Towards Gluing KKL Observer for Hybrid Systems with Unknown Jump Times (Full Version)

Gia Quoc Bao Tran, Sergio Garcia, Pauline Bernard, Florent Di Meglio, and Ricardo G. Sanfelice

Abstract— **This work proposes an observer design for general hybrid systems, whose outputs are continuous at jumps, and whose jump times are unknown. Inspired by the gluing approach and the Kravaris-Kazantzis/Luenberger (KKL) paradigm, we present conditions under which the hybrid dynamics can be transformed into continuous-time dynamics that take the form of a filter of the output and for which an observer can be readily designed. The possibility of recovering the estimate in the original coordinates is guaranteed outside of the jump times, under a mild backward distinguishability condition that ensures injectivity away from the jump set, assuming sufficient regularity of the transformation. Contrary to previous gluing results, the design of the gluing transformation and the observer is systematic with a well-identified target form of dynamics. While the theoretical conditions are validated on an academic bouncing ball system, we illustrate our method on an application concerning dry friction parameter estimation in the presence of stick-slip, using neural networks to learn a numerical model of the inverse transformation.**

I. INTRODUCTION

Observer design for hybrid systems remains a challenge due to the interconnection of continuous-time and discrete-time dynamics as well as the complex dependence of the solutions' time domain on their (unknown) initial conditions. When the jump times are assumed to be known or detected, which includes mechanical systems with impact sensors, sampled continuous-time systems, or switched systems with known switching times, the jumps of the observer can be triggered at the same time when the system jumps, thereby facilitating convergence and stability analysis (see [1] and the references therein). Observer designs for this class are rich in the literature, such as [2], [3] in the context of either impulsive or switched systems, [4], [5] for continuous-time systems with sporadic measurements, or [6], [1] for the general hybrid context, among many others, with designs including the Kalman-like design [7] or the coupling of flow- and jump-based observers under a unified Lyapunov analysis [8].

*Research supported by the French National Research Agency (ANR-23-CE48-0013-01). Research by R. G. Sanfelice partially supported by NSF Grants no. CNS-2039054 and CNS-2111688, by AFOSR Grants nos. FA9550-19-1-0169, FA9550-20-1-0238, FA9550-23-1-0145, and FA9550-23-1-0313, by AFRL Grant nos. FA8651-22-1-0017 and FA8651-23-1-0004, by ARO Grant no. W911NF-20-1-0253, and by DoD Grant no. W911NF-23-1-0158.

Gia Quoc Bao Tran, Pauline Bernard, and Florent Di Meglio are with Centre Automatique et Systèmes (CAS), Mines Paris - PSL, Paris, France {firstname,lastname}@minesparis.psl.eu. Sergio Garcia is with Mines Paris - PSL, Paris, France sergio.garcia_castro@etu.minesparis.psl.eu. Ricardo G. Sanfelice is with Department of Electrical and Computer Engineering, University of California, Santa Cruz, CA 95064, USA ricardo@ucsc.edu.

On the other hand, when the jump times are unknown, the definition of an observer and its convergence, or even of observability, is no longer straightforward because the observer jumps are not synchronized with those of the system—the observer cannot converge to the system state in standard distances around the jump times, because of non-simultaneous (in time) discontinuities [9]. Such a phenomenon also appears in the context of contraction and trajectory tracking [10], [11], [12]. Concerning observer design, very few results exist for general hybrid systems [13] apart from: (i) mode location observers for switched systems with unknown switching times, but mainly in the linear case (either by running parallel observers and monitoring the residual output error [14], [15], or via LMI design [16], or through optimization algorithms to find the switching signal matching most of the output signal over a time window [17]); (ii) a semi-global design based on an arbitrarily fast high-gain observer during flow that is "disconnected" around jump times, for systems with instantaneously observable flow dynamics [18]; (iii) a "gluing" design for hybrid systems whose outputs are continuous at jumps, based on transforming the hybrid dynamics into continuous-time dynamics where an observer may be designed [19] (see also [20] for mechanical systems with impacts). In [19], the existence of such a gluing change of coordinates that is injective except on the jump set is shown to exist for a broad class of hybrid systems with appropriate manifold structure of the flow and jump sets. Then, if an observer can be designed in the new coordinates, the state estimate is obtained by running this continuous-time observer and inverting the transformation, with asymptotic convergence except during increasingly smaller intervals around the jump times. However, [19] does not provide constructive methods to find the transformation and does not guarantee that an observer can be designed in the target coordinates.

On the other hand, the Kravaris-Kazantzis/Luenberger (KKL) paradigm provides a universal idea of observer design for nonlinear systems. This framework relies on transforming the dynamics into some linear stable filter of the output, where the observer is straightforwardly designed, and the estimate is recovered in the original coordinates by inverting this transformation under a weak backward distinguishability condition [21], [22]. The rich literature of KKL observers for many classes of systems (continuous-/discrete-time, autonomous/time-varying, etc.), including the closed form of the KKL transformation and observability condition for each class, is summarized in [23]. While the implementation of these remains a challenge due to the difficulty in computing the transformation, systematic numerical methods are being developed to learn models of the (inverse) transformation using neural networks (NNs) [24], [25], [26].

In this work, we propose to use the KKL transformation as a gluing function for general hybrid systems of form [13]

$$\dot{x} = f(x) \ \ x \in C, \quad x^+ = g(x) \ \ x \in D, \quad y = h(x), \quad (1)$$

with state $x \in \mathbb{R}^{n_x}$, where $C, D \subset \mathbb{R}^{n_x}$ are the flow and jump sets, $f, g, h : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ are the flow, jump and output maps, respectively, such that the output is continuous at jumps. Thus, the jumps are not immediately visible in the output, and pairs of points $(x, g(x))$ for x in D are not distinguishable. Still, in the spirit of [19], we show that there exists a map $T : \mathrm{cl}(C) \cup D \to \mathbb{R}^{n_z}$ such that the image by T of solutions to system (1) follows the continuous-time dynamics

$$\dot{z} = Az + By \quad (2)$$

for some pairs $(A, B) \in \mathbb{R}^{n_z \times n_z} \times \mathbb{R}^{n_z}$ with A Hurwitz, and T is injective on $\mathrm{int}(C \setminus D)$, under backward distinguishability of the system outside of the jump set and some regularity conditions. Then, as in (27) below, a trivial observer for the target form (2) is obtained by running system (2) from any initial condition and an estimate for x can be recovered by a left inversion of T, with asymptotic convergence except in smaller and smaller intervals around the jump times. The available tools for the numerical approximation of T and its left inverse generalize to this hybrid framework with some adaptation, taking into account the loss of injectivity of T on D, thus providing a systematic observer design for hybrid systems with unknown jump times. Models of the form (1) cover dynamical systems with state-triggered changes, which includes state-triggered switched systems or hybrid automata, with state $x = (x_c, q)$ where x_c is the physical state and $q \in \mathbb{N}$ encodes the modes, $f = (f_q, 0)$ the continuous dynamics in each mode, and g and D the transitions from each mode to the others, that can also depend on x_c. While the theoretical assumptions and results are illustrated on the classical bouncing ball toy problem, the gluing KKL observer is then applied to the problem of dry friction parameter estimation on a drilling mechanical system exhibiting a switching stick-slip behavior. A hybrid model of dimension 7 with unknown jump times is proposed, for which an analytical gluing function would be otherwise very difficult to find.

After setting the technical background and assumptions of this paper in Section II, we propose in Section III a systematic change of coordinates into the continuous-time system (2) and study its injectivity on $C \setminus D$. The possibilities of returning to the initial x-coordinates are discussed in Section IV by revisiting the convergence result of [19] under milder assumptions. Finally, a numerical implementation of this gluing KKL observer is shown in Section V on an application featuring the stick-slip phenomenon.

Notations: Let $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{N}$ (resp., $\mathbb{Z}$) denote the set of real, complex, and natural numbers (resp., integers). Let $\mathbb{R}_{\geq 0}$ (resp., $\mathbb{R}_{\leq 0}$) denote $[0, +\infty)$ (resp., $(-\infty, 0]$), and similarly for $\mathbb{Z}_{\geq 0}$ and $\mathbb{Z}_{\leq 0}$. Let $\Re(z)$ be the real part of $z \in \mathbb{C}$. Given $\rho > 0$, define $\mathbb{R}_\rho = \{\lambda \in \mathbb{R} : \lambda < -\rho\}$ and $\mathbb{C}_\rho = \{\lambda \in \mathbb{C} : \Re(\lambda) < -\rho\}$. Denote $\mathbb{R}^{m \times n}$ as the set of real $(m \times n)$-dimensional matrices. Given a set S, $\mathrm{cl}(S)$ is its closure, and $\mathrm{int}(S)$ denotes its interior. Denote $\mathcal{S} + \delta$ as the set of points within a distance smaller than or equal to some $\delta > 0$ from a point in $\mathcal{S}$. Let $|\cdot|$ be the Euclidean norm. Denote $\langle a, b \rangle$ as the scalar product of vectors a and b. Let $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the diagonal matrix with entries λ_i, $i = 1, 2, \ldots, n$. The notions of class-$\mathcal{K}$ and class-$\mathcal{K}_\infty$ functions are from [27, Definitions 4.2 and 4.3]. For the dynamics $\dot{x} = f(x)$, $\Psi_f(x_0, \tau)$ is the associated flow operator from initial value x_0 evaluated after τ time unit(s). For a solution $(t, j) \mapsto \phi(t, j)$ to a hybrid system, we denote by $\mathrm{dom}\, x$ its domain [13], $\mathrm{dom}_t\, \phi$ (resp., $\mathrm{dom}_j\, \phi$) the domain's projection on the ordinary time (resp., jump) component, for $j \in \mathbb{N}$, $t_j(\phi)$ the unique time such that $(t_j(\phi), j) \in \mathrm{dom}\, \phi$ and $(t_j(\phi), j-1) \in \mathrm{dom}\, \phi$. The mention of ϕ is omitted when no confusion is possible. Given a hybrid arc ϕ defined on $\mathrm{dom}\, \phi$, let $\phi_{|\mathcal{D}}$ be the restriction of ϕ to $\mathcal{D} \subset \mathrm{dom}\, \phi$.

II. TECHNICAL ASSUMPTIONS

In [13, Definition 2.6], solutions to system (1) are defined in *forward positive* hybrid time, i.e., on a hybrid time domain subset of $\mathbb{R}_{\geq 0} \times \mathbb{N}$. Here, to define our transformation later, we consider solutions defined in both forward and backward time, namely on a hybrid time domain subset of $\mathbb{R} \times \mathbb{Z}$. Thus, we generalize the notion of a hybrid time domain.

Definition 1: A subset $E \subset \mathbb{R} \times \mathbb{Z}$ is a *compact hybrid time domain* if, denoting $E_{\geq 0} = E \cap (\mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0})$ and $E_{\leq 0} = E \cap (\mathbb{R}_{\leq 0} \times \mathbb{Z}_{\leq 0})$, we have (when not empty)

$$E_{\geq 0} = \bigcup_{j=0}^{J_M - 1} \left([t_j, t_{j+1}] \times \{j\} \right), \quad (3a)$$

$$E_{\leq 0} = \bigcup_{j=J_m}^{-1} \left([t_j, t_{j+1}] \times \{j+1\} \right), \quad (3b)$$

for some integers $J_m \leq 0$, $J_M \geq 0$ and a finite sequence of times $t_{J_m} \leq t_{J_m+1} \leq \ldots \leq t_0 = 0 \leq \ldots \leq t_{J_M-1} \leq t_{J_M}$ in $\mathbb{R}$. A set $E \subset \mathbb{R} \times \mathbb{Z}$ is a *hybrid time domain* if it is the union of a non-decreasing sequence of compact hybrid time domains, namely, E is the union of compact hybrid time domains E_j such that $\ldots \subset E_{j-1} \subset E_j \subset E_{j+1} \ldots$.

This definition corresponds to that of [13, Definition 2.6] when $J_m = 0$. More generally, it coincides with the notion of hybrid time domain with memory introduced in [28], but where $E_{\leq 0}$ is rather written as

$$E_{\leq 0} = \bigcup_{k=1}^{K} \left([s_k, s_{k-1}] \times \{-k+1\} \right), \quad (4)$$

with the convention that $s_k = t_{-k}$. Inspired from [29, Definitions 6 and 8], we now define solutions to system (1) in forward and backward time.

Definition 2: Given system (1), we define its backward counterpart as

$$\dot{x} = -f(x) \qquad x \in C, \tag{5a}$$

$$x^+ \in \{x' \in D : x = g(x')\} \qquad x \in g(D), \tag{5b}$$

$$y = h(x). \tag{5c}$$

Then, $\phi : \operatorname{dom}\phi \to \mathbb{R}^{n_x}$ is solution to system (1) if $\operatorname{dom}\phi$ is a hybrid time domain and if, denoting $\mathcal{D}_{\geq 0} = \operatorname{dom}\phi \cap (\mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0})$ and $\mathcal{D}_{\leq 0} = \operatorname{dom}\phi \cap (\mathbb{R}_{\leq 0} \times \mathbb{Z}_{\leq 0})$,

1) $\phi_{fw} := \phi_{|\mathcal{D}_{\geq 0}}$ is solution to (1) on $\mathcal{D}_{fw} := \mathcal{D}_{\geq 0}$ in the sense of [13, Definition 2.6];
2) ϕ_{bw} defined on $\mathcal{D}_{bw} := -\mathcal{D}_{\leq 0}$ as

$$\phi_{bw}(t,j) = \phi(-t,-j), \qquad \forall (t,j) \in \mathcal{D}_{bw}, \tag{6}$$

is a solution to (5) in the sense of [13, Definition 2.6].

Remark 1: Another way of defining solutions without using [13] would be to say that : $\phi : \operatorname{dom}\phi \to \mathbb{R}^n$ is a solution to system (1) if $\operatorname{dom}\phi$ is a hybrid time domain, the map $t \mapsto \phi(t,j)$ is locally absolutely continuous on each interval $I_j = \{t : (t,j) \in \operatorname{dom}\phi\}$ for all $j \in \operatorname{dom}_j \phi$, and it satisfies the following properties:

(a) $\phi(0,0) \in \operatorname{cl}(C) \cup D$;

(b) For each $j \in \operatorname{dom}_j \phi$ such that $\operatorname{int}(I_j) \neq 0$ we have

$$\begin{aligned} \dot{\phi}(t,j) &= f(\phi(t,j)), && \text{for almost all } t \in I_j, \\ \phi(t,j) &\in C, && \text{for all } t \in \operatorname{int}(I_j) \setminus \{0\}; \end{aligned} \tag{7}$$

(c) For each $(t,j) \in \operatorname{dom}\phi$ such that $(t,j+1) \in \operatorname{dom}\phi$, we have

$$\begin{aligned} \phi(t,j+1) &= g(\phi(t,j)), \\ \phi(t,j) &\in D. \end{aligned} \tag{8}$$

A solution ϕ to system (1) is maximal if there does not exist any other solution ϕ' to system (1) such that $\operatorname{dom}\phi$ is a strictly proper subset of $\operatorname{dom}\phi'$ and $\phi = \phi'$ on $\operatorname{dom}\phi$. Besides, ϕ is forward (resp., backward) complete if $\operatorname{dom}\phi \cap (\mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0})$ (resp., $\operatorname{dom}\phi \cap (\mathbb{R}_{\leq 0} \times \mathbb{Z}_{\leq 0})$) is unbounded, ϕ is t-forward (resp., t-backward) complete if $\operatorname{dom}_t \phi \cap \mathbb{R}_{\geq t}$ (resp., $\operatorname{dom}_t \phi \cap \mathbb{R}_{\leq t}$) is unbounded (and similarly, for j-forward and backward completeness). Last, ϕ is Zeno in forward (resp., backward) time if it is forward (resp., backward) complete and $\sup \operatorname{dom}_t \phi$ (resp., $\inf \operatorname{dom}_t \phi$) is bounded.

We are interested in estimating the state of system (1) from the knowledge of the measurement y. To that end, we make the following assumptions.

Assumption 1: For system (1), assume that:

(A1.1) There exist sets $\mathcal{X}_0 \subset \mathcal{X} \subset \operatorname{cl}(C) \cup D$ such that all maximal solutions to system (1) initialized in $\mathcal{X}_0$ are t-forward complete and remain in $\mathcal{X}$ in forward time;

(A1.2) For any $x \in \operatorname{cl}(C) \cup D$, there exists a unique maximal solution ϕ to system (1) such that $\phi(0,0) = x$;

(A1.3) The maps f, g, and h are continuous, and

$$h(g(x)) = h(x), \qquad \forall x \in D. \tag{9}$$

In Assumption 1, the sets $\mathcal{X}_0$ and $\mathcal{X}$ in Item (A1.1) could be $\operatorname{cl}(C) \cup D$ if no extra information is available about the solutions of interest, but sometimes we know from physical knowledge that they remain within some bounds that can be encoded in $\mathcal{X}$, from a certain set of initial conditions $\mathcal{X}_0$ that may be unknown. The t-forward completeness condition of the maximal solutions of interest is needed because we propose in this paper an *asymptotic* observer design exploiting (2). With Item (A1.3), we assume that the output y does not change at jumps, which in turn means that jumps in the solutions cannot be detected by the jumps of the measurement. Item (A1.2) then allows us to uniquely define an output map as follows.

Definition 3: Suppose Items (A1.2) and (A1.3) of Assumption 1 hold. For every $x \in \operatorname{cl}(C) \cup D$, given the maximal solution ϕ to system (1) initialized as x, define

$$t^-(x) := \inf \operatorname{dom}_t \phi, \qquad t^+(x) := \sup \operatorname{dom}_t \phi, \tag{10}$$

and $Y(x,\cdot) : (t^-(x), t^+(x)) \to \mathbb{R}$ as follows

$$Y(x,t) = h(\phi(t,j(t))), \qquad \forall t : (t,j(t)) \in \operatorname{dom}\phi, \tag{11}$$

where $j(t) = \min\{j' : (t,j') \in \operatorname{dom}\phi\}$.

Remark 2: With (9) in Item (A1.3) of Assumption 1, $j(t)$ could be replaced by any j such that $(t,j) \in \operatorname{dom}\phi$ in (11). The map $Y(x,\cdot)$ is well-defined and continuous for all $x \in \operatorname{cl}(C) \cup D$ thanks to (9). Indeed, for any $t \in \operatorname{dom}_t \phi$, $h(\phi(t,j')) = h(g^{j'-j}(\phi(t_j,j))) = h(\phi(t_j,j))$ for any $j < j'$ such that (t,j) and (t,j') in $\operatorname{dom}\phi$.

We now propose a KKL-based approach to estimate the state of system (1), which does not require jump detection.

III. CHANGE OF COORDINATES INTO CONTINUOUS-TIME DYNAMICS

Following the KKL approach [22] and the gluing idea in [19], we wish to find a map $T : \operatorname{cl}(C) \cup D \to \mathbb{R}^{n_z}$ such that 1) the image of solutions to system (1) under T follows continuous-time dynamics of the form $\dot{z} = Az + By$, for appropriate matrices A and B, and such that 2) its restriction to the set $C \backslash D$ is injective, in order to reconstruct from z an estimate $\hat{x}$ of x that converges in the x-coordinates, at least outside of the jump times.

A. *Definition of T*

In order to define T, we make the following assumption.

Assumption 2: There exists $\rho > 0$ such that for every $x \in \operatorname{cl}(C) \cup D$, the map $Y(x,\cdot)$ introduced in Definition 3 verifies:

(a) If $t^-(x) \neq -\infty$, $\lim\limits_{s \to t^-(x)} Y(x,s)$ exists and is finite;

(b) If $t^-(x) = -\infty$, $s \mapsto e^{\rho s} Y(x,s)$ is integrable on $\mathbb{R}_{\leq 0}$.

Example 1: Consider a bouncing ball with height x_1, velocity x_2, and restitution coefficient $c > 0$, described by

$$\begin{cases} \dot{x} = (x_2, -dx_2^p - a_g), & \text{if } x_1 \geq 0 \\ x^+ = (x_1, -cx_2 + \mu), & \text{if } x_1 = 0 \text{ and } x_2 \leq 0, \end{cases} \tag{12}$$

where $a_g = 9.8$ (m/s^2) is the gravitational acceleration, $c > 0$ is a restitution coefficent, $\mu > 0$ is some constant jump input, $d > 0$ and $p \in \mathbb{R}_{\geq 0}$ are friction parameters, and with output $y = x_1$. Firstly, maximal solutions to system (12) initialized in $\mathcal{X}_0 = \mathbb{R}_{\geq 0} \times \mathbb{R}$ are both t- and j-forward complete and remain in $\mathcal{X} = \mathcal{X}_0$ in forward time. Secondly, considering the backward counterpart of system (12)

$$\begin{cases} \dot{x} = (-x_2, dx_2^p + a_g), & \text{if } x_1 \geq 0 \\ x^+ = \left(x_1, \frac{-x_2+\mu}{c}\right), & \text{if } x_1 = 0 \text{ and } x_2 \geq \mu, \end{cases} \tag{13}$$

we deduce that maximal solutions to system (12) are unique (in both forward and backward time). Thirdly, (9) holds. Therefore, system (12) satisfies Assumption 1. If $p = 1$ and $c \leq 1$, all maximal solutions to (12) in backward time are either t-backward complete or with a bounded and closed time domain (indeed, no Zeno phenomenon can happen in system (13)[1] and solutions either end with flow at $x = (0, x_2)$ with $0 \leq x_2 < \mu$, or end with one jump to $(0, 0)$). For the former type, thanks to linearity in the maps and a_g being constant, backward solutions explode at most exponentially during flow and linearly at jumps (with a dwell time) so that Item (b) of Assumption 2 holds. For the latter, Item (a) of Assumption 2 holds by continuity. Therefore, Assumption 2 is satisfied in this case. However, with $p > 1$, we may have a finite-time escape in backward time, so solutions may not be t-backward complete and $Y(x, s)$ may not have a finite limit as $s \to t^-(x)$, so Assumption 2 may not hold.

With ρ defined in Assumption 2, consider $n_z \in \mathbb{N}$, $A \in \mathbb{R}^{n_z \times n_z}$ such that $A + \rho I$ is Hurwitz, and $B \in \mathbb{R}^{n_z}$. We define $T : C \cup D \to \mathbb{R}^{n_z}$ as

$$T(x) = \int_{-\infty}^{0} e^{-As} B \breve{Y}(x, s) ds, \tag{14}$$

where, for every $x \in C \cup D$ and $s \in \mathbb{R}$,

$$\breve{Y}(x, s) = \begin{cases} Y(x, s), & \text{if } s > t^-(x), \\ \lim\limits_{\tau \to t^-(x)} Y(x, \tau), & \text{otherwise.} \end{cases} \tag{15}$$

Remark 3: While for every $x \in \mathrm{cl}(C) \cup D$, $s \mapsto Y(x, s)$ is defined on $(t^-(x), t^+(x))$ only, $s \mapsto \breve{Y}(x, s)$ is defined and continuous on $\mathbb{R}_{\leq 0}$ according to Assumption 2. Besides, still under Assumption 2, for all $x \in \mathrm{cl}(C) \cup D$, the function $s \mapsto e^{-As} B \breve{Y}(x, s)$ is integrable on $\mathbb{R}_{\leq 0}$. It is thus a continuous extension of $Y(x, \cdot)$, allowing us to define T as in (14).

B. Continuous-time Dynamics in the z-Coordinates

Now we prove that the image by T of solutions to system (1) satisfies (2).

Lemma 1: Suppose Assumptions 1 and 2 hold. For any maximal solution ϕ to system (1) initialized in $\mathcal{X}_0$, there exist a C^1 map $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_z}$ and a continuous map $y : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_y}$ such that

$$T(\phi(t, j)) = z(t), h(\phi(t, j)) = y(t), \forall (t, j) \in \operatorname{dom} \phi : t \geq 0, \tag{16}$$

and

$$\dot{z}(t) = Az(t) + By(t), \qquad \forall t \in \mathbb{R}_{\geq 0}, \tag{17}$$

where T is defined in (14) with $A \in \mathbb{R}^{n_z \times n_z}$ such that $A + \rho I$ is Hurwitz for ρ in Item (b) of Assumption 2, and $B \in \mathbb{R}^{n_z}$.

[1]This is because $x_2 \geq \mu$ before a jump and $x_2 \leq 0$ after, and $\mu > 0$ along with the definition of the flow map, giving us a uniform bound between these values.

Proof: Let ϕ be a maximal solution to system (1). The map $y : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_y}$ defined by $y(t) = h(\phi(t, j))$ for all $(t, j) \in \operatorname{dom} \phi$, $t \geq 0$, is well-defined and continuous according to Items (A1.2) and (A1.3) of Assumption 1. Then, we exploit Lemma 3 to deduce that

$$T(g(x)) = T(x), \qquad \forall x \in D. \tag{18}$$

It follows that there exists a continuous map $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_z}$ such that $z(t) = T(\phi(t, j))$ for all $(t, j) \in \operatorname{dom} \phi$ with $t \geq 0$. Let us show that z is C^1. Consider first $t \in \mathbb{R}_{\geq 0}$ such that $t \in (t_j, t_{j+1})$ for some $j \in \operatorname{dom}_j \phi$ and a scalar $\Delta \neq 0$ small enough such that $t + \Delta \in (t_j, t_{j+1})$. Exploiting the uniqueness of solutions from Item (A1.2) of Assumption 1 and the fact that for all $x \in \mathrm{cl}(C) \cup D$ and $s \in [0, \Delta]$, we have $\breve{Y}(x, s) = Y(x, s)$, we get

$$\begin{aligned} z(t + \Delta) &= \int_{-\infty}^{0} e^{-As} B \breve{Y}(\phi(t + \Delta, j), s) ds \\ &= \int_{-\infty}^{0} e^{-As} B \breve{Y}(\phi(t, j), s + \Delta) ds \\ &= \int_{-\infty}^{\Delta} e^{-A(s' - \Delta)} B \breve{Y}(\phi(t, j), s') ds' \\ &= e^{A\Delta} \bigg(\int_{-\infty}^{0} e^{-As'} B \breve{Y}(\phi(t, j), s') ds' \\ &\qquad + \int_{0}^{\Delta} e^{-As'} B \breve{Y}(\phi(t, j), s') ds' \bigg) \\ &= e^{A\Delta} z(t) + e^{A\Delta} \int_{0}^{\Delta} e^{-As} B Y(\phi(t, j), s) ds. \end{aligned}$$

Rearranging the terms, we get

$$\begin{aligned} \frac{z(t + \Delta) - z(t)}{\Delta} &= \frac{(e^{A\Delta} - I)}{\Delta} z(t) \\ &\quad + \frac{1}{\Delta} e^{A\Delta} \int_{0}^{\Delta} e^{-As} B Y(\phi(t, j), s) ds. \end{aligned}$$

Taking the limit as $\Delta \to 0$, we obtain that z is differentiable at time t and

$$\begin{aligned} \dot{z}(t) &= Az(t) + BY(\phi(t, j), 0) \\ &= Az(t) + Bh(\phi(t, j)) = Az(t) + By(t). \end{aligned}$$

Now consider any $t > 0$. By the t-forward completeness in Item (A1.1) of Assumption 1, there exist $j_1, j_2 \in \operatorname{dom}_j \phi$ such that $(t - \Delta, j_1) \in \operatorname{dom} \phi$ and $(t + \Delta, j_2) \in \operatorname{dom} \phi$, for all $\Delta > 0$ sufficiently small. Reproducing the same computations as before, we get

$$\begin{aligned} \lim_{\Delta \to 0^-} \frac{z(t + \Delta) - z(t)}{\Delta} &= Az(t) + BY(\phi(t, j_1), 0) \\ &= Az(t) + By(t), \\ \lim_{\Delta \to 0^+} \frac{z(t + \Delta) - z(t)}{\Delta} &= Az(t) + BY(\phi(t, j_2), 0) \\ &= Az(t) + By(t). \end{aligned}$$

Similarly, at $t = 0$, reasoning with $\Delta > 0$, we get that z is continuously differentiable on $\mathbb{R}_{\geq 0}$ and verifies (17). ■

In Lemma 1, we have shown the existence of a map T transforming system (1) into the continuous-time dynamics (17). It follows that for any solution ϕ to system (1), implementing (17) and

$$\dot{\hat{z}}(t) = A\hat{z}(t) + By(t) \tag{19}$$

fed with the measured output y, from any initial condition, gives us

$$\lim_{t+j\to+\infty} |T(\phi(t,j)) - \hat{z}(t)| = 0, \tag{20}$$

namely, $\hat{z}(t)$ provides an asymptotic estimate of $T(\phi(t,j))$ *without any jump detection*. The great advantage of this approach compared to [19] is that an observer is directly available in the z-coordinates, given the specific target form of the dynamics (17). Moreover, the change of coordinates given by T is guaranteed to exist, with a systematic approach for constructing a numerical model of it (see Section V).

Example 2: Consider the bouncing ball in Example 1 with $d = 0.01$ (m^{-1}), $p = 2$, $c = 0.8$, and $\mu = 2$ (m/s). Exploiting [19] only, we are not able to find an analytic gluing function T for this system. Instead, we follow the KKL route of this paper. Taking advantage of the system's low dimension, we propose to approximate T using a look-up table. To do this, we simulate the interconnection (1)-(17) from any initial conditions in $\mathcal{X}_0 \times \mathbb{R}^{n_z}$, where $n_z = 3$ with $A = \operatorname{diag}(-1, -2, -3)$ and $B = (1, 1, 1)$. After a long enough time compared to the eigenvalues of A, z approximates $T(x)$ and we can form a look-up table consisting of the $(x, T(x))$ pairs by storing these points taken from a large number of simulations. In Figure 1, we show (left) the data points of the first component of $z = (z_1, z_2, z_3)$ as a function of $x = (x_1, x_2)$ taken from the look-up table and plot, along a particular system solution $(t, j) \mapsto x(t, j)$, the value $z_1(t)$ in the look-up table (approximating $T(x(t, j))$) such that its corresponding x component matches most $x(t, j)$ in the Euclidean norm. It is confirmed that the values of z_1 before and after each jump are the same, which is consistent with the continuity of z at the jumps. On the right, we compare $z_1(t)$ from the look-up table and $\hat{z}_1(t)$ obtained by running observer (19) from some arbitrary initial condition, showing the convergence in the z-coordinates. Conditions to deduce an estimate in the x-coordinates are given in the next section.

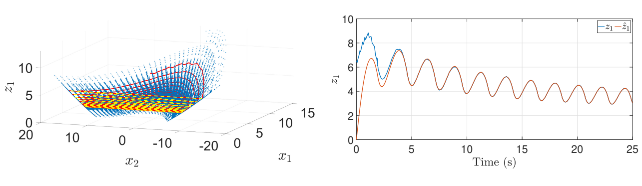

Fig. 1: Left: Data points (x, z_1) in look-up table (blue), and $t \mapsto (x(t,j), z_1(t))$ along a solution (red-yellow), where $z_1(t)$ is fetched as the closest point in the look-up table; Right: $t \mapsto z_1(t)$ from look-up table obtained from $t \mapsto x(t,j)$ vs. $t \mapsto \hat{z}_1(t)$ solution to observer (19).

C. Injectivity of T

As it is usually done in KKL theory, we now exploit a distinguishability property to study the injectivity of T, and thus the ability to reconstruct ϕ from the knowledge of $\hat{z}$. Of course, due to (18), T is not injective on D unless the map g is identity, and we thus focus on $C \setminus D$.

Assumption 3: Any distinct points x_a, x_b in $\operatorname{int}(C\backslash D)$ are backward distinguishable, namely there exists $t \in (\max\{t^-(x_a), t^-(x_b)\}, 0]$ such that $Y(x_a, t) \neq Y(x_b, t)$, with Y introduced in Definition 3.

Using Assumption 3, we now exploit the tools developed in [22] (and the references therein) to show that the map T is injective on $\operatorname{int}(C\backslash D)$.

Lemma 2: Suppose Assumptions 1, 2, and 3 hold. Assume that, for all $\lambda \in \mathbb{C}_\rho$ with ρ from Assumption 2, the map

$$x \mapsto T_0(\lambda, x) := \int_{-\infty}^{0} e^{-\lambda s} \breve{Y}(x, s) ds \tag{21}$$

is C^1 on $\operatorname{int}(C\backslash D)$; moreover, for all $\lambda \in \mathbb{R}_\rho$ and for all $k \in \mathbb{N}$, the map $x \mapsto \frac{\partial^k T_0}{\partial \lambda^k}(\lambda, x)$ exists and is C^1 on $\operatorname{int}(C\backslash D)$. Define $m_0 = 2n_x + 1$. For almost any pair of matrices $(A_0, B_0) \in \mathbb{R}^{m_0 \times m_0} \times \mathbb{R}^{m_0}$ with $A_0 + \rho I$ Hurwitz for ρ in Item (b) of Assumption 2, the map $T : C \cup D \to \mathbb{R}^{n_z}$ defined in (14) with $A = A_0 \otimes I_{n_y}$ and $B = B_0 \otimes I_{n_y}$ is injective on $\operatorname{int}(C\backslash D)$.

Proof: Similar to [22, Theorem 3.4]. See in [30]. Similarly to [22, Appendix B.1], we show that the result is equivalent to showing that for all $l \in \{0, 1, \ldots, n_x\}$ and for almost all $(\lambda_1, \lambda_2, \ldots, \lambda_{2n_x - l + 1}) \in \Omega_{l,\rho}$, with $\Omega_{l,\rho} = \mathbb{C}_\rho^l \times \mathbb{R}_\rho$, the map

$$x \mapsto T_{\text{diag}}(x) = (T_0(\lambda_1, x), T_0(\lambda_2, x), \ldots, T_0(\lambda_{2n_x - l + 1}, x)) \tag{22}$$

is injective on $\operatorname{int}(C\backslash D)$. Now we adapt [22, Appendix B.2.3]. Since $\operatorname{int}(C\backslash D)$ is open, we use $\Upsilon = \{(x_a, x_b) \in \operatorname{int}(C\backslash D) \times \operatorname{int}(C\backslash D) : x_a \neq x_b\}$ an open subset of $\mathbb{R}^{2n_x}$, and define the same Θ_i, the same $g_i(\lambda, x_a, x_b) = T_0(\lambda, x_a) - T_0(\lambda, x_b) = \int_{-\infty}^{0} e^{-(\lambda+\rho)s} \Delta(x_a, x_b, s) ds$ with

$$\Delta(x_a, x_b, s) = e^{\rho s}(\breve{Y}(x_a, s) - \breve{Y}(x_b, s)). \tag{23}$$

From Assumption 3, for all $(x_a, x_b) \in \Upsilon$, by the definition of $\breve{Y}$ in (15), there exists $s \leq 0$ such that $\Delta(x_a, x_b, s) \neq 0$. By properties of the Laplace transform and continuity of $s \mapsto \breve{Y}(x, s)$, we deduce that for all $(x_a, x_b) \in \Upsilon$, $\lambda \mapsto g_i(\lambda, x_a, x_b)$ cannot be identically zero on $\Omega_{\ell,\rho}$. Moreover, we can check the regularity conditions of [22, Lemma B.3]: (i) for all $x \in \operatorname{int}(C \setminus D)$, $T_0(\cdot, x)$ is holomorphic on $\mathbb{C}_\rho$ and is C^∞ on $\mathbb{R}_\rho$, and (ii) we have the required regularity of T_0 with respect to x by assumption. Applying the generalized Coron's lemma [22, Lemma B.3], we get the results. ■

Remark 4: If the regularity of $\frac{\partial^k T_0}{\partial \lambda^k}$ is not guaranteed, we can still achieve injectivity of T_{diag} for almost any $(\lambda_1, \lambda_2, \ldots, \lambda_{n_x+1}) \in \mathbb{C}_\rho^{n_x+1}$ as in [21, Theorem 3], for $A_0 = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_x+1})$ and $B_0 = (1, \ldots, 1) \in \mathbb{R}^{n_x+1}$.

The continuity of Y and thus $\breve{Y}$ and thus T should be ensured under the hybrid basic conditions and some uniformity in Assumption 2, by exploiting [13, Proposition 6.14] as well as the continuity of h at jumps. However, the continuous differentiability of T is left to further study, although evidence of this regularity is obtained in simulations and from preliminary theoretical results under reasonable conditions on the data of the system.

The reason for considering $\text{int}(C \backslash D)$ instead of $C \backslash D$ or $\text{cl}(C) \setminus D$ is that an open set is needed to apply Coron's lemma, the key tool of [21], [22]. A way around this would be to manage to define T on a larger open set containing $C \backslash D$, but this typically requires us to extend Y outside of $\text{cl}(C) \cup D$ while preserving its regularity.

Remark 5: It is interesting to note that, when $n_y = 1$, the generic dimension $n_z = 2n_x + 1$ providing injectivity of T according to Lemma 2 corresponds to the dimension of the gluing function guaranteed to exist in [19, Remark 2] through proper embedding arguments and under additional smoothness and manifold assumptions on the data. This dimension is conservative, to guarantee a generic result impervious to the data of the hybrid system, but it is not necessary as will be illustrated in the examples.

Back to our estimation problem, we are interested in reconstructing $x(t, j)$ from the knowledge of $T(x(t, j))$. The injectivity of T on $\text{int}(C \backslash D)$ suggests it is possible except around the jump times. To formalize the notion of convergence, we use the concept of the gluing function introduced in [19].

Example 3: Consider the bouncing ball in Example 1 with parameters in Example 2. Since the whole state is instantaneously observable during flows, Assumption 3 is satisfied. We take $n_z = 3$ which is seen to give injectivity of T on $C \setminus D$, run observer (19), and recover the estimate in the x-coordinates by the look-up table built in Example 2. In Figure 2, convergence is recovered in the x-coordinates, but outside of the jump times when we cannot distinguish between the x before and after the jumps. Note that for systems of large dimensions, the look-up table approach does not give satisfactory performance due to memory limitations and the curse of dimensionality. Therefore, an NN-based approach is proposed in Section V.

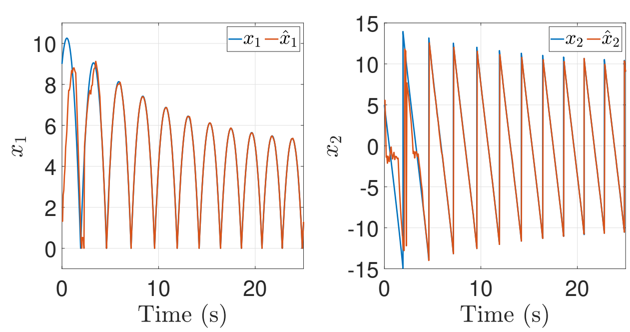

Fig. 2: Estimation results with the bouncing ball system.

IV. CONVERGENCE IN THE x-COORDINATES

A gluing function is essentially a function that transforms the hybrid system (1) into continuous-time dynamics, by "gluing" the jump set D with its image $g(D)$ while preserving injectivity on the rest of the domain. The key gluing properties are thus as follows.

Definition 4: A function $T : C \cup D \to \mathbb{R}^{n_z}$, with $m \geq n$, is called a *gluing function* for system (1) if it satisfies:

(G.1) $T(x) = T(g(x))$ for all $x \in D$;

(G.2) T is injective on $C \setminus D$.

In [19, Definition 1], a gluing function is additionally required to be C^1, with a Jacobian that is full-rank on C. We show that this assumption is not needed here to demonstrate the convergence of the estimate, its only impact being to 1) study the dynamics of $T(x)$ along solutions, but this is done here in Lemma 1 without any regularity condition on T, and 2) provide a *linear* modulus of injectivity of T, instead of $\mathcal{K}^\infty$ one in (33), but we show this is not needed for the proof of Theorem 1.

According to (18) and Lemma 2, the function $T : C \cup D \to \mathbb{R}^{n_z}$ defined in (14) is *almost* a gluing function (as in Definition 4) for system (1): it is if the injectivity is ensured on the entire $C \setminus D$ (not only on its interior).

By implementing (17) from any initial condition, we know that $\hat{z}$ converges arbitrarily close to $T(x)$. It is thus tempting to apply a left inverse of T to $\hat{z}$. But by both Items (G.1) and (G.2) of Definition 4, this is only possible in $T(C \setminus D)$. In [19], it is proposed to project $\hat{z}$ onto $T(C \setminus D)$ and sufficient conditions are given to ensure that x persistently and uniformly stays away from D and $g(D)$.

Assumption 4: For system (1), assume that:

(A4.1) The set $\mathcal{X}$ defined in Item (A1.1) of Assumption 1, in which the solutions of interest remain in forward time, is compact;

(A4.2) $D \cap g(D) = \varnothing$;

(A4.3) There exist smooth maps $r_D : \mathbb{R}^{n_x} \to \mathbb{R}$ and $r_{g(D)} : \mathbb{R}^{n_x} \to \mathbb{R}$ satisfying that

$$\begin{aligned} D &= \{x \in C : r_D(x) = 0\}, \\ g(D) &= \{x \in C : r_{g(D)}(x) = 0\}, \\ C &\subset \{x \in \mathbb{R}^n : r_D(x) \leq 0 \text{ and } r_{g(D)}(x) \geq 0\}, \end{aligned}$$

(A4.4) $\begin{cases} \langle \nabla r_D(x), f(x) \rangle > 0, & \forall x \in D, \\ \langle \nabla r_{g(D)}(x), f(x) \rangle > 0, & \forall x \in g(D). \end{cases}$

Remark 6: In Assumption 4, Item (A4.1) allows us to have uniform injectivity and continuity properties along solutions. Item (A4.2) ensures that no consecutive jumps can happen. Items (A4.3) and (A4.4) guarantee that a solution x of the system cannot stay inside of D or $g(D)$ during flows and also forbid the solution from leaving $C \cup D$ after a jump.

As done in [19], $\mathcal{X}$ being compact, we introduce a projection map $\Pi_{T(\mathcal{X})} : \mathbb{R}^{n_z} \to T(\mathcal{X})$ that satisfies

$$\Pi_{T(\mathcal{X})}(z) \in \left\{ z' : \underset{z' \in T(\mathcal{X})}{\text{argmin}} |z - z'| \right\}, \qquad \forall z \in \mathbb{R}^{n_z}. \quad (24)$$

We know from Item (G.2) of Definition 4 that the restriction of T to $C \backslash D$ is injective, so $T|_{C \backslash D}$ admits a left inverse on[2] $T(C \backslash D) = T(C) = T(C \cup D)$. We thus define T^{inv} :

[2]By Item (G.1) of Definition 4, $T(D) = T(g(D))$, and by Assumption 4, $D \subset C$ and $g(D) \subset C \backslash D$, so that $T(D) \subset T(C \backslash D)$, and $T(C \backslash D) = T(C \backslash D) \cup T(D) = T(C)$.

$\mathbb{R}^{n_z} \to (C \setminus D) \cap \mathcal{X}$ as

$$T^{\text{inv}}(z) = T\big|_{C\setminus D}^{-1}\left(\Pi_{T(\mathcal{X})}(z)\right), \tag{25}$$

which verifies

$$T^{\text{inv}}(T(x)) = x, \qquad \forall x \in (C \setminus D) \cap \mathcal{X}. \tag{26}$$

We then get the following result.

Theorem 1: Suppose Assumptions 1, 2, and 4 hold. Pick $n_z \in \mathbb{N}$, $(A, B) \in \mathbb{R}^{n_z \times n_z} \times \mathbb{R}^{n_z}$ such that $A + \rho I$ Hurwitz for ρ in Item (b) of Assumption 2, and $T : C \cup D \to \mathbb{R}^{n_z}$ such that the conclusion of Lemma 1, both Items (G.1) and (G.2) of Definition 4 hold. There exists a class-$\mathcal{K}$ function α and a positive scalar $\epsilon^\star$ such that for any $0 < \epsilon < \epsilon^\star$, there exists $t_\epsilon \geq 0$ such that for any solution x to system (1) initialized in $\mathcal{X}_0$ and any solution to

$$\dot{\hat{z}} = A\hat{z} + By, \qquad \hat{x} = T^{\text{inv}}(\hat{z}), \tag{27}$$

where y is the output of system (1), we have

$$|x(t, j) - \hat{x}(t)| < \epsilon, \qquad \forall (t, j) \in \operatorname{dom} x : t \geq t_\epsilon, t \in \tau_\alpha(\epsilon), \tag{28}$$

where $\tau_\alpha(\epsilon) = \mathbb{R}_{\geq 0} \setminus \bigcup_{j \in \operatorname{dom}_j x} [t_j - \alpha(\epsilon), t_j + \alpha(\epsilon)]$.

Proof: We follow the same ideas as in the proof of [19, Theorem 1]. Let x be a solution to system (1), initialized in $\mathcal{X}_0$. For $\epsilon > 0$, define

$$\mathcal{O}_{g(D)}(\epsilon) := \{x \in \mathcal{X} : d_{g(D)\cap\mathcal{X}}(x) < \epsilon\},$$
$$\mathcal{O}_D(\epsilon) := \{x \in \mathcal{X} : d_{D\cap\mathcal{X}}(x) < \epsilon\}.$$

Following [19, Lemma 2], relying on Assumption 4, there exists a class-$\mathcal{K}$ function α such that for all $\epsilon < \epsilon^\star$,

$$x(t, j) \notin \mathcal{O}_{g(D)}(\epsilon) \cup \mathcal{O}_D(\epsilon), \qquad \forall t \in \tau_\alpha(\epsilon). \tag{29}$$

Assume $\epsilon > 0$ is small enough such that $\mathcal{O}_D(\epsilon) \cap \mathcal{O}_{g(D)}(\epsilon) = \varnothing$, this is possible because $\mathcal{O}_D(\epsilon), \mathcal{O}_{g(D)}(\epsilon) \subset \mathcal{X}$ and $(D \cap \mathcal{X}) \cap (g(D) \cap \mathcal{X}) = \varnothing$. Now, since $\mathcal{O}_D(\epsilon)$ is open relative to $\mathcal{X}$ and $\mathcal{X}$ is compact, we get that $\mathcal{X}\setminus\mathcal{O}_D(\epsilon)$ is compact, so by Lemma 4, there exists a class-$\mathcal{K}$ function $\rho_{\epsilon,1}$ such that

$$|x_a - x_b| \leq \rho_{\epsilon,1}(|T(x_a) - T(x_b)|), \quad \forall x_a, x_b \in \mathcal{X}\setminus\mathcal{O}_D(\epsilon).$$

A similar reasoning with $\mathcal{O}_{g(D)}(\epsilon)$ shows that there exists a class-$\mathcal{K}$ function $\rho_{\epsilon,2}$ such that

$$|x_a - x_b| \leq \rho_{\epsilon,2}(|T(x_a) - T(x_b)|), \quad \forall x_a, x_b \in \mathcal{X}\setminus\mathcal{O}_{g(D)}(\epsilon).$$

Let $\rho_\epsilon(\cdot) := \max\{\rho_{\epsilon,1}(\cdot), \rho_{\epsilon,2}(\cdot)\}$, which is also a class-$\mathcal{K}$ function.

By the conclusion of Lemma 1, there exists a C^1 map $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_z}$ such that (16) and (17) hold. Since A is Hurwitz, any solution to (27) is such that $\lim_{t\to+\infty} |z(t) - \hat{z}(t)| = 0$, therefore, there exists $t_\epsilon \geq 0$ such that

$$|z(t) - \hat{z}(t)| < \frac{1}{2}\rho_\epsilon^{-1}(\epsilon), \qquad \forall t > t_\epsilon,$$

and thus, since $z(t) \in T(\mathcal{X})$ by (16) and Item (A4.1) of Assumption 4, for all $t > t_\epsilon$,

$$\begin{aligned}\left|z(t) - \Pi_{T(\mathcal{X})}(\hat{z}(t))\right| &\leq |z(t) - \hat{z}(t)| + \left|\hat{z}(t) - \Pi_{T(\mathcal{X})}(\hat{z}(t))\right| \\ &\leq |z(t) - \hat{z}(t)| + |z(t) - \hat{z}(t)| \\ &< \frac{1}{2}\rho_\epsilon^{-1}(\epsilon) + \frac{1}{2}\rho_\epsilon^{-1}(\epsilon) = \rho_\epsilon^{-1}(\epsilon).\end{aligned}$$

Also by (29), for any $(t, j) \in \operatorname{dom} x$ with $t \in \tau_\alpha(\epsilon)$, we have $x(t, j) \in \mathcal{X}\setminus(\mathcal{O}_D(\epsilon) \cup \mathcal{O}_{g(D)}(\epsilon)) = (\mathcal{X}\setminus\mathcal{O}_D(\epsilon)) \cap (\mathcal{X}\setminus\mathcal{O}_{g(D)}(\epsilon))$, and $\hat{x}(t) \in \mathcal{X} = (\mathcal{X}\setminus\mathcal{O}_D(\epsilon)) \cup (\mathcal{X}\setminus\mathcal{O}_{g(D)}(\epsilon))$. Thus $x(t, j)$ and $\hat{x}(t)$ belong together to either $\mathcal{X}\setminus\mathcal{O}_D(\epsilon)$ or $\mathcal{X}\setminus\mathcal{O}_{g(D)}(\epsilon)$. Therefore, for any $(t, j) \in \operatorname{dom} x$ with $t \in \tau_\alpha(\epsilon) \cap (t_\epsilon, +\infty)$, if $x(t, j), \hat{x}(t) \in \mathcal{X}\setminus\mathcal{O}_D(\epsilon)$ we have

$$\begin{aligned}|x(t, j) - \hat{x}(t)| &\leq \rho_{\epsilon,1}(|T(x(t, j)) - T(\hat{x}(t))|) \\ &\leq \rho_\epsilon(|T(x(t, j)) - T(\hat{x}(t))|) \\ &\leq \rho_\epsilon(|z(t) - \Pi_{T(\mathcal{X})}(\hat{z}(t))|) \\ &< \rho_\epsilon(\rho_\epsilon^{-1}(\epsilon)) < \epsilon.\end{aligned}$$

And if $x(t, j), \hat{x}(t) \in \mathcal{X}\setminus\mathcal{O}_{g(D)}(\epsilon)$, we have the same conclusion with $\rho_{\epsilon,2}$ instead of $\rho_{\epsilon,1}$. ■

This result proves the convergence of the estimation $\hat{x}$ given by (27) to the real solution x of system (1), but outside some intervals around the jump times t_j, whose length tends to zero as time goes to infinity.

V. APPLICATION TO STICK-SLIP ESTIMATION

We study the stick-slip phenomenon encountered, e.g., in rotary drilling [31]. In this process, a hole is created several kilometers into the ground by a *drill bit* connected to the surface actuators by a series of pipes called the *drill string*. Only surface real-time measurements are usually available, and the estimation of downhole conditions is of paramount importance to improve efficiency and reduce failure. A simplified model of the rotational dynamics consists of two masses (top and bottom) connected by a torsional spring of stiffness k. The equations of motion and the output read

$$\begin{cases}\dot{\Delta\theta} = \omega_1 - \omega_2 \\ \dot{\omega}_1 = -k\Delta\theta + u \\ \dot{\omega}_2 = \gamma k\Delta\theta + \text{Dry Friction}\end{cases} \qquad y = \omega_1, \tag{30}$$

where ω_1 (resp., ω_2) is the top (resp., bottom) velocity of the bit string, and $\Delta\theta$ models the distortion of the string. At the top, ω_1 is measured and regulated to ω_{ref} via a PI controller $u = -k_p(\omega_1 - \omega_{\text{ref}}) - k_i\eta$ where $\dot{\eta} = \omega_1 - \omega_{\text{ref}}$. To model the dry friction between the string and the walls, we use a 2-parameter *stiction* model which cannot be seen as a differential inclusion and requires some switching/hybrid logic [32, Section 4.2]. While $\omega_2 > 0$, the friction equals $-F_d$, and when the velocity decreases to 0 (with $\dot{\omega}_2 < 0$ and thus $\gamma k\Delta\theta \leq F_d$), ω_2 may either *stick* with $\omega_2 = 0$, if the external force $\gamma k\Delta\theta \in [-F_s, F_d]$ is not high enough to win over friction, or *slip* with $\omega_2 < 0$ if $\gamma k\Delta\theta \leq -F_s$. And symmetrically for $\omega_2 < 0$. Then, once it has stuck, ω_2 may slip again with $\omega_2 > 0$ (resp., $\omega_2 < 0$) only if the external force $\gamma k\Delta\theta$ overcomes static friction, i.e., becomes larger than F_s (resp., smaller than $-F_s$). High-gain PI controllers and large static-to-dynamic friction ratios typically induce undesirable stick-slip limit cycles, i.e., periodic trajectories alternating stick and slip phases.

All in all, these dynamics can be modeled using a hybrid system with state $x = (\Delta\theta, \omega_1, \omega_2, q, \eta, F_s, F_d)$, where q is a logic variable that is 0 in the stick phase, 1 in the forward slip phase, and -1 in the backward slip phase, with the flow dynamics

$$\dot{x} = \begin{cases} f_{\pm 1}(x), & \text{if } x \in C_1 \cup C_{-1} \\ f_0(x), & \text{if } x \in C_0, \end{cases} \tag{31a}$$

where

$$\begin{aligned} f_{\pm 1}(x) = (&\omega_1 - \omega_2, -k\Delta\theta - k_p(\omega_1 - \omega_{\text{ref}}) - k_i\eta, \\ &\gamma k \Delta\theta - qF_d, 0, \omega_1 - \omega_{\text{ref}}, 0, 0), \\ f_0(x) = (&\omega_1 - \omega_2, -k\Delta\theta - k_p(\omega_1 - \omega_{\text{ref}}) - k_i\eta, \\ &0, 0, \omega_1 - \omega_{\text{ref}}, 0, 0), \end{aligned}$$

with the flow sets

$$\begin{aligned} C_0 &= \left[-\frac{F_s}{\gamma k}, \frac{F_s}{\gamma k}\right] \times \mathbb{R} \times \{0\} \times \{0\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \\ C_1 &= \mathbb{R} \times \mathbb{R} \times [0, +\infty) \times \{1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \\ C_{-1} &= \mathbb{R} \times \mathbb{R} \times (-\infty, 0] \times \{-1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \end{aligned}$$

with the jump dynamics

$$x^+ = \begin{cases} g_{0\pm 1}(x), & \text{if } x \in D_{0\pm 1} \\ g_{\pm 10}(x), & \text{if } x \in D_{\pm 10} \\ g_{1-1}(x), & \text{if } x \in D_{1-1} \\ g_{-11}(x), & \text{if } x \in D_{-11}, \end{cases} \tag{31b}$$

where

$$\begin{aligned} g_{0\pm 1}(x) &= (\Delta\theta, \omega_1, \omega_2, \text{sign}(\Delta\theta), \eta, F_s, F_d), \\ g_{\pm 10}(x) &= (\Delta\theta, \omega_1, 0, 0, \eta, F_s, F_d), \\ g_{1-1}(x) &= (\Delta\theta, \omega_1, \omega_2, -1, \eta, F_s, F_d), \\ g_{-11}(x) &= (\Delta\theta, \omega_1, \omega_2, 1, \eta, F_s, F_d), \end{aligned}$$

with the jump sets

$$\begin{aligned} D_{0\pm 1} &= \left(\mathbb{R} \setminus \left(-\frac{F_s}{\gamma k}, \frac{F_s}{\gamma k}\right)\right) \times \mathbb{R} \times \{0\} \times \{0\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \\ D_{\pm 10} &= \left(\left[-\frac{F_s}{\gamma k}, \frac{F_d}{\gamma k}\right] \times \mathbb{R} \times \{0\} \times \{1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}\right) \\ &\quad \cup \left(\left[-\frac{F_d}{\gamma k}, \frac{F_s}{\gamma k}\right] \times \mathbb{R} \times \{0\} \times \{-1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}\right), \\ D_{1-1} &= \left(-\infty, -\frac{F_s}{\gamma k}\right] \times \mathbb{R} \times \{0\} \times \{1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \\ D_{-11} &= \left[\frac{F_s}{\gamma k}, +\infty\right) \times \mathbb{R} \times \{0\} \times \{-1\} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \end{aligned}$$

and the output

$$y = h(x) = \omega_1. \tag{31c}$$

Besides, η is known from controller design. We observe that solutions are t-forward complete and the output is continuous at jumps. However, solutions are non-unique, neither in forward nor in backward time, so Item (A1.2) of Assumption 1 does not hold. Still, we use the KKL-based gluing approach presented in this paper to estimate the full state, as well as the friction parameters (F_s, F_d). Note that the jump/switching times modeling mode changes are unknown. The parameters that we use correspond to a 2700-meter long inclined well and read $\gamma = 3.25$, $k = 0.08$ (s^{-1}), $k_p = 1.87$ (s^{-1}), $k_i = 7$ (s^{-2}), and $\omega_{\text{ref}} = 1$ (rad/s).

To implement the observer, we follow the approach of [24]. First, note that, from the successive derivatives of ω_1 and knowing η, $(\Delta\theta, \omega_2)$ are instantaneously observable during flows, F_d is instantaneously observable during flows in the slip mode, while F_s is visible when switching from stick to slip. Therefore, the system is backward distinguishable along solutions exhibiting stick-slip, with the information of F_s "hidden" but present in the full output trajectory.

As in Example 2, we perform 20000 simulations of the interconnection (31)-(17), with $A = 0.01\,\text{diag}(5, 6, 7, 8, 9, 10)$ and $B = (1, \ldots, 1)$, randomly initialized in $[-5, 5] \times [0, 2] \times [0, 2] \times \{0, 1\} \times [-2, 2] \times [0, 10] \times [0, 5]$ (with the constraint that $F_d \leq 0.5F_s$) and store 10 pairs $(x, T(x))$ in each simulation (after 100 units of time to get past the transient of z). Note that the eigenvalues in A are picked to be sufficiently slow to preserve the information of the stick-slip phenomenon. Then, regression is performed using a 3-layer Multi-Layer Perceptron (MLP) to compute $(\Delta\theta, \omega_2, q)$ from (z, η, ω_1), and two separate 10-layer MLPs to compute F_s and F_d from (z, η, ω_1) (each layer contains a few hundred neurons). Given the discrete nature of q, the q-output of the first MLP is quantified through a threshold at 0.5 (the case where $q = -1$ is so rare that it is neglected as an outlier). The results of a simulation of the observer (27) (with T^{inv} provided by the obtained NNs) are given in Figure 3, where $(\hat{F}_s, \hat{F}_q)$ are filtered to take into account their constant nature.

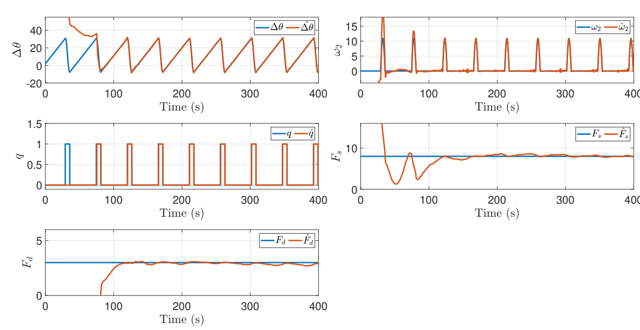

Fig. 3: Estimation results for the stick-slip system.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have shown the possibility of using the KKL paradigm to systematically construct a gluing function into continuous-time dynamics admitting an observer in the form of a filter of the output. This allows us to build an observer for general hybrid systems with continuous outputs at jumps and unknown jump times. This is illustrated numerically by application to stick-slip parameter estimation.

Future work includes the study of the regularity of the change of coordinates as well as developing systematic numerical schemes to learn numerical models of the (discontinuous) inverse of the gluing function.

APPENDIX

Lemma 3: Suppose Assumptions 1 and 2 hold. For every $x \in D$, we have

$$\breve{Y}(g(x), s) = \breve{Y}(x, s), \qquad \forall s \in (-\infty, t^+(x)). \tag{32}$$

Proof: Let $x \in D$. First we prove that $t^-(g(x)) = t^-(x)$. For this, let ϕ_1 and ϕ_2 be the maximal solutions to system (1) given in Item (A1.2) of Assumption 1 such that $\phi_1(0,0) = x$ and $\phi_2(0,0) = g(x)$, respectively. By uniqueness of solutions in Item (A1.2) of Assumption 1, $(0,1) \in \operatorname{dom} \phi_1$ and $\phi_1(0,1) = g(x)$. Define the function ϕ_2' on $\{(t,j) \in \mathbb{R} \times \mathbb{Z} : (t, j+1) \in \operatorname{dom} \phi_1\}$ as

$$\phi_2'(t,j) = \phi_1(t, j+1).$$

We observe that $(0,0) \in \operatorname{dom} \phi_2'$ since $(0,1) \in \operatorname{dom} \phi_1$ and $\phi_2'(0,0) = \phi_1(0,1) = g(\phi_1(0,0)) = g(x)$, therefore ϕ_2' is a solution to system (1) starting at $g(x)$, and satisfies $\inf \operatorname{dom}_t \phi_2' = \inf \operatorname{dom}_t \phi_1 = t^-(x)$. By the maximality of ϕ_2, we conclude that $t^-(g(x)) \leq t^-(x)$ and thus by the uniqueness of solutions, $\phi_2'(t,j) = \phi_2(t,j)$ for all $(t,j) \in \operatorname{dom} \phi_2'$. Now by contradiction, suppose that $t^-(g(x)) < t^-(x)$, then there exists $\tau \in \operatorname{dom}_t \phi_2 \setminus \operatorname{dom}_t \phi_1$, with $t^-(g(x)) < \tau < t^-(x)$. Consider the trajectory ϕ_1' defined on $\{(t,j) \in \mathbb{R} \times \mathbb{Z} : (t, j-1) \in \operatorname{dom} \phi_2\}$ as

$$\phi_1'(t,j) = \phi_2(t, j-1).$$

By a similar argument as before, we see that ϕ_1' is a solution to system (1) initialized as x such that $\operatorname{dom}_t \phi_1' = \operatorname{dom}_t \phi_2$. By the maximality of ϕ_1, we have $\operatorname{dom}_t \phi_1' \subset \operatorname{dom}_t \phi_1$, but this is a contradiction because $\tau \in \operatorname{dom}_t \phi_1' \setminus \operatorname{dom}_t \phi_1$. Therefore we conclude that $t^-(g(x)) = t^-(x)$. Similarly, $t^+(g(x)) = t^+(x)$ and $\operatorname{dom}_t \phi_1 = \operatorname{dom}_t \phi_2$.

Now we show that $\breve{Y}(x,s) = \breve{Y}(g(x), s)$ for all $s \in (-\infty, t^+(x))$. As noted before, $\phi_2'(s,j) = \phi_2(s,j)$ for all $(s,j) \in \operatorname{dom} \phi_2'$. By definition of ϕ_2', we have

$$\phi_1(s, j+1) = \phi_2(s,j), \qquad \forall (s,j) \in \operatorname{dom} \phi_2$$

so that

$$h(\phi_1(s, j+1)) = h(\phi_2(s,j)), \qquad \forall (s,j) \in \operatorname{dom} \phi_2.$$

By definition in (11), we have $Y(x,s) = Y(g(x), s)$ for all $s \in (t^-(x), , t^+(x))$. It follows then directly that for all $s \in (-\infty, t^+(x))$,

$$\begin{aligned} \breve{Y}(x,s) = \lim_{\tau^+ \to t^-(x)} Y(x,\tau) \qquad &= \lim_{\tau^+ \to t^-(x)} Y(g(x), \tau) \\ = \lim_{\tau^+ \to t^-(g(x))} Y(g(x), \tau) \quad &= \breve{Y}(g(x), s). \end{aligned}$$

The conclusion follows. ■

Lemma 4: Assume g is injective on D and $g(D) \subset C \setminus D$. Consider a gluing function T in the sense of Definition 4. Then for any compact set $\mathcal{M} \subset C$ satisfying $\mathcal{M} \cap D = \varnothing$ or $\mathcal{M} \cap g(D) = \varnothing$, the map T is injective on $\mathcal{M}$ and there exists a class-$\mathcal{K}$ function ρ such that

$$|x_a - x_b| \leq \rho(|T(x_a) - T(x_b)|), \forall (x_a, x_b) \in \mathcal{M} \times \mathcal{M}. \tag{33}$$

Proof: Let T be a gluing function for system (1) and let $\mathcal{M} \subset C$ be compact. For the injectivity of T on $\mathcal{M}$ we consider the following cases. Case 1: $\mathcal{M} \cap D = \varnothing$. Then we have $\mathcal{M} \subset C \setminus D$, and by Item (G.2) of Definition 4, T is injective on $\mathcal{M}$. Case 2: $\mathcal{M} \cap g(D) = \varnothing$. Assume there exist $x_a, x_b \in \mathcal{M}$ such that $T(x_a) = T(x_b)$. If both $x_a, x_b \in C \setminus D$, by Item (G.2) of Definition 4, we get $x_a = x_b$. Then, if $x_a \in D$ and $x_b \in C \setminus D$, we have $T(g(x_a)) = T(x_b)$ by Item (G.1) of Definition 4, and since $g(x_a) \in g(D) \subset C \setminus D$, it follows from Item (G.2) of Definition 4 that $x_b = g(x_a)$. But then, $x_b \in g(D)$ and thus $x_b \in \mathcal{M} \cap g(D)$, which contradicts the fact that $\mathcal{M} \cap g(D) = \varnothing$, so this case cannot happen. Finally, if both $x_a, x_b \in D$, we have from Item (G.1) of Definition 4, $T(x_a) = T(g(x_a)) = T(x_b) = T(g(x_b))$, and since T is injective on $g(D) \subset C \setminus D$, we get that $g(x_a) = g(x_b)$, and thus $x_a = x_b$ by injectivity of g on D. We conclude that T is injective on $\mathcal{M}$. Then, the existence of a class-$\mathcal{K}$ function satisfying (33) follows classically from the injectivity of T on a compact set (see e.g. [33, Lemma A.12]). ■

REFERENCES

[1] P. Bernard and R. G. Sanfelice, "Observer Design for Hybrid Dynamical Systems with Approximately Known Jump Times," *Automatica*, vol. 141, p. 110225, 2022.

[2] E. A. Medina and D. A. Lawrence, "State Estimation for Linear Impulsive Systems," *Annual American Control Conference*, pp. 1183–1188, 2009.

[3] A. Tanwani, H. Shim, and D. Liberzon, "Comments on "Observability of Switched Linear Systems: Characterization and Observer Design"," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, pp. 3396–3400, 2015.

[4] F. Deza, E. Busvelle, J. Gauthier, and D. Rakotopara, "High Gain Estimation for Nonlinear Systems," *Systems & Control Letters*, vol. 18, no. 4, pp. 295 – 299, 1992.

[5] T. Raff and F. Allgöwer, "Observers with Impulsive Dynamical Behavior for Linear and Nonlinear Continuous-Time Systems," *IEEE Conference on Decision and Control*, pp. 4287–4292, 2007.

[6] H. Ríos, J. Dávila, and A. R. Teel, "State Estimation for Linear Hybrid Systems with Periodic Jumps and Unknown Inputs," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 15, pp. 5966–5988, 2020.

[7] G. Q. B. Tran and P. Bernard, "Kalman-like Observer for Hybrid Systems with Linear Maps and Known Jump Times," in *62nd IEEE Conference on Decision and Control*, Singapore, Singapore, Dec. 2023.

[8] G. Q. B. Tran, P. Bernard, and R. G. Sanfelice, "Coupling Flow and Jump Observers for Hybrid Systems with Known Jump Times," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 6958–6963, 2023, 22nd IFAC World Congress.

[9] P. Bernard and R. Sanfelice, "On Notions of Detectability and Observers for Hybrid Systems," *IEEE Conference on Decision and Control*, pp. 5767–5772, 2020.

[10] B. Biemond, M. Heemels, R. G. Sanfelice, and N. van de Wouw, "Distance Function Design and Lyapunov Techniques for the Stability of Hybrid Trajectories," *Automatica*, vol. 73, pp. 38–46, 2016.

[11] J. B. Biemond, R. Postoyan, W. Heemels, and N. V. de Wouw, "Incremental Stability of Hybrid Dynamical Systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4094–4109, 2018.

[12] M. Rijnen, J. Biemond, N. van de Wouw, A. Saccon, and H. Nijmeijer, "Hybrid Systems With State-Triggered Jumps: Sensitivity-Based Stability Analysis With Application to Trajectory Tracking," *IEEE Trans. on Automatic Control*, vol. 65, no. 11, pp. 4568–4583, 2020.

[13] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid Dynamical Systems: Modeling Stability, and Robustness*. Princeton University Press Princeton, NJ, 2012.

[14] A. Balluchi, L. Benvenutia, M. D. D. Benedetto, and A. Sangiovanni-Vincentelli, "The Design of Dynamical Observers for Hybrid Systems: Theory and Application to an Automotive Control Problem," *Automatica*, vol. 49, no. 4, pp. 915–925, 2013.

[15] Z. Ping, C. Lee, and H. Shim, "Robust Estimation Algorithm for both Switching Signal and State of Switched Linear Systems," *Int. Journal of Control, Automation and Systems*, vol. 15, no. 1, pp. 95–103, 2017.

[16] W. Xiang, J. Xiao, and M. N. Iqbal, "Robust Observer Design for Nonlinear Uncertain Switched Systems under Asynchronous Switching," *Nonlinear Analysis: Hybrid Sys.*, vol. 6, no. 1, pp. 754–773, 2012.

[17] S. C. Shankar, G. Yang, and J. P. Hespanha, "State Estimation for Asynchronously Switched Sampled-Data Systems," in *IEEE Conference on Decision and Control*, 2022, pp. 1–7.

[18] P. Bernard and R. G. Sanfelice, "Semiglobal High-Gain Hybrid Observer for a Class of Hybrid Dynamical Systems with Unknown Jump Times," *IEEE Transactions on Automatic Control*, vol. 69, no. 9, pp. 5804 – 5819, 2024.

[19] J. Kim, H. Shim, and J. H. Seo, "State Estimation and Tracking Control for Hybrid Systems by Gluing the Domains," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 3026–3033, 2019.

[20] L. Menini and A. Tornambè, "State Immersion Observers for Mechanical Systems with Impacts," *IEEE Conference on Decision and Control*, pp. 7117–7122, 2016.

[21] V. Andrieu and L. Praly, "On the Existence of a Kazantzis–Kravaris/Luenberger Observer," *SIAM Journal on Control and Optimization*, vol. 45, no. 2, pp. 432–456, 2006.

[22] L. Brivadis, V. Andrieu, P. Bernard, and U. Serres, "Further Remarks on KKL Observers," *Sys. & Control Letters*, vol. 172, p. 105429, 2023.

[23] G. Q. B. Tran and P. Bernard, "Arbitrarily Fast Robust KKL Observer for Nonlinear Time-Varying Discrete Systems," *IEEE Transactions on Automatic Control*, pp. 1–16, 2023.

[24] L. Ramos, F. D. Meglio, V. Morgenthaler, L. Silva, and P. Bernard, "Numerical Design of Luenberger Observers for Nonlinear Systems," *IEEE Conference on Decision and Control*, pp. 5435–5442, 2020.

[25] M. Buisson-Fenet, L. Bahr, V. Morgenthaler, and F. Di Meglio, "Towards Gain Tuning for Numerical KKL Observers," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 4061–4067, 2023, 22nd IFAC World Congress.

[26] M. U. B. Niazi, J. Cao, X. Sun, A. Das, and K. H. Johansson, "Learning-based Design of Luenberger Observers for Autonomous Nonlinear Systems," in *2023 American Control Conference (ACC)*, 2023, pp. 3048–3055.

[27] H. K. Khalil, "Nonlinear Systems, Prentice-Hall," *Upper Saddle River, NJ*, vol. 3, 1996.

[28] J. Liu and A. R. Teel, "Hybrid Systems with Memory: Existence and Well-posedness of Generalized Solutions," *SIAM Journal on Control and Optimization*, vol. 56, no. 2, pp. 1011–1037, 2018.

[29] N. Wang and R. G. Sanfelice, "HyRRT-Connect: An Efficient Bidirectional Rapidly-Exploring Random Trees Motion Planning Algorithm for Hybrid Dynamical Systems," *IFAC-PapersOnLine*, vol. 58, no. 11, pp. 51–56, 2024, 8th IFAC Conference on Analysis and Design of Hybrid Systems ADHS 2024.

[30] G. Q. B. Tran, S. Garcia, P. Bernard, F. Di Meglio, and R. G. Sanfelice, "Towards Gluing KKL Observer for Hybrid Systems with Unknown Jump Times (Full Version)," in *63rd IEEE Conference on Decision and Control*, 2024. [Online]. Available: https://hal.science/hal-04685195

[31] B. Van de Vrande, D. Van Campen, and A. De Kraker, "An Approximate Analysis of Dry-Friction-Induced Stick-Slip Vibrations by a Smoothing Procedure," *Nonlinear Dynamics*, vol. 19, pp. 159–171, 1999.

[32] R. I. Leine and H. Nijmeijer, *Modelling of Dry Friction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 39–46.

[33] P. Bernard, *Observer Design for Nonlinear Systems*, L. N. on Control and I. Sciences, Eds. Springer, 2019, vol. 4.