^{l\times n}$ satisfies (5), and $S(t) \subseteq \mathbb{R}^l$ is contained in the range space of H for all $t \geq t_0$.

The following example motivates how the stability of a constrained dynamical system raises some interesting issues that are not seen in the case of unconstrained dynamical systems.

Example 1. In this example, we consider S to be a complement of an open disk such that the origin is contained in the boundary of S. For example, $S := \{(x_1, x_2) \in \mathbb{R}^2 \,|\, (x_1 - 1)^2 + (x_2 - 1)^2 - 2 \geq 0\}$, which is r-prox-regular for $r < \sqrt{2}$. Take $A := \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $G = H = I_{2\times 2}$. Then it is seen that no state trajectory starting from the initial condition $x_1(0) = x_2(0) > 0$ converges to the origin. This is because when such a state trajectory hits the boundary of the set, the only way it can remain in the set is if it acquires static equilibrium at the point of contact with the boundary of the set. But from any other initial condition, the resulting state trajectory continues to slide along the boundary of the set, hence continuing its motion towards the origin. See Figure 2(a) for illustration of sample trajectories of this system.

It is useful to recall at this point that even in the case when A is Hurwitz, and S is convex, it may be that the system is not asymptotically stable. For example, take $S := \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_2 \leq x_1\}$, and $A := \begin{bmatrix} \gamma & -1 \\ 2 & \gamma \end{bmatrix}$, for sufficiently small $\gamma < 0$. Then, even though A is Hurwitz, the trajectories of system (9)

are not necessarily bounded, see Figure 2(b). The case of S being a time-invariant, closed-convex cone was studied thoroughly in [21] where one can find an example of a matrix A with eigenvalues on imaginary axis, and a closed convex cone S that result in an unstable system (see [21, Table I]).

Using these examples as motivation, we propose sufficient conditions for asymptotic stability of system (9) where we allow $S(\cdot)$ to be time-varying and nonconvex-valued. The primary difference compared to the stability conditions proposed for static, and convex valued S is that the asymptotic stability no longer holds globally, and in our work we also compute estimates for the basin of attraction of system (9).

3.1 Sufficient Conditions for Stability

Before stating the result on stability of system (9), we recall some terminology. The system is called asymptotically stable (with respect to the origin) if *a)* it is Lyapunov stable, that is, for every $\epsilon > 0$ there exists $\delta > 0$ such that $|x(t)| < \epsilon$ for all $t \geq t_0$, whenever $|x(t_0)| < \delta$, and *b)* $\lim_{t\to\infty} |x(t)| = 0$. An asymptotically stable system has the basin of attraction $\mathcal{R}$, if for every $x(t_0) \in \mathcal{R}$, the corresponding trajectory converges to the origin.

Theorem 2. *Consider system* (9) *under the hypotheses* **(H1)**, **(H2)** *and Assumption 1. Suppose there exists a symmetric positive definite matrix P that satisfies the following for some $\theta > 0$:*

$$A^\top P + PA \leq -\theta P \tag{10}$$

$$PG = H^\top. \tag{11}$$

For $0 < \beta < 1$, define

$$\mathcal{R}_\rho := \left\{x \in \mathbb{R}^n \,|\, x^\top P x \leq \rho^2\right\}, \qquad \rho := \frac{\beta\,\theta\,r}{b\,\|RAR^{-1}\|}, \tag{12}$$

where R is the symmetric positive definite matrix such that $P = R^2$, $\overline{H} := HR^{-1}$, and $b := \frac{\|\overline{H}\|^2}{\sigma^+_{\overline{H}}}$. If θ is large enough such that

$$(1-\beta)\theta > \varepsilon + \frac{b}{r\sigma^+_{\overline{H}}} v, \tag{13}$$

and $0 \in S(t)$ for all $t \geq t_0$, then system (9) *is asymptotically stable and the basin of attraction contains the set $\mathcal{R}_\rho \cap H^{-1}(S(t_0))$.*

Proof. We proceed with the proof in four steps.
Step 1: Let $z := Rx$, then system (9) in new coordinates becomes

$$\dot{z}(t) = RAR^{-1}z(t) - RGw_t$$
$$w_t \in \mathcal{N}(S(t); HR^{-1}z(t)).$$

From (11), we have $RG = R^{-1}H^\top$, so that system (9) is equivalently written as:

$$\dot{z}(t) = RAR^{-1}z(t) - \overline{w}_t \tag{14a}$$
$$\overline{w}_t \in \mathcal{N}(S'(t); z(t)), \tag{14b}$$

where $S'(t) = \{z \in \mathbb{R}^n \,|\, HR^{-1}z \in S(t)\}$ is r'-prox-regular with $r' := r\sigma^+_{\overline{H}}/\|\overline{H}\|^2$ due to Lemma 3. From Theorem 1, it follows that (14) admits a unique locally absolutely continuous solution over $[t_0, \infty)$ and from (8), we have

$$|\overline{w}_t| \leq |RAR^{-1}z(t)| + |\dot{v}_{S'}(t)|. \tag{15}$$

Step 2: Consider the Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}_+$ defined as $V(z) = z^\top z$, then $V(\cdot)$ is continuously differentiable and its derivative along the trajectories of (14) satisfies the following for almost all $t \in [t_0, \infty)$:

$$\begin{aligned}\dot{V}(z(t)) &= z(t)^\top (R^{-1}A^\top R + RAR^{-1})z(t) - 2z(t)^\top \overline{w}_t \\ &\leq -z(t)^\top (R^{-1}A^\top R + RAR^{-1})z(t) + \frac{1}{r'}|\overline{w}_t| \cdot |z(t)|^2,\end{aligned}$$

where the last inequality was obtained by applying (4), and using the fact that $0 \in S'(t)$ by assumption, and $z(t) \in S'(t)$ for $z(\cdot)$ satisfying (14). Since equation (10) is equivalent to $R^{-1}A^\top R + RAR^{-1} \leq -\theta I$, using the bound on $|\overline{w}_t|$ from (15) and $|\dot{v}_{S'}|$ from Lemma 4a, we get

$$\begin{aligned}\dot{V}(z(t)) &\leq -\theta z(t)^\top z(t) + \frac{1}{r'}(|RAR^{-1}z(t)| + |\dot{v}_{S'}(t)|) \cdot |z(t)|^2 \\ &\leq -\theta |z(t)|^2 + \frac{b}{r}\left(\|RAR^{-1}\| \cdot |z(t)| + \frac{1}{\sigma^+_{\overline{H}}}|\dot{v}_S(t)|\right)|z(t)|^2 \\ &\leq -\left(\theta - \frac{bv}{r\sigma^+_{\overline{H}}}\right)|z(t)|^2 + \frac{b}{r}\|RAR^{-1}\| \cdot |z(t)|^3 \\ &\leq -(\varepsilon + \beta\,\theta)\,|z(t)|^2 + \frac{b}{r}\|RAR^{-1}\| \cdot |z(t)|^3 \qquad (16)\end{aligned}$$

where we substituted $r' = r\sigma^+_{\overline{H}}/\|\overline{H}\|^2 = r/b$ in the second inequality, and (13) was used to derive the last inequality.

Step 3: If $R^{-1}z(t_0) \in \mathcal{R}_\rho$, then $R^{-1}z(t) \in \mathcal{R}_\rho$, for all $t \geq t_0$. This follows due to absolute continuity of $V(z(\cdot))$. Assume this is not the case, then there exist $0 < \delta < 1$, and a time $\bar{t} > t_0$ such that $V(z(\bar{t})) = \rho^2 + \frac{\delta^2 r^2 \varepsilon^2}{4b^2\|RAR^{-1}\|^2}$. Let $\bar{t}$ be the minimal such time for a fixed δ. Then, for every t in a neighborhood of $\bar{t}$, it holds that $V(z(t)) \leq \rho^2 + \frac{r^2\varepsilon^2}{4b^2\|RAR^{-1}\|^2}$, and hence $|z(t)| \leq \rho + \frac{r\,\varepsilon}{2\,b\|RAR^{-1}\|}$, which in turn implies using (12) and (16) that

$$\dot{V}(z(t)) \leq -\frac{\varepsilon}{2}|z(t)|^2$$

for almost all t in a neighborhood of $\bar{t}$. It then follows that there exists $t \in (t_0, \bar{t})$ such that

$$V(z(t)) > V(z(\bar{t}))$$

which contradicts the minimality of $\bar{t}$.

Step 4: For $x(t_0) \in H^{-1}(S(t_0)) \cap \mathcal{R}_\rho$, it follows from the previous step that $|z(t)| \leq \rho$, for all $t \geq t_0$, and for almost all $t \geq t_0$, (16) yields

$$\dot{V}(z(t)) \leq -\varepsilon V(z(t)).$$

By comparison lemma and integration, $V(z(t)) \leq e^{-\varepsilon(t-t_0)}V(z(t_0))$, for $t \geq t_0$ and the solution $z(\cdot)$ of system (14) with initial condition $R^{-1}z(t_0) \in H^{-1}(S(t_0)) \cap \mathcal{R}_\rho$. The foregoing relation guarantees that (14) is stable in the sense of Lyapunov, and also $\lim_{t\to\infty} z(t) = 0$; hence (14) is asymptotically stable. The matrix P being positive definite guarantees that R is invertible, so that asymptotic stability is preserved under the proposed change of coordinates, and the basin of attraction of system (9) contains the set $\mathcal{R}_\rho$ as claimed in the theorem statement. □

The conditions for stability given in (10), (11) mean that the triplet (A, G, H) is dissipative in Willem's sense [45]. Dissipativity has always been an essential property of the linear part in Lur'e systems depicted in Fig. 1. This dissipative relation allows the change of variables $z = Rx$ in the proof of Theorem 2, which was first introduced in [7] in the context of differential inclusions and used for the sake of analysis in [8, 10, 21].

Example 2 (Revisiting Example 1)**.** For this example, we may take $P = I$, so that $R = I$, and $\theta = 2$. This gives $\|RAR^{-1}\| = 1$, and since the set is not moving, we let $\dot{v}_S = 0$. Condition (13) is satisfied for any $\beta \in (0, 1)$. It then follows that we may take

$$\mathcal{R}_\rho := \{x \in \mathbb{R}^n \mid |x| \leq 2\sqrt{2}\,\beta\}.$$

Thus, the region of attraction for this problem is $S \cap \mathcal{R}_\rho$.

3.1.1 Convex Case

As already pointed out, every convex set is an r-prox-regular set with r being arbitrarily large. Thus, we would like to see what shape the result of Theorem 2 acquires when the underlying sets are convex. It is noted that if $S(\cdot)$ is constant and convex-valued, then the dynamics in (9) are equivalently written as a system of differential variational inequalities, for which the stability has been studied in [21].

Corollary 1. *Consider system* (9) *under the hypotheses* **(H1)**, **(H2)** *and Assumption 1. Suppose that* $S(t)$ *is convex,* $0 \in S(t)$*, for each* $t \geq t_0$*. If there exist positive definite matrices* $P > 0$*, and a scalar* $\theta > 0$ *that satisfy* (10) *and* (11)*, then the origin of* (9) *is asymptotically stable with the basin of attraction equal to* $H^{-1}(S(t_0))$*.*

The proof is a direct consequence of Theorem 2 since one can let $r \to \infty$ for convex sets. As a result, $\mathcal{R}_\rho$ is the entire state space, so that every admissible initial condition leads to a trajectory converging to the origin.

3.1.2 Linear Complementarity Systems

Linear complementarity systems form an important class of nonsmooth dynamical systems and have been widely used in the modeling of physical systems such as electrical circuits. The general framework proposed in this paper covers a certain subclass of such dynamical systems described as follows:

$$\dot{x}(t) = Ax(t) + H^\top w_t \tag{17a}$$

$$0 \leq w_t \perp v(t) = Hx(t) + Du(t) \geq 0. \tag{17b}$$

Using a basic result from convex analysis, one may write

$$0 \leq w_t \perp Hx(t) + Du(t) \geq 0 \Leftrightarrow -w_t \in \partial\psi_Q(Hx(t) + Du(t)),$$

where $Q = \mathbb{R}^m_+$, and $\psi_Q(\cdot)$ denotes the indicator function of the set Q, i.e., $\psi_Q(x) = 0$ if $x \in Q$ and $\psi_Q(x) = +\infty$ otherwise, while ∂ denotes the subdifferential of convex analysis. For each $t \in [0, \infty)$, the closed set $S(t) := \{x \in \mathbb{R}^n \mid Hx + Du(t) \geq 0\}$ and $\mathbb{R}^+_m$ are convex polyhedral sets and $\psi_{S(t)}(x) = (\psi_{\mathbb{R}^+_m - Du(t)} \circ H)(x)$. Thus, using the chain rule, $H^\top \partial\psi_{\mathbb{R}^+_m}(Hx(t) + Du(t)) = \partial\psi_{S(t)}(x(t)) = \mathcal{N}(S(t); x(t))$. Using this relation, system (17) can be equivalently described by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) - w_t, \\ w_t &\in \mathcal{N}(S(t), x(t)). \end{aligned}$$

which is of the same form as (9). It is noted that $v_S(\cdot)$ is a locally absolutely continuous function when $u(\cdot)$ is locally absolutely continuous and that $S(t)$ is a polyhedral set, and hence convex, for each t. The stability conditions in Theorem 2 boil down to the existence of a solution to equations (10),(11), which in turn requires the symmetric part of A to be Hurwitz. This observation is consistent with the stability results obtained in [11, Remark 3.1]. In general, a complementarity system with $v(t)$ in (17b) of the form $v(t) = Hx(t) + Ew(t) + Du(t)$ can also be described as a sweeping process, where $E \geq 0$ has a special structure [8].

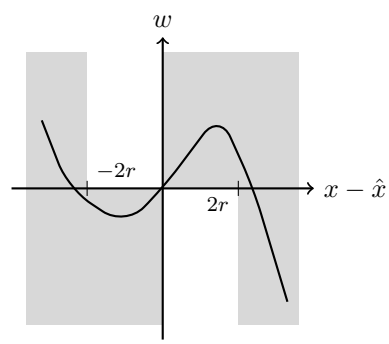


Figure 3: Each element of the Fréchet normal cone satisfies the sector condition locally in scalar case.

3.1.3 Scalar Case

Making further comparisons with the convex case, we recall that a normal vector $w \in \mathcal{N}(S; x)$, when S is convex, satisfies the following inequality:

$$\langle w, x' - x \rangle \leq 0 \qquad \forall x' \in S,$$

which is similar to the sector condition imposed on the feedback nonlinearity in determining the absolute stability of a Lur'e system.

For an r-prox-regular set, we merely have the characterization (4) for normal vectors, and it follows from (4) that[1], in scalar case, for $x \in S \subseteq \mathbb{R}$,

$$\langle -w, x' - x \rangle \geq 0 \qquad \text{if } |x - x'| < 2r.$$

Thus, the nonlinearity in the system due to the state constraints actually satisfies the monotonocity condition locally. Also, in the scalar case, a closed set is r-prox-regular if it is a disjoint union of closed intervals where the distance between one interval and the other is greater than $2r$. Thus, in the light of Theorem 2, and the above observation, a scalar system modeled as (9) is asymptotically stable if one of the components contains the origin at all times and the initial condition is within the interval containing the origin.

3.2 Observer Design

We now shift our focus to designing observers for the differential inclusions considered in (9). We suppose that the output equation associated with system (9) is the following one:

$$y(t) = Cx(t) \tag{18}$$

where $C \in \mathbb{R}^{p \times n}$ and it is assumed that $p \leq n$. For our observer design, it will be assumed throughout that the state trajectory stays bounded.

Assumption 2. The state trajectory $x(\cdot)$, that satisfies (9), is bounded by a constant M at all times, that is, $\sup_{t \geq t_0} |x(t)| \leq M$.

Two different approaches are adopted for observer design: In the first case (Section 3.2.1), we obtain an estimate $\hat{x}(\cdot)$ such that $\lim_{t \to \infty} |x(t) - \hat{x}(t)| = 0$, under the assumption that the initial value of the state estimation error $|x(t_0) - \hat{x}(t_0)|$ can be chosen to be small enough (in the sense to be made precise later).

[1] If $w, x, x' \in \mathbb{R}$ and $w \neq 0$, then inequality (4) can be equally written as

$$-w(x' - x) + \frac{\text{sign}(w)w}{2r}(x' - x)^2 \geq 0.$$

In the second approach (Section 3.2.2), there is no a priori bound on the initial value of state estimation error, and the observer is designed to achieve *practical* convergence, that is, for every $\epsilon > 0$, there exist an estimator and $T > t_0$ such that for all $t \geq T$, we have $|x(t) - \hat{x}(t)| < \epsilon$. These two approaches are then combined in Section 3.2.3 to obtain an observer that results in asymptotic convergence of the error dynamics to the origin without any bounds on the initial value of the state estimation error.

3.2.1 Locally convergent observer

In the previous section, we saw that even if A is Hurwitz, the system is asymptotically stable only locally in the presence of non-convex state constraints. For this reason, the Luenberger-like observer proposed in this section generates converging estimates only when the initial error is small enough. The observer we propose, is the following dynamical system:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + L(y(t) - C\hat{x}(t)) - G\hat{w}_t \tag{19a}$$

$$\hat{w}_t \in \mathcal{N}(S(t); H\hat{x}(t)) \tag{19b}$$

where the initial condition satisfies $H\hat{x}(t_0) \in S(t_0)$. We next state the criteria to compute the output-injection gain L, and derive the bound on the initial value of the error between $x(\cdot)$ and $\hat{x}(\cdot)$ so that the state estimate of (19) converges to the true state. In order to state the result, we let $\widetilde{x} := x - \hat{x}$ denote the state estimation, whose time derivative satisfies the following equation for almost all $t \in [t_0, \infty)$:

$$\dot{\tilde{x}}(t) = (A - LC)\tilde{x}(t) - G(w_t - \hat{w}_t), \tag{20}$$

where $w_t, \hat{w}_t$ are given by (9b) and (19b), respectively.

Proposition 1. *Consider system* (9) *under the hypotheses* **(H1)**, **(H2)** *and suppose that Assumptions 1 and 2 hold. If there exist some constants* $\varrho, \theta, \varepsilon, c_1, c_2 > 0$, $\beta \in (0, 1)$, *and a matrix* P *that satisfy the following inequalities:*

$$c_1 I \leq P \leq c_2 I \tag{21a}$$

$$A^\top P + PA - 2\varrho C^\top C \leq -\theta P \tag{21b}$$

$$PG = H^\top \tag{21c}$$

$$(1 - \beta)\theta > \varepsilon + \frac{c_2}{c_1}\frac{2\|H\|^2}{r\sigma_H^+}\left(M\|A\| + \frac{v}{\sigma_H^+}\right), \tag{21d}$$

then choosing $L = \varrho P^{-1}C^\top$ *renders the error dynamics* (20) *asymptotically stable, and for every* $\tilde{x}(t_0) \in \mathcal{R}_{\widetilde{\rho}}$, *defined as*

$$\mathcal{R}_{\widetilde{\rho}} := \left\{\widetilde{x} \in \mathbb{R}^n \,|\, \widetilde{x}^\top P \widetilde{x} \leq \widetilde{\rho}^2\right\}, \quad \text{and} \quad \widetilde{\rho} := \frac{\beta\,\theta\,r}{b\,\|R(A - LC)R^{-1}\|} \tag{22}$$

in which $R = P^{1/2}$, $\overline{H} = HR^{-1}$, *and* $b := \frac{\|\overline{H}\|^2}{\sigma_{\overline{H}}^+}$, *we have* $\lim_{t\to\infty} |\tilde{x}(t)| = 0$.

Proof. Choosing R such that $R^2 = P$, and introducing the coordinate transformation $z = Rx$ yields (14), along with $y(t) = CR^{-1}z(t)$. A similar transformation for the observer with $\hat{z} = R\hat{x}$ gives:

$$\dot{\hat{z}}(t) = R(A - LC)R^{-1}\hat{z}(t) + RLCR^{-1}z(t) - \widehat{\overline{w}}_t \tag{23a}$$

$$\widehat{\overline{w}}_t \in \mathcal{N}(S'; \hat{z}(t)), \tag{23b}$$

where $S'(t) := (HR^{-1})^{-1}(S(t)) = RH^{-1}(S(t))$ is r'-prox-regular, for $r' = r\sigma_{\overline{H}}^+/\|\overline{H}\|^2 = r/b$, due to Lemma 3. Under Theorem 2, the observer is thus well-posed and admits a unique locally absolutely continuous solution since $H\hat{x}(t_0) \in S(t_0)$, or equivalently, $\hat{z}(t_0) \in S'(t_0)$.

Choose the candidate Lyapunov function to be $V(\tilde{z}) = \tilde{z}^\top \tilde{z}$, where $\tilde{z} = R\tilde{x}$, then using (21b), the time derivative of $V(\cdot)$ along the the trajectories of (20) for almost all $t \geq t_0$ is computed as follows:

$$\dot{V}(\tilde{z}(t)) = -\tilde{z}(t)^\top \theta \tilde{z}(t) - 2(z(t) - \hat{z}(t))^\top (\overline{w}_t - \widehat{\overline{w}}_t) \tag{24a}$$

$$\leq -\theta |\tilde{z}(t)|^2 + \frac{1}{r'}(|\overline{w}_t| + |\widehat{\overline{w}}_t|)|z(t) - \hat{z}(t)|^2 \tag{24b}$$

$$\leq -\theta |\tilde{z}(t)|^2 + \frac{b}{r}\big(|RAR^{-1}z(t)| + |RAR^{-1}\hat{z}(t) + RLCR^{-1}\tilde{z}(t)| + \frac{2}{\sigma_{\overline{H}}^+}\dot{v}_S(t)\big)|\tilde{z}(t)|^2, \tag{24c}$$

where we arrived at (24b) using (4), and (24c) was obtained by substituting the value of r' obtained from Lemma 3. Also, in arriving at (24c), the upper bounds on $|\overline{w}_t|$, and $|\widehat{\overline{w}}_t|$, were obtained by applying (8) to system (14), and (23), respectively. Using the inequalities [6, Corollary 9.6.6], $\sigma_H^+ \sigma_{\min}(R^{-1}) \leq \sigma_{\overline{H}}^+ \leq \sigma_H^+ \sigma_{\max}(R^{-1})$, and the fact that the eigenvalues of R and R^{-1} are square roots of eigenvalues of P and P^{-1}, respectively, we obtain the following using (21a) and (21d):

$$\dot{V}(\tilde{z}(t)) \leq -\theta |\tilde{z}(t)|^2 + 2\frac{b}{r}\big(|RAR^{-1}z| + \frac{v}{\sigma_{\overline{H}}^+}\big)|\tilde{z}(t)|^2 + \frac{b}{r}\|R(A - LC)R^{-1}\| \cdot |\tilde{z}(t)|^3 \tag{24d}$$

$$\leq -\theta |\tilde{z}(t)|^2 + 2\frac{c_2}{c_1}\frac{\|H\|^2}{r\sigma_H^+}\big(\|A\|M + \frac{v}{\sigma_H^+}\big)|\tilde{z}(t)|^2 + \frac{b}{r}\|R(A - LC)R^{-1}\| \cdot |\tilde{z}(t)|^3 \tag{24e}$$

$$\leq -(\varepsilon + \beta\theta)|\tilde{z}(t)|^2 + \frac{b}{r}\|R(A - LC)R^{-1}\| \cdot |\tilde{z}(t)|^3, \tag{24f}$$

where we used $|R^{-1}z(t)| = |x(t)| \leq M$ in (24e) and the condition (21d) given in the theorem statement in (24f). One can now follow the same reasoning as demonstrated in Step 3 of the proof of Theorem 2 to show that $|\tilde{z}(t_0)| \leq \tilde{\rho}$ implies that $|\tilde{z}(t)| \leq \tilde{\rho}$ for all $t \geq t_0$, where $\tilde{\rho}$ is defined in (22). Hence, we have the following expression for $\dot{V}(\tilde{z}(t))$ for every trajectory $\tilde{z}(\cdot)$ starting with the initial condition $|\tilde{z}(t_0)| \leq \tilde{\rho}$:

$$\dot{V}(\tilde{z}(t)) \leq -\varepsilon V(\tilde{z}(t)),$$

and hence $\tilde{z}(\cdot)$ dynamics are asymptotically stable. Since the stability is preserved under coordinate transformation, and $\tilde{x}(t_0) \in \mathcal{R}_{\tilde{\rho}}$ implies $|\tilde{z}(t_0)| \leq \tilde{\rho}$, it follows that the error dynamics (20) are asymptotically stable, and every trajectory $\tilde{x}$ starting with the initial condition $\tilde{x}(t_0) \in \mathcal{R}_{\tilde{\rho}}$ converges to the origin. □

Remark 4. In the convex case with $r \to \infty$, the lower bound on the constant c_3 shrinks and (21a), (21b) admit a solution for every observable pair (A, C). Moreover the set $\mathcal{R}_e$ is the entire state space $\mathbb{R}^n$, so that one recovers the results given in [9]. Also, for a certain class of linear complementarity systems specified in Section 3.1.2, the proposed state estimator coincides with the observer studied in [22].

Remark 5 (Positive observers for positive systems)**.** One thing to note in the proposed observer is that the state estimate obtained from the observer respects the same constraints as the actual state. As an implication, if the original system is positive, that is, $S(t) = \mathbb{R}_+^n$, for all $t \geq t_0$, and hence convex at each time instant with $G = H = I$, then the state estimate obtained from (19) has the property that each component of $\hat{x}(t)$ is nonnegative for $t \geq t_0$.

Example 3 (Local observer for motion on a circle)**.** Consider system (9) with $S := \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1^2 + x_2^2 - r^2 = 0\}$, for some $r > 0$, which represents a circle of radius r centered at the origin and is, by definition, r-prox-regular. Let the system matrices be $A = \left[\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ and $C = [1\ 1]$. To represent the motion on the circle defined by S, one may take $G = H = I_{2\times 2}$. We now use the estimator (19) and proceed to compute the output injection gain L and the basin of attraction for error dynamics $\mathcal{R}_{\tilde{\rho}}$. In order to satisfy (21c), we pick $P = I_{2\times 2}$, and hence $c_1 = c_2 = 1$. Note that, once the system is initialized in S, the state $x(t) \in S$, $\forall t \geq t_0$, and thus one may take the bound on the state trajectory x to be $M = r$, see Assumption 2. Letting $\varrho = 1$ allows us to choose $\theta = 4$ in order to satisfy (21b). Since, $\|A\| = \sqrt{2}$, we must pick β so that $4(1 - \beta) > 2\sqrt{2}$. Thus, with $L = [1\ 0]^\top$, and $\tilde{\rho} < \frac{2\beta}{r}$, Proposition 1 guarantees that the state estimate obtained from (19) converges to actual state of the system.

3.2.2 Practically convergent observer

The state estimator proposed in the previous section converges to the actual state only when the initial estimation error is small enough, and it follows the same constraints as the system state. When the initial error is not small, we are interested in designing an estimator which caters for large (but bounded) values of $|\tilde{x}(t_0)|$, and has the property that $\tilde{x}(t) \in \mathcal{R}_{\tilde{\rho}}$ after some finite time, for $\mathcal{R}_{\tilde{\rho}}$ defined in (22). If this is achieved, then one can run the local observer and obtain the asymptotic convergence to the actual state. This motivates us to design a nonlinear state estimator to reduce the state estimation error up to the desired accuracy in finite time, which is primarily adopted from the idea of applying high-gain control to attenuate disturbances [14, 44]. However, in order to implement that idea, the system is required to have relative degree one between w_t and the output y, which may be a very strict requirement. A relaxation of this condition was proposed in [19] by letting some higher order derivative of the outputs depend on w_t. The following assumption, which includes this condition, is now introduced for the development of results in this section.

Assumption 3. There exist positive integers $d_1, \ldots, d_p$ such that $d_i \geq 1$ is the smallest positive integer that satisfies $c_i A^{d_i-1} G \neq 0$, where c_i denotes the i-th row vector of matrix C. Also, letting $C_d := \mathrm{col}(c_1, \ldots, c_1 A^{d_1-1}, \ldots, c_p, \ldots, c_p A^{d_p-1})$, it is assumed that $\mathrm{rank}(C_d G) = \mathrm{rank}\, G$, where $\mathrm{col}(x, y) := \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^{1\times 2n}$ for $x, y \in \mathbb{R}^{1\times n}$. Furthermore, the matrix $H \in \mathbb{R}^{l\times n}$ has rank l and there exists $P > 0$ such that $PG = H^\top$.

The vector $y_d = C_d x$ acts as an auxiliary output, whose derivative depends upon w_t, and if y_d were directly available, we could have directly implemented the disturbance-attenuating observer. Since this is not the case, we first use a high-gain observer to estimate the derivatives of the output and then use these estimated derivatives in a disturbance-attenuating observer to reduce the state-estimation error within some ball around the origin in finite time. Inspired by the development in [27], we proceed with the following high-gain observer to estimate the vector $y_d(\cdot)$:

$$\dot{\bar{y}}_d(t) = \bar{A}\bar{y}_d(t) + \bar{L}(y(t) - \bar{C}\bar{y}_d(t)) \tag{25}$$

$$\bar{A} = \text{block diag}\,(\bar{A}_1, \ldots, \bar{A}_p), A_i \in \mathbb{R}^{d_i\times d_i}, \qquad \bar{C} = \text{block diag}\,(\bar{C}_1, \ldots, \bar{C}_p), C_i \in \mathbb{R}^{1\times d_i},$$

where $(\bar{A}_i, \bar{C}_i)$ are in Brunovsky form. The gain $\bar{L}$ is designed as

$$\bar{L} = \text{block diag}\,(\bar{L}_1, \ldots, \bar{L}_2), \qquad \bar{L}_i = \mathrm{col}(l_1^i/\epsilon, l_2^i/\epsilon^2, \ldots, l_{d_i}^i/\epsilon^{d_i}) \in \mathbb{R}^{d_i\times 1}, \tag{26}$$

where the coefficients l_j^i are defined such that the roots of $s^{d_i} + l_1^i s^{d_i-1} + \cdots + l_{d_i}^i = 0$ are in the open left-half plane for each $i = 1, \ldots, p$, and $\epsilon > 0$ is a design parameter to be specified. Using $y_{d_{ij}}(t)$ to denote the $(d_{i-1} + j)$-th element of $y_d(t)$, with $d_0 = 0$, the scaled state estimation error is introduced as follows:

$$\eta_{ij}(t) := \frac{y_{d_{ij}}(t) - \bar{y}_{d_{ij}}(t)}{\epsilon^{d_i-j}}, \quad 1 \leq i \leq p, 1 \leq j \leq d_i.$$

Define

$$\bar{D} := \text{block diag}\,(D_1, \cdots, D_p), \qquad D_i = \text{diag}\,(\epsilon^{d_i-1}, \ldots, 1) \in \mathbb{R}^{d_i\times d_i},$$

which gives

$$y_d(t) - \bar{y}_d(t) = \bar{D}\eta(t)$$

where $\eta := \mathrm{col}(\eta_{11}, \ldots, \eta_{1d_1}, \ldots, \eta_{p1}, \ldots, \eta_{pd_p})$, with $\|\bar{D}\| = 1$. Following the idea in [31, Lemma 3], we show that there exists a time $T_1(\epsilon)$, such that $|\eta(t)| \leq c\epsilon$, for some $c \geq 0$ and each $t \geq T_1(\epsilon)$. Towards this end, we start by writing the dynamics for η as follows [27]:

$$\epsilon\dot{\eta}(t) = (\bar{A} - \bar{L}\bar{C})\eta(t) + \epsilon\bar{B}w_t,$$

where $\bar{B} = \text{block diag}\,(\bar{B}_1, \ldots, \bar{B}_p)$, with $\bar{B}_i := (0, \ldots, 0, c_i A^{d_i-1} G) \in \mathbb{R}^{d_i\times 1}$. In the sequel, we would need the bound on $|w_t|$ which is derived in terms of $|\overline{w}_t|$. Under Assumption 3, there exists $P > 0$ that

satisfies (11), so that one can introduce the coordinate transformation to arrive at (14) where $|\overline{w}_t|$ admits the bound (15). Since H is assumed to have full row rank, and we have the relation $\overline{w}_t = HR^{-1}w_t =: \overline{H}w_t$ from Lemma 1, we get $|w_t| \leq \frac{1}{\sigma^+_{\overline{H}}}|\overline{w}_t|$, and hence

$$|w_t| \leq \frac{1}{\sigma^+_{\overline{H}}}\left(\|RA\| \cdot M + \frac{1}{\sigma^+_{\overline{H}}}v\right) =: \phi, \tag{27}$$

where we recall that M is the bound on $x(\cdot)$ and $v := \operatorname{ess\,sup}_{t\geq t_0} |\dot{v}_S(t)|$. Since $(\bar{A} - \bar{L}\bar{C})$ is Hurwitz, we take $\bar{P}$ to be the solution of the following Lyapunov equation:

$$(\bar{A} - \bar{L}\bar{C})^\top \bar{P} + \bar{P}(\bar{A} - \bar{L}\bar{C}) = -I,$$

and let $\bar{V}(\eta) := \eta^\top \bar{P}\eta$, whose derivative satisfies the following inequalities:

$$\begin{aligned}\dot{\bar{V}}(\eta(t)) &\leq -\frac{\sigma_{\min}(\bar{P})}{\epsilon}\bar{V}(\eta(t)) + 2\sqrt{\sigma_{\max}(\bar{P})}\|\bar{P}\bar{B}\||w_t|\sqrt{\bar{V}(\eta(t))} \\ &\leq -\frac{\sigma_{\min}(\bar{P})}{2\epsilon}\bar{V}(\eta(t)), \text{ for } \bar{V}(\eta(t)) \geq \epsilon^2 k_1,\end{aligned}$$

where $k_1 := 16\|\bar{P}\bar{B}\|^2\sigma_{\max}(\bar{P})|w_t|^2/\sigma_{\min}(\bar{P})^2$. Using the notation $d := \max_{1\leq i\leq p} d_i$, the following inequality then results from the Gronwall-Bellman lemma:

$$\bar{V}(\eta(t)) \leq \bar{V}(\eta(t_0))e^{-\sigma_{\min}(P)t/2\epsilon} \leq \frac{k_2}{\epsilon^{2d-2}}e^{-\sigma_{\min}(P)t/2\epsilon},$$

where $k_2 > 0$ is the upper bound on $\sigma_{\max}(\bar{P})|y_d(t_0) - \bar{y}_d(t_0)|^2$, which is finite since $|y_d(t)| = |C_d x(t)| \leq \|C_d\|M$ by Assumption 2. Thus, for $t > T_1(\epsilon)$, we have $\bar{V}(\eta(t)) \leq \epsilon^2 k_1$, or equivalently, $|\eta(t)| \leq \bar{c}\epsilon$, for all $t \geq T_1(\epsilon)$, where $\bar{c} := \sqrt{k_1}/\sigma_{\min}(\bar{P})$, and

$$T_1(\epsilon) = \frac{2\epsilon}{\sigma_{\min}(\bar{P})}\log\left(\frac{k_2}{k_1\epsilon^{2d}}\right).$$

We now inject $\bar{y}_d(\cdot)$ obtained from (25) in the following state estimator which is run over an interval $[t_0, T)$:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + L_1(y(t) - C\hat{x}(t)) + Gg_\delta(\bar{y}_d(t), \hat{x}(t)), \qquad \hat{x}(t_0) \in \mathbb{R}^n, \tag{28}$$

where

$$g_\delta(\bar{y}_d, \hat{x}) := \frac{F_1(\bar{y}_d - \hat{y}_d)}{|F_1(\bar{y}_d - \hat{y}_d)|\phi + \delta}\phi^2, \tag{29}$$

for some $\delta > 0$, $\hat{y}_d = C_d\hat{x}$, and $\phi \geq |w_t|$ defined in (27). The matrix L_1 in (28) is chosen to satisfy

$$(A - L_1C)^\top P_1 + P_1(A - L_1C) \leq -Q_1 \tag{30}$$

for some $Q_1 > 0$, and the matrix F_1 in (29) is chosen such that $P_1G = (F_1C_d)^\top$, which always exists under Assumption 3. The dynamics of the error $\tilde{x} = x - \hat{x}$ are:

$$\dot{\tilde{x}}(t) = (A - L_1C)\tilde{x}(t) + Gw_t - Gg_\delta(\bar{y}_d(t), \hat{x}(t)).$$

Choosing the Lyapunov function $V_1(\tilde{x}) = \tilde{x}^\top P_1\tilde{x}$, we get the following inequalities for $\dot{V}(\cdot)$ in which we have suppressed the arguments of the functions for brevity:

$$\begin{aligned}\dot{V}_1 &= \tilde{x}^\top((A - L_1C)^\top P_1 + P_1(A - LC))\tilde{x} + \tilde{x}^\top P_1Gw_t - \tilde{x}^\top P_1Gg_\delta(\bar{y}_d, \hat{x}) \\ &= -\tilde{x}^\top Q_1\tilde{x} + (F_1C_d\tilde{x} - F_1\bar{D}\eta)^\top w_t - (F_1C_d\tilde{x} - F_1\bar{D}\eta)^\top g_\delta + (F_1\bar{D}\eta)^\top w_t - (F_1\bar{D}\eta)^\top g_\delta \\ &\leq -\tilde{x}^\top Q_1\tilde{x} + \frac{|F_1C_d\tilde{x} - F_1\bar{D}\eta||w_t|(|F_1C_d\tilde{x} - F_1\bar{D}\eta|\phi + \delta) - |F_1C_d\tilde{x} - F_1\bar{D}\eta|^2\phi^2}{|F_1C_d\tilde{x} - F_1\bar{D}\eta|\phi + \delta} + (F_1\bar{D}\eta)^\top(w_t - g_\delta) \\ &\leq -\tilde{x}^\top Q_1\tilde{x} + \delta + (F_1\bar{D}\eta)^\top|w_t - g_\delta|,\end{aligned}$$

where we used that $F_1(\bar{y}_d - \hat{y}_d) = F_1(\bar{y}_d - y_d + y_d - \hat{y}_d) = F_1(-\bar{D}\eta + C_d(x - \hat{x})) = F_1 C_d \tilde{x} - F_1 \bar{D}\eta$ in (29), and the fact that $|w_t| \leq \phi$. Since $|\eta(t)| < c\epsilon$ for $t \geq T_1(\epsilon)$, it follows using $\|\bar{D}\| = 1$ and $|w_t - g_\delta| \leq 2\phi$ that

$$\dot{V}_1(\tilde{x}(t)) \leq -\sigma_{\min}(Q_1)|\tilde{x}(t)|^2 + \delta' = -\sigma_{\min}(Q_1)|\tilde{x}(t)| \left(|\tilde{x}(t)| - \frac{\delta'}{\sigma_{\min}(Q_1)} \right) \qquad \text{for} \quad t \geq T_1(\epsilon),$$

where $\delta' := \delta + 2\bar{c}\|F_1\|\epsilon\phi$, which can be made arbitrarily small by picking δ and ϵ small. Thus, $\dot{V}_1(\tilde{x}(t)) < 0$ if $|\tilde{x}(t)| > \frac{\delta'}{\sigma_{\min}(Q_1)}$. If for some pre-assigned $\rho^*, \bar{\delta} > 0$, the constants ϵ in (26) and δ in (29) are chosen such that

$$0 < \delta' < \bar{\delta} \leq \frac{\rho^{*2}\sigma_{\min}(P_1)\sigma_{\min}(Q_1)}{\sigma_{\max}(P_1)}, \tag{31}$$

then $|\tilde{x}(t)| \leq \rho^*$, for all $t \geq T$, where

$$T > T_1(\epsilon) + \frac{k_3\sigma_{\max}(P_1) - \rho^{*2}\sigma_{\min}(P_1)}{\bar{\delta} - \delta'}, \tag{32}$$

and k_3 is the finite upper bound on $|\tilde{x}(t_0 + T(\epsilon))|^2$.

To summarize, the estimator proposed in (25) and (28) have the property that $|\tilde{x}(t)| < \rho^*$ after some finite time, for some pre-specified $\rho^* > 0$. The development of this section thus leads to the following result:

Proposition 2. *Consider system* (9) *under Assumptions 1 – 3, and hypotheses* **(H1)**, **(H2)**. *The state estimate* $\hat{x}(\cdot)$ *obtained from* (25) – (29), *using the design parameters* ϵ, δ, *and* ρ^* *that satisfy* (31), *has the property that* $|x(t) - \hat{x}(t)| < \rho^*$ *for all* $t \geq T$, *where* T *is given by* (32).

3.2.3 Global Asymptotic convergence

Running the state estimator (25) – (29) makes the state estimation error small after some finite time, and in order to achieve convergence to the actual state, we can now activate the locally convergent observer (19) for the interval $[T, \infty)$. However, to run (19), we must have $\hat{x}(T) \in H^{-1}(S(T))$, which can be done by picking the nearest point to the vector $\hat{x}(T^-)$ in the set $\bar{S}(T) := H^{-1}(S(T))$. As $S(T)$ is r-prox-regular, $\bar{S}(T)$ is $\bar{r}$-prox-regular with $\bar{r} = r\sigma_H^+/\|H\|^2$ (derived in Lemma 3), and the nearest point in $\bar{S}(T)$ to the vector $\hat{x}(T^-)$ is uniquely defined if $d(\hat{x}(T^-), \bar{S}(T)) < \bar{r}$. This is done by choosing an appropriate value of ρ^* in (33). Noting from Lemma 2 that $d(\hat{x}(T^-), \bar{S}(T)) \leq \frac{1}{\sigma_H^+} d(H\hat{x}(T^-), S(T)) \leq \frac{\|H\|}{\sigma_H^+} \cdot |x(T^-) - \hat{x}(T^-)| \leq \frac{\|H\|}{\sigma_H^+}\rho^*$, and recalling the definition of $\widetilde{\rho}$ in (22), if ρ^* is chosen such that

$$\rho^* < \frac{\sigma_H^+}{\|H\| + \sigma_H^+} \cdot \min\left\{ \bar{r}, \frac{\widetilde{\rho}}{\sqrt{\sigma_{\max}(P)}} \right\}, \tag{33}$$

then $d(\hat{x}(T^-), \bar{S}(T)) < \bar{r}$, so that

$$\hat{x}(T) = \text{prox}_{\bar{S}(T)}(\hat{x}(T^-)) \tag{34}$$

is well-defined, and

$$|\hat{x}(T) - x(T)| \leq |x(T^-) - \hat{x}(T^-)| + d(\hat{x}(T^-), \bar{S}(T)) \leq \rho^* + \frac{\|H\|}{\sigma_H^+}\rho^* \leq \left(\frac{\sigma_H^+}{\|H\| + \sigma_H^+} \right) \rho^* < \frac{\widetilde{\rho}}{\sqrt{\sigma_{\max(P)}}},$$

which in turn implies that $\tilde{x}(T) \in \mathcal{R}_{\widetilde{\rho}}$, with $\mathcal{R}_{\widetilde{\rho}}$ defined in (22). Thus running the estimator (19) for the time interval $[T, \infty)$, guarantees that $\hat{x}(t) \to x(t)$ under the hypotheses of Proposition 1 without the constraints on initial value of the state estimation error. This way combining the two estimators results in semiglobal (with respect to state estimation error) convergence, and this result is summed-up as follows.

Theorem 3. *Consider system* (9) *under Assumptions 1 – 3, and hypotheses* **(H1)**, **(H2)**. *Let us suppose that there exists* P *satisfying* (21). *If the state estimate* $\hat{x}(\cdot)$ *is constructed as follows:*

- $\hat{x}(\cdot)$ *is obtained from* (25) – (29) *over a time interval* $[t_0, T)$ *using the parameters* $\bar{\delta}$ *and* ρ^* *as in* (31) *and* (33),
- $\hat{x}(T)$ *is given by* (34), *for* T *satisfying* (32),
- $\hat{x}(\cdot)$ *satisfies* (19) *over the interval* $[T, \infty)$,

then $\lim_{t\to\infty} |\hat{x}(t) - x(t)| = 0$.

4 Bounded Variations and State Jumps

In this section, we relax **(H2)** to allow $v_S(\cdot)$ to be a function of bounded variation, which may introduce discontinuities in the state trajectory. If $x : I \to \mathbb{R}^n$ is a function of bounded variation, then one can associate with it a Lebesgue-Stieltjes measure or the so-called *differential measure* dx on I. In addition, if $x(\cdot)$ is right-continuous, then we have the relation that $x(t) = x(s) + \int_{(s,t]} dx$, for $[s, t] \subset I$. Since the derivatives of functions of bounded variation do not exist in the classical sense, we use the notion of differential measure to describe the dynamics of the state trajectory[2]:

$$dx \in f(x)d\lambda - \mathcal{N}(S(t); x), \quad \text{with} \quad x(t_0) = x_0 \in S(t_0), \tag{35}$$

where λ denotes the Lebesgue measure on the interval $[t_0, \infty)$, and it is assumed that the variation of the set-valued map $S : [t_0, \infty) \rightrightarrows \mathbb{R}^n$ satisfies the following hypothesis:

(H3) The variation function associated with the set-valued map $S(\cdot)$, $v_S : [t_0, \infty) \to \mathbb{R}_+$ is *locally rcbv*.

The function $v_S(\cdot)$, being *locally rcbv*, has countably many discontinuities over the interval $[t_0, \infty)$, and we let $\mathcal{I}$ denote a countable set that indexes the discontinuities of $v_S(\cdot)$. The measure μ associated with $v_S(\cdot)$, i.e. $\mu = dv_S$, admits the following decomposition [41, Theorem 6.10]: $\mu = \mu_{\text{ac}} + \mu_{\text{sing}} + \mu_d$, where μ_{ac} is absolutely continuous with respect to Lebesgue measure, μ_{sing} is associated with a singular function which is continuous everywhere and differentiable almost everywhere with zero derivative (e.g., the Cantor function). We use the notation $\mu_c := \mu_{\text{ac}} + \mu_{\text{sing}}$ to denote the continuous part of μ. The discontinuous part is given by $\mu_d := \sum_{i\in\mathcal{I}} \mu(\{t_i\})$, where $\mu(\{t_i\})$ denotes the measure of the singleton $\{t_i\}$ and corresponds to the size of jump in $v_S(\cdot)$ at t_i. To clarify some notation later and better understand the definition of the measure μ, note that if $v_S(\cdot)$ is *locally rcbv* on $[t_0, \infty)$, and has a single discontinuity at t_i, $t_0 \le s < t_i < t < \infty$, then $\mu_d([s, t]) = \mu(\{t_i\}) = v_S(t_i^+) - v_S(t_i^-)$, and $\mu_c([s, t]) = v_S(t) - v_S(t_i^+) + v_S(t_i^-) - v_S(s)$. We remark that μ is non-negative since $v_S(\cdot)$ is nondecreasing.

By introducing **(H3)**, we allow jumps in the variation of the set-valued function $S(\cdot)$ which in turn induces jumps in the state trajectory $x(\cdot)$. The framework of absolutely continuous solutions does not capture this richer class of state trajectories, thus motivating us to consider solutions with bounded variation. It is noteworthy that **(H3)** also allows for Zeno behavior, i.e., an infinite number of jumps in the state trajectory over finite time intervals (with the accumulation of jumps to the left of accumulation times). Such functions are used to model behavior of nonsmooth mechanical systems and there are several references [7, 30] which adopt the control theoretic framework to analyze systems with solutions of bounded variation. The major difference between the stability analysis of measure differential inclusions adopted in [7, 30] and the framework of this paper is that the earlier approach uses maximal monotonicity of the multivalued map, and in our case the multivalued operator is not maximal monotone.

4.1 Solution Concepts

The details on existence and precise notion of solutions for system (35) appear in [10, 17], which we now briefly recall. The density of dx with respect to a positive Radon measure ν over an interval I is defined as:

$$\frac{dx}{d\nu}(t) := \lim_{\varepsilon\to 0} \frac{dx(I(t,\varepsilon))}{\nu(I(t,\varepsilon))}, \tag{36}$$

[2]The notation $d\lambda$ corresponds to the usual notation dt and has been used to avoid confusion later on.

where $I(t,\varepsilon) := I \cap [t-\varepsilon, t+\varepsilon]$.

Definition 3. A mapping $x : [t_0, \infty) \to \mathbb{R}^n$ is called a solution of (35) if:

(S1) the map $x(\cdot)$ is *locally rcbv*, $x(t_0) = x_0$, and $x(t) \in S(t)$ for all $t \in [t_0, \infty)$,

(S2) there exists a Radon measure ν absolutely continuously equivalent[3] to $\mu + \lambda$ such that the differential measure dx is absolutely continuous with respect to ν, $\frac{dx}{d\nu}(\cdot) \in \mathcal{L}_1^{loc}(I, \mathbb{R}^n; \nu)$, and

$$-\frac{dx}{d\nu}(t) + f(x(t))\frac{d\lambda}{d\nu}(t) \in \mathcal{N}(S(t); x(t)), \qquad \nu - \text{a.e.},\ t \in [t_0, \infty) \tag{37}$$

where

$$\frac{d\lambda}{d\nu}(t) = \lim_{\varepsilon \to 0} \frac{\lambda(I(t,\varepsilon))}{\nu(I(t,\varepsilon))} \tag{38}$$

denotes the density of Lebesgue measure λ relative to ν.

Observations: We remark some properties of the foregoing solution framework to give a better understanding of how the state trajectories of (35) evolve:

1. The state $x(\cdot)$ is discontinuous at a time instant t_i only if there is a jump in $v_S(\cdot)$ at that time. If not, then there exists t_i such that $\lim_{\varepsilon\to 0} \nu(I(t_i,\varepsilon)) = 0$, whereas $\lim_{\varepsilon\to 0} dx(I(t_i,\varepsilon))$ denotes the size of jump which does not equal zero, and (S2) is violated as $\frac{dx}{d\nu}(\cdot) \notin \mathcal{L}_1^{loc}(I, \mathbb{R}^n; \nu)$.

2. If the set valued mapping $S(\cdot)$ is not continuous at t_i, then $\mu(\{t_i\}) = d_H(S(t_i^-), S(t_i))$. In that case, $\frac{d\lambda}{d\nu}(t_i) = 0$, and (37) becomes
$$-\frac{dx}{d\nu}(t_i) \in \mathcal{N}(S(t_i); x(t_i)).$$
From (36) and recalling $x(\cdot)$ is *locally rcbv* so that $x(t_i^+) = x(t_i)$, it then follows that
$$\frac{x(t_i) - x(t_i^-)}{\mu(\{t_i\})} \in -\mathcal{N}(S(t_i), x(t_i)).$$
Since the right-hand side is a cone and $\mu(\{t_i\}) > 0$, we obtain
$$x(t_i^-) - x(t_i) \in \mathcal{N}(S(t_i), x(t_i)).$$
Thus, a possible choice for $x(t_i)$ is
$$x(t_i) \in \text{prox}_{S(t_i)}(x(t_i^-)),$$
where $\text{prox}(x, S) := \arg\min_{z\in S} |x - z|$ is uniquely defined for an r-prox-regular set S provided that $d(x, S) < r$.

3. The solution concept adopted in Definition 3 is a natural extension of the absolutely continuous case. If $v_S(\cdot)$ is absolutely continuous, then we may take $\mu = \dot{v}_S\lambda$, so that $\mu + \lambda = (1 + \dot{v}_S)\lambda$ is absolutely continuously equivalent to λ, and (37) corresponds to the usual differential inclusion (1).

Let us use these observations to formally state the conditions for existence of solution to system (35). The following result basically follows from [10, Theorem 4.4(b)] and [17, Theorem 3.1 and Corollary 3.1] where the existence of solution is proved. The property (39) stated here is not explicitly mentioned in those references, so the proof has been included in Appendix A.

[3] A measure μ is absolutely continuous with respect to another measure ν if $\nu(I) = 0$ implies $\mu(I) = 0$. The measures μ and ν are absolutely equivalent if they are absolutely continuous with respect to each other.

Theorem 4. *Consider system* (35) *over the interval* $I := [t_0, \infty)$ *where* $f(t,x)$ *satisfies the assumptions* **(A1)** – **(A3)** *listed in Theorem 1. If* **(H1)**, **(H3)** *hold, and* $\sup_{s\geq t_0} \mu(\{s\}) < r/2$, *then, for each* $x_0 \in S(t_0)$, *there exists a* locally rcbv *solution* $x(\cdot)$ *to system* (35) *that satisfies*

$$\left|\frac{dx}{d\nu}(t) - f(t,x(t))\frac{d\lambda}{d\nu}(t)\right| \leq |f(t,x(t))|\,\frac{d\lambda}{d\nu}(t) + \frac{d\mu}{d\nu}(t), \quad \nu - a.e.\ t \in I \tag{39}$$

where $\nu := \mu + \lambda$. *In addition, if* $\sup_{s\geq t_0} \mu(\{s\}) < r/4$, *then the solution satisfying* (39) *is unique.*

Remark 6. The existence of the solution claimed in Theorem 4 is a direct consequence of [17, Theorem 3.1], as it can be verified that the vector field $f(\cdot,\cdot)$ satisfies the required assumptions. The inequality (39) is stated differently here because of the different choice of ν; and in Appendix A.3 , we show how our choice of ν leads to the bound given in (39). The uniqueness result is given in [17, Corollary 3.1] for the case $f \equiv 0$, and could be modified in a straightforward manner (by redefining the function $g(\cdot)$ used in their proof to handle $f(\cdot,\cdot)$ that satisfies the assumptions **(A1)** – **(A3)**).

Remark 7. In the formulation of Theorem 4, we assume a bound on $\mu(\{t_i\})$ that constrains the size of jumps in the variation of the set-valued mapping $S(\cdot)$. This is done so that $x(t_i)$ could be obtained by projecting $x(t_i^-)$ onto $S(t_i)$ because the projections are uniquely defined *only* locally for prox-regular sets. If $\mu(\{t_i\})$ is too large, then there may be more than one possible choice for $x(t_i) \in S(t_i)$ that solves (35). Even with the constraint $\mu(\{t_i\}) < r/4$, it is possible that there is more than one solution to (35), but this constraint ensures that there is only one solution that satisfies (39). For such a solution, $x(t_i)$ is obtained by projecting $x(t_i^-)$ on the set $S(t_i)$ at jump instants, and hence this choice of $x(t_i) \in S(t_i)$ minimizes $|x(t_i) - x(t_i^-)|$. To illustrate these arguments, we recall an example from [17] in which the set-valued mapping $S(\cdot)$ is given by:

$$S(t) = \begin{cases} [0,1] \cup \{10\}, & t \in [0,1) \\ \{1, 10\}, & t \geq 1 \end{cases}$$

so that $S(t)$, for each $t \geq 0$, is r-prox regular for $r < r^* := 4.5$. Moreover, $\sup_{t\geq 0} \mu(\{t\}) = \mu(\{1\}) = 1 < \frac{r^*}{4}$, so that the condition imposed on the measure associated with the variation of $S(\cdot)$ in Theorem 4 holds. Let us consider the two functions $x^1, x^2 : [t_0, \infty) \to \mathbb{R}$ defined as:

$$x^1(t) = \begin{cases} 0, & t \in [0,1) \\ 1, & t \geq 1 \end{cases} \qquad \text{and} \qquad x^2(t) = \begin{cases} 0, & t \in [0,1) \\ 10, & t \geq 1 \end{cases}$$

then $x^1(\cdot), x^2(\cdot)$ are two different solutions to the inclusion

$$dx^i \in -\mathcal{N}(S(t); x^i) \quad \text{with} \quad x^i(0) = 0 \in S(0), \qquad i = 1, 2.$$

With ν fixed as sum of μ and λ, inequality (39) at the discontinuity instants t_i, implies that

$$|x(t_i) - x(t_i^-)| \leq \mu(\{t_i\})),$$

and for this example, it is noted that,

$$\sup_{t\geq 0} |x^1(t^+) - x^1(t^-)| = 1 < \frac{r^*}{4} \quad \text{and} \quad \sup_{t\geq 0} |x^2(t^+) - x^2(t^-)| = 10 \geq \frac{r^*}{4}.$$

Thus, *only one* solution satisfies the inequality (39) appearing in the statement of Theorem 4.

Another way to state this result is to say that *the solution with the least variation is unique*. In the rest of the paper, whenever we talk about the solution of (35), we only consider the unique solution with least variation.

4.2 Stability Considerations

From stability viewpoint, we need to analyze whether the jumps introduce any destabilizing effect. The fundamental difference between the convex and prox-regular sets is that the projection onto a convex set is a non-expansive map, that is, a Lipschitz map with Lipschitz constant at most 1. However, the projection on a prox-regular set doesn't satisfy such nice property: It is single-valued and Lipschitz only within a neighborhood of the set, with the Lipschitz constant depending on the size of neighborhood under consideration, which is greater than one (proved in Lemma 5 below). Thus, the jumps in systems with prox-regular constraints introduce discontinuities that may increase the norm of the state. The basic idea in studying the stability of such systems is to limit the number of jumps over an interval such that the system has enough time in-between the discontinuities to compensate for the destabilizing jumps. As a corollary to the main result, we use the notion of average dwell-time to characterize the frequency of jumps that preserve stability of the system.

The system class we consider is defined by the following equations:

$$dx(t) = Ax(t)d\lambda - Gw_t \tag{40a}$$

$$w_t \in \mathcal{N}(S(t); Hx(t)), \tag{40b}$$

with initial condition $x(t_0) = x_0$ satisfying $Hx_0 \in S(t_0)$. Before proceeding with the stability result, we first state a result parallel to Lemma 4a to describe a bound on the variation of the preimage of a set in the case of bounded variation. Its proof is given in Appendix A.

Lemma 4b. *Let $S(\cdot), S'(\cdot)$ and H be as in Lemma 4a. If $v_S(\cdot)$ is* locally rcbv*, then $v_{S'}(\cdot)$ is also* locally rcbv*. Moreover, $\frac{d\mu'_c}{d\nu}(t) \leq \frac{1}{\sigma_H^+}\frac{d\mu_c}{d\nu}(t)$ for ν-almost every $t \geq t_0$, and $\mu'_d(\{t_i\}) \leq \frac{1}{\sigma_H^+}\mu_d(\{t_i\})$ for every $i \in \mathcal{I}$, where $\mu := dv_S$, $\mu' := dv_{S'}$, and $\nu := \mu' + \lambda$.*

The next two lemmas characterize the size and effect of jumps in state trajectories due to discontinuities in the variation of the set valued map $S(\cdot)$, and would be used in analyzing the stability of system (40) at the jump instants.

Lemma 5. *Let S^a, S^b be r-prox-regular sets, and $\alpha := d_H(S^a, S^b) < \frac{r}{4}$. Suppose that $z_i^a \in S^a$, and $z_i^b := \mathrm{prox}_{S^b}(z_i^a) \in S^b$, for $i = 1, 2$. Then, using the notation $\gamma := \left(1 - \frac{\alpha}{r}\right)^{-2}$, we have*

$$|z_1^b - z_2^b|^2 \leq \gamma\, |z_1^a - z_2^a|^2. \tag{41}$$

Proof. Since the distance between S^a and S^b is less than r by hypothesis, the nearest point to $z_i^a \in S^a$ in the set S^b is uniquely defined and given by z_i^b, for $i = 1, 2$. We can write $z_i^b = z_i^a + \xi_i$, where $\xi_i \in \mathcal{N}(S^b, z_i^b)$, and $|\xi_i| \leq \alpha$. Then,

$$\begin{aligned}
|z_1^b - z_2^b| \cdot |z_1^a - z_2^a| &\geq \left\langle z_1^b - z_2^b, z_1^a - z_2^a \right\rangle = \left\langle z_1^b - z_2^b, \xi_1 - \xi_2 \right\rangle + |z_1^b - z_2^b|^2 \\
&\geq -\frac{1}{2r}(|\xi_1| + |\xi_2|)|z_1^b - z_2^b|^2 + |z_1^b - z_2^b|^2 \\
&\geq \left(1 - \frac{\alpha}{r}\right)|z_1^b - z_2^b|^2,
\end{aligned}$$

where (4) has been used and whence (41) follows. □

Lemma 6. *Let $\alpha_i := \mu(\{t_i\}) = d_H(S(t_i^+), S(t_i^-)) < \frac{r}{4}$, and $\gamma_i := \left(1 - \frac{\alpha_i}{r}\right)^{-2}$ for each i belonging to a countable set $\mathcal{I}$ that indexes the discontinuities of $v_S(\cdot)$ over the interval I. Using the notation $\mu_d(I) := \sum_{i \in \mathcal{I}} \alpha_i$, there exists a constant $b_d > 0$ such that*

$$\prod_{i \in \mathcal{I}} \gamma_i \leq e^{\frac{b_d}{r}\mu_d(I)}. \tag{42}$$

Proof. For the desired claim, we first find an upper bound on the series $\sum_{i\in\mathcal{I}} \frac{2r\alpha_i-\alpha_i^2}{(r-\alpha_i)^2}$. We use the fact that $0<\alpha_i<\frac{r}{4}$ for each $i\in\mathcal{I}$, so that $(r-\alpha_i)>\frac{3r}{4}$ and $(2r-\alpha_i)<2r$. This in turn gives:

$$\sum_{i\in\mathcal{I}} \frac{2r\alpha_i-\alpha_i^2}{(r-\alpha_i)^2} < \frac{32}{9r}\sum_{i=\mathcal{I}}\alpha_i = \frac{b_d}{r}\mu_d(I),$$

where $b_d=\frac{32}{9}$, which in general depends on the upper bound on $\mu(\{t_i\})$. This inequality is now used to derive the desired result as follows:

$$\prod_{i\in\mathcal{I}}\gamma_i = \prod_{i\in\mathcal{I}}\left(1-\frac{\alpha_i}{r}\right)^{-2} = \prod_{i\in\mathcal{I}}\left(1+\frac{2r\alpha_i-\alpha_i^2}{(r-\alpha_i)^2}\right) \le e^{\sum_{i\in\mathcal{I}}\frac{2r\alpha_i-\alpha_i^2}{(r-\alpha_i)^2}} \le e^{\frac{b_d}{r}\mu_d(I)}.$$

where the exponential bound was obtained using the fact that $\ln(1+x)\le x$ for $x\ge 0$. □

The bound obtained in (42) in terms of a constant b_d depends upon the upper bound of $\mu(\{t_i\})$; the smaller the value of $\mu(\{t_i\})$, the smaller the constant b_d.

Theorem 5. *Consider system* (40) *under the hypotheses* **(H1)**, **(H3)** *and Assumption 1. Suppose there exists a symmetric positive definite matrix P that satisfies* (10)–(11) *for some constant $\theta>0$, and take $R:=P^{1/2}$, $\overline{H}:=HR^{-1}$. Let*

$$\kappa(t_0,t) := -a(t-t_0) + \frac{b_c}{r\sigma^+_{\overline{H}}}\mu_c([t_0,t]) + \frac{b_d b_c}{r\sigma^+_{\overline{H}}}\mu_d([t_0,t]), \tag{43}$$

where $a:=(1-\beta)\theta$ for $0<\beta<1$, $b_c:=\frac{\|\overline{H}\|^2}{\sigma^+_{\overline{H}}}$, and b_d satisfies (42)*. If $\sup_{s\ge t_0}\mu(\{s\})<\frac{r\sigma^+_{\overline{H}}}{4b_c}$, $0\in S(t)$ for all $t\ge t_0$, and for some $\varepsilon,\bar{\kappa}>0$,*

$$\kappa(t_0,t)\le -\varepsilon(t-t_0)+2\bar{\kappa}, \tag{44}$$

then system (40) *is asymptotically stable with $H^{-1}(S(t_0))\cap\mathcal{R}_{e^{-\bar{\kappa}}\rho}$ contained in the basin of attraction, where*

$$\mathcal{R}_{e^{-\bar{\kappa}}\rho} := \{x\in\mathbb{R}^n \,|\, x^\top Px \le e^{-2\bar{\kappa}}\rho^2\}, \quad \textit{and} \quad \rho := \frac{\beta\theta r}{b_c\,\|RAR^{-1}\|}. \tag{45}$$

Proof. We proceed with the proof in several steps.
Step 1: Using (11), we have $RG=R^{-1}H^\top =: \overline{H}^\top$, where R is such that $R^2=P$. Letting $z:=Rx$, system (40) in new coordinates is written as:

$$dz(t) = RAR^{-1}z(t) - \overline{w}_t \tag{46a}$$
$$\overline{w}_t \in \mathcal{N}(S'(t); z(t)), \tag{46b}$$

where $S'(t)=\overline{H}^{-1}(S(t))$ is r'-prox-regular with $r':=\frac{r}{b_c}$ as a result of Lemma 3. The solution of this system is taken in terms of densities with respect to the measure $\nu=\mu'+\lambda$, where μ' is the measure associated with the variation of $S'(\cdot)$. Under the hypothesis that $\sup_{s\ge t_0}\mu(\{s\})<\frac{r\sigma^+_{\overline{H}}}{4b_c}$, Lemma 4b guarantees that $\sup_{s\ge t_0}\mu'(\{s\})\le\frac{r}{4b_c}=\frac{r'}{4}$, and from Theorem 4, we obtain

$$|\overline{w}_t| \le |RAR^{-1}z(t)|\frac{d\lambda}{d\nu}(t) + \frac{d\mu'}{d\nu}(t). \tag{47}$$

Step 2: Consider the Lyapunov function $V:\mathbb{R}^n\to\mathbb{R}_+$ defined as $V(z)=z^\top z$, then $V\circ z$ is locally *rcbv* and it has at most countably many jump discontinuities at time instants t_i, where i belongs to the countable

set $\mathcal{I}$ that indexes the discontinuities in $v_S(\cdot)$. The density of the differential measure of $V \circ z$ relative to ν for each $t \in (t_i, t_{i+1})$ is computed as [36, Section 11]:

$$\begin{aligned}
\frac{dV}{d\nu}(z(t)) &= (z(t^+) + z(t^-))^\top \frac{dz}{d\nu}(t) \\
&= z(t)(R^{-1}A^\top R + RAR^{-1})z(t)\frac{d\lambda}{d\nu}(t) - 2z(t)^\top \overline{w}_t \\
&\leq -\theta z(t)^\top z(t)\frac{d\lambda}{d\nu}(t) + \frac{1}{r'}|\overline{w}_t| \cdot |z(t)|^2 \qquad (48a)\\
&\leq -\theta V(z(t))\frac{d\lambda}{d\nu}(t) + \frac{b_c}{r}\left(\|RAR^{-1}\| \cdot |z(t)|\frac{d\lambda}{d\nu}(t) + \frac{1}{\sigma_{\overline{H}}^+}\frac{d\mu_c}{d\nu}(t)\right)V(z(t)), \qquad (48b)
\end{aligned}$$

where we arrived at (48a) using (4) and (10) under the hypothesis that $0 \in S(t)$ for all $t \geq t_0$. Equation (48b) was obtained using (47), the value of r', and the bound on $\frac{d\mu_c'}{d\nu}$ given in Lemma 4b.

Step 3: If $R^{-1}z(t_0) \in \mathcal{R}_{e^{-\kappa}\rho}$, then $R^{-1}z(t) \in \mathcal{R}_\rho$, for all $t \geq t_0$. Assume this is not the case, then there exists $0 < \delta < 1$, and a time $\bar{t} > t_0$ such that $V(z(\bar{t})) \geq \rho^2 + \delta^2 \frac{r^2\varepsilon^2}{4b_c^2\|RAR^{-1}\|^2}$. Let $\bar{t}$ be the minimal such time for a fixed δ. Then, due to minimality of $\bar{t}$, it holds that for every $t \in [t_0, \bar{t})$, $V(z(t)) \leq \rho^2 + \frac{r^2\varepsilon^2}{4\,b_c^2\|RAR^{-1}\|^2}$, and hence $|z(t)| \leq \rho + \frac{r\varepsilon}{2\,b_c\|RAR^{-1}\|}$, which in turn implies that

$$\frac{dV}{d\nu}(z(t)) \leq -\bar{a}V(z(t))\frac{d\lambda}{d\nu}(t) + \frac{\bar{b}_c}{r}V(z(t))\frac{d\mu_c}{d\nu}(t), \qquad (49)$$

for ν-almost all $t \in [t_0, \bar{t})$, in which $\bar{a} := \left(\theta - \beta\theta - \frac{\varepsilon}{2}\right)$, and $\bar{b}_c := \frac{b_c}{\sigma_{\overline{H}}^+}$. Now, for each $i \in \mathcal{I}$ satisfying $t_i < \bar{t}$, define for $t \in (t_{i-1}, t_i)$,

$$W(t) := e^{\bar{a}(t-t_{i-1}) - \frac{\bar{b}_c}{r}\mu_c([t_{i-1},s])}V(z(t)).$$

The use of product rule and chain rule for differential of functions of bounded variation (given in [33, Page 8] and [37, Theorem 3], respectively) gives:

$$\frac{dW}{d\nu}(t) = e^{\bar{a}(t-t_{i-1}) - \frac{\bar{b}_c}{r}\mu_c([t_{i-1},t])}\frac{dV}{d\nu}(z(t)) + V(z(t))e^{\bar{a}(t-t_{i-1}) - \frac{\bar{b}_c}{r}\mu_c([t_{i-1},t])}\left(\bar{a}\frac{d\lambda}{d\nu}(s) - \frac{\bar{b}_c}{r}\frac{d\mu_c}{d\nu}(s)\right) \leq 0.$$

Thus, $W(\cdot)$ is non-increasing on the interval (t_{i-1}, t_i) for each $i \in \mathcal{I}$, and since $W(t)$ is right-continuous, $W(t) \leq W(t_{i-1})$, which implies that

$$V(z(t)) \leq e^{-\bar{a}(t-t_i) + \frac{\bar{b}_c}{r}\mu_c([t_i,s])}V(z(t_i)), \quad s \in [t_{i-1}, t_i). \qquad (50)$$

Also, for each t_i, $i \in \mathcal{I}$, it follows from Lemma 4b that $\alpha_i' := \mu'(\{t_i\}) = d_H(S'(t_i^+), S'(t_i^-)) < \frac{\mu(\{t_i\})}{\sigma_{\overline{H}}^+} < \frac{r'}{4}$, which combined with Lemma 5 gives

$$V(z(t_i)) \leq \gamma_i V(z(t_i^-)), \qquad (51)$$

where $\gamma_i = \left(1 - \frac{\alpha_i'}{r'}\right)^{-2}$. Combining (50) and (51), we get:

$$V(z(t)) \leq \prod_{i \in \overline{\mathcal{I}}} \gamma_i \cdot e^{-\bar{a}(t-t_0) + \frac{\bar{b}_c}{r}\mu_c([t_0,t])}V(z(t_0)), \qquad t \in [t_0, \bar{t}).$$

where $\overline{\mathcal{I}}$ indexes the discontinuities of $v_S(\cdot)$ over the interval $[t_0, t]$. From Lemma 6, $\prod_{i \in \overline{\mathcal{I}}} \gamma_i \leq e^{\frac{b_d}{r'}\mu_d'([t_0,t])}$.

Using $\mu_d'([t_0,t]) \leq \frac{1}{\sigma_{\bar{H}}^{+}}\mu_d([t_0,t])$ (due to Lemma 4b), and the value of r', it follows that for each $t \in [t_0,\bar{t})$:

$$V(z(t)) \leq e^{-\bar{a}(t-t_0)+\frac{\bar{b}_c}{r}\mu_c([t_0,t])+\frac{b_d\bar{b}_c}{r}\mu_d([t_0,t])}V(z(t_0)) \tag{52a}$$

$$\leq e^{a(\bar{t}-t)+\frac{\varepsilon}{2}(\bar{t}-t_0)+\kappa(t_0,\bar{t})-\frac{b_d\bar{b}_c}{r}\mu_d((t,\bar{t}])}V(z(t_0)) \tag{52b}$$

$$\leq e^{a(\bar{t}-t)-\frac{b_d\bar{b}_c}{r}\mu_d((t,\bar{t}])+2\bar{\kappa}-\frac{\varepsilon}{2}(t-t_0)}V(z(t_0)) \tag{52c}$$

$$\leq e^{a(\bar{t}-t)-\frac{b_d\bar{b}_c}{r}\mu_d((t,\bar{t}])-\frac{\varepsilon}{2}(t-t_0)}\rho^2 \tag{52d}$$

where we used (43) and (44) in deriving (52b) and (52c), and for (52d) we used the fact that $V(z(t_0)) < e^{-2\bar{\kappa}}\rho^2$. In (52d), taking the limit as t approaches $\bar{t}$ from left, it is seen that $\mu_d((t,\bar{t}])$ converges to $\mu_d(\{\bar{t}\})$, and since

$$V(z(\bar{t})) \leq e^{\frac{b_d\bar{b}_c}{r}\mu_d(\{\bar{t}\})}V(z(\bar{t}^-))$$

it follows that that $V(z(\bar{t})) < \rho^2$ which is a contradiction to our initial hypothesis introduced in Step 3.

Step 4: Having shown that $|z(t)| < \rho$, $t \geq t_0$, for every trajectory $z(\cdot)$ starting from the initial condition $R^{-1}z(t_0) \in \mathcal{R}_{e^{-\bar{\kappa}}\rho}$, one can follow the same procedure as in Step 3 to show that (48b) leads to

$$\begin{aligned} V(z(t)) &\leq e^{-a(t-t_0)+\frac{b_c}{r\sigma_{\bar{H}}^{+}}\mu_c([t_0,t])+\frac{b_d b_c}{r\sigma_{\bar{H}}^{+}}\mu_d([t_0,t])}V(z(t_0)) \\ &= e^{\kappa(t_0,t)}V(z(t_0)) \leq e^{-\varepsilon(t-t_0)+2\bar{\kappa}}V(z(t_0)) \end{aligned} \tag{53}$$

for all $t \geq t_0$. From the above inequality, it follows that system (46) is asymptotically stable, and hence (40) is also asymptotically stable since stability is preserved under coordinate transformation. The set $\mathcal{R}_{e^{-\bar{\kappa}}\rho} \cap H^{-1}(S(t_0))$ is contained in the basin of attraction because every initial condition within the set $H^{-1}(S(t_0))$ is guaranteed to generate a solution, and $x(t_0) \in \mathcal{R}_{e^{-\bar{\kappa}}\rho}$ guarantees that $|R^{-1}z(t_0)| < e^{-\bar{\kappa}}\rho$. □

Remark 8. In the proof of Theorem 5, we used the differential measure of the Lyapunov function $V(\cdot)$ to analyze the stability in-between the discontinuities even though $V(\cdot)$ is continuous and its derivative exists almost everywhere on such intervals. That was done to handle the singular part of the measure μ_c. To understand this, consider the function $f_i : [0,1] \to [0,1]$, $i = 1,2$, such that $f_1(x) = -\alpha x$, where $\alpha \in (0,1)$ and $f_2(\cdot)$ is the Cantor function on the interval $[0,1]$. Let $f = f_1 + f_2$, then $f(\cdot)$ is a continuous function of bounded variation, and $\dot{f} = -\alpha < 0$ almost everywhere with respect to Lebesgue measure, but $f(\cdot)$ is monotonically increasing. However, $df([0,1]) = (1-\alpha) > 0$, that is, the differential measure of the function shows that the function is increasing on the interval $[0,1]$.

Remark 9 (Absolutely Continuous v_S)**.** The condition (12) in Theorem 2 is slightly stronger than the condition of Theorem (5). The difference is essentially due to the fact that for absolutely continuous $v_S(\cdot)$, whose derivative is locally essentially bounded by v, it holds that, $\mu_d([t_0,t]) = 0$ and $\mu_c([t_0,t]) = v_S(t) - v_S(t_0) \leq v\,(t-t_0)$, for each $t \geq 0$.

As stated earlier, for the convex sets, we let $r \to \infty$ and the stability results given in Theorem 5 are in coherence with the existing results in the literature. By letting r go to infinity, we see that ρ can be chosen to be arbitrarily large. The positive terms in the expression for $\kappa(t)$ in (43) vanish to zero and hence the system is asymptotically stable starting from every initial condition in $H^{-1}(S(t_0))$ as long as (10) and (11) are satisfied.

4.2.1 Special Cases

Requiring the value of κ in (43) to be negative or decrease in time represents a tradeoff between the *stable* and *unstable* elements of the system. The instability in the system is due to non-dissipative jumps that are represented by μ_d, and the variation of the set $S(\cdot)$ denoted by μ_c, as both these terms increase the value of κ. It is also noted that increase in the value of parameter r not only enlarges the basin of attraction but

also shrinks down the effect of μ_d and μ_c. Several special cases of Theorem 5 can be derived depending on how much the hypotheses are strengthened.

We consider two special cases where the variation in $S(\cdot)$ is only due to jump discontinuities so that $\mu_c = 0$. In essence, system (40) then behaves like an impulsive system with state constraints and in this case we arrive at a result similar to stability of impulsive systems reported in [23, Theorem 1].

Corollary 2 (Jumps in Stationary Sets)**.** *Consider system* (40) *under the hypotheses* **(H1)**, **(H3)** *and Assumption 1 with* $\mu_c([t_0,\infty)) = 0$*. Let the matrices* $P, R, \overline{H}$ *and the constants* a, b_c, b_d *be defined as in the statement of Theorem 5. It follows that for* $|x(t_0)|$ *sufficiently small, and* $V(x) = x^\top Px$*, we have*

$$\frac{dV}{d\lambda}(x(t)) \leq -aV(x(t)), \qquad t \in (t_i, t_{i+1}), \tag{54}$$

and at jump instants t_i*,*

$$V(x(t_i)) \leq \gamma_i V(x(t_i^-)), \tag{55}$$

where γ_i *satisfy* $\prod_{i\in\overline{\mathcal{I}}} \gamma_i \leq e^{\frac{b_d b_c}{r\sigma^+_{\overline{H}}}\mu_d([t_0,t])}$*, and* $\overline{\mathcal{I}}$ *indexes the discontinuities of* v_S *in* $[t_0, t]$*. If, for all* $t \geq t_0$*,* $0 \in S(t)$*,* $\mu(\{t_i\}) < \frac{r\sigma^+_{\overline{H}}}{4b_c}$*, and the following holds:*

$$\kappa_d(t_0, t) := \frac{b_d b_c}{r\sigma^+_{\overline{H}}}\mu_d([t_0, t]) - a(t - t_0) \leq 2\bar{\kappa}_d - \varepsilon(t - t_0) \tag{56}$$

for some $\bar{\kappa}_d, \varepsilon > 0$*, then* $V(x(t)) \leq e^{2\bar{\kappa}_d - \varepsilon(t-t_0)}V(x(t_0))$*, and hence* $\lim_{t\to\infty} V(x(t)) = 0$*.*

In the above result, the assumptions on system data and the initial condition allow us to arrive at (54) and (55). In general, if a system with *locally rcbv* state trajectories admits a function $V(\cdot)$ which is continuously differentiable with respect to the state variable, and satisfies (54), (55) then condition (56) gives a maximum bound on the transient response of the state trajectories, and guarantees asymptotic convergence of $V(\cdot)$ to the origin. Note that the above result includes the case where the state trajectories may exhibit Zeno phenomenon and thus it can be used independently for systems without state constraints provided the flow and jumps in Lyapunov function satisfy the required assumptions stated in (54) and (55) with some finite bound on the product of γ_i. The result is also in contrast to the existing conditions for stability of an accumulation point (or, Zeno equilibrium) reported in [20, 29], since our work deals with the stability of the origin where the system trajectories may exhibit infinite jumps in finite time away from the origin.

The next special case of Theorem 5 is obtained by excluding the accumulation of discontinuities and assuming that there is an upper bound on average number of discontinuities over each interval. The notion was originally introduced in [24] and is widely used in stability of switched systems where the discrete part tends to bring instability to the system.

Corollary 3 (Average dwell-time condition)**.** *Consider system* (40) *under the hypotheses* **(H1)**, **(H3)** *and Assumption 1. Let the matrices* $P, R, \overline{H}$ *and the constants* a, b_c, b_d *be defined as in the statement of Theorem 5. Assume that the discontinuities of* $v_S(\cdot)$ *satisfy the following inequality:*

$$N_\mu(t, t_0) \leq N_0 + \frac{t - t_0}{\tau_a}, \tag{57}$$

where $N_\mu(t, t_0)$ *denotes the number of jumps in* $v_S(\cdot)$ *over the open interval* (t_0, t) *and* $N_0 > 0$ *is a constant. If* $0 \in S(t)$ *and* $\mu_c([t_0, t]) = 0$ *for all* $t \geq t_0$*;* $\alpha := \sup_{s\geq t_0} \mu(\{s\}) < \frac{r\sigma^+_{\overline{H}}}{4b_c}$*; and* τ_a *is such that*

$$\tau_a > \frac{\log\gamma}{a - \varepsilon}, \tag{58}$$

for some $\varepsilon > 0$*, and* $\gamma := \left(1 - \frac{\alpha b_c}{r\sigma^+_{\overline{H}}}\right)^{-2}$*, then system* (40) *is asymptotically stable with* $\mathcal{R}_{e^{-\bar{\kappa}}\rho} \cap H^{-1}(S(t_0))$ *contained in the basin of attraction, and* $\mathcal{R}_{e^{-\bar{\kappa}}\rho}$ *defined in* (45) *using* $\bar{\kappa} = \frac{N_0 \log\gamma}{2}$*.*

Proof. Consider γ_i as introduced in (51). Since $\mu(\{t_i\}) < \alpha$, for each t_i where $v_S(\cdot)$ is discontinuous, we get $1 < \gamma_i = \left(1 - \frac{\mu'(\{t_i\})}{r'}\right)^{-2} \leq \left(1 - \frac{\mu(\{t_i\})b_c}{r\sigma_{\overline{H}}^{+}}\right)^{-2} < \gamma$, and hence

$$\prod_{i \in \overline{\mathcal{I}}} \gamma_i < \gamma^{N_\mu(t,t_0)} = e^{N_\mu(t,t_0)\log\gamma},$$

where $\overline{\mathcal{I}}$ indexes the discontinuities of v_S in $[t_0, t]$. Thus, equation (53) gets modified as follows:

$$V(z(t)) \leq e^{N_\mu(t,t_0)\log\gamma - a(t-t_0) + \frac{b_c}{r\sigma_{\overline{H}}^{+}}\mu_c([t_0,t])} V(z(t_0)).$$

Since the variation is constant between two discontinuities, i.e., $\mu_c([t_0, t]) = 0$, the right-hand side simplifies as:

$$V(z(t)) \leq e^{N_\mu(t,t_0)\log\gamma - a(t-t_0)} V(z(t_0)).$$

The bound on $N_\mu(t, t_0)$ given in (57) then results in

$$V(z(t)) \leq e^{N_0 \log\gamma} e^{\left(\frac{\log\gamma}{\tau_a} - a\right)(t-t_0)} V(z(t_0)).$$

It follows from (58) that $\frac{\log\gamma}{\tau_a} - a < -\varepsilon$, hence $V(z(t)) \leq e^{N_0\log\gamma - \varepsilon(t-t_0)} V(z(t_0))$ and $V(z(t)) \to 0$ as $t \to \infty$. □

4.3 Observer Design

We now consider the problem of state estimation for system (40) using the output equation (18). The observer construction intrinsically follows the same design procedure as discussed in the absolutely continuous case. The care must be taken however due to the destabilizing effect of state jumps. Recent works dealing with the observers for discontinuous systems [42, 43] assume that the jump maps are globally Lipschitz, so that the state estimation error is scaled by the corresponding Lipschitz constant and its effect can be minimized by choosing fast enough convergence rate between the jumps. In our setup, the jumps are introduced by taking the nearest point to $x(t^-)$ on the set $H^{-1}(S(t))$. Since these sets are prox-regular, and not necessarily convex, this map is only locally Lipschitz and not even uniquely defined for large jumps in the sets. This intrinsic difficulty in the problem setup allows us to look at the local results for the state estimators when the sets evolve with bounded variation. For that, we propose the following estimator:

$$d\hat{x} = (A - LC)\hat{x}d\lambda + Ly(t)d\lambda - G\hat{w}_t \tag{59a}$$

$$\hat{w}_t \in \mathcal{N}(S(t); Hx(t)), \tag{59b}$$

with initial condition satisfying $H\hat{x}(t_0) \in S(t_0)$. Consider the state estimation error $\tilde{x} := x - \hat{x}$, which is locally *rcbv*. The differential measure of $\tilde{x}(\cdot)$ satisfies the following inclusion:

$$d\tilde{x} \in (A - LC)\tilde{x}\, d\lambda - G(w_t - \hat{w}_t), \tag{60}$$

where $w_t, \hat{w}_t$ are given by (40b) and (59b), respectively. The result concerning the convergence of this state estimator is stated next.

Proposition 3. *Consider system* (40) *under the hypotheses* **(H1)**, **(H2)** *and suppose that Assumptions 1 and 2 hold. Assume that there exist some constants* $\varrho, \theta, c_1, c_2 > 0$, *and a matrix* P *that satisfy the following inequalities:*

$$c_1 I \leq P \leq c_2 I \tag{61a}$$

$$A^\top P + PA - 2\varrho C^\top C \leq -(\theta + c_3)P \tag{61b}$$

$$PG = H^\top \tag{61c}$$

where $c_3 > \frac{c_2}{c_1}\frac{2\|H\|^2\|A\|M}{r\sigma_{\overline{H}}^{+}}$. *Letting* $R := P^{1/2}$, $\overline{H} = HR^{-1}$, *if for some* $\varepsilon, \bar{\kappa}_e > 0$, $\beta \in (0,1)$,

$$\kappa_e(t_0,t) := -a(t-t_0) + \frac{2b_c}{r\sigma_{\overline{H}}^{+}}\mu_c([t_0,t]) + \frac{b_d b_c}{r\sigma_{\overline{H}}^{+}}\mu_d([t_0,t]) \le 2\bar{\kappa}_e - \varepsilon(t-t_0)$$

with $a := (1-\beta)\theta$, $b_c := \frac{\|\overline{H}\|^2}{\sigma_{\overline{H}}^{+}}$, $b_d := \frac{32}{9}$, *then choosing* $L = \varrho P^{-1}C^\top$ *renders the error dynamics* (60) *asymptotically stable, and for every* $\tilde{x}(t_0) \in \mathcal{R}_{e^{-\bar{\kappa}_e}\widetilde{\rho}}$, *defined as*

$$\mathcal{R}_{e^{-\bar{\kappa}_e}\widetilde{\rho}} := \left\{\widetilde{x} \in \mathbb{R}^n \,|\, \widetilde{x}^\top P \widetilde{x} \le e^{-2\bar{\kappa}_e}\widetilde{\rho}^2\right\}, \quad \textit{and} \quad \widetilde{\rho} := \frac{\beta\,\theta\,r}{b_c\,\|R(A-LC)R^{-1}\|}, \tag{62}$$

we have $\lim_{t\to\infty}|\tilde{x}(t)| = 0$.

Proof. Picking $R = P^{1/2}$ and introducing the coordinate transformation $z = Rx$, and $\hat{z} = R\hat{x}$ shows that the system and observer dynamics are well-posed and yield a unique solution in the sense of Definition 3 using Theorem 4. Choose the candidate Lyapunov function to be $V(\tilde{z}) = \widetilde{z}^\top\widetilde{z}$, where $\widetilde{z} = R\tilde{x}$, then using the computations similar to (48) and (24), the differential measure of $V \circ \widetilde{z}$ between two discontinuities over the interval (t_i, t_{i+1}) is given by:

$$\begin{aligned}
\frac{dV}{d\nu}(\widetilde{z}(t)) &= -(\theta + c_3)\widetilde{z}(t)^\top\widetilde{z}(t)\frac{d\lambda}{d\nu} + 2(z(t)-\hat{z}(t))^\top(\overline{w}_t - \hat{\overline{w}}_t) \\
&\le -(\theta+c_3)|\widetilde{z}(t)|^2\frac{d\lambda}{d\nu}(t) + 2\frac{c_2}{c_1}\frac{\|H\|^2}{r\sigma_{\overline{H}}^{+}}\|A\|M|\widetilde{z}(t)|^2\frac{d\lambda}{d\nu}(t) \\
&\qquad + \frac{\|\overline{H}\|^2}{r\sigma_{\overline{H}}^{+}}\left(\|R(A-LC)R^{-1}\|\cdot|\widetilde{z}(t)|\frac{d\lambda}{d\nu}(t) + \frac{2}{\sigma_{\overline{H}}^{+}}\frac{d\mu_c}{d\nu}(t)\right)|\widetilde{z}(t)|^2 \\
&\le -\theta|\widetilde{z}(t)|^2\frac{d\lambda}{d\nu}(t) + \frac{b_c}{r}\left(\|R(A-LC)R^{-1}\|\cdot|\widetilde{z}(t)|)\frac{d\lambda}{d\nu}(t) + \frac{2}{\sigma_{\overline{H}}^{+}}\frac{d\mu_c}{d\nu}(t)\right)|\widetilde{z}(t)|^2.
\end{aligned}$$

From here onwards, one can follow the same arguments as in the Step 3 and Step 4 of Theorem 5. That is, under the condition that $\tilde{x}(t_0) \in \mathcal{R}_{e^{-\bar{\kappa}_e}\widetilde{\rho}}$, we have $|z(t_0)| < e^{-\bar{\kappa}_e}\widetilde{\rho}$, which leads to the following bound:

$$V(\widetilde{z}(t)) \le e^{-a(t-t_0)+\frac{2b_c}{r\sigma_{\overline{H}}^{+}}\mu_c([t_0,t])+\frac{b_d b_c}{r\sigma_{\overline{H}}^{+}}\mu_d([t_0,t])}V(\widetilde{z}(t_0)) =\le e^{-\varepsilon(t-t_0)+2\bar{\kappa}_e}V(\widetilde{z}(t_0)) \tag{63}$$

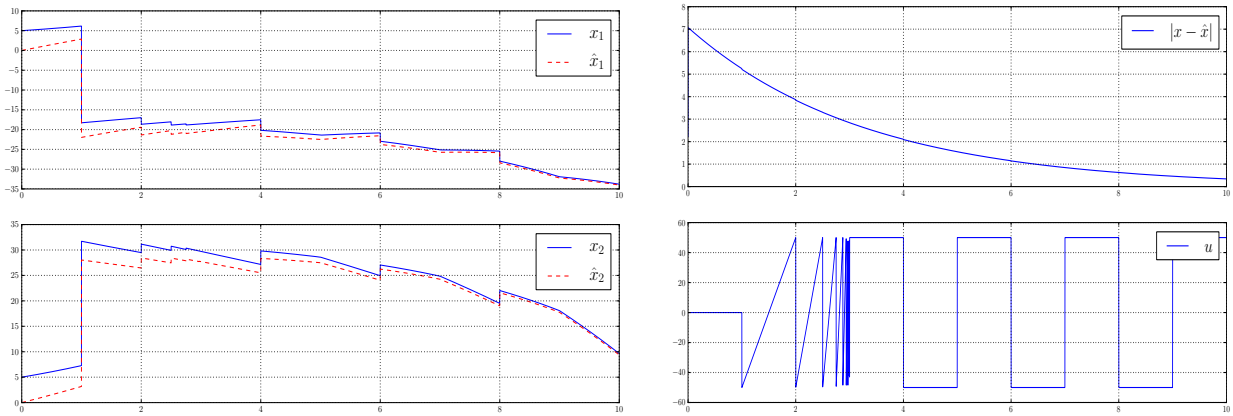
from where the asymptotic stability of the error dynamics (60) follows, and in particular $\tilde{x}(\cdot)$ converges to zero whenever $\tilde{x}(t_0) \in \mathcal{R}_{e^{-\bar{\kappa}_e}\widetilde{\rho}}$. □

Example 4 (A linear complementary system with inputs of locally bounded variation)**.** We apply our observer design to linear complementarity systems defined in (17) in which we take,

$$A = \begin{bmatrix} 0.1 & 0.1 \\ 0.3 & 0.1 \end{bmatrix}, \quad H = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad D = 1$$

along with the output equation $y = Cx = [1\ 1]x$. The input $u(\cdot)$ is chosen to be a locally *rcbv* function, defined as,

$$u(t) = \begin{cases} 0 & t \in [0,1), \\ c_u\,2^{k+1}\,t - c_u(1+2^{k+1}(3-2^{1-k})) & t \in [3-2^{1-k}, 3-2^{-k}), \quad 0 \le k \le N, \\ 0 & t \in [3-2^{-N}, 3) \\ c_u & t \in [3+2k, 4+2k), \quad k \ge 0 \\ -c_u & t \in [4+2k, 5+2k), \quad k \ge 0 \end{cases} \tag{64}$$



(a) The two plots show the evolution of the state $x(\cdot)$ and its estimate $\hat{x}(\cdot)$ for each of their components, where the estimate converges to the actual value.

(b) The top plot shows the norm of the state estimation error converging to zero. The bottom plot shows the input $u(\cdot)$ in (64) with parameters $c_u = 50$ and $N = 100$.

Figure 4: Simulation results for state estimation in linear complementarity systems with locally *rcbv* inputs.

where $c_u > 0$ is some constant and N can be an arbitrarily large positive integer. Using the same arguments as in Section 3.1.2, such systems can be rewritten in the form (40) with $S(t) := \{x \in \mathbb{R}^n \mid Hx + Du(t) \geq 0\}$ and $G = H^\top$. It is verified that $v_S(\cdot)$ is indeed locally *rcbv* because $u(\cdot)$ in (17) is locally *rcbv* [10]. We can thus realize the observer (59), in which we pick $L = [0.3\ 0.3]^\top$, so that $A - LC = \left[\begin{smallmatrix} -0.2 & -0.2 \\ 0 & -0.2 \end{smallmatrix}\right]$ and $(A - LC) + (A - LC)^\top = \left[\begin{smallmatrix} -0.4 & -0.2 \\ -0.2 & -0.4 \end{smallmatrix}\right] \leq -0.2\, I_{2\times 2}$. Since $S(t)$ is a polyhedral, and hence convex set, for each $t \geq 0$, we can let $r \to \infty$ and the solution of the estimator (59) initialized with $\hat{x}_0$ satisfying $H\hat{x}_0 \in S(0)$ converges to the actual state of the system.

The simulations for this example were carried out on SICONOS platform [2] and the results are shown in Figure 4. The numerical integration schemes implemented in SICONOS are based on Moreau's time-stepping algorithm which were also used to prove the existence of solutions for system (40) in [17]. Because these algorithms are numerically convergent, we can numerically approximate the accumulation of jumps in our simulations. It is seen that the jumps in the input $u(\cdot)$ introduce the jumps in the state trajectories and despite the accumulation of jumps, the proposed estimator continues to converge to the actual state.

Remark 10. If there is no knowledge about the initial state of the system, so that the initial error does not satisfy the bound specified in Proposition 3, then the transposition of high-gain observer (proposed in Section 3.2.3 for the absolutely continuous case) for the systems with bounded variation is not straightforward. The primary reason is that we cannot replicate the jump map with the high-gain approach because, in that case, the state estimate does not follow the set constraints of the actual system and the projection map (which causes jumps in the state of the system) is well-defined only within the neighborhood of the sets $S(\cdot)$. Furthermore, if we estimate the state of system (40) using an absolutely continuous trajectory following the same approach, as in Section 3.2.3, we observe that the state estimation error would decrease over an interval without jumps, but due to the jump in the state $x(\cdot)$ at some time instant $\{t_i\}$, it can be shown that

$$\bar{V}(t_i^+) \leq \bar{V}(t_i^-) + 2k_2\sqrt{\bar{V}(t_i^-)}\mu(\{t_i\}) + k_1\mu(\{t_i\})^2. \tag{65}$$

The last term in the above inequality shows that the error increases by an amount corresponding to the size of jump in the sets, even though the error before t_i was small. To avoid this, one may be able to construct a high-gain observer to approximate the state up to the desired accuracy (given by $\mathcal{R}_e$) in finite time under the added assumption that initially $v_S(\cdot)$ is absolutely continuous, by repeating the procedure outlined in Section 3.2.3. Once the state estimation is made small enough using such techniques, one may project the estimate on the set S at that time and then run the observer (59) to obtain asymptotic convergence to the

actual state. This approach is just an extension of the results already developed in this paper, and hence we avoid the details.

5 Conclusions

In this paper, we addressed the problem of stability and observer design for a class of measure differential inclusions that model the evolution of state trajectories constrained by non-convex and time-varying sets. The framework of this paper generalizes some previous approach for systems constrained by convex sets. The stability results are local where the basin of attraction is parameterized by a scalar r which, in some sense, characterizes the lack of convexity such that when $r \to \infty$, we recover the global results of the convex case. We apply the stability analysis to design a locally convergent observer. In addition, a practically convergent observer is designed based on the high-gain approach, which in combination with the locally convergent observer yields asymptotic convergence of the state estimate to the actual state.

The work could be extended in several directions. In system described by (1), the multivalued part is due to the subdifferential of the *indicator* function associated with a set-valued mapping and one could investigate how far these ideas extend when treating the subdifferentials of general time-varying nonconvex set-valued mappings. Another possible extension could be to address the performance of the state estimators in closed-loop with state feedback. Also, as a possible future work, one could consider set-valued maps which are not functions of time but rather of the state. This would allow for discontinuities in the state trajectories which are not known to the estimator, and hence a different approach needs to be adopted. A classical example of such a system would be a bouncing ball, which is modeled as a second order sweeping process. Designing estimators for such kind of mechanical systems is an ongoing work.

Acknowledgements: The authors are grateful to the anonymous reviewers for their careful reading of the manuscript and their useful suggestions to improve the presentation of this paper. We also thank Olivier Huber for his help with the simulations in Example 4.

A Proofs of Intermediate Results

In this Appendix, we collect the proofs of Lemmas 3 – 4b, and the derivation of equation (39) in Theorem 4.

A.1 Proof of Lemma 3

Proof. To prove the desired result, we introduce the following definition: A vector w belongs to the Fréchet subdifferential of a lower semicontinuous function $f(\cdot)$, denoted $w \in \partial_F f(x)$, if

$$\liminf_{x' \to x} \frac{f(x') - f(x) - \langle w, x' - x \rangle}{|x' - x|} \geq 0.$$

Let $\mathbb{B}$ denote the closed unit ball in $\mathbb{R}^n$ with respect to the Euclidean norm and consider $\overline{w} \in \mathcal{N}(S'; \bar{z})$, then $\partial_F d_{S'}(\bar{z}) = \mathcal{N}(S'; \bar{z}) \cap \mathbb{B}$, so that $\frac{\overline{w}}{|\overline{w}|} =: \hat{w} \in \partial_F d_{S'}(\bar{z})$, where $d_{S'}$ denotes the distance function from the set S', $d_{S'}(z) = d(S', z)$, $\forall z \in \mathbb{R}^n$. Using the definition of Fréchet subdifferential, it follows that $\sigma_H^+ \hat{w} \in \partial_F (d_S \circ H)(\bar{z})$. Indeed, since $\hat{w}$ satisfies

$$\liminf_{z \to \bar{z}} \frac{d_{S'}(z) - d_{S'}(\bar{z}) - \langle \hat{w}, z - \bar{z} \rangle}{|z - \bar{z}|} \geq 0,$$

where $d_{S'}(\bar{z}) = 0 = (d_S \circ H)(\bar{z})$ for all $\bar{z} \in S'$ and $(d_S \circ H)(z) \geq \sigma_H^+ d_{S'}(z)$ due to (7), it follows that

$$\liminf_{z \to \bar{z}} \frac{(d_S \circ H)(z) - (d_S \circ H)(\bar{z}) - \langle \sigma_H^+ \hat{w}, z - \bar{z} \rangle}{|z - \bar{z}|} \geq \sigma_H^+ \liminf_{z \to \bar{z}} \frac{d_{S'}(z) - d_{S'}(\bar{z}) - \langle \hat{w}, z - \bar{z} \rangle}{|z - \bar{z}|} \geq 0.$$

From the basic subdifferential chain rule [40, Theorem 10.6], it follows under (5) that $\partial_F(d_S \circ H)(\bar{z}) = H^\top \partial_F d_S(H\bar{z})$. Hence there exists $w \in \mathcal{N}(S; H\bar{z})$, with $|w| \leq 1$, such that $\sigma_H^+ \hat{w} = H^\top w$. Using the prox-regularity of S, we combine the aforementioned arguments to arrive at the desired result using the following steps: for any $z' \in S'$,

$$\begin{aligned}\langle \overline{w}, \bar{z} - z' \rangle &= \frac{|\overline{w}|}{\sigma_H^+} \left\langle H^\top w, \bar{z} - z' \right\rangle = \frac{|\overline{w}|}{\sigma_H^+} \langle w, H(\bar{z} - z') \rangle \\ &\geq -\frac{|\overline{w}|}{2r\sigma_H^+} |H(\bar{z} - z')|^2.\end{aligned}$$

The above inequality holds for all $\overline{w} \in \mathcal{N}(S'; \bar{z})$, and each $\bar{z} \in S'$. From the definition of prox-regular sets, S' is prox-regular with $r' > 0$. □

A.2 Proof of Lemmas 4a and 4b

Proof. Consider the value of $S(\cdot)$ at two time instants $t_{i-1} \neq t_i$, and without loss of generality, assume that $S(t_{i-1}) \neq S(t_i)$. Under the given hypotheses, we obtain

$$\sup_{w \in S(t_{i-1})} d(w, S(t_i)) = \sup_{Hy \in S(t_{i-1})} d(Hy, S(t_i)) = \sup_{y \in S'(t_{i-1})} d(Hy, S(t_i)) \geq \sigma_H^+ \sup_{y \in S'(t_{i-1})} d(y, S'(t_i)),$$

where the first and second equalities followed from the fact that $S(t)$ is contained in the range space of H, for all t, and the last inequality is a direct implication of Lemma 2. From the definition of Hausdorff distance (2), it then follows that $d_H(S'(t_{i-1}), S'(t_i)) \leq \frac{1}{\sigma_H^+} d_H(S(t_{i-1}), S(t_i))$. Next, let $\mathcal{M}_\lambda$ denote the set of points where $v_S(\cdot)$ is not differentiable, so that $\mathcal{M}_\lambda$ has Lebesgue measure zero. For a fixed $t \in (t_0, \infty) \setminus \mathcal{M}_\lambda$, consider $h > 0$ to be small enough such that $t_0 \leq t - h$, and sequences $\{t_i\}_{i=1}^k \in [t-h, t+h]$, for $k \in \mathbb{N}$, then

$$\begin{aligned}v_{S'}(t+h) - v_{S'}(t-h) &= \sup_{\{t_i\}_{i=1}^k \in [t-h, t+h]} d_H(S'(t_i), S'(t_{i-1})) \\ &\leq \frac{1}{\sigma_H^+} \sup_{\{t_i\}_{i=1}^k \in [t-h, t+h]} d_H(S(t_i), S(t_{i-1})) \\ &= \frac{1}{\sigma_H^+} (v_S(t+h) - v_S(t-h)).\end{aligned}$$

Dividing both sides of the above expression by h, and letting h go to zero, we get $\dot{v}_{S'}(t) \leq \frac{1}{\sigma_H^+} \dot{v}_S(t)$ which proves Lemma 4a. To prove Lemma 4b, we let $\nu := \mu' + \lambda$. Using the definition of density of measures (38), the following holds for ν-almost every $t \in [t_0, \infty)$:

$$\begin{aligned}\frac{d\mu'}{d\nu}(t) &= \lim_{\varepsilon \to 0} \frac{\mu'([t-\varepsilon, t+\varepsilon])}{\nu([t-\varepsilon, t+\varepsilon])} = \lim_{\varepsilon \to 0} \frac{v_{S'}(t+\varepsilon) - v_{S'}(t-\varepsilon)}{\nu([t-\varepsilon, t+\varepsilon])} \\ &\leq \frac{1}{\sigma_H^+} \lim_{\varepsilon \to 0} \frac{v_S(t+\varepsilon) - v_S(t-\varepsilon)}{\nu([t-\varepsilon, t+\varepsilon])} = \frac{1}{\sigma_H^+} \lim_{\varepsilon \to 0} \frac{\mu([t-\varepsilon, t+\varepsilon])}{\nu([t-\varepsilon, t+\varepsilon])} = \frac{1}{\sigma_H^+} \frac{d\mu}{d\nu}(t),\end{aligned}$$

where the limit in the last equality is well-defined since μ is absolutely continuous with respect to ν. In the above inequality, if t is a time instant between two consecutive discontinuities of $v_S(\cdot)$, then we have $\frac{d\mu'_c}{d\nu}(t) \leq \frac{1}{\sigma_H^+} \frac{d\mu_c}{d\nu}(t)$ for ν-almost every $t \geq t_0$. If there is a jump discontinuity in $v_S(\cdot)$ at time t_i, then integration with respect to $d\nu$ yields $\mu'_d(\{t_i\}) \leq \frac{1}{\sigma_H^+} \mu_d(\{t_i\})$. □

A.3 Derivation of Equation (39)

The existence of solution is proved in [17, Theorem 3.1] with respect to a certain measure which is equivalent to $\nu := \mu + \lambda$, and based on their solution concept, for $s < t$ sufficiently close to t, we have

$$x(t) = \mathrm{prox}_{S(t)}(x(s) + (t-s)f(s, x(s)))$$

which gives

$$\begin{aligned} |x(t) - x(s) - (t-s)f(s,x(s))| &\leq d(x(s) + (t-s)f(s,x(s)), S(t)) \\ &\leq d(x(s), S(t)) + (t-s)|f(s,x(s))| \\ &\leq \mu([s,t]) + (t-s)|f(s,x(s))|. \end{aligned}$$

Using the Cauchy-Schwarz inequality, for $\xi \neq 0$ we get

$$\begin{aligned} \left\langle \frac{\xi(t)}{|\xi(t)|}, x(t) - x(s) - (t-s)f(s,x(s)) \right\rangle &\leq |x(t) - x(s) - (t-s)f(s,x(s))| \\ &\leq \mu([s,t]) + (t-s)|f(s,x(s))|. \end{aligned}$$

Dividing both sides by $\nu([s,t])$, we get

$$\left\langle \frac{\xi(t)}{|\xi(t)|}, \frac{x(t)-x(s)}{\nu([s,t])} - \frac{(t-s)}{\nu([s,t])} f(s,x(s)) \right\rangle \leq \frac{\mu([s,t])}{\nu([s,t])} + \frac{(t-s)}{\nu([s,t])}|f(s,x(s))|$$

Since $\nu([s,t]) \geq \lambda([s,t]) = t - s$, taking the limit $s \to t$ gives

$$\left\langle \frac{\xi(t)}{|\xi(t)|}, \frac{dx}{d\nu} - f(t,x(t))\frac{d\lambda}{d\nu}(t) \right\rangle \leq \frac{d\mu}{d\nu}(t) + |f(t,x(t))|\frac{d\lambda}{d\nu}(t)$$

Letting

$$\xi(t) := \frac{dx}{d\nu}(t) - f(t,x(t))\frac{d\lambda}{d\nu}(t),$$

we get

$$|\xi(t)| = \left\langle \frac{\xi(t)}{|\xi(t)|}, \xi(t) \right\rangle \leq \frac{d\mu}{d\nu}(t) + |f(t,x(t))|\frac{d\lambda}{d\nu}(t).$$

If t_i is a point of discontinuity in the variation of $S(\cdot)$, then $d\lambda(t_i) = 0$, and

$$|\xi(t_i)| = \left| \frac{dx}{d\nu}(t_i) \right| \leq \frac{d\mu}{d\nu}(t_i)$$

and hence

$$|x(t_i^+) - x(t_i^-)| = \left| \int_{\{t_i\}} dx \right| \leq \int_{\{t_i\}} \left| \frac{dx}{d\nu} \right| d\nu \leq \int_{\{t_i\}} d\mu = \mu(\{t_i\}).$$

References

[1] V. Acary, O. Bonnefon, and B. Brogliato. *Nonsmooth Modeling and Simulation for Switched Circuits*, volume 69 of *Lecture Notes in Electrical Engineering*. Springer, Heidelberg, 2011.

[2] V. Acary and F. Pérignon. An introduction to SICONOS. Technical Report RT-0340, INRIA, Rhône-Alpes, 2007. Project website: `http://siconos.gforge.inria.fr`.

[3] M. Arcak and P. Kokotović. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37:1923 – 1930, 2001.

[4] E. G. Baltazar and J. A. Moreno. Dissipative design of adaptive observers for systems with multivalued nonlinearities. In *Proc. 49th IEEE Conf. Decision & Control*, pages 2625 – 2630, Atlanta (GA), USA, December 2010.

[5] F. Bernard, L. Thibault, and N. Zlateva. Prox-regular sets and epigraphs in uniformly convex Banach spaces: Various regularities and other properties. *Transactions of the American Mathematical Society*, 363(4):2211 – 2247, 2011.

[6] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas With Application to Linear Systems Theory.* Princeton University Press, New Jersey, 2005.

[7] B. Brogliato. Absolute stability and the Lagrange–Dirichlet theorem with monotone multivalued mappings. *Systems & Control Letters*, 51:343 – 353, 2004.

[8] B. Brogliato and D. Goeleven. Well-posedness, stability and invariance results for a class of multivalued Lur'e dynamical systems. *Nonlinear Analysis Series A: Theory, Methods & Applications*, 74:195 – 212, 2011.

[9] B. Brogliato and W.P.M.H. Heemels. Observer design for Lur'e systems with multivalued mappings: A passivity approach. *IEEE Trans. Automatic Control*, 54(8):1996 – 2001, 2009.

[10] B. Brogliato and L. Thibault. Existence and uniqueness of solutions for non-autonomous complementarity dynamical systems. *Journal of Convex Analysis*, 17(3 & 4):961 – 990, 2010.

[11] M. K. Camlibel, J. S. Pang, and J. Shen. Lyapunov stability of complementarity and extended systems. *SIAM J. Optim.*, 17(4):1056 – 1101, 2006.

[12] D. Chen, G. Yang, and Z. Han. Impulsive observer for input-to-state stability based synchronization of Lur'e differential inclusion system. *Commun. Nonlinear Sci. Numer. Simulat.*, 17:2990 – 2996, 2012.

[13] G. Colombo and L. Thibault. Prox-regular sets and applications. In D. Y. Gao and D. Motreanu, editors, *Handbook of Noncovex Analysis and Applications.* International Press, Massachusetts, 2010.

[14] M. J. Corless and G. Leitmann. Continuous state feedback guaranteeing uniform ultimate boundedness for uncertain dynamic systems. *IEEE Trans. Automatic Control*, 26(5):1139 – 1144, 1981.

[15] A. L. Dontchev, A. S. Lewis, and R. T. Rockafellar. The radius of metric regularity. *Transactions of the American Mathematical Society*, 355(2):493 – 517, 2002.

[16] J. F. Edmond and L. Thibault. Relaxation of an optimal control problem involving a perturbed sweeping process. *Math. Program., Ser. B*, 104:347 – 373, 2005.

[17] J. F. Edmond and L. Thibault. BV solutions of nonconvex sweeping process differential inclusion with perturbation. *J. Differential Equations*, 226:135 – 179, 2006.

[18] X. Fan and M. Arcak. Observer design for systems with multivariable monotone nonlinearities. *Systems & Control Letters*, 50:319 – 330, 2003.

[19] T. Floquet, C. Edwards, and S. K. Spurgeon. On sliding mode observers for systems with unknown inputs. *Int. J. Adaptive Control and Signal Processing*, 21:638–656, 2007.

[20] R. Goebel and A. Teel. Lyapunov characterization of Zeno behavior in hybrid systems. In *Proc. 47th IEEE Conf. Decision & Control*, pages 2752 – 2757, Cancun, Mexico, December 2008.

[21] D. Goeleven and B. Brogliato. Stability and instability matrices for linear evolution variational inequalities. *IEEE Trans. Automatic Control*, 49(4):521 – 534, 2004.

[22] W. P. M. H. Heemels, M. K. Camlibel, J. M. Schumacher, and B. Brogliato. Observer-based control of linear complementarity systems. *Int. J. Robust & Nonlinear Control*, 21(10):1193 – 1218, 2011.

[23] J. P. Hespanha, D. Liberzon, and A. Teel. Lyapunov conditions for input-to-state stability of impulsive systems. *Automatica*, 44:2735 – 2744, 2008.

[24] J. P. Hespanha and A. S. Morse. Stability of switched systems with average dwell-time. In *Proc. 38th IEEE Conf. Decision & Control*, pages 2655 – 2660, Phoenix (AZ), USA, December 1999.

[25] J. Huang and Z. Han. Adaptive non-fragile observer design for the uncertain Lur'e differential inclusion system. *Applied Mathematical Modelling*, 2012.

[26] J. Huang, Z. Han, X. Cai, and L. Liu. Adaptive full-order and reduced-order observers for the Lur'e differential inclusion system. *Commun. Nonlinear Sci. Numer. Simulat.*, 16:2869 – 2879, 2011.

[27] H. K. Khalil. High-gain observers in nonlinear feedback control. In H. Nijmeijer and T. I. Fossen, editors, *New Directions in Nonlinear Observer Design*, volume 244 of *Lecture Notes in Control and Information Sciences*, pages 249 – 268. Springer, London, 1999.

[28] M. Kunze and M. D. P. Monteiro Marques. An introduction to Moreau's sweeping process. In B. Brogliato, editor, *Impacts in Mechanical Systems: Analysis and Modelling*, volume 551 of *Lecture Notes in Physics*, pages 1 – 60. Springer-Verlag, Berlin, 2000.

[29] A. Lamperski and A. D. Ames. Lyapunov theory for Zeno stability. *IEEE Trans. Automatic Control*, 58(1):100 – 112, 2013.

[30] R. I. Leine and N. van de Wouw. *Stability and Convergence of Mechanical Systems with Unilateral Constraints*, volume 36 of *Lecture Notes in Applied and Computational Mechanics*. Springer-Verlag, 2008.

[31] N. Mahmoud and H. Khalil. Asymptotic regulation of minimum phase nonlinear systems using output feedback. *IEEE Trans. Automatic Control*, 41(10):1402 – 1412, 1996.

[32] B. Maury and J. Venel. A discrete contact model for crowd motion. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(1):145 – 168, 2011.

[33] M. D. P. Monteiro Marques. *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, volume 9 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser, 1993.

[34] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*, volume 330 of *Gundlehren der mathematischen Wissenchaften*. Springer-Verlag, Berlin, 2006.

[35] J. J. Moreau. Evolution problem associated with a moving convex set in Hilbert space. *J. Differential Equations*, 26:347 – 374, 1977.

[36] J. J. Moreau. Bounded variation in time. In J. J. Moreau, P. D. Panagiotopoulos, and G. Strang, editors, *Topics in Nonsmooth Mechanics*. Birkhäuser Verlag, 1988.

[37] J. J. Moreau and M. Valadier. A chain rule involving vector functions of bounded variation. *Journal of Functional Analysis*, 74(2):333 – 345, 1987.

[38] J. A. Osorio and J. A. Moreno. Dissipative design of observers for multivalued nonlinear systems. In *Proc. 45th IEEE Conf. Decision & Control*, pages 5400 – 5405, San Diego (CA), USA, December 2006.

[39] R. A. Poliquin, R. T. Rockafellar, and L. Thibault. Local differentiability of distance functions. *Transactions of the American Mathematical Society*, 352(11):5231 – 5249, 2000.

[40] R. T. Rockafellar and R. J-B Wets. *Variational Analysis*, volume 317 of *Gundlehren der mathematischen Wissenchaften*. Springer-Verlag, Berlin, 3rd printing, 2009 edition, 1998.

[41] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 3rd edition, 1987.

[42] H. Shim and A. Tanwani. Hybrid-type observer design based on a sufficient condition for observability in switched nonlinear systems. *Int. J. Robust & Nonlinear Control: Special Issue on High Gain Observers and Nonlinear Output Feedback Control*, 2012. DOI: 10.1002/rnc.2901, To appear.

[43] A. Tanwani, H. Shim, and D. Liberzon. Observability for switched linear systems: characterization and observer design. *IEEE Trans. Automatic Control*, 58(4):891 – 904, 2013.

[44] B. L. Walcott and S. H. Zak. State observation of nonlinear uncertain dynamical systems. *IEEE Trans. Automatic Control*, 32(2):166 – 170, 1987.

[45] J. C. Willems. Dissipative dynamical systems; Part I: General theory; Part II: Linear systems with quadratic supply rates. *Archive for Rational Mechanics and Analysis*, 45(5):321 – 393, 1972.