Observer Design for Unilaterally Constrained Lagrangian Systems: A Passivity-Based Approach

Aneel Tanwani, Bernard Brogliato, Christophe Prieur

Observer Design for Unilaterally Constrained Lagrangian Systems: A Passivity-Based Approach

Aneel Tanwani, *Member, IEEE*, Bernard Brogliato, and Christophe Prieur, *Member, IEEE*

***Abstract*—This paper addresses the problem of state estimation in nonlinear Lagrangian dynamical systems with frictionless unilateral constraints using the position measurement as output. The discontinuous velocity variable in such systems is modeled as a function of bounded variation (so that Zeno phenomenon is not ruled out). Since the derivative of such functions is represented with the Lebesgue-Stieltjes measure, the framework of measure differential inclusions (MDIs) is used to describe the dynamics. A class of estimators is proposed, which also uses the framework of MDIs, and is shown to generate asymptotically converging state estimates. The existence and uniqueness of solutions for the proposed estimators is rigorously proven. The global stability of error dynamics is analyzed using the generalized Lyapunov methods for functions of bounded variation. As particular cases of our estimators, we provide an explicit construction of a full-order observer, and a reduced-order observer.**

***Index Terms*—Complementarity systems, functions of bounded variation, convex analysis, systems with impacts, Lagrangian systems, Lyapunov methods, Moreau's sweeping process, observers, state estimation.**

I. Introduction

In this paper, we consider Lagrangian mechanical systems with unilateral constraints (without friction) on the position of a moving point. The position and velocity of this point is denoted by q and $\dot{q}$, respectively. Assuming the mass matrix $M(q)$ to be symmetric and positive definite, the unconstrained motion of the system satisfies the equation

$$M(q)\ddot{q} + F(t, q, \dot{q}) = 0 \tag{1a}$$

and the position q is constrained by

$$h_i(q) \geq 0, \quad i \in \{1, \dots, m\} \tag{1b}$$

where $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ denotes a vector field of generalized forces, and $h_i : \mathbb{R}^n \to \mathbb{R}$ represent the unilateral constraints imposed on the system's motion. Mechanical systems with impacts, such as robots and colliding rigid bodies could be seen as systems with unilateral constraints. In general, the trajectories of such systems are algebraically constrained and exhibit continuous as well as discrete dynamics; hence, forming an important class of nonsmooth systems. When none of the constraints $h_i(\cdot)$ are active, that is, $h_i(q) > 0$ for every $1 \leq i \leq m$, then q and $\dot{q}$ are obtained simply by integrating (1a) and are absolutely continuous. The discontinuity in the velocity $\dot{q}$ may appear in such systems when one of the constraints is active, that is, $h_i(q) = 0$, and the velocity points outside the admissible domain, that is, $\nabla h_i^\top(q)\dot{q} < 0$, where we use the notation $\nabla h_i(\cdot)$ to denote the gradient of the function $h_i(\cdot)$. This is because the velocity must change its direction instantaneously to keep the moving point inside the admissible set. In case $\nabla h_i^\top(q)\dot{q} \geq 0$, and $h_i(q) = 0$, there are no discontinuities and one only observes continuous motion on the constraint surface (of reduced dimension) defined by $h_i(q) = 0$. There are several modeling frameworks for such nonsmooth systems; one such modeling framework, which is used in general to model the motion of state-constrained trajectories is the so-called *sweeping process* [19], [22], [24], [25]. The term so-coined because it represents the motion of a point inside a closed set. As the set moves, the point is swept across by the moving set. If for such processes, the constraint set is parameterized by time *only*, then we call it a *time-dependent* sweeping process. However, for system (1), we first define an admissible set for velocity $\dot{q}(\cdot)$ which is parameterized by the position $q(\cdot)$, and this formulation leads to a *state-dependent* sweeping process.

This paper is concerned with the design of observers for estimating the velocity $\dot{q}(\cdot)$ using the position $q(\cdot)$ as the output, while using the sweeping process formulation to describe the dynamics of the system and the observer. The construction of observers, or state estimators, is a classical problem in the design of control systems and several estimation techniques have been established for smooth and unconstrained Lagrangian systems with application to output feedback control, see for example, [6], [7], [27], [42]. A common element of these designs is to assume that the velocity $\dot{q}(\cdot)$ is uniformly bounded (in time) which is primarily because $F(t, q, \cdot)$ is quadratic in general for mechanical systems.

Lately, however, the researchers have started looking at the state-estimation problems in nonsmooth systems. In this regard, we mention the recent work on observer design of switched systems with ordinary differential equations [34], [38], [39], switched differential-algebraic equations [40], [41], certain classes of differential inclusions [10], [29], [36], complementarity systems [17], and the references therein for more details. Classical approaches for observer design are based on constructing an auxiliary dynamical system driven by the error between the measured output and the estimated output, where it is shown that the resulting dynamics of the state estimation error converge to the origin. However, for nonsmooth systems subjected to impacts, such schemes are not easily implementable since the impacts, or discrete dynamics, are not influenced by error injection and hence destroy the integration effect.

For nonsmooth Lagrangian systems with impacts, the problem of state estimation has been considered in [21] under certain restrictive assumptions, and state estimation with tracking control in [16] for motions restricted within a convex polyhedral domain. The work of [5] also deals with the problem of tracking

A. Tanwani and B. Brogliato are with the Team BipOp (INRIA), 38334 Montbonnot Cedex, France (e-mail: atanwani@ieee.org; bernard.brogliato@inria.fr).

C. Prieur is with Gipsa-lab (CNRS), 38402 SaintMartin d'Hères, France.

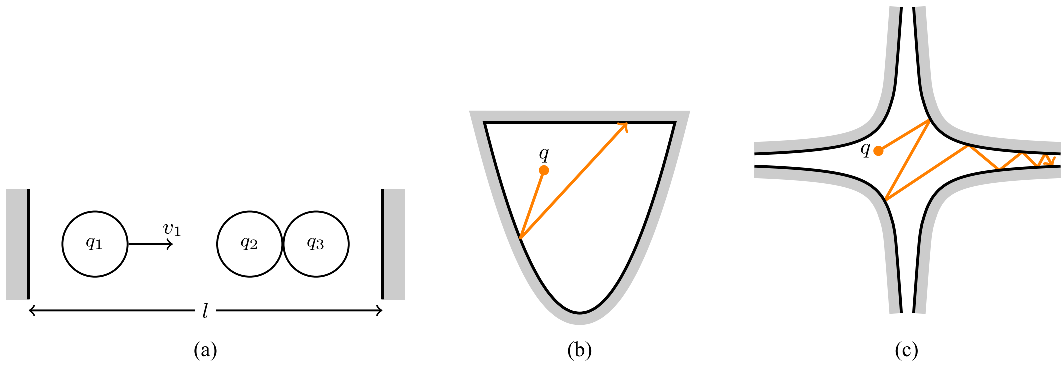


Fig. 1. Examples of systems with unilateral constraints. (a) 3-ball chain with walls. (b) Nonlinear billiard. (c) Non-convex billiard.

control (without estimation) for similar kind of systems. This article, however, deals with a more general class of nonsmooth Lagrangian dynamics, that allow more general admissible domains for position variable q using the formalism of differential inclusions. Some examples that can be treated within our setup are given in Fig. 1. The interesting aspects of these examples are

- In Fig. 1(a), when the ball q_1 collides with q_2 and q_3 stacked together at rest, then q_2 and q_3 may remain glued after the impact, and hence one of the constraints causes discontinuities in the velocities of q_2 and q_3, whereas another constraint only allows continuous motion on its boundary. The same happens when q_3 collides with the wall.
- In Fig. 1(b), the point mass is subjected to downward gravitational force only. After multiple impacts initially with the two boundaries of the constraints, one sees an accumulation of jumps, followed by a continuous motion on the constraint parabolic surface, which is a surface of reduced dimension than the state space.
- In Fig. 1(c), there are many impacts in short time-intervals, the domain is nonconvex, and there is possibly a chaotic behavior due to increased frequency of impacts.[1]

Our goal is to design estimators for the particles subjected to unilateral constraints of the form mentioned in Fig. 1, which in general may depict all the above complexities. From technical standpoint, systems with impacts, and state-dependent sweeping processes, in general, do not exhibit continuity of solutions with respect to initial conditions. The reason being, for state-dependent sweeping processes, the state trajectories starting from different initial conditions are not contained in the same set at all times because of which, the monotonicity argument cannot be invoked. When the system is governed by a time-dependent sweeping process, the trajectories of the estimator can be constrained within the same set as the plant, and then under appropriate passivity assumption on system data, the convergence of the estimate is obtained due to monotonicity of the normal cone operators of convex sets, as done in [10].

In this paper, however, the systems under consideration involve a state-dependent sweeping process because the constraint set for velocities is parameterized by the position variable, which is measured as the output of the system. However, by measuring position we can provide the system constraints to the estimator, and with an appropriate design, impose the monotone relation on the dynamics of the estimation error. We develop this intuitive idea to propose a class of observers which generates exponentially converging state estimates. Because the constraint set is only a lower semicontinuous function of the position variable, the proof for existence of solutions for observers is technically involved. The studied system class allows for state-trajectories of locally bounded variation (BV), which may contain countably many discontinuities in finite time (and hence the Zeno phenomenon is not ruled out).

The article is organized as follows: in Section II some useful mathematical definitions are recalled. The Moreau's sweeping process in which we embed Lagrangian nonsmooth mechanical systems, and the definition of its solutions are described in Section III. The proposed class of estimators is described in Section IV, and rigorous analysis is carried out in Sections V and VI, for existence of solutions, and convergence of estimation error, respectively. In Section VII, numerical algorithm for implementing the proposed estimators, and simulation results obtained with the INRIA software package SICONOS are presented for the system shown in Fig. 1(b), and for the rest, see [35]. Section VIII is dedicated to one of the main results of this article, i.e., the proof of the well-posedness (existence and uniqueness of solutions) of the observer. Conclusions end the paper in Section IX.

II. Preliminaries

In this section, we collect some basic definitions and notation that will be used later on.

Functions of Bounded Variation: 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 on* I, if it is of bounded variation on each compact subinterval of I. The variation of $f(\cdot)$ over an interval $[0, t]$ is denoted by $\mathrm{var}_f(t)$. 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 $f(\cdot)$ means that $f(t^+) = f(t)$.

Locally Integrable Functions: We denote by $\mathcal{L}_1(I, \mathbb{R}^n; d\mu)$ and $\mathcal{L}_1^{\mathrm{loc}}(I, \mathbb{R}^n; d\mu)$ the space of integrable and locally integrable functions, respectively, from I to $\mathbb{R}^n$ with respect to the

[1]This example was pointed to the authors by the associate editor responsible for handling this paper, L. Menini.

measure $d\mu$. 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. The space of continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ is denoted by $\mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$, for $m, n \in \mathbb{N}$.

Lebesgue-Stieltjes Measure Associated With BV Functions: If $v : I \to \mathbb{R}^n$ is a function of bounded variation, then one can associate with it a Lebesgue-Stieltjes measure or the so-called *differential measure* dv on I. Also, if $v(\cdot)$ is *rcbv* on $[a, b]$, then we have the relation that $v(t) = v(a) + \int_{]a,b]} dv$.

The density of the measure dv with respect to a positive Radon measure $d\mu$ over an interval I is defined as:

$$\frac{dv}{d\mu}(t) := \lim_{\varepsilon \to 0} \frac{dv\,(I(t,\varepsilon))}{d\mu\,(I(t,\varepsilon))} \tag{2}$$

where $I(t, \varepsilon) := I \cap [t - \varepsilon, t + \varepsilon]$. Similarly, one can define the density of the Lebesgue measure dt with respect to the Radon measure $d\mu$. A Radon measure $d\nu$ is absolutely continuous with respect to $d\mu$ if for every measurable set $\mathcal{A}$, $d\mu(\mathcal{A}) = 0$ implies that $d\nu(\mathcal{A}) = 0$. Further, the measure $d\nu$ is absolutely continuous with respect to $d\mu$ if and only if the density function $(d\nu/d\mu)(\cdot)$ is well-defined (finite μ-almost everywhere) and is $d\mu$ integrable.

Convex Analysis: For a set $V \subset \mathbb{R}^n$, we will denote its interior by int V, and the boundary of this set is denoted by $bd(V)$. If V is closed convex, then $\mathcal{N}_V(v)$ denotes the normal cone to V at $v \in V$ and is defined as

$$\mathcal{N}_V(v) := \{w \in \mathbb{R}^n \mid \langle w, x - v\rangle \leq 0 \; \forall\, x \in V\} \tag{3}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{R}^n$. We adopt the convention that $\mathcal{N}_V(v) = \emptyset$ if $v \notin V$. Obvious from the definition, the normal cone to a closed convex set is a monotone operator, that is, if $w_i \in \mathcal{N}_V(v_i)$, $i = 1, 2$, then

$$\langle w_1 - w_2, v_1 - v_2\rangle = \underbrace{\langle w_1, v_1 - v_2\rangle}_{(3)\Rightarrow \geq 0} - \underbrace{\langle w_2, v_1 - v_2\rangle}_{(3)\Rightarrow \leq 0} \geq 0.$$

When V is a closed convex cone, we denote by V° the closed convex polyhedral cone polar to V with respect to usual inner product on $\mathbb{R}^n$, which is defined as

$$V^\circ := \{w \in \mathbb{R}^n \mid w^\top v \leq 0,\; \forall\, v \in V\}. \tag{4}$$

III. Dynamic Model for Constrained Lagrangian Systems

In this section, we will describe the dynamics of nonsmooth Lagrangian systems using differential inclusions and briefly talk about their solutions. The observer will then be designed using this formalism.

A. Mathematical Description

We consider mechanical systems with a finite number of degrees of freedom that are subjected to the unilateral constraints described in (1b). The position variable $q \in \mathbb{R}^n$ is thus assumed to evolve in a set that admits the following form:

$$\Phi := \{q \in \mathbb{R}^n \mid h_i(q) \geq 0,\; i = 1, 2, \ldots, m\}. \tag{5}$$

The geometry of the set Φ is determined by the functions $h_i(\cdot)$, and the only condition we will impose on the functions $h_i(\cdot)$ is that they are continuously differentiable so that $\nabla h_i(\cdot)$ is continuous for each i. This allows us to model a large number of closed domains which may even be nonconvex.

The convex polyhedral tangent cone $V(q)$ to the region Φ at a point q is given by

$$V(q) := \left\{v \in \mathbb{R}^n \mid v^\top \nabla h_i(q) \geq 0,\; \forall\, i \in \mathcal{J}(q)\right\} \tag{6}$$

where the set $\mathcal{J}(q)$ denotes the set of active constraints at q, i.e.,

$$\mathcal{J}(q) := \{i \in \{1, \ldots, m\} \mid h_i(q) = 0\}\,.$$

One can think of the set $V(q(t))$ as the set of admissible velocities that keep the position variable $q(t)$ inside the set Φ. In what follows, the notion of normal cone to the set $V(q)$, denoted $\mathcal{N}_{V(q)}$, is instrumental. For $v \in V(q)$, we have

$$\mathcal{N}_{V(q)}(v) = \left\{ w \in \mathbb{R}^n \;\middle|\; w = -\sum_{i \in \mathcal{K}(v)} \lambda_i \nabla h_i(q),\, \lambda_i \geq 0 \right\} \tag{7a}$$

$$= \left\{ w \in \mathbb{R}^n \;\middle|\; w = -\sum_{i \in \mathcal{J}(q)} \lambda_i \nabla h_i(q),\; 0 \leq \lambda_i \perp v^\top \nabla h_i(q) \geq 0 \right\} \tag{7b}$$

where $\mathcal{K}(v) := \{i \in \mathcal{J}(q) | v^\top \nabla h_i(q) = 0\}$. It is seen that (7b), at once, shows the link with complementarity framework.

As a graphical illustration of the normal cone $\mathcal{N}_{V(q)}$, we consider the example given in Fig. 1(b), where the two constraint functions are $h_1(q_x, q_y) = -q_y + 4 \geq 0$, and $h_2(q_x, q_y) = q_y - q_x^2 \geq 0$. Three different scenarios are depicted in Fig. 2 for this example corresponding to $\mathcal{J}(q) = \{1\}$ in Fig. 2(a), $\mathcal{J}(q) = \{2\}$ in Fig. 2(b), and $\mathcal{J}(q) = \{1, 2\}$ in Fig. 2(c).

We now formulate the dynamics of system (1) as a measure differential inclusion

$$dq = v dt \tag{8a}$$

$$M(q)dv + F(t, q, v)dt \in -\mathcal{N}_{V(q)}(v_e) \tag{8b}$$

where

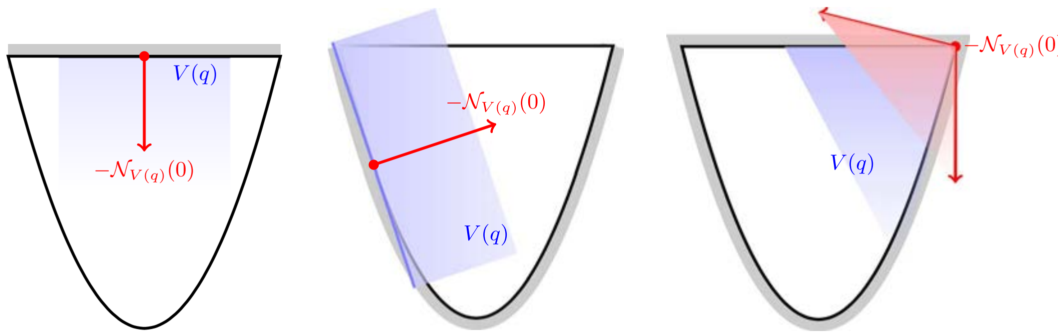
$$v_e(t) := \frac{v(t^+) + ev(t^-)}{1 + e} \tag{8c}$$

and $e \in [0, 1]$ is the coefficient of restitution. The initial condition is assumed to satisfy $q_0 := q(0) \in \Phi$, and $v_0 := v(0)$ is such that $v_e(0) \in V(q_0)$. The motivation for working with the MDI is that we are seeking a solution to the evolution problem in the space of locally *rcbv* functions to deal with possible collisions with the boundary of the admissible set. Functions which are locally *rcbv* possess generalized derivatives that can be identified with Stieltjes measure and (8b) precisely describes the inclusion of the measure dv, associated with $v(\cdot)$, into a normal cone described by the constraint set Φ.

B. Interpreting MDI (8)

Let us first provide some explanations to understand (8).

a) When no constraints are active: It is noted that, if $q \in$ int Φ, that is, $h_i(q) > 0$, for each $1 \leq i \leq m$, so that $\mathcal{J}(q) = \emptyset$,



(a) (b) (c)

Fig. 2. Polyhedral cones for two constraints: $h_1(q) = -q_y + 4 \geq 0$, $h_2(q) = q_y - q_x^2 \geq 0$. (a) Cones for the case $h_1(q) \leq 0$. (b) Cones for the case $h_2(q) \leq 0$. (c) Corner case: $h_1(q) \leq 0 \wedge h_2(q) \leq 0$.

then $V(q) = \mathbb{R}^n$ and consequently $\mathcal{N}_{V(q)}(\cdot) = \{0\}$. This reduces (8) to ordinary differential equations described by $\dot{q} = v$ and $M(q)\dot{v} + F(t, q, v) = 0$.

b) Post-impact velocities: It is also noted that the post-impact velocity determined according to Moreau's collision rule (or Newton's impact law) is directly encoded in the MDI (8). To derive an explicit expression for the post-impact velocity, we proceed as follows:

$$M\left(q(t_k)\right)\left[v\left(t_k^+\right) - v\left(t_k^-\right)\right] \in -\mathcal{N}_{V(q(t_k))}\left(v_e(t_k)\right) \tag{9a}$$

$$\Leftrightarrow\ v\left(t_k^+\right) - v\left(t_k^-\right) \in -M\left(q(t_k)\right)^{-1}\mathcal{N}_{V(q(t_k))}\left(v_e(t_k)\right) \tag{9b}$$

$$\Leftrightarrow\ \frac{1}{1+e}\left[v(t_k^+) - v(t_k^-)\right] \in -M(q(t_k))^{-1}\mathcal{N}_{V(q(t_k))}(v_e(t_k)) \tag{9c}$$

$$\Leftrightarrow\ v_e(t_k) - v\left(t_k^-\right) \in -M\left(q(t_k)\right)^{-1}\mathcal{N}_{V(q(t_k))}\left(v_e(t_k)\right) \tag{9d}$$

$$\Leftrightarrow\ v_e(t_k) = \text{proj}_{M(q(t_k))}\left(V\left(q(t_k)\right); v\left(t_k^-\right)\right) \tag{9e}$$

$$\Leftrightarrow\ v\left(t_k^+\right) = -ev\left(t_k^-\right) + (1+e) \times \text{proj}_{M(q(t_k))}\left(V\left(q(t_k)\right); v\left(t_k^-\right)\right) \tag{9f}$$

where $\text{proj}_{M(q)}(V(q); v)$ denotes the projection of v on the set $V(q)$ according to the kinetic metric at q, which is defined by the inner product $\langle v, w\rangle_{M(q)} = \langle v, M(q)w\rangle = \langle M(q)v, w\rangle$. In the above expression, it is used that a normal cone is invariant under multiplication by a nonnegative scalar in (9c). Expression (9e) is obtained using a well-known result from convex analysis that relates the projection of a point on a convex set with the normal cone to the convex set at the projected point.

One can also interpret MDI (8) at impact times in the sense that, we want to compute $v(t_k^+)$ such that $v_e(t_k)$ belongs to the set $V(q(t_k))$ while minimizing $|v_e(t_k) - v(t_k^-)|_{M(q(t_k))}$. Thus, there is an optimization problem to be solved in order to compute $v(t_k^+)$. One typically reformulates this optimization problem using the framework of complementarity program [13] for which there are already some efficient solvers. More precisely, we let

$$M(q_k)\left[v\left(t_k^+\right) - v\left(t_k^-\right)\right] = -\sum_{\alpha\in\mathcal{J}(q_k)} \lambda_\alpha \nabla h_\alpha(q_k),\ \lambda_\alpha \geq 0$$

where $q_k := q(t_k)$, and λ_α is computed from

$$0 \leq \lambda_\alpha \perp \left\langle \nabla h_\alpha(q_k), v\left(t_k^+\right) + ev\left(t_k^-\right)\right\rangle \geq 0,\ \ \alpha \in \mathcal{J}(q_k). \tag{10}$$

The above complementarity relation is an equivalent way of writing

$$v_e(t_k) \in V(q_k)$$

$$\text{and } \left\langle v\left(t_k^+\right) - v\left(t_k^-\right), M(q_k)\left(v\left(t_k^+\right) + ev\left(t_k^-\right)\right)\right\rangle = 0$$

which are the relations encoded in the MDI (8).

c) Continuous motion on constraint surfaces: Contact with the surface will not always result in the discontinuities of the velocity variable v. From (9f), it is seen that, if $v(t_k^-) \in V(q(t_k))$, then we have $v(t_k^+) = v(t_k^-)$. The MDI (8) in case of continuous motion along the boundary of the constraint is written as

$$M(q)\dot{v} + F(t, q, v) = -\sum_{\alpha\in\mathcal{J}(q)} \lambda_\alpha \nabla h_\alpha(q),\ \ \lambda_\alpha \geq 0$$

where λ_α are again obtained through the relation (10). We see that (8) encapsulates switches to lower dimensional systems, thanks to the existence of suitable multipliers (i.e., contact forces) calculated from a complementarity problem.

The formulation for constrained mechanical systems, as in (8), was pioneered by J. J. Moreau [24], and the MDI (8) describes a state-dependent sweeping process as $V(q)$, the constraint set for velocities appearing in (8b), depends on the state variable $q(\cdot)$. Further details on inclusions of type (8) and comparisons with other modeling frameworks could be found in [8, Section 5.3]. For our purpose, it is seen that the observer design given in Section IV is partially aided by this compact formulation. It is noteworthy that there is a close link between the sweeping process in (8) and so-called complementarity Lagrangian systems, see, e.g., [2, Section 3.6].

C. Assumptions on System Data and Solutions

The solution of MDI (8) is considered in the following sense:

Definition 1: A solution to the Cauchy problem (8) with initial data $(q_0, v_0) \in \Phi \times V(q_0)$, over an interval $I = [0, T]$, is a pair (q, v) such that $v(\cdot)$ is *rcbv* on I; $q(t) = q_0 + \int_0^\tau v(s)ds$; $q(t) \in \Phi$ and $v_e(t) \in V(q(t))$ for all $t \geq 0$; and furthermore, there exists a positive measure (represented by) $d\mu$ such that both dt and dv possess densities with respect to $d\mu$, denoted by $dt/d\mu$, and $dv/d\mu$ respectively, such that

$$M(q)\frac{dv}{d\mu}(t) + F(t,q,v)\frac{dt}{d\mu}(t) \in - \mathcal{N}_{V(q(t))}(v_e),\ d\mu - \text{a.e. on } I. \quad (11)$$

The choice of the measure $d\mu$ is not unique since the right-hand side of (8) is a cone. However, by Lebesgue-Radon-Nikodym theorem, the functions $dt/d\mu(\cdot) \in \mathcal{L}_1(I, R; d\mu)$ and $dv/d\mu(\cdot) \in \mathcal{L}_1(I, \mathbb{R}^n; d\mu)$ are uniquely determined for a given $d\mu$.

The problem of existence of solutions for evolution problems (1) has been studied for a long time. Earlier results on this problem dealt with the single constraint case ($m = 1$) and one may refer to [22, Ch. 3], [32] for results in this direction. The basic idea in these works is to introduce a time discretization scheme, either at position level [32] or velocity level [22] to construct a sequence of approximate solutions which is shown to converge as the step size converges to zero. For several unilateral constraints ($m \geq 2$), the existence and uniqueness has been proved in [3] under analytic assumptions on the data using the solution theory for differential equations and variational inequalities. Building on the results derived in [20], the most relaxed conditions, under which the existence of solutions has been proved using discretization at velocity level, have appeared recently in [14] for the inelastic case ($e = 0$), and in [15], [31] for general values of $e \in [0, 1]$. Based on the work of [31], the following regularity assumptions are required on the system data for the existence of solution, and are also needed for the observer design:

- (**H1**) The function $F(\cdot,\cdot,\cdot)$ is continuous and is continuously differentiable ($\mathcal{C}^1$) with respect to its second and third arguments.
- (**H2**) The mapping $M(\cdot)$, from $\mathbb{R}^n$ to the set of symmetric positive definite matrices, belongs to class $\mathcal{C}^1$ and there exists $0 < \underline{\lambda}_M \leq \overline{\lambda}_M$ such that
$$\underline{\lambda}_M |v|^2 \leq v^\top M(q) v \leq \overline{\lambda}_M |v|^2 \quad \forall (q,v) \in \Phi \times \mathbb{R}^n. \quad (12)$$
- (**H3**) For each $i \in \{1,\dots,m\}$, the function $h_i \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$, its Euclidean gradient $\nabla h_i(q)$ is locally Lipschitz continuous and does not vanish in a neighborhood of $\{q \in \mathbb{R}^n | h_i(q) = 0\}$.
- (**H4**) The active constraints are functionally independent, i.e., $\{\nabla h_i(q)\}_{i\in\mathcal{J}(q)}$ is linearly independent for all $q \in \Phi$.

Without recalling the formal result on existence and assuming that a solution exists in the sense of Definition 1 under hypotheses (**H1**)–(**H4**), we only collect the properties of the solutions to system (8) which provide more insight.

D. Solution Characteristics

1) Regularity of State Trajectories: The function q is absolutely continuous, but not necessarily everywhere differentiable. The velocity $v(\cdot)$ is a locally *rcbv* function, for which the left and right limits are defined everywhere. The acceleration is represented by the measure dv and can be decomposed as a sum of three measures: an atomic measure $d\mu_a$, Lebesgue measure dt, and a measure associated with singularly continuous function $d\mu_{sc}$ i.e., we may write $dv = d\mu_a + \dot{v}dt + d\mu_{sc}$.

2) Countably Many Impacts: The set of impact times, at which $v(\cdot)$ is discontinuous, is at most countable. One may simply take $d\mu_a = \sum_{k\geq 0}[v(t_k^+) - v(t_k^-)]\delta_{t_k}$, where δ_{t_k} is the Dirac impulse at time t_k and $\{t_k\}_{k\geq 0}$ is an ordered sequence of impact times. Thus, the formulation (8) does not exclude the Zeno phenomenon (with a finite or infinite number of left accumulation points). However, if $e = 1$, then it is shown in [4] that there exists a constant $\rho_T(q(0), v(0)) > 0$ such that $t_{k+1} - t_k > \rho_T(q(0), v(0))$, for each t_k, t_{k+1} belonging to a compact interval $[0, T]$.

3) Non-Uniqueness and Continuity of Solutions: The solution of system (8) is unique if the system data is analytic [3], but in general, it may not be the case. Even under the analyticity assumption, the solutions may not vary continuously with respect to initial conditions under the hypotheses (**H1**)–(**H4**). For this to hold, there is an additional condition on the set $\{\nabla h_i(q)\}_{i\in\mathcal{J}(q)}$ given in [30], which states that for $e = 0$, the active constraints must satisfy $\langle \nabla h_i(q), M^{-1}(q)\nabla h_j(q)\rangle \leq 0$, and for $e \in (0, 1]$, $\langle \nabla h_i(q), M^{-1}(q)\nabla h_j(q)\rangle = 0$.

IV. Observer Design

We now address the problem of designing observers for the systems considered in Section III. It will be assumed that the position $q(\cdot)$ is the measured variable, and the objective is then to design an estimator which either estimates the full state (q, v), or only the unknown velocity v of the moving point. The class of state estimators that we propose for this purpose comprises a differential inclusion with state $z := (z_1^\top, z_2^\top)^\top \in \mathbb{R}^{d_z}$ satisfying

$$\dot{z}_1 = F_1(t, q, z) \quad (13a)$$
$$M(q)dz_2 + F_2(t, q, z)dt \in -\mathcal{N}_{V(q)}(\hat{v}_e) \quad (13b)$$

where $\hat{v}_e(t) \in \mathbb{R}^n$ is given by

$$\hat{v}_e(t) = \frac{\hat{v}(t^+) + e\hat{v}(t^-)}{1+e}. \quad (14)$$

The state estimate $(\hat{q}(t), \hat{v}(t)) \in \mathbb{R}^{2n}$ is defined as:

$$\hat{q} = f_1(z_1, q) \quad (15a)$$
$$\hat{v} = z_2 + f_2(z_1, q) \quad (15b)$$

and the function $\binom{z_1}{z_2} \mapsto f(q, z) := \binom{f_1(z_1,q)}{z_2 + f_2(z_1,q)}$ is assumed to be a diffeomorphism for each $q \in \mathbb{R}^n$, so that the function $f^{-1}(q, \cdot)$ is well-defined and continuously differentiable. Choosing the functions F_1, F_2 and f_1, f_2 is a part of the design procedure and we will give two possible ways of choosing these functions so that the estimate $(\hat{q}, \hat{v})$ converges asymptotically to the actual state (q, v). Moreover, it will be shown that, under certain regularity assumptions on the functions F_1, F_2 and f_1, f_2, there exists a unique solution to the proposed observer (13).

Before proceeding towards these main results, we choose to rewrite the observer dynamics in $(\hat{q}, \hat{v})$ coordinates

$$\dot{\hat{q}}(t) = \widehat{F}_1(t, x, \hat{x}) \quad (16a)$$
$$M(q)d\hat{v} + \widehat{F}_2(t, x, \hat{x}) \in -\mathcal{N}_{V(q)}(\hat{v}_e) \quad (16b)$$

where, for brevity, we let

$$\widehat{F}_1(t, x, \hat{x}) := \frac{\partial f_1}{\partial z_1} F_1\left(t, q, f^{-1}(q, x)\right) + \frac{\partial f_1}{\partial q} v \quad (16c)$$
$$\widehat{F}_2(t, x, \hat{x}) := F_2\left(t, q, f^{-1}(q, x)\right) + M(q)\frac{\partial f_2}{\partial z_1} F_1\left(t, q, f^{-1}(q, x)\right) - M(q)\frac{\partial f_2}{\partial q} v. \quad (16d)$$

This new description of the observer dynamics also provides an insight about its mechanism. Equation (16b) basically tells us that the estimate $\hat{v}$ is constrained in the same way as the actual velocity v. The nonsmooth behavior in the velocity variable is due to the forces that belong to the set $\mathcal{N}_{V(q)}(v)$. By measuring the position variable, the set $V(q)$ can be computed at each time. One then uses the monotonicity property of the normal cone operator $\mathcal{N}_{V(q)}(\cdot)$ in analyzing the error dynamics to show convergence.

In Sections V and VI, we will show that the proposed observer (16) has the following two properties, respectively:

- **Well-posedness**: For each absolutely continuous function $q(\cdot)$, there exists a unique locally *rcbv* function $\hat{v}(\cdot)$ obtained from (13)–(15).
- **Error convergence**: The estimates $\hat{q}(\cdot), \hat{v}(\cdot)$ converge to $q(\cdot), v(\cdot)$ asymptotically.

Before proceeding with these technical results, note that the original system may not have unique solutions, but the observer has the property that it generates a unique trajectory corresponding to each function $q(\cdot)$ observed as an output of system (8) ; see [10, Remark 3.3] for further explanation along these lines.

V. Observer Well-Posedness

The estimator (13) is actually an evolution inclusion in which the multi-valued function $\mathcal{N}_{V(q)}(\cdot)$ is closed and convex valued. It is noted that the function $q(\cdot)$ is seen as an external "input" by the observer and hence $V(q(\cdot))$ is seen as a time-parameterized multi-valued function that does not depend on any of the internal states of the estimator. We will basically prove the well-posedness result for the differential inclusion

$$M\left(q(t)\right) d\hat{v} + g(t,\hat{v})dt \in -\mathcal{N}_{V(q(t))}(\hat{v}_e) \tag{17}$$

under certain regularity assumption on the function $g(t,\hat{v})$. Using this result, it will be shown that the observer (13) can be transformed into a system of form (17). The solution to system (13) is interpreted in a sense similar to Definition 1.

Let us now state the following result on existence and uniqueness of solution to (17). This is a fundamental step since the existence of a solution secures that the error stability analysis is meaningful, while uniqueness property secures that the observer output is unique for a given plant trajectory, as stated earlier.

Theorem 1: Consider the differential inclusion (17) under the hypotheses (**H2**)–(**H4**), and $V(q)$ defined in (6). Assume that the function $q:[0,T]\to\mathbb{R}^n$ is absolutely continuous, and that the function $g:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n$ satisfies

$$|g(t,\hat{v}_1)-g(t,\hat{v}_2)| \le C_{g,l}|\hat{v}_1-\hat{v}_2| \qquad \forall\, \hat{v}_1,\hat{v}_2\in\mathbb{R}^n, \quad \forall\, t\in[0,T] \tag{18}$$

$$|g(t,\hat{v})| \le C_{g,b}\left(1+|\hat{v}|\right), \quad \forall\, t\in[0,T] \tag{19}$$

for some constants $C_{g,l}, C_{g,b}>0$. Then the system (17) is well-posed, that is, there exists a unique solution $\hat{v}\in BV([0,T];\mathbb{R}^n)$ for any initial condition $\hat{v}(0)\in V(q(0))$. Moreover, it holds that

$$\hat{v}_e(t)\in V\left(q(t)\right), \quad \forall t\in[0,T]. \tag{20}$$

The result on existence and uniqueness of solutions for MDI (17), stated in Theorem 1, is important in several respects:

(a) The multivalued operator on the right-hand side is non-compact, time-varying and the variation of this set-valued map (measured using Hausdorff-distance) is not bounded.

(b) Even though the interior of $V(q)$ for each q is nonempty, in general, there does not exist any *common* open ball which is contained in $V(q)$, for each $q\in\Phi$.

(c) The argument of $\mathcal{N}_{V(q)}(\cdot)$ is not simply the state $\hat{v}$ but rather a weighted sum of pre- and post-impact values of $\hat{v}(\cdot)$.

(d) The mapping $t\mapsto V(q(t))$ is lower semicontinuous (because $t\mapsto q(t)$ is absolutely continuous and $q\mapsto V(q)$ is lower semicontinuous).

Because of these reasons, we cannot use the existing results on solutions of time-dependent sweeping processes, for example [22, Ch. 2], in a straightforward manner. Moreover, the numerical implementation of the examples considered in this paper (see Section VII) is based on a time-discretization procedure and the proof of Theorem 1 shows that the proposed sequence of discretized solutions indeed converges to a unique solution of system (17). With this motivation, we work out a formal proof of Theorem 1 in this paper. In this section, we will only develop an outline which shows all the steps involved in the proof and for some of these steps, detailed calculations are given in Section VIII. Before discussing the proof, we first show how the result of Theorem 1 can be used to study the well-posedness of the observer class (13).

A. Applying Theorem 1 to Observer (13)

Our goal is to show that the proposed observer (13) can be written in the form of (17), and that the hypotheses of Theorem 1 hold in this case. To see this, one can rewrite the description of the observer in (16) as follows:

$$\begin{bmatrix} I & 0 \\ 0 & M\left(q(t)\right)\end{bmatrix}\begin{pmatrix} d\hat{q} \\ d\hat{v}\end{pmatrix} + \begin{pmatrix}\widehat{F}_1\left(t,q(t),v(t),\hat{q},\hat{v}\right) \\ \widehat{F}_2\left(t,q(t),v(t),\hat{q},\hat{v}\right)\end{pmatrix} \in -\mathcal{N}_{\mathbb{R}^n\times V(q)}\begin{pmatrix}\hat{q} \\ \hat{v}_e\end{pmatrix}. \tag{21}$$

The underlying reasoning behind this transformation is that $\mathcal{N}_{\mathcal{S}_1\times\mathcal{S}_2}(\hat{q},\hat{v}) = \mathcal{N}_{\mathcal{S}_1}(\hat{q})\times\mathcal{N}_{\mathcal{S}_2}(\hat{v})$, for $\hat{q}\in\mathcal{S}_1, \hat{v}\in\mathcal{S}_2$, and $\mathcal{S}_1,\mathcal{S}_2$ being closed, convex subsets of $\mathbb{R}^n$. Let $\widehat{F}(t,\hat{x}) := \left(\begin{smallmatrix}\widehat{F}_1(t,q(t),v(t),\hat{q},\hat{v}) \\ -\widehat{F}_2(t,q(t),v(t),\hat{q},\hat{v})\end{smallmatrix}\right)$, where we see q, v as functions of time, and use the notation $\hat{x} := \left(\hat{q}^\top,\hat{v}^\top\right)^\top$. We now have the following corollary:

Corollary 1: Consider the differential inclusion (21) under the hypotheses (**H2**)–(**H4**), and $V(q)$ defined in (6). Assume that the function $q:[0,T]\to\mathbb{R}^n$ is absolutely continuous, $v:[0,T]\to\mathbb{R}^n$ is a function of bounded variation, and that the function $\widehat{F}:[0,T]\times\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ satisfies

$$\left|\widehat{F}(t,\hat{x}_1)-\widehat{F}(t,\hat{x}_2)\right| \le C_{\widehat{F},l}|\hat{x}_1-\hat{x}_2| \qquad \forall\, \hat{x}_1,\hat{x}_2\in\mathbb{R}^{2n}, \quad \forall\, t\in[0,T]$$

$$\left|\widehat{F}(t,\hat{x})\right| \le C_{\widehat{F},b}\left(1+|\hat{x}|\right) \qquad \forall\, \hat{x}\in\mathbb{R}^{2n}, \quad \forall\, t\in[0,T]$$

for some constants $C_{\widehat{F},l}, C_{\widehat{F},b} > 0$. Then the system (21) is well-posed, that is, there exists a unique solution $\hat{q} \in AC([0,T], \mathbb{R}^n)$, $\hat{v} \in BV([0,T];\mathbb{R}^n)$ for any initial condition $(\hat{q}(0), \hat{v}(0)) \in \mathbb{R}^n \times V(q(0))$. Moreover, it holds that

$$\hat{v}_e(t) \in V\left(q(t)\right) \quad \forall\, t \in [0,T]. \tag{22}$$

The proof of this corollary is a direct application of Theorem 1 where we work with the augmented variable $(\hat{q}^\top, \hat{v}^\top)^\top$. The hypotheses stated in Theorem 1 for (17) also hold for (21). However, Corollary 1 claims that $\hat{q}$ is absolutely continuous, whereas Theorem 1 only guarantees solutions in the class of bounded variation functions. This extra regularity on $\hat{q}$ follows due to the fact that $\hat{q}$ dynamics are basically unconstrained and are obtained by integrating $\widehat{F}_1(t, q(t), v(t), \hat{q}, \hat{v})$.

B. *Proof Outline for Theorem 1*

Consider a partition $\mathcal{P}$ of the interval $[0,T]$ given by

$$\mathcal{P} := \{t_{\mathcal{P},i}, 0 \le i \le N_\mathcal{P}\}$$
$$0 = t_{\mathcal{P},0} < t_{\mathcal{P},1} < t_{\mathcal{P},2} < \cdots < t_{\mathcal{P},N_\mathcal{P}} = T$$

and let

$$\hat{v}_{\mathcal{P},0} = \hat{v}_0 \tag{23a}$$
$$\hat{v}_{\mathcal{P},i} = -e\hat{v}_{\mathcal{P},i-1} + (1+e)\mathrm{proj}_{M_{\mathcal{P},i}}\left[\hat{v}_{\mathcal{P},i-1} - \frac{1}{1+e} M_{\mathcal{P},i}^{-1} G_{\mathcal{P},i}, V_{\mathcal{P},i}\right] \tag{23b}$$

where $G_{\mathcal{P},i} := \int_{t_{\mathcal{P},i-1}}^{t_{\mathcal{P},i}} g(s, \hat{v}_{\mathcal{P},i-1}) ds$ and $M_{\mathcal{P},i} := M(q(t_{\mathcal{P},i}))$. One can then define a piecewise constant solution $\hat{v}_\mathcal{P}(\cdot)$ for each partition $\mathcal{P}$ as follows:

$$\hat{v}_\mathcal{P}(t) := \begin{cases} \hat{v}_{\mathcal{P},i} & t \in [t_{\mathcal{P},i}, t_{\mathcal{P},i+1}) \\ \hat{v}_{\mathcal{P},N_\mathcal{P}} & t = t_{N_\mathcal{P}}. \end{cases} \tag{24}$$

The motivation behind defining the successive elements of a piecewise constant solution using (23b) is that it gives[2]

$$\frac{\hat{v}_{\mathcal{P},i} + e\hat{v}_{\mathcal{P},i-1}}{1+e} = \mathrm{proj}_{M_{\mathcal{P},i}}\left[\hat{v}_{\mathcal{P},i-1} - \frac{1}{1+e} M_{\mathcal{P},i}^{-1} G_{\mathcal{P},i}, V_{\mathcal{P},i}\right]$$
$$\iff M_{\mathcal{P},i}(\hat{v}_{\mathcal{P},i} - \hat{v}_{\mathcal{P},i-1}) + G_{\mathcal{P},i} \in -\mathcal{N}_{V_{\mathcal{P},i}}\left(\frac{\hat{v}_{\mathcal{P},i} + e\hat{v}_{\mathcal{P},i-1}}{1+e}\right)$$

which is a quite natural discretization of (17).

In the sequel:

- a uniform bound (with respect to $\mathcal{P}$) is derived on $|\hat{v}_\mathcal{P}|$ in Section VIII-A;
- an estimate of the total variation of $\hat{v}_\mathcal{P}$ over a compact interval is computed in Section VIII-B.

Using these bounds to invoke a generalized version of Helly's first theorem (see Theorem A.2 in Appendix A), there exists a filter $\mathcal{F}$ finer than the filter of sections of $\mathcal{P}$, and a function of bounded variation $\hat{v} : [0,T] \to \mathbb{R}^n$ which is the weak pointwise limit of $\hat{v}_\mathcal{P}(\cdot)$ with respect to $\mathcal{F}$. Since we are working in the finite-dimensional setup, $\hat{v}(\cdot)$ is a strong pointwise generalized sublimit of $\hat{v}_\mathcal{P}$:

$$\lim_{\mathcal{F}} |\hat{v}(t) - \hat{v}_\mathcal{P}(t)| = 0 \qquad \forall\, t \in [0,T]. \tag{25}$$

The next step is to show that $\hat{v}(\cdot)$ obtained above is indeed a solution to system (17). We demonstrate it by showing that:

- the differential inclusion (17) holds at continuity points of $\hat{v}$ (Section VIII-C);
- the inclusion (17) is satisfied at discontinuity points of $\hat{v}$ (Section VIII-D).

The fact that $\hat{v}_e(t) \in V(q(t))$ follows due to closedness of $V(q(t))$. To complete the proof, it remains to show that the solution to (17) is unique, which basically follows due to convexity of $V(q(t))$ and Lipschitz continuity of $g(t,\cdot)$. To see that, let $\hat{v}^1(\cdot)$, $\hat{v}^2(\cdot)$ be two solutions to (17) with $\hat{v}^1(0) = \hat{v}^2(0)$, then there exists a measure $d\hat{\mu}$ such that

$$M\left(q(t)\right) \frac{d\hat{v}^i}{d\hat{\mu}}(t) + g(t, \hat{v}^i)\frac{dt}{d\hat{\mu}}(t) \in -\mathcal{N}_{V(q(t))}(\hat{v}^i), \quad i = 1,2.$$

Using the monotonicity property of the normal cone, we get

$$\left\langle M\left(q(t)\right)\left(\frac{d\hat{v}^1}{d\hat{\mu}}(t) - \frac{d\hat{v}^2}{d\hat{\mu}}(t)\right),\ \hat{v}^1(t) - \hat{v}^2(t)\right\rangle$$
$$\le \left|g\left(t, \hat{v}^2(t)\right) - g\left(t, \hat{v}^1(t)\right)\right| \cdot \left|\hat{v}^1(t) - \hat{v}^2(t)\right|.$$

Since $g(t,\cdot)$ is Lipschitz, and $\hat{v}^1(0) = \hat{v}^2(0)$, the above inequality becomes

$$|\hat{v}_1 - \hat{v}_2|^2 \le \frac{2C_{g,l}}{\underline{\lambda}_M} \int_{]0,t]} |\hat{v}_1(s) - \hat{v}_2(s)|^2\, d\hat{\mu}(s).$$

One can now invoke the Gronwall-Bellman like lemma for functions of bounded variation [18, Lemma 4], to get

$$|\hat{v}_1 - \hat{v}_2|^2 \le 0$$

whence it follows that $\hat{v}^1(t) = \hat{v}^2(t)$, for $t \in [0,T]$.

VI. Error Stability Analysis

In this section, we address the convergence of the state estimation error to zero. In what follows, let $x := (q^\top, v^\top)^\top$, $\hat{x} := (\hat{q}^\top, \hat{v}^\top)^\top$, and let the state estimation error be denoted by $\tilde{x} := (\tilde{q}^\top, \tilde{v}^\top)^\top := x - \hat{x}$. The main result on convergence of error now follows:

Theorem 2: Assume that there exists a symmetric positive definite matrix-valued function $R : \mathbb{R}^n \to \mathbb{R}^{2n \times 2n}$, $q \mapsto R(q)$

$$R(q) = \begin{bmatrix} R_{11} & 0 \\ 0 & M(q) \end{bmatrix}$$

and a constant $\beta > 0$ such that

$$2\tilde{x}^\top R(q)\begin{pmatrix} v - \widehat{F}_1(t,x,\hat{x}) \\ -F(t,x) + \widehat{F}_2(t,x,\hat{x}) \end{pmatrix} + \underbrace{\tilde{x}^\top \dot{R}(q,v)\tilde{x}}_{=\tilde{v}^\top \dot{M}(q,v)\tilde{v}} \le -\beta\tilde{x}^\top R(q)\tilde{x} \tag{26}$$

then the state estimation error decays exponentially, that is

$$|\tilde{x}(t)| \le \mathsf{e}^{-\beta\, t}\, |\tilde{x}(0)|.$$

[2] We use the fact that for a convex set V, it holds that $x = \arg\min_{y \in V} |z - y|_M$, that is, x is the projection of z onto V with respect to the norm induced by a symmetric positive definite matrix M, *if and only if* $\langle M(z-x), y-x\rangle \le 0$, $\forall\, y \in V \Leftrightarrow M(z-x) \in \mathcal{N}_V(x)$.

The statement of Theorem 2 basically requires us to choose a state estimator where the unconstrained ODEs result in error dynamics which are dissipative with respect to a quadratic Lyapunov function. The matrix that determines this quadratic form has some structure described by $R(q)$. We will show in Section VI-A and B that two possible observer design techniques for unconstrained Lagrangian systems could be tailored into the framework of (13), and satisfy the conditions for well-posedness listed in Theorem 1 and the stability requirements given in Theorem 2.

Proof of Theorem 2: The error dynamics are defined as

$$\dot{\tilde{q}} = v - \widehat{F}_1(t,x,\hat{x}) \tag{27a}$$

$$M(q)d\tilde{v} + F(t,x) - \widehat{F}_2(t,x,\hat{x}) \in -(\eta - \hat{\eta}) \tag{27b}$$

where

$$\eta \in \mathcal{N}_{V(q)}(v_e) \quad \text{and} \quad \hat{\eta} \in \mathcal{N}_{V(q)}(\hat{v}_e)$$

and $v_e, \hat{v}_e$ are defined as in (8c) and (14), respectively.

In what follows, we fix $d\mu = d\mu_c + d\mu_a + d\hat{\mu}_a$, where μ_c denotes the continuous part of the measure μ, and we chose μ_c to be the sum of the Lebesgue measure dt, and the singularly continuous component μ_{sc}, that is, $d\mu_c = dt + d\mu_{sc}$. The atomic measure $d\mu_a$ (respectively, $d\hat{\mu}_a$) is supported by the time instants at which $v(\cdot)$ (respectively, $\hat{v}(\cdot)$) is discontinuous. It is seen that $d\mu_c + d\mu_a$, and $d\mu_c + d\hat{\mu}_a$ are absolutely continuous with respect to $d\mu$ and hence the densities $(dv/d\mu)(\cdot)$ and $(d\hat{v}/d\mu)(\cdot)$ are well-defined on the complement of a $d\mu$-null set.

Pick $W(q,\tilde{x}) = \tilde{x}^\top R(q)\tilde{x}$, then $W(\cdot)$ is locally *rcbv* using the chain rule [26, Theorem 3], and its differential is computed as follows (see also [9, Eq. (13)]):

$$\begin{aligned}\frac{dW}{d\mu}(t) &= \left(\tilde{x}(t^+) + \tilde{x}(t^-)\right)^\top R(q)\frac{d\tilde{x}}{d\mu}(t)\\ &\quad + \frac{\partial}{\partial q}\left(\tilde{x}(t^+)^\top R\left(q(t)\right)\tilde{x}(t^+)\right)\frac{dq}{d\mu}(t).\end{aligned}$$

If there is a jump in $v(\cdot)$ or $\hat{v}(\cdot)$ at t_k, then $(d\mu_c/d\mu)(t_k) = 0$, which also implies that $(dq/d\mu)(t_k) = (dq/d\mu_c)(t_k)\cdot(d\mu_c/d\mu)(t_k) = 0$, and $(d\tilde{q}/d\mu)(t_k) = 0$ from (27a). We thus obtain

$$\begin{aligned}\frac{dW}{d\mu}(t_k) &= \left(\tilde{v}_k^+ + \tilde{v}_k^-\right)^\top M_k \frac{d\tilde{v}}{d\mu}(t_k)\\ &= \left(\tilde{v}_k^+ + \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right)\end{aligned}$$

where we used the notation $\tilde{v}_k^+ := \tilde{v}(t_k^+)$, $\tilde{v}_k^- := \tilde{v}(t_k^-)$, and $M_k := M(q(t_k))$. We can rewrite the above expression as

$$\begin{aligned}\frac{dW}{d\mu}(t_k) &= \tilde{v}_k^{+\top} M_k \tilde{v}_k^+ - \tilde{v}_k^{-\top} M_k \tilde{v}_k^-\\ &\quad \pm \frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right)\\ &= \frac{1}{(1+e)}\Big[(1+e)\left(\tilde{v}_k^{+\top} M_k \tilde{v}_k^+ - \tilde{v}_k^{-\top} M_k \tilde{v}_k^-\right)\\ &\qquad + (1-e)\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right)\Big]\\ &\quad - \frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right)\\ &= \frac{2}{(1+e)}\left[\tilde{v}_k^{+\top} M_k \tilde{v}_k^+ - e\tilde{v}_k^{-\top} M_k \tilde{v}_k^-\right.\\ &\qquad \left. - (1-e)\tilde{v}_k^{+\top} M_k \tilde{v}_k^-\right]\\ &\quad - \frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right)\\ &= \frac{2}{(1+e)}\left\langle M_k\left(\tilde{v}_k^+ - \tilde{v}_k^-\right), \tilde{v}_k^+ + e\tilde{v}_k^-\right\rangle\\ &\quad - \frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right).\end{aligned}$$

Substituting $M(q(t_k))(\tilde{v}_k^+ - \tilde{v}_k^-) = -(\eta_k - \hat{\eta}_k)$, the above equation becomes

$$\begin{aligned}\frac{dW}{d\mu}(t_k) &= -2\left\langle \eta_k - \hat{\eta}_k, \frac{\tilde{v}_k^+ + e\tilde{v}_k^-}{1+e}\right\rangle\\ &\quad - \frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right).\end{aligned} \tag{28}$$

By definition [see (3)], it follows that:

$$\eta_k \in \mathcal{N}_{V(q(t_k))}\left(v_e(t_k)\right) \Longleftrightarrow \langle \eta_k, v_e(t_k) - \hat{v}_e(t_k)\rangle \geq 0 \tag{29}$$

$$\hat{\eta}_k \in \mathcal{N}_{V(q(t_k))}\left(\hat{v}_e(t_k)\right) \Longleftrightarrow \langle \hat{\eta}_k, v_e(t_k) - \hat{v}_e(t_k)\rangle \leq 0 \tag{30}$$

which in turn implies that:

$$\langle \eta_k - \hat{\eta}_k, v_e(t_k) - \hat{v}_e(t_k)\rangle \geq 0$$

or equivalently

$$\left\langle \eta_k - \hat{\eta}_k, \frac{\tilde{v}_k^+ + e\tilde{v}_k^-}{1+e}\right\rangle \geq 0. \tag{31}$$

Using the inequality (31) in (28), we get

$$\frac{dW}{d\mu}(t_k) \leq -\frac{(1-e)}{(1+e)}\left(\tilde{v}_k^+ - \tilde{v}_k^-\right)^\top M_k \left(\tilde{v}_k^+ - \tilde{v}_k^-\right) \leq 0. \tag{32}$$

Thus, when $0 \leq e < 1$, we have a strict decrease in the value of Lyapunov function $W(\cdot)$ at jump instants, and $W(\cdot)$ at most remains constant for the case $e = 1$.

If $t \neq t_k$, then we are interested in computing $(dW/d\mu_c)(t)$. By definition, $(d\mu_c/d\mu)(t) = 1$ when $t \neq t_k$, and

$$\begin{aligned}\frac{dW}{d\mu}(t) &= \left\langle R(q)\tilde{x}, \frac{dx}{d\mu}(t) - \frac{d\hat{x}}{d\mu}(t)\right\rangle\\ &\quad + \frac{\partial}{\partial q}\left(\tilde{x}(t)^\top R\left(q(t)\right)\tilde{x}(t)\right)\frac{dq}{d\mu}(t)\\ &= 2\tilde{x}^\top(t)R(q)\begin{pmatrix} v - \widehat{F}_1(t,x)\\ -F(t,x) + \widehat{F}_2(t,x,\hat{x})\end{pmatrix}\frac{dt}{d\mu}(t)\\ &\quad - \tilde{v}^\top(t)(\eta - \hat{\eta}) + \tilde{x}(t)^\top \dot{R}\left(q(t),v(t)\right)\tilde{x}(t)\frac{dt}{d\mu}(t).\end{aligned}$$

In the above expression, $\tilde{v}^\top(\eta - \hat{\eta}) \geq 0$ because $v(t), \hat{v}(t) \in V(q(t))$ for all t at which v, $\hat{v}$ are continuous, due to which $\langle \eta, v - \hat{v}\rangle \geq 0$, and $\langle \hat{\eta}, v - \hat{v}\rangle \leq 0$. It now follows under condition (26) that

$$\frac{dW}{d\mu}(t) \leq -\beta W(t)\frac{dt}{d\mu}(t), \quad t \neq t_k. \tag{33}$$

Since we fixed $d\mu = d\mu_c + d\mu_a + d\hat{\mu}_a$, and W is non-increasing at the atoms of $d\mu_a$ and $d\hat{\mu}_a$ because of (32), and decreasing exponentially with respect to continuous measure due to (33). One can now invoke the chain rule for differential of bounded variation functions [26] to arrive at the following inequality (the formal arguments can also be found in our recent work [37, Proof of Theorem 1]):

$$W(t) \leq \mathrm{e}^{-\beta t} W(0) \Rightarrow |\tilde{v}(t)| \leq \sqrt{\frac{\overline{\lambda}_M}{\underline{\lambda}_M}} \mathrm{e}^{-\beta t} |\tilde{v}(0)|$$

for all $t \geq 0$. □

It is worth mentioning that, in order to deal with discontinuities of v, we do not just consider the classical derivative of the storage function, but instead compute the density of dW with respect to $d\mu$. We also remark that the condition (26) was introduced explicitly to obtain dissipation of smooth part of the error dynamics with respect to kinetic metric. It basically highlights the fact that if there is any observer available in the literature for smooth Lagrangian systems for which the continuous error dynamics admit $\tilde{x}^\top R(q)\tilde{x}$ as the Lyapunov function, then those designs could be embedded into the formalism of (13) to arrive at a different criteria for error convergence. We now show two particular instances of how observers in the literature for unconstrained Lagrangian systems can be modified to fit in the framework of (13), and satisfy the conditions required for well-posedness and convergence of state estimation error.

A. Full-Order Observer

To arrive at a result on convergence of state estimation error, we introduce additional structure on the nonlinear term $F(t, q, v)$ which is natural for Lagrangian dynamical systems. We suppose that the following assumption holds:

Assumption 1: The velocity $v(\cdot)$ obtained as a solution to (8) stays bounded, that is

$$v(t) \in \mathcal{B}_v := \{v \in \mathbb{R}^n \mid |v| \leq C_v\} \quad \forall\, t \geq 0. \tag{34}$$

The following properties are satisfied by such systems [28]:

- **(P1)** If $\dot{M}(q, v)$ denotes the derivative of the mass matrix, then $\dot{M}(q, v) - 2C(q, v)$ is a skew-symmetric operator, that is, $\tilde{v}^\top(\dot{M}(q, v) - 2C(q, v))\tilde{v} = 0,\ \forall\, \tilde{v} \in \mathbb{R}^n$. Here, $C(q, v)v$ is defined using Christofel symbols and denotes the Coriolis and centrifugal torques.
- **(P2)** There exists a constant $C_M(q) > 0$ s.t.

$$\|C(q, v)\| \leq C_M(q)|v|, \quad \forall\, v \in \mathcal{B}_v. \tag{35}$$

Before describing the observer dynamics, we let $\overline{F}(t, q, \cdot)$ denote the Lipschitz extension[3] of $F(t, q, \cdot)$ from $\mathcal{B}_v$ such that there exists $C_F(t, q)$ satisfying

$$\left|\overline{F}(t, q, v_1) - \overline{F}(t, q, v_2)\right| \leq C_F(t, q) \cdot |v_1 - v_2|, \quad \forall\, v_1, v_2 \in \mathbb{R}^n$$

and it is understood by definition that $\overline{F}(t, q, v) = F(t, q, v)$ for $v \in \mathcal{B}_v$. The idea of using Lipschitz extension of the system vector fields for state estimators appeared in [33].

[3]For a locally Lipschitz function $F(t, q, \cdot) : \mathcal{B} \to \mathbb{R}^n$, the function $\overline{F}(t, q, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is called the Lipschitz extension of $F(t, q, \cdot)$ from $\mathcal{B} \subset \mathbb{R}^n$ if $\overline{F}(t, q, \cdot)$ is globally Lipschitz over $\mathbb{R}^n$ and $\overline{F}(t, q, v) = F(t, q, v)$ for all $(t, q) \in \mathbb{R} \times \mathbb{R}^n$ and $v \in \mathcal{B}$.

The following full-state observer is an adaptation of the design presented in [6] for unconstrained Lagrangian systems:

$$\dot{z}_1 = z_2 + L_d(q - z_1) \tag{36a}$$

$$M(q)dz_2 + \overline{F}(t, q, \hat{v})dt - (L_{\rho 1} + M(q)L_{\rho 2})(q - z_1)dt \in -\mathcal{N}_{V(q)}(\hat{v}_e) \tag{36b}$$

where we let the estimates to be[4]

$$\hat{q} = z_1 \tag{36c}$$

$$\hat{v} = z_2 + l_d \tilde{q}. \tag{36d}$$

It is seen that the observer (36) indeed fits within the general framework proposed in (13) and satisfies the assumptions required for well-posedness.

The matrices $L_{\rho 1}$ and Λ are symmetric, positive definite, and the matrices L_d and $L_{\rho 2}$ are defined as follows:

$$L_d := l_d I + \Lambda, \qquad L_{\rho 2} := l_d \Lambda$$

for some scalar $l_d > 0$. One can equally write

$$\dot{\hat{q}} = \hat{v} - \Lambda(q - \hat{q})$$

$$M(q)d\hat{v} + \overline{F}(t, q, \hat{v})dt - L_{\rho 1}(q - \hat{q})dt - l_d M(q)(v - \hat{v})dt \in -(\eta - \hat{\eta}).$$

Corollary 2: Consider system (8) under hypotheses (**H1**)–(**H4**) and assume that the properties (**P1**), (**P2**), and Assumption 1 hold. For the estimator (36), if $l_d > 0$ is chosen such that the condition

$$\underline{\lambda}_M l_d > C_F(t, q) + C_M(q)C_v + \beta$$

is satisfied for all $(t, q) \in \mathbb{R}_+ \times \Phi$, and some constant $\beta > 0$, then the estimates $\hat{q}(\cdot), \hat{v}(\cdot)$ given by (36c) and (36d), respectively, converge to $q(\cdot), v(\cdot)$ exponentially, that is, for some $c > 0$

$$|x(t) - \hat{x}(t)| \leq c\, \mathrm{e}^{-\beta t} |x(0) - \hat{x}(0)|. \tag{37}$$

Proof: To show that (26) holds, we let $R_{11} := L_{\rho 1}$, $\widehat{F}_1(t, x, \hat{x}) = \hat{v} - \Lambda(q - \hat{q})$, and $\widehat{F}_2(t, x, \hat{x}) = \overline{F}(t, q, \hat{v}) - L_{\rho 1}(q - \hat{q}) - l_d M(q)(v - \hat{v})$, and observe that

$$\begin{aligned}
&\tilde{q}^\top L_{\rho_1}(v - \Lambda\tilde{q}) + \tilde{v}^\top\big(-F(t, q, v) + \overline{F}(t, q, \hat{v}) - L_{\rho 1}\tilde{q} - l_d M(q)\tilde{v}\big) \\
&\quad + \tilde{v}^\top \dot{M}(q, v)\tilde{v} \\
&\qquad \leq -\tilde{q}^\top L_{\rho 1}\Lambda\tilde{q} - l_d\, \tilde{v}^\top M(q)\tilde{v} \\
&\qquad\quad + \tilde{v}^\top \left(\overline{F}(t, q, \hat{v}) - F(t, q, v) + C(q, v)\tilde{v}\right) \\
&\qquad \leq -\tilde{q}^\top L_{\rho 1}\Lambda\tilde{q} - l_d \underline{\lambda}_M |\tilde{v}|^2 \\
&\qquad\quad + |\tilde{v}| \left(C_{F(t,q)}(q)|\tilde{v}| + C_M(q)C_v|\tilde{v}|\right) \\
&\qquad \leq -\tilde{q}^\top L_{\rho 1}\Lambda\tilde{q} - \beta|\tilde{v}|^2
\end{aligned}$$

and the exponential decay of the state estimation error now follows from Theorem 2. □

[4]The definition of $\hat{v}$ considered in [6] is different than the definition of $\hat{v}$ considered here. In [6], the authors take $\hat{v} = z_2 + L_d\tilde{q}$, whereas in our definition $\hat{v} = z_2 + l_d\tilde{q}$. Due to this difference, the error variable $(\tilde{q}^\top, \tilde{v}^\top)^\top$ in our calculations is a linear transformation of the error variable considered in [6]. The reason for introducing this linear transformation is that it allows us to work with a quadratic Lyapunov function $\tilde{x}^\top R(q)\tilde{x}$ where $R(q)$ is block-diagonal as required in the statement of Theorem 2.

B. Partial-Order Observer

One can also design a reduced order observer to show that the conditions of Theorem 2 hold in such case. Consider the following state estimator:

$$M(q)dz + \overline{F}(t,q,\hat{v})dt - l_d M(q)\hat{v}dt \in -\mathcal{N}_{V(q)}(\hat{v}_e) \quad (38a)$$

where we let

$$\hat{q}(t) = q(t) \quad (38b)$$

$$\hat{v}(t) := z(t) - l_d q(t) \quad (38c)$$

$$\hat{v}_e(t) := \frac{\hat{v}(t^+) + e\hat{v}(t^-)}{1+e} \quad (38d)$$

and $q(\cdot)$ in (38a) and (38c) is an absolutely continuous function of time which is obtained from (8) as the measured output. The initial condition $\hat{v}(0) \in V(q(0))$. Once again, it is seen that the observer (38) falls under the class of estimators proposed in (13), and satisfies the regularity conditions stated in Theorem 1 for well-posedness. For error convergence, we have the following result similar to Corollary 2.

Corollary 3: Consider system (8) under hypotheses (**H1**)–(**H4**) and assume that the properties (**P1**), (**P2**), and Assumption 1 hold. For the estimator (38), if the constant $l_d > 0$ is chosen such that the condition

$$l_d \underline{\lambda}_M \geq 2C_v C_M(q) + 2C_F(t,q) + \beta \quad (39)$$

for all $(t,q) \in \mathbb{R}_+ \times \Phi$, and some constant $\beta > 0$, then the velocity estimate $\hat{v}(\cdot)$ given by (38c) converges to $v(\cdot)$ exponentially, that is, for some $c > 0$

$$|v(t) - \hat{v}(t)| \leq c\,\mathrm{e}^{-\beta t} |v(0) - \hat{v}(0)|. \quad (40)$$

Since we have chosen $\hat{q}(t) = q(t)$, we have $\tilde{q}(t) \equiv 0$. We let $\widehat{F}_1(t) = v(t)$, and observe that

$$\begin{aligned}
&\tilde{v}^\top \left(-F(t,q,v) + \overline{F}(t,q,\hat{v}) - l_d M(q)\tilde{v}\right) + \tilde{v}^\top \dot{M}(q,v)\tilde{v} \\
&\quad\leq -l_d \tilde{v}^\top M(q)\tilde{v} + \tilde{v}^\top \left(\overline{F}(t,q,\hat{v}) - F(t,q,v) + C(q,v)\tilde{v}\right) \\
&\quad\leq -l_d \underline{\lambda}_M |\tilde{v}|^2 + |\tilde{v}| \left(C_{F(t,q)}(q)|\tilde{v}| + C_M(q)C_v|\tilde{v}|\right) \\
&\quad\leq -\beta|\tilde{v}|^2
\end{aligned}$$

from where (26) follows and Theorem 2 can now be invoked to show the exponential convergence of the state estimate.

C. Passivity Interpretation

Lagrangian systems are basically modeled such that the total energy, that is, the sum of kinetic and potential energy, of the system decreases with the passage of time. The kinetic energy is obtained by the quadratic form of v induced by the symmetric positive definite mass matrix $M(q)$. When dealing with impacts, the kinetic energy actually dissipates at each impact. This allows one to state the dissipativity of Lagrangian systems subjected to unilateral constraints and impacts, see [8, Sect. 6.8.2 and 7.2.4].

Inspired by these preliminary results, the basic idea behind the observer design is to realize an interconnection of three passive blocks as shown in Fig. 3. To analyze the passivity of each one of these blocks, we introduce the variables

$$\chi := R(q)\frac{d\tilde{x}}{dt} + \dot{R}(q,v)\tilde{x}\frac{dt}{d\mu}, \quad \nu := \begin{pmatrix} 0_{1\times n} \\ -(\eta - \hat{\eta}) \end{pmatrix}.$$

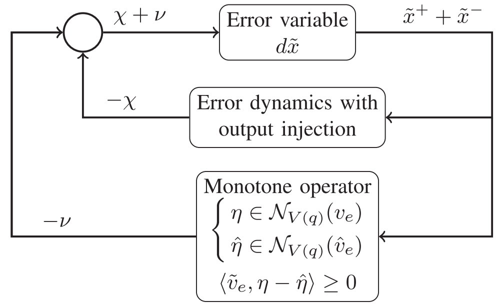


Fig. 3. Interpretation of error dynamics (27) in terms of passivity.

It is then seen that:

- We have a passive interconnection from $\tilde{x}^+ + \tilde{x}^-$ to ν because it follows from calculations in the proof of Theorem 2 that:

$$\langle -\nu, \tilde{x}^+ + \tilde{x}^- \rangle = \langle \eta - \hat{\eta}, \tilde{v}^+ + \tilde{v}^- \rangle \geq 0.$$

- Also, $\chi = R(q)\left(\begin{smallmatrix} v - \widehat{F}_1(t,x,\hat{x}) \\ -F(t,x) + \widehat{F}_2(t,x,\hat{x}) \end{smallmatrix}\right)(dt/d\mu) + \dot{R}(q,v)\tilde{x}(dt/d\mu)$ and the output injection gain is chosen such that the condition (26) holds, so that

$$\begin{aligned}
\langle -\chi, \tilde{x}^+ + \tilde{x}^- \rangle_{L_2} &= -2\int_0^T \chi(s)^\top \tilde{x}(s)ds \\
&\geq \beta \int_0^T \tilde{x}^\top(s) R\left(q(s)\right)\tilde{x}(s)ds \geq c\|\tilde{x}\|_2^2
\end{aligned}$$

  where $c > 0$ is some constant, and $\|\tilde{x}\|_2$ denotes the L_2 norm.
- Lastly, a passive relation from $\chi + \nu$ to $\tilde{x}^+ + \tilde{x}^-$ is observed since

$$\begin{aligned}
\langle \tilde{x}^+ + \tilde{x}^-, \chi + \nu \rangle_{L_2} &= \int_0^T \frac{dW}{d\mu}d\mu \\
&\geq -W\left(\tilde{q}(0), \tilde{v}(0)\right)
\end{aligned}$$

  where we recall that $W(q,\tilde{x}) = \tilde{x}^\top R(q)\tilde{x}$.

VII. Numerical Implementation

We now discuss the numerical implementation of our estimators and the simulation of examples given in the introduction. The basic idea in simulating inclusions of type (13) is to rewrite the system as a combination of differential equation and complementarity relations. For simulation of differential equations, classical numerical techniques may be used whereas the complementarity relations are handled by optimization algorithms which are now commercially available. Further details can be found in [35].

We have applied our results to the three systems described in Fig. 1 given in the introduction. Since the velocity variable is the only quantity of interest that needs to be estimated, we will only implement the observer (38) for these systems. This observer has been implemented in the software platform SICONOS [1]. Due to space constraints, we only present the

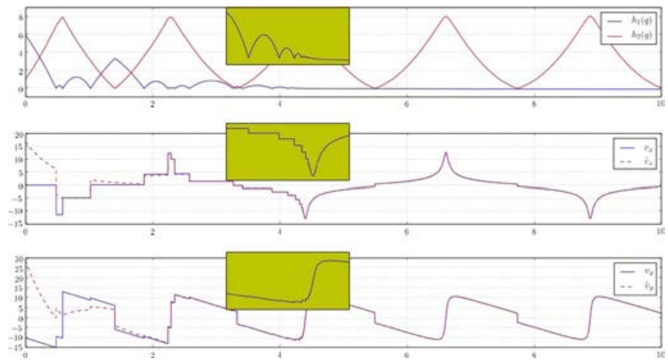


Fig. 4. Velocity estimation error profile for system in Fig. 1(b). The top plot shows the evolution of two gap functions $h_1(q)$ and $h_2(q)$ with time. The middle and bottom plot show the velocity components v_x and v_y, along with their estimates, respectively. The boxes in the center of these plots provides a magnified image of the quantity of interest around the accumulation point $t = 4.5$ s.

simulation of one example in this manuscript, given in Fig. 1(b). For the other two examples given in Fig. 1, and the related animations with several different initial conditions, please refer to the links provided in [35].

Example 1: We consider a particle with $M(q) = I$ bouncing in two-dimensional plane with a parabolic and linear constraint, see Fig. 1(b). We denote the position variable by $\binom{q_x}{q_y}$ and the corresponding velocity vector by $\binom{v_x}{v_y}$. The constraint relations are defined as

$$\begin{aligned} h_1(q) &= q_y - q_x^2 \geq 0 \\ h_2(q) &= c - q_y \geq 0 \end{aligned}$$

where we choose $c = 8$ for the sake of simulations. Using the notation of (8), we choose $F(t, q, v) = 9.81$, that is, the point mass is subjected to the gravitational force, and the coefficient of restitution at impacts is $e = 0.9$. This choice eventually leads to accumulation of impacts in finite time, as one sees from the plots in Fig. 4. Initially, when either of the constraints h_1 or h_2 becomes zero, a jump in at least one of the velocity components is observed. The accumulation of impacts is observed around $t = 4.5$ s, because the mass is being pulled downward continuously by gravity, and the dissipative reaction force (that acts on the particle to maintain the constraint $h_1(q) \geq 0$) reduces the norm of the velocity at each impact. Eventually, after the accumulation point, we see that $h_1(q)$ remains identically zero and sliding of the particle on the reduced-order surface $\{q | h_1(q) = 0\}$ is observed. Our velocity estimator replicates this phenomenon, and after the initial transients, the estimates converge to the actual velocity of the particle, see Fig. 4.

VIII. Calculations for Theorem 1

In this section, we show calculations for the claims made in the proof of Theorem 1.

A. Estimate of a Uniform Bound on $\hat{v}_{\mathcal{P}}$

It is assumed that $|v(t)| \leq C_v$ for all $t \in [0, T]$, so that $q(t) \in \mathcal{B}(q_0, C_v T)$. Let $C_{M^{1/2}}$ be the Lipschitz constant associated with the mapping $q \mapsto M^{1/2}(q)$ on $\mathcal{B}(q_0, C_v T)$. The projection, with respect to the norm induced by $M_{\mathcal{P},i} := M(q(t_{\mathcal{P},i}))$, on the set $V(q(t_{\mathcal{P},i}))$ is denoted by $P_{\mathcal{P},i}$ and on the set $M^{-1}(q(t_{\mathcal{P},i}))V^\circ(q(t_{\mathcal{P},i}))$ by $Q_{\mathcal{P},i}$. We denote by $\overline{\lambda}_M, \underline{\lambda}_M$ the constants introduced in (12) for the compact set $\mathcal{B}(q_0, C_v T)$. Let $u_{\mathcal{P},i}$ be defined as

$$u_{\mathcal{P},i} := \hat{v}_{\mathcal{P},i-1} - \frac{1}{1+e} M_{\mathcal{P},i}^{-1} G_{\mathcal{P},i}. \tag{41}$$

These notations are now used in deriving a bound on $\hat{v}_{\mathcal{P}}$. Using Moreau's two-cone lemma (see Lemma A.1 in Appendix A), we first get

$$\begin{aligned} &|\hat{v}_{\mathcal{P},i}|_{M_{\mathcal{P},i}} \\ &\quad = \left| P_{\mathcal{P},i}(u_{\mathcal{P},i}) - e Q_{\mathcal{P},i}(u_{\mathcal{P},i}) - \frac{e}{1+e} M_{\mathcal{P},i}^{-1} G_{\mathcal{P},i} \right|_{M_{\mathcal{P},i}} \\ &\quad \leq |u_{\mathcal{P},i}|_{M_{\mathcal{P},i}} + \frac{e}{1+e} \left\| M_{\mathcal{P},i}^{-1/2} \right\| \cdot |G_{\mathcal{P},i}| \\ &\quad \leq |\hat{v}_{\mathcal{P},i-1}|_{M_{\mathcal{P},i}} + \frac{1}{\sqrt{\underline{\lambda}_M}} |G_{\mathcal{P},i}| \\ &\quad \leq |\hat{v}_{\mathcal{P},i-1}|_{M_{\mathcal{P},i-1}} + \left\| M_{\mathcal{P},i}^{\frac{1}{2}} - M_{\mathcal{P},i-1}^{\frac{1}{2}} \right\| \cdot |\hat{v}_{\mathcal{P},i-1}| \\ &\qquad + \frac{1}{\sqrt{\underline{\lambda}_M}} |G_{\mathcal{P},i}| \\ &\quad \leq |\hat{v}_{\mathcal{P},i-1}|_{M_{\mathcal{P},i-1}} + C_{M^{\frac{1}{2}}} \int_{t_{\mathcal{P},i-1}}^{t_{\mathcal{P},i}} |v(s)|\, ds \cdot |\hat{v}_{\mathcal{P},i-1}| \\ &\qquad + \frac{1}{\sqrt{\underline{\lambda}_M}} |G_{\mathcal{P},i}|. \end{aligned}$$

Since $\hat{v}_{\mathcal{P}}(\cdot)$ has the constant value $\hat{v}_{\mathcal{P},i}$ on the interval $[t_{\mathcal{P},i}, t_{\mathcal{P},i+1})$, the above inequality results in

$$\begin{aligned} &\sqrt{\underline{\lambda}_M}\, |\hat{v}_{\mathcal{P}}(t_{\mathcal{P},i})| - \sqrt{\overline{\lambda}_M}\, |\hat{v}_{\mathcal{P}}(0)| \\ &\quad \leq \sum_{j=1}^{i} |\hat{v}_{\mathcal{P},i}|_{M_{\mathcal{P},i}} - |\hat{v}_{\mathcal{P},i-1}|_{M_{\mathcal{P},i-1}} \\ &\quad \leq C_{M^{\frac{1}{2}}} C_v \int_0^{t_{\mathcal{P},i}} |\hat{v}_{\mathcal{P}}(s)|\, ds + \frac{C_{g,b}}{\sqrt{\underline{\lambda}_M}} \int_0^{t_{\mathcal{P},i}} (1 + |\hat{v}_{\mathcal{P}}(s)|\, ds) \end{aligned}$$

that is, for each $t \in [0, T]$

$$|\hat{v}_{\mathcal{P}}(t)| \leq \sqrt{\frac{\overline{\lambda}_M}{\underline{\lambda}_M}} |\hat{v}_0| + \frac{C_{g,b}}{\underline{\lambda}_M} t + \left(\frac{C_{M^{\frac{1}{2}}} C_v}{\sqrt{\underline{\lambda}_M}} + \frac{C_{g,b}}{\underline{\lambda}_M} \right) \int_0^t |\hat{v}_{\mathcal{P}}(s)| ds.$$

Applying the Gronwall-Bellman inequality for discontinuous functions [18, Lemma 1] to (24), the following bound on $|\hat{v}_{\mathcal{P}}(t)|$ is obtained for each $t \in [0, T]$, which does not depend on the partition $\mathcal{P}$:

$$|\hat{v}_{\mathcal{P}}(t)| \leq \sqrt{\frac{\overline{\lambda}_M}{\underline{\lambda}_M}} |\hat{v}_0| + \frac{C_{g,b}}{\underline{\lambda}_M} t + A_1 \exp(A_2 t) \leq C_{\sup} \tag{42}$$

where

$$\begin{aligned} A_1 &:= \left(\sqrt{\frac{\overline{\lambda}_M}{\underline{\lambda}_M}} |\hat{v}_0| + \frac{C_{g,b}}{C_{M^{\frac{1}{2}}} C_v \sqrt{\underline{\lambda}_M} + C_{g,b}} \right) \\ A_2 &:= \left(\left(\frac{C_{M^{\frac{1}{2}}} C_v}{\sqrt{\underline{\lambda}_M}} + \frac{C_{g,b}}{\underline{\lambda}_M} \right) t \right) \end{aligned}$$

and $C_{\sup}$ is obtained by evaluating the right-hand side of the first inequality in (42) at $t = T$.

B. Estimates on the Variation

For a fixed partition $\mathcal{P}$ of the interval $[0,T]$, we now compute the total variation of $\hat{v}_{\mathcal{P}}(\cdot)$. For conciseness, we drop the subscript $\mathcal{P}$ in the quantities appearing in (23) and (24). By definition, we have

$$\hat{v}_i = -e\hat{v}_{i-1} + (1+e)P_i(u_i).$$

Using Moreau's two-cone lemma, we can write $u_i := P_i(u_i) + Q_i(u_i)$, so that

$$\begin{aligned}\hat{v}_i - \hat{v}_{i-1} &= -(1+e)u_i + (1+e)P_i(u_i) - M_i^{-1}G_i \\ &= -(1+e)Q_i(u_i) - M_i^{-1}G_i. \end{aligned} \tag{43}$$

Since $Q_i(\cdot)$ denotes the projection on $V_i^* := M_i^{-1}V^\circ(q_i)$ with respect to the kinetic metric, we take $Q_i(u) = 0$ if $\mathcal{J}(q_i) = \emptyset$, in which case

$$|\hat{v}_i - \hat{v}_{i-1}| = \left|M_i^{-1}G_i\right| \le \frac{|G_i|}{\underline{\lambda}_M}.$$

Otherwise, if $\mathcal{J}(q_i) \neq \emptyset$, we have

$$Q_i(u) := \sum_{\alpha\in\mathcal{J}(q_i)} \langle u, \nabla h_\alpha(q_i)\rangle^- M_i^{-1}H_\alpha(q_i)$$

where

$$H_\alpha(q_i) := \frac{\nabla h_\alpha(q_i)}{\nabla h_\alpha^\top(q_i)M_i^{-1}\nabla h_\alpha(q_i)}$$

and $\langle u, \nabla h_\alpha(q)\rangle^- := \min\{\langle u, \nabla h_\alpha(q)\rangle, 0\}$. One may rewrite (43) as

$$\begin{aligned}\hat{v}_i - \hat{v}_{i-1} = &-(1+e)\left(Q_i(u_i) - Q_i(\hat{v}_{i-1})\right) \\ &-(1+e)\left(Q_i(\hat{v}_{i-1}) - \tilde{Q}_{i-1}(\hat{v}_{i-1})\right) \\ &-(1+e)\tilde{Q}_{i-1}(\hat{v}_{i-1}) - M_i^{-1}G_i \end{aligned} \tag{44}$$

where

$$\tilde{Q}_{i-1}(u) := \sum_{\alpha\in\mathcal{J}(q_i)} \langle u, \nabla h_\alpha(q_{i-1})\rangle^- M_{i-1}^{-1}H_\alpha(q_{i-1}).$$

We now compute an upper bound on the norm of the right-hand side of (44).

First term: It is noted using the contraction property of the projection map and the bound derived in (42) that

$$\begin{aligned}|Q_i(u_i) - Q_i(\hat{v}_{i-1})| &\le \frac{1}{\sqrt{\underline{\lambda}_M}}|u_i - \hat{v}_{i-1}|_i \\ &\le \frac{1}{(1+e)\sqrt{\underline{\lambda}_M}}\left|M_i^{-1}G_i\right|_i \le \frac{1}{1+e}\frac{|G_i|}{\underline{\lambda}_M}\end{aligned} \tag{45}$$

where $|\cdot|_i$ is used as a short-hand for $|\cdot|_{M_{\mathcal{P},i}}$ when the partition $\mathcal{P}$ is considered to be fixed.

Second term: Under the hypothesis that $\nabla h_\alpha(\cdot)$ is locally Lipschitz continuous, for each $\alpha = 1,\dots,m$, and that $q(\cdot)$ evolves within a compact set over the interval $[0,T]$, there exists a constant C_h such that

$$|\nabla h_\alpha(q_i) - \nabla h_\alpha(q_{i-1})| \le C_h|q_i - q_{i-1}| \le C_hC_v|t_i - t_{i-1}| \tag{46}$$

for all $t_i, t_{i-1} \in [0,T]$. Similarly, since $M^{-1/2}(\cdot)$ is locally Lipschitz continuous, there exists $C_H > 0$ such that

$$|H_\alpha(q_i) - H_\alpha(q_{i-1})| \le C_H|q_i - q_{i-1}| \le C_HC_v(t_i - t_{i-1}) \tag{47}$$

for all $t_i, t_{i-1} \in [0,T]$. Thus, we get

$$\begin{aligned}&Q_i(\hat{v}_{i-1}) - \tilde{Q}_{i-1}(\hat{v}_{i-1}) \\ &\quad= \sum_{\alpha\in\mathcal{J}(q_i)} \langle \hat{v}_{i-1}, \nabla h_\alpha(q_i)\rangle^- M_i^{-1}H_\alpha(q_i) \\ &\qquad- \sum_{\alpha\in\mathcal{J}(q_i)} \langle \hat{v}_{i-1}, \nabla h_\alpha(q_{i-1})\rangle^- M_{i-1}^{-1}H_\alpha(q_{i-1}) \\ &\quad= \sum_{\alpha\in\mathcal{J}(q_i)} \left(\langle \hat{v}_{i-1}, \nabla h_\alpha(q_i)\rangle^- - \langle \hat{v}_{i-1}, \nabla h_\alpha(q_{i-1})\rangle^-\right) \\ &\qquad\times M_i^{-1}H_\alpha(q_i) \\ &\qquad+ \sum_{\alpha\in\mathcal{J}(q_i)} \langle \hat{v}_{i-1}, \nabla h_\alpha(q_{i-1})\rangle^- \\ &\qquad\times \left[\left(M_i^{-1} - M_{i-1}^{-1}\right)H_\alpha(q_i) \right.\\ &\qquad\qquad \left.+M_{i-1}^{-1}\left(H_\alpha(q_i) - H_\alpha(q_{i-1})\right)\right].\end{aligned}$$

This further leads to

$$\begin{aligned}&\left|Q_i(\hat{v}_{i-1}) - \tilde{Q}_{i-1}(\hat{v}_{i-1})\right| \\ &\quad\le \sum_{\alpha\in\mathcal{J}(q_i)} \left|\langle \hat{v}_{i-1}, \nabla h_\alpha(q_i)\rangle^- - \langle \hat{v}_{i-1}, \nabla h_\alpha(q_{i-1})\rangle^-\right| \\ &\qquad\times \left|M_i^{-1}H_\alpha(q_i)\right| \\ &\qquad+ \sum_{\alpha\in\mathcal{J}(q_i)} |\langle \hat{v}_{i-1}, \nabla h_\alpha(q_{i-1})\rangle| \\ &\qquad\times \left[\left\|M_i^{-1} - M_{i-1}^{-1}\right\| \cdot |H_\alpha(q_i)| \right.\\ &\qquad\qquad\left.+ \left\|M_{i-1}^{-1}\right\| \cdot |H_\alpha(q_i) - H_\alpha(q_{i-1})|\right] \\ &\quad\le \frac{m}{\underline{\lambda}_M} C_{\sup}C_hC_v\tilde{C}_{H_\alpha}(t_i - t_{i-1}) \\ &\qquad+ mC_{\sup}\tilde{C}_{h_\alpha}\left(C_{M^{-1}}\tilde{C}_{H_\alpha} + C_H\right)C_v(t_i - t_{i-1}) \\ &\quad=: C_{\text{proj}}(t_i - t_{i-1})\end{aligned} \tag{48}$$

where $\tilde{C}_{h_\alpha} = \sup_{q\in\mathcal{B}(q_0,C_vT)}|\nabla h_\alpha(q)|$ and $\tilde{C}_{H_\alpha} := \sqrt{(\overline{\lambda}_M/\underline{\lambda}_M)}\,C_{h_\alpha} \ge \sup_{q\in\mathcal{B}(q_0,C_vT)}|H_\alpha|$. The constants C_v, $C_{\sup}$, C_h, and C_H were introduced in (34), (42), (46), and (47), respectively. The constant $C_{M^{-1}}$ is chosen such that $\|M_i - M_{i-1}\| \le C_{M^{-1}}$, for each $i \in \mathbb{N}$.

Third term: We have

$$\begin{aligned}\hat{v}_i = -e\hat{v}_{i-1} + (1+e)P_i(u_i) = P_i(u_i) \\ - eQ_i(u_i) - \frac{e}{1+e}M_i^{-1}G_i\end{aligned}$$

which gives

$$\hat{v}_i + \frac{e}{1+e}M_i^{-1}G_i = P_i(u_i) - eQ_i(u_i) \in V(q_i). \tag{49}$$

This further leads to

$$\left|\tilde{Q}_{i-1}(\hat{v}_{i-1})\right| \leq \sum_{\alpha\in\mathcal{J}(q_i)} \left|M_{i-1}^{-1}H_\alpha(q_{i-1})\right|$$
$$\times\left|\left\langle \hat{v}_{i-1}+\frac{e}{1+e}M_{i-1}^{-1}G_{i-1}-\frac{e}{1+e}M_{i-1}^{-1}G_{i-1}, \nabla h_\alpha(q_{i-1})\right\rangle^-\right|.$$

From (49), we have

$$\left\langle \hat{v}_{i-1}+\frac{e}{1+e}M_{i-1}^{-1}G_{i-1}, \nabla h_\alpha(q_{i-1})\right\rangle \geq 0, \quad \forall\, j=1,\ldots,m$$

and hence[5]

$$\begin{aligned}\left|\tilde{Q}_{i-1}(\hat{v}_{i-1})\right| &\leq \sum_{\alpha\in\mathcal{J}(q_i)} \frac{e}{1+e}\left|\langle M_{i-1}^{-1}G_{i-1}, \nabla h_\alpha(q_{i-1})\rangle\right| \\ &\quad\times \frac{\left|M_{i-1}^{-1}\nabla h_\alpha(q_{i-1})\right|_{i-1}}{\sqrt{\underline{\lambda}_M}\nabla h_\alpha^\top(q_{i-1})M_{i-1}^{-1}\nabla h_\alpha(q_{i-1})} \\ &\leq \frac{e}{1+e}\sum_{\alpha\in\mathcal{J}(q_i)}\left|M_{i-1}^{-1}G_{i-1}\right|_{i-1} \\ &\quad\times \frac{\left|M_{i-1}^{-1}\nabla h_\alpha(q_{i-1})\right|_{i-1}}{\sqrt{\underline{\lambda}_M}\left|M_{i-1}^{-1}\nabla h_\alpha(q_{i-1})\right|_{i-1}} \\ &\leq \frac{me}{1+e}\frac{|G_{i-1}|}{\underline{\lambda}_M}. \end{aligned} \tag{50}$$

Plugging the bounds from (45), (48), and (50) into (44), we obtain

$$|\hat{v}_i-\hat{v}_{i-1}| \leq \frac{1}{\underline{\lambda}_M}|G_i| + 2C_{\text{proj}}(t_i-t_{i-1}) + \frac{1}{\underline{\lambda}_M}|G_i| + \frac{m}{\underline{\lambda}_M}|G_{i-1}|.$$

Using the norm estimate on $\hat{v}_{\mathcal{P}}(\cdot)$, we have $|G_i| \leq (1+C_{\text{sup}})(t_i-t_{i-1})$ and thus for $0\leq s<t\leq T$, it follows that:

$$\operatorname{Var}(\hat{v}_{\mathcal{P}};[s,\,t]) \leq C_{\text{var}}(t-s) \tag{51}$$

where $C_{\text{var}} := (1/\underline{\lambda}_M)((m+2)(1+C_{\text{sup}})+2C_{\text{proj}})$.

C. Continuity Points of the Limit Function

Assume that $y\in V(q(\tau))$ for all $\tau\in[s,t]\subseteq[0,T]$, then it is claimed that

$$\begin{aligned}&\int_s^t \langle g(\sigma,\hat{v}),(y-\hat{v})\rangle + \left\langle \dot{M}(\sigma)\hat{v}, \left(y-\frac{\hat{v}}{2}\right)\right\rangle d\sigma \\ &\quad\leq \langle M(q(t))\,\hat{v}(t) - M(q(s))\,\hat{v}(s), y\rangle \\ &\quad\quad - \frac{1}{2}\left(|\hat{v}(t)|^2_{M(q(t))} - |\hat{v}(s)|^2_{M(q(s))}\right).\end{aligned} \tag{52}$$

In the sequel, we proceed to prove this claim:

Consider a partition $\mathcal{P}$ of the interval $[0,T]$ that contains the nodes $t_{\mathcal{P},j}=s$ and $t_{\mathcal{P},k}=t$ for some $j,k\in\mathbb{N}$. From the discretization scheme (23), for $j+1\leq i\leq k$, we have

$$\frac{\hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}}{1+e} = P_{\mathcal{P},i}\left(\hat{v}_{\mathcal{P},i-1} - \frac{1}{1+e}M_{\mathcal{P},i}^{-1}G_{\mathcal{P},i}\right).$$

[5]We use the fact that for $a,b,c\in\mathbb{R}^n$, satisfying $\langle a,c\rangle\geq 0$, we have $|\langle a+b,c\rangle^-| = |\min\{0,\langle a+b,c\rangle\}| \leq |\min\{0,\langle a,c\rangle\}| + |\min\{0,\langle b,c\rangle\}| = |\langle b,c\rangle|$.

Using the definition of the projection operator, we get, $\forall\, y\in V_{\mathcal{P},i}$,

$$\left\langle \hat{v}_{\mathcal{P},i-1} - \hat{v}_{\mathcal{P},i} - M_{\mathcal{P},i}^{-1}G_{\mathcal{P},i}, (1+e)y - (\hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1})\right\rangle_{M_{\mathcal{P},i}} \leq 0.$$

The above inequality is equivalently written as

$$\begin{aligned}&\left\langle -M_{\mathcal{P},i}^{-1}G_{\mathcal{P},i},(1+e)y\right\rangle_{M_{\mathcal{P},i}} + \left\langle M_{\mathcal{P},i}^{-1}G_{\mathcal{P},i}, \hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}\right\rangle_{M_{\mathcal{P},i}} \\ &\quad\leq (1+e)\langle \hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}, y\rangle_{M_{\mathcal{P},i}} \\ &\quad\quad - \langle \hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}, \hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}\rangle_{M_{\mathcal{P},i}} \qquad (53a) \\ &\quad\leq (1+e)\langle M_{\mathcal{P},i}\hat{v}_{\mathcal{P},i} - M_{\mathcal{P},i-1}\hat{v}_{\mathcal{P},i-1}, y\rangle \qquad (53b) \\ &\quad\quad - \frac{1+e}{2}\left(|\hat{v}_{\mathcal{P},i}|^2_{M_{\mathcal{P},i}} - |\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i-1}}\right) \\ &\quad\quad - (1+e)\left\langle (M_{\mathcal{P},i}-M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}, y - \frac{\hat{v}_{\mathcal{P},i-1}}{2}\right\rangle. \qquad (53c)\end{aligned}$$

To arrive at (53c), the last term in (53a) is rewritten as

$$\begin{aligned}&\langle \hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}, \hat{v}_{\mathcal{P},i}\rangle_{M_{\mathcal{P},i}} \\ &\quad= \frac{1}{2}\Big[|\hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i}} + |\hat{v}_{\mathcal{P},i}|^2_{M_{\mathcal{P},i}} - |\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i-1}} \\ &\quad\quad - \hat{v}_{\mathcal{P},i-1}^\top(M_{\mathcal{P},i}-M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}\Big] \\ &\langle \hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}, e\hat{v}_{\mathcal{P},i-1}\rangle_{M_{\mathcal{P},i}} \\ &\quad= \frac{e}{2}\Big[-|\hat{v}_{\mathcal{P},i}-\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i}} + |\hat{v}_{\mathcal{P},i}|^2_{M_{\mathcal{P},i}} - |\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i-1}} \\ &\quad\quad - \hat{v}_{\mathcal{P},i-1}^\top(M_{\mathcal{P},i}-M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}\Big].\end{aligned} \tag{54}$$

Inequality (53c) now leads to

$$\begin{aligned}&\left\langle -G_{\mathcal{P},i}, y - \frac{\hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}}{1+e}\right\rangle \\ &\quad + \left\langle (M_{\mathcal{P},i}-M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}, y - \frac{\hat{v}_{\mathcal{P},i-1}}{2}\right\rangle \\ &\quad\leq \langle M_{\mathcal{P},i}\hat{v}_{\mathcal{P},i} - M_{\mathcal{P},i-1}\hat{v}_{\mathcal{P},i-1}, y\rangle \\ &\quad\quad - \frac{1}{2}\left(|\hat{v}_{\mathcal{P},i}|^2_{M_{\mathcal{P},i}} - |\hat{v}_{\mathcal{P},i-1}|^2_{M_{\mathcal{P},i-1}}\right)\end{aligned}$$

which further yields

$$\begin{aligned}&\sum_{i=j+1}^{k}\left\langle -G_{\mathcal{P},i}, y - \frac{\hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}}{1+e}\right\rangle \\ &\quad + \left\langle (M_{\mathcal{P},i}-M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}, y - \frac{\hat{v}_{\mathcal{P},i-1}}{2}\right\rangle \\ &\quad\leq \langle M_{\mathcal{P},k}\hat{v}_{\mathcal{P},k} - M_{\mathcal{P},j}\hat{v}_{\mathcal{P},j}, y\rangle \\ &\quad\quad - \frac{1}{2}\left(|\hat{v}_{\mathcal{P},k}|^2_{M_{\mathcal{P},k}} - |\hat{v}_{\mathcal{P},j}|^2_{M_{\mathcal{P},j}}\right).\end{aligned} \tag{55}$$

Using the fact that $((\hat{v}_{\mathcal{P}}(\tau)+e\hat{v}_{\mathcal{P}}(\tau))/(1+e))$ converges to $\hat{v}(\tau)$ for Lebesgue almost all $\tau\in[0,T]$, it follows that:

$$\sum_{i=j+1}^{k}\left\langle G_{\mathcal{P},i}, y - \frac{\hat{v}_{\mathcal{P},i}+e\hat{v}_{\mathcal{P},i-1}}{1+e}\right\rangle \to \int_s^t \langle g(\tau,\hat{v}(\tau)), y-\hat{v}(\tau)\rangle\, d\tau. \tag{56}$$

Since $t \mapsto M(q(t))$ is an absolutely continuous function, $\dot{M}(\tau) := (d/dt)M(q(t))|_{t=\tau}$ exists for Lebesgue almost-all τ, and we have

$$\left\langle (M_{\mathcal{P},i} - M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}, y - \frac{\hat{v}_{\mathcal{P},i-1}}{2} \right\rangle = \int_{t_{\mathcal{P},i-1}}^{t_{\mathcal{P},i}} \left\langle \dot{M}(\tau)\hat{v}_{\mathcal{P}}(\tau), y - \frac{\hat{v}_{\mathcal{P}}(\tau)}{2} \right\rangle d\tau$$

$$\sum_{i=j+1}^{k} \left\langle (M_{\mathcal{P},i} - M_{\mathcal{P},i-1})\hat{v}_{\mathcal{P},i-1}, y - \frac{\hat{v}_{\mathcal{P},i-1}}{2} \right\rangle = \int_{t_{\mathcal{P},j}}^{t_{\mathcal{P},k}} \left\langle \dot{M}(\tau)\hat{v}_{\mathcal{P}}(\tau), y - \frac{\hat{v}_{\mathcal{P}}(\tau)}{2} \right\rangle d\tau. \tag{57}$$

For the terms on the right-hand side, we have the following convergence:

$$\begin{aligned} &\langle M_{\mathcal{P},k}\hat{v}_{\mathcal{P},k} - M_{\mathcal{P},j}\hat{v}_{\mathcal{P},j}, y\rangle - \frac{1}{2}\left(|\hat{v}_{\mathcal{P},k}|^2_{M_{\mathcal{P},k}} - |\hat{v}_{\mathcal{P},j}|^2_{M_{\mathcal{P},j}}\right) \\ &\quad \to \langle M\left(q(t)\right)\hat{v}(t) - M(q(s))\hat{v}(s), y\rangle \\ &\quad - \frac{1}{2}\left(|\hat{v}(t)|^2_{M(q(t))} - |\hat{v}(s)|^2_{M(q(s))}\right). \end{aligned} \tag{58}$$

The desired inequality (52) now follows by taking the limit in (55) along all partitions finer than $\mathcal{P}$ and using (56)–(58).

Let μ be the measure defined by $d\mu = |d\hat{v}| + dt$. Since $d\hat{v}$ and dt are absolutely continuous with respect to $d\mu$ there exists a $d\mu$ negligible set A such that, for all $t \in [0,T] \setminus A$:

$$\frac{dt}{d\mu}(t) = \lim_{\varepsilon\to 0^+} \frac{dt([t,t+\varepsilon])}{d\mu\left([t,t+\varepsilon]\right)} \text{ and } \frac{d\hat{v}}{d\mu}(t) = \lim_{\varepsilon\to 0^+} \frac{d\hat{v}\left([t,t+\varepsilon]\right)}{d\mu\left([t,t+\varepsilon]\right)}.$$

Assume that $\hat{v}$ is continuous at t and let $y \in \text{int } V(q(t))$. Then due to lower semicontinuity of $q \mapsto V(q)$ and absolute continuity of $t \mapsto q(t)$, $y \in V(q(\tau))$ for all $\tau \in I_\varepsilon := [t, t+\varepsilon]$. Due to the variational inequality (52), we get

$$\begin{aligned} &\int_t^{t+\varepsilon} \langle -g(\sigma,\hat{v}), (y-\hat{v})\rangle + \left\langle \dot{M}(\sigma)\hat{v}(\sigma), \left(y - \frac{\hat{v}}{2}\right)\right\rangle d\sigma \\ &\quad \le \langle M\left(q(t+\varepsilon)\right)\hat{v}(t+\varepsilon) - M\left(q(t)\right)\hat{v}(t), y\rangle \\ &\quad - \frac{1}{2}\left(|\hat{v}(t+\varepsilon)|^2_{M(q(t+\varepsilon))} - |\hat{v}(t)|^2_{M(q(t))}\right). \end{aligned} \tag{59}$$

Divide both sides by $d\mu([t,t+\varepsilon])$. When $\varepsilon \to 0$, the left-hand side of (59) converges to

$$\frac{dt}{d\mu}\left[\langle -g\left(t,\hat{v}(t)\right), y - \hat{v}(t)\rangle + \left\langle \dot{M}(t)\hat{v}(t), y - \frac{1}{2}\hat{v}(t)\right\rangle\right].$$

The first term on the right-hand side of (59) becomes

$$\begin{aligned} &\frac{1}{d\mu(I_\varepsilon)}\langle M\left(q(t+\varepsilon)\right)\hat{v}(t+\varepsilon) - M\left(q(t)\right)\hat{v}(t), y\rangle \\ &= \frac{\langle M\left(q(t)\right)d\hat{v}(I_\varepsilon), y\rangle}{d\mu(I_\epsilon)} \\ &\quad + \frac{1}{d\mu(I_\epsilon)}\left\langle \int_t^{t+\varepsilon} \dot{M}(s)ds\ \hat{v}(t+\varepsilon), y\right\rangle \\ &\xrightarrow{\varepsilon\to 0} \left\langle M\left(q(t)\right)\frac{d\hat{v}}{d\mu}(t), y\right\rangle + \left\langle \dot{M}(t)\hat{v}(t), y\right\rangle\frac{dt}{d\mu} \end{aligned}$$

and the second term on the right-hand side of (59) becomes

$$\begin{aligned} &\frac{1}{2d\mu(I_\varepsilon)}\Bigg[\langle M\left(q(t+\varepsilon)\right)\hat{v}(t+\varepsilon), d\hat{v}(I_\varepsilon)\rangle \\ &\quad + \langle M\left(q(t)\right)d\hat{v}(I_\varepsilon), \hat{v}(t)\rangle \\ &\quad + \left\langle \int_t^{t+\varepsilon}\dot{M}(s)ds\ \hat{v}(t+\varepsilon), \hat{v}(t)\right\rangle\Bigg] \\ &\xrightarrow{\varepsilon\to 0} \left\langle M\left(q(t)\right)\frac{d\hat{v}}{d\mu}(t), \hat{v}(t)\right\rangle + \left\langle \dot{M}(t)\hat{v}(t), \frac{\hat{v}(t)}{2}\right\rangle\frac{dt}{d\mu}. \end{aligned}$$

Thus, in the limit, (59) leads to, $\forall\ y \in \text{int } V(q(t))$

$$\left\langle M\left(q(t)\right)\frac{d\hat{v}}{d\mu}(t) + g\left(t,\hat{v}(t)\right)\frac{dt}{d\mu}(t), y - \hat{v}(t)\right\rangle \le 0.$$

From the definition of $\mathcal{N}_{V(q(t))}(\cdot)$, and using the density argument, it follows that $\hat{v}(\cdot)$ satisfies the differential inclusion (17) at the continuity points of $\hat{v}(\cdot)$.

D. Impact Characterization of the Limit Solution

Let $t_k \in [0,T]$ be the time instant at which $V(q(t_k)) \neq \mathbb{R}^n$. If $\hat{v}(t_k^-) \in V(q(t_k))$, then $\hat{v}(t_k^+) = \hat{v}(t_k^-)$. This is a straightforward consequence of the variational inequality (52). Indeed, let $y \in \text{int } V(q(t_k))$ then $y \in V(q(t))$ for all $t \in [t_k - \delta, t_k + \delta]$ and some $\delta > 0$. Applying the inequality (52) with $s = t_k - \delta$, $t = t_k + \delta$, and letting $\delta \to 0$, we get

$$\begin{aligned} &\langle M\left(q(t_k)\right)\left(\hat{v}\left(t_k^+\right) - \hat{v}\left(t_k^-\right)\right), y\rangle \\ &\quad - \frac{1}{2}\left(\left|\hat{v}\left(t_k^+\right)\right|_{M(q(t_k))} - \left|\hat{v}\left(t_k^-\right)\right|_{M(q(t_k))}\right) \ge 0. \end{aligned}$$

By density, the same inequality holds for all $y \in \text{int } V(q(t_k))$. Picking $y = \hat{v}(t_k^-)$ gives

$$\left|\hat{v}\left(t_k^+\right) - \hat{v}\left(t_k^-\right)\right|^2 \le 0 \quad \Rightarrow \quad \hat{v}\left(t_k^+\right) = \hat{v}\left(t_k^-\right).$$

Next, consider the case where $\hat{v}(t_k^-) \notin V(q(t_k))$. We use the shorthand notation $\hat{v}^+ := \hat{v}(t_k^+)$ and $\hat{v}^- := \hat{v}(t_k^-)$, and show that the following impact law holds:

$$\hat{v}^+ = -e\hat{v}^- + (1+e)\,\text{proj}_{M(q(t_k))}\left(\hat{v}^-, V\left(q(t_k)\right)\right).$$

Define $\tilde{u}_k := -e\hat{v}(t_k^-) + (1+e)\text{proj}_{M(q(t_k))}(\hat{v}(t_k^-), V(q(t_k)))$.

Consider the partition $\mathcal{P}$ that contains the node $t_{\mathcal{P},k_d} = t_k$ for some $k_d \in \mathbb{R}$. By definition

$$\begin{aligned}
&|\hat{v}_{\mathcal{P}}(t_k) - \tilde{u}| \\
&\quad = e\left|\hat{v}_{\mathcal{P},k_d-1} - v^-\right| + (1+e) \\
&\qquad \times \left|\mathrm{proj}_{M_{\mathcal{P},k_d}}\left[\hat{v}_{\mathcal{P},k_d-1} - \frac{1}{1+e}M^{-1}_{\mathcal{P},k_d}G_{\mathcal{P},k_d}; V\left(q(t_k)\right)\right]\right. \\
&\qquad\qquad \left. - \mathrm{proj}_{M_{\mathcal{P},k_d}}\left[\hat{v}^-; V\left(q(t_k)\right)\right]\right| \\
&\quad \le e\left|\hat{v}_{\mathcal{P},k_d-1} - v^-\right| \\
&\qquad + (1+e)\sqrt{\overline{\lambda}_M}\left|\hat{v}_{\mathcal{P},k_d-1} - \frac{1}{1+e}M^{-1}_{\mathcal{P},k_d}G_{\mathcal{P},k_d} - v^-\right| \\
&\quad \le \left(\sqrt{\overline{\lambda}_M} + e\right)\left|\hat{v}_{\mathcal{P},k_d-1} - \hat{v}^-\right| + \sqrt{\frac{\overline{\lambda}_M}{\underline{\lambda}_M}}\left|G_{\mathcal{P},k_d}\right|. \quad (60)
\end{aligned}$$

The pointwise convergence (25) implies the existence of some filter $\mathcal{F}$ such that, for all $\mathcal{P} \in \mathcal{F}$, we have $t', t_k \in \mathcal{P}$ where t' is such that $|\hat{v}(t') - \hat{v}(t_k^-)| < (\varepsilon/3)$, $|\hat{v}_{\mathcal{P}}(t') - \hat{v}(t')| < (\varepsilon/3)$ and $t_k - t' < (\varepsilon/3C_{\mathrm{var}})$; then

$$\begin{aligned}
&\left|\hat{v}_{\mathcal{P},k_d-1} - \hat{v}\left(t_k^-\right)\right| \\
&\quad \le |\hat{v}_{\mathcal{P},k_d-1} - \hat{v}_{\mathcal{P}}(t')| + |\hat{v}_{\mathcal{P}}(t') - \hat{v}(t')| + \left|\hat{v}(t') - \hat{v}\left(t_k^-\right)\right| \\
&\quad \le \mathrm{Var}\left(\hat{v}_{\mathcal{P}}; [t', t_k)\right) + \frac{2}{3}\varepsilon \le C_{\mathrm{var}}(t_k - t') + \frac{2}{3}\varepsilon < \varepsilon. \quad (61)
\end{aligned}$$

Substituting (61) in (60), and taking the limit, we obtain

$$|\hat{v}(t_k) - \tilde{u}_k| < \varepsilon$$

for every $\varepsilon > 0$, whence the desired result follows.

IX. Conclusion

The problem of designing asymptotically convergent state estimators for nonsmooth mechanical systems with frictionless unilateral constraints and impacts is considered in this paper. As a solution, we propose a class of estimators described by differential inclusions. The existence and uniqueness of solutions for these estimators is rigorously established. The error analysis (for the convergence of velocity estimate) is based on generalizing the Lyapunov techniques to functions of locally bounded variation, which also allow for accumulations of impacts (Zeno phenomenon). Also, under the umbrella of these general estimators, we design a full-order observer that constructs estimates of position and velocity variables, and a reduced-order observer for estimation of the velocity variable only.

Appendix

Lemma A.1 (Moreau's Two-Cone Lemma): If V and V° denote a pair of mutually polar closed convex cones of a Euclidean linear space $\mathbb{R}^n$, then the following statements are equivalent for $x, y, z \in \mathbb{R}^n$:

- $x = \mathrm{proj}(z; V)$ and $y = \mathrm{proj}(z, V^\circ)$
- $z = x + y$, $x \in V$, $y \in V^\circ$, and $\langle x, y\rangle = 0$.

Theorem A.2 (Generalization of Helly's First Theorem [19, Theorem 0.2.2]): Let (u_α) be a generalized sequence or net of functions of bounded variation from the interval $[0, T]$ to a Hilbert space H. Assume that the norm and the variation of u_α are uniformly bounded, that is, there exist $C_{\max}$ and C_{var} such that

$$\|u_\alpha\| \le C_{\max} \quad \text{and} \quad \mathrm{Var}\left(u_\alpha; [0, T]\right) \le C_{\mathrm{var}}$$

then there is a filter $\mathcal{F}$ finer than the filter of the sections of the index set (that is, there exists a subnet extracted from the given net) and there exists a function of bounded variation $u : [0, T] \to H$ that satisfies

$$\operatorname*{weak-\lim}_{\mathcal{F}} u_\alpha = u \quad \text{and} \quad \mathrm{Var}\left(u; [0, T]\right) \le C_{\mathrm{var}}.$$

In the foregoing result if the Hilbert space H into consideration is finite dimensional then the convergence is uniform, that is

$$\lim \|u_\alpha - u\| = 0.$$

References

[1] V. Acary and F. Pérignon, "An Introduction to SICONOS," INRIA, Rhône-Alpes, Tech. Rep. RR-0340, 2007. [Online]. Available: http://hal.inria.fr/inria-00162911, Project website: [Online]. Available: http://siconos.gforge.inria.fr

[2] V. Acary and B. Brogliato, *Numerical Methods for Nonsmooth Dynamical Systems*, ser. Lecture Notes in Applied and Computational Mechanics. Heidelberg, Germany: Springer-Verlag, 2008, vol. 35.

[3] P. Ballard, "The dynamics of discrete mechanical systems with perfect unilateral constraints," *Arch. Rational Mech. Anal.*, vol. 154, pp. 199–274, 2000.

[4] P. Ballard, "Formulation and well-posedness of the dynamics of rigid-body systems with perfect unilateral constraints," *Phil. Trans. R. Soc. Lond. A: Math., Phys. & Eng. Sci.*, vol. 359, no. 1789, pp. 2327–2346, 2001.

[5] J. J. Benjamin Biemond, N. van de Wouw, W. P. M. H. Heemels, and H. Nijmeijer, "Tracking control for hybrid systems with state-triggered jumps," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 876–890, Apr. 2013.

[6] H. Berghuis and H. Nijmeijer, "A passivity approach to controller-observer design for robots," *IEEE Trans. Robot. Autom.*, vol. 9, no. 6, pp. 740–754, 1993.

[7] G. Besancon, "Global output feedback tracking control for a class of Lagrangian systems," *Automatica*, vol. 38, pp. 1915–1921, 2002.

[8] B. Brogliato, *Nonsmooth Mechanics*, 2nd ed. London, U.K.: Springer-Verlag, 1999.

[9] B. Brogliato, "Absolute stability and the Lagrange-Dirichlet theorem with monotone multivalued mappings," *Syst. Control Lett.*, vol. 51, no. 5, pp. 343–353, 2004.

[10] B. Brogliato and W. P. M. H. Heemels, "Observer design for Lur'e systems with multivalued mappings: A passivity approach," *IEEE Trans. Autom. Control*, vol. 54, no. 8, pp. 1996–2001, Aug. 2009.

[11] B. Brogliato, R. Lozano, B. Maschke, and O. Egeland, *Dissipative Systems Analysis and Control*, 2nd ed. London, U.K.: Springer-Verlag, 2007.

[12] C. Castaing and M. D. P. Monteiro Marques, "BV periodic solutions of an evolution problem associated with continuous moving convex sets," *Set-Valued Anal.*, vol. 3, pp. 381–399, 1995.

[13] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Philadelphia, PA, USA: SIAM Classics in Applied Mathematics, 2009.

[14] R. Dzonou and M. D. P. Monteiro Marques, "A sweeping process approach to inelastic contact problems with general inertia operators," *Eur. J. Mechan., A/Solids*, vol. 26, pp. 474–490, 2007.

[15] R. Dzonou, M. D. P. Monteiro Marques, and L. Paoli, "A convergence result for a vibro-impact problem with a general inertia operator," *Nonlin. Dynam.*, vol. 58, pp. 361–384, 2009.

[16] F. Forni, A. Teel, and L. Zaccarian, "Follow the bouncing ball: Global results on tracking and state estimation with impacts," *IEEE Trans. Autom. Control*, vol. 58, no. 6, pp. 1470–1485, Jun. 2013.

[17] W. P. M. H. Heemels, M. K. Camlibel, J. M. Schumacher, and B. Brogliato, "Observer-based control of linear complementarity systems," *Int. J. Robust & Nonlin. Control*, vol. 21, no. 10, pp. 1193–1218, 2011.

[18] G. S. Jones, "Fundamental inequalities for discrete and discontinuous functional equations," *J. Soc. Indust. Appl. Math.*, vol. 12, no. 1, pp. 43–57, 1964.

[19] M. Kunze and M. D. P. Monteiro Marques, "An introduction to Moreau's sweeping process," in *Impacts in Mechanical Systems: Analysis and Modelling* B. Brogliato, Ed. Berlin, Germany: Springer-Verlag, 2000, vol. 551, Lecture Notes in Physics, pp. 1–60.

[20] M. Mabrouk, "A unified variational model for the dynamics of perfect unilateral constraints," *Eur. J. Mechan.—A/Solids*, vol. 17, no. 5, pp. 819–842, 1998.

[21] L. Menini and A. Tornambé, "Velocity observers for non-linear mechanical systems subject to non-smooth impacts," *Automatica*, vol. 38, pp. 2169–2175, 2002.

[22] M. D. P. Monteiro Marques, *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*. Basel, Switzerland: Birkhäuser, 1993, vol. 9, Progress in Nonlinear Differential Equations and their Applicationsm.

[23] C. I. Morarescu and B. Brogliato, "Trajectory tracking control of multiconstraint complementarity Lagrangian systems," *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1300–1313, Jun. 2010.

[24] J. J. Moreau, "Liaisons unilatérales sans frottement et chocs inélastiques," *C. R. Acad. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre*, vol. 296, no. 19, pp. 1473–1476, 1983.

[25] J. J. Moreau, "Unilateral contact and dry friction in finite freedom dynamics," in *Nonsmooth Mechan. and Applic.* J. J. Moreau, and P. D. Panagiotopoulos, Eds. Berlin, Germany: Springer, 1988, vol. 302, CISM Courses and Lectures, International Centre for Mechanical Sciences, pp. 1–82.

[26] J. J. Moreau and M. Valadier, "A chain rule involving vector functions of bounded variation," *J. Funct. Anal.*, vol. 74, no. 2, pp. 333–345, 1987.

[27] S. Nicosia and P. Tomei, "Robot control by using only joint position measurements," *IEEE Trans. Autom. Control*, vol. 35, no. 9, pp. 1058–1061, Sep. 1990.

[28] R. Ortega and M. W. Spong, "Adaptive motion control of rigid robots: A tutorial," *Automatica*, vol. 25, no. 6, pp. 877–888, 1989.

[29] J. A. Osorio and J. A. Moreno, "Dissipative design of observers for multivalued nonlinear systems," in *Proc. 45th IEEE Conf. Decision & Control*, San Diego, CA, USA, Dec. 2006, pp. 5400–5405.

[30] L. Paoli, "Continuous dependence on data for vibro-impact problems," *Math. Models and Methods in Appl. Sci.*, vol. 15, no. 1, pp. 53–93, 2005.

[31] L. Paoli, "A proximal-like method for a class of second order measure-differential inclusions describing vibro-impact problems," *J. Different. Equat.*, vol. 250, pp. 476–514, 2011.

[32] L. Paoli and M. Schatzman, "A numerical scheme for impact problems I and II," *SIAM J. Numer. Anal.*, vol. 40, no. 2, pp. 702–768, 2002.

[33] H. Shim, Y. Son, and J. Seo, "Semi-global observer for multi-output nonlinear systems," *Sys. & Control Lett.*, vol. 42, no. 3, pp. 233–244, 2001.

[34] H. Shim and A. Tanwani, "Hybrid-type observer design based on a sufficient condition for observability in switched nonlinear systems," *Int. J. Robust & Nonlin. Control: Special Issue on High Gain Observers and Nonlinear Output Feedback Control*, vol. 24, no. 6, pp. 1064–1089, 2014.

[35] A. Tanwani, B. Brogliato, and C. Prieur, "Observer Design for Frictionless and Unilaterally Constrained Mechanical Systems: A Passivity-Based Approach," Full report, HAL-INRIA, 2015, PDF link: [Online]. Available: https://hal.inria.fr/hal-01113344v2, Simulation gallery: [Online]. Available: http://siconos.gforge.inria.fr/Examples/PassivityBasedObserver.html

[36] A. Tanwani, B. Brogliato, and C. Prieur, " Stability and observer design for multivalued Lur'e systems with non-monotone, time-varying nonlinearities and state jumps," *SIAM J. Control and Optimiz.*, vol. 52, no. 6, pp. 3639–3672, 2014.

[37] A. Tanwani, B. Brogliato, and C. Prieur, "Stability notions for a class of nonlinear systems with measure controls," *Math. Controls, Signals & Syst.*, vol. 27, no. 2, pp. 245–275, 2015.

[38] A. Tanwani, H. Shim, and D. Liberzon, "Observability implies observer design for switched linear systems," in *Proc. ACM Conf. Hybrid Syst.: Comput. Control*, 2011, pp. 3–12.

[39] A. Tanwani, H. Shim, and D. Liberzon, "Observability for switched linear systems: Characterization and observer design," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 891–904, Apr. 2013.

[40] A. Tanwani and S. Trenn, "Observability of switched differential-algebraic equations for general switching signals," in *Proc. 51st IEEE Conf. Decision Control*, 2012, pp. 2648–2653.

[41] A. Tanwani and S. Trenn, "An observer for switched differential-algebraic equations based on geometric characterization of observability," in *Proc. 52nd IEEE Conf. Decision Control*, 2013, pp. 5981–5986.

[42] B. Xian, M. S. de Queiroz, D. M. Dawson, and M. L. McIntyre, "A discontinuous output feedback controller and velocity observer for nonlinear mechanical systems," *Automatica*, vol. 40, pp. 695–700, 2004.