Optimization-Based Safe Stabilizing Feedback with Guaranteed Region of Attraction

Pol Mestres   Jorge Cortés

Abstract— **This paper proposes an optimization with penalty-based feedback design framework for safe stabilization of control affine systems. Our starting point is the availability of a control Lyapunov function (CLF) and a control barrier function (CBF) defining affine-in-the-input inequalities that certify, respectively, the stability and safety objectives for the dynamics. Leveraging ideas from penalty methods for constrained optimization, the proposed design framework imposes one of the inequalities as a hard constraint and the other one as a soft constraint. We study the properties of the closed-loop system under the resulting feedback controller and identify conditions on the penalty parameter to eliminate undesired equilibria that might arise. Going beyond the local stability guarantees available in the literature, we are able to provide an inner approximation of the region of attraction of the equilibrium, and identify conditions under which the whole safe set belongs to it. Simulations illustrate our results.**

I. Introduction

Safety-critical control has garnered a lot of attention in the controls and robotics communities motivated by applications to many different classes of engineered and natural systems. Safety refers to the ability to ensure by design that the evolution of the dynamics stays within a desired set. Control barrier functions (CBFs) are a useful tool to deal with safety specifications that do not require addressing the difficult task of computing the system's reachable set. In many scenarios, safety must be achieved together with some stabilization goal, and this raises interesting challenges for control design in order to ensure that both are achieved via feedback controllers that are easily implemented and have appropriate smoothness guarantees. These challenges motivate us to develop here an optimization with penalty-based feedback design framework for safe stabilization of control affine systems.

Literature Review: We rely on ideas from two different bodies of work. The first one is CLFs [1], which have been successfully used in the control design for stabilization of nonlinear systems. Of particular interest to this work is the pointwise-minimum norm (PMN) formula [2], that uses a CLF to compute a stabilizing controller. The second relevant body of work pertains to CBFs [3], [4], whose aim is to render a certain predefined safe set forward invariant. However, in applications where both safety and stability must be certified, CBFs fall short of providing provable stability guarantees. To tackle this issue, [5] combines a CLF and a CBF into a so-called CLBF, and then uses Sontag's universal formula to derive a smooth controller. However, in general it might be difficult to satisfy the conditions required for the existence of such a CLBF [6]. Another approach is the universal formula for smooth safe stabilization from [7]. However, this formula is only applicable in a set where both the CLF and the CBF are compatible (i.e., there exists a control satisfying their associated inequalities at every point of the set). An alternative approach [4] to tackle joint safety and stability specifications is to combine the CLF and the CBF in a quadratic program (QP). To guarantee the feasibility of the program when the functions are not compatible and to avoid the resulting controller to be non-Lipschitz when they are [8], the stability constraint is often relaxed. This results in a lack of guarantee of stability, even for arbitrarily large penalties in the relaxation parameter [9]. Moreover, as shown in [10], [11], this QP formulation can introduce undesired equilibria beyond the original equilibrium, which can even be asymptotically stable. This line of work [9], [11] then identifies conditions under which local stability guarantees of the equilibrium can be given. Although the region of attraction is not explicitly characterized, a strategy similar to the one pursed here could be employed. An alternative design, e.g., [12], assumes a priori knowledge of a CBF and a nominal (possibly unsafe) stabilizing controller. Then, a safety filter is applied to this nominal controller. As a result, the filtered controller generally lacks stability guarantees. The recent paper [13] gives an estimate of the region of attraction of the closed-loop system obtained by using such a filtered controller.

Statement of Contributions: We consider the problem of safe stabilization of control affine systems. Given a control Lyapunov function and a control barrier function whose 0-superlevel set defines an arbitrary, possibly non-convex safe set, we aim to synthesize a safe, stabilizing feedback and identify the region of attraction of the origin for the resulting closed-loop system. In particular, we study under what conditions such region of attraction contains the safe set. The contributions of this paper are the following. Given the safety and stability objectives, our first contribution designs an optimization with penalty-based controller that has one of the objectives as a hard constraint and the other as a soft constraint. The controller depends on a penalty parameter that can be tuned to enhance the soft objective at the cost of reduced optimality, while guaranteeing the satisfaction of the hard constraint. An advantage of the proposed design is that the controller is automatically Lipschitz and has a closed-form expression. Our second contribution shows that the controller can introduce undesired equilibrium points

This work was supported by NSF Award IIS-2007141.

P. Mestres and J. Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {pomestre,cortes}@ucsd.edu

different from the origin. By choosing the penalty parameter appropriately, and under some technical conditions, these undesired equilibria can be eliminated. Finally, our third contribution shows that the proposed controller can be tuned to provide an inner approximation of the region of attraction of the origin for the closed-loop system. As a consequence of this analysis, we provide conditions under which all of the safe set belongs to the region of attraction of the origin for the closed-loop system. Simulations on a planar system compare our design with other approaches in the literature.

II. Preliminaries on CLFs and CBFs

This section presents[1] preliminaries on control Lyapunov and barrier functions. Consider a control-affine system

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

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are locally Lipschitz functions, with $x \in \mathbb{R}^n$ the state and $u \in \mathbb{R}^m$ the input. Throughout the paper, and without loss of generality, we assume $f(0) = 0$, so that the origin $x = 0$ is the desired equilibrium point of the (unforced) system.

We start by recalling the notion of Control Lyapunov function (CLF) [1], [2].

Definition 1: (Control Lyapunov Function): Given an open set $\mathcal{D} \subseteq \mathbb{R}^n$, with $0 \in \mathcal{D}$, a continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ is a **CLF** on $\mathcal{D}$ for system (1) if

- V is proper in $\mathcal{D}$, i.e., $\{x \in \mathcal{D} \ : \ V(x) \leq c\}$ is a compact set for all $c > 0$,
- V is positive definite,
- there exists a continuous positive function $W : \mathbb{R}^n \to \mathbb{R}$ such that, for each $x \in \mathcal{D}\backslash\{0\}$, there exists a control $u \in \mathbb{R}^m$ satisfying

$$L_fV(x) + L_gV(x)u \leq -W(x). \tag{2}$$

CLFs provide a way to guarantee asymptotic stability of the origin. Namely, if a Lipschitz controller u satisfies (2) for all $x \in \mathcal{D}\backslash\{0\}$, then the origin of the closed-loop system is asymptotically stable [1]. If $W(x)$ in (2) is replaced by $\gamma(V(x))$, where γ is a class $\mathcal{K}$ function, then such Lipschitz controller makes the origin exponentially stable. Such controllers can be synthesized by means of the pointwise minimum-norm (PMN) control optimization [2, Chapter 4.2],

$$u(x) = \arg\min_{u\in\mathbb{R}^m} \frac{1}{2}\|u\|^2 \quad \text{s.t. (2) holds.}$$

Note that, at each $x \in \mathbb{R}^n$, this is a quadratic program in u.

Next we recall the notion of Control Barrier Function (CBF) [4]. Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a closed set describing the safe states for the system (1).

Definition 2: (Control Barrier Function): Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function such that $\mathcal{C} = \{x \in \mathbb{R}^n \ : \ h(x) \geq 0\}$. The function h is a **CBF** of $\mathcal{C}$ for system (1) if there exists a class $\mathcal{K}_\infty$ function α such that, for all $x \in \mathcal{C}$, there exists a control $u \in \mathbb{R}^m$ satisfying

$$L_fh(x) + L_gh(x)u + \alpha(h(x)) \geq 0. \tag{3}$$

CBFs can be used to guarantee safety, i.e., forward invariance of $\mathcal{C}$ under the dynamics (1). Namely, if a Lipschitz continuous controller satisfies (3) for all $x \in \mathcal{C}$, then $\mathcal{C}$ is forward invariant [4, Theorem 2]. Similar to the PMN controller above, a common design methodology [4] is via the optimization

$$u(x) = \arg\min_{u\in\mathbb{R}^m} \frac{1}{2}\|u\|^2 \quad \text{s.t. (3) holds,}$$

which results in a Lipschitz controller [14, Theorem 2].

When dealing with both the stability and safety of system trajectories under the dynamics (1), it is important to note that an input u might satisfy (2) but not (3), or vice versa. The following notion, adapted from [7, Definition 2.3], captures when the CLF and the CBF are compatible.

Definition 3: (Compatibility of CLF-CBF pair): Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $\mathcal{C} \subset \mathcal{D}$ closed, V a CLF on $\mathcal{D}$ and h a CBF of $\mathcal{C}$. Then, V and h are **compatible at** $x \in \mathcal{C}$ if there exists $u \in \mathbb{R}^m$ satisfying (2) and (3) simultaneously. We refer to both functions as compatible if V and h are compatible at every point of $\mathcal{C}$.

III. Problem Statement

We are interested in designing controllers that are both stabilizing and safe. We also require them to be Lipschitz in order to guarantee existence and uniqueness of solutions of the closed-loop system. Formally, consider a control-affine system of the form (1). Let $V : \mathbb{R}^n \to \mathbb{R}$ be a CLF on the open set $\mathcal{D} \subseteq \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$ be a CBF of the closed set $\mathcal{C} \subset \mathcal{D}$. We assume the origin belongs to $\mathcal{C}$. Given the availability of these functions, it seems reasonable to employ V to ensure the stabilizing aspect of the controller and h to ensure safety. We also seek to provide formal characterizations of the region of attraction of the equilibrium for the resulting closed-loop system. If V and h are compatible at every point in the safe set, one option is to find the control through pointwise optimization with (2) and (3) as constraints. However, [8] gives a counterexample that shows that this pointwise minimization can result in a non-Lipschitz

[1] We denote by $\mathbb{Z}_{>0}, \mathbb{R}$ and $\mathbb{R}_{\geq 0}$ the set of positive integers, real, and nonnegative real numbers, resp. We write $\text{int}(\mathcal{S}), \partial\mathcal{S}$ for the interior and the boundary of the set $\mathcal{S}$, resp. Given $x \in \mathbb{R}^n$, $\|x\|$ denotes its Euclidean norm. Given $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ and a smooth function $W : \mathbb{R}^n \to \mathbb{R}$, the notation $L_fW : \mathbb{R}^n \to \mathbb{R}$ (resp. $L_gW : \mathbb{R}^n \to \mathbb{R}^m$) denotes the Lie derivative of W with respect to f (resp. g), that is $L_fW = \nabla W^T f$ (resp. $\nabla W^T g$). We denote by $\mathcal{C}^1(\mathbb{R}^n)$ and $\mathcal{C}^2(\mathbb{R}^n)$ the set of continuously differentiable and twice continuously differentiable functions in $\mathbb{R}^n$, respectively. Given $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, let H denote the hyperplane defined by $H = \{x \in \mathbb{R}^n \ : \ \langle a, x\rangle = b\}$. We denote the projection of $v \in \mathbb{R}^n$ onto H by $P_H(v) = v - \frac{\langle a,v\rangle - b}{\|a\|^2}a$. A function $\beta : \mathbb{R} \to \mathbb{R}$ is of class $\mathcal{K}$ if $\beta(0) = 0$ and β is strictly increasing. If moreover $\lim_{t\to\infty} \beta(t) = \infty$, β is of class $\mathcal{K}_\infty$. A function $V : \mathbb{R}^n \to \mathbb{R}$ is positive definite if $V(0) = 0$ and $V(x) > 0$ for $x \neq 0$. Given a matrix $M \in \mathbb{R}^{n\times m}$, $\ker(M) = \{x \in \mathbb{R}^m \ : \ Mx = \mathbf{0}_n\}$. Given a square matrix $A \in \mathbb{R}^{n\times n}$ with eigenvectors $\{v_j\}_{j=1}^n$ and corresponding eigenvalues $\{\lambda_j\}_{j=1}^n$, the stable subspace of A is defined as $\mathcal{V}_s(A) = \text{span}(\{v_j \ : \ \Re(\lambda_j) < 0, \ j = 1, \dots, n\})$, where $\Re(\lambda_j)$ denotes the real part of λ_j. We denote by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ the smallest non-zero and largest real parts of the eigenvalues of A, respectively.

controller. To remedy this, and also to extend the design to scenarios where V and h might not be compatible at some points in the safe set, a popular approach [4] is to relax one of the inequalities (2), (3) (in safety-critical applications, the CLF constraint (2)), and formulate a QP that penalizes the relaxation parameter:

$$u(x) = \arg \min_{(u,\delta)\in\mathbb{R}^{m+1}} \frac{1}{2}\|u\|^2 + p\delta^2, \quad (4)$$
$$\text{s.t.} \quad (3),\ L_f V(x) + L_g V(x)u \leq -W(x) + \delta.$$

Nevertheless, even in the case where the CLF and the CBF are compatible at all points in the safe set, the resulting controller might not be stabilizing even for arbitrarily large values of p [9]. Moreover, as pointed out in [10], [11], this design might introduce undesired equilibria in the closed-loop system, which can even be asymptotically stable. To the best of the authors' knowledge, only local stability guarantees exist [11, Theorem 3], [9, Theorem 1], and no estimates of the region of attraction are available in the literature.

An alternative design, e.g., [12], assumes a nominal (possibly unsafe) stabilizing controller u_{nom} is available, and seeks to modify it as little as possible while guaranteeing safety. This can be done by solving the following QP:

$$u(x) = \arg \min_{u\in\mathbb{R}^m} \frac{1}{2}\|u - u_{\text{nom}}(x)\|^2, \quad (5)$$
$$\text{s.t.} \quad (3).$$

In general, the resulting modified controller might not retain the stability properties of the original nominal controller but, under certain conditions [13], one can provide an estimate of the region of attraction of the equilibrium. Interestingly, nominal controllers other than the given one might result in larger regions of attraction, so in this sense the design directly with the CLF offers greater flexibility.

We are interested in building an alternative to the designs (4), (5) to solve the aforementioned issues. In particular, we tackle the following problem:

Problem 1: Determine a Lipschitz control law u and a region of attraction $\Gamma \subseteq \mathbb{R}^n$, $\Gamma \cap \mathcal{C} \neq 0$ such that for all $x(0) \in \Gamma \cap \mathcal{C}$, $x(t) \in \mathcal{C}$ for all $t \geq 0$ and the system (1) in closed-loop with u is asymptotically stable with respect to the origin. △

IV. Safety and Stability via QP with Penalty

In this section we design a candidate control law to solve Problem 1 by leveraging the CLF V and the CBF h. We first present our exposition in a general context, then particularize to our setting. Consider general Lipschitz functions $a, c : \mathbb{R}^n \to \mathbb{R}$ and $b, d : \mathbb{R}^n \to \mathbb{R}^m$. Consider the following two affine inequalities in $u \in \mathbb{R}^m$,

$$a(x) + b(x)u \leq 0, \quad c(x) + d(x)u \leq 0.$$

Given a neighborhood $\bar{\mathcal{C}}$ of $\mathcal{C}$, we assume that for every $x \in \bar{\mathcal{C}}$, there exist $u_1, u_2 \in \mathbb{R}^m$ such that $a(x) + b(x)u_1 \leq 0$ and $c(x) + d(x)u_2 \leq 0$. To select u, we regard at the first inequality as a *soft constraint* and the second as a *hard constraint*. Inspired by the theory of penalty methods for constrained optimization [15, Chapter 13], we formulate a QP where we include the soft constraint in the objective function with a penalty parameter ($\epsilon > 0$) and enforce the hard constraint. The resulting solution of the QP is parametrized by $x \in \mathbb{R}^n$ and ϵ:

$$u_\epsilon(x) := \arg \min_{u\in\mathbb{R}^m} \frac{1}{2}\|u\|^2 + \frac{1}{\epsilon}(a(x) + b(x)u),$$
$$\text{s.t.} \quad c(x) + d(x)u \leq 0. \quad (6)$$

Since this optimization problem is a QP, it is convex. The following result gives a closed-form expression for u_ϵ and establishes that it is Lipschitz.

Proposition 4.1: (Closed-form expression for Lipschitz controller): Let $a, c : \mathbb{R}^n \to \mathbb{R}$ and $b, d : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $\bar{\mathcal{C}}$ a neighborhood of $\mathcal{C}$ and assume that for every $x \in \bar{\mathcal{C}}$, there exist $u_1, u_2 \in \mathbb{R}^m$ such that $a(x) + b(x)u_1 \leq 0$ and $c(x) + d(x)u_2 \leq 0$. For each $x \in \mathcal{C}$, let $H(x) := \{u \in \mathbb{R}^m : c(x) + d(x)u = 0\}$ and $e(x) := c(x) - \frac{1}{\epsilon}d(x)b(x)$. Then,

$$u_\epsilon(x) = \begin{cases} -\frac{1}{\epsilon}b(x) & \text{if } e(x) \leq 0, \\ P_{H(x)}(-\frac{1}{\epsilon}b(x)) & \text{if } e(x) > 0, \end{cases} \quad (7)$$

and u_ϵ is Lipschitz on $\bar{\mathcal{C}}\backslash\{0\}$. Moreover, if $d(0) \neq 0$, u_ϵ is Lipschitz at 0.

Proof: The expression (7) follows by calculating the KKT points of (6). Note that (7) is well defined because if $d(x) = 0$, necessarily $e(x) = c(x) \leq 0$. Lipschitzness of $u_\epsilon(x)$ follows from [16, Section 3.10, Theorem 2], which as a special case includes the minimization of a quadratic cost function subject to affine inequality constraints. ■

We next particularize the general design (6) to our setup. We consider two cases:

Safety QP with stability penalty: The selection $a(x) = L_f V(x) + W(x)$, $b(x) = L_g V(x)$, $c(x) = -L_f h(x) - \alpha(h(x))$, and $d(x) = -L_g h(x)$ makes the CLF inequality (2) a soft constraint and the CBF inequality (3) a hard one. We denote by u_ϵ^{safe} the controller resulting from (6). If $L_g h(0) \neq 0$, Proposition 4.1 guarantees that u_ϵ^{safe} is Lipschitz on $\mathcal{C}$. Moreover, since it satisfies the CBF inequality (3) for all $x \in \mathcal{C}$, the resulting closed-loop system is safe for all $\epsilon > 0$;

Stability QP with safety penalty: Alternatively, the selection $a(x) = -L_f h(x) - \alpha(h(x))$, $b(x) = -L_g h(x)$, $c(x) = L_f V(x) + W(x)$, and $d(x) = L_g V(x)$, makes the CBF inequality (3) a soft constraint and the CLF inequality (2) a hard one. We denote by $u_\epsilon^{\text{stable}}$ the resulting controller from (6). In this case, $d(0) = 0$ and hence Proposition 4.1 only guarantees that $u_\epsilon^{\text{stable}}$ is Lipschitz in $\bar{\mathcal{C}}\backslash\{0\}$. Moreover, since (2) is satisfied for all $x \in \bar{\mathcal{C}}\backslash\{0\}$, the origin is asymptotically stable for the resulting closed-loop system.

From this point onwards, we formulate the results for the controller u_ϵ^{safe}. With minor modifications, similar results can be stated for $u_\epsilon^{\text{stable}}$. Note also that Proposition 4.1 provides a closed-form expression for the controllers. This allows the closed-loop system to be implemented without having to continuously solve the optimization (6), which is something one faces with (4), e.g., [4]. The expression (7) indicates that smaller ϵ lead to controllers with larger norms. Even though here the input is unconstrained, this should be taken into account in applications with limited actuation power.

Remark 1: (Nominal Controller): Our framework can be adapted to the scenario described in (5), where instead of a CLF, one has access to a nominal stabilizing controller $u_{\rm nom}$ and a certificate of stability in the form of a Lyapunov function V satisfying $L_fV(x)+L_gV(x)u_{\rm nom}(x)+W(x)\leq 0$ for $x\in\mathcal{D}$, with $\mathcal{D}$ some open set. To design a control u as close as possible to $u_{\rm nom}$ that is safe and stabilizing, one can set $v=u-u_{\rm nom}$. Then, it is easy to check that V is a CLF for $\dot{x}=\bar{f}(x)+g(x)v$, where $\bar{f}(x)=f(x)+g(x)u_{\rm nom}(x)$. In this case, one could use the safety QP with stability penalty setting $a(x)=L_{\bar{f}}V(x)+W(x)$, $b(x)=L_gV(x)$, $c(x)=-L_{\bar{f}}h(x)-\alpha(h(x))$, and $d(x)=-L_gh(x)$. •

V. Analysis of Safety QP with Stability Penalty

Here, we analyze the closed-loop properties of (1) under $u_\epsilon^{\rm safe}$. We first show how to choose ϵ to avoid undesired equilibria of the closed-loop system and then go on to solve Problem 1. Throughout the section,

$$e(x)=-L_fh(x)+\frac{1}{\epsilon}L_gh(x)^TL_gV(x)-\alpha(h(x)).$$

A. Ruling out Undesired Equilibrium Points

Here we show that the closed-loop implementation of the safety QP with stability penalty controller might introduce new equilibria other than the origin. The next result characterizes such equilibria and shows that, under some conditions, they can be confined to an arbitrarily small neighborhood of the origin for small enough ϵ.

Proposition 5.1: (Characterization of Equilibria): For $\epsilon>0$, the set of equilibrium points of the closed-loop system $\dot{x}=f(x)+g(x)u_\epsilon^{\rm safe}(x)$ in $\mathcal{C}$ is $\mathcal{Q}=\mathcal{Q}_1^\epsilon\cup\mathcal{Q}_2^\epsilon$, with

$$\mathcal{Q}_1^\epsilon:=\{x\in\mathcal{C}\;:\;e(x)\leq 0,\;f(x)=\frac{1}{\epsilon}g(x)L_gV(x)\},$$

$$\mathcal{Q}_2^\epsilon:=\{x\in\partial\mathcal{C}\;:\;e(x)>0,\;f(x)=\frac{L_fh(x)}{\|L_gh(x)\|^2}g(x)L_gh(x)$$

$$+\frac{g(x)}{\epsilon}(L_gV(x)-\frac{L_gh(x)^TL_gV(x)}{\|L_gh(x)\|^2}L_gh(x))\},$$

and $0\in\mathcal{Q}_1^\epsilon$. Let $\mathcal{V}$ be a neighborhood of the origin, $\bar{\mathcal{V}}$ a neighborhood of $P_g:=\{x\in\mathcal{C}\backslash\{0\}\;:\;L_gV(x)=0\}$ and let N_1,N_2, $N_3^{\mathcal{V},\bar{\mathcal{V}}}$ and N_4 be defined by

$$N_1:=\sup_{x\in\mathcal{C}}\|f(x)\|,$$

$$N_2:=\sup_{\substack{x\in\partial\mathcal{C}\\ e(x)>0}}\left\|f(x)-\frac{L_fh(x)}{\|L_gh(x)\|^2}g(x)L_gh(x)\right\|,$$

$$N_3^{\mathcal{V},\bar{\mathcal{V}}}:=\inf_{x\in\mathcal{C}\backslash(\mathcal{V}\cup\bar{\mathcal{V}})}\|g(x)L_gV(x)\|.$$

$$N_4:=\inf_{\substack{x\in\partial\mathcal{C}\\ e(x)>0}}\left\|g(x)(L_gV(x)-\frac{L_gh(x)^TL_gV(x)}{\|L_gh(x)\|^2}L_gh(x))\right\|.$$

then,

- if N_1 is finite, then $\mathcal{Q}_1^\epsilon\subseteq\mathcal{V}$ for all $0<\epsilon<\frac{N_3^{\mathcal{V},\bar{\mathcal{V}}}}{N_1}$,
- if N_2 is finite and N_4 is positive, then $\mathcal{Q}_2^\epsilon=\emptyset$ for $0<\epsilon<\frac{N_4}{N_2}$.

Proof: Since $u_\epsilon^{\rm safe}(x)$ takes a different form depending on the sign of $e(x)$, we distinguish two cases:

Case 1: $e(x)\leq 0$: In this case, the equilibrium points of the closed-loop system satisfy $f(x)=\frac{1}{\epsilon}g(x)L_gV(x)$. Note that if $g(x)L_gV(x)=0$, by multiplying on the left by $\nabla V(x)^T$ we obtain $L_gV(x)=0$. Since V is a CLF, $L_fV(x)<0$ if $x\neq 0$. This implies that $f(x)\neq 0$ and hence x is not an equilibrium point. Hence, no point other than the origin satisfies $L_gV(x)=0$ and $f(x)=\frac{1}{\epsilon}g(x)L_gV(x)$, and we can choose a neighborhood $\bar{\mathcal{V}}$ of P_g with $\mathcal{Q}_1^\epsilon\cap\bar{\mathcal{V}}=\emptyset$. Now, by taking any neighborhood $\mathcal{V}$ of the origin, the choice $\epsilon<\frac{N_3^{\mathcal{V},\bar{\mathcal{V}}}}{N_1}$ rules out any equilibrium of this kind in $\mathcal{C}\backslash\mathcal{V}$. Note that, since $f(0)=0$ and $\nabla V(0)=0$, we have $e(0)=-\alpha(h(0))\leq 0$, and hence $0\in\mathcal{Q}_1^\epsilon$.

Case 2: $e(x)>0$: In this case the equilibrium points of the closed-loop system satisfy

$$f(x)-\frac{L_fh(x)+\alpha(h(x))}{\|L_gh(x)\|^2}g(x)L_gh(x)=$$
$$=\frac{g(x)}{\epsilon}(L_gV(x)-\frac{L_gh(x)^TL_gV(x)}{\|L_gh(x)\|^2}L_gh(x)).\quad(8)$$

Let us show that these equilibria can only occur in $\partial\mathcal{C}$. Multiplying both sides of (8) by $\nabla h(x)^T$, we obtain $-\alpha(h(x))=0$. Since α is a class $\mathcal{K}_\infty$ function, this can only occur when $h(x)=0$, i.e., $x\in\partial\mathcal{C}$. Now, by taking $\epsilon<\frac{N_4}{N_2}$, all equilibrium points of these kind are ruled out. ■

Note that the assumption that N_1 and N_2 are finite in Proposition 5.1 is satisfied if $\mathcal{C}$ is bounded. The neighborhood $\mathcal{V}$ of the origin in the statement can be taken arbitrarily small and, consequently, if N_4 is positive, the controller $u_\epsilon^{\rm safe}$ with sufficiently small ϵ confines the equilibria of the closed-loop system arbitrarily close to the origin. However, as $\mathcal{V}$ gets arbitrarily small, $N_3^{\mathcal{V}}$ (and hence ϵ) could also get arbitrarily small. In Corollary 5.4 later, we give sufficient conditions to ensure that this does not happen.

Remark 2: (Existence of boundary equilibria): The assumption that N_4 is positive is not satisfied if $g(x)L_gV(x)$ and $g(x)L_gh(x)$ are linearly dependent. In this scenario, using condition (8), we infer that the equilibrium points in $\partial\mathcal{C}$ that cannot be removed by tuning ϵ are those where $f(x)$, $g(x)L_gV(x)$ and $g(x)L_gh(x)$ are collinear and $e(x)>0$ for all ϵ. •

B. Incompatibility and Region of Attraction

Here we show that $u_\epsilon^{\rm safe}$ solves Problem 1. The flexibility provided by the design parameter ϵ is instrumental in doing so. We first introduce a characterization of points where the CLF and the CBF are incompatible, the proof of which follows as a special case of [17, Theorem 1].

Lemma 5.2: (Characterization of incompatible points): Let $\mathcal{D}\subseteq\mathbb{R}^n$ be open, $\mathcal{C}\subset\mathcal{D}$ closed, V a CLF on $\mathcal{D}$ and h a CBF of $\mathcal{C}$. V and h are incompatible at $x\in\mathcal{C}$ if and only if $L_gV(x)$ and $L_gh(x)$ are linearly dependent, $L_gV(x)^T L_gh(x) > 0$ and $L_fV(x) + W(x) > \frac{L_gV(x)^T L_gh(x)}{\|L_gh(x)\|^2}(L_fh(x) + \alpha(h(x)))$.

The next result shows that, by taking ϵ sufficiently small for the closed-loop system, any level set of V that does not contain incompatible points is a region of attraction of a neighborhood of the origin.

Theorem 5.3: (Parameter tuning for guaranteed region of attraction): Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $\mathcal{C} \subset \mathcal{D}$ closed, V a CLF on $\mathcal{D}$ and h a CBF of $\mathcal{C}$. Let $\nu > 0$ be such that the sublevel set $\Gamma_\nu = \{x \in \mathbb{R}^n \ : \ V(x) \leq \nu\}$ does not contain any incompatible points. For x such that $e(x) > 0$ (which implies $L_gh(x) \neq 0$ since h is a CBF), define

$$B(x) := L_fV(x) + W(x) - \frac{L_fh(x) + \alpha(h(x))}{\|L_gh(x)\|^2} L_gV(x)^T L_gh(x),$$

$$C(x) := \frac{(L_gV(x)^T L_gh(x))^2}{\|L_gh(x)\|^2} - \|L_gV(x)\|^2 .$$

Let $\mathcal{V}$ be a neighborhood of the origin, $\bar{\mathcal{V}}$ a neighborhood of $P_g := \{x \in \mathcal{C}\backslash\{0\} \ : \ L_gV(x) = 0\}$ such that $L_fV(x) + W(x) \leq 0$ for all $x \in \bar{\mathcal{V}}$ and $\mathcal{W}$ a neighborhood of $P_\nu = \{x \in \Gamma_\nu \ : \ e(x) > 0, C(x) = 0\}$ such that $e(x) > 0$ and $B(x) \leq 0$ for all $x \in \mathcal{W}\backslash\{0\}$. Define constants M_1^ν, M_2^ν, $M_3^{\nu,\mathcal{V},\bar{\mathcal{V}}}$ and $M_4^{\nu,\mathcal{V},\mathcal{W}}$ by

$$M_1^\nu := \sup_{x\in\Gamma_\nu} |L_fV(x) + W(x)|,$$

$$M_2^\nu := \sup_{\substack{x\in\Gamma_\nu \\ e(x)>0}} |\frac{L_fh(x) + \alpha(h(x))}{\|L_gh(x)\|^2} L_gh(x)^T L_gV(x)|,$$

$$M_3^{\nu,\mathcal{V},\bar{\mathcal{V}}} := \inf_{x\in\Gamma_\nu\backslash(\mathcal{V}\cup\bar{\mathcal{V}})} \|L_gV(x)\|^2 ,$$

$$M_4^{\nu,\mathcal{V},\mathcal{W}} := \inf_{\substack{x\in\Gamma_\nu\backslash(\mathcal{W}\cup\mathcal{V}) \\ e(x)>0}} |C(x)|.$$

Then, for $\epsilon < \bar{\epsilon} := \min\{\frac{M_4^{\nu,\mathcal{V},\mathcal{W}}}{M_1^\nu + M_2^\nu}, \frac{M_3^{\nu,\mathcal{V},\bar{\mathcal{V}}}}{M_1^\nu}\}$, $\mathcal{V}$ is asymptotically stable and $\Gamma_\nu \cap \mathcal{C}$ is forward invariant and a subset of the region of attraction of $\mathcal{V}$.

Proof: Let $z_\epsilon(x) := L_fV(x) + L_gV(x)u_\epsilon^{\text{safe}}(x) + W(x)$. It follows from (7) that

$$z_\epsilon(x) = \begin{cases} L_fV(x) + W(x) - \frac{1}{\epsilon}\|L_gV(x)\|^2 & \text{if } e(x) \leq 0, \\ B(x) + \frac{1}{\epsilon}C(x) & \text{if } e(x) > 0. \end{cases}$$

We show that $z_\epsilon(x) \leq 0$ for all $x \in \mathcal{C}\backslash\mathcal{V}$ if $\epsilon < \bar{\epsilon}$, from which the result follows. First, note that $\bar{\mathcal{V}}$ as required in the statement exists because V is a CLF and hence, any point $x \neq 0$ that satisfies $L_gV(x) = 0$ is such that $L_fV(x) + W(x) < 0$ (without loss of generality, since if $L_fV(x) + W(x) = 0$ we can take $\tilde{W}(x) = \frac{1}{2}W(x)$). Hence, by continuity there exists a neighborhood $\bar{\mathcal{V}}$ of P_g where $L_fV(x) + W(x) - \frac{1}{\epsilon}\|L_gV(x)\|^2 \leq L_fV(x) + W(x) < 0$ for all $x \in \bar{\mathcal{V}}$, for any $\epsilon > 0$. Hence by taking $\epsilon < \bar{\epsilon}$, we ensure that $z_\epsilon(x) \leq 0$ for all $x \in \bar{\mathcal{V}}$ independently of the sign of $e(x)$. Note also that $\mathcal{W}$ as required in the statement exists because Γ_ν does not contain any point where V and h are incompatible and therefore by Lemma 5.2, all points in Γ_ν satisfying $C(x) = 0$ necessarily also satisfy $B(x) < 0$ (without loss of generality, using a similar argument as above). By continuity of $B(x)$, for any $\epsilon > 0$ we can take a neighborhood $\mathcal{W}$ around P_ν so that $B(x) + \frac{1}{\epsilon}C(x) \leq B(x) \leq 0$ for all $x \in \mathcal{W}$ (by Cauchy-Schwartz's inequality, $C(x) \leq 0$). Hence, by taking $\epsilon < \bar{\epsilon}$, independently of whether $e(x) \leq 0$ or $e(x) > 0$ we ensure that $z_\epsilon(x) \leq 0$ for all $x \in \mathcal{W} \cup \bar{\mathcal{V}}$. Now we argue that if $\epsilon < \bar{\epsilon}$, $z_\epsilon(x) \leq 0$ for all $x \in \Gamma_\nu\backslash(\mathcal{W} \cup \mathcal{V} \cup \bar{\mathcal{V}})$. Note that $\Gamma_\nu\backslash(\mathcal{W} \cup \mathcal{V} \cup \bar{\mathcal{V}})$ does not contain any points where $L_gV(x)$ and $L_gW(x)$ are linearly dependent, since that would imply $C(x) = 0$ and hence $x \in \mathcal{W}$. Thus, by Cauchy-Schwartz's inequality, $C(x) < 0$ for all $x \in \Gamma_\nu\backslash(\mathcal{W}\cup\mathcal{V}\cup\bar{\mathcal{V}})$. Hence, $M_4^{\nu,\mathcal{V},\mathcal{W}} > 0$. Note also that $M_3^{\nu,\mathcal{V},\bar{\mathcal{V}}} > 0$. Therefore, regardless of whether $e(x) \leq 0$ or $e(x) > 0$, by taking $\epsilon < \bar{\epsilon}$ we ensure that $z_\epsilon(x) \leq 0$ for all $x \in \Gamma_\nu\backslash(\mathcal{W} \cup \mathcal{V} \cup \bar{\mathcal{V}})$, as claimed. Moreover, since by construction u_ϵ^{safe} satisfies (3) and is Lipschitz, by [4, Theorem 2], trajectories stay inside $\mathcal{C}$ for all $t \geq 0$. ∎

Note that in the statement of Theorem 5.3, one can pick $\mathcal{V}$ arbitrarily small, which might require an arbitrarily small ϵ. The next result states that under some additional reasonable assumptions, this does not happen and hence there exists a finite ϵ for which trajectories converge to the origin.

Corollary 5.4: (Convergence to the origin): Under the same assumptions and notation of Theorem 5.3, assume additionally that $f, g \in \mathcal{C}^1(\mathbb{R}^n)$, $V \in \mathcal{C}^2(\mathbb{R}^n)$, $0 \in \text{int}(\mathcal{C})$ and $\ker(g(0)^T) \subseteq \mathcal{V}_s(\frac{\partial f}{\partial x}(0))$. Then, for $\epsilon < \hat{\epsilon} := \min\{\frac{\bar{\lambda}_{\min}(g(0)g(0)^T\nabla^2V(0))}{|\lambda_{\max}(\frac{\partial f}{\partial x}(0))|}, \bar{\epsilon}\}$, the origin is asymptotically stable and $\Gamma_\nu \cap \mathcal{C}$ is forward invariant and a subset of the region of attraction of the origin.

Proof: Since $0 \in \text{int}(\mathcal{C})$, $e(0) < 0$ and the Jacobian at 0 of the closed-loop system is $J = \frac{\partial f}{\partial x}(0) - \frac{1}{\epsilon}g(0)g(0)^T\nabla^2V(0)$. We show that, with $\epsilon < \hat{\epsilon}$, one has $v^TJv < 0$ for $v \in \mathbb{R}^n\backslash\{0\}$. First, consider $v \in \ker(g(0)^T)$. By assumption, $v \in \mathcal{V}_s(\frac{\partial f}{\partial x}(0))$, and hence $v^TJv < 0$. Now, assume $v \notin \ker(g(0)^T)$. Since $\nabla^2V(0)$ is positive definite and $g(0)g(0)^T$ is positive semidefinite, $\ker(g(0)g(0)^T\nabla^2V(0)) = \ker(g(0)g(0)^T)$ and $g(0)g(0)^T\nabla^2V(0)$ has non-negative eigenvalues [18, 7.2.P21]. Hence, $v^TJv \leq (\lambda_{\max}(\frac{\partial f}{\partial x}(0)) - \frac{1}{\epsilon}\bar{\lambda}_{\min}(g(0)g(0)^T\nabla^2V(0)))\|v\|^2$. This implies that J is Hurwitz, because $J + J^T$ is negative definite and the real parts of its eigenvalues are twice those of J. Thus, we can take $\mathcal{V}$ in Theorem 5.3 such that the closed-loop trajectories with $\epsilon < \frac{\bar{\lambda}_{\min}(g(0)g(0)^T\nabla^2V(0))}{|\lambda_{\max}(\frac{\partial f}{\partial x}(0))|}$ starting at $\mathcal{V}$ converge to 0. Finally, arguing as in Theorem 5.3, V decreases on $\Gamma_\nu\backslash\mathcal{V}$, and the result follows. ∎

Under the assumptions of Corollary 5.4, by ensuring that the origin is asymptotically stable in Γ_ν, we rule out the existence of equilibrium points in Γ_ν other than the origin. Theorem 5.3 and Corollary 5.4 solve Problem 1. Under the stated assumptions, by taking u_ϵ^{safe} with $\epsilon < \hat{\epsilon}$ as a safe stabilizing controller, an inner approximation of the region of attraction of the origin is the largest level set of V that does not contain any incompatible points inside it. In particular, if there exists a sublevel set of V that contains $\mathcal{C}$, u_ϵ^{safe} with $\epsilon < \hat{\epsilon}$ safely stabilizes the origin and the whole safe set $\mathcal{C}$ is in its region of attraction.

VI. SIMULATIONS

Here, we compare the stability QP with safety penalty controller with the CLF-CBF QP (4) and its modification, M-CLF-CBF QP, introduced in [11, Theorem 3] to avoid undesired equilibria. We focus on the following planar system

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = \begin{pmatrix}x_1\\ x_2\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} u. \quad (9)$$

For this system, $V(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$ is a CLF. The safe set $\mathcal{C}$ is the complement of the ball $\{x \in \mathbb{R}^2 : \|x - (0, 4)\| \leq 2\}$, and we use the CBF $h(x_1, x_2) = x_1^2 + (x_2 - 4)^2 - 4$, with $\alpha(s) = s$. According to [11], the CLF-CBF QP (4) creates undesired equilibria in $\text{int}(\mathcal{C})$ for all values of p. Instead, both M-CLF-CBF QP and the stability QP with safety penalty controller u_ϵ^{safe}, with $\epsilon \neq 1$, do not introduce undesired equilibria in $\text{int}(\mathcal{C})$. The latter can be checked from the definition of $\mathcal{Q}_1^\epsilon$ given in Proposition 5.1. In this example, the incompatible points are given by $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 = 0, x_2 > 4\}$. Therefore, the approximation of the region of attraction given by Theorem 5.3 is $\Gamma_2 = \{x \in \mathbb{R}^2 : \|x\|^2 < 4\}$. Figure 1 shows that the stability QP with safety penalty controller and M-CLF-CBF QP behave similarly, whereas CLF-CBF QP (4) fails to stabilize the origin. The plot also illustrates that trajectories starting at $(0, 9)$ converge to the boundary equilibrium point at $(0, 6)$ for all three approaches (this corresponds to a point where f, gL_gV, and gL_gh are collinear, cf. Remark 2). This is not surprising since, for scenarios where the unsafe set is bounded, global convergence with a smooth vector field is impossible due to topological obstructions [6]. An advantage of the approach proposed here is the explicit inner approximation of the region of attraction which, as Figure 1 shows, is conservative.

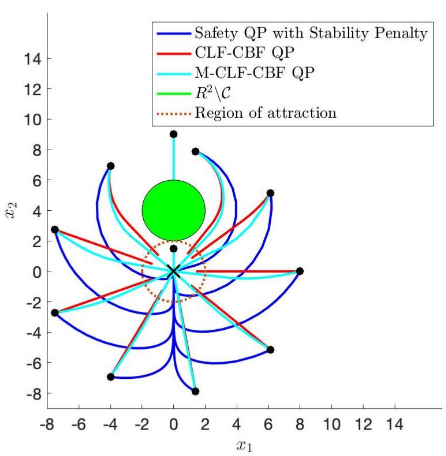

Fig. 1: Safe stabilization of a planar system. The green ball is the set of unsafe states and the small dots display ten initial conditions for the system trajectories under the CLF-CBF QP, the M-CLF-CBF QP, and the safety QP with stability penalty controllers. The orange dotted curve marks the boundary of the estimate Γ_2 of the region of attraction. The CLF-CBF QP controller (with $p = 1$) preserves safety but does not reach the origin because of undesired equilibrium points. The safety QP with stability penalty (with $\epsilon = 0.01$) and the M-CLF-CBF QP (with $p = 1$) preserve safety and have trajectories converge to the origin, except for the one starting at $(0, 9)$.

VII. CONCLUSIONS

We have addressed the problem of safe stabilization of nonlinear affine control systems by proposing an optimization-based feedback design framework inspired by penalty methods for constrained optimization. Our design enforces strictly either stability or safety via a hard constraint while promoting the satisfaction of the other property via a soft constraint. We have characterized the equilibria of the closed-loop system under the proposed controllers. We have shown how to tune the penalty parameter to eliminate spurious equilibria and to increase the region of attraction to all Lyapunov level sets that do not include points where the CLF and the CBF are not compatible. Future work will develop tighter estimates of the region of attraction, consider extra design parameters and explore the extension of the proposed framework to generalized notions of CBFs.

REFERENCES

[1] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed., ser. TAM. Springer, 1998, vol. 6.
[2] R. A. Freeman and P. V. Kototovic, *Robust Nonlinear Control Design: State-space and Lyapunov Techniques*. Cambridge, MA, USA: Birkhauser Boston Inc., 1996.
[3] P. Wieland and F. Allgöwer, "Constructive safety using control barrier functions," *IFAC Proceedings Volumes*, vol. 40, no. 12, pp. 462–467, 2007.
[4] 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, June 2019, pp. 3420–3431.
[5] M. Z. Romdlony and B. Jayawardhana, "Stabilization with guaranteed safety using control Lyapunov-barrier function," *Automatica*, vol. 66, pp. 39–47, 2016.
[6] P. Braun and C. M. Kellett, "On (the existence of) control Lyapunov barrier functions," Newcastle, Australia, 2017.
[7] P. Ong and J. Cortés, "Universal formula for smooth safe stabilization," in *IEEE Conf. on Decision and Control*, Nice, France, Dec. 2019, pp. 2373–2378.
[8] B. J. Morris, M. J. Powell, and A. D. Ames, "Continuity and smoothness properties of nonlinear optimization-based feedback controllers," in *IEEE Conf. on Decision and Control*, Osaka, Japan, Dec 2015, pp. 151–158.
[9] M. Jankovic, "Robust control barrier functions for constrained stabilization of nonlinear systems," *Automatica*, vol. 96, pp. 359–367, 2018.
[10] 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.
[11] X. Tan and D. V. Dimarogonas, "On the undesired equilibria induced by control barrier function based quadratic programs," *arXiv preprint arXiv:2104.14895*, 2021.
[12] 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.
[13] W. S. Cortez and D. V. Dimarogonas, "On compatibility and region of attraction for safe, stabilizing control laws," *arXiv preprint arXiv:2008.12179*, 2021.
[14] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames, "Robustness of control barrier functions for safety critical control," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 54–61, 2015.
[15] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Addison-Wesley, 1984.
[16] ——, *Optimization by Vector Space Methods*. Wiley, 1969.
[17] X. Xu, "Constrained control of input-output linearizable systems using control sharing barrier functions," *Automatica*, vol. 87, pp. 195–201, 2018.
[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.