Characterization of the Dynamical Properties of Safety Filters for Linear Planar Systems

Yiting Chen* Pol Mestres* Emiliano Dall'Anese Jorge Cortés

Abstract— **This paper studies the dynamical properties of closed-loop systems obtained from control barrier function-based safety filters. We provide a sufficient and necessary condition for the existence of undesirable equilibria and show that the Jacobian matrix of the closed-loop system evaluated at an undesirable equilibrium always has a nonpositive eigenvalue. In the special case of linear planar systems and ellipsoidal obstacles, we give a complete characterization of the dynamical properties of the corresponding closed-loop system. We show that for underactuated systems, the safety filter always introduces a single undesired equilibrium, which is a saddle point. We prove that all trajectories outside the global stable manifold of such equilibrium converge to the origin. In the fully actuated case, we discuss how the choice of nominal controller affects the stability properties of the closed-loop system. Various simulations illustrate our results.**

I. Introduction

Modern autonomous systems and cyber-physical systems – from self-driving vehicles and robotic systems to critical infrastructures – must provide safety guarantees while performing complex operational tasks [1]. A popular approach to promote safety, where the term "safety" here refers to the ability to render a predefined set of states forward invariant, relies on the so-called *safety filters*; these filters take a potentially unsafe nominal controller, designed to provide stability or optimality guarantees, and minimally modify it to account for safety constraints. While the filtered controller ensures safety, it may not preserve the stability or optimality properties of the nominal controller. This challenge is the main motivation for this work.

Literature Review: One of the main approaches for rendering a given set forward invariant is via Control Barrier Functions (CBFs) [2]–[5]. Given a nominal controller with desirable properties such as asymptotic stability of an equilibrium, CBFs acts on top of the nominal controller to ensure safety. This technique is often referred to as a *safety filter* [6]. The main research question here is whether the closed-loop system with safety filters retains the stability guarantees of the nominal controller. This was studied in, e.g., [7], which provides an estimate of the region of attraction of the equilibrium. However, it is unclear how conservative such estimate may be for general systems. The seminal works in [8]–[12] show that designs similar to safety filters can introduce undesired equilibria that may be stable or unstable.

*Equal contribution of the authors. Y. Chen and E. Dall'Anese are with the Department of Electrical and Computer Engineering at Boston University; P. Mestres and J. Cortés are with the Department of Mechanical and Aerospace Engineering at the University of California San Diego.

This work was supported by the AFOSR Award FA9550-23-1-0740.

Statement of Contributions: The goal of this paper is to advance the understanding of the dynamical properties of closed-loop systems obtained from CBF-based safety filters. The main contribution of the paper is two-fold:
(i) Our first contribution is to characterize the undesired equilibria that emerge in the closed-loop system formed by a control-affine dynamical system, a stabilizing nominal controller, and a CBF-based safety filter. General obstacles are considered (this is the subject of Section III).
(ii) Next, we focus our attention to linear time-invariant (LTI) planar systems (Section IV). We show that, for these systems, the dynamical properties of systems with ellipsoidal obstacles are equivalent to those of systems with circular obstacles. For underactuated LTI planar systems, we give a complete characterization of the trajectories of the closed-loop system. We show that such systems always have a single undesired equilibrium. Moreover, we show that such undesired equilibrium is a saddle point and show that all trajectories that lie outside the global stable manifold of this equilibrium converge to the origin. For fully actuated LTI planar systems, we show that the closed-loop system can have up to three undesired equilibria, and characterize their stability properties.

Additionally, we show that in the fully actuated case there always exists a nominal controller (which can be explicitly computed) that makes the closed-loop system have a single undesired saddle point equilibrium. Therefore, our findings can be used to inform the design of the nominal controller.

II. Preliminaries and Problem Statement

Notation. We denote by $\mathbb{N}_{>0}$, $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ the set of positive integers, real, and nonnegative numbers. We use bold symbols to represent vectors and non-bold symbols to represent scalar quantities; $\mathbf{0}_n$ represents the n-dimensional zero vector. Given $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|$ denotes its Euclidean norm. Given a matrix $G \in \mathbb{R}^{n \times n}$, $\|\mathbf{x}\|_G = \sqrt{\mathbf{x}^T G \mathbf{x}}$. A function $\beta : \mathbb{R} \to \mathbb{R}$ is of extended class $\mathcal{K}_\infty$ if $\beta(0) = 0$, β is strictly increasing and $\lim_{s \to \pm\infty} \beta(s) = \pm\infty$. Given a set $S \subset \mathbb{R}^n$, we denote by Int(S) and ∂S the interior and boundary of S, respectively. For a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$, $\nabla h(\mathbf{x})$ denotes its gradient at $\mathbf{x}$.

Consider the system $\dot{\mathbf{x}} = f(\mathbf{x})$, with $f : \mathbb{R}^n \to \mathbb{R}^n$ locally Lipschitz. Then, for any initial condition $\mathbf{x}_0 \in \mathbb{R}^n$ at time t_0, there exists a maximal interval of existence $[t_0, t_1)$ such that $\mathbf{x}(t; \mathbf{x}_0)$ is the unique solution to $\dot{\mathbf{x}} = f(\mathbf{x})$ on $[t_0, t_1)$, cf. [13]. For f continuously differentiable and $\mathbf{x}^*$ an equilibrium point of f (i.e., $f(\mathbf{x}^*) = \mathbf{0}_n$), $\mathbf{x}^*$ is *degenerate* if the Jacobian of f evaluated at $\mathbf{x}^*$ has at least one eigenvalue with real part equal to zero (otherwise, we refer to $\mathbf{x}^*$ as *hyperbolic*). Given a hyperbolic equilibrium point with $k \in \mathbb{Z}_{>0}$ eigenvalues with negative real part, the Stable Manifold Theorem [14, Section 2.7] ensures that there exists an invariant k-dimensional manifold S for which all trajectories with initial conditions lying on S converge to $\mathbf{x}^*$. The global stable manifold at $\mathbf{x}^*$ is defined as $W_s(\mathbf{x}^*) = \bigcup_{\{t \le 0,\ \mathbf{x}_0 \in S\}} \mathbf{x}(t; \mathbf{x}_0)$.

A. Control barrier functions and safety filters

Consider a control-affine dynamical system of the form

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are locally Lipschitz functions, $\mathbf{x} \in \mathbb{R}^n$ is the state, and $\mathbf{u} \in \mathbb{R}^m$ is the input.

Definition 1 (Control Barrier Function): Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, and define the set $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) \ge 0\}$. The function h is a **CBF** of $\mathcal{C}$ for the system (1) if there exists an extended class $\mathcal{K}_\infty$ function α such that, for all $\mathbf{x} \in \mathcal{C}$, there exists $\mathbf{u} \in \mathbb{R}^m$ satisfying $\nabla h(\mathbf{x})^\top (f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}) + \alpha(h(\mathbf{x})) \ge 0$. □

Suppose that a nominal controller $\mathbf{u} = k(\mathbf{x})$ is designed so that the system $\dot{\mathbf{x}} = \tilde{f}(\mathbf{x}) := f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$ renders the origin globally asymptotically stable (this is without loss of generality). Consider the system

$$\dot{\mathbf{x}} = \tilde{f}(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x}), \tag{2}$$

where the map $\mathbf{x} \mapsto v(\mathbf{x})$ is defined as:

$$\begin{aligned} v(\mathbf{x}) = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^m} & \|\boldsymbol{\theta}\|^2_{G(\mathbf{x})} \\ \text{s.t. } & \nabla h(\mathbf{x})^\top (f(\mathbf{x}) + g(\mathbf{x})(k(\mathbf{x}) + \boldsymbol{\theta})) + \alpha(h(\mathbf{x})) \ge 0 \end{aligned} \tag{3}$$

with $G : \mathbb{R}^n \to \mathbb{R}^{m \times m}$ continuously differentiable and positive definite for all $\mathbf{x} \in \mathbb{R}^n$. We assume the following.

Assumption 1 (Origin in the interior of $\mathcal{C}$): The set $\{\mathbf{x} \in \mathbb{R}^n : \ h(\mathbf{x}) = 0, \ \tilde{f}(\mathbf{x}) = \mathbf{0}_n\}$ is empty and $h(\mathbf{0}_n) > 0$. □

Assumption 2 (Feasibility): There exists an extended class $\mathcal{K}_\infty$ function α such that $g(\mathbf{x})^\top \nabla h(\mathbf{x}) \neq 0$ for all $\mathbf{x}$ in $\{\mathbf{x} \in \mathbb{R}^n : h(x) \ge 0, \ \nabla h(\mathbf{x})^\top f(\mathbf{x}) + \alpha(h(\mathbf{x})) \le 0\}$. □

Assumption 2 ensures that (3) is feasible for all $\mathbf{x} \in \mathbb{R}^n$ and therefore $v(\mathbf{x})$ is well-defined for all $\mathbf{x} \in \mathbb{R}^n$. Moreover, under Assumption 2, and using arguments similar to [15, Lemma III.2], one can show that $v(\mathbf{x})$ is locally Lipschitz. Assumption 2 also ensures that $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}) \neq \mathbf{0}_n$ for all $\mathbf{x} \in \partial\mathcal{C}$. From [3, Thm. 2], it follows that the system (2) with the controller $v(\mathbf{x})$ renders the set $\mathcal{C}$ forward invariant. Because of this feature, and because $\mathcal{C}$ is modeling a set of *safe* states, $v(\mathbf{x})$ is typically referred to as *safety filter*.

B. Problem Statement

We consider a control-affine dynamical system as in (1) and a safe set $\mathcal{C} \subset \mathbb{R}^n$ defined as the 0-superlevel set of a differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. Assume that h is a CBF of $\mathcal{C}$ for system (1), and that Assumptions 1 and 2 hold. Studying the dynamical behavior of (2) is challenging. Indeed, as noted in [7], it does not inherit the global asymptotic stability properties of the controller k, and can even have undesired equilibria [8]–[11]. However, most of these works focus on studying conditions under which such undesired equilibria exist or can be confined to specific regions of interest, but do not study dynamical properties of the closed-loop system. Hence the goal of this paper is as follows:

Problem 1: Given system (1) with a stabilizing nominal controller $k(\mathbf{x})$ and the safety filter $v(\mathbf{x})$, characterize the dynamical properties of (2) (such as undesirable equilibria and their regions of attraction, limit cycles and region of attraction of the origin) and investigate how these properties are determined by the original closed-loop system $\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$. □

In the following section, we consider the system (2) and characterize its undesired equilibria. In Section IV, given the complexity of solving Problem 1, we then restrict our attention to linear planar systems.

III. Characterization of undesirable equilibria

We start by reformulating the expression for the unique optimal solution $v(\mathbf{x})$ of the quadratic program (3). Let $\eta(\mathbf{x}) = \nabla h(\mathbf{x})^T (f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})) + \alpha(h(\mathbf{x}))$. Then,

$$v(\mathbf{x}) = \begin{cases} \mathbf{0}_m, & \text{if } \eta(\mathbf{x}) \ge 0, \\ \bar{u}(\mathbf{x}), & \text{if } \eta(\mathbf{x}) < 0, \end{cases} \tag{4}$$

where $\bar{u}(\mathbf{x}) := -\frac{\eta(\mathbf{x}) G(\mathbf{x})^{-1} g(\mathbf{x})^\top \nabla h(\mathbf{x})}{\|g(\mathbf{x})^\top \nabla h(\mathbf{x})\|^2_{G(\mathbf{x})^{-1}}}$. We use this expression in the following result, which provides a necessary and sufficient condition for undesirable equilibria of (2).

Lemma 1: (Conditions for undesirable equilibria): Let Assumptions 1 and 2 be satisfied. Let $\mathbf{p}_0 \in \mathbb{R}^n$ be such that $\tilde{f}(\mathbf{p}_0) \neq \mathbf{0}_n$. Then, $\mathbf{p}_0$ is an equilibrium of (2) if and only if there exists $\delta < 0$ such that

$$\begin{aligned} & h(\mathbf{p}_0) = 0 \text{ and} \\ & \tilde{f}(\mathbf{p}_0) = \delta g(\mathbf{p}_0) G(\mathbf{p}_0)^{-1} g(\mathbf{p}_0)^\top \nabla h(\mathbf{p}_0)\,. \end{aligned} \tag{5}$$ □

Proof: If there exists $\mathbf{p}_0$ and $\delta < 0$ satisfying (5), then it holds that that $\delta = \frac{\nabla h(\mathbf{p}_0)^\top \tilde{f}(\mathbf{p}_0)}{\|g(\mathbf{p}_0)^\top \nabla h(\mathbf{p}_0)\|^2_{G(\mathbf{x})^{-1}}}$, from which it follows that $\tilde{f}(\mathbf{p}_0) + g(\mathbf{p}_0)\bar{u}(\mathbf{p}_0) = \mathbf{0}_n$ and $\eta(\mathbf{p}_0) < 0$. Hence, $\mathbf{p}_0$ is an equilibrium of (2).

On the other hand, if $\mathbf{p}_0$ is an equilibrium of (2) and $\tilde{f}(\mathbf{p}_0) \neq \mathbf{0}_n$, then $\tilde{f}(\mathbf{p}_0) + g(\mathbf{p}_0)\bar{u}(\mathbf{p}_0) = \mathbf{0}_n$. It follows that

$$0 = \nabla h(\mathbf{p}_0)^\top (\tilde{f}(\mathbf{p}_0) + g(\mathbf{p}_0)\bar{u}(\mathbf{p}_0)) = -\alpha(h(\mathbf{p}_0)),$$

by which we get the first equation.

In addition, we get that $\eta(\mathbf{p}_0) < 0$ and $f(\mathbf{p}_0) = \frac{\eta(\mathbf{p}_0)}{\|g(\mathbf{p}_0)^\top \nabla h(\mathbf{p}_0)\|^2_{G(\mathbf{x})^{-1}}} g(\mathbf{p}_0) G(\mathbf{p}_0)^{-1} g(\mathbf{p}_0)^\top \nabla h(\mathbf{p}_0)$, which implies the second equation. ■

This result has the same flavor as [9, Theorem 2] and [10, Proposition 5.1], which characterize the undesired equilibria for related, but different, safety filter designs.

By Lemma 1, we can define the set of *potential undesirable equilibria* of (2) as:

$$\mathcal{E} := \{\mathbf{x} : \ \exists\ \delta \in \mathbb{R} \text{ s.t. } (\mathbf{x}, \delta) \text{ solves (5)}\}\,.$$

On the other hand, the set of *undesirable equilibria* is:

$$\hat{\mathcal{E}} := \{\mathbf{x} : \exists\, \delta < 0 \text{ s.t. } (\mathbf{x}, \delta) \text{ solves (5)}\} \subset \mathcal{E}.$$

The term *undesirable* stems from the fact that these equilibria are different from the origin, which is the equilibrium point where the system needs to be stabilized. By Lemma 1, it follows that determining the equilibrium points of system (2) is equivalent to solving (5) and checking the sign of δ. For a solution $(\mathbf{p}_0, \delta_{\mathbf{p}_0})$ to (5), we refer to $\delta_{\mathbf{p}_0}$ as the **indicator** of $\mathbf{p}_0$, since the sign of $\delta_{\mathbf{p}_0}$ determines whether $\mathbf{p}_0$ is a new, *undesirable*, equilibrium of the system with the CBF filter. Additionally, we show that the value of the indicator is useful for determining the stability properties of the undesirable equilibrium. For a given undesirable equilibrium $\mathbf{p}_0$ of (2), the indicator can be computed as $\delta_{\mathbf{p}_0} = \frac{\nabla h(\mathbf{p}_0)^\top \tilde{f}(\mathbf{p}_0)}{\|g(\mathbf{p}_0)^\top \nabla h(\mathbf{p}_0)\|^2_{G(\mathbf{x})^{-1}}}$. In addition, Assumption 1 ensures that no solution of (5) has $\delta = 0$.

Under appropriate conditions, the next result shows that we can compute the Jacobian of $\tilde{f}(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x})$ at $\mathbf{x} \in \hat{\mathcal{E}}$ and find one of its eigenvalues.

Lemma 2: (Jacobian at the undesirable equilibrium): Let Assumptions 1 and 2 be satisfied and assume that $D = g(\mathbf{x})G(\mathbf{x})^{-1}g(\mathbf{x})^\top$ is a constant matrix, $\tilde{f}(\mathbf{x})$, $\alpha(\cdot)$ are differentiable and $h(\mathbf{x})$ is twice differentiable. For any $\mathbf{x} \in \hat{\mathcal{E}}$, the Jacobian of $\tilde{f}(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x})$ evaluated at $\mathbf{x}$ is

$$J\,|_{\mathbf{x}\in\hat{\mathcal{E}}} = J_{\tilde{f}} - \frac{D\nabla h(\mathbf{x})\nabla h(\mathbf{x})^\top}{\nabla h(\mathbf{x})^\top D \nabla h(\mathbf{x})}[J_{\tilde{f}} + \alpha'(0)\mathbf{I}_n]$$
$$- \frac{D}{\nabla h(\mathbf{x})^\top D \nabla h(\mathbf{x})}[H_h \nabla h(\mathbf{x})^\top \tilde{f}(\mathbf{x}) - \nabla h(\mathbf{x})\tilde{f}(\mathbf{x})^\top H_h],$$

where $J_{\tilde{f}}$ is the Jacobian matrix of $\tilde{f}(\mathbf{x})$ and H_h is the Hessian of $h(\mathbf{x})$. Moreover, for any $\mathbf{x} \in \hat{\mathcal{E}}$, it holds that

$$(J\,|_{\mathbf{x}\in\hat{\mathcal{E}}})^\top \nabla h(\mathbf{x}) = -\alpha'(0)\nabla h(\mathbf{x}),$$

the algebraic multiplicity of $-\alpha'(0)$ is 1, and all the other eigenvalues of $J\,|_{\mathbf{x}\in\hat{\mathcal{E}}}$ do not change when $\alpha(\cdot)$ changes. □

The proof of Lemma 2 follows from a careful computation and it is omitted due to space constraints. Note that J always has an eigenvalue $-\alpha'(0)$; it follows that all the undesirable equilibria are degenerate if $\alpha'(0) = 0$, which complicates the stability analysis. If $\alpha'(0) > 0$, the Jacobian evaluated at $\mathbf{x} \in \hat{\mathcal{E}}$ always has a negative eigenvalue. Lemmas 1 and 2 show that the extended $\mathcal{K}_\infty$ function $\alpha(\cdot)$ does not play a role in the existence of undesirable equilibria. Additionally, changing $\alpha(\cdot)$ will only affect one eigenvalue of the Jacobian evaluated at $\mathbf{x} \in \hat{\mathcal{E}}$. The assumption that $g(\mathbf{x})G(\mathbf{x})^{-1}g(\mathbf{x})^\top$ is constant is satisfied for several classes of systems, such as mechanical systems, like the ones considered in [7, Section III.B].

IV. LTI Planar Systems with Safety Filters

Since Problem 1 is difficult to solve in general, here we provide a solution for it for planar LTI dynamics and ellipsoidal obstacles. Consider the LTI planar system

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \tag{6}$$

with $\mathbf{x} = [x_1, x_2]^\top \in \mathbb{R}^2$, $\mathbf{u} \in \mathbb{R}^m$, with $m \in \{1, 2\}$, $A \in \mathbb{R}^{2\times 2}$, and with $B \in \mathbb{R}^{2\times m}$ full column rank. We make the following assumption on (6).

Assumption 3 (Stabilizability): The system (6) is stabilizable. Moreover, let $\mathbf{u} = -K\mathbf{x}$, $K \in \mathbb{R}^{2\times m}$, be any stabilizing controller such that $\tilde{A} = A - BK$ is Hurwitz. □

In this setup, the system (2) is then customized as follows:

$$\dot{\mathbf{x}} = F(\mathbf{x}) := (A - BK)\mathbf{x} + Bv(\mathbf{x}), \tag{7}$$

where the safety filter is given by

$$v(\mathbf{x}) = \begin{cases} 0, & \text{if } \eta(\mathbf{x}) \geq 0, \\ -\frac{\eta(\mathbf{x})G(\mathbf{x})^{-1}B^T\nabla h(\mathbf{x})}{\|B^T\nabla h(\mathbf{x})\|^2_{G(\mathbf{x})^{-1}}}, & \text{if } \eta(\mathbf{x}) < 0. \end{cases} \tag{8}$$

In the following, we show that the undesired equilibria and their stability properties of (7) with ellipsoidal obstacles are equivalent to those of a system with circular obstacles.

Proposition 1: (Safety filters with ellipsoidal and circular obstacles have the same dynamical properties): Let $\mathbf{x}_c \in \mathbb{R}^2$, $P \in \mathbb{R}^{2\times 2}$ positive definite, $h(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_c)^T P(\mathbf{x}-\mathbf{x}_c) - 1$, $\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$. Suppose that $P = E^T E$, with $E \in \mathbb{R}^{2\times 2}$ also positive definite, and define $\hat{\mathbf{x}}_c = E\mathbf{x}_c$, $\hat{h}(\hat{\mathbf{x}}) = (\hat{\mathbf{x}} - \hat{\mathbf{x}}_c)^T(\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) - 1$ and $\hat{\mathcal{C}} = \{\mathbf{x} \in \mathbb{R}^n : \hat{h}(\mathbf{x}) \geq 0\}$. Moreover, let $\hat{A} = EAE^{-1}$, $\hat{B} = EB$, $\hat{G}(\hat{\mathbf{x}}) = G(E^{-1}\hat{\mathbf{x}})$ and $\hat{\eta}(\hat{\mathbf{x}}) = \nabla\hat{h}(\hat{\mathbf{x}})^T(\hat{A} - \hat{B}KE^{-1})\hat{\mathbf{x}} + \alpha(\hat{h}(\hat{\mathbf{x}}))$. Consider the system

$$\dot{\hat{\mathbf{x}}} = \hat{F}(\hat{\mathbf{x}}) := (\hat{A} - \hat{B}KE^{-1})\hat{\mathbf{x}} + \hat{B}\hat{v}(\hat{\mathbf{x}}), \tag{9}$$

where

$$\hat{v}(\hat{\mathbf{x}}) = \begin{cases} 0, & \text{if } \hat{\eta}(\hat{\mathbf{x}}) \geq 0, \\ -\frac{\hat{\eta}(\hat{\mathbf{x}})\hat{G}(\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}})\hat{B}^T\nabla\hat{h}(\hat{\mathbf{x}})}{\|\hat{B}^T\nabla\hat{h}(\hat{\mathbf{x}})\|^2_{\hat{G}(\hat{\mathbf{x}})^{-1}}}, & \text{if } \hat{\eta}(\hat{\mathbf{x}}) < 0 \end{cases} \tag{10}$$

Then,

i) $\hat{\mathcal{C}}$ is forward invariant under system (9) and $\mathcal{C}$ is forward invariant under system (7);

ii) system (9) is locally Lipschitz and system (7) is locally Lipschitz;

iii) (A, B) is stabilizable if and only if $(\hat{A}, \hat{B})$ is stabilizable;

iv) $\hat{\mathbf{p}} \in \mathbb{R}^2$ is an undesired equilibrium of (9) if and only if $\mathbf{p} := E^{-1}\hat{\mathbf{p}}$ is an undesired equilibrium of (7);

v) the Jacobian of $\hat{F}$ at $\hat{\mathbf{p}}$ and the Jacobian of F at $\mathbf{p}$ are similar. □

Proof: To show i), note that system (9) satisfies $\nabla\hat{h}(\hat{\mathbf{x}})^T\hat{F}(\hat{\mathbf{x}}) + \alpha(\hat{h}(\hat{\mathbf{x}})) \geq 0$ and system (7) satisfies $\nabla h(\mathbf{x})^T F(\mathbf{x}) + \alpha(h(\mathbf{x})) \geq 0$. These two conditions imply, respectively, that $\hat{\mathcal{C}}$ is forward invariant under system (9) and $\mathcal{C}$ is forward invariant under system (7) [3, Theorem 2]. To show ii), recall that Assumption 2 implies the right-hand-side of (7) is locally Lipschitz. Now let us show that Assumption 2 also holds for $\hat{h}$ and system (9), from which the result follows. Indeed, suppose that $(EB)^\top(\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) = 0$. Since for any $\hat{\mathbf{x}} \in \mathbb{R}^n$ there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\hat{\mathbf{x}} = E\mathbf{x}$ and Assumption 2 holds, we have $0 = (EB)^\top(\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) = B^\top E^\top E(\mathbf{x} - \mathbf{x}_c) = BP(\mathbf{x} - \mathbf{x}_c) = 0$, which means that $(\hat{\mathbf{x}}-\hat{\mathbf{x}}_c)^\top EAE^{-1}\hat{\mathbf{x}} = (\mathbf{x}-\mathbf{x}_c)^\top E^\top EA\mathbf{x} = (\mathbf{x}-\mathbf{x}_c)^\top P^\top A\mathbf{x} > 0$. Hence Assumption 2 holds for $\hat{h}$ and (9), from which it follows that (9) is locally Lipschitz. Item iii) follows from the observation that if $A-BK$ is Hurwitz then $\hat{A}-\hat{B}KE^{-1} = E(A-BK)E^{-1}$ is also Hurwitz. For iv), it follows that $\hat{\mathbf{p}}$ satisfies the conditions in Lemma 1 for (9) if and only if $\mathbf{p}$ satisfies the conditions in Lemma 1 for (7). For v), we note that $F(E^{-1}\hat{\mathbf{x}}) = E^{-1}\hat{F}(\hat{\mathbf{x}})$ for any $\hat{\mathbf{x}} \in \mathbb{R}^2$. Since the safety filter is active at undesired equilibria, $\eta(\mathbf{p}) < 0$. Now, let J be the Jacobian of (7) at $\mathbf{p}$, and let $\hat{J}$ be the Jacobian of (9) at $\hat{\mathbf{p}}$. By the chain rule, $\hat{J} = EJE^{-1}$, which implies that J and $\hat{J}$ are similar. ■

Given that Proposition 1 ensures that undesired equilibria for general ellipsoidal obstacles have the same stability properties as undesired equilibria for circular obstacles, in the following we focus on studying the dynamical properties of safety filters for LTI systems and circular obstacles.

Accordingly, we consider the circular unsafe set:

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 \;:\; h(\mathbf{x}) = \|\mathbf{x}-\mathbf{x}_c\|^2 - r^2 \geq 0\},$$

with $\mathbf{x}_c \in \mathbb{R}^2$ the center. We take the extended class $\mathcal{K}_\infty$ function in Definition 1 to be linear and with slope $\alpha_0 > 0$, so that $\alpha'(0) > 0$. We denote the eigenvalues of $\tilde{A}$ as λ_1, $\lambda_2 \in \mathbb{C}^2$. Let $V(\mathbf{x}) = \mathbf{x}^\top Q\mathbf{x}$ be the associated Lyapunov function, with a positive definite symmetric matrix Q, such that $\mathbf{x}^\top Q\tilde{A}\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}_2$. Additionally, we pick $G(\mathbf{x}) = B^\top B$.

The first result rules out the existence of limit cycles.

Proposition 2: (Non-existence of limit cycles): Suppose that Assumptions 1–3 hold for the closed-loop system (7). Assume that for (7), $\hat{\mathcal{E}} = \{\hat{\mathbf{x}}^*\}$, with $\hat{\mathbf{x}}^*$ a saddle point. Then, there exist $\alpha_1^* > 0$ such that for any $\alpha(s) = \alpha_0 s$ with $\alpha_0 \geq \alpha_1^*$, the closed-loop system does not have limit cycles in $\mathcal{C}$. □

Proof: First, let us show that the closed-loop system does not have limit cycles not circling the origin. Let $\gamma \subset \mathcal{C}$ be such limit cycle and let us reach a contradiction. Note that since $\mathcal{C}$ is forward invariant, such limit cycle has to be contained in $\mathcal{C}$. Now, we consider two cases. First, assume that γ is not circling the obstacle. Note that γ can not contain an undesired equilibrium, since otherwise it would not be a limit cycle. Since the undesired equilibria of the closed-loop system are located at the boundary of the obstacle, γ does not encircle any equilibrium point. However, this contradicts [16, Corollary 6.26]. Next, consider the case where γ encircles the obstacle. Once again, γ can not contain an undesired equilibrium. Therefore, the undesired equilibrium of the closed-loop system is encircled by γ. However, as shown in [14, Theorem 6, Section 3.12], the index of a saddle point is -1 and therefore by [14, Section 3.12, Theorem 2 and Theorem 3], γ can not contain a saddle point. Therefore, there does not exist a limit cycle not circling the origin.

Now, let $\gamma \subset \mathcal{C}$ be a limit cycle encircling the origin and the obstacle. In this case, γ encircles two equilibria: the origin (which is asymptotically stable, because the safety filter is inactive in a neighborhood of it, and therefore has index 1 [14, Thm. 6, Section 3.12]) and the undesired equilibrium in the boundary of $\mathcal{C}$, which is a saddle point and therefore has index 1 [14, Thm. 6, Section 3.12]). However, this contradicts [14, Section 3.12, Thm. 2 and Thm. 3].

Finally, consider the case where γ encircles the origin, but not the obstacle. Recall that we have the Lyapunov function $V(\mathbf{x}) = \mathbf{x}^\top Q\mathbf{x}$ with symmetric Q, such that $\mathbf{x}^\top Q\tilde{A}\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}_2$. Then there exists $\mathbf{q} \in \partial\mathcal{C} = \{\mathbf{x} :\; h(x) = 0\}$ such that $Q\mathbf{q} = -\xi(\mathbf{q}-\mathbf{x}_c)$, $\xi > 0$. It follows that $\eta(\mathbf{q}) = \nabla h(\mathbf{q})\tilde{A}\mathbf{q} = -\frac{2}{\xi}\mathbf{q}^\top Q\tilde{A}\mathbf{q} > 0$, since $\mathbf{q} \neq \mathbf{0}_2$ by Assumption 1. By continuity, there exists an open neighborhood $N(\mathbf{q})$ of $\mathbf{q}$, such that $\eta(\mathbf{x}) > 0$ for all $\mathbf{x} \in N(\mathbf{q})$. Next we show that $\{\mathbf{x} :\; V(\mathbf{x}) \leq V(\mathbf{q}), h(\mathbf{x}) = 0\} = \{\mathbf{q}\}$. Define $H_1(\mathbf{x}) = V(\mathbf{x}) + \xi h(\mathbf{x})$, then $H_1(\mathbf{x})$ is strongly convex in $\mathbb{R}^2$ and $\nabla H_1(\mathbf{q}) = 0$. It follows that $\mathbf{q}$ is the unique global minimizer of H_1 in $\mathbb{R}^2$, which implies that $\{\mathbf{x} :\; V(\mathbf{x}) \leq V(\mathbf{q}), h(\mathbf{x}) = 0\} = \{\mathbf{q}\}$. Then the set $\Gamma_{V(\mathbf{q})} \setminus N(\mathbf{q})$ is a compact subset of $\mathrm{int}(\mathcal{C})$, where $\Gamma_{V(\mathbf{q})} := \{\mathbf{x} :\; V(\mathbf{x}) \leq V(\mathbf{q})\}$. It follows that $\exists\ \alpha_1^* > 0$, such that $\alpha_1^* \geq \min_{\mathbf{x} \in \Gamma_{V(\mathbf{q})} \setminus N(\mathbf{q})} \frac{-\nabla h(\mathbf{x})^\top \tilde{A}\mathbf{x}}{h(\mathbf{x})}$. Hence if $\alpha(s) = \alpha_0 s$ with $\alpha_0 \geq \alpha_1^*$, $\eta(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \Gamma_{V(\mathbf{q})}$. Thus for any solution $\mathbf{x}(t)$ with $\mathbf{x}(0) \in \Gamma_{V(\mathbf{q})}$, we have $\lim_{t\to+\infty} \mathbf{x}(t;\mathbf{x}_0) = \mathbf{0}_2$. However, as γ encircles the origin, but not the obstacle, and $\Gamma_{V(\mathbf{q})} \cap \partial\mathcal{C} \neq \varnothing$, it follows that γ must intersect with $\Gamma_{V(\mathbf{q})}$. Let $\mathbf{q}_0 \in \gamma \cap \Gamma_{V(\mathbf{q})}$, then the solution $\mathbf{x}(t)$ with $\mathbf{x}(0) = \mathbf{q}_0$ converge to the origin, which contradicts with γ is a limit cycle. ■

By combining the results in this section, we have the following.

Theorem 1 (Global behavior analysis): Suppose that the Assumptions 1–3 hold for the closed-loop system (7). Assume that $\hat{\mathcal{E}} = \{\hat{\mathbf{x}}^*\}$ and $\hat{\mathbf{x}}^*$ is a saddle point. Then, there exists $\alpha_2^* > 0$ such that for any $\alpha(s) = \alpha_0 s$ with $\alpha_0 \geq \alpha_2^*$, if $W_s(\hat{\mathbf{x}}^*)$ denotes the global stable manifold of $\hat{\mathbf{x}}^*$ it holds that:

1) if $\mathbf{x}_0 \in W_s(\hat{\mathbf{x}}^*)$, then $\lim\limits_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \hat{\mathbf{x}}^*$;
2) if $\mathbf{x}_0 \notin W_s(\hat{\mathbf{x}}^*)$, then $\lim\limits_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \mathbf{0}_2$. □

Proof: Let α_1^* be as described in Proposition 2. We first note that there exists $\alpha_2^* \geq \alpha_1^*$ such that for all $\alpha_0 \geq \alpha_2^*$, $\{\mathbf{x} :\; \eta(\mathbf{x}) \leq 0\}$ is compact, as $\eta(\mathbf{x}) = 2(\mathbf{x}-\mathbf{x}_c)^\top A\mathbf{x} + \alpha_0(\|\mathbf{x}-\mathbf{x}_c\|^2 - r^2)$ is radially unbounded for α_0 large enough. Next, recalling that $V(\mathbf{x}) = \mathbf{x}^\top P\mathbf{x}$, then there exists a sublevel set Γ_0 of V with $(\mathbb{R}^n \backslash \mathcal{C}) \cup \{\mathbf{x} :\; \eta(\mathbf{x}) \leq 0, \alpha(s) = \alpha_2^* s\} \subset \mathrm{Int}(\Gamma_0)$. It follows that for all $\alpha_0 \geq \alpha_2^*$, $(\mathbb{R}^n \backslash \mathcal{C}) \cup \{\mathbf{x} :\; \eta(\mathbf{x}) \leq 0, \alpha(s) = \alpha_0 s\} \subset \mathrm{Int}(\Gamma_0)$. This implies that by taking $\alpha_0 \geq \alpha_2^*$, Γ_0 is forward-invariant. Similarly, we have that any sublevel set of V Γ_1 satisfying $\Gamma_0 \subseteq \Gamma_1$, is forward-invariant. Hence no trajectory goes to infinity and the solution $\mathbf{x}(t;\mathbf{x}_0)$ is defined for all $t \geq 0$ and is therefore maximal. Now, by Proposition 2, for all $\alpha_0 \geq \alpha_2^*$, we have $\alpha_0 \geq \alpha_1^*$ and hence all trajectories with initial conditions in $\mathcal{C}$ are not limits cycles. Moreover, the Stable Manifold Theorem [14, Ch. 2.7] and the definition of global stable manifold [14, Ch. 2.7, Def. 3] ensure that $\lim\limits_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \hat{\mathbf{x}}^*$ if and only if $\mathbf{x}_0 \in W_s(\hat{\mathbf{x}}^*)$. Therefore, by taking $\alpha_0 \geq \alpha_2^*$, the Poincaré-Bendixson Theorem [17, Chapter 7, Thm. 4.1] ensures that all trajectories with initial condition outside of $W_s(\hat{\mathbf{x}}^*)$ converge to the origin. ■

Remark 1 (Almost global asymptotic stability): The Stable Manifold Theorem [14, Ch. 2.7] ensures that if $\hat{\mathbf{x}}^*$ is a saddle point in $\mathbb{R}^2$, the local stable manifold is 1-dimensional. Therefore, it has measure of zero. Moreover, the global stable manifold must also have measure of zero. If this were not the case, solutions would have to intersect. However this is not possible due to the uniqueness of solutions. Hence $\{\mathbf{x}_0 \in \mathbb{R}^n \ : \ \lim_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \mathbf{0}_n\} = S$. It follows that the set of initial conditions whose associated trajectory converges to $\hat{\mathbf{x}}^*$ has measure zero. □

A. Under-actuated LTI Planar Systems

In the under-actuated case, we write

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad (11)$$

Throughout this section, we denote $\mathbf{x}_c = [x_{c,1}, x_{c,2}]^\top$ and let $\beta = a_{11}b_2 - b_1 a_{21}$, $\gamma = a_{22}b_1 - b_2 a_{12}$, and $T_3 = -\gamma x_{c,2} + \beta x_{c,1}$ and assume that $k : \mathbb{R}^2 \to \mathbb{R}$ is a linear stabilizing controller of the form $k(\mathbf{x}) = -K\mathbf{x} = -k_1 x_1 - k_1 x_2$ for some $k_1, k_2 \in \mathbb{R}$. We note also that since in this case G is a scalar, (7) is independent of G.

The following results give conditions on h and system (11) that ensure that Assumptions 1 and 2 hold.

Lemma 3 (Conditions for Assumption 1): Assumption 1 holds if and only if $\|\mathbf{x}_c\|^2 > r^2$. □

The proof of Lemma 3 follows from the observation that $\|\mathbf{x}_c\|^2 > r^2$ guarantees that the origin is safe.

Proposition 3 (Conditions for Assumption 2): Let $\alpha_0 > 0$, $T_1 := b_2\beta + b_1\gamma + \frac{1}{2}\alpha_0(b_2^2 + b_1^2)$, and $T_2 := (\beta x_{c,1} - \gamma x_{c,2})^2 + 2\alpha_0 r^2 T_1$. Suppose that $r > 0$, $b_1^2 + b_2^2 > 0$, $T_1 > 0$, and

$$\frac{r}{\sqrt{b_2^2 + b_1^2}} > \frac{|T_3| + \sqrt{T_2}}{2T_1}.$$

Then, Assumption 2 holds with the linear extended class $\mathcal{K}$ function $\alpha(s) = \alpha_0 s$. □

Proof: We need to ensure that all $\mathbf{x} \in \mathbb{R}^2$ such that $h(\mathbf{x}) \geq 0$ and $B^T(\mathbf{x} - \mathbf{x}_c) = 0$, satisfy $2(\mathbf{x} - \mathbf{x}_c)A\mathbf{x} + \alpha(h(\mathbf{x})) > 0$. First suppose $b_1 \neq 0$. Equivalently, we need to ensure that

$$\begin{aligned}&(x_2 - x_{c,2})^2\Big((a_{11} + \frac{\alpha_0}{2})\frac{b_2^2}{b_1^2} - \frac{b_2}{b_1}(a_{12} + a_{21}) + a_{22} + \frac{\alpha_0}{2}\Big) \\ &+(x_2 - x_{c,2})\Big(a_{22}x_{c,2} - \frac{b_2}{b_1}a_{11}x_{c,1} - \frac{b_2}{b_1}a_{12}x_{c,2} + a_{21}x_{c,1}\Big) \\ &- \frac{1}{2}\alpha_0 r^2 > 0 \end{aligned} \qquad (12)$$

whenever $(x_2 - x_{c,2})^2 \geq r^2/((b_2^2/b_1^2) + 1)$. This follows by assumption. The condition $T_1 > 0$ ensures that the coefficient of $x_2 - x_{c,2}$ of (12) is positive, and the condition $T_2 > 0$ ensures that the discriminant of (12) is positive. Now, by calculating the roots of the quadratic equation in $x_2 - x_{c,2}$ we observe that the rest of the conditions in the statement ensure that (12) holds whenever $(x_2 - x_{c,2})^2 \geq r^2/((b_2^2/b_1^2) + 1)$. The case $b_1 = 0$ follows by a similar argument. ■

We next give a result that will be used later in the paper.

Lemma 4 (Conditions for β and γ): Let Assumption 3 hold, then $\gamma^2 + \beta^2 > 0$. Furthermore, suppose that the conditions in Proposition 3 hold. Then, $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$. Moreover, if the Assumption 1 holds, then $\gamma x_{c,1} + \beta x_{c,2} \neq 0$.

Proof: First, note that if $\gamma^2 + \beta^2 = 0$, $\gamma = \beta = 0$. This implies that the determinant of the controllability matrix associated with (11) is zero, which contradicts Assumption 3. Now let us show that $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$. By noting that $|T_3| + \sqrt{T_2} > 0$, and squaring both sides of the last two conditions in Proposition 3, we get:

$$|T_3| < \frac{(b_2\beta + b_1\gamma)r}{\sqrt{b_1^2 + b_2^2}}. \qquad (13)$$

Note that (13) requires $b_2\beta + b_1\gamma > 0$ since otherwise the conditions in (13) would not be feasible for any T_3. Now, by using condition (13) and applying the Cauchy-Schwartz inequality, we get $T_3 > -\sqrt{b_1^2 + b_2^2}r$, $T_3 < \sqrt{b_1^2 + b_2^2}r$, from which it follows that $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$. Finally suppose that $\|\mathbf{x}_c\|^2 > r^2$ and $\gamma x_{c,1} + \beta x_{c,2} = 0$. Note that $T_3^2 = (-\gamma x_{c,2} + \beta x_{c,1})^2 = (-\gamma x_{c,2} + \beta x_{c,1})^2 + (\gamma x_{c,1} + \beta x_{c,2})^2 = (\gamma^2 + \beta^2)\|\mathbf{x}_c\|^2$. Since $\|\mathbf{x}_c\|^2 > r^2$, this implies that $r^2(\gamma^2 + \beta^2) - T_3^2 < 0$, which is a contradiction. ■

The following result characterizes the undesired equilibria of the closed-loop system (7) with (11).

Proposition 4: (Equilibria in Under-actuated Systems): Suppose that Assumptions 1, 3 and the conditions in Proposition 3 hold. Define $\mathbf{p}_+ := (\gamma z_+, \beta z_+)$, and $\mathbf{p}_- := (\gamma z_-, \beta z_-)$, where

$$z_+ = \frac{\gamma x_{c,1} + \beta x_{c,2} + \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2}}{\gamma^2 + \beta^2},$$
$$z_- = \frac{\gamma x_{c,1} + \beta x_{c,2} - \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2}}{\gamma^2 + \beta^2}.$$

Then,

i) if $\gamma x_{c,1} + \beta x_{c,2} < 0$, $\mathbf{p}_+$ is the only undesired equilibrium of the closed-loop system (7) with (11);
ii) if $\gamma x_{c,1} + \beta x_{c,2} > 0$, $\mathbf{p}_-$ is the only undesired equilibrium of the closed-loop system (7) with (11).

Proof: By Lemma 4, the expressions for $\mathbf{p}_+$ and $\mathbf{p}_-$ are well-defined (note that if $\gamma^2 + \beta^2 = 0$ the result in Assumption 3 would not hold). Moreover, it follows from Lemma 1 that $\mathbf{p}_+$ and $\mathbf{p}_-$ are the only two potential undesired equilibria for system (7) with (11). In order to ensure that $\mathbf{p}_+$ is an undesired equilibrium of the closed-loop system, the condition $(\mathbf{x} - \mathbf{x}_c)^T(A - BK)\mathbf{x}|_{\mathbf{x}=\mathbf{p}_+} < 0$ should hold. By using the expression of $\mathbf{p}_+$, the condition is equivalent to

$$z_+ T_4\Big(b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2})\Big) < 0, \qquad (14)$$

where $T_4 = a_{11}a_{22} - a_{12}a_{21} - k_1\gamma - k_2\beta$. Since $A - BK$ is Hurwitz, $a_{11}a_{22} - a_{12}a_{21} - \gamma k_1 - \beta k_2 > 0$. This implies that $T_4 > 0$ and therefore (14) is equivalent to

$$z_+(b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2})) < 0. \qquad (15)$$

Now, let us show that $b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2}) > 0$. Indeed, this is equivalent to

$$T_3(\gamma b_2 - \beta b_1) + (\gamma b_1 + \beta b_2)\sqrt{r^2(\gamma^2+\beta^2) - T_3^2} > 0,$$

and since $\gamma b_1 + \beta b_2 > 0$ as argued in the proof of Lemma 4, this could only not hold if $T_3(\gamma b_2 - \beta b_1) < 0$ and $(\gamma b_1 + \beta b_2)^2(r^2(\gamma^2+\beta^2) - T_3^2) \leq T_3^2(\gamma b_2 - \beta b_1)^2$. However, this last condition enters in contradiction with (13). Therefore, (15) holds if and only if $z_+ < 0$, which is equivalent to: $\gamma x_{c,1} + \beta x_{c,2} < 0$ and

$$|\gamma x_{c,1} + \beta x_{c,2}| > \sqrt{r^2(\gamma^2+\beta^2) - T_3^2}. \tag{16}$$

Note that since $r^2(\gamma^2+\beta^2) - T_3^2 = (x_{c,1}\gamma + x_{c,2}\beta)^2 - (\gamma^2+\beta^2)(x_{c,1}^2 + x_{c,2}^2 - r^2) < (x_{c,1}\gamma + x_{c,2}\beta)^2$ (where in the last inequality we have used the fact that $x_{c,1}^2 + x_{c,2}^2 > r^2$), it follows that the last of the inequalities in (16) always holds. Therefore, $\mathbf{p}_+$ is an undesired equilibrium of the closed-loop system if and only if $\gamma x_{c,1} + \beta x_{c,2} < 0$. Also by following an analogous argument, we can show that $\mathbf{p}_-$ is an undesired equilibrium if and only if $\gamma x_{c,1} + \beta x_{c,2} > 0$. ■

Note that by Lemma 4, $\gamma x_{c,1} + \beta x_{c,2} \neq 0$. Therefore Proposition 4 shows that for linear, planar, underactuated and stabilizable linear systems, (2) has exactly one undesired equilibrium. Note also that the result in Proposition 4 is independent of the linear stabilizing controller k and the extended class $\mathcal{K}$ function α chosen.

The following result establishes that the undesired equilibrium of the closed-loop system is always a saddle point.

Proposition 5: (Undesired Equilibria are Saddle Points): Suppose that Assumptions 1, 3 and the conditions in Proposition 3 hold. Then there always exists one and only one undesirable equilibrium, which must be a saddle point.

Proof: To analyze the stability of $\mathbf{p}_+$, we compute the Jacobian of (2) at $\mathbf{p}_+$, which is given by:

$$J^+ = A - \frac{B}{(\mathbf{p}_+ - \mathbf{x}_c)^\top B}((\mathbf{p}_+ - \mathbf{x}_c)^\top (A + \alpha_0 \mathbf{I}_n)),$$

where we have used the fact that $\mathbf{p}_+$ satisfies the conditions in Lemma 1. To study the eigenvalues of J^+, we examine its determinant. After some calculations, we get that $\mathrm{Det}(J^+) = -\alpha_0 T^+$, where

$$T^+ := \frac{\sqrt{r^2(\gamma^2+\beta^2) - T_3^2}(\beta^2+\gamma^2)}{T_3(b_2\gamma - b_1\beta) + (b_1\gamma + b_2\beta)\sqrt{r^2(\gamma^2+\beta^2) - T_3^2}},$$

By Lemma 4, $T^+ \neq 0$. Moreover, as shown in the proof of Proposition 4, $T_3(b_2\gamma - b_1\beta) + (b_1\gamma + b_2\beta)\sqrt{r^2(\gamma^2+\beta^2) - T_3^2} > 0$. Since $\alpha_0 > 0$, this implies that $\mathrm{Det}(J^+) < 0$, which implies that J^+ has two real eigenvalues with opposite sign and hence if $\mathbf{p}_+$ is an undesired equilibrium, it is a saddle point. An analogous argument shows that if $\mathbf{p}_-$ is an undesired equilibrium, it is a saddle point. ■

Note that the results in Propositions 4 and 5 are independent of the choice of weighting matrix G, nominal controller k or extended class $\mathcal{K}$ function α. The combination of Propositions 4 and 5 with Theorem 1 provide a complete picture of the under-actuated case, which we summarize as follows.

Corollary 1: (Characterization of trajectories for linear planar underactuated systems): Suppose that Assumptions 1, 3 and the conditions in Proposition 3 hold. Then, the closed-loop system (7) obtained from (11) has one and only one undesired equilibrium $\hat{\mathbf{x}}^*$ equal to either $\mathbf{p}_+$ or $\mathbf{p}_-$. Additionally, there exists $\alpha_2^* > 0$ such that for any $\alpha(s) = \alpha_0 s$ with $\alpha_0 \geq \alpha_2^*$, if $W_s(\hat{\mathbf{x}}^*)$ denotes the global stable manifold of $\hat{\mathbf{x}}^*$ it holds that:

1) if $\mathbf{x}_0 \in W_s(\hat{\mathbf{x}}^*)$, then $\lim_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \hat{\mathbf{x}}^*$;
2) if $\mathbf{x}_0 \notin W_s(\hat{\mathbf{x}}^*)$, then $\lim_{t\to\infty} \mathbf{x}(t;\mathbf{x}_0) = \mathbf{0}_2$. □

B. Fully Actuated LTI Planar Systems

We now consider the case where B is invertible; in this case, Assumptions 2 and 3 are satisfied. The proofs for the results presented in this section are postponed to the appendix.

1) $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$: We start by considering two conditions for the case where $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$.

Condition 1. $\lambda_1 < \lambda_2 < 0$, $\tilde{A}\mathbf{x}_c = \lambda_2 \mathbf{x}_c$, $\tilde{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\mathbf{v}_2 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2\|\mathbf{x}_c\|^2} = 0$, $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2\|\mathbf{x}_c\|^2}$.

Condition 2. $\lambda_1 < \lambda_2 < 0$, $\tilde{A}\mathbf{x}_c = \lambda_2 \mathbf{x}_c$, $\tilde{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\mathbf{v}_2 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2\|\mathbf{x}_c\|^2} = 0$, $(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2\|\mathbf{x}_c\|^2}$.

We have that there exists only one undesirable equilibrium and it is a degenerate equilibrium if and only if *Condition 1* is true. If *Condition 2* is true, there are two undesirable equilibria, one of which is a saddle point and the other one is a degenerate equilibrium.

If neither *Condition 1* nor *Condition 2* is true, we summarize the results about undesirable equilibria for the case that $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$ in Tables I and II. We gather all the cases in the following result.

Proposition 6: (Characterization of undesired equilibria): Let Assumptions 1 be satisfied and B be invertible. Given that $\tilde{A}$ is stable and $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$, then one of the following is true:

(i) $|\mathcal{E}| = 2$, $|\hat{\mathcal{E}}| = 1$, $\mathbf{x} \in \hat{\mathcal{E}}$ is a degenerate equilibrium.
(ii) $|\mathcal{E}| = 2$, $|\hat{\mathcal{E}}| = 1$, $\mathbf{x} \in \hat{\mathcal{E}}$ is a saddle point.
(iii) $|\mathcal{E}| = 3$, $|\hat{\mathcal{E}}| = 2$, one point in $\hat{\mathcal{E}}$ is a saddle point and the other point in $\hat{\mathcal{E}}$ is a degenerate equilibrium.
(iv) $|\mathcal{E}| = 4$, $|\hat{\mathcal{E}}| = 3$, two points in $\hat{\mathcal{E}}$ are saddle points and the other point in $\hat{\mathcal{E}}$ is asymptotically stable. □

Proposition 6 asserts that the number and the stability property of the undesirable equilibria are determined by the number of solutions of (5), if $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$.

Proposition 7: (Spectrum of $\tilde{A}$ does not determine stability properties of undesired equilibria): Let Assumptions 1 be satisfied and B be invertible. Then for any given negative λ_1 and λ_2, there exists K_1 and K_2 in the set $\{K : \lambda_1, \lambda_2 = \mathrm{spec}(A - BK)\}$, such that there is an undesirable asymptotically stable equilibrium after applying the CBF filter with $u = -K_1\mathbf{x}$; and there is only one undesirable equilibrium and it is a saddle point after applying the CBF filter with $u = -K_2\mathbf{x}$. □

	SP	DE	ASE
$(\mathbf{v}_1^\top \mathbf{v}_2)^2 < 1 - \frac{r^2}{\lambda^2 \|\mathbf{x}_c^2\|}$	1	0	0
$(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - \frac{r^2}{\lambda^2 \|\mathbf{x}_c^2\|}$	1	1	0
$(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - \frac{r^2}{\lambda^2 \|\mathbf{x}_c^2\|}$	2	0	1

TABLE I

$\tilde{A}$ STABLE, $\tilde{A}\mathbf{v}_2 = \lambda \mathbf{v}_2 + \mathbf{v}_1$, $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\tilde{A}\mathbf{x}_c = \lambda \mathbf{x}_c$, $\|v_2\| = 1$. SP: SADDLE POINT, DE: DEGENERATE EQUILIBRIUM, ASE: UNDESIRABLE ASYMPTOTICALLY STABLE EQUILIBRIUM.

	SP	DE	ASE
$(\mathbf{v}_i^\top \mathbf{v}_j)^2 < 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2}$	1	0	0
$(\mathbf{v}_i^\top \mathbf{v}_j)^2 = 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2}$	1	1	0
$(\mathbf{v}_i^\top \mathbf{v}_j)^2 > 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2}$	2	0	1

TABLE II

$\tilde{A}$ STABLE, $\tilde{A}\mathbf{x}_c = \lambda_i \mathbf{x}_c$, $\mathbf{v}_i = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\tilde{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$, $\|\mathbf{v}_j\| = 1$, $i, j = \{1, 2\}$, $\{\mathbf{v}_i, \mathbf{v}_j\}$ LINEARLY INDEPENDENT.

Note that one can characterize the global stability property of the origin based on the eigenvalues of $A - BK$. However, based on Proposition 7, the eigenvalues of $A - BK$ do not fully determine the global stability property of the origin. On the other hand, Proposition 7 shows that there always exists a nominal controller $\mathbf{u} = -K\mathbf{x}$ such that $\tilde{A}$ has negative eigenvalues and the set of trajectories of (7) that do not converge to the origin has measure zero (cf. Theorem 1). Note that as shown in Lemma 2 and Tables I, II, the class $\mathcal{K}$ function only affects the rate of decay in the stable manifold of the undesirable equilibria and it does not affect the existence and stability of undesirable equilibria. Therefore, the choice of nominal controller $\mathbf{u} = -K\mathbf{x}$ determines in which of the cases we fall into. Ideally, the controller should be designed to that there exists only one undesirable equilibrium and it is a saddle point.

2) $\mathbf{x}_c$ is not an eigenvector of $\tilde{A}$: Next, we analyze the number of undesirable equilibria when $\mathbf{x}_c$ is not an eigenvector of $\tilde{A}$. In this case, the analysis is more involved and we only study the stability properties of undesired equilibria under some sufficient conditions.

Proposition 8 (Number of undesired equilibria): Let Assumptions 1 be satisfied and B be invertible. Given that $G(\mathbf{x}) = B^\top B$ and $\tilde{A}$ is stable and $\mathbf{x}_c$ is not an eigenvector of $\tilde{A}$, then $1 \leq |\hat{\mathcal{E}}| \leq 3$ and $|\mathcal{E} \setminus \hat{\mathcal{E}}| \geq 1$. In addition, if $\lambda_1 \leq \lambda_2$, there exists $\mathbf{x} \in \hat{\mathcal{E}}$ with indicator $\delta < \frac{\lambda_1}{2}$. □

Combining Propositions 1, 4, 8 and Table I, II, it follows that applying the CBF filter to a LTI planar system (either under or fully actuated) with a linear stabilizing controller always introduces at least one undesirable equilibrium when the obstacle is ellipsoidal. By [18, Thm. 9.5] and Lemma 2, there exists at least one trajectory converging to the undesirable equilibrium. This result is consistent with [19], which states that given a local Lipschitz dynamical system and a compact unsafe set, if the safe set is forward invariant then there exists at least one trajectory that does not converge to the origin. Theorem 1 ensures that if there is only one undesirable equilibrium and it is a saddle point, then there is only one such trajectory and it corresponds to the global stable manifold of the undesired equilibrium.

To analyze the stability of undesirable equilibria in the case that $\mathbf{x}_c$ is not an eigenvector of $\tilde{A}$, we need to determine the eigenvalue of $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}}$. By Lemma 2, we know that $-\alpha'(0) = -\alpha_0$ is always an eigenvalue of $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}}$. In addition, for any linearly independent vectors $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathbb{C}^2$, we can compute $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}} \mathbf{v}_1 = d_{11}\mathbf{v}_1 + d_{21}\mathbf{v}_2$ and $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}} \mathbf{v}_2 = d_{12}\mathbf{v}_1 + d_{22}\mathbf{v}_2$. Then the trace of $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}}$ is equal to $d_{11} + d_{22}$ and the other eigenvalue of $J\mid_{\mathbf{x}\in\hat{\mathcal{E}}}$ is equal to $d_{11} + d_{22} + \alpha'(0)$. By choosing some specific $\mathbf{v}_1$ and $\mathbf{v}_2$, we obtain the following result.

Lemma 5: (The other eigenvalue of the Jacobian): Assume that $\lambda_1 \neq \lambda_2$, then there exists $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathbb{C}^2$ such that $\|\mathbf{v}_1\|_2 = \|\mathbf{v}_2\|_2 = 1$, $\tilde{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and $\tilde{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$. For any $\mathbf{x} \in \hat{\mathcal{E}}$, if the associated indicator $\delta_\mathbf{x} \notin \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$, $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$ and $\mathbf{x} = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$, then it holds that $\beta_3 - \beta_1 = \frac{-\lambda_1 \beta_1}{\lambda_1 - 2\delta_\mathbf{x}}$, $\beta_4 - \beta_2 = \frac{-\lambda_2 \beta_2}{\lambda_2 - 2\delta_\mathbf{x}}$ and the eigenvalue other than $-\alpha'(0)$ of $J\mid_\mathbf{x}$ is $\lambda_1 + \lambda_2 - 2\delta_\mathbf{x} - \frac{(\beta_3 - \beta_1)\lambda_1}{r^2}\Delta_1 - \frac{(\beta_4 - \beta_2)\lambda_2}{r^2}\Delta_2$, where $\Delta_1 := (\beta_3 - \beta_1)^* + (\beta_4 - \beta_2)^* \mathbf{v}_2^* \mathbf{v}_1$, $\Delta_2 := (\beta_3 - \beta_1)^* \mathbf{v}_1^* \mathbf{v}_2 + (\beta_4 - \beta_2)^*$. □

Using Lemma 5, we get the following result.

Proposition 9: (Sufficient conditions for undesired equilibria): Let Assumptions 1 be satisfied and B be invertible. Given that $G(\mathbf{x}) = B^\top B$, $\tilde{A}$ is stable with two real eigenvalues $\lambda_1 < \lambda_2$ and $\mathbf{x}_c$ is not an eigenvector of $\tilde{A}$, then there is no undesirable equilibrium with indicator $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. Besides, let $\mathbf{v}_1$ and $\mathbf{v}_2$ be the eigenvectors associated with λ_1 and λ_2, respectively, and $\mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$; and then we can write $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$. Then, the following holds.

i) If $\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, then for any undesirable equilibrium $\mathbf{x}$ with indicator δ such that $\delta < \frac{\lambda_1}{2}$, $\mathbf{x}$ is a saddle point.

ii) If $\beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 + \beta_2^2 \geq 0$, then for any undesirable equilibrium $\mathbf{x}$ with indicator δ such that $\frac{\lambda_2}{2} < \delta < 0$, $\mathbf{x}$ is asymptotically stable.

iii) Define $F_1 : \mathbb{R} \to \mathbb{R}$ as:

$$\begin{aligned} F_1(\delta) := &-|\lambda_1 - 2\delta|^2 |\lambda_2 - 2\delta|^2 r^2 \\ &+ |\lambda_1 \beta_1|^2 |\lambda_2 - 2\delta|^2 + |\lambda_2 \beta_2|^2 |\lambda_1 - 2\delta|^2 \\ &+ \lambda_1^* \beta_1^* \lambda_2 \beta_2 (\lambda_2 - 2\delta)^* (\lambda_1 - 2\delta) \mathbf{v}_1^* \mathbf{v}_2 \\ &+ \lambda_1 \beta_1 \lambda_2^* \beta_2^* (\lambda_2 - 2\delta)(\lambda_1 - 2\delta)^* \mathbf{v}_2^* \mathbf{v}_1. \end{aligned} \tag{17}$$

If the third order polynomial $\frac{dF_1(\delta)}{d\delta}$ has only one real root[1] and $\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, then there exists only one undesirable equilibrium and it is a saddle point. □

If $|\beta_1| \gg |\beta_2|$ and $|\mathbf{v}_1^\top \mathbf{v}_2|$ is small, then the case $\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \geq 0$ is a generalized version of the case in the first row of Table II. If $|\beta_2| \gg |\beta_1|$ (i.e., $\mathbf{x}_c$ is "essentially" eigenvector associated with λ_2) and $\lambda_1 \ll \lambda_2$, then the case $\beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 + \beta_2^2 \geq 0$ is a generalized version of the case in the last row of Table II, as $1 - \frac{(\lambda_2 - \lambda_1)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} < 0$ with $\lambda_1 \ll \lambda_2$.

[1]For third-order polynomial $ax^3 + bx^2 + cx + d$, its discriminant is defined as $18abcd - 4b^3 d + b^2 c^2 - 4ac^3 - 27a^2 d^2$. If $a \neq 0$ and the discriminant is negative, the third-order polynomial only has one real root.

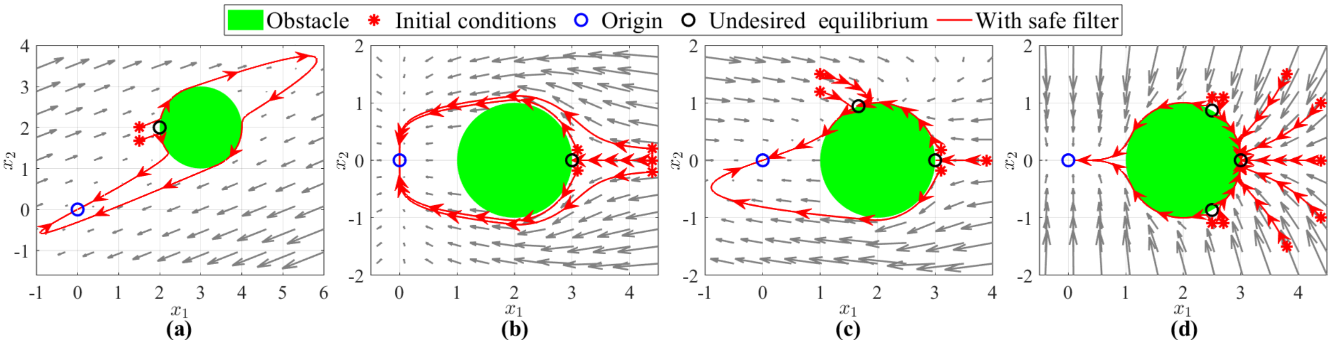

Fig. 1. Examples of trajectories of an LTI planar system with a safety filter for a circular obstacle; the figures show the vector fields, the undesired equilibria, and the desired equilibrium (which is the origin). (a): Under-actuated system. (b)-(c)-(d): Fully actuated system, corresponding to the three rows of Table II respectively. In (a) and (b) the undesirable equilibrium is a saddle point. In (c) there is one degenerate equilibrium and one saddle point. In (d) there are three undesirable equilibria, one is asymptotically stable while the others are saddle points.

V. Numerical Experiments

As a first experiment, we consider the safety set $\mathcal{C} = \{\mathbf{x} : \ \|\mathbf{x} - (3,2)^\top\|^2 - 1 \geq 0\}$ and the under-actuated system $\dot{\mathbf{x}} = \begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} \mathbf{u}$ with nominal controller $\mathbf{u} = -\begin{bmatrix} 3 & -2 \end{bmatrix} \mathbf{x}$. Once the CBF-based filter is applied, there is one undesirable equilibrium $(2,2)^\top$, as guaranteed by Proposition 4. Examples of trajectories of the system with the safety filter are shown in Figure 1(a), along with the vector field, the spurious undesired equilibrium, and the desired equilibrium (which is the origin).

In Figures 1(b), (c) and (d), we consider a safety set $\mathcal{C} = \{\mathbf{x} : \ \|\mathbf{x} - (2,0)^\top\|^2 - 1 \geq 0\}$, and the integrator dynamics $\dot{\mathbf{x}} = \mathbf{u}$ as an example of (6).

In Figure 1(b), we show the results for the integrator dynamics with $K = \begin{bmatrix} -5 & 0 \\ 0 & -1 \end{bmatrix}$, $G(\mathbf{x}) = B^\top B$, and the safety filter with $\alpha(s) = \alpha_0 s$, $\alpha_0 = 10$. There is one undesirable equilibrium $(3,0)^\top$. We note that for both the setups in Figures 1 (a) and (b), there is only one undesirable equilibrium and it is a saddle point. Only one trajectory converges to the undesirable equilibrium and all other trajectories converge to the origin.

In Figure 1(c), we show the results for (7) with $K = \begin{bmatrix} -3 & 4\sqrt{2} \\ 0 & -1 \end{bmatrix}$, $G(\mathbf{x}) = B^\top B$ and $\alpha(s) = \alpha_0 s$, $\alpha_0 = 10$. There two undesirable equilibria, which are $(\frac{5}{3}, \frac{2\sqrt{2}}{3})^\top$ (degenerate equilibrium) and $(3,0)^\top$ (saddle point). Only one trajectory converges to $(3,0)^\top$. The measure of the stable set of the degenerate equilibrium is positive (in fact, the measure is $+\infty$), although the degenerate equilibrium is unstable.

In Figure 1(d), we show that results for (7) with $K = \begin{bmatrix} -1 & 0 \\ 0 & -5 \end{bmatrix}$, $G(\mathbf{x}) = B^\top B$ and $\alpha(s) = \alpha_0 s$, $\alpha_0 = 10$. There are three undesirable equilibria: $(\frac{5}{2}, \frac{\sqrt{3}}{2})^\top$, $(\frac{5}{2}, -\frac{\sqrt{3}}{2})^\top$ and $(3,0)^\top$; the last one is asymptotically stable and the first two are saddle points. The two trajectories converging to $(\frac{5}{2}, \frac{\sqrt{3}}{2})^\top$, $(\frac{5}{2}, -\frac{\sqrt{3}}{2})^\top$ and part of the obstacle constitute the boundary of the region of attraction of $(3,0)^\top$. Since the examples in Figure 1(b), (c) and (d) all satisfy that $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$, these results are consistent with Proposition 6 (ii), (iii), (iv), respectively.

References

[1] N. Matni, A. D. Ames, and J. C. Doyle, "Towards a theory of control architecture: A quantitative framework for layered multi-rate control," *arXiv preprint arXiv:2401.15185*, 2024.

[2] P. Wieland and F. Allgöwer, "Constructive safety using control barrier functions," *IFAC Proceedings Volumes*, vol. 40, no. 12, pp. 462–467, 2007.

[3] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: theory and applications," in *European Control Conference*, Naples, Italy, 2019, pp. 3420–3431.

[4] W. Xiao and C. Belta, "High-order control barrier functions," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3655–3662, 2022.

[5] W. Xiao, C. G. Cassandras, and C. Belta, *Safe Autonomy with Control Barrier Functions: Theory and Applications*. Springer, 2023.

[6] L. Wang, A. Ames, and M. Egerstedt, "Safety barrier certificates for collisions-free multirobot systems," *IEEE Transactions on Robotics*, vol. 33, no. 3, pp. 661–674, 2017.

[7] W. S. Cortez and D. V. Dimarogonas, "On compatibility and region of attraction for safe, stabilizing control laws," *IEEE Transactions on Automatic Control*, vol. 67, no. 9, pp. 7706–7712, 2022.

[8] M. F. Reis, A. P. Aguilar, and P. Tabuada, "Control barrier function-based quadratic programs introduce undesirable asymptotically stable equilibria," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 731–736, 2021.

[9] X. Tan and D. V. Dimarogonas, "On the undesired equilibria induced by control barrier function based quadratic programs," *arXiv preprint arXiv:2104.14895*, 2021.

[10] P. Mestres and J. Cortés, "Optimization-based safe stabilizing feedback with guaranteed region of attraction," *IEEE Control Systems Letters*, vol. 7, pp. 367–372, 2023.

[11] Y. Yi, S. Koga, B. Gavrea, and N. Atanasov, "Control synthesis for stability and safety by differential complementarity problem," *IEEE Control Systems Letters*, vol. 7, pp. 895–900, 2023.

[12] G. Notomista and M. Saveriano, "Safety of dynamical systems with multiple non-convex unsafe sets using control barrier functions," *IEEE Control Systems Letters*, vol. 6, pp. 1136–1141, 2022.

[13] E. D. Sontag, *Mathematical control theory: deterministic finite dimensional systems*. Springer Science & Business Media, 2013, vol. 6.

[14] L. Perko, *Differential Equations and Dynamical Systems*, 3rd ed., ser. Texts in Applied Mathematics. New York: Springer, 2000, vol. 7.

[15] M. Alyaseen, N. Atanasov, and J. Cortés, "Continuity and boundedness of minimum-norm CBF-safe controllers," *arXiv preprint arXiv:2306.07398*, 2023.

[16] J. D. Meiss, *Differential Dynamical Systems*. Society for Industrial and Applied Mathematics, 2007.

[17] P. Hartman, *Ordinary Differential Equations*, 2nd ed. SIAM, 2002, no. 38.

[18] G. Teschl, *Ordinary Differential Equations And Dynamical Systems*, ser. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2012, vol. 140.

[19] P. Braun and C. M. Kellett, "Comment on "stabilization with guaranteed safety using control Lyapunov–barrier function"," *Automatica*, vol. 122, p. 109225, 2020.

APPENDIX

Denote the eigenvalues of $A-BK$ as $\lambda_1, \lambda_2 \in \mathbb{C}$. We note that the (5), which is used to check for potential undesirable equilibria), can be rewritten as follows:

$$(\tilde{A} - 2\delta \mathbf{I}_{2\times 2})(\mathbf{x} - \mathbf{x}_c) = -\tilde{A}\mathbf{x}_c \text{ and,} \quad (18)$$
$$\|\mathbf{x} - \mathbf{x}_c\|^2 - r^2 = 0 . \quad \square$$

A. *Proof of the results of the case where $\mathbf{x}_c$ is an eigenvector of $\tilde{A}$*

Denote λ the eigenvalue associated with $\mathbf{x}_c$. Then $\lambda = \lambda_i$, $i = 1$ or 2; and both λ_1 and λ_2 are real. We first determine the solution for (18) with $\delta \notin \{\frac{\lambda_1}{2},\ \frac{\lambda_2}{2}\}$.

Since $\delta \notin \{\frac{\lambda_1}{2},\ \frac{\lambda_2}{2}\}$, by the first equation in (18), it follows that $\mathbf{x} = \frac{2\delta}{2\delta-\lambda}\mathbf{x}_c$. Plugging this in the second equation, we can solve δ. The solutions for (18) are $(\mathbf{x}_{*,-}, \frac{\lambda}{2} + \frac{\lambda\|\mathbf{x}_c\|}{2r})$ and $(\mathbf{x}_{*,+}, \frac{\lambda}{2} - \frac{\lambda\|\mathbf{x}_c\|}{2r})$, where $\mathbf{x}_{*,-} := (1 + \frac{r}{\|\mathbf{x}_c\|})\mathbf{x}_c)$ and $\mathbf{x}_{*,+} := (1 - \frac{r}{\|\mathbf{x}_c\|})\mathbf{x}_c)$.

Notice that $\mathbf{x}_{*,-}$ comes with a negative δ and $\mathbf{x}_{*,+}$ comes with a positive δ, so $\mathbf{x}_{*,-}$ is an undesirable equilibrium while $\mathbf{x}_{*,+}$ is not an undesirable equilibrium. By Lemma 2, the Jacobian at $\mathbf{x}_{*,-}$ is

$$J \mid_{\mathbf{x}_{*,-}} = \tilde{A} - 2\delta \mathbf{I} - \frac{\mathbf{x}_c\mathbf{x}_c^\top}{\|\mathbf{x}_c\|^2}(\tilde{A} - (2\delta - \alpha'(0))\mathbf{I}).$$

where $\delta = \frac{\lambda_1}{2} + \frac{\lambda_1\|\mathbf{x}_c\|}{2r}$.

In the following, we will determine if there exists solutions for (18) with $\delta \in \{\frac{\lambda_1}{2},\ \frac{\lambda_2}{2}\}$ and discuss the stability of undesirable equilibrium case by case.

Case 1 $\tilde{A}$ is not diagonalizable

In this case, we must have $\lambda_1 = \lambda_2$. Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\mathbf{v}_2$ be a vector such that $\|\mathbf{v}_2\| = 1$, $\tilde{A}\mathbf{v}_2 = \lambda_1\mathbf{v}_2 + \mathbf{v}_1$. If we write $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$, then $\beta_1 = \|\mathbf{x}_c\|$ and $\beta_2 = 0$. Using the same technique in the proof of Lemma 5, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_2 - 2\delta_{\mathbf{x}_{*,-}} = \lambda_1 - \lambda_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{r} > 0$, implying that $\mathbf{x}_{*,-}$ is a saddle point.

Next, we determine if there exists a solution with $\delta = \frac{\lambda_1}{2}$. We write $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$. Hence the first equation of (18) with $\delta = \frac{\lambda_1}{2}$ can be rewritten as

$$\beta_4 = -\lambda_1\|\mathbf{x}_c\|. \quad (19)$$

Plugging the value of β_4 into the second equation of (18), it follows that

$$\|(\beta_3 - \|\mathbf{x}_c\|)\mathbf{v}_1 - \lambda_1\|\mathbf{x}_c\|\mathbf{v}_2\|^2 - r^2 = 0.$$

Define $\hat{\beta}_3 := \beta_3 - \|\mathbf{x}_c\|$ and $\tau_1 := \lambda_1\|\mathbf{x}_c\|$, then

$$\hat{\beta}_3^2 - 2\tau_1\mathbf{v}_1^\top\mathbf{v}_2\hat{\beta}_3 + \tau_1^2 - r^2 = 0. \quad (20)$$

Note that the discriminant of quadratic equation (20) is

$$\Delta := 4\left(\tau_1^2(\mathbf{v}_1^\top\mathbf{v}_2)^2 - \tau_1^2 + r^2\right) = 4\left(\tau_1^2((\mathbf{v}_1^\top\mathbf{v}_2)^2 - 1) + r^2\right)$$

Case 1.1 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 < 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c^2\|}$, there doesn't exist a solution associated with $\delta = \frac{\lambda_1}{2}$.

Hence in **Case 1.1**, there is only one undesirable equilibrium, which is a saddle point.

Case 1.2 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 = 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, then there exists one solution with $\delta_{\hat{\mathbf{x}}} = \frac{\lambda_1}{2}$, which is

$$(\hat{\mathbf{x}}, \delta_{\hat{\mathbf{x}}}) = \left((\tau_1\mathbf{v}_1^\top\mathbf{v}_2 + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2, \frac{\lambda_1}{2}\right).$$

We note that $(\hat{\mathbf{x}} - \mathbf{x}_c)^\top\mathbf{v}_1 = 0$ and $(\tilde{A} - 2\delta_{\hat{\mathbf{x}}}\mathbf{I})\mathbf{v}_1 = 0$. Hence $J \mid_{\hat{\mathbf{x}}} \mathbf{v}_1 = 0$.

Thus in **Case 1.2**, there is another one undesirable equilibrium, at which the Jacobian has a negative eigenvalue and a zero eigenvalue.

Case 1.3 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 > 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, there exists two solutions $\hat{\beta}_3 = \hat{\beta}_3^{(1)}$ and $\hat{\beta}_3 = \hat{\beta}_3^{(2)}$ for (20). Then there exists two extra solutions for (18): $\left((\hat{\beta}_3^{(1)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2, \frac{\lambda_1}{2}\right)$ and $\left((\hat{\beta}_3^{(2)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2, \frac{\lambda_1}{2}\right)$ and both of them are undesirable equilibrium.

Notice that in this sub-case, $\hat{\beta}_3^{(1)} + \hat{\beta}_3^{(2)} = 2\tau_1\mathbf{v}_1^\top\mathbf{v}_2$, we can assume that $\hat{\beta}_3^{(1)} < \tau_1\mathbf{v}_1^\top\mathbf{v}_2$ and $\hat{\beta}_3^{(1)} > \tau_1\mathbf{v}_1^\top\mathbf{v}_2$.

Using the same technique in the proof of Lemma 5, we can show that $J \mid_{\mathbf{x}}$, with $\mathbf{x} = \left((\hat{\beta}_3^{(1)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2\right)$, has a eigenvalue $\frac{\tau_1}{r^2}(\hat{\beta}_3^{(1)} - \tau_1\mathbf{v}_1^\top\mathbf{v}_2) > 0$; and $J \mid_{\mathbf{x}}$, with $\mathbf{x} = \left((\hat{\beta}_3^{(2)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2\right)$, has an eigenvalue $\frac{\tau_1}{r^2}(\hat{\beta}_3^{(2)} - \tau_1\mathbf{v}_1^\top\mathbf{v}_2) < 0$.

Hence in **Case 1.3**, there are another two undesirable equilibria, one of which is stable and the other one is saddle point.

Case 2 $\lambda_1 \leq \lambda_2 < 0$, $\tilde{A}\mathbf{x}_c = \lambda_1\mathbf{x}_c$

Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\mathbf{v}_2$ be an eigenvector associated with λ_2 and $\|\mathbf{v}_2\| = 1$, $\mathbf{v}_1^\top\mathbf{v}_2 \geq 0$. If we write $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$, then $\beta_1 = \|\mathbf{x}_c\|$ and $\beta_2 = 0$. By Lemma 5, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_2 - 2\delta_{\mathbf{x}_{*,-}} = \lambda_2 - \lambda_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{r} > 0$, implying that $\mathbf{x}_{*,-}$ is a saddle point.

Next, we determine if there exists a solution with $\delta \in \{\frac{\lambda_1}{2},\ \frac{\lambda_2}{2}\}$. We write $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$ and then the first equation of (18) can be rewritten as

$$\begin{aligned}&(\lambda_1 - 2\delta)(\beta_3 - \|\mathbf{x}_c\|) = -\lambda_1\|\mathbf{x}_c\| \\ &(\lambda_2 - 2\delta)\beta_4 = 0\end{aligned} \quad (21)$$

which follows that $\delta \neq \frac{\lambda_1}{2}$.

If $\delta = \frac{\lambda_2}{2}$, it follows that $\beta_3 = \frac{-\lambda_2\|\mathbf{x}_c\|}{\lambda_1-\lambda_2}$. Plugging the value of β_3 into the second equation of (18), it follows that

$$\|-\tau_2\mathbf{v}_1 + \beta_4\mathbf{v}_2\|^2 - r^2 = 0.$$

where $\tau_2 := \frac{\lambda_1\|\mathbf{x}_c\|}{\lambda_1-\lambda_2}$. Then

$$\beta_4^2 - 2\tau_2\mathbf{v}_1^\top\mathbf{v}_2\beta_4 + \tau_2^2 - r^2 = 0 \quad (22)$$

Note that the discriminant of quadratic equation (22) is

$$\Delta := 4\left(\tau_2^2(\mathbf{v}_1^\top\mathbf{v}_2)^2 - \tau_2^2 + r^2\right) = 4\left(\tau_2^2((\mathbf{v}_1^\top\mathbf{v}_2)^2 - 1) + r^2\right)$$

Case 2.1 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 < 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1-\lambda_2)^2r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, there doesn't exist a solution associated with $\delta = \frac{\lambda_2}{2}$.

Hence in **Case 2.1**, there is only one undesirable equilibrium, which is a saddle point.

Case 2.2 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$, then there exists one solution with $\delta_{\hat{\mathbf{x}}} = \frac{\lambda_2}{2}$, which is

$$(\hat{\mathbf{x}}, \delta_{\hat{\mathbf{x}}}) = \left(\frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1^\top \mathbf{v}_2 \mathbf{v}_2, \frac{\lambda_2}{2} \right).$$

We note that $(\hat{\mathbf{x}} - \mathbf{x}_c)^\top \mathbf{v}_2 = 0$ and $(\tilde{A} - 2\delta_{\hat{\mathbf{x}}} \mathbf{I})\mathbf{v}_2 = 0$. Hence $J \mid_{\hat{\mathbf{x}}} \mathbf{v}_2 = 0$.

Thus in **Case 2.2**, there is another one undesirable equilibrium, at which the Jacobian has a negative eigenvalue and a zero eigenvalue.

Case 2.3 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$, there exists two solutions $\beta_4 = \beta_4^{(1)}$ and $\beta_4 = \beta_4^{(2)}$ for (21). Then there exists two extra solutions for (18): $\left(\frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(1)} \mathbf{v}_2, \frac{\lambda_2}{2} \right)$ and $\left(\frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(2)} \mathbf{v}_2, \frac{\lambda_2}{2} \right)$ and both of them are undesirable equilibrium. Notice that in this sub-case, $\beta_4^{(1)} + \beta_4^{(2)} = 2\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 > 0$ and $\beta_4^{(1)} \beta_4^{(2)} = \tau_2^2 - r^2 > 0$, we can assume that $0 < \beta_4^{(1)} < \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$ and $\beta_4^{(2)} > \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$. It follows that $-\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} + \tau_2^2 - r^2 = -\beta_4^{(1)} \beta_4^{(1)} + \tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} > 0$ and $-\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} + \tau_2^2 - r^2 = -\beta_4^{(2)} \beta_4^{(2)} + \tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} < 0$.

Using the same technique in the proof of Lemma 5, we can show that $J \mid_{\mathbf{x}}$, with $\mathbf{x} = \left(\frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(1)} \mathbf{v}_2 \right)$, has a eigenvalue $\frac{\lambda_2 - \lambda_1}{r^2} (\tau_2^2 - \tau_2^2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} - r^2) > 0$.; and $J \mid_{\mathbf{x}}$, with $\mathbf{x} = \left(\frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(2)} \mathbf{v}_2 \right)$, has a eigenvalue $\frac{\lambda_2 - \lambda_1}{r^2} (\tau_2^2 - \tau_2^2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} - r^2) < 0$.

Hence in **Case 2.3**, there are another two undesirable equilibria, one of which is stable and the other one is saddle point.

Case 3 $\lambda_1 < \lambda_2 < 0$, x_0 is an eigenvector of $A - BK$ associated with λ_2

Let $\mathbf{v}_2 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\mathbf{v}_1$ be an eigenvector associated with λ_1 and $\|\mathbf{v}_1\| = 1$, $\mathbf{v}_1^\top \mathbf{v}_2 \geq 0$. If we write $x_0 = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$, then $\beta_2 = \|\mathbf{x}_c\|$ and $\beta_1 = 0$. By Lemma 5, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_1 - 2\delta_{\mathbf{x}_{*,-}} = \lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r}$.

We will determine the sign of $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r}$ later.

Next, we determine if there exists a solution with $\delta \in \{\frac{\lambda_1}{2}, \ \frac{\lambda_2}{2}\}$. We write $\mathbf{x}_c = \|\mathbf{x}_c\| \mathbf{v}_2$, $\mathbf{x} = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$ and then the first equation of (18) can be rewritten as

$$\begin{aligned} &(\lambda_1 - 2\delta)\beta_3 = 0 \\ &(\lambda_2 - 2\delta)(\beta_4 - \|\mathbf{x}_c\|) = -\lambda_2 \|\mathbf{x}_c\|, \end{aligned} \tag{23}$$

which follows that $\delta \neq \frac{\lambda_2}{2}$.

When $\delta = \frac{\lambda_1}{2}$, it follows that $\beta_4 = \frac{-\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1}$. Plugging the value of β_4 into the second equation of (18), it follows that $\|\beta_3 \mathbf{v}_1 - \tau_3 \mathbf{v}_2\|^2 - r^2 = 0$, where $\tau_3 := \frac{\lambda_2 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1}$. Then

$$\beta_3^2 - 2\tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3 + \tau_3^2 - r^2 = 0. \tag{24}$$

Note that the discriminant of quadratic equation (24) is

$$\Delta := 4\left(\tau_3^2 (\mathbf{v}_1^\top \mathbf{v}_2)^2 - \tau_3^2 + r^2\right) = 4\left(\tau_3^2 ((\mathbf{v}_1^\top \mathbf{v}_2)^2 - 1) + r^2\right)$$

Case 3.1 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 < 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$, there doesn't exist a solution associated with $\delta = \frac{\lambda_1}{2}$.

In addition, we recall that the eigenvalue (other than $-\alpha'(0)$) of Jacobian at $\mathbf{x}_{*,-}$ is $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r}$. In this case, we have $1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} > (\mathbf{v}_1^\top \mathbf{v}_2)^2 \geq 0$, which implies that $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} > 0$. Hence in **Case 3.1**, we only have one undesirable equilibrium, which is a saddle point.

Case 3.2 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$ there exists one solution associated with $\delta = \frac{\lambda_1}{2}$, which is

$$\hat{\mathbf{x}} = \left(\frac{\lambda_2 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_1^\top \mathbf{v}_2 \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2, \frac{\lambda_1}{2} \right).$$

We note that $(\hat{\mathbf{x}} - \mathbf{x}_c)^\top \mathbf{v}_1 = 0$ and $(\tilde{A} - 2\delta_{\hat{\mathbf{x}}} \mathbf{I})\mathbf{v}_2 = 0$. Hence $J \mid_{\hat{\mathbf{x}}} \mathbf{v}_1 = 0$, i.e. $\hat{\mathbf{x}}$ is an undersirable equilibrium at which the Jacobian has a negative eigenvalue and a zero eigenvalue.

In addition, we recall that the eigenvalue (other than $-\alpha'(0)$) of Jacobian at $\mathbf{x}_{*,-}$ is $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r}$.

In this case, we still have $1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} = (\mathbf{v}_1^\top \mathbf{v}_2)^2 > 0$, which implies that $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} > 0$. Hence in this case, $\mathbf{x}_{*,-}$ is an saddle point.

Thus in **Case 3.2**, besides $\mathbf{x}_{*,-}$, there is another one undesirable equilibrium, at which the Jacobian has a negative eigenvalue and a zero eigenvalue.

Case 3.3 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$, $\mathbf{v}_1^\top \mathbf{v}_2 = 0$, there exists one solution associated with $\delta = \frac{\lambda_1}{2}$, which is

$$\hat{\mathbf{x}} = \left(-\frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2, \frac{\lambda_1}{2} \right).$$

Notice that $1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} = 0$ implies that $\frac{-\lambda_2}{\lambda_2 - \lambda_1} = \frac{r}{\|\mathbf{x}_c\|}$, which follows that

$$-\frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2 = (1 - \frac{\lambda_2}{\lambda_2 - \lambda_1}) \mathbf{x}_c = (1 + \frac{r}{\|\mathbf{x}_c\|}) \mathbf{x}_c = \mathbf{x}_{*,-},$$

and $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} = 0$.

Thus in **Case 3.3**, there is only one undesirable equilibrium, at which the Jacobian has a zero eigenvalue.

Case 3.4 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1-\lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$, there exists two solutions $\beta_3 = \beta_3^{(1)}$ and $\beta_3 = \beta_3^{(2)}$ for (23). Then there exists two extra solutions for (18): $\left(\beta_3^{(1)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2, \frac{\lambda_1}{2} \right)$ and $\left(\beta_3^{(2)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2, \frac{\lambda_1}{2} \right)$ and both of them are undesirable equilibrium. Notice that in this sub-case, we have $\beta_3^{(1)} + \beta_3^{(2)} = 2\tau_3 \mathbf{v}_1^\top \mathbf{v}_2 < 0$. Then we can assume that $\beta_3^{(1)} < \tau_3 \mathbf{v}_1^\top \mathbf{v}_2$ and $\beta_3^{(2)} > \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$.

Using the same technique in the proof of Lemma 5, we can show that the Jacobian evaluated at $\left(\beta_3^{(1)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2 \right)$ has an eigenvalue

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(1)} - r^2);$$

and the Jacobian evaluated at $\left(\beta_3^{(2)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2 \right)$ has an eigenvalue

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(2)} - r^2).$$

Recall that the Jacobian evaluated at $\mathbf{x}_{*,-}$ has an eigenvalue $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r}$, and then we only need to determine the sign of these three eigenvalues case by case.

Case 3.4.1 If $0 < \tau_3^2 - r^2 = \frac{\lambda_2^2\|\mathbf{x}_c\|^2}{(\lambda_2-\lambda_1)^2} - r^2$, it is easy to check that $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r} > 0$. In addition, similar to **Case 3.3**, we can show that $\{\frac{\lambda_1-\lambda_2}{r^2}(\tau_3^2 - \tau_3\mathbf{v}_1^\top\mathbf{v}_2\beta_3^{(i)} - r^2) : i = 1, 2\}$ contains one positive number and one negative number.

Thus **Case 3.4.1**, there are three undesirable equilibria in total, two of which are saddle point and one of which is asymptotically stable.

Case 3.4.2 If $0 = \tau_3^2 - r^2 = \frac{\lambda_2^2\|\mathbf{x}_c\|^2}{(\lambda_2-\lambda_1)^2} - r^2$, it is easy to check that $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r} = 0$. In addition, we have $\beta_3^{(2)} = 0$ and the point $\left(\beta_3^{(2)}\mathbf{v}_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{\lambda_2-\lambda_1}\mathbf{v}_2\right)$ is equal to $\mathbf{x}_{*,-}$.

The point $\left(\beta_3^{(1)}\mathbf{v}_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{\lambda_2-\lambda_1}\mathbf{v}_2\right)$ is a saddle point since the eigenvalue

$$\frac{\lambda_1 - \lambda_2}{r^2}(\tau_3^2 - \tau_3\mathbf{v}_1^\top\mathbf{v}_2\beta_3^{(1)} - r^2) = 2\frac{\lambda_2 - \lambda_1}{r^2}\tau_3^2(\mathbf{v}_1^\top\mathbf{v}_2)^2$$
$$>2\frac{\lambda_2 - \lambda_1}{r^2}\tau_3^2(1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2\|\mathbf{x}_c\|^2}) > 0.$$

Thus in **Case 3.4.2**, there are two undesirable equilibria in total, one of which is a saddle point and the other one results in the Jacobian having a zero eigenvalue.

Case 3.4.3 If $0 > \tau_3^2 - r^2 = \frac{\lambda_2^2\|\mathbf{x}_c\|^2}{(\lambda_2-\lambda_1)^2} - r^2$, it is easy to check that $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r} < 0$, which implies that $\mathbf{x}_{*,-}$ is asymptotically stable.

By $\beta_3^{(1)}\beta_3^{(2)} = \tau_3^2 - r^2 < 0$ and $\beta_3^{(1)} < \tau_3\mathbf{v}_1^\top\mathbf{v}_2 < 0$, it follows that $\beta_3^{(2)} > 0$.

Using the fact that $\beta_3^{(1)} < \tau_3\mathbf{v}_1^\top\mathbf{v}_2 < 0$, we can show that

$$\frac{\lambda_1 - \lambda_2}{r^2}(\tau_3^2 - \tau_3\mathbf{v}_1^\top\mathbf{v}_2\beta_3^{(1)} - r^2) > 0.$$

Using the fact that $\beta_3^{(1)} > 0 > \tau_3\mathbf{v}_1^\top\mathbf{v}_2$, we can show that

$$\frac{\lambda_1 - \lambda_2}{r^2}(\tau_3^2 - \tau_3\mathbf{v}_1^\top\mathbf{v}_2\beta_3^{(2)} - r^2) > 0.$$

Thus in **Case 3.4.3**, there are three undesirable equilibria in total, two of which are saddle point and one of which is asymptotically stable.

We note that the above is the collection of all possible cases, by which we conclude Proposition 6. Next, we can see that there exists only one undesirable equilibrium and it is a degenerate equilibrium if and only if **Case 3.3** occurs, which corresponds to *Condition 1*. If **Case 3.4.2** (corresponding to *Condition 2*) occurs, there are two undesirable equilibria, one of which is a saddle point and the other one is a degenerate equilibrium. Next, **Case 1.1, 1.2** and **1.3** correspond to the first, second and third columns of Table I respectively. In addition, the first column of Table II includes **Case 2.1** and **3.1**; the second column of Table II includes **Case 2.2** and **3.2**; the third column of Table II includes **Case 2.3, 3.4.1** and **3.4.3**.

B. *Proof of Proposition 7*

If $\lambda_1 = \lambda_2$, we let $\tilde{A}_1$ and $\tilde{A}_2$ be any two matrices satisfying the conditions of the third row and first row of Table I, respectively. Then $K_1 = B^{-1}(A - \tilde{A}_1)$ and $K_2 = B^{-1}(A - \tilde{A}_2)$ satisfy our requirement.

If $\lambda_1 \neq \lambda_2$, we let $\tilde{A}_3$ and $\tilde{A}_4$ be any two matrices satisfying the conditions of the third row and first row of Table II, respectively. Then $K_1 = B^{-1}(A - \tilde{A}_3)$ and $K_2 = B^{-1}(A - \tilde{A}_4)$ satisfy our requirement.

C. *Proof of Proposition 8*

Recall that the equation for undesirable equilibria can be rewritten as (18), we consider two cases.

• *Case* #1: $\lambda_1 \neq \lambda_2$. Let $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathbb{C}^n$ be eigenvectors such that $\tilde{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $\tilde{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$. Write $\mathbf{x}_c$ as $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$ and $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$. Hence, the first equation in (18) can be rewritten as:

$$\begin{aligned}(\lambda_1 - 2\delta)(\beta_3 - \beta_1) &= -\lambda_1\beta_1\\ (\lambda_2 - 2\delta)(\beta_4 - \beta_2) &= -\lambda_2\beta_2\,.\end{aligned} \tag{25}$$

Note that $\beta_1 \neq 0$ and $\beta_2 \neq 0$ as $\mathbf{x}_c$ is not an eigenvector of $A - BK$; it follows that there is no solution with $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. For any solution $(\mathbf{x}, \delta_{\mathbf{x}})$ of (18), we have that $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1-2\delta_{\mathbf{x}}}$, $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2-2\delta_{\mathbf{x}}}$ and $\left\|\frac{-\lambda_1\beta_1}{\lambda_1-2\delta_{\mathbf{x}}}\mathbf{v}_1 + \frac{-\lambda_2\beta_2}{\lambda_2-2\delta_{\mathbf{x}}}\mathbf{v}_2\right\|^2 - r^2 = 0$, which is equivalent to $F_1(\delta) = 0$, where $F_1(\delta)$ is defined in (17).

We first note that $F_1(\delta) = 0$ can have at most 4 solutions. Therefore, there are four solutions at most for (18). In addition, notice that $F_1(-\infty) < 0$, $F_1(+\infty) < 0$ and $F_1(0) = (\|\mathbf{x}_c\|^2 - r^2)\|\lambda_1\lambda_2\|^2 > 0$, it follows that there exists at least solution for (18) with positive δ and at least one solution with negative δ.

If $\lambda_1 \leq \lambda_2$, we have $F_1(-\infty) < 0$, and $F_1(\frac{\lambda_1}{2}) > 0$; there exists at least one solution for (18) with $\delta < \frac{\lambda_1}{2}$.

• *Case* #2: $\lambda_1 = \lambda_2$. Note that both eigenvalues are negative and $\mathbf{x}_c$ is not an eigenvector; we let $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathbb{R}^n$ be vectors of length 1, such that $\tilde{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $\tilde{A}\mathbf{v}_2 = \lambda_1\mathbf{v}_2 + \mathbf{v}_1$. We write $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$ and $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$.

Hence, the first equation in (18) can be rewritten as

$$\begin{aligned}(\lambda_1 - 2\delta)(\beta_3 - \beta_1) + (\beta_4 - \beta_2) &= -\lambda_1\beta_1 - \beta_2\\ (\lambda_2 - 2\delta)(\beta_4 - \beta_2) &= -\lambda_2\beta_2\,.\end{aligned} \tag{26}$$

Note that $\beta_2 \neq 0$ as $\mathbf{x}_c$ is not an eigenvector of $A - BK$; it follows that there is no solution with $2\delta = \lambda_1$. For any solution $(\mathbf{x}, t_{\mathbf{x}})$ of equation (18), we have $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1-2\delta_{\mathbf{x}}} + \frac{2\delta_{\mathbf{x}}\beta_2}{(\lambda_1-2\delta_{\mathbf{x}})^2}$, $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2-2\delta_{\mathbf{x}}}$ and $\left\|\left(\frac{-\lambda_1\beta_1}{\lambda_1-2\delta_{\mathbf{x}}} + \frac{2\delta_{\mathbf{x}}\beta_2}{(\lambda_1-2\delta_{\mathbf{x}})^2}\right)\mathbf{v}_1 + \frac{-\lambda_2\beta_2}{\lambda_2-2\delta_{\mathbf{x}}}\mathbf{v}_2\right\|^2 - r^2 = 0$, which is equivalent to $F_2(\delta) = 0$, where $F_2(\delta)$ is defined as

$$\begin{aligned}F_2(\delta) :=& -(\lambda_1 - 2\delta)^4 r^2 + (\lambda_1\beta_2)^2(\lambda_1 - 2\delta)^2\\ &+ 2(\lambda_1\beta_1(\lambda_1 - 2\delta)^2 - 2\delta(\lambda_1 - 2\delta)\beta_2)\lambda_1\beta_2\mathbf{v}_1^\top\mathbf{v}_2\\ &+ (2\delta\beta_2 - \lambda_1(\lambda_1 - 2\delta)\beta_1)^2.\end{aligned}$$

We first note that $F_2(\delta) = 0$ can have at most 4 solutions. Therefore, there are four solutions at most for (18). In addition, notice that $F_2(+\infty) < 0$ and $F_2(0) = (\|\mathbf{x}_c\|^2 - r^2)\lambda_1^4 >$

0; it follows that there exists at least a solution for (18) with positive δ. Similarly, we have that $\frac{1}{(\lambda_1-2\delta)^4}F_2(\delta) < 0$ as $\delta \to -\infty$, and $\frac{1}{(\lambda_1-2\delta)^4}F_2(\delta) \to +\infty$ as $\delta \to \frac{\lambda_1^-}{2}$; then. there exists at least one solution for (18) with negative $\delta < \frac{\lambda_1}{2}$.

To conclude, there always exists at least one solution with negative δ and at least one solution with positive δ for (18). In addition, (18) can have four solutions at most. If $\lambda_1 \le \lambda_2$, there exists a solution for (18) with indicator $\delta < \frac{\lambda_1}{2}$.

D. Proof of Proposition 9

Let $\mathbf{x} \in \hat{\mathcal{E}}$ with indicator $\delta_{\mathbf{x}} < \frac{\lambda_1}{2}$, and write $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$; then, it follows by Lemma 5 that the Jacobian evaluated at $\mathbf{x}$ has an eigenvalue greater than $\frac{(\lambda_2-\lambda_1)}{r^2}\frac{-\lambda_1}{\lambda_1-2\delta_{\mathbf{x}}}((\beta_3 - \beta_1)\beta_1 + (\beta_4 - \beta_2)\beta_1\mathbf{v}_2^\top\mathbf{v}_1)$. Notice that $\frac{(\lambda_2-\lambda_1)}{r^2}\frac{-\lambda_1}{\lambda_1-2\delta_{\mathbf{x}}} > 0$ and

$$\begin{aligned}(\beta_3 - \beta_1)\beta_1 &+ (\beta_4 - \beta_2)\beta_1\mathbf{v}_2^\top\mathbf{v}_1\\ &= \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}}}\beta_1^2 + \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}}}\beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1\\ &\ge \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}}}(\beta_1^2 + \beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1) \ge 0\,.\end{aligned}$$

Hence, the Jacobian evaluated at x has an positive eigenvalue and, thus, x is a saddle point.

For any $\mathbf{x} \in \hat{\mathcal{E}}$ with indicator $\frac{\lambda_2}{2} < \delta_{\mathbf{x}} < 0$, write $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$; then, by Lemma 5, it follows that the Jacobian evaluated at $\mathbf{x}$ has an eigenvalue less than $\frac{(\lambda_2-\lambda_1)}{r^2}\frac{2\lambda_2}{\lambda_2-2\delta_x}((\beta_3 - \beta_1)\beta_2\mathbf{v}_1^\top\mathbf{v}_2 + (\beta_4 - \beta_2)\beta_2)$. Notice that $\frac{(\lambda_2-\lambda_1)}{r^2}\frac{2\lambda_2}{\lambda_1-2\delta_{\mathbf{x}}} > 0$ and

$$\begin{aligned}(\beta_3 - \beta_1)\beta_2\mathbf{v}_2^\top\mathbf{v}_1 &+ (\beta_4 - \beta_2)\beta_2\\ &= \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}}}\beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1 + \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}}}\beta_2^2\\ &\le \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}}}(\beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1 + \beta_2^2) \le 0.\end{aligned}$$

Besides, by Lemma 2, $-\alpha_0$ is another eigenvalue. Hence, all the eigenvalues of the Jacobian evaluated at $\mathbf{x}$ are negative, which means that $\mathbf{x}$ is an undesirable asymptotically stable equilibrium.

To prove the last claim, let δ_0 denote the only real root of the third-order polynomial $\frac{dF_1(\delta)}{d\delta}$. It follows that $F_1(\delta)$ is monotonically increasing on $(-\infty, \delta_0)$ and monotonically decreasing on $(\delta_0, +\infty)$; this implies that $F_1(\delta) = 0$ only has two solutions. By Lemma 5, there is only one undesirable equilibrium and its indicator satisfies $\delta < \frac{\lambda_1}{2}$. Since $\beta_1^2 + \beta_1\beta_2\mathbf{v}_1^\top\mathbf{v}_2 \ge 0$, there is only one undesirable equilibrium and it is a saddle point.