Inverse-Optimal Safety Control for Hybrid Systems

Carlos A. Montenegro G.
University of California, Santa Cruz,
Santa Cruz, CA 95064, USA
camonten@ucsc.edu

Santiago J. Leudo
University of California, Santa Cruz,
Santa Cruz, CA 95064, USA
sjimen28@ucsc.edu

Ricardo G. Sanfelice
University of California, Santa Cruz,
Santa Cruz, CA 95064, USA
ricardo@ucsc.edu

ABSTRACT

We study control design methods to endow hybrid systems under disturbances with safety guarantees as an inverse-optimality problem. First, we provide sufficient conditions to guarantee input-to-state safety of a hybrid system with disturbance inputs only. Next, given a nominal feedback law, we show that a hybrid system, with inputs and disturbances, can be rendered input-to-state controlled safe under the existence of a control barrier function (CBF) using pointwise min-norm safeguarding feedback laws. Finally, we demonstrate that every CBF is a meaningful value function for a two-player zero-sum hybrid game in the context of safety, and that every pointwise min-norm safeguarding feedback law is optimal for such a game, even though its design is independent of any cost functional. The main results are illustrated in an example.

CCS CONCEPTS

• **Theory of computation → Mathematical optimization**; • **Computer systems organization → Robotic autonomy**; • **Information systems → Process control systems**.

KEYWORDS

Hybrid Systems, Input-to-State Safety, Conditional Invariance, Cost Evaluation, Inverse-Optimal Control

ACM Reference Format:

Carlos A. Montenegro G., Santiago J. Leudo, and Ricardo G. Sanfelice. 2025. Inverse-Optimal Safety Control for Hybrid Systems. In *28th ACM International Conference on Hybrid Systems: Computation and Control, May 6–9, 2025, Irvine, California.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3716863.3718044

1 INTRODUCTION

Designing autonomous decision-making processes for dynamical systems that exhibit both continuous-time and discrete-time behavior under the presence of adversarial actions is an active area of research. Applications in close proximity to humans demand formal safety guarantees to characterize the potential impact of the disturbances. Notions that relate inputs and the state of the system conveniently allow the design of control strategies to guarantee a nominal property, e.g., stability of a set of interest, and simultaneously reduce the effect of a disturbance that potentially drives the system's trajectories into an unsafe set. Applying only continuous-time results fails when seeking to measure and optimize the performance of the system at discrete-time events. Systems that interconnect physical and computational components or include continuous dynamics plus timers that expire, communication switches, impacts, or resets require a specialized control design formulation to address the effect of adversarial actions.

These observations led us to write a paper on the design of safety filters for *hybrid systems* under disturbances as a two-player zero-sum game where, following the framework in [4, 9], a player P_1 selects the control input to minimize a cost functional, while a player P_2 designs a disturbance to maximize it. In addition, an important link between stability and optimality is presented; the value function for a meaningful optimal stabilization problem is also a Lyapunov function for the closed-loop system. Given that the computation of the optimal strategy (known as the *saddle-point equilibrium*) and the optimal cost via solving Hamilton-Jacobi-Bellman-Isaacs (HJBI) equations represents a nontrivial task, an inverse optimality approach is proposed in this paper.

In the context of safety, barrier function-based control design methods have been shown to be suitable for safety-critical scenarios. In [1, 2], control Lyapunov function-based quadratic programs (CLF-QP) with constraints are solved online to synthesize controllers to perform locomotion and manipulation tasks. Safety guarantees are handled by including control barrier functions (CBFs) constraints in the optimization problem (CBF-CLF-QP). In addition, input-to-state safety notions trace back to [7] and [11] for continuous-time systems, and [14] considers the case of compositional input-to-state safety for nonlinear systems given as an interconnection of subsystems. The inverse optimal design of CBF-based safety filters for continuous-time systems with disturbances was studied in [12] with a two-player zero-sum formulation of a differential game.

To the best of our knowledge, no work has addressed the design of safety filters for hybrid systems as in [4] under disturbances as an inverse optimal problem. To close this gap, as an extension of the work in [13], this paper presents an approach to formulate such a problem as a two-player zero-sum game, as in [8–10]. In addition, we provide input-to-state safety guarantees through different families of feedback laws. Specifically, the main contributions of this paper are a formulation of two-player zero-sum games with hybrid constraints, following the framework in [4], encoding the design of safety filters. In Section 3, Theorem 3.6 provides sufficient conditions to guarantee input-to-state safety of a hybrid system with disturbance inputs only. In addition, Theorem 5.4 shows that a hybrid system, with inputs and disturbances, can be rendered input-to-state controlled safe under the existence of a CBF, whereas Theorem 5.6 and Theorem 5.8 prove that this can also be achieved

HSCC '25, May 6–9, 2025, Irvine, California
2025. ACM ISBN 979-8-4007-1504-4/2025/05.
https://doi.org/10.1145/3716863.3718044

with a pointwise min-norm continuous feedback law. Finally, in Theorem 6.3, we show that every CBF is a meaningful value function for the two-player zero-sum hybrid game in the context of safety, and that every pointwise min-norm feedback law is optimal for such meaningful game, even though its construction is independent of any cost functional or HJBI equation.

The remainder of the paper is structured as follows: in Section 2, we present preliminary concepts. A framework for the study of input-to-state safety for hybrid systems with disturbances is presented in Section 3. In Section 4, we state the problem addressed. In Section 5, we present results on input-to-state safety filters based on control barrier functions for hybrid systems. The main results of the paper on inverse-optimal safety filters are presented in Section 6, and a numerical application of the algorithm is presented in Section 7. Conclusions and future work are presented in Section 8.

Notation. The symbol $\mathbb{N}$ denotes the set of natural numbers including zero. The symbol $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_{\geq 0}$ denotes the set of nonnegative reals, and $\mathbb{R}_{>0}$ the set of positive reals. Given two vectors $x, y \in \mathbb{R}^n$, we use the equivalent notation $(x, y) = [x^\top y^\top]^\top$ and $\langle x, y\rangle$ denotes the Euclidean inner product. Given a nonempty set C, denote by $\operatorname{int} C$ its interior and by $\overline{C}$ its closure. A function $\alpha : [0, a) \to \mathbb{R}_{\geq 0}$ is a class-$\mathcal{K}$ function, also written as $\alpha \in \mathcal{K}$, if α is zero at zero, continuous, and strictly increasing. It is a class-$\mathcal{K}_\infty$, also denoted as $\alpha \in \mathcal{K}_\infty$, if it is class-$\mathcal{K}$, it is such that $a = \infty$, and $\lim_{r\to\infty} \alpha(r) = \infty$. Given a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ and a vector field $f : \mathbb{R}^n \to \mathbb{R}^n$, we denote the Lie derivative of h along f as $L_f h(x) := \langle \nabla h(x), f(x)\rangle$.

2 PRELIMINARIES

2.1 Hybrid Dynamical Systems

This paper considers hybrid systems modeled based on the framework in [4]. In such a framework, the continuous dynamics of the system are modeled by differential equations, while the discrete dynamics are modeled by difference equations. Based on this, a hybrid dynamical system $\mathcal{H}$ affine in the input $(u, w) = ((u_C, w_C), (u_D, w_D)) \in \mathbb{R}^{m_C} \times \mathbb{R}^{m_D} = \mathbb{R}^m$, where $u := (u_C, u_D) \in \mathbb{R}^{m_{C_u}} \times \mathbb{R}^{m_{D_u}} = \mathbb{R}^{m_u}$ is a control input and $w := (w_C, w_D) \in \mathbb{R}^{m_{C_w}} \times \mathbb{R}^{m_{D_w}} = \mathbb{R}^{m_w}$ is a disturbance, is defined as

$$\mathcal{H} : \begin{cases} \dot{x} = F(x, (u_C, w_C)) := f(x) + f_u(x)u_C + f_w(x)w_C \\ \qquad\qquad\qquad\qquad\qquad (x, (u_C, w_C)) \in C, \\ x^+ = G(x, (u_D, w_D)) := g(x) + g_u(x)u_D + g_w(x)w_D \\ \qquad\qquad\qquad\qquad\qquad (x, (u_D, w_D)) \in D \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state. The *flow map* $F : \mathbb{R}^n \times \mathbb{R}^{m_C} \to \mathbb{R}^n$ captures the continuous evolution of the system when the state and the input (u_C, w_C) are in the *flow set* C. The discrete evolution of the system is captured by the *jump map* $G : \mathbb{R}^n \times \mathbb{R}^{m_D} \to \mathbb{R}^n$ when the state and the input (u_D, w_D) are in the *jump set* D.

Since solutions to $\mathcal{H}$ in (1) can exhibit both continuous and discrete behavior, we use ordinary time t to determine the amount of flow elapsed and a counter $j \in \mathbb{N}$ that keeps track of the number of jumps that have occurred. Based on this parametrization of time, the concept of hybrid time domain, in which solutions to $\mathcal{H}$ are fully described, is proposed as follows.

Definition 2.1. (Hybrid time domain) *A set $\tilde{E} \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ is a compact hybrid time domain if*

$$\tilde{E} = \bigcup_{j=0}^{J-1} \left([t_j, t_{j+1}] \times \{j\}\right) \tag{2}$$

where $J \in \mathbb{N}$ and $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_J$. A set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ is a hybrid time domain if it is the union of a non decreasing sequence $E_1 \subset E_2 \subset E_3 \subset \ldots$ of compact hybrid time domains.

For a hybrid time domain E, notice that each element $(t, j) \in E$ denotes the elapsed hybrid time, which indicates that t seconds of flow time and j jumps have occurred. A hybrid signal is a function defined on a hybrid time domain. Given a hybrid signal ϕ and $j \in \mathbb{N}$, we define $I^j := \{t : (t, j) \in \operatorname{dom} \phi\}$.

Definition 2.2. (Hybrid arc) *A hybrid signal $\phi : \operatorname{dom} \phi \to \mathbb{R}^n$ is called a hybrid arc if, for each $j \in \mathbb{N}$, the function $t \mapsto \phi(t, j)$ is locally absolutely continuous on the interval I^j. A hybrid arc ϕ is compact if* $\operatorname{dom} \phi$ *is compact.*

Definition 2.3. (Hybrid Input) *A hybrid signal u is a hybrid input if, for each $j \in \mathbb{N}$, the function $t \mapsto u(t, j)$ is Lebesgue measurable and locally essentially bounded on the interval $I_u^j := \{t : (t, j) \in \operatorname{dom} u\}$.*

Let $\mathcal{X}$ be the set of hybrid arcs $\phi : \operatorname{dom} \phi \to \mathbb{R}^n$ and $\mathcal{U} \times \mathcal{W}$ the set of hybrid inputs $(u, w) : \operatorname{dom}(u, w) \to \mathbb{R}^{m_u} \times \mathbb{R}^{m_w}$, where $u = (u_C, u_D)$ and $w = (w_C, w_D)$. A solution to the hybrid system $\mathcal{H}$ in (1) is defined as follows.

Definition 2.4. (Solution to $\mathcal{H}$) *A pair $(\phi, (u, w))$ defines a solution to (1) if $\phi \in \mathcal{X}$, $u = (u_C, u_D) \in \mathcal{U}$, $w = (w_C, w_D) \in \mathcal{W}$,* $\operatorname{dom}\phi = \operatorname{dom}(u, w)$*, and*

- $(\phi(0, 0), (u_C(0, 0), w_C(0, 0))) \in \overline{C}$ *or* $(\phi(0, 0), (u_D(0, 0), w_D(0, 0))) \in D$,
- *For each $j \in \mathbb{N}$ such that I^j has a nonempty interior* $\operatorname{int} I^j$*, we have, for all $t \in \operatorname{int} I^j$,*
$$(\phi(t, j), (u_C(t, j), w_C(t, j))) \in C$$
and, for almost all $t \in I^j$,
$$\frac{d\phi}{dt}(t, j) = F(\phi(t, j), (u_C(t, j), w_C(t, j)))$$
- *For each $(t, j) \in \operatorname{dom} \phi$ such that $(t, j + 1) \in \operatorname{dom} \phi$,*

$$\begin{aligned} (\phi(t, j), (u_D(t, j), w_D(t, j))) &\in D \\ \phi(t, j + 1) &= G(\phi(t, j)), (u_D(t, j), w_D(t, j))) \end{aligned}$$

A solution $(\phi, (u, w))$ is a compact solution if $\operatorname{dom} \phi$ *is compact.*

The $\mathcal{L}^\infty$ norm of a hybrid signal $r = (r_C, r_D)$ is given by

$$\|r\|_{(t,j)} := \max\left\{\|r_C\|_{(t,j)}, \|r_D\|_{(t,j)}\right\} \tag{3a}$$

$$\|r_C\|_{(t,j)} := \max_{j' \leq j} \operatorname*{ess\,sup}_{t' \text{s.t.} (t', j') \in \operatorname{dom} r} \left|r(t', j')\right| \tag{3b}$$

$$\|r_D\|_{(t,j)} := \sup_{(t', j') \in \Gamma(r), t'+j' \leq t+j} \left|r(t', j')\right| \tag{3c}$$

where $\Gamma(r) := \{(t,j) \in \operatorname{dom} r : (t,j+1) \in \operatorname{dom} r\}$. For notational convenience, $\|r\|_{\#}$ denotes $\lim_{t+j\to N} \|r\|_{(t,j)}$, where $N = \sup_{(t,j)\in\operatorname{dom} r} t+j \in [0,+\infty]$. We say a solution $(\phi,(u,w))$ to $\mathcal{H}$ is maximal if it cannot be extended, and we say it is complete if $\operatorname{dom}\phi$ is unbounded. We denote by $\widehat{\mathcal{S}}_{\mathcal{H}}(M)$ the set of solutions $(\phi,(u,w))$ to $\mathcal{H}$ as in (1) such that $\phi(0,0) \in M$. The set $\mathcal{S}_{\mathcal{H}}(M) \subset \widehat{\mathcal{S}}_{\mathcal{H}}(M)$ denotes all maximal solutions from M. We define $\operatorname{dom}_t \phi := \{t \in \mathbb{R}_{\geq 0} : \exists j \text{ s.t. } (t,j) \in \operatorname{dom}\phi\}$, $\operatorname{dom}_j \phi := \{j \in \mathbb{N}_{\geq 0} : \exists t \text{ s.t. } (t,j) \in \operatorname{dom}\phi\}$, $\sup_t \operatorname{dom}\phi := \sup \operatorname{dom}_t \phi$ and $\sup_j \operatorname{dom}\phi := \sup \operatorname{dom}_j \phi$. We define the projections of C and D onto $\mathbb{R}^n$, respectively, as

$$\Pi(C) := \{\xi \in \mathbb{R}^n : \exists (u_C, w_C) \text{ s.t. } (\xi,(u_C,w_C)) \in C\}$$
$$\Pi(D) := \{\xi \in \mathbb{R}^n : \exists (u_D, w_D) \text{ s.t. } (\xi,(u_D,w_D)) \in D\}.$$

Well-posed dynamical systems refer to a class of dynamical systems where the solutions enjoy very useful structural properties [5]. A hybrid system $\mathcal{H}$ as in (1) is well-posed if the hybrid basic conditions hold [18, Lemma 2.21].

2.2 Hybrid Systems with Disturbances

Consider the hybrid system resulting from assigning the control input u of $\mathcal{H}$ as in (1) to a given feedback law $\kappa := (\kappa_C,\kappa_D) : \mathbb{R}^n \to \mathbb{R}^{m_{C_u}} \times \mathbb{R}^{m_{D_u}}$ and with disturbance input $w \in \mathbb{R}^{m_{C_w}} \times \mathbb{R}^{m_{D_w}}$, namely

$$\mathcal{H}_\kappa : \begin{cases} \dot{x} = F(x,(\kappa_C(x),w_C)) =: F_\kappa(x,w_C) & (x,w_C) \in C_\kappa \\ x^+ = G(x,(\kappa_D(x),w_D)) =: G_\kappa(x,w_D) & (x,w_D) \in D_\kappa \end{cases} \quad (4)$$

where $C_\kappa := \{(x,w_C) \in \mathbb{R}^n \times \mathbb{R}^{m_{C_w}} : (x,(\kappa_C(x),w_C)) \in C\}$ and $D_\kappa := \{(x,w_D) \in \mathbb{R}^n \times \mathbb{R}^{m_{D_w}} : (x,(\kappa_D(x),w_D)) \in D\}$.

Recall that $\mathcal{X}$ is the set of hybrid arcs $\phi : \operatorname{dom}\phi \to \mathbb{R}^n$ and $\mathcal{W}$ is the set of hybrid inputs $w : \operatorname{dom} w \to \mathbb{R}^{m_w}$, where $w = (w_C, w_D)$. We define a solution to $\mathcal{H}_\kappa$ as follows.

Definition 2.5. (Solution to $\mathcal{H}_\kappa$) *A pair (ϕ,w) defines a solution to $\mathcal{H}_\kappa$ as in (4) if $\phi \in \mathcal{X}$, $w = (w_C,w_D) \in \mathcal{W}$, $\operatorname{dom}\phi = \operatorname{dom} w$, and*

- $(\phi(0,0), w_C(0,0)) \in \overline{C_\kappa}$ *or* $(\phi(0,0), w_D(0,0)) \in D_\kappa$,
- *For each $j \in \mathbb{N}$ such that I^j_ϕ has a nonempty interior $\operatorname{int} I^j_\phi$, we have, for all $t \in \operatorname{int} I^j_\phi$,*
$$(\phi(t,j), w_C(t,j)) \in C_\kappa$$
and, for almost all $t \in I^j_\phi$,
$$\frac{d\phi}{dt}(t,j) = F(\phi(t,j),(\kappa_C(\phi(t,j)), w_C(t,j)))$$
- *For each $(t,j) \in \operatorname{dom}\phi$ such that $(t,j+1) \in \operatorname{dom}\phi$,*
$$(\phi(t,j), w_D(t,j)) \in D_\kappa$$
$$\phi(t,j+1) = G(\phi(t,j),(\kappa_D(\phi(t,j)), w_D(t,j)))$$

A solution pair (ϕ,w) is a compact solution if $\operatorname{dom}\phi$ is compact.

Given a solution pair (ϕ,w) to $\mathcal{H}_\kappa$, the component ϕ is referred to as the state trajectory. We say that the hybrid closed-loop system with disturbances, namely $\mathcal{H}_\kappa$ as in (4), results from assigning the input u of $\mathcal{H}$ in (1) to a feedback law κ.

Finally, we define the projections of C_κ and D_κ onto $\mathbb{R}^n$, respectively, as

$$\Pi(C_\kappa) := \{\xi \in \mathbb{R}^n : \exists w_C \text{ s.t. } (\xi, w_C) \in C_\kappa\}$$
$$\Pi(D_\kappa) := \{\xi \in \mathbb{R}^n : \exists w_D \text{ s.t. } (\xi, w_D) \in D_\kappa\}.$$

3 INPUT-TO-STATE SAFETY FOR HYBRID SYSTEMS

Given a hybrid system $\mathcal{H} = (C,F,G,D)$ as in (1) and a feedback law κ, we formulate conditions guaranteeing that every state trajectory of the resulting closed-loop system $\mathcal{H}_\kappa$ that starts in a closed set $K \subset \mathbb{R}^n$ remains close to K under the presence of a disturbance $w = (w_C, w_D)$, where the closeness to K depends on the size of w. For this purpose, following [11], we use the notion of input-to-state safety (ISSf) to guarantee that a larger set containing K is conditionally invariant for $\mathcal{H}_\kappa$ with respect to w and K. We introduce the following definitions of invariance and safety.

Definition 3.1. (Conditional pre-invariance with disturbances) *Given a feedback law κ, a set $S \subset \mathbb{R}^n$ is said to be conditionally pre-invariant for $\mathcal{H}_\kappa$ in (4) with respect to the disturbance w and the set $K \subset S$ if each $(\phi,w) \in \mathcal{S}_{\mathcal{H}_\kappa}(K)$ is such that $\phi(t,j) \in S$ for all $(t,j) \in \operatorname{dom}\phi$.*

Barrier functions (BFs) serve as a synthesis tool to guarantee invariance of a set of interest, see, e.g., [1] and [19]. In the context of safety, given an unsafe set $X_u \subset \Pi(C) \cup \Pi(D) \cup G(D)$ and a continuous function $B : \mathbb{R}^n \to \mathbb{R}$ such that $B(x) > 0$ for all $x \in X_u$, we define a set K as the zero-sublevel set of B restricted to $\Pi(C)\cup\Pi(D)$, i.e.,

$$K := \{x \in \Pi(C) \cup \Pi(D) : B(x) \leq 0\} \quad (5)$$

which is closed when $\Pi(C) \cup \Pi(D)$ is closed.

The following definition introduces the notion of safety for hybrid systems with disturbance inputs.

Definition 3.2. (Input-to-state safety) *Consider a closed set $K \subset \mathbb{R}^n$ defined by a function $B : \operatorname{dom} B \to \mathbb{R}$ as in (5), and a feedback law κ defining the hybrid closed-loop system $\mathcal{H}_\kappa$ as in (4). If there exist $\bar{w} \geq 0$ and $\rho \in \mathcal{K}_\infty$ such that*

$$\begin{aligned} &(\phi,w) \in \mathcal{S}_{\mathcal{H}_\kappa}(K), \|w\|_{\#} \leq \bar{w} \\ &\qquad \Rightarrow B(\phi(t,j)) \leq \rho(\bar{w}) \quad \forall (t,j) \in \operatorname{dom}\phi \end{aligned} \quad (6)$$

where the function ρ is referred to as the ISSf gain, then the system $\mathcal{H}_\kappa$ is $\bar{w}$-small-input input-to-state safe ($\bar{w}$-small-input ISSf) with respect to the disturbance w and the set K.

Notice that Definition 3.1 and Definition 3.2 do not require maximal solutions to be complete, for which we employ the prefix 'pre-', for more details, see [3, 16, 17]. In addition, observe that, from the construction of K in (5) and the properties of the barrier function B, it follows that

$$K \subset (\Pi(C) \cup \Pi(D) \cup G(D)) \setminus X_u. \quad (7)$$

Small-input ISSf is strengthened to ISSf if (6) holds for arbitrary large $\bar{w}$. In addition, small-input ISSf resembles the notion of safety in [16] when $\bar{w} = 0$.

Remark 3.3. (Safety and invariance) *It is immediate that the system $\mathcal{H}_\kappa$ is $\bar{w}$-small-input ISSf with respect to w and K if and only if there exist $\bar{w} \geq 0$ and $\rho \in \mathcal{K}_\infty$ such that the set $K_d(\bar{w}) \supset K$ defined as*

$$K_d(\bar{w}) := \{x \in \Pi(C) \cup \Pi(D) : B(x) - \rho(\bar{w}) \leq 0\} \tag{8}$$

is conditionally pre-invariant for $\mathcal{H}_\kappa$ with respect to w and K.

Remark 3.4. (Connections with the literature) *In this work, we are interested in characterizing the ISSf property in Definition 3.2 for hybrid systems so that we can guarantee that, under the worst-case disturbance w, trajectories starting from K do not reach the unsafe set X_u.*

- *Selection of a finite $\bar{w}$: following [11], in which a connection between safety and conditional invariance of a set is established in terms of an upper bound on the disturbances, the notion of input-to-state safety herein relies on a similar approach. In the context of robust safety for continuous-time systems, previous work, such as [6], considers disturbances bounded by a known constant to design feedback laws that robustly stabilize the system while rendering a set of interest forward invariant. Notice that for $\mathcal{H}_\kappa$ to be w-robustly safe[1] with respect to (K, X_u), it is sufficient to find a finite*

$$\begin{aligned} 0 \leq \bar{w} \leq v^* := \arg\sup_{v>0} \quad & v \\ \text{subject to} \quad & K_d(v) \cap X_u = \emptyset \end{aligned} \tag{9}$$

  and κ such that $K_d(\bar{w}) \supset K$ in (8) is conditionally pre-invariant for $\mathcal{H}_\kappa$ with respect to w and K. Thus, if $\mathcal{H}_\kappa$ is w-robustly safe with respect to (K, X_u), then it is $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K satisfying (7). Furthermore, when $\bar{w} = 0$, $\bar{w}-$small-input ISSf of the system $\mathcal{H}_\kappa$ with no disturbances implies that each $(\phi, w) \in \mathcal{S}_{\mathcal{H}_\kappa}(K)$ is such that $\phi(t, j) \in K$ for all $(t, j) \in \operatorname{dom} \phi$.
- *Existing notions of ISSf with respect to disturbances: a version of Definition 3.2 was presented in [12] for continuous-time systems. The $\mathcal{KL}$ bound therein accounts for solutions that start outside of K, case we do not consider in this work. Compared to [12], the set K in (7) is defined following an opposite sign convention, namely, K is defined as the zero-sublevel set of B (contrary to being defined as the zero-superlevel set of h in [12]). Without loss of generality, (5) relies on an upper bound $\bar{w}$ on disturbances, which can be conveniently chosen to resemble (7) in [12].*

The following sections provide sufficient conditions to establish input-to-state safety using barrier functions to guarantee that a larger set containing K as in (5) is conditionally invariant for $\mathcal{H}_\kappa$ with respect to w and K. To this end, we formalize the notion of an input-to-state safety barrier function candidate in the next definition, where we impose conditions on a function B such that the set K can be defined as in (5) and conditionally invariance can be satisfied as well.

Definition 3.5. (ISSf barrier function candidate) *Consider a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1). The function $B : \operatorname{dom} B \to \mathbb{R}$ and the sets $K \subset K_i \subset \mathbb{R}^n$ define an ISSf barrier function (ISSf-BF) candidate for $\mathcal{H}$ with respect to (K, K_i) if the following conditions hold:*

1) *$\overline{\Pi(C)} \cup \Pi(D) \cup G(D) \subset \operatorname{dom} B$ and $K_i \subset \Pi(C) \cup \Pi(D)$;*
2) *for some open set $\mathcal{V}$ containing an open neighborhood of K_i, B is continuously differentiable on $(\mathcal{V} \setminus K_i) \cap \overline{\Pi(C)}$;*
3) *$B(x) > 0$ for all $x \in \left(\overline{\Pi(C)} \cup \Pi(D)\right) \setminus K$;*
4) *$B(x) \leq 0$ for all $x \in K$.*

Notice that $K_i \supset K$ in Definition 3.5 is the set we aim to render invariant, whose role will be played by $K_d(\bar{w})$ in the following results .

Theorem 3.6. (ISSf using a barrier function candidate) *Given a hybrid system $\mathcal{H}$ as in (1), a closed set $K \subset \mathbb{R}^n$, and a feedback law $\kappa = (\kappa_C, \kappa_D)$ defining the closed-loop system $\mathcal{H}_\kappa = (C_\kappa, F_\kappa, D_\kappa, G_\kappa)$ as in (4) with disturbance $w = (w_C, w_D)$, suppose B is an ISSf-BF candidate for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, and let $\mathcal{V}$ be an open set containing an open neighborhood of $K_d(\bar{w})$. If there exist $\alpha_C \geq 0$ and $\alpha_D \in [0, 1]$ such that*

$$\begin{aligned} &\langle \nabla B(x), F_\kappa(x, w_C) \rangle \leq -\alpha_C B(x) \\ &\qquad \forall (x, w_C) \in C_\kappa : x \in \mathcal{V} \setminus K_d(\bar{w}), |w_C| \leq \bar{w} \end{aligned} \tag{10a}$$

$$\begin{aligned} &B(G_\kappa(x, w_D)) - B(x) \leq -\alpha_D (B(x) - \rho(\bar{w})) \\ &\qquad \forall (x, w_D) \in D_\kappa : x \in K_d(\bar{w}), |w_D| \leq \bar{w} \end{aligned} \tag{10b}$$

$$G_\kappa(\tilde{D}_\kappa) \subset \Pi(C_\kappa) \cup \Pi(D_\kappa) \tag{10c}$$

where $\tilde{D}_\kappa := \{(x, w_D) \in D_\kappa : x \in K_d(\bar{w})\}$, then $\mathcal{H}_\kappa$ is $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K, as in Definition 3.2.

Notice that (10c) requires that G_κ does not map the state x outside of $\Pi(C_\kappa) \cup \Pi(D_\kappa)$ after a jump, case in which $K_d(\bar{w})$ would not be conditionally pre-invariant for K. Therefore, we restrict our attention to a family of hybrid systems that satisfy the next assumption.

Assumption 3.7. *Consider a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1) and a closed set $K \subset \mathbb{R}^n$. Let B be an ISSf-BF candidate for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} > 0$. Suppose that, for all $(x, (u_D, w_D)) \in D$ such that $x \in K_d(\bar{w})$, the following holds*

$$G(x, (u_D, w_D)) \subset \Pi(C) \cup \Pi(D).$$

4 PROBLEM STATEMENT

Consider the system $\mathcal{H}$ in (1), with the feedback law $\kappa = (\kappa_C, \kappa_D)$ assigning values to the control input u, the disturbance input $w \in \mathcal{W}$, and an unsafe set $X_u \subset \mathbb{R}^n$. We consider the case when the feedback κ is the sum of a given nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$ capturing desired properties, referred to as *uncertified* objectives, such as rendering a set asymptotically stable for $\mathcal{H}$, and a safeguarding feedback law $\widehat{\kappa}$. We say $\mathcal{H}$ is $\bar{w}$-small-input input-to-state

[1] Following [15, 18], the system $\mathcal{H}_\kappa$ is said to be w-robustly safe with respect to (K, X_u) if each $(\phi, w) \in \mathcal{S}_{\mathcal{H}_\kappa}(K)$ is such that $\phi(t, j) \in \mathbb{R}^n \setminus X_u$ for all $(t, j) \in \operatorname{dom} \phi$.

controlled safe when the corresponding closed-loop system $\mathcal{H}_\kappa$ as in (4) is $\bar{w}$-small-input ISSf.

In this paper, we address the problem of designing the safeguarding feedback law $\widehat{\kappa}$ that not only renders $\mathcal{H}$ $\bar{w}$-small-input input-to-state controlled safe but also solves a zero-sum hybrid game. We use a continuous function $x \mapsto B(x)$ defining a barrier function candidate, leading to K given in (5) such that (7) holds, to guarantee that state trajectories starting in K never reach X_u. Specifically, we seek the existence of $\rho \in \mathcal{K}_\infty$ such that every $(\phi, (u, w)) \in \mathcal{S}_\mathcal{H}(\Pi(C) \cup \Pi(D))$, with input u given by $\operatorname{dom}\phi \ni (t, j) \mapsto u(t, j) = \kappa(\phi(t, j)) = \bar{\kappa}(\phi(t, j)) + \widehat{\kappa}(\phi(t, j))$, satisfies (6) for all $(t, j) \in \operatorname{dom}\phi$. This objective[2] is attained by considering the corresponding closed-loop system $\mathcal{H}_\kappa$ and solving the following problem.

Problem 4.1. (Inverse-Optimal Safety Filter) *Given a closed set $K \subset \mathbb{R}^n$ and an uncertified nominal feedback law $\bar{\kappa}$, design a safeguarding feedback law $\widehat{\kappa}$ that renders the corresponding hybrid closed-loop system $\mathcal{H}_\kappa$ $\bar{w}$-small-input input-to-state safe with respect to the disturbance w and the set K. In addition, determine the cost functional that $\widehat{\kappa}$ minimizes under the worst-case disturbance w.*

Remark 4.2. (Relation to the literature) *A version of Problem 4.1 was solved in [12] for continuous-time systems without constraints, i.e., the case in which $\mathcal{H} = (\mathbb{R}^n \times \mathbb{R}^m, F, \emptyset, \star)$, where $\star$ denotes an arbitrary jump map, and K in (5) is defined as $K := \{x \in \mathbb{R}^n \mid B(x) \geq 0\}$.*

5 INPUT-TO-STATE SAFETY FILTERS

In this section, we address the first part of Problem 4.1 by using ISSf control barrier functions (ISSf-CBFs) as a synthesis tool to guarantee safety of a hybrid system. First, we introduce definitions and preliminary results on ISSf-CBFs for hybrid systems with disturbances.

5.1 Input-to-State Safety Control Barrier Functions

Given a hybrid system $\mathcal{H} = (C, F, G, D)$ we define the projection of $\star$ onto $\mathbb{R}^n \times \mathbb{R}^{m_{\star w}}$, for each $\star \in \{C, D\}$, as

$$\Pi_{u_\star}(\star) := \{(x, w_\star) : \exists u_\star \text{ s.t. } (x, (u_\star, w_\star)) \in \star\}.$$

In addition, we define, for each $\star \in \{C, D\}$, the set of admissible control inputs during flows ($\star = C$) and during jumps ($\star = D$) at each state and disturbance as

$$\Psi_\star(x, w_\star) := \left\{u_\star \in \mathbb{R}^{m_{\star u}} : (x, (u_\star, w_\star)) \in \star\right\}.$$

Definition 5.1. (ISSf-CBF with respect to disturbances) *Given a system $\mathcal{H} = (C, F, D, G)$ as in (1) and a closed set K, suppose B is an ISSf-BF candidate for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$. Let $\mathcal{V}$ be an open set containing an open neighborhood of $K_d(\bar{w})$. We say that B is an ISSf-control barrier function (ISSf-CBF) for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$ if there exist $\alpha_C \geq 0$ and $\alpha_D \in [0, 1]$ such that*

$$\begin{aligned} \inf_{u_C \in \Psi_C(x, w_C)} \langle \nabla B(x), F(x, (u_C, w_C))\rangle \leq -\alpha_C B(x) \\ \forall (x, w_C) \in \Pi_{u_C}(C) : x \in \mathcal{V}\backslash K_d(\bar{w}), |w_C| \leq \bar{w} \end{aligned} \tag{11a}$$

$$\begin{aligned} \inf_{u_D \in \Psi_D(x, w_D)} B(G(x, (u_D, w_D))) - B(x) \\ \leq -\alpha_D(B(x) - \rho(\bar{w})) \\ \forall (x, w_D) \in \Pi_{u_D}(D) : x \in K_d(\bar{w}), |w_D| \leq \bar{w}. \end{aligned} \tag{11b}$$

The following results are used to establish a connection between the existence of an ISSf-CBF and a feedback law that renders the hybrid closed-loop system ISSf. To characterize the effect of inputs in the safety conditions at jumps, we restrict our attention to a family of systems and barrier functions that satisfy the next assumption.

Assumption 5.2. *Given a system $\mathcal{H} = (C, F, D, G)$ as in (1) and a function $B : \mathbb{R}^n \to \mathbb{R}$, suppose there exist functions $\widehat{B}_{Lu} : \Pi(D) \to \mathbb{R}^{m_{Du}}$ and $\widehat{B}_{Lw} : \Pi(D) \to \mathbb{R}^{m_{Dw}}$ such that, for all $(x, (u_D, w_D)) \in D$,*

$$\begin{aligned} B(G(x, u_D)) &= B(g(x) + g_u(x)u_D + g_w(x)w_D) \\ &\leq B(g(x)) + \widehat{B}_{Lu}(x)u_D + \widehat{B}_{Lw}(x)w_D. \end{aligned} \tag{12}$$

Lemma 5.3. (Equivalent ISSf conditions) *Given a system $\mathcal{H} = (C, F, D, G)$ as in (1) and a closed set $K \subset \mathbb{R}^n$, suppose B is an ISSf-CBF for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, $\mathcal{V}$ is an open set containing an open neighborhood of $K_d(\bar{w})$, $\alpha_C \geq 0$, and $\alpha_D \in [0, 1]$. The tuple $(B, \rho, \bar{w})$ satisfies (11a) if and only if*

$$L_{f_u}B(x) = 0, x \in (\mathcal{V} \setminus K_d(\bar{w})) \cap \Pi(C) = 0 \Rightarrow \omega_C(x) \leq 0 \tag{13a}$$

where

$$\omega_C(x) := L_f B(x) + \alpha_C B(x) + |L_{f_w}B(x)|\rho^{-1}(B(x)) \tag{13b}$$

and, under Assumption 5.2, satisfies (11b) if and only if

$$\widehat{B}_{Lu}(x) = 0, x \in K_d(\bar{w}) \cap \Pi(D) \Rightarrow \omega_D(x) \leq 0 \tag{14a}$$

where

$$\omega_D(x) := B(g(x)) - B(x) + \alpha_D(B(x) - \rho(\bar{w})) + |\widehat{B}_{Lw}(x)|\bar{w}. \tag{14b}$$

Theorem 5.4. (ISSf-CBF Sontag-like formula) *Under Assumption 5.2, suppose that there exists an ISSf-CBF B for $\mathcal{H} = (C, F, D, G)$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, and that Assumption 3.7 is satisfied. Then, the feedback law $\widehat{\kappa}_S = (\widehat{\kappa}_{SC}, \widehat{\kappa}_{SD})$, with*

$$\widehat{\kappa}_{SC}(x) := \begin{cases} L_{f_u}B(x)\kappa_{SC}(x) & \text{if} \quad L_{f_u}B(x) \neq 0 \\ 0 & \text{if} \quad L_{f_u}B(x) = 0 \end{cases} \tag{15a}$$

where

$$\kappa_{SC}(x) := \frac{-\omega_C(x) - \sqrt{\omega_C^2(x) + |L_{f_u}B(x)|^4}}{|L_{f_u}B(x)|^2}, \tag{15b}$$

ω_C defined in (13b), and

$$\widehat{\kappa}_{SD}(x) := \begin{cases} \widehat{B}_{Lu}(x)\kappa_{SD}(x) & \text{if} \quad \widehat{B}_{Lu}(x) \neq 0 \\ 0 & \text{if} \quad \widehat{B}_{Lu}(x) = 0 \end{cases} \tag{16a}$$

where

$$\kappa_{SD}(x) := \frac{-\omega_D(x) - \sqrt{\omega_D^2(x) + |\widehat{B}_{Lu}(x)|^4}}{|\widehat{B}_{Lu}(x)|^2}, \tag{16b}$$

[2] Notice that the safeguarding map $\widehat{\kappa}$ plays the role of a filter that shall be zero when (6) is satisfied by the nominal feedback law $\bar{\kappa}$.

ω_D defined in (14b), renders the resulting hybrid closed-loop system $\mathcal{H}_{\widehat{\kappa}_S}$ as in (4) $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K.

5.2 Input-to-State Safety QP Filter

Given a nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$, we endow a system $\mathcal{H}$ with an input-to-state safety property by solving a quadratic program (QP) problem in terms of an ISSf control barrier function.

Let $\mathcal{V}$ be an open set containing an open neighborhood of $K_d(\bar{w})$. Given $\alpha_C \geq 0$, we define

$$\omega_C(x) := L_{f+f_u\bar{\kappa}_C}B(x) + |L_{f_w}B(x)|\rho^{-1}(B(x)) + \alpha_C B(x) \quad (17)$$

for all $x \in \mathcal{V} \cap \Pi(C)$ and introduce the following QP:

$$\begin{aligned} \widehat{\kappa}_{C_{QP}}(x) = \underset{v\in\mathbb{R}^{m_{C_u}}}{\arg\min} \quad & |v|^2 \\ \text{subject to} \quad & L_{f_u}B(x)v \leq -\omega_C(x). \end{aligned} \quad (18)$$

whose closed-form solution is given by

$$\widehat{\kappa}_{C_{QP}}(x) := \begin{cases} -\dfrac{\max\{0, \omega_C(x)\}}{|L_{f_u}B(x)|^2} L_{f_u}B(x) & \text{if } L_{f_u}B(x) \neq 0 \\ 0 & \text{if } L_{f_u}B(x) = 0. \end{cases} \quad (19)$$

Similar to Assumption 5.2, to characterize the effect of the QP filter and the disturbance in the safety conditions at jumps, we impose the following assumption to upper bound the growth of B by a control-affine function.

Assumption 5.5. *Consider a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1), a feedback law $\kappa := (\kappa_C, \kappa_D) = (\bar{\kappa}_C + \widehat{\kappa}_C, \bar{\kappa}_D + \widehat{\kappa}_D)$, and a function $B : \mathbb{R}^n \to \mathbb{R}$. Suppose that there exist functions $\widehat{B}_{Lu} : \Pi(D) \to \mathbb{R}^{m_{D_u}}$ and $\widehat{B}_{Lw} : \Pi(D) \to \mathbb{R}^{m_{D_w}}$ such that, for all $(x, (\kappa_D(x), w_D)) \in D$,*

$$\begin{aligned} B(G(x, (\kappa_D(x), w_D))) &= B\big(g(x) + g_u(x)\kappa_D(x) + g_w(x)w_D\big) \\ &\leq B\big(g(x) + g_u(x)\bar{\kappa}_D(x)\big) \\ &\quad + \widehat{B}_{Lu}(x)\widehat{\kappa}_D(x) + \widehat{B}_{Lw}(x)w_D. \end{aligned} \quad (20)$$

Similarly, given $\alpha_D \in [0, 1]$, under Assumption 5.5, we define

$$\begin{aligned} \omega_D(x) := & B\big(g(x) + g_u(x)\bar{\kappa}_D(x)\big) - B(x) \\ & + |\widehat{B}_{Lw}(x)|\bar{w} + \alpha_D(B(x) - \rho(\bar{w})) \end{aligned} \quad (21)$$

for all $x \in \Pi(D) \cap K_d(\bar{w})$, and introduce the following QP:

$$\begin{aligned} \widehat{\kappa}_{D_{QP}}(x) = \underset{v\in\mathbb{R}^{m_{D_u}}}{\arg\min} \quad & |v|^2 \\ \text{subject to} \quad & \widehat{B}_{Lu}(x)v \leq -\omega_D(x) \end{aligned} \quad (22)$$

whose closed-form solution is expressed

$$\widehat{\kappa}_{D_{QP}}(x) := \begin{cases} -\dfrac{\max\{0, \omega_D(x)\}}{|\widehat{B}_{Lu}(x)|^2}\widehat{B}_{Lu}(x) & \text{if } \widehat{B}_{Lu}(x) \neq 0 \\ 0 & \text{if } \widehat{B}_{Lu}(x) = 0. \end{cases} \quad (23)$$

With the QP safety filters (18) and (22) we establish the following result.

Theorem 5.6. (ISSf QP filter via CBFs) *Consider a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1), a nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$, and a closed set $K \subset \mathbb{R}^n$. Suppose that there exists an ISSf-CBF B for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, such that Assumption 3.7 is satisfied, and consider the feedback law $\kappa = (\bar{\kappa}_C + \widehat{\kappa}_{C_{QP}}, \bar{\kappa}_D + \widehat{\kappa}_{D_{QP}})$, with $\widehat{\kappa}_{C_{QP}}$ as in (19) and $\widehat{\kappa}_{D_{QP}}$ as in (23) defining the corresponding hybrid closed-loop system $\mathcal{H}_\kappa = (C_\kappa, F_\kappa, D_\kappa, G_\kappa)$ as in (4). If there exist functions $\widehat{B}_{Lu} : \Pi(D) \to \mathbb{R}^{m_{D_u}}$ and $\widehat{B}_{Lw} : \Pi(D) \to \mathbb{R}^{m_{D_w}}$ such that Assumption 5.5 holds, then κ renders $\mathcal{H}_\kappa$ $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K.*

Remark 5.7. (Noncompleteness of solutions under QP control) *Notice that the optimization in (18) and (22) is carried over $\mathbb{R}^{m_{C_u}}$ and $\mathbb{R}^{m_{D_u}}$, respectively, instead of over the constrain sets $\Psi_\star$, $\star \in \{C, D\}$, as in Definition 5.1. This allows to compute the closed-form safeguarding feedback law $\widehat{\kappa} = (\widehat{\kappa}_{C_{QP}}, \widehat{\kappa}_{D_{QP}})$, which may potentially lead to maximal solutions to $\mathcal{H}_\kappa$ that are not complete. The "pre" term in the results accounts for this trade-off. We refer the reader to [18, Prop. 2.34] for sufficient conditions to assure completeness of solutions for the hybrid closed-loop system $\mathcal{H}_\kappa$.*

The next result shows that the "half-Sontag" formula [12] not only guarantees input-to-state safety, but also generates a pointwise min-norm feedback law.

Theorem 5.8. (ISSf-CBF "half-Sontag" formula) *Consider a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1), a nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$, and a closed set $K \subset \mathbb{R}^n$. Suppose that there exists an ISSf-CBF B for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, such that Assumption 3.7 is satisfied, and consider the feedback law $\kappa = (\bar{\kappa}_C + \frac{1}{2}\widehat{\kappa}_{SC}, \bar{\kappa}_D + \frac{1}{2}\widehat{\kappa}_{SD})$, with $\widehat{\kappa}_{SC}$ as in (15) and $\widehat{\kappa}_{SD}$ as in (16) defining the corresponding hybrid closed-loop system $\mathcal{H}_{\widehat{\kappa}}$ as in (4). If there exist functions $\widehat{B}_{Lu} : \Pi(D) \to \mathbb{R}^{m_{D_u}}$ and $\widehat{B}_{Lw} : \Pi(D) \to \mathbb{R}^{m_{D_w}}$ such that Assumption 5.5 holds, then κ renders $\mathcal{H}_\kappa$ $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K. In addition, let $\mathcal{V}$ be an open set containing $K_d(\bar{w})$. Then, for all $x \in (\mathcal{V} \setminus K_d(\bar{w})) \cap \Pi(C)$, the feedback law $\frac{1}{2}\widehat{\kappa}_{SC}$ is the pointwise minimizer of*

$$\begin{aligned} \underset{v\in\mathbb{R}^{m_{C_u}}}{\arg\min} \quad & |v|^2 \\ \textit{subject to} \quad & L_{f_u}B(x)v \leq \frac{1}{2}|L_{f_u}B(x)|^2\kappa_{SC}(x). \end{aligned} \quad (24)$$

Similarly, for all $x \in K_d(\bar{w}) \cap \Pi(D)$, the feedback law $\frac{1}{2}\widehat{\kappa}_{SD}$ is the pointwise minimizer of

$$\begin{aligned} \underset{v\in\mathbb{R}^{m_{D_u}}}{\arg\min} \quad & |v|^2 \\ \textit{subject to} \quad & \widehat{B}_{Lu}(x)v \leq \frac{1}{2}|\widehat{B}_{Lu}(x)|^2\kappa_{SD}(x). \end{aligned} \quad (25)$$

6 INVERSE-OPTIMAL SAFETY FILTERS

Given that the control input u defined in terms of a safeguarding feedback law $\widehat{\kappa}$ aims to keep state trajectories of $\mathcal{H}$ from the set K close to K, but the disturbance w seeks to prevent it, we formulate a zero-sum hybrid game that captures such a setting. As stated in

Problem 4.1, given a feedback law κ, which is the sum of a nominal feedback law $\bar{\kappa}$ and a safeguarding feedback law $\widehat{\kappa}$, that renders $\mathcal{H}_\kappa$ $\bar{w}$-small-input input-to-state safe with respect to the disturbance w and the set K, we are interested in determining the cost functional that makes the feedback control action κ optimal. For starters, following [9], we formulate a zero-sum hybrid game. Given $\xi \in K$, an input action $(u, w) = ((u_C, u_D), (w_C, w_D)) \in \mathcal{U} \times \mathcal{W}$, the stage cost for flows $L_C : \mathbb{R}^n \times \mathbb{R}^{m_C} \to \mathbb{R}_{\geq 0}$, the stage cost for jumps $L_D : \mathbb{R}^n \times \mathbb{R}^{m_D} \to \mathbb{R}_{\geq 0}$, and the terminal cost $q : \mathbb{R}^n \to \mathbb{R}$, we define the cost associated to the solution $(\phi, (u, w))$ to $\mathcal{H}$ from ξ as

$$\begin{aligned}\mathcal{J}(\xi, (u, w)) := &\sum_{j=0}^{\sup_j \operatorname{dom}\phi} \int_{t_j}^{t_{j+1}} L_C(\phi(t, j), (u_C(t, j), w_C(t, j)))dt \\ &+ \sum_{j=0}^{\sup_j \operatorname{dom}\phi - 1} L_D(\phi(t_{j+1}, j), (u_D(t_{j+1}, j), w_D(t_{j+1}, j))) \\ &+ \limsup_{\substack{t+j \to \sup_t \operatorname{dom}\phi + \sup_j \operatorname{dom}\phi \\ (t,j)\in \operatorname{dom}\phi}} q(\phi(t, j))\end{aligned} \tag{26}$$

where $t_{\sup_j \operatorname{dom}\phi+1} := \sup_t \operatorname{dom}\phi$ defines the upper limit of the last integral, and $\{t_j\}_{j=0}^{\sup_j \operatorname{dom}\phi}$ is a nondecreasing sequence associated to the definition of the hybrid time domain of ϕ; see Definition 2.1. The terminal cost in (26) is captured by the third term therein and given by q at the value of the state trajectory ϕ at the terminal time.

Given a system $\mathcal{H} = (C, F, D, G)$ as in (1), a closed set $K \subset \mathbb{R}^n$, an ISSf-CBF B for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$, $\bar{w} \geq 0$, a nominal feedback law $\bar{\kappa}$, and $\xi \in K$, we consider the following optimization problem:

$$\underset{\substack{u \\ u\in\mathcal{U}_{\mathcal{H}}(\bar{\kappa},\bar{w})}}{\text{minimize}}\ \underset{w}{\text{maximize}} \quad \mathcal{J}(\xi, (u, w)) \tag{27}$$

where

$$\begin{aligned}\mathcal{U}_{\mathcal{H}}(\bar{\kappa}, \bar{w}) := \Big\{&u \in \mathcal{U} : \exists \widehat{\kappa}, (\phi, (u, w)) \in \mathcal{S}_{\mathcal{H}}(\xi), \\ &\operatorname{dom}\phi \ni (t, j) \mapsto u(t, j) = \bar{\kappa}(\phi(t, j)) + \widehat{\kappa}(\phi(t, j)), \\ &\operatorname{dom}\phi \ni (t, j) \mapsto \phi(t, j) \in K_d(\bar{w})\Big\}.\end{aligned}$$

The next definition introduces the notion of value function for the hybrid game in (27).

Definition 6.1. (Value function) *Given $\xi \in K$ and a nominal feedback law $\bar{\kappa}$, the value function at ξ is given by*

$$\mathcal{J}^*(\xi) := \min_{\substack{u \\ u=(u,w)\in\mathcal{U}_{\mathcal{H}}(\bar{\kappa},\bar{w})}} \max_{w} \mathcal{J}(\xi, (u, w)). \tag{28}$$

6.1 Inverse-Optimal QP Filter

In this section, we provide sufficient conditions to solve Problem 4.1 when the safeguarding controller is expressed as the pointwise solution to a QP, as introduced in Section 5.2. Notice that the min-norm safeguarding feedback law $\widehat{\kappa}_{QP} = (\widehat{\kappa}_{C_{QP}}, \widehat{\kappa}_{D_{QP}})$, with values as in (19) and (23), guarantees input-to-state safety and makes the

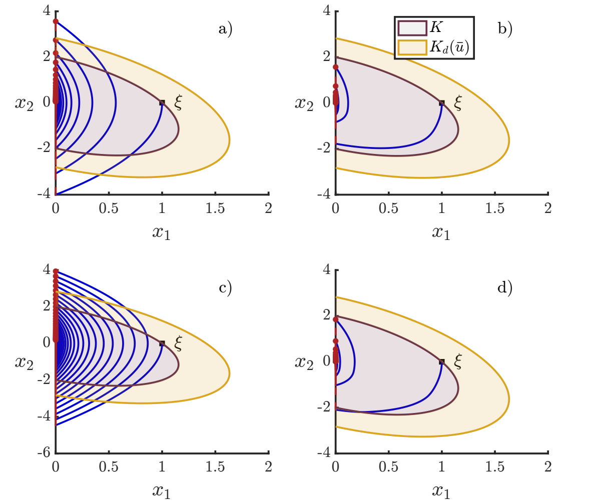

Figure 1 – Consider the hybrid system $\mathcal{H}$ given by (33), and the sets K and $K_d(\bar{w})$ as in (35) and (36), respectively. We show the phase portraits of the state trajectory ϕ of $\mathcal{H}$ from $\xi \in \partial K$ under different settings: a) without disturbance and without the ISSf QP filter, b) without disturbance and with the ISSf QP filter, c) with disturbance and without the ISSf QP filter, and d) with disturbance and with the ISSf QP filter.

feedback law $\kappa = \bar{\kappa} + \widehat{\kappa}_{QP}$ deviate as little as possible from the given nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$.

Given $\xi \in \Pi(\overline{C}) \cup \Pi(D)$, an input action $u = (u_C, u_D) \in \mathcal{U}$, a nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$, the stage cost for flows $L_C : \mathbb{R}^n \times \mathbb{R}^{m_C} \to \mathbb{R}_{\geq 0}$, the stage cost for jumps $L_D : \mathbb{R}^n \times \mathbb{R}^{m_D} \to \mathbb{R}_{\geq 0}$, and the terminal cost $q : \mathbb{R}^n \to \mathbb{R}$, we define the cost associated to the solution $(\phi, (u, w))$ to $\mathcal{H}$ from ξ as in (26), where

- for all $(x, (u_C, w_C)) \in C : x \in \mathcal{V} \setminus K_d(\bar{w})$

$$\begin{aligned}L_C(x, (u_C, w_C)) :=& L_{1C}(x) - \lambda\gamma\Big(\frac{|w_C|}{\lambda}\Big) \\ &+ \frac{1}{2}\frac{|L_{f_u}B(x)|^2}{\max\{0, \omega_C(x)\}}|u_C - \bar{\kappa}_C(x)|^2\end{aligned} \tag{29a}$$

- for all $(x, (u_D, w_D)) \in D : x \in K_d(\bar{w})$

$$\begin{aligned}L_D(x, (u_D, w_D)) :=& L_{1D}(x) - \lambda\gamma\Big(\frac{|w_D|}{\lambda}\Big) \\ &+ \frac{1}{2}\frac{|\widehat{B}_{Lu}(x)|^2}{\max\{0, \omega_D(x)\}}|u_D - \bar{\kappa}_D(x)|^2\end{aligned} \tag{29b}$$

- for all $x \in (\Pi(C) \cup \Pi(D)) \cap \mathcal{V}$

$$q(x) = B(x) \tag{29c}$$

where $\gamma \in \mathcal{K}_\infty$, $\lambda \in (0, 1]$, and $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, and $\mathcal{V}$ is an open set containing and open neighborhood of $K_d(\bar{w})$. We approach the optimization problem in (27) as an inverse problem: we design the optimal safeguarding feedback law, and the stage costs L_{1C} and L_{1D} in (29).

A key tool in our analysis to solve the hybrid game in (27) is the Legendre–Fenchel transform, which is defined as follows.

Definition 6.2. (Legendre-Fenchel transform of a class-$\mathcal{K}_\infty$ function [13, Lemma A.1]) *For a class-$\mathcal{K}_\infty$ function γ whose derivative exists and is also a class-$\mathcal{K}_\infty$ function, the Legendre–Fenchel transform of γ is defined as*

$$\bar{\gamma}(r) = r(\gamma')^{-1}(r) - \gamma((\gamma')^{-1}(r)) \quad \forall r \geq 0 \tag{30}$$

where $(\gamma')^{-1}$ stands for the inverse function of $\gamma' := \frac{d\gamma}{dr}$.

Finally, the next result shows that every pointwise min-norm feedback law is optimal for a meaningful game.

Theorem 6.3. (Inverse-optimal QP safety filter) *Consider the hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1), a nominal feedback law $\bar{\kappa} = (\bar{\kappa}_C, \bar{\kappa}_D)$, and a closed set $K \subset \mathbb{R}^n$. Suppose that there exists an ISSf-CBF B for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, where $K_d(\bar{w})$ is defined as in (8) for some $\rho \in \mathcal{K}_\infty$ and $\bar{w} \geq 0$, such that Assumption 3.7 is satisfied. Consider the feedback law $\kappa = \bar{\kappa} + \widehat{\kappa}_{QP}$, where $\widehat{\kappa}_{QP} = (\widehat{\kappa}_{C_{QP}}, \widehat{\kappa}_{D_{QP}})$ has values as in (19) and (23), defining the corresponding hybrid closed-loop system $\mathcal{H}_\kappa = (C_\kappa, F_\kappa, D_\kappa, G_\kappa)$. If there exist functions $\widehat{B}_{Lu} : \Pi(D) \to \mathbb{R}^{m_{D_u}}$ and $\widehat{B}_{Lw} : \Pi(D) \to \mathbb{R}^{m_{D_w}}$ such that Assumption 5.5 holds, then κ renders $\mathcal{H}_\kappa$ $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K, and minimizes, for any $\xi \in K$, the cost $\mathcal{J}$ in (26) with*

$$L_{1C}(x) := -\Big(L_{f+L_{f_u}\bar{\kappa}_C}B(x) - \frac{1}{2}\max\{0, \omega_C(x)\} + \lambda\bar{\gamma}(|L_{f_w}B(x)|)\Big) \quad \forall x \in \Pi(C_\kappa) \cap \mathcal{V} \tag{31}$$

where $\mathcal{V} \subset \mathbb{R}^n$ is an open set containing an open neighborhood of $K_d(\bar{w})$, and

$$L_{1D}(x) := -\Big(B(g(x) + g_u(x)\bar{\kappa}_D(x)) - B(x) - \frac{1}{2}\max\{0, \omega_D(x)\} + \lambda\bar{\gamma}(|\widehat{B}_{Lw}(x)|)\Big) \quad \forall x \in \Pi(D_\kappa) \cap K_d(\bar{w}) \tag{32}$$

with ω_C defined in (17) and ω_D in (21).

7 NUMERICAL EXAMPLE

To illustrate our results, consider the following oscillator with impacts with dynamics given by

$$\mathcal{H} : \begin{cases} \dot{x} = \begin{pmatrix} x_2 \\ -\zeta_C \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u_C + \begin{pmatrix} 0 \\ 2 \end{pmatrix} w_C & (x, (u_C, w_C)) \in C \\ x^+ = \begin{pmatrix} x_1 \\ -\zeta_D x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \eta_u \end{pmatrix} u_D + \begin{pmatrix} 0 \\ \eta_w \end{pmatrix} w_D & (x, (u_D, w_D)) \in D \end{cases} \tag{33}$$

where $\zeta_C, \eta_u, \eta_w \geq 0$, $\zeta_D \in (0, 1]$, and

$$C := \{(x, (u_C, w_C)) \in \mathbb{R}^2 \times \mathbb{R}^2 : x_1 \geq 0\}$$
$$D := \{(x, (u_D, w_D)) \in \mathbb{R}^2 \times \mathbb{R}^2 : x_1 = 0,\ x_2 \leq 0\}$$

with $(u_\star, w_\star) \in \mathbb{R}^2$, for $\star \in \{C, D\}$. Now, consider the following nominal feedback law:

$$\bar{\kappa}(x) = (\bar{\kappa}_C(x), \bar{\kappa}_D(x)) := \left(-\frac{1}{2}r_C x_2, \frac{\zeta_D x_2}{1 + 2r_D}\right) \tag{34}$$

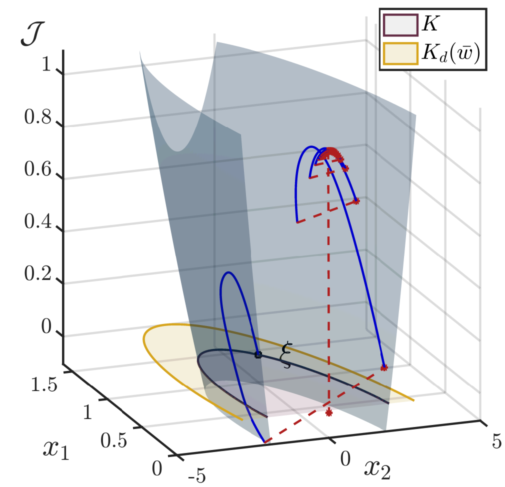

Figure 2 – Consider the hybrid system $\mathcal{H}$ given by (33), and the sets K and $K_d(\bar{w})$ as in (35) and (36), respectively. We show that phase portrait of the state trajectory ϕ of $\mathcal{H}$ from $\xi \in \partial K$. In addition, the surface plot of the value function (dark gray) and the cost of solution (blue-red), rendered by the saddle-point equilibrium strategy, are shown. Notice that the $\mathcal{J}^*(\xi) = B(\xi) = 0$, as guaranteed by Theorem 6.3.

where $r_C > 0$ and $r_D \in \left(-\infty, \frac{1}{2\zeta_D - 2}\right) \cup \left(-\frac{1}{2\zeta_D + 2}, \infty\right)$. Next, consider the following set

$$K = \left\{x \in \mathbb{R}^2 : (x_1 \geq 0 \text{ or } x_2 \leq 0),\ \left(\frac{x_1}{a}\right)^2 + \frac{x_1 x_2}{ab} + \left(\frac{x_2}{b}\right)^2 \leq 1\right\} \tag{35}$$

for some $a, b \in \mathbb{R}_{>0}$ such that $b > a$ and $5a > 2b$. Pick $r \mapsto \rho(r) = r^3$ and $\bar{w} = 1$. Then,

$$K_d(\bar{w}) = \left\{x \in \mathbb{R}^2 : (x_1 \geq 0 \text{ or } x_2 \leq 0),\ \left(\frac{x_1}{a}\right)^2 + \frac{x_1 x_2}{ab} + \left(\frac{x_2}{b}\right)^2 \leq 2\right\} \tag{36}$$

which are depicted in Fig. 1 and Fig. 2. From this choice, we have

$$B(x) = \left(\frac{x_1}{a}\right)^2 + \frac{x_1 x_2}{ab} + \left(\frac{x_2}{b}\right)^2 - 1. \tag{37}$$

To show that B is an ISSf-BF candidate[3] for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, notice that: i) $\overline{\Pi(C)} \cup \Pi(D) \cup G(D) = \{x \in \mathbb{R}^2 : x_1 \geq 0\} \subset \operatorname{dom} B = \mathbb{R}^2$ and $K_d(\bar{w}) \subset \Pi(C) \cup \Pi(D) = \{x \in \mathbb{R}^2 : x_1 \geq 0\}$, ii) B is continuously differentiable everywhere on $\mathbb{R}^2$, iii) $B(x) > 0$ for all $x \in \{\xi \in \mathbb{R}^2 : \xi_1 \geq 0\} \setminus K$, and iv) $B(x) \leq 0$ for all $x \in K$. Then, B is an ISSf-BF candidate for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$, according to Definition 3.5. Also, notice that

$$G\left(D \cap \left(K_d(\bar{w}) \times \mathbb{R}^2\right)\right) = \{x \in \mathbb{R}^2 : x_1 = 0\} \subset \Pi(C) \cup \Pi(D)$$

[3]Notice that the dynamics $\mathcal{H}$ and B satisfy Assumption 5.5, for each $\star \in \{u, w\}$, with

$$\widehat{B}_{L\star}(x) = \frac{2\eta_\star x_2}{b^2} \qquad \forall x \in \mathbb{R}^2 : x_1 = 0, x_2 \leq 0.$$

and, as a result, Assumption 3.7 is satisfied. The ISSf-BF candidate B in (37) is also an ISSf-CBF for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$. To see this, pick $\alpha_C = 1$ and check the ISSf-CBF condition during flows in (13a), namely

$$L_{f_u} B(x) = 0 \quad \Rightarrow \quad x_1 = -\frac{2a}{b} x_2.$$

This equation defines a line on the plane along which ω_C, defined in (13b), satisfies

$$\omega_C(x) = \frac{x_2^2}{b^2}\left(\frac{b}{a} - 1\right) - 1.$$

Now, let

$$\mathcal{V} := \left\{ x \in \mathbb{R}^2 : (x_1 \geq 0 \text{ or } x_2 \leq 0),\ B(x) < \frac{3a}{b-a} - 1 \right\}$$

be an open set containing an open neighborhood of the boundary of $K_d(\bar{w})$. Then, we see that:

$$\omega_C(x) \leq 0 \quad \forall x \in (\mathcal{V} \setminus K_d(\bar{w})) \cap \Pi(C) : x_1 = -\frac{2a}{b} x_2.$$

Thus, (13a) is satisfied. Similarly, notice that $\widehat{B}_{Lu}(x) = 0 \Rightarrow x_2 = 0$ which, in turn, implies that ω_D, defined in (14b), satisfies

$$\omega_D(x) = \alpha_D(B(x) - \rho(\bar{w})) \quad \forall x \in \mathbb{R}^2 : x_2 = 0$$

for any $\alpha_D \in [0, 1]$. It is then immediate that $\omega_D(x) \leq 0$ for all $x \in K_d(\bar{w}) \cap \Pi(D)$ such that $x_2 = 0$, that is (14a) holds. As a result, B is an ISSf-CBF for $\mathcal{H}$ with respect to $(K, K_d(\bar{w}))$.

Following Section 5.2, the pointwise min-norm QP safeguarding feedback law $\widehat{\kappa}_{QP}(x) = (\widehat{\kappa}_{C_{QP}}(x), \widehat{\kappa}_{D_{QP}}(x))$ has values as in (19) and (23), and, from Theorem 5.6, we conclude that the feedback law $\kappa = (\bar{\kappa}_C + \widehat{\kappa}_{C_{QP}}, \bar{\kappa}_D + \widehat{\kappa}_{D_{QP}})$ renders the resulting closed-loop system $\mathcal{H}_\kappa$ $\bar{w}$-small-input ISSf with respect to the disturbance w and the set K. In particular, we can see that in Fig. 1d) the set $K_d(\bar{w})$ is conditionally invariant for $\mathcal{H}_\kappa$ with respect to w and K, as opposed to Fig. 1c) where the ISSf QP filter is not active. Also, notice that from Fig. 1a) and Fig. 1b), when disturbances are not considered, the set K is forward invariant for $\mathcal{H}_\kappa$, as discussed in Remark 3.4.

In addition, pick $r \mapsto \gamma(r) = r^2$ and invoking Theorem 6.3, we have that κ minimizes, for any initial condition $\xi \in K$, the cost $\mathcal{J}$ in (26). This can also be seen from Fig. 2, where $\mathcal{J}$ attains the value of the value function evaluated at the initial condition. In this particular example, we chose $\xi \in \partial K$, therefore, it follows that $J^*(\xi) = B(\xi) = 0$.

8 CONCLUSIONS AND FUTURE WORK

In this paper, we study the problem of designing safety filters for hybrid systems under disturbances as an inverse optimal problem. Via a characterization of safeguarding feedback law that assigns the control input, we formulate the problem as a two-player zero-sum hybrid game to minimize the effect of the worst-case disturbance. Instead of solving the game for a given cost functional, we design the cost functional that the safeguarding feedback law minimizes. A QP formulation is shown to solve the problem. Future work includes designing projection tools to deal with feedback laws that render maximal solutions not complete and force them to satisfy the constraints specified by the hybrid dynamics.

ACKNOWLEDGMENTS

This research has been 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.

REFERENCES

[1] A. D. Ames, J. W. Grizzle, and P. Tabuada. 2014. Control barrier function based quadratic programs with application to adaptive cruise control. *Proceedings of the IEEE Conference on Decision and Control* 2015-Febru, February (2014), 6271–6278.
[2] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada. 2017. Control Barrier Function Based Quadratic Programs for Safety Critical Systems. *IEEE Trans. Automat. Control* 62, 8 (2017), 3861–3876.
[3] M. Ghanbarpour, A. Isaly, R. G. Sanfelice, and W. Dixon. 2022. Optimal Safety for Constrained Differential Inclusions using Nonsmooth Control Barrier Functions. *IEEE Control Systems Letters (L-CSS)* 7 (December 2022), 1303–1308.
[4] R. Goebel, R. G. Sanfelice, and A. R. Teel. 2009. Hybrid dynamical systems. *IEEE Control Systems Magazine* 29, 2 (2009), 28–93.
[5] R. Goebel, R. G. Sanfelice, and A. R. Teel. 2012. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness.* Princeton University Press, New Jersey.
[6] Mrdjan Jankovic. 2018. Robust Control Barrier Functions for Constrained Stabilization of Nonlinear Systems. *Automatica* 96 (2018), 359–367.
[7] B. Jayawardhana and M. Z. Romdlony. 2019. Robustness analysis of systems' safety through a new notion of input-to-state safety. *Int. J. Robust Nonlinear Control* 29, 7 (2019), 2125–2136.
[8] S. Jimenez Leudo, F. Ferrante, and R. G. Sanfelice. 2024. On the Optimal Cost and Asymptotic Stability in Two-Player Zero-Sum Set-Valued Hybrid Games. In *Proceedings of the 2024 American Control Conference.*
[9] S. Jimenez Leudo and R. G. Sanfelice. 2022. Sufficient Conditions for Optimality and Asymptotic Stability in Two-Player Zero-Sum Hybrid Games. In *Proceedings of the 25th ACM International Conference on Hybrid Systems: Computation and Control* (Milan, Italy) *(HSCC '22).* Article 18, 11 pages.
[10] S. Jimenez Leudo and R G Sanfelice. 2022. Sufficient Conditions for Optimality in Finite-Horizon Two-Player Zero-Sum Hybrid Games. In *Proceedings of the 2022 61th IEEE Conference on Decision and Control (CDC).* 3268–3273.
[11] S. Kolathaya and A. D. Ames. 2019. Input-to-State Safety With Control Barrier Functions. *IEEE Control Systems Letters* 3, 1 (2019), 108–113.
[12] M. Krstic. 2024. Inverse Optimal Safety Filters. *IEEE Trans. Automat. Control* 69, 1 (2024), 16–31.
[13] M. Krstic and Zhong-Hua Li. 1998. Inverse optimal design of input-to-state stabilizing nonlinear controllers. *IEEE Trans. Automat. Control* 43, 3 (1998), 336–350.
[14] L. Long and X. Huang. 2024. On sufficient and necessary conditions of compositional input-to-state safety for nonlinear systems. *Systems & Control Letters* 191 (2024), 105886.
[15] M. Maghenem and D. Karaki. 2023. On a Strong Robust-Safety Notion for Differential Inclusions. *IEEE Trans. Automat. Control* (2023), 1–12.
[16] M. Maghenem and R. G. Sanfelice. 2019. Characterizations of safety and conditional invariance in dynamical systems. In *Proceedings of the American Control Conference.* 5039–5044.
[17] M. Maghenem and R. G. Sanfelice. 2023. On the Converse Safety Problem for Differential Inclusions: Solutions, Regularity, and Time-Varying Barrier Functions. *IEEE Trans. Automat. Control* 68, 1 (January 2023), 172–187.
[18] R. G. Sanfelice. 2021. *Hybrid Feedback Control.* Princeton University Press.
[19] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames. 2015. Robustness of Control Barrier Functions for Safety Critical Control. *IFAC-PapersOnLine* 48, 27 (2015), 54–61.