Control Barrier Function-Based Design of Gradient Flows for Constrained Nonlinear Programming

Ahmed Allibhoy Jorge Cortés

***Abstract*— This paper considers the problem of designing a continuous-time dynamical system that solves constrained nonlinear optimization problems and makes the feasible set forward invariant and asymptotically stable. The invariance of the feasible set makes the dynamics anytime, when viewed as an algorithm, meaning it returns a feasible solution regardless of when it is terminated. Our approach augments the gradient flow of the objective function with inputs defined by the constraint functions, treats the feasible set as a safe set, and synthesizes a safe feedback controller using techniques from the theory of control barrier functions. The resulting closed-loop system, termed safe gradient flow, can be viewed as a primal-dual flow, where the state corresponds to the primal variables and the inputs correspond to the dual ones. We provide a detailed suite of conditions based on constraint qualification under which (both isolated and nonisolated) local minimizers are asymptotically stable with respect to the feasible set and the whole state space. Comparisons with other continuous-time methods for optimization in a simple example illustrate the advantages of the safe gradient flow.**

I. Introduction

Optimization problems are ubiquitous in engineering and applied science. The traditional emphasis on the numerical analysis of algorithms is motivated by the implementation on digital platforms. The alternate viewpoint of optimization algorithms as continuous-time dynamical systems taken here also has a long history, often as a precursor of the synthesis of discrete-time algorithms. This viewpoint has been fruitful for gaining insight into qualitative properties such as stability and convergence.

For constrained optimization problems, the picture is complicated by the fact that algorithms may need to ensure convergence to the optimizer as well as enforce feasibility of the iterates. The latter is important in real-time applications, when feasibility guarantees may be required at all times in case the algorithm is terminated before completion, or when the algorithm is implemented on a physical plant where the constraints encode its safe operation. In this paper, we show that, just as unconstrained optimization algorithms can be viewed as dynamical systems, constrained optimization algorithms can be viewed as control systems. Within this framework, the task of designing an optimization algorithm for a constrained problem is equivalent to that of designing a feedback controller for a nonlinear system. We use this connection to derive a novel control-theoretic algorithm for solving constrained nonlinear programs that combines continuous-time gradient flows to optimize the objective function with techniques from control barrier functions to maintain invariance of the feasible set.

Related Work: Dynamical systems and optimization are closely intertwined disciplines [2]–[4]. The work [5] provides a contemporary review of the dynamical systems approach to optimization for both constrained an unconstrained problems, with an emphasis on applications where the optimization problem is in a feedback loop with a plant, see e.g. [6]–[8]. Examples of such scenarios are numerous, including power systems [7], [8], network congestion [9], and transportation [10].

Flows for Equality Constrained Problems: For problems involving only equality constraints, [11], [12] employ differential geometric techniques to design a vector field that maintains feasibility along the flow, makes the constraint set asymptotically stable, and whose solutions converge to critical points of the objective function. The work [13] introduces a generalized form of this vector field to deal with inequality constraints in the form of a differential algebraic equation, and explores links with sequential quadratic programming.

Projected Gradient Methods: Another approach to solving nonlinear programs in continuous time makes use of projected dynamical systems [14] by projecting the gradient of the objective function onto the cone of feasible descent directions, see e.g., [15]. However, projected dynamical systems are, in general, discontinuous, which from an analysis viewpoint requires properly dealing with notions and existence of solutions, cf. [16]. The work [17] proposes a continuous modification of the projected gradient method, whose stability is analyzed in [18]. However, this method projects onto the constraint set itself, rather than the tangent cone, and may fail when it is nonconvex. Another modification is the "constrained gradient flow" proposed in [19], derived using insights from nonsmooth mechanics, and is well-defined outside the feasible set. The resulting method is related to the one presented here and converges to critical points, though the dynamics are once again discontinuous, and stability guarantees are only provided in the case of convexity, which we do not assume.

Saddle-Point Dynamics: Convex optimization problems can be solved by searching for saddle points of the associated Lagrangian via a primal-dual dynamics consisting of a gradient descent in the primal variable and a gradient ascent in the dual one. The analysis of stability and convergence of this method has a long history [2], [20], with more recent accounts provided for discrete-time implementations [21] and continuous-time

A preliminary version of this work appeared as [1] at the IEEE Conference on Decision and Control.

A. Allibhoy and J. Cortés are with the Department of Mechanical and Aerospace Engineering, UC San Diego, {aallibho,cortes}@ucsd.edu

ones [22]–[24]. These methods are particularly well suited for distributed implementation on network optimization problems, but they do not leave the feasible set invariant.

Contributions: We consider the synthesis of continuous-time dynamical systems that solve constrained optimization problems while making the feasible set forward invariant and asymptotically stable. Our first contribution is the design of the safe gradient flow for constrained optimization using the framework of safety-critical control. The basic intuition is to combine the standard gradient flow to optimize the objective function with the idea of keeping the feasible set safe. To maintain safety, we augment the gradient dynamics with inputs associated with the constraint functions and use a control barrier function approach to design an optimization-based feedback controller that ensures forward invariance and asymptotic stability of the feasible set. The approach is primal-dual, in the sense that the states correspond to the primal variables and the inputs correspond to the dual variables.

Our second contribution unveils the connection of the proposed dynamics with the projected gradient flow. Specifically, we provide an alternate derivation of the safe gradient flow as a continuous modification of the projected gradient flow, based on a design parameter. We show that, as the parameter grows to ∞, the safe gradient flow becomes the projected gradient flow.

In addition to establishing an interesting parallelism, we build on this equivalence in our third contribution for understanding the regularity and stability properties of the safe gradient flow. We show that the flow is locally Lipschitz (ensuring the existence and uniqueness of classical solutions), well defined on an open set containing the feasibility region (which allows for the possibility of infeasible initial conditions), that its equilibria exactly correspond to the critical points of the optimization problem, and that the objective function is monotonically decreasing along the feasible set of the optimization problem. Lastly, we prove that the feasible set is forward invariant and asymptotically stable.

Our fourth contribution consists of a thorough stability analysis of the critical points of the optimization problem under the safe gradient flow. We provide a suite of constraint qualification-based conditions under which isolated local minimizers are either locally asymptotically stable with respect to the feasible set, locally asymptotically stable with respect to the global state space, or locally exponentially stable. We also provide conditions for semistability of nonisolated local minimizers and establish global convergence to critical points of the optimization problem. Our technical analysis for this builds on a combination of the Kurdyaka-Łojasiewicz inequality with a novel angle-condition Lyapunov test to establish the finite arclength of trajectories, which we present in the appendix.

A preliminary version of this work appeared previously as [1]. The present work significantly expands the scope of the stability analysis of isolated local minimizers under weaker assumptions, as well as characterizes the stability of nonisolated local minimizers, global convergence to critical points, and highlights the advantages of the safe gradient flow over other continuous-time methods in optimization.

Notation: We let $\mathbb{R}$ denote the set of real numbers. For $v, w \in \mathbb{R}^n$, $v \leq w$ (resp. $v < w$) denotes $v_i \leq w_i$ (resp. $v_i < w_i$) for $i \in \{1, \dots, n\}$. We let $\|v\|$ denote the Euclidean norm and $\|v\|_\infty = \max_{1\leq i\leq n} |v_i|$ the infinity norm. For $y \in \mathbb{R}$, we denote $[y]_+ = \max\{0, y\}$, and $\operatorname{sgn}(y) = 1$ if $y > 0$, $\operatorname{sgn}(y) = -1$ if $y < 0$ and $\operatorname{sgn}(y) = 0$ if $y = 0$. We let $\mathbf{1}_m \in \mathbb{R}^m$ denote the vector of all ones. For a matrix $A \in \mathbb{R}^{n\times m}$, we use $\rho(A)$ and $A^\dagger$ to denote its spectral radius and its Moore-Penrose pseudoinverse, respectively. We write $A \succeq 0$ (resp., $A \succ 0$) to denote A is positive semidefinite (resp., A is positive definite). Given a subset $\mathcal{C} \subset \mathbb{R}^n$, the distance of $x \in \mathbb{R}^n$ to $\mathcal{C}$ is $\operatorname{dist}_\mathcal{C}(x) = \inf_{y\in\mathcal{C}}\|x - y\|$. We let $\overline{\mathcal{C}}$, $\operatorname{int}(\mathcal{C})$, and $\partial\mathcal{C}$ denote the closure, interior, and boundary of $\mathcal{C}$, respectively. Given $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}^m$, the graph of f is $\operatorname{graph}(f) = \{(x, f(x)) \mid x \in X\}$. Similarly, given a set-valued map $\mathcal{F} : X \rightrightarrows \mathbb{R}^m$, its graph is $\operatorname{graph}(\mathcal{F}) = \{(x, y) \mid x \in X, y \in \mathcal{F}(x)\}$. Given $g : \mathbb{R}^n \to \mathbb{R}$, we denote its gradient by ∇g and its Hessian by $\nabla^2 g$. For $g : \mathbb{R}^n \to \mathbb{R}^m$, $\frac{\partial g(x)}{\partial x}$ denotes its Jacobian. For $I \subset \{1, 2, \dots, m\}$, we denote by $\frac{\partial g_I(x)}{\partial x}$ the matrix whose rows are $\{\nabla g_i(x)^\top\}_{i\in I}$.

II. Preliminaries

We present notions on invariance, stability, variational analysis, control barrier functions, and nonlinear programming. The reader familiar with the material can safely skip the section.

A. Invariance and Stability Notions

We recall basic definitions from the theory of ordinary differential equations [25]. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz vector field and consider the dynamical system $\dot{x} = F(x)$. Local Lipschitzness ensures that, for every initial condition $x_0 \in \mathbb{R}^n$, there exists $T > 0$ and a unique trajectory $x : [0, T] \to \mathbb{R}^n$ such that $x(0) = x_0$ and $\dot{x}(t) = F(x(t))$. If the solution exists for all $t \geq 0$, then it is *forward complete*. In this case, the *flow map* is defined by $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ such that $\Phi_t(x) = x(t)$, where $x(t)$ is the unique solution with $x(0) = x$. The *positive limit set* of $x \in \mathbb{R}^n$ is

$$\omega(x) = \bigcap_{T\geq 0} \overline{\{\Phi_t(x) \mid t > T\}}.$$

A set $\mathcal{K} \subset \mathbb{R}^n$ is *forward invariant* if $x \in \mathcal{K}$ implies that $\Phi_t(x) \in \mathcal{K}$ for all $t \geq 0$. If $\mathcal{K}$ is forward invariant and $x^* \in \mathcal{K}$ is an equilibrium, x^* is *Lyapunov stable relative to* $\mathcal{K}$ if for every open set U containing x^*, there exists an open set $\tilde{U}$ also containing x^* such that for all $x \in \tilde{U} \cap \mathcal{K}$, $\Phi_t(x) \in U \cap \mathcal{K}$ for all $t > 0$. The equilibrium x^* is *asymptotically stable relative to* $\mathcal{K}$ if it is Lyapunov stable relative to $\mathcal{K}$ and there is an open set U containing x^* such that $\Phi_t(x) \to x^*$ as $t \to \infty$ for all $x \in U \cap \mathcal{K}$. We say x^* is *exponentially stable relative to* $\mathcal{K}$ if it is asymptotically stable relative to $\mathcal{K}$ and there exists $\mu > 0$ and an open set U containing x^* such that for all $x \in U \cap \mathcal{K}$, $\|\Phi_t(x) - x^*\| \leq e^{-\mu t}\|x - x^*\|$. Analogous definitions of Lyapunov stability and asymptotically stability can be made for sets, instead of individual points.

Consider a forward invariant set $\mathcal{K}$ and a set of equilibria $\mathcal{S}$ contained in it, $\mathcal{S} \subset \mathcal{K}$. We say $x^* \in \mathcal{S}$ is *semistable relative to* $\mathcal{K}$ if x^* is Lyapunov stable and, for any open

set U containing x^*, there is an open set $\tilde{U}$ such that for every $x \in \tilde{U} \cap \mathcal{K}$, the trajectory starting at x converges to a Lyapunov stable equilibrium in $U \cap \mathcal{S}$. Note that if x^* is an isolated equilibrium, then semistability is equivalent to asymptotic stability. For all the concepts introduced here, when the invariant set is unspecified, we mean $\mathcal{K} = \mathbb{R}^n$.

B. Variational Analysis

We review basic notions from variational analysis following [26]. The *extended real line* is $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, its *domain* is $\mathrm{dom}(f) = \{x \in \mathbb{R}^n \mid f(x) \neq \infty, -\infty\}$. The *indicator function of* $\mathcal{C} \subset \mathbb{R}^n$ is $\delta_{\mathcal{C}} : \mathbb{R}^n \to \overline{\mathbb{R}}$,

$$\delta_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C}, \\ \infty & \text{if } x \notin \mathcal{C}. \end{cases}$$

Note that $\mathrm{dom}(\delta_{\mathcal{C}}) = \mathcal{C}$. For $x \in \mathrm{dom}(f)$ and $d \in \mathbb{R}^n$, consider the following limits

$$f'(x;d) = \lim_{(h,y)\to(0^+,x)} \frac{f(y+hd) - f(x)}{h}, \tag{1a}$$

$$f''(x;d) = \lim_{(h,y)\to(0^+,x)} \frac{f(y+hd) - f(x) - hf'(y;d)}{h^2}. \tag{1b}$$

If the limit in (1a) (resp. (1b)) exists, f is *directionally differentiable in the direction* d (resp. *twice directionally differentiable in the direction* d). By definition, $f'(x;d) = \nabla f(x)^\top d$ if f is continuously differentiable at x and $f''(x;d) = d^\top \nabla^2 f(x) d$ if f is twice continuously differentiable at x.

Given a dynamical system $\dot{x} = F(x)$ and a function $V : \mathbb{R}^n \to \mathbb{R}$, the *upper-right Dini derivative* of V along solutions of the system is

$$D_F^+ V(x) = \limsup_{h\to 0^+} \frac{1}{h} \left[V(\Phi_h(x)) - V(x) \right],$$

where Φ_h is the flow map of the system. If V is directionally differentiable then $D_F^+ V(x) = V'(x; F(x))$, and if V is differentiable then $D_F^+ V(x) = \nabla V(x)^\top F(x)$.

The *tangent cone* to $\mathcal{C} \subset \mathbb{R}^n$ at $x \in \mathbb{R}^n$ is

$$T_{\mathcal{C}}(x) = \big\{ d \in \mathbb{R}^n \mid \exists \{t^\nu\}_{\nu=1}^\infty \subset (0,\infty), \{x^\nu\}_{\nu=1}^\infty \subset \mathcal{C} \\ t^\nu \to 0^+, x^\nu \to x, \frac{x^\nu - x}{t^\nu} \to d \text{ as } \nu \to \infty \big\}.$$

If $\mathcal{C}$ is an embedded submanifold of $\mathbb{R}^n$, then the tangent cone coincides with the usual differential geometric notion of tangent space. Let $\Pi_{\mathcal{C}} : \mathbb{R}^n \rightrightarrows \overline{\mathcal{C}}$, $\Pi_{\mathcal{C}}(x) = \big\{ y \in \overline{\mathcal{C}} \mid \|x - y\| = \mathrm{dist}_{\mathcal{C}}(x) \big\}$, be the projection map onto $\overline{\mathcal{C}}$. The *proximal normal cone* to $\mathcal{C}$ at x is

$$\mathcal{N}_{\mathcal{C}}^{\mathrm{prox}}(x) = \big\{ d \in \mathbb{R}^n \mid \exists \{t^\nu\}_{\nu=1}^\infty \subset (0,\infty), \\ \{(x^\nu, y^\nu)\}_{\nu=1}^\infty \subset \mathrm{graph}(\Pi_{\mathcal{C}}), \\ t^\nu \to 0^+, x^\nu \to x, \frac{x^\nu - y^\nu}{t^\nu} \to d \text{ as } \nu \to \infty \big\}.$$

C. Safety Critical Control via Control Barrier Functions

We introduce here basic concepts from safety and a method for synthesizing safe controllers using vector control barrier functions. Our exposition here slightly generalizes [27], [28] to set the stage for dealing with constrained optimization problems later. Consider a control-affine system

$$\dot{x} = F_0(x) + \sum_{i=1}^r u_i F_i(x), \tag{2}$$

with locally Lipschitz vector fields $F_i : \mathbb{R}^n \to \mathbb{R}^n$, for $i \in \{0, \ldots, r\}$, and a set $\mathcal{U} \subset \mathbb{R}^m$ of valid control inputs. Let $\mathcal{C} \subset \mathbb{R}^n$ represent the set of states where the system can operate safely and $u : X \to \mathcal{U}$ be a locally Lipschitz feedback controller, with $X \subset \mathbb{R}^n$ a set containing $\mathcal{C}$. The closed-loop system (2) under u is *safe* with respect to $\mathcal{C}$ if $\mathcal{C}$ is forward invariant under the closed-loop system.

Feedback controllers can be certified to be safe by resorting to the notion of control barrier function, which we here generalize for convenience. Let $\mathcal{C} \subset X \subset \mathbb{R}^n$ and $m, k \in \mathbb{Z}_{\geq 0}$. A (m,k)*-vector control barrier function* (VCBF) of $\mathcal{C}$ on X relative to $\mathcal{U}$ is a continuously differentiable function $\phi : \mathbb{R}^n \to \mathbb{R}^{m+k}$ such that the following properties hold:

(i) The safe set can be expressed using m inequality constraints and k equality constraints:

$$\mathcal{C} = \{ x \in \mathbb{R}^n \mid \phi_i(x) \leq 0,\ 1 \leq i \leq m, \\ \phi_j(x) = 0,\ m+1 \leq j \leq m+k \};$$

(ii) there exists $\alpha > 0$ such that the map $K : \mathbb{R}^n \rightrightarrows \mathcal{U}$,

$$K_\alpha(x) = \big\{ u \in \mathcal{U} \mid \\ D_{F_0}^+ \phi_i(x) + \sum_{\ell=1}^r u_\ell D_{F_\ell}^+ \phi_i(x) + \alpha \phi_i(x) \leq 0, \\ D_{F_0}^+ \phi_j(x) + \sum_{\ell=1}^r u_\ell D_{F_\ell}^+ \phi_j(x) + \alpha \phi_j(x) = 0, \\ 1 \leq i \leq m,\ m+1 \leq j \leq m+k \big\},$$

takes nonempty values for all $x \in X$.

In the special case where $m = 1$ and $k = 0$, this definition coincides with the usual notion of control barrier function [28, Definition 2], where the class $\mathcal{K}$ function is linear, and the Lie derivative has been replaced with the upper-right Dini derivative. In general, the problem of finding a suitable VCBF ϕ is problem-specific: in many cases, the function naturally emerges from formalizing mathematically the safety specifications one seeks to enforce. The use of vector-valued functions instead of scalar-valued ones allows to consider a broader class of safe sets. If ϕ is a VCBF and u is a feedback where $u(x) \in K_\alpha(x)$, it follows that along solutions to (2), $\frac{d}{dt}\phi_i(x) \leq -\alpha\phi_i(x)$ for $1 \leq i \leq m$ and $\frac{d}{dt}\phi_j(x) = -\alpha\phi_j(x)$ for $m+1 \leq j \leq m+k$, which implies safety of $\mathcal{C}$. This is stated formally in the next generalization of [28, Theorem 2].

Lemma 2.1 (Safe feedback control): Consider the system (2) with safety set $\mathcal{C}$ and let ϕ be a vector control barrier function for $\mathcal{C}$ on X. Then, any feedback controller $u : X \to \mathcal{U}$ satisfying $u(x) \in K_\alpha(x)$ for all $x \in X$ and such that

$x \mapsto F_0(x) + \sum_{i=0}^{m} u_i(x)F_i(x)$ is locally Lipschitz renders $\mathcal{C}$ forward invariant and asymptotically stable.

While Lemma 2.1 provides sufficient conditions for feedback controller to be safe, it does not specify how to synthesize it. A common technique [27] is, for each $x \in X$, to define $u(x)$ as the minimum-norm element of $K_\alpha(x)$. Here, we pursue an alternative design of the form:

$$u(x) \in \underset{u \in K_\alpha(x)}{\operatorname{argmin}} \left\{ \left\| \sum_{i=1}^{r} u_i F_i(x) \right\|^2 \right\}. \tag{3}$$

This design has the interpretation of finding a controller which guarantees safety while modifying the drift term in (2) as little as possible. In general, Lipschitz continuity of the closed-loop dynamics under either design is not guaranteed, cf. [29], so additional assumptions may be needed to establish safety via Lemma 2.1.

D. Optimality Conditions for Nonlinear Programming

We present the basic background on necessary conditions for optimality [30]. Consider a nonlinear program of the form:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g(x) \leq 0 \\ & h(x) = 0, \end{aligned} \tag{4}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^k$ are continuously differentiable. Let

$$\mathcal{C} = \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}, \tag{5}$$

denote its feasible set. Necessary conditions for optimality can be derived provided that the feasible set satisfies appropriate constraint qualification conditions. Let the active constraint, constraint violation, and inactive constraint sets be

$$\begin{aligned} I_0(x) &= \{1 \leq i \leq m \mid g_i(x) = 0\}, \\ I_+(x) &= \{1 \leq i \leq m \mid g_i(x) > 0\}, \\ I_-(x) &= \{1 \leq i \leq m \mid g_i(x) < 0\}, \end{aligned}$$

respectively. We say that the optimization problem (4) satisfies

- the Mangasarian-Fromovitz Constraint Qualification (MFCQ) if $\{\nabla h_j(x)\}_{j=1}^k$ are linearly independent and there exists $\xi \in \mathbb{R}^n$ such that $\nabla h_j(x)^\top \xi = 0$ for all $j \in \{1, \ldots, k\}$ and $\nabla g_i(x)^\top \xi < 0$ for all $i \in I_0(x)$;
- the Extended Mangasarian-Fromovitz Constraint Qualification (EMFCQ) if $\{\nabla h_j(x)\}_{j=1}^k$ are linearly independent and there exists $\xi \in \mathbb{R}^n$ such that $\nabla h_j(x)^\top \xi = 0$ for all $j \in \{1, \ldots, k\}$ and $\nabla g_i(x)^\top \xi < 0$ for all $i \in I_0(x) \cup I_+(x)$;
- the Linear Independence Constraint Qualification (LICQ) at x, if $\{\nabla g_i(x)\}_{i \in I_0(x)} \cup \{\nabla h_j(x)\}_{j=1}^k$ are linearly independent.

If $x^* \in \mathcal{C}$ is a local minimizer, and any of the above constraint qualification conditions hold at x^*, then there exists $u^* \in \mathbb{R}^m$ and $v^* \in \mathbb{R}^k$ such that the *Karash-Kuhn-Tucker* conditions hold,

$$\nabla f(x^*) + \frac{\partial g(x^*)}{\partial x}^\top u^* + \frac{\partial h(x^*)}{\partial x}^\top v^* = 0, \tag{6a}$$

$$g(x^*) \leq 0, \tag{6b}$$

$$h(x^*) = 0, \tag{6c}$$

$$u^* \geq 0, \tag{6d}$$

$$(u^*)^\top g(x^*) = 0. \tag{6e}$$

The pair (u^*, v^*) are called Lagrange multipliers, and the triple (x^*, u^*, v^*) satisfying (6) is referred to as a *KKT triple*. We denote the set of KKT points of (4) by

$$X_{\text{KKT}} = \{x^* \in \mathbb{R}^n \mid \exists (u^*, v^*) \in \mathbb{R}^m \times \mathbb{R}^k \text{ such that } (x^*, u^*, v^*) \text{ solves (6)}\}.$$

III. Problem Formulation

Our goal is to solve the optimization problem (4) by designing a dynamical system $\dot{x} = F(x)$ that converges to its solutions. The dynamics should enjoy the following properties:

(i) trajectories should remain feasible if they start from a feasible point. This can be formalized by asking the feasible set $\mathcal{C}$, defined in (5), to be forward invariant;

(ii) trajectories that start from an infeasible point should converge to the set of feasible points. This can be formalized by requiring that F is well defined on an open set containing $\mathcal{C}$, and that $\mathcal{C}$ as a set is asymptotically stable with respect to the dynamics.

The requirement (i) ensures that, when viewed as an algorithm, the dynamics is *anytime*, meaning that it is guaranteed to return a feasible solution regardless of when it is terminated. The requirement (ii) ensures in particular that trajectories beginning from infeasible initial conditions approach the feasible set and, if the solutions of the optimization (4) belong to the interior of the feasible set, such trajectories enter it in finite time, never to leave it again. The problem is summarized below:

Problem 1: Find an open set X containing $\mathcal{C}$ and design a vector field $F : X \to \mathbb{R}^n$ such that the system $\dot{x} = F(x)$ satisfies the following properties:

(i) F is locally Lipschitz on X;

(ii) $\mathcal{C}$ is forward invariant and asymptotically stable;

(iii) x^* is an equilibrium if and only if $x^* \in X_{\text{KKT}}$;

(iv) x^* is asymptotically stable if x^* is a isolated local minimizer.

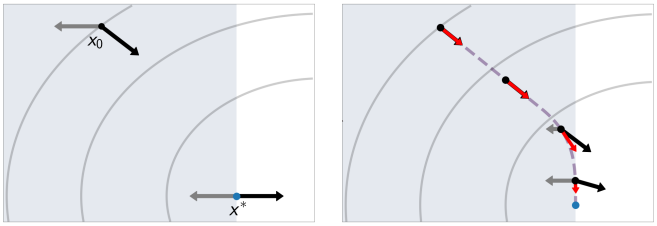
IV. Constrained Nonlinear Programming via Safe Gradient Flow

In this section we introduce our solution to Problem 1 in the form of a dynamical system called the *safe gradient flow*. We present two interpretations of this system: the first is from the perspective of safety critical control, where we augment the standard gradient flow with an input and design a feedback controller using the procedure outlined in Section II-C. The second is as an approximation of the projected gradient flow. Interestingly, we show that both interpretations are equivalent.

A. Safe Gradient Flow via Feedback Control

Consider the control-affine system

$$\dot{x} = -\nabla f(x) - \frac{\partial g(x)}{\partial x}^\top u - \frac{\partial h(x)}{\partial x}^\top v. \tag{7}$$



(a) (b)

Fig. 1: Intuition behind the design of the safe gradient flow. Grey lines are the level curves of the objective function and the shaded region is $\mathcal{C}$. In (a), the initial condition is x_0 and the minimizer is x^*, with $-\nabla f(x)$ in black and $-\nabla g(x)$ in gray at both points. In (b), the dashed line is a trajectory of $\dot{x} = -\nabla f(x) - u\nabla g(x)$ starting from x_0. The black vectors are $-\nabla f(x)$, the gray vectors are $-u\nabla g(x)$, and the red vectors are $\dot{x}$. Deep in the interior of the feasible set, one has $u \approx 0$, as following the gradient of f does not jeopardize feasibility while minimizing it. As the trajectory approaches the boundary, u increases to keep the trajectory in $\mathcal{C}$.

One can interpret this system as the standard gradient flow of f modified by a "control action". The intuition is that the drift term takes care of optimizing f toward the minimizer, and this direction can be modified with the input if the trajectory gets close to the boundary of the feasible set, cf. Figure 1.

Our idea for the controller design is to only modify the drift when the feasibility of the state is endangered. We accomplish this by looking at the feasible set $\mathcal{C}$ as a safe set and using $\phi = (g, h) : \mathbb{R}^n \to \mathbb{R}^{m+k}$ as an (m, k)-vector control barrier function to synthesize the feedback controller, as described next.

Let $\alpha > 0$ be a design parameter. For reasons of space, we sometimes omit the arguments of functions when they are clear from the context. Following Section II-C, define the admissible control set as

$$\begin{aligned} K_\alpha(x) = \Big\{ (u, v) \in \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k \,\Big|\, & \\ -\frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top v &\leq \frac{\partial g}{\partial x}\nabla f(x) - \alpha g(x) \\ -\frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top v &= \frac{\partial h}{\partial x}\nabla f(x) - \alpha h(x) \Big\}. \end{aligned} \tag{8}$$

The next result shows that ϕ is a valid VCBF for (7).

Lemma 4.1: (Vector control barrier function for (7)*):* Consider the optimization problem (4). If MFCQ holds for all $x \in \mathcal{C}$, then there exists an open set X containing $\mathcal{C}$ such that the function $\phi = (g, h) : \mathbb{R}^n \to \mathbb{R}^{m+k}$ is a valid (m, k)-VCBF for (7) on X relative to $\mathcal{U} = \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k$.

Proof: We begin by showing that inequalities parameterizing $K_\alpha(x)$ are strictly feasible for all $x \in \mathcal{C}$, i.e., for each $x \in \mathcal{C}$, there exists $\epsilon > 0$ and $(u, v) \in \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k$ such that

$$-\frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top v \leq \frac{\partial g}{\partial x}\nabla f(x) - \alpha g(x) - \epsilon \mathbf{1}_m \tag{9a}$$

$$-\frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top v = \frac{\partial h}{\partial x}\nabla f(x) - \alpha h(x). \tag{9b}$$

Let $\tilde{g} = g(x) + \frac{\epsilon}{\alpha}\mathbf{1}_m$. By Farka's Lemma [31], (9) is infeasible if and only if there exists a solution (u, v) to

$$-\frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top v \geq 0 \tag{10a}$$

$$-\frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top u - \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top v = 0 \tag{10b}$$

$$u \geq 0 \tag{10c}$$

$$u^\top\left(\frac{\partial g}{\partial x}\nabla f - \alpha\tilde{g}\right) + v^\top\left(\frac{\partial h}{\partial x}\nabla f - \alpha h(x)\right) < 0. \tag{10d}$$

Then (10a), (10b), and (10c) imply that

$$\begin{bmatrix} u \\ v \end{bmatrix}^\top \begin{bmatrix} \frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top \\ \frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \leq 0$$

but, since the matrix is positive semidefinite,

$$(u, v) \in \ker \begin{bmatrix} \frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top \\ \frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top \end{bmatrix} = \ker \begin{bmatrix} \frac{\partial g}{\partial x}^\top & \frac{\partial h}{\partial x}^\top \end{bmatrix}. \tag{11}$$

Next, by (11) and that $x \in \mathcal{C}$, (10d) reduces to

$$-u^\top(\alpha g(x) - \epsilon \mathbf{1}_m) < 0, \tag{12}$$

and by a second application of Farka's Lemma, we see that (10c), (11) and (12) are feasible if and only if the following system is infeasible:

$$\frac{\partial g(x)}{\partial x}\xi \leq -\alpha g(x) - \epsilon \mathbf{1}_m \qquad \frac{\partial h(x)}{\partial x}\xi = 0. \tag{13a}$$

We claim that a solution to (13) can be constructed if MFCQ holds at x. Indeed, by MFCQ, there exists $\tilde{\xi} \in \mathbb{R}^n$ such that $\frac{\partial g_{I_0}}{\partial x}\tilde{\xi} < 0$ and $\frac{\partial h}{\partial x}\tilde{\xi} = 0$, and for ϵ sufficiently small, there exists $\gamma > 0$ such that $\xi = \gamma\tilde{\xi}$ solves (13). Thus (10) is infeasible, and therefore (9) is feasible.

By strict feasibility and the fact that the matrix $\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top$ has full rank, it can be shown by [32, Theorem 2] that, for all $x \in \mathcal{C}$, the affine inequalities that parameterize $K_\alpha(x)$ are *regular*[1]. Finally, since the affine inequalities parameterizing K_α are continuous, $K_\alpha(y)$ is nonempty for any y sufficiently close to x. Hence there exists an open set X such that K_α takes nonempty values on X. ■

Since ϕ is a VCBF, we can design a feedback of the form (3) to maintain safety of $\mathcal{C}$ while modifying the drift term as little as possible. Formally,

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} \in \underset{u, v \in K_\alpha(x)}{\operatorname{argmin}} \left\{ \left\| \frac{\partial g(x)}{\partial x}^\top u + \frac{\partial h(x)}{\partial x}^\top v \right\|^2 \right\}. \tag{14}$$

We refer to the closed-loop system (7) under the controller (14) as the *safe gradient flow*. In general, the solution to (14) might not be unique. Nevertheless, as we show later, the safe gradient flow is well-defined because, the closed-loop behavior of the system is independent of the chosen solution.

Comparing (7) with the KKT equation (6a) suggests that $(u(x), v(x))$ can be interpreted as approximations of the dual variables of the problem. With this interpretation, the safe gradient flow can be viewed as a primal-dual method. We use this viewpoint later to establish connections between the proposed method and the projected gradient flow.

[1]Consider a linear system of inequalities of the form $Cz \leq c$, $Dz = d$, and a solution z_0. The system is regular (c.f. [32]) if for C', c', D', d' sufficiently close to C, c, D, d, the perturbed system $C'z \leq c'$, $D'z = d'$ remains feasible, and the distance of z_0 to the solution set of the perturbed system is proportional to the magnitude of the perturbation.

Remark 4.2: (Connection with the Literature): The work [11] considers the problem of designing a dynamical system to solve (4) when only equality constraints are present using a differential geometric approach. Here, we show that the safe gradient flow generalizes the solution proposed in [11]. Under the assumption that $h \in \mathcal{C}^r$ and LICQ holds, the feasible set $\mathcal{C} = \{x \in \mathbb{R}^n \mid h(x) = 0\}$ is an embedded C^r submanifold of $\mathbb{R}^n$ of codimension k. The approach in [11] proceeds by identifying a vector field $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfying: (i) $F \in \mathcal{C}^r$ and $F(x) \in T_{\mathcal{C}}(x)$ for all $x \in \mathcal{C}$; and (ii) $\dot{h}(x) = -\alpha h(x)$ along the trajectories of $\dot{x} = F(x)$, where $\alpha > 0$ is a design parameter. The proposed vector field satisfying both properties is

$$F(x) = -\Big(I - \frac{\partial h}{\partial x}^{\dagger}\frac{\partial h}{\partial x}\Big)\nabla f(x) - \alpha\frac{\partial h}{\partial x}^{\dagger} h(x). \tag{15}$$

To see that this corresponds to the safe gradient flow, note that the admissible control set (8) in this case is

$$K_\alpha(x) = \Big\{v \in \mathbb{R}^k \mid -\frac{\partial h}{\partial x}\nabla f(x) - \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^{\top} v = -\alpha h(x)\Big\}.$$

By the LICQ assumption, $K_\alpha(x)$ is a singleton whose unique element is

$$v(x) = -\Big(\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^{\top}\Big)^{-1}\Big(\frac{\partial h}{\partial x}\nabla f(x) - \alpha h(x)\Big).$$

Substituting this into (7), we obtain the expression (15). This provides an alternative interpretation from a control-theoretic perspective of the differential-geometric design in [11], and justifies viewing the safe gradient flow as the natural extension to the case with both inequality and equality constraints. •

Remark 4.3: (Inequality Constraints via Quadratic Slack Variables): The work [13] pursues a different approach that the one taken here to deal with inequality constraints by reducing them to equality constraints. This is accomplished introducing quadratic slack variables. Formally, for each $i \in \{1, \dots, m\}$, one replaces the constraint $g_i(x) \leq 0$ with the equality constraint $g_i(x) = -y_i^2$, and solves the equality-constrained optimization problem in the variables $(x, y) \in \mathbb{R}^{n+m}$ with a flow of the form (15). While this method can be expressed in closed form, there are several drawbacks with it. First, this increases the dimensionality of the problem, which can be problematic when there are a large number of inequality constraints. Second, adding quadratic slack variables introduces equilibrium points to the resulting flow which do not correspond to KKT points of the original problem. •

B. Safe Gradient Flow as an Approximation of the Projected Gradient Flow

Here, we introduce an alternative design in terms of a continuous approximation of the projected gradient flow. The latter is a discontinuous dynamical system obtained by projecting the gradient of the objective function onto the tangent cone of the feasible set. Later, we show that this continuous approximation is in fact equivalent to the safe gradient flow.

Let $x \in \mathcal{C}$ and suppose that MFCQ holds at x. Then the tangent cone of $\mathcal{C}$ at x is

$$T_{\mathcal{C}}(x) = \Big\{\xi \in \mathbb{R}^n \,\Big|\, \frac{\partial h(x)}{\partial x}\xi = 0, \frac{\partial g_{I_0}(x)}{\partial x}\xi \leq 0\Big\}.$$

For $x \in \mathcal{C}$, let $\Pi_{T_{\mathcal{C}}(x)}$ be the projection onto $T_{\mathcal{C}}(x)$. In general, the projection is a set-valued map, but the fact that $T_{\mathcal{C}}(x)$ is closed and convex makes the projection onto $T_{\mathcal{C}}(x)$ unique in this case. The projected gradient flow is

$$\begin{aligned} \dot{x} &= \Pi_{T_{\mathcal{C}}(x)}(-\nabla f(x)) \\ &= \underset{\xi \in \mathbb{R}^n}{\operatorname{argmin}} \quad \frac{1}{2}\|\xi + \nabla f(x)\|^2 \\ &\qquad \text{subject to} \quad \frac{\partial g_{I_0}(x)}{\partial x}\xi \leq 0, \frac{\partial h(x)}{\partial x}\xi = 0. \end{aligned} \tag{16}$$

In general, this system is discontinuous, so one must resort to appropriate notions of solution trajectories and establish their existence, see e.g., [16]. Here, we consider Carathéodory solutions, which are absolutely continuous functions that satisfy (16) almost everywhere. When Carathéodory solutions exist in $\mathcal{C}$, then the KKT points of (4) are equilibria of (16), and isolated local minimizers are asymptotically stable.

Consider the following continuous approximation of (16) by letting $\alpha > 0$ and defining $\mathcal{G}_\alpha$ by

$$\begin{aligned} \mathcal{G}_\alpha(x) = \underset{\xi \in \mathbb{R}^n}{\operatorname{argmin}} \quad & \frac{1}{2}\|\xi + \nabla f(x)\|^2 \\ \text{subject to} \quad & \frac{\partial g(x)}{\partial x}\xi \leq -\alpha g(x) \\ & \frac{\partial h(x)}{\partial x}\xi = -\alpha h(x). \end{aligned} \tag{17}$$

Note that (17) has a similar form to (16), and has a unique solution if one exists. However, as we show later, unlike the projected gradient flow, the vector field $\mathcal{G}_\alpha$ is well defined outside $\mathcal{C}$ and is Lipschitz.

We now show that $\mathcal{G}_\alpha$ approximates the projected gradient flow. Intuitively, this is because for inactive constraints $j \notin I_0(x)$, one has $g_j(x) < 0$ and hence the jth inequality constraint in (17), $\nabla g_j(x)^\top \xi \leq -\alpha g_j(x)$, becomes $\nabla g_j(x)^\top \xi \leq \infty$ as $\alpha \to \infty$ and the constraint is effectively removed, reducing the problem to (16). This is formalized next.

Proposition 4.4 ($\mathcal{G}_\alpha$ approximates the projected gradient): Let $x \in \mathcal{C}$ and suppose MFCQ holds. Then

(i) $\mathcal{G}_\alpha(x) \in T_{\mathcal{C}}(x)$.
(ii) $\lim_{\alpha\to\infty} \mathcal{G}_\alpha(x) = \Pi_{T_{\mathcal{C}}(x)}(-\nabla f(x))$.

Proof: To show (i), note that if $x \in \mathcal{C}$, then $h(x) = 0$ and $g_{I_0}(x) = 0$, so the constraints in (17) imply that $\frac{\partial h(x)}{\partial x}\mathcal{G}_\alpha(x) = 0$ and $\frac{\partial g_{I_0}(x)}{\partial x}\mathcal{G}_\alpha(x) \leq 0$, and therefore $\mathcal{G}_\alpha(x) \in T_{\mathcal{C}}(x)$.

Regarding (ii), for fixed $x \in \mathcal{C}$, let $J = I_-(x)$ and consider the following quadratic program

$$\begin{aligned} P_x(\epsilon) = \underset{\xi \in \mathbb{R}^n}{\operatorname{argmin}} \quad & \frac{1}{2}\|\xi + \nabla f(x)\|^2 \\ \text{subject to} \quad & \frac{\partial g_{I_0}(x)}{\partial x}\xi \leq 0, \frac{\partial h(x)}{\partial x}\xi = 0 \\ & \epsilon\frac{\partial g_J(x)}{\partial x}\xi \leq -g_J(x). \end{aligned} \tag{18}$$

When $\epsilon = 0$, the feasible sets of (18) and (16) are the same. Since the objective functions are also the same, $P_x(0) = \Pi_{T_{\mathcal{C}}(x)}(-\nabla f(x))$. Furthermore, for all $\alpha > 0$, $P_x(\frac{1}{\alpha}) = \mathcal{G}_\alpha(x)$. Finally, since the QP defining P_x has a unique solution, and satisfies the regularity conditions in

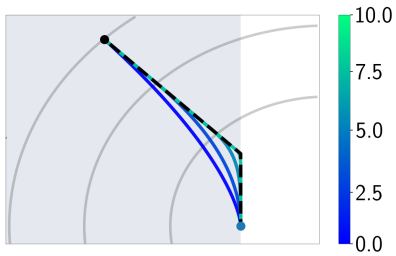


Fig. 2: Projected gradient flow versus continuous approximation. The solution of the projected gradient flow is in black and solutions of $\dot{x} = \mathcal{G}_\alpha(x)$ for varying values of α are in the colors corresponding to the colorbar. All solutions start from the same initial condition, marked by the black dot.

[33, Definition 2.1], P_x is continuous at $\epsilon = 0$ by [33, Theorem 2.2]. Hence $\lim_{\alpha\to\infty} \mathcal{G}_\alpha(x) = \lim_{\epsilon\to 0^+} P_x(\epsilon) = P_x(0) = \Pi_{T_{\mathcal{C}}(x)}(-\nabla f(x))$. ■

A consequence of Proposition 4.4 is that solutions of $\dot{x} = \mathcal{G}_\alpha(x)$ approximate the solutions of the projected gradient flow, with decreasing error as α increases, cf. Figure 2.

C. Equivalence Between the Two Interpretations

Here we establish the equivalence between the two interpretations of the safe gradient flow. Specifically, we show that the control barrier function quadratic program (14) can be interpreted as a dual program corresponding to the continuous approximation of the projected gradient flow in (17).

Let $L : \mathbb{R}^n \times \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}$ be

$$L(\xi, u, v; x) = \frac{1}{2}\|\xi + \nabla f(x)\|^2 + u^\top\Big(\frac{\partial g(x)}{\partial x}\xi + \alpha g(x)\Big) + v^\top\Big(\frac{\partial h(x)}{\partial x}\xi + \alpha h(x)\Big).$$

Then for each $x \in \mathbb{R}^n$, the Lagrangian of (17) is $(\xi, u, v) \mapsto L(\xi, u, v; x)$.

For each $x \in \mathbb{R}^n$, the KKT conditions corresponding to the optimization (17) are:

$$\xi + \nabla f(x) + \frac{\partial g(x)}{\partial x}^\top u + \frac{\partial h(x)}{\partial x}^\top v = 0 \tag{19a}$$

$$\frac{\partial g(x)}{\partial x}\xi + \alpha g(x) \leq 0 \tag{19b}$$

$$\frac{\partial h(x)}{\partial x}\xi + \alpha h(x) = 0 \tag{19c}$$

$$u \geq 0 \tag{19d}$$

$$u^\top\left(\frac{\partial g(x)}{\partial x}\xi + \alpha g(x)\right) = 0 \tag{19e}$$

Because the (17) is strongly convex, the existence of a triple (ξ, u, v) satisfying (19) is sufficient for optimality of ξ. Since the optimizer is unique, for any triple (ξ, u, v) satisfying these conditions, $\xi = \mathcal{G}_\alpha(x)$.

Let $\Lambda_\alpha : \mathbb{R}^n \rightrightarrows \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k$ be defined by

$$\Lambda_\alpha(x) = \{(u, v) \in \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k \,|\exists \xi \in \mathbb{R}^n \text{ such that } (\xi, u, v) \text{ solves (19)}\}. \tag{20}$$

By definition, $\Lambda_\alpha(x)$ is the set of Lagrange multipliers of (17) at $x \in \mathbb{R}^n$. When $\Lambda_\alpha(x) \neq \emptyset$, then the conditions (19) are also necessary for optimality of (17). As we show next, this necessity follows as a consequence of the constraint qualification conditions.

Lemma 4.5 (Necessity of optimality conditions): For $\alpha > 0$, if (4) satisfies MFCQ at $x \in \mathcal{C}$ (resp. EMFCQ at $x \in \mathbb{R}^n$), then there is an open set U containing x such that $\Lambda_\alpha(x') \neq \emptyset$ for all $x' \in U \cap \mathcal{C}$ (resp. $x' \in U$).

Proof: Let $I = I_0(x)$. If MFCQ holds at $x \in \mathcal{C}$, there exists $\xi \in \mathbb{R}^n$ such that $\nabla g_i(x)^\top \xi < 0$ for all $i \in I$ and $\nabla h_j(x)^\top \xi = 0$ for all $j \in \{1, \ldots, k\}$. Next, for every $j \in I_-(x)$, let $\epsilon_j > 0$ be defined as

$$\epsilon_j \leq \begin{cases} \frac{-\alpha g_j(x)}{\nabla g_j(x)^\top \xi} & \nabla g_j(x)^\top \xi > 0, \\ 1 & \nabla g_j(x)^\top \xi \leq 0. \end{cases}$$

Then taking $0 < \epsilon \leq \min_{j \in I_-(x)}\{\epsilon_j\}$ and $\tilde{\xi} = \epsilon\xi$, satisfies

$$\frac{\partial g(x)}{\partial x}\tilde{\xi} < -\alpha g(x) \qquad \frac{\partial h(x)}{\partial x}\tilde{\xi} = -\alpha h(x). \tag{21}$$

The above means that the constraints of (17) satisfy Slater's condition [34, Chapter 5.2.3] at x, so the affine constraints are *regular* [32, Theorem 2]. This implies that there exists an open set U containing x on which (17) is feasible and $\Lambda_\alpha(x') \neq \emptyset$ for all $x' \in U$. The proof for the case where EMFCQ holds at $x \in \mathbb{R}^n$ is identical to the above reasoning with $I = I_0(x) \cup I_+(x)$. ■

We use the optimality conditions to show that (14) is actually the dual problem corresponding to (17) in the appropriate sense.

Proposition 4.6: (Equivalence of two constructions of the safe gradient flow): If $\Lambda_\alpha(x) \neq \emptyset$,

(i) If $(u, v) \in \Lambda_\alpha(x)$, then (u, v) solves (14);
(ii) $\mathcal{G}_\alpha$ is the closed-loop dynamics corresponding to the implementation of (14) over (7).

Proof: To show (i), let $(u, v) \in \Lambda_\alpha(x)$. Then there is $\xi \in \mathbb{R}^n$ such that (ξ, u, v) solves (19). By (19a), $\xi = -\nabla f(x) - \frac{\partial g(x)}{\partial x}^\top u - \frac{\partial h(x)}{\partial x}^\top v$ and substituting ξ into the constraints of (17), it follows immediately that $(u, v) \in K_\alpha(x)$, defined in (8). We claim that (u, v) is also optimal for (14). To prove this, let (u', v') be a solution of (14) and, reasoning by contradiction, suppose

$$\left\|\frac{\partial g(x)}{\partial x}^\top u + \frac{\partial h(x)}{\partial x}^\top v\right\|^2 > \left\|\frac{\partial g(x)}{\partial x}^\top u' + \frac{\partial h(x)}{\partial x}^\top v'\right\|^2.$$

Then, $\xi' = -\nabla f(x) - \frac{\partial g(x)}{\partial x}^\top u' - \frac{\partial h(x)}{\partial x}^\top v'$ satisfies the constraints in (17) and $\|\xi' + \nabla f(x)\| < \|\xi + \nabla f(x)\|$, which contradicts the fact that ξ is optimal for (17).

To show (ii), suppose that (u, v) solves (14), and $\xi = -\nabla f(x) - \frac{\partial g(x)}{\partial x}^\top u - \frac{\partial h(x)}{\partial x}^\top v$. We claim that ξ is optimal for (17). Indeed, if $\tilde{\xi}$ is the optimizer of (17), since $\Lambda_\alpha(x) \neq \emptyset$, there exists $(\tilde{u}, \tilde{v}) \in \Lambda_\alpha(x)$ such that $(\tilde{\xi}, \tilde{u}, \tilde{v})$ solves (19). Note that $(\tilde{u}, \tilde{v})$ is feasible for (14), and

$$\|\xi + \nabla f(x)\|^2 = \left\|\frac{\partial g(x)}{\partial x}^\top u + \frac{\partial h(x)}{\partial x}^\top v\right\|^2 \leq \left\|\frac{\partial g(x)}{\partial x}^\top \tilde{u} + \frac{\partial h(x)}{\partial x}^\top \tilde{v}\right\|^2 = \|\tilde{\xi} + \nabla f(x)\|^2,$$

where the inequality follows by optimality of (u, v). It follows that ξ is optimal, but since the optimizer of (17) is unique,

$\xi = \mathcal{G}_\alpha(x)$. Hence, $\mathcal{G}_\alpha(x) = -\nabla f(x) - \frac{\partial g(x)}{\partial x}^\top u - \frac{\partial h(x)}{\partial x}^\top v$, which is the closed-loop implementation of (14) over (7). ■

Remark 4.7: (Lagrange Multipliers of Continuous Approximation to Projected Gradient): The notion of duality in Proposition 4.6 is weaker than the usual notion of Lagrangian duality. While the result ensures that the Lagrange multipliers of (17) are solutions to (14), the converse is not true in general. This is because if (u, v) solves (14), then $(\mathcal{G}_\alpha(x), u, v)$ might not satisfy the transversality condition (19e), in which case $(u, v) \notin \Lambda_\alpha(x)$. An example of this is given by the following constrained problem with objective f and inequality constraints $g(x) \leq 0$, where

$$f(x) = \|x\|^2 \qquad g(x) = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x - \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The constraints satisfy LICQ for all $x \in \mathbb{R}^n$. The solution is $x^* = 0$ and $\Lambda_\alpha(x^*) = \{(0, 0)\}$. However, $(1, 1)$ is an optimizer of (14), even though $(1, 1) \notin \Lambda_\alpha(x^*)$. •

Remark 4.8: (Lagrangian Dual of Continuous Approximation to Projected Gradient): The safe gradient flow can also be implemented using the Lagrangian dual of (17). This is obtained by replacing the feedback controller (14) with

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} \in \underset{(u,v)\in\mathbb{R}^m_{\geq 0}\times\mathbb{R}^k}{\operatorname{argmin}} \left\{ \frac{1}{2} \begin{bmatrix} u \\ v \end{bmatrix}^\top \begin{bmatrix} \frac{\partial g}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial g}{\partial x}\frac{\partial h}{\partial x}^\top \\ \frac{\partial h}{\partial x}\frac{\partial g}{\partial x}^\top & \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}^\top \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + u^\top \left(\frac{\partial g}{\partial x} \nabla f - \alpha g(x) \right) + v^\top \left(\frac{\partial h}{\partial x} \nabla f - \alpha h(x) \right) \right\}$$

and considering its closed-loop implementation over (7). Though this controller no longer has the same intuitive interpretation as the CBF-QP (14), it has the advantage that its values correspond exactly with $\Lambda_\alpha(x)$. •

Proposition 4.6 shows that there are two equivalent interpretations of the safe gradient flow. The first is as the closed-loop system corresponding to (7) with the controller (14), which maintains forward invariance of the feasible set $\mathcal{C}$ while ensuring the dynamics is as close as possible to the gradient flow of the objective function. The second interpretation is as an approximation of the projection of the gradient flow of the objective function onto the tangent cone of the feasible set. Both interpretations are related by the fact that the Lagrange multipliers corresponding to the approximate projection are the optimal control inputs solving (7). Beyond the interesting theoretical parallelism, this interpretation is instrumental in our ensuing discussion when characterizing the equilibria, regularity, and stability properties of the safe gradient flow.

V. Stability Analysis of the Safe Gradient Flow

Here we conduct a thorough analysis of the stability properties of the safe gradient flow and show that it solves Problem 1. We start by characterizing its equilibria and regularity properties, then focus on establishing the stability properties of local minimizers, and finally characterize the global convergence properties of the flow.

A. Equilibria, Regularity, and Safety

We rely on the necessary optimality conditions introduced in Section IV-C to characterize the equilibria of $\mathcal{G}_\alpha$.

Proposition 5.1: (Equilibria of safe gradient flow correspond to KKT points): If MFCQ holds at $x^* \in \mathcal{C}$, then

(i) $\mathcal{G}_\alpha(x^*) = 0$ if and only if $x^* \in X_{\text{KKT}}$;

(ii) If $x^* \in X_{\text{KKT}}$, then $\Lambda_\alpha(x^*)$ is the set of Lagrange multipliers of (4) at x^*.

Proof: Suppose that $\mathcal{G}_\alpha(x^*) = 0$. By Lemma 4.5, there exists $(u^*, v^*) \in \Lambda_\alpha(x^*)$ such that $(0, u^*, v^*)$ satisfies the necessary optimality conditions in (19), which reduce to

$$\nabla f(x^*) + \frac{\partial g(x^*)}{\partial x}^\top u^* + \frac{\partial h(x^*)}{\partial x}^\top v^* = 0 \tag{22a}$$

$$\alpha g(x^*) \leq 0 \tag{22b}$$

$$\alpha h(x^*) = 0 \tag{22c}$$

$$u^* \geq 0 \tag{22d}$$

$$(u^*)^\top (\alpha g(x^*)) = 0 \tag{22e}$$

Because $\alpha > 0$, it follows immediately that (22) implies that (x^*, u^*, v^*) satisfy (6) and $x^* \in X_{\text{KKT}}$.

Conversely, if $x^* \in X_{\text{KKT}}$, then for any (u^*, v^*) such that (x^*, u^*, v^*) solves (6), we have that $(0, u^*, v^*)$ solves (19), which implies that $\mathcal{G}_\alpha(x^*) = 0$ and $(u^*, v^*) \in \Lambda_\alpha(x^*)$. ■

Proposition 5.1(i) shows that the safe gradient flow meets Problem 1(iii). The correspondence in Proposition 5.1(ii) between the Lagrange multipliers of (17) and the Lagrange multipliers of (4) means that the proposed method can be interpreted as a primal-dual method when implemented via (17). This is because the state of the system (7) corresponds to the primal variable of (4), and the inputs to the system (7) correspond to the dual variables.

We next establish that $\mathcal{G}_\alpha$ is locally Lipschitz on an open set containing $\mathcal{C}$ when the MFCQ condition holds. This ensures the existence and uniqueness of classical solutions to the safe gradient flow.

Proposition 5.2 (Lipschitzness of safe gradient flow): Let $\alpha > 0$ and suppose that (4) satisfies MFCQ for all $x \in \mathcal{C}$, f, g and h are continuously differentiable, and their derivatives are locally Lipschitz. Then $\mathcal{G}_\alpha$ is well defined and locally Lipschitz on an open set X containing $\mathcal{C}$.

Proof: By the proof of Lemma 4.5, if MFCQ holds at $x \in \mathcal{C}$, there is an open neighborhood U_x containing x on which the constraints of (17) satisfy Slater's condition. Then, $\mathcal{G}_\alpha$ is the unique solution to (17) on U_x, and by [35, Theorem 3.10], $\mathcal{G}_\alpha$ is Lipschitz on U_x. The desired result follows by letting $X = \bigcup_{x\in\mathcal{C}} U_x$. ■

Proposition 5.2 verifies that the safe gradient flow meets Problem 1(i). Next, we show that under slightly stronger constraint qualification conditions at KKT points, the triple satisfying (19) is unique and Lipschitz near them.

Proposition 5.3 (Lipschitzness of the solution to (19)*):* Let $x^* \in X_{\text{KKT}}$ and suppose (4) satisfies LICQ at x^*. Then, there exists an open set U containing x^* and Lipschitz functions $u : U \to \mathbb{R}^m_{\geq 0}$, $v : U \to \mathbb{R}^m$ such that $(\mathcal{G}_\alpha(x), u(x), v(x))$ is the unique solution to (19) for all $x \in U$.

Proof: We claim that the variational equation (19) is *strongly regular* [36] for all $x^* \in X_{\text{KKT}}$. Strong regularity

implies, cf. [36, Corollary 2.1], that there exists an open set U containing x^* and Lipschitz functions $\xi : U \to \mathbb{R}^n$, $u : U \to \mathbb{R}^m_{\geq 0}$, $v : U \to \mathbb{R}^k$ such that $(\xi(x), u(x), v(x))$ is the unique triple solving (17). Since the solution (17) is unique, if such a triple exists, then $\xi(x) = \mathcal{G}_\alpha(x)$. To prove the claim, we begin by noting that (17) satisfies the strong second-order sufficient condition since $\nabla^2_{\xi\xi} L(\xi, u, v; x) = I \succ 0$. Let (x^*, u^*, v^*) be a KKT triple of (4). By Proposition 5.1, $(0, u^*, v^*)$ satisfies (19). Since the ith inequality constraint of (17) is $\nabla g_i(x^*)^\top \xi + \alpha g_i(x^*) \leq 0$, when $\xi = 0$ the constraint is active if and only if $g_i(x^*) = 0$. It follows that when $x^* \in X_{\text{KKT}}$, the indices of the active constraints of (17) are the same as those of (4). Moreover, for all $\xi \in \mathbb{R}^n$,

$$\frac{\partial}{\partial \xi}(\nabla g_i(x^*)^\top \xi + \alpha g_i(x^*)) = \nabla g_i(x^*)^\top,$$
$$\frac{\partial}{\partial \xi}(\nabla h_j(x^*)^\top \xi + \alpha h_j(x^*)) = \nabla h_j(x^*)^\top,$$

so the gradients of the binding (i.e., the active inequality and equality) constraints of (17) and (4) are also the same. By LICQ, the gradients of the binding constraints are linearly independent, which along with the strong second-order condition implies that (19) is strongly regular by [36, Theorem 4.1]. ■

The significance of Proposition 5.3 is twofold. First, it establishes that, under certain conditions, the Lagrange multipliers of (17) are Lipschitz as a function of x, which ensures the existence of a locally Lipschitz continuous feedback solving (14). Secondly, the result establishes conditions for uniqueness of the Lagrange multipliers in a neighborhood of an equilibrium x^*. These facts will play an important role in the stability analysis of local minimizers in the sequel.

We are now ready to show that the safe gradient flow also meets Problem 1(ii).

Theorem 5.4: (Safety of feasible set under safe gradient flow): Consider the optimization problem (4). If MFCQ is satisfied on $\mathcal{C}$, then $\mathcal{C}$ is forward invariant and asymptotically stable under the safe gradient flow.

The proof of this result follows from applying Lemma 2.1 with $\phi = (g, h)$ as a VCBF, cf. Lemma 4.1, and the construction of the feedback controller (14) yielding the closed-loop system locally Lipschitz, cf. Proposition 5.2.

Remark 5.5: (Advantages of safe gradient flow over projected gradient flow): Unlike the projected gradient flow, the vector field $\mathcal{G}_\alpha$ is locally Lipschitz, so classical solutions to $\dot{x} = \mathcal{G}_\alpha(x)$ exist, and the continuous-time flow can be numerically solved using standard ODE discretization schemes. Secondly, under mild conditions, $\mathcal{G}_\alpha$ is well defined for initial conditions outside $\mathcal{C}$, allowing us to guarantee convergence to a local minimizer starting from infeasible initial conditions. Finally, because both (16) and (17) are least-squares problems of the same dimension subject to affine constraints, the computational complexity of solving either one is equivalent. •

Remark 5.6: (Discretization of safe gradient flow and role of parameter α): When considering discretizations of the safe gradient flow, the parameter α plays an important role since its value impacts the behavior of its trajectories. By construction, trajectories of the safe gradient flow beginning at infeasible initial conditions converge to $\mathcal{C}$ at an exponential rate $\alpha > 0$, meaning that larger values of α ensure faster convergence toward the feasible set. On the other hand, smaller values of α result in a design that enforces safety more conservatively and hence, intuitively, this should allow for larger stepsizes. Our preliminary numerical experiments with the forward-Euler discretization $x^+ = x + h\mathcal{G}_\alpha(x)$ confirm these intuitions, showing that larger choices of α reduce the range of allowable stepsizes h that preserve the invariance of the feasible set $\mathcal{C}$ and stability of local minimizers. In particular, we have noticed that the maximal allowable stepsize h^*_α such that $0 < h < h^*_\alpha$ ensures stability and approximate safety, satisfies $h^*_\alpha \to 0$ as $\alpha \to \infty$. For space reasons, we leave to future work the formal characterization of suitable stepsizes. •

B. Stability of Isolated Local Minimizers

Here we characterize the stability properties of isolated local minimizers under the safe gradient flow. The following result shows that the safe gradient flow meets Problem 1(iv).

Theorem 5.7 (Stability of isolated local minimizers): Consider the optimization problem (4). Let x^* be a local minimizer and an isolated KKT point, and let U be an open set such that x^* is the only KKT point contained in U. Then,

(i) If MFCQ holds for all $x \in U \cap \mathcal{C}$, then x^* is asymptotically stable relative to $\mathcal{C}$;
(ii) If EMFCQ holds for all $x \in \bar{U}$, then x^* is asymptotically stable relative to $\mathbb{R}^n$;
(iii) If LICQ, strict complementarity, and the second-order sufficient condition hold at x^*, then x^* is exponentially stable relative to $\mathbb{R}^n$.

We divide the technical discussion leading up to the proof of the result in three parts, corresponding to each statement.

1) Stability of Isolated Local Minimizers Relative to $\mathcal{C}$: Here we analyze the stability of local minimizers relative to the feasible set. We start by characterizing the growth of the objective function along solutions of the safe gradient flow.

Lemma 5.8: (Growth of objective function along safe gradient flow): Let $x \in \mathbb{R}^n$ such that $\Lambda_\alpha(x) \neq \emptyset$. Then,

- For all $(u, v) \in \Lambda_\alpha(x)$,
$$D^+_{\mathcal{G}_\alpha} f(x) = -\|\mathcal{G}_\alpha(x)\|^2 + \alpha u^\top g(x) + \alpha v^\top h(x).$$
- If $x \in \mathcal{C}$ then,
$$D^+_{\mathcal{G}_\alpha} f(x) \leq 0,$$
with equality if and only if $x \in X_{\text{KKT}}$.

Proof: For $x \in X$ (with X given by Proposition 5.2) such that $(u, v) \in \Lambda_\alpha(x) \neq \emptyset$, $(\mathcal{G}_\alpha(x), u, v)$ solves (19). Next,

$$\begin{aligned} D^+_{\mathcal{G}_\alpha} f(x) &= \mathcal{G}_\alpha(x)^\top \nabla f(x) \\ &\overset{(a)}{=} -\mathcal{G}_\alpha(x)^\top \Big(\mathcal{G}_\alpha(x) + \frac{\partial g(x)}{\partial x}^\top u + \frac{\partial h(x)}{\partial x}^\top v\Big) \\ &\overset{(b)}{=} -\|\mathcal{G}_\alpha(x)\|^2 + \alpha u^\top g(x) + \alpha v^\top h(x), \end{aligned}$$

where (a) follows by rearranging (19a), and (b) follows from (19c) and (19e).

To show the second statement, note that if $x \in \mathcal{C}$, then $g(x) \leq 0$ and $h(x) = 0$. Since $u \geq 0$, it follows $\alpha u^\top g(x) + \alpha v^\top h(x) \leq 0$ and therefore

$$D^+_{\mathcal{G}_\alpha} f(x) \leq -\|\mathcal{G}_\alpha(x)\|^2 \leq 0.$$

Finally, $D^+_{\mathcal{G}_\alpha} f(x) = 0$ if and only if $\mathcal{G}_\alpha(x) = 0$, which by Proposition 5.1, is equivalent to $x \in X_{\text{KKT}}$. ■

As a consequence of Lemma 5.8, the objective function decreases monotonically along the solutions starting in $\mathcal{C}$. We use this fact to show that isolated local minimizers that are isolated equilibria are asymptotically stable relative to $\mathcal{C}$.

Proof: [Proof of Theorem 5.7(i)] By hypothesis and using Lemma 4.5, $\Lambda_\alpha(x) \neq \emptyset$ for all $x \in U \cap \mathcal{C}$. Because x^* is the unique strict minimizer of f on $U \cap \mathcal{C}$, and by Lemma 5.8, $D^+_{\mathcal{G}_\alpha} f(x) < 0$ for all $x \in U \cap \mathcal{C} \setminus \{x^*\}$, it follows by Lemma A.1 that x^* is asymptotically stable relative to $\mathcal{C}$. ■

2) Stability of Isolated Local Minimizers Relative to $\mathbb{R}^n$: Here we establish the asymptotic stability of isolated local minima relative to $\mathbb{R}^n$. To do so, we cannot rely any more on the objective function f as a Lyapunov function. This is because outside of $\mathcal{C}$, there may exist points $x \in \mathbb{R}^n \setminus \mathcal{C}$ where $f(x) < f(x^*)$. Therefore, to show stability relative to $\mathbb{R}^n$, we need to identify an alternative function whose unconstrained minimizer is x^*. In fact, the problem of finding a function whose unconstrained minimizers correspond to the local minimizers of a nonlinear program is well studied in the optimization literature [37]: such functions are called *exact penalty functions*. Our discussion proceeds by constructing an exact penalty function that is also a Lyapunov function for the safe gradient flow.

Let $\Omega \subset \mathbb{R}^n$ be a compact set. A function $V : \Omega \times (0, \infty) \to \mathbb{R}$ is a *strong exact penalty function* relative to Ω if there exists $\epsilon^* > 0$ such that for all $0 < \epsilon < \epsilon^*$, $x^* \in \text{int}(\Omega)$ is a local minimizer of $V_\epsilon(x) := V(x, \epsilon)$ if and only if x^* is a local minimizer of (4). The following result gives a strong exact penalty function for (4) whose upper-right Dini derivative along $\mathcal{G}_\alpha$ is well defined on Ω.

Lemma 5.9: (Existence of strong exact penalty function): Let $\Omega \subset \mathbb{R}^n$ be compact such that $\text{int}(\Omega) \cap \mathcal{C} \neq \emptyset$. Suppose (4) satisfies EMFCQ at every $x \in \Omega$ and let $V : \Omega \times (0, \infty) \to \mathbb{R}$,

$$V(x, \epsilon) = f(x) + \frac{1}{\epsilon} \sum_{i=1}^{m} [g_i(x)]_+ + \frac{1}{\epsilon} \sum_{j=1}^{k} |h_j(x)|. \tag{23}$$

Then, V is a strong exact penalty function relative to Ω, V is directionally differentiable on Ω, and

$$D^+_{\mathcal{G}_\alpha} V_\epsilon(x) = D^+_{\mathcal{G}_\alpha} f(x) + \frac{1}{\epsilon} \sum_{i \in I_+(x)} D^+_{\mathcal{G}_\alpha} g_i(x) + \frac{1}{\epsilon} \sum_{j=1}^{k} \text{sgn}(h_j(x)) D^+_{\mathcal{G}_\alpha} h_j(x), \tag{24}$$

for all $x \in \Omega$.

Proof: The fact that V is a strong exact penalty function relative to Ω readily follows from [37, Theorem 4]. From [37, Proposition 3], V_ϵ is directionally differentiable on Ω and its directional derivative in the direction $\xi \in \mathbb{R}^n$ is

$$\begin{aligned} V'_\epsilon(x; \xi) &= \nabla f(x)^\top \xi \\ &+ \frac{1}{\epsilon} \sum_{i \in I_+(x)} \nabla g_i(x)^\top \xi + \frac{1}{\epsilon} \sum_{i \in I_0(x)} [\nabla g_i(x)^\top \xi]_+ \\ &+ \frac{1}{\epsilon} \sum_{\substack{j \text{ such that} \\ h_j(x) \neq 0}} \text{sgn}(h_j(x)) \nabla h_j(x)^\top \xi + \frac{1}{\epsilon} \sum_{\substack{j \text{ such that} \\ h_j(x) = 0}} |\nabla h_j(x)^\top \xi|. \end{aligned}$$

We examine this expression in the case $V'_\epsilon(x; \mathcal{G}_\alpha(x)) = D^+_{\mathcal{G}_\alpha} V_\epsilon(x)$. For any $1 \leq i \leq m$, the definition of $\mathcal{G}_\alpha$ implies

$$\nabla g_i(x)^\top \mathcal{G}_\alpha(x) = D^+_{\mathcal{G}_\alpha} g_i(x) \leq -\alpha g_i(x).$$

Therefore, if $i \in I_0(x)$, then $[\nabla g_i(x)^\top \mathcal{G}_\alpha(x)]_+ = 0$. Similarly, for any $1 \leq j \leq k$, the definition of $\mathcal{G}_\alpha$ implies that

$$\nabla h_j(x)^\top \mathcal{G}_\alpha(x) = D^+_{\mathcal{G}_\alpha} h_j(x) = -\alpha h_j(x),$$

so if $h_j(x) = 0$, then $|\nabla h_j(x)^\top \mathcal{G}_\alpha(x)| = 0$, and the result follows. ■

We now show that V_ϵ is a Lyapunov function for ϵ sufficiently small and use this fact to certify the asymptotic stability of isolated local minimizers.

Proof: [Proof of Theorem 5.7(ii)] Assume, without loss of generality, that U is bounded. By Lemma 5.9, the function V_ϵ defined in (23) is a strong exact penalty relative to $\overline{U}$. By definition, this means that there exists $\epsilon_1 > 0$ such that when $\epsilon < \epsilon_1$, x^* is the only minimizer of V_ϵ in U. Let $x \in U$ and $(u, v) \in \Lambda_\alpha(x)$. Then, using Lemmas 5.8 and 5.9 and the definition of $\mathcal{G}_\alpha$, we have

$$\begin{aligned} D^+_{\mathcal{G}_\alpha} V_\epsilon(x) \leq &- \|\mathcal{G}_\alpha(x)\|^2 + \alpha u^\top g(x) + \alpha v^\top h(x) \\ &- \frac{1}{\epsilon} \sum_{i \in I_+(x)} \alpha g_i(x) - \frac{1}{\epsilon} \sum_{j=1}^{k} \alpha |h_j(x)|. \end{aligned}$$

Let $I = I_0(x) \cup I_-(x)$. It follows that,

$$\begin{aligned} D^+_{\mathcal{G}_\alpha} V_\epsilon(x) \leq &-\|\mathcal{G}_\alpha(x)\|^2 + \alpha \sum_{i \in I} u_i g_i(x) + \\ &\alpha \sum_{i \in I_+(x)} \Big(u_i - \frac{1}{\epsilon}\Big) g_i(x) + \alpha \sum_{j=1}^{k} \Big(|v_j| - \frac{1}{\epsilon}\Big) |h_j(x)|. \end{aligned}$$

Next, choose $0 < \epsilon_2 < \frac{1}{B}$ where $B > 0$ satisfies the bound given by Lemma B.1. Then, for $\epsilon < \epsilon_2$,

$$\sum_{i \in I_+(x)} \Big(u_i - \frac{1}{\epsilon}\Big) g_i(x) + \sum_{j=1}^{k} \Big(|v_j| - \frac{1}{\epsilon}\Big) |h_j(x)| < 0,$$

Finally, since $u \geq 0$, we have $\alpha \sum_{i \in I} u_i g_i(x) \leq 0$. Thus,

$$D^+_{\mathcal{G}_\alpha} V_\epsilon(x) \leq -\|\mathcal{G}_\alpha(x)\|^2 < 0,$$

for all $x \in U \setminus \{x^*\}$, whenever $\epsilon < \min\{\epsilon_1, \epsilon_2\}$. Therefore V_ϵ is a Lyapunov function on U and asymptotic stability of x^* follows by Lemma A.1. ■

Remark 5.10: (Relationship to merit functions in numerical optimization): In numerical optimization, the ℓ^1 penalty function in (23) is often used as a *merit function*, i.e., a function that quantifies how well a single iteration of an optimization algorithm balances the two goals of reducing the value of the objective function and reducing the constraint violation (cf. [38, Section 15.4]). Typically, the stepsize on each iteration is chosen so that the merit function is nonincreasing. Thus, if the algorithm is viewed as a discrete-time dynamical system, the merit function is a Lyapunov function. The ℓ^1 penalty plays a similar role for the continuous-time system described here. •

3) Exponential Stability of Isolated Local Minimizers: We now discuss the exponentially stability of isolated local minimizers. Our first step is to identify conditions under which the safe gradient flow is differentiable. To do so, we introduce the notions of strict complementarity and second-order condition on the optimization problem.

Definition 5.11: (Strict complementarity and second-order sufficient conditions): Let (x^*, u^*, v^*) be a KKT triple of (4).

- The strict complementarity condition holds if $u_i^* > 0$ for all $i \in I_0(x^*)$;
- The second-order sufficient condition holds if $z^\top Q z > 0$ for all $z \in \ker \frac{\partial g_{I_0}(x^*)}{\partial x} \cap \ker \frac{\partial h(x^*)}{\partial x}$, where

$$Q = \nabla^2 f(x^*) + \sum_{i=1}^{m} u_i^* \nabla^2 g_i(x^*) + \sum_{j=1}^{k} v_i^* \nabla^2 h_j(x^*). \quad (25)$$

When LICQ holds, strict complementarity together with the second-order sufficient condition can be used to establish the differentiability of a KKT triple of a nonlinear parametric program with respect to the parameters [39]. When these conditions are satisfied by (4), we show next that (17) is differentiable and provide an expression for its Jacobian.

Lemma 5.12 (Jacobian of safe gradient flow): Let $x^* \in X_{\text{KKT}}$ and u^*, v^* be the associated Lagrange multipliers. Suppose:

- LICQ holds at x^*;
- (x^*, u^*, v^*) satisfies the strict complementarity condition;
- (x^*, u^*, v^*) satisfies the second-order sufficient condition.

Then $\mathcal{G}_\alpha$ is differentiable at x^* and

$$\frac{\partial \mathcal{G}_\alpha(x^*)}{\partial x} = -PQ - \alpha(\mathcal{I} - P),$$

where $\mathcal{I}$ is the $n \times n$ identity matrix, P is the orthogonal projection matrix onto $\ker \frac{\partial g_{I_0}(x^*)}{\partial x} \cap \ker \frac{\partial h(x^*)}{\partial x}$, and Q is defined in (25).

Proof: By Proposition 5.3, there is a neighborhood U of x^* where the unique KKT triple of (17) corresponding to $x \in U$ is $(\mathcal{G}_\alpha(x), u(x), v(x))$. From the proof of that result, LICQ and the strong second-order sufficient condition hold for (17) at x^*. Further, the indices of the active constraints of (4) are the same as those of (17). Because (x^*, u^*, v^*) satisfies the strict complementarity condition for (4), $(0, u^*, v^*)$ satisfies the strict complementarity condition for (17). Thus, by [40, Corollary 1], the KKT triple is continuously differentiable with respect to x, and the Jacobian of $\mathcal{G}_\alpha$ can be computed using the given expression for the derivative. ■

Using the result in Lemma 5.12, stability of an isolated local minimizer can be inferred by showing that the eigenvalues of the Jacobian of the safe gradient flow are all strictly negative.

Proof: [Proof of Theorem 5.7(iii)] By the second-order sufficient condition, $z^\top PQPz > 0$ for all $z \in \operatorname{im} P \setminus \{0\}$. It follows that $PQPz = 0$ if and only if $z \in \ker P$. Therefore 0 is an eigenvalue of PQP with multiplicity r and PQP has $n - r$ strictly positive eigenvalues, where $r = \dim \ker P$. Let $z_1, \ldots, z_r$ be the eigenvectors corresponding to the zero eigenvalues, and $z_{r+1}, \ldots, z_n$ be eigenvectors corresponding to the positive eigenvalues, denoted $\lambda_{r+1}, \ldots, \lambda_n$. Then

$$Pz_i = \begin{cases} 0 & i = 1, \ldots, r, \\ z_i & i = r+1, \ldots, n. \end{cases}$$

Let

$$\mu_i = \begin{cases} 0 & i = 1, \ldots, r, \\ \lambda_i - \alpha & i = r+1, \ldots, n. \end{cases}$$

Then, it follows that $(PQP - \alpha P)z_i = \mu_i z_i$ for all $1 \le i \le n$. Observe that $PQP - \alpha P = (PQ - \alpha\mathcal{I})P$ has precisely the same eigenvalues as $P(PQ - \alpha\mathcal{I}) = PQ - \alpha P$. Therefore, since μ_i is an eigenvalue of $PQ - \alpha P$, it follows that $\mu_i + \alpha$ is an eigenvalue of

$$PQ - \alpha P + \alpha\mathcal{I} = PQ + \alpha(\mathcal{I} - P) = -\frac{\partial \mathcal{G}_\alpha(x^*)}{\partial x}.$$

Hence the eigenvalues of $\frac{\partial \mathcal{G}_\alpha(x^*)}{\partial x}$ are

$$\{-\alpha, -\alpha, \ldots, -\alpha, -\lambda_{r+1}, \ldots, -\lambda_n\},$$

where $-\alpha$ appears with multiplicity r. Since all the eigenvalues are strictly negative, x^* is exponentially stable. ■

C. Stability of Nonisolated Local Minimizers

We have characterized in Section V-B the stability under the safe gradient flow of local minimizers that are isolated KKT points. In general, if x^* is strict local minimizer that is not an isolated KKT point (for example, if there are an infinite number of local maximizers arbitrarily close to x^*, cf. [41, page 5]), or if x^* is only a local minimizer, then there are no guarantees on Lyapunov stability. However, as we show here, nonisolated minimizers are stable under the safe gradient flow under additional assumptions on the problem data.

When there are no constraints, the safe gradient flow reduces to the classical gradient flow, where conditions for semistability of local minimizers are well known: if the objective function is real-analytic, then all trajectories of the gradient flow have finite arclength, cf. [42], in which case the objective function can be used to construct an arclength-based Lyapunov function satisfying the hypotheses of Lemma A.2 to establish semistability. In this section, we conduct a similar analysis for the constrained case. Our main result is as follows.

Theorem 5.13: (Stability of nonisolated local minima): Consider the optimization problem (4), and assume f, g and h are real-analytic. Let $\mathcal{S}$ be a bounded set of local minimizers on which f is constant and equal to f^* such that

(i) There is an open set U and $\beta > 0$ such that $U \cap X_{\text{KKT}} = \mathcal{S}$ and $f(x) - f^* \ge \beta \operatorname{dist}_{\mathcal{S}}(x)^2$ for all $x \in U \cap \mathcal{C}$;
(ii) LICQ is satisfied at all $x^* \in \mathcal{S}$;
(iii) $T_{\mathcal{S}}(x^*) \cap \mathcal{N}_{\mathcal{S}}^{\text{prox}}(x^*) = \{0\}$ for all $x^* \in \mathcal{S}$.

Then there is $\alpha^* > 0$ such that every $x^* \in \mathcal{S}$ is semistable relative to $\mathbb{R}^n$ under the safe gradient flow $\mathcal{G}_\alpha$, for $\alpha > \alpha^*$.

To prove this result, we first discuss various intermediate results. In particular, the growth condition in Theorem 5.13(i) plays a crucial role in the construction of a Lyapunov function to prove the result. Any $x^* \in \mathcal{S}$ satisfying this property is called a *weak sharp minimizer of f relative to $\mathcal{S}$*. Weak sharp minimizers play an important role in sensitivity analysis for nonlinear programs as well as convergence analysis for numerical methods in optimization [43], [44].

We review second-order optimality conditions for weak sharp minimizers. Let $x^* \in X_{\text{KKT}}$, suppose that LICQ holds

at x^*, and let (u^*, v^*) be the unique Lagrange multipliers of (4) associated to x^*. Define the index set of *strongly active* constraints as

$$I_0^+(x^*) = \{1 \le i \le m \,|\, u_i^* > 0\}.$$

The *critical cone* is

$$\begin{aligned}\Gamma(x^*) = \{d \in \mathbb{R}^n \mid \nabla h_j(x^*)^\top d = 0, j = 1, \dots k,\\ \nabla g_i(x^*)^\top d = 0, i \in I_0^+(x^*),\\ \nabla g_j(x^*)^\top d \le 0, j \in I_0(x^*) \setminus I_0^+(x)\}.\end{aligned} \tag{26}$$

Lemma 5.14: (Second-order necessary condition for constrained weak sharp minima [44, Proposition 3.5]): Consider (4) and let $\mathcal{S} \subset \mathcal{C}$ be a set on which f is constant. Suppose that $x^* \in \partial\mathcal{S}$ is a weak sharp local minimizer of f relative to $\mathcal{S}$ and LICQ is satisfied at x^*. Let u^*, v^* be the Lagrange multipliers and define $\ell(x) = f(x) + (u^*)^\top g(x) + (v^*)^\top h(x)$. Then, there exists $\gamma > 0$ such that, for all $d \in \Gamma(x^*)$,

$$\ell''(x^*; d) \ge \gamma \mathrm{dist}_{T_{\mathcal{S}}(x^*)}(d)^2.$$

Lemma 5.15: (Second-order sufficient condition for unconstrained weak sharp minima [44, Theorem 2.5]): Consider $W : \mathbb{R}^n \to \mathbb{R}$ and suppose that W is constant on $\mathcal{S}$. Suppose $x^* \in \partial\mathcal{S}$ and $W''(x^*; d) > 0$ for all $d \in \mathcal{N}_{\mathcal{S}}^{\text{prox}}(x^*) \setminus \{0\}$, then x^* is a weak sharp local minimizer of W relative to $\mathcal{S}$.

We now proceed with the construction of the Lyapunov function. Let $T_{\mathcal{C}}^{(\alpha)} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be the set-valued map where, for each $x \in \mathbb{R}^n$, $T_{\mathcal{C}}^{(\alpha)}(x)$ is the constraint set of (17). Let $J_\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be

$$J_\alpha(x, \xi) = \alpha f(x) + \nabla f(x)^\top \xi + \frac{1}{2}\|\xi\|^2.$$

Consider the optimization problem

$$\underset{\xi \in T_{\mathcal{C}}^{(\alpha)}(x)}{\text{minimize}} \quad J_\alpha(x, \xi) \tag{27}$$

As we show next, the solution to (27) is (17).

Lemma 5.16: (Correspondence between (27) *and* (17)*):* Let $x \in \mathbb{R}^n$. Then the program (17) has a solution at x if and only if (27) has a solution, in which case $\mathcal{G}_\alpha(x) = \arg\min_{\xi \in T_{\mathcal{C}}^{(\alpha)}(x)} \{J_\alpha(x, \xi)\}$.

Proof: Note that the feasible sets of (27) and (17) coincide. Next, for all $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$:

$$J_\alpha(x, \xi) - \frac{1}{2}\|\xi + \nabla f(x)\|^2 = \alpha f(x) - \frac{1}{2}\|\nabla f(x)\|^2.$$

Since the difference of the objectives in (27) and (17) does not depend on ξ, both problems have the same optimizer. ■

Lemma 5.16 shows that (27) is another characterization of the safe gradient flow in terms of a parametric quadratic program. Let $W_\alpha : X \to \mathbb{R}$ be the value function:

$$\begin{aligned}W_\alpha(x) &= \inf_{\xi \in T_{\mathcal{C}}^{(\alpha)}(x)} \{J_\alpha(x, \xi)\}\\ &= \alpha f(x) + \nabla f(x)^\top \mathcal{G}_\alpha(x) + \frac{1}{2}\|\mathcal{G}_\alpha(x)\|^2.\end{aligned} \tag{28}$$

Our strategy to prove Theorem 5.13 consists of showing that W_α is a Lyapunov function satisfying the hypotheses in Lemma A.3 whenever α is sufficiently large. Towards this end, we begin by computing the directional derivative of W_α. Let $Q : X \times \mathbb{R}^m_{\ge 0} \times \mathbb{R}^k \to \mathbb{R}^{n \times n}$ be the matrix-valued function,

$$Q(x, u, v) = \nabla^2 f(x) + \sum_{i=1}^m u_i \nabla^2 g_i(x) + \sum_{j=1}^k v_j \nabla^2 h_j(x).$$

Since the Lagrange multipliers, $(u(x), v(x))$ are unique in a neighborhood of $\mathcal{S}$, we slightly abuse notation by defining $Q(x) := Q(x, u(x), v(x))$. By Lipschitzness of u and v, Q is continuous on X. The proof of the next result follows from [40, Theorem 2] and [45, Corollary 4.1] and is omitted for brevity.

Lemma 5.17 (Differentiability of W_α): Suppose that $\mathcal{S}$ satisfies the hypotheses in Theorem 5.13, and X is an open set containing $\mathcal{S}$ on which $(\mathcal{G}_\alpha(x), u(x), v(x))$ is the unique solution to (19). Then

(i) For all $x \in X$, W_α is differentiable with

$$\nabla W_\alpha(x) = -(\alpha I - Q(x))\mathcal{G}_\alpha(x); \tag{29}$$

(ii) For all $x^* \in \mathcal{S}$, W_α is twice directionally differentiable in any direction $d \in \mathbb{R}^n$, where

$$\begin{aligned}W''_\alpha(x^*; d) = \min_{\zeta \in \mathbb{R}^n} \quad & \begin{bmatrix} d \\ \zeta \end{bmatrix}^\top \begin{bmatrix} \alpha Q(x^*) & Q(x^*) \\ Q(x^*) & I \end{bmatrix} \begin{bmatrix} d \\ \zeta \end{bmatrix}\\ \text{s.t.} \quad & \alpha \nabla h_j(x^*)^\top d + \nabla h_j(x^*)^\top \zeta = 0,\\ & \quad \forall j = 1, \dots, k,\\ & \alpha \nabla g_i(x^*)^\top d + \nabla g_i(x^*)^\top \zeta = 0,\\ & \quad \forall i \in I_0^+(x^*),\\ & \alpha \nabla g_s(x^*)^\top d + \nabla g_s(x^*)^\top \zeta \le 0,\\ & \quad \forall s \in I_0(x^*) \setminus I_0^+(x^*).\end{aligned} \tag{30}$$

Remark 5.18: (Dependence of $Q(x)$ on α): In general, for $x \in X$, the value of $Q(x)$ depends on the choice of α, since $u(x)$ and $v(x)$ depend on α. However, if $x^* \in X_{\text{KKT}}$, then $u(x^*), v(x^*)$ correspond to the Lagrange multipliers of (4) and $Q(x^*)$ is the Hessian of the Lagrangian of (4). In particular, this means that for all $x^* \in X_{\text{KKT}}$, the value of $Q(x^*)$ depends only on the problem data and is independent of α. •

We now proceed with the proof of Theorem 5.13.

Proof: [Proof of Theorem 5.13] Let $\alpha^* = \sup_{x^* \in \mathcal{S}}\{\rho(Q(x^*))\}$. For $\alpha > \alpha^*$, we have $\alpha I - Q(x^*) \succ 0$ for all $x^* \in \mathcal{S}$. Assume without loss of generality that $\alpha I - Q(x) \succ 0$ for all $x \in U$ (if not, since Q is continuous, we can always find an open subset of U containing $\mathcal{S}$ for which these property holds). We claim that W_α satisfies each of the conditions (i)-(iii) in Lemma A.3 with $\mathcal{K} = \mathbb{R}^n$.

We begin by showing condition (iii). If $x^* \in U$ is a local minimizer of W_α, then $\nabla W_\alpha(x^*) = (\alpha I - Q(x^*))\mathcal{G}_\alpha(x^*) = 0$. Since $\alpha I - Q(x^*) \succ 0$, from (29) we deduce $\mathcal{G}_\alpha(x^*) = 0$, so $x^* \in X_{\text{KKT}}$ and therefore $x^* \in U \cap X_{\text{KKT}} = \mathcal{S}$.

Conversely, suppose that $x^* \in \mathcal{S}$. Note that, by Proposition 5.1, $W_\alpha(x) = \alpha f(x)$ for all $x \in \mathcal{S}$. Therefore, if $x^* \in \text{int}(\mathcal{S})$, it follows that x^* is a local minimizer of W_α. Suppose instead that $x^* \in \partial\mathcal{S}$. For $d \in \mathbb{R}^n$, let ζ_d be the unique optimizer of (30). Then

$$W''_\alpha(x^*; d) = \alpha d^\top Q(x^*) d + 2\zeta_d^\top Q(x^*) d + \|\zeta_d\|^2. \tag{31}$$

From the constraints in (30), $\zeta_d + \alpha d \in \Gamma(x^*)$. Because $x^* \in \partial\mathcal{S}$ is a weak sharp minimizer of f relative to $\mathcal{S}$, by

Lemma 5.14, there exists $\gamma > 0$ such that

$$\begin{aligned}\ell''(x^*;\zeta_d+\alpha d) &= (\zeta_d+\alpha d)^\top \nabla^2\ell(x^*)(\zeta_d+\alpha d),\\ &\geq \gamma \mathrm{dist}_{T_{\mathcal{S}}(x^*)}(\zeta_d+\alpha d)^2,\ \forall d\in\mathbb{R}^n.\end{aligned} \tag{32}$$

Since $\nabla^2\ell(x^*) = Q(x^*)$, we combine (31) and (32) to get

$$\alpha W''_\alpha(x^*;d) \geq \zeta_d^\top(\alpha I - Q(x^*))\zeta_d + \gamma \mathrm{dist}_{T_{\mathcal{S}}(x^*)}(\zeta_d+\alpha d)^2.$$

Because $\alpha I - Q(x^*) \succ 0$, if $W''_\alpha(x^*;d) = 0$, then $\zeta_d = 0$ and $d \in T_S(x^*)$. But $T_{\mathcal{S}}(x^*) \cap \mathcal{N}_{\mathcal{S}}^{\text{prox}}(x^*) = \{0\}$, which means $W''_\alpha(x^*;d) > 0$ for all $d \in \mathcal{N}_{\mathcal{S}}^{\text{prox}}(x^*)\backslash\{0\}$, so by Lemma 5.15, x^* is a weak sharp local minimizer of W_α.

Next we verify condition (ii) in Lemma A.3. For all $x \in U$,

$$\begin{aligned}D^+_{\mathcal{G}_\alpha} W_\alpha(x) &= \nabla W_\alpha(x)^\top \mathcal{G}_\alpha(x)\\ &= -\mathcal{G}_\alpha(x)^\top(\alpha I - Q(x))\mathcal{G}_\alpha(x).\end{aligned}$$

Without loss of generality, we can assume that U is bounded. Then, we can choose $c_1, c_2 > 0$ so that

$$\begin{aligned}c_1 &< \inf_{x\in U}\{\lambda_{\min}(\alpha I - Q(x))\}\\ c_2 &> \sup_{x\in U}\{\lambda_{\max}(\alpha I - Q(x))\}.\end{aligned}$$

It follows that $D^+_{\mathcal{G}_\alpha} W_\alpha(x) \leq -c_1\|\mathcal{G}_\alpha(x)\|^2$ for all $x \in U$, but since $\|\nabla W_\alpha(x)\| \leq c_2\|\mathcal{G}_\alpha(x)\|$, we have for all $x \in U$,

$$D^+_{\mathcal{G}_\alpha} W_\alpha(x) \leq -\frac{c_1}{c_2}\|\nabla W_\alpha(x)\|\|\mathcal{G}_\alpha(x)\|.$$

Finally, we claim that $W_\alpha|_U$ is a globally subanalytic function, and therefore condition (i) holds by [46, Theorem 1], and the fact that the class of globally subanalytic sets is an o-minimal structure (cf. [46, Definition 1]). To prove the claim, first note that, since f is real-analytic, J_α is real-analytic, and therefore subanalytic [47, Definition 3.1]. Since U is bounded, and the restriction of any subanalytic function to a bounded open set is globally subanalytic [48], it follows that $J_\alpha|_U$ is globally subanalytic. Finally, since $T_{\mathcal{C}}^{(\alpha)}|_U : U \rightrightarrows \mathbb{R}^n$ is a globally subanalytic set valued map, and

$$W_\alpha|_U(x) = \inf_{\xi\in T_{\mathcal{C}}^{(\alpha)}|_U(x)}\{J_\alpha|_U(x,\xi)\},$$

it follows that $W_\alpha|_U$ is globally subanalytic. The statement follows by applying Lemma A.3 with $\mathcal{K} = \mathbb{R}^n$. ■

D. Global Convergence

Finally, we turn to the characterization of the global convergence properties of the safe gradient flow. We show that when the problem data are real-analytic and the feasible set is bounded, every trajectory converges to a KKT point.

Theorem 5.19: (Global convergence properties): Consider the optimization problem (4), and assume $\mathcal{C}$ is bounded, f, g, and h are real-analytic functions, and LICQ holds everywhere on $\mathcal{C}$. Let X be an open set containing $\mathcal{C}$ on which the safe gradient flow is well defined. Then there is $\alpha^* > 0$ such that for $\alpha > \alpha^*$, every trajectory of the safe gradient flow starting in X converges to some KKT point.

To prove Theorem 5.19, we use the next result characterizing the positive limit set of solutions of the safe gradient flow.

Lemma 5.20: (Convergence to connected component): Consider the optimization problem (4), and assume $\mathcal{C}$ is bounded, f, g, and h are real-analytic functions, and MFCQ holds everywhere on $\mathcal{C}$. Let X be an open set containing $\mathcal{C}$ on which the safe gradient flow is well defined. Then for all $x \in X$, $\omega(x)$ is contained in a unique connected component of X_{KKT}.

Proof: By Theorem 5.4, $\mathcal{C}$ is asymptotically stable and forward invariant on X, and by Lemma 5.8, $D^+_{\mathcal{G}_\alpha} f(x) \leq 0$ for all $x \in \mathcal{C}$. Using the terminology from [49], f is a *height function* of the pair $(\mathcal{C}, \mathcal{G}_\alpha)$.

Because f, g, and h are real-analytic and $\mathcal{C}$ is bounded, $\mathcal{C}$ is a globally subanalytic set. Let $\hat{f} = f + \delta_{\mathcal{C}}$. Then $\hat{f}$ is a globally subanalytic function, $\hat{f}$ is continuous on $\mathrm{dom}(\hat{f}) = \mathcal{C}$, and X_{KKT} is precisely the set of critical points of $\hat{f}$. By the Morse-Sard Theorem for subanalytic functions [50, Theorem 14], X_{KKT} has at most a countable number of connected components, and $\hat{f}$ is constant on each connected component. Since $f(x) = \hat{f}(x)$ for all $x \in \mathcal{C}$, f is also constant on each connected component of X_{KKT}, meaning that the connected components of X_{KKT} are *contained in* f (cf. [49, Definition 5]).

Hence, we can apply [49, Theorem 6], and conclude that for all $x \in X$, the positive limit set $\omega(x)$ is nonempty and contained in a unique connected component of $E = \{x \in \mathcal{C} \mid D^+_{\mathcal{G}_\alpha} f(x) = 0\}$. However, by Lemma 5.8, $E = X_{\text{KKT}}$, concluding the result. ■

We are ready to prove Theorem 5.19.

Proof: [Proof of Theorem 5.19] By Lemma 5.20, for $x \in X$, there is a connected component $\mathcal{S} \subset X_{\text{KKT}}$ such that $\omega(x) \subset \mathcal{S}$. Since LICQ holds on $\mathcal{S}$, by Proposition 5.3 there is an open set U containing $\mathcal{S}$ and Lipschitz functions $(u, v) : U \to \mathbb{R}^m_{\geq 0} \times \mathbb{R}^k$ such that $U \cap X_{\text{KKT}} = \mathcal{S}$ and $(\mathcal{G}_\alpha(x), u(x), v(x))$ is the unique solution to (19) on U.

Let W_α be given by (28). By Lemma 5.17, W_α is differentiable on U, and using the same reasoning as in the proof of Theorem 5.13, W_α is a globally subanalytic function, and satisfies the Kurdyka-Łojasiewicz inequality. Furthermore, if $\alpha > \alpha^* = \sup_{x^*\in\mathcal{S}}\{\rho(Q(x^*))\}$, then there is some $c > 0$ such that $D^+_{\mathcal{G}_\alpha} W_\alpha(y) \leq -c\|\nabla W_\alpha(y)\|\|\mathcal{G}_\alpha(y)\|$ for all $y \in U$.

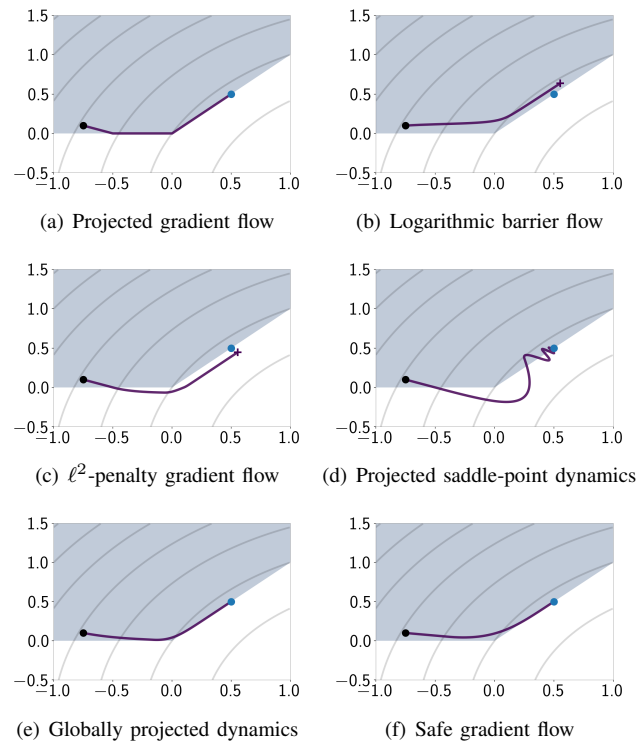
Thus, we can apply Lemma A.3 with $\mathcal{K} = \mathbb{R}^n$ to conclude that there exists a neighborhood $\tilde{U}$ containing $\mathcal{S}$ such that every trajectory starting in $\tilde{U}$ that remains in $\tilde{U}$ for all time converges to a point in $\mathcal{S}$. Finally, since $\omega(x) \subset \mathcal{S}$, there exists $T > 0$ such that $\Phi_T(x) \in \tilde{U}$ and $\Phi_t(\Phi_T(x)) = \Phi_{T+t}(x) \in \tilde{U}$ for all $t > 0$. Thus, there exists $x^* \in \mathcal{S}$ such that $\Phi_{T+t}(x) \to x^*$ as $t \to \infty$, and the trajectory starting at x converges to x^*. ■

Remark 5.21: (Lower bounds on the parameter α to ensure global convergence): Note that the proof of Theorem 5.19 yields the expression $\alpha^* = \sup_{x^*\in\mathcal{S}}\{\rho(Q(x^*))\}$ for the lower bound on α that guarantees global convergence. In general, computing this expression requires knowledge of the primal and dual optimizers of the original problem. However, reasonable assumptions on f, g, and h allow us to obtain upper bounds of α^*. For instance, if $\mathcal{C}$ is polyhedral and ∇f is ℓ_f-Lipschitz on $\mathcal{C}$, it follows that $\|\nabla^2 f(x)\| \leq \ell_f$, and $\nabla^2 g_i(x) = 0$ and $\nabla^2 h_j(x) = 0$ for all $i = 1, \ldots m$ and $j = 1, \ldots k$. Therefore, $\alpha^* \leq \ell_f$, and ℓ_f can be used instead as a lower bound on α to ensure global convergence. •

VI. Comparison With Other Optimization Methods

Here we compare the safe gradient flow with other continuous-time flows to solve optimization problems. For simplicity, we restrict our attention to an inequality constrained convex program. Figure 3 shows the outcome of the comparison on the same example problem taken from [5]. The methods compared are the projected gradient flow, the logarithmic barrier method (see e.g. [51, Section 3]), the ℓ^2-penalty gradient flow (see e.g. [52, Chapter 4]), the projected saddle-point dynamics (see e.g., [24]), the globally projected dynamics (see e.g., [18]), and the safe gradient flow.

Under the logarithmic barrier method, the feasible set is forward invariant and the minimizer of the logarithmic barrier penalty $f_{\text{barrier}}(x;\mu) = f(x) - \mu\sum_{i=1}^{m}\log(-g_i(x))$, with $\mu > 0$, does not correspond to the minimizer of (4). Under the unconstrained minimizer of the ℓ^2-penalty, $f_{\text{penalty}}(x;\epsilon) = f(x) + \frac{\epsilon}{2}\sum_{i=1}^{m}[g_i(x)]_+^2$, with $\epsilon > 0$, does not correspond to the minimizer of (4), and the feasible set is not forward invariant under the gradient flow of f_{penalty}. Under the projected saddle-point dynamics, the feasible set is not forward invariant, but each trajectory converges to x^*. Under the globally projected dynamics, the feasible set is forward invariant, trajectories converge to x^*, and trajectories are smooth. However, unlike the safe gradient flow, the globally projected dynamics is undefined when the constraints are not convex.



(a) Projected gradient flow (b) Logarithmic barrier flow

(c) ℓ^2-penalty gradient flow (d) Projected saddle-point dynamics

(e) Globally projected dynamics (f) Safe gradient flow

Fig. 3: Comparison of methods minimizing $f(x) = 0.25\|x\|^2 - 0.5x_1 + 0.25x_2$ subject to $x_2 \geq 0$ and $x_1 \leq x_2$ (see also [5, Figure 8] for a comparison of additional methods). The blue-shaded region is the feasible set and the grey curves are level sets of the objective function. The initial condition is denoted by the purple dot, and the global minimizer is denoted by a blue dot. (a) The trajectory converges to the global minimizer, and the trajectory remains inside the feasible set for all time but it is nonsmooth. (b) The trajectory is smooth and remains inside the feasible set but does not converge to the global minimizer. However, by choosing μ small enough, the trajectory can be made to converge arbitrarily close to the minimizer. (c) The trajectory is smooth, but does not remain inside the feasible set or converge to the global minimizer. However, by choosing ϵ small enough, the trajectory can be made to converge arbitrarily close to the minimizer. (d) Initialized with $u(0) = 0$, the trajectory does not remain inside the feasible set, but it converges to the global minimum. (e) The trajectory is smooth, converges to the global minimizer, and remains inside the feasible set. However, this method may not be well-defined for nonconvex problems (f) The trajectory is smooth, converges to the global minimizer, and remains inside the feasible set. Of the methods implemented here and in [5, Figure 8], the safe gradient flow is the only nonconvex method that satisfies all of these properties.

VII. Conclusions

We have introduced the safe gradient flow, a continuous-time dynamical system to solve constrained optimization problems that makes the feasible set forward invariant. The system can be derived either as a continuous approximation of the projected gradient flow or by augmenting the gradient flow of the objective function with inputs, then using a control barrier function-based QP to ensure safety of the feasible set. The equilibria are exactly the critical points of the optimization problem, and the steady-state inputs at the equilibria correspond to the dual optimizers of the program. We have conducted a thorough stability analysis of the dynamics, identified conditions under which isolated local minimizers are asymptotically stable and nonisolated local minimizers are semistable. Future work will generalize the safe gradient flow to nonsmooth problems and distributed problems, explore its robustness properties, and leverage convexity to obtain stronger global convergence guarantees. Further, we hope to explore issues related to the practical implementation of the safe gradient flow, including interconnections of the optimizing dynamics with a physical system, where the optimization problem is in a feedback loop with a dynamically evolving plant, develop discretizations of the dynamics and study their relationship with discrete-time iterative methods for nonlinear programming, and extend the framework to Newton-like flows for nonlinear programs which incorporate higher-order information.

Appendix A
Lyapunov Tests for Stability

Here we present Lyapunov based tests for stability of an equilibrium. The first result is a special case of [53, Corollary 7.1], and establishes the stability of an isolated equilibrium.

Lemma A.1 (Lyapunov test for relative stability): Let $\mathcal{K}$ be a forward invariant set of $\dot{x} = F(x)$ and x^* an isolated equilibrium. Let $U \subset \mathbb{R}^n$ be an open set containing x^* and suppose that $V : U \cap \mathcal{K} \to \mathbb{R}$ is a directionally differentiable function such that

- x^* is the unique minimizer of V on $U \cap \mathcal{K}$.
- $D_F^+ V(x) < 0$ for all $x \in U \cap \mathcal{K} \setminus \{x^*\}$.

Then x^* is asymptotically stable relative to $\mathcal{K}$.

The next results provides a test for attractivity and stability of a set of nonisolated equilibria, using an "arclength"-based Lyapunov test [54, Theorem 4.3 and Theorem 5.2].

Lemma A.2 (Arclength-based Lyapunov test): Let $\mathcal{K}$ be a forward invariant set of $\dot{x} = F(x)$. Let $\mathcal{S} \subset \mathcal{K}$ be a set of equilibria and $U \subset \mathbb{R}^n$ an open set containing $\mathcal{S}$ where $U \cap F^{-1}(\{0\}) = \mathcal{S}$. Let $V : U \cap \mathcal{K} \to \mathbb{R}$ be a continuously differentiable function such that

(i) There exists a $c > 0$ such that for all $x \in U \cap \mathcal{K}$,

$$D_F^+ V(x) \leq -c\|F(x)\|. \tag{33}$$

Then every bounded trajectory that starts in $U \cap \mathcal{K}$ and remains in $U \cap \mathcal{K}$ for all time has finite arclength and converges to a point in $\mathcal{S}$. If in addition,

(ii) x^* is a minimizer of V if and only if $x^* \in \mathcal{S}$.

then every $x^* \in \mathcal{S}$ is semistable relative to $\mathcal{K}$.

Kurdyka-Łojasiewicz inequality[2] (condition (i) in Lemma A.3 below), we show that the condition in (33) for the arclength-based Lyapunov test can be replaced with $D_F^+ V(x) \leq -c\|F(x)\|\|\nabla V(x)\|$. This is referred to as the "angle-condition" and has been exploited [56], [57] to show convergence of descent methods to solve nonlinear programming problems. The name arises from the fact that the inequality implies that the angle between $F(x)$ and $\nabla V(x)$ remains bounded in a neighborhood of the equilibrium. In the next result, we show that the angle condition, together with the Kurdyaka-Łojasiewicz inequality, implies that every trajectories of the system have finite arclength.

Lemma A.3: (Angle-condition-based Lyapunov test):. Let $\mathcal{K}$ be a forward invariant set of $\dot{x} = F(x)$. Let $\mathcal{S} \subset \mathcal{K}$ be a bounded set of equilibria and $U \subset \mathbb{R}^n$ a bounded open set containing $\mathcal{S}$ where $U \cap F^{-1}(\{0\}) = \mathcal{S}$. Let $V : U \cap \mathcal{K} \to \mathbb{R}$ be a differentiable function such that

(i) There exists $V^* \in \mathbb{R}$ such that $V(x) = V^*$ if and only if $x \in \mathcal{S}$, and there exists $c_1 > 0$ and a strictly increasing, differentiable function $\psi : [0, \infty) \to \mathbb{R}$ such that for all $x \in (U \cap \mathcal{K}) \setminus \mathcal{S}$:

$$\psi'(|V(x) - V^*|)\|\nabla V(x)\| \geq c.$$

(ii) There is $c_2 > 0$ such that for all $x \in U \cap \mathcal{K}$,

$$D_F^+ V(x) \leq -c_2\|\nabla V(x)\|\|F(x)\|.$$

Then every trajectory that starts in $U \cap \mathcal{K}$ and remains in $U \cap \mathcal{K}$ for all time has finite arclength and converges to a point in $\mathcal{S}$. If, in addition,

(iii) x^* is a minimizer of V if and only if $x^* \in \mathcal{S}$.

then every $x^* \in \mathcal{S}$ is semistable relative to $\mathcal{K}$.

Proof: Suppose (i) and (ii) hold and assume without loss of generality that $\psi(0) = 0$. Define $\tilde{V} : U \cap \mathcal{K} \to \mathbb{R}$ by

$$\tilde{V}(x) = \begin{cases} \psi(V(x) - V^*) & V(x) > V^* \\ 0 & V(x) = V^* \\ -\psi(V^* - V(x)) & V(x) < V^*. \end{cases}$$

Then for all $x \in U$ with $V(x) > V^*$, we have

$$\begin{aligned} D_F^+ \tilde{V}(x) &= \psi'(V(x) - V^*) D_F^+ V(x) \\ &\leq -c_2 \psi'(V(x) - V^*)\|\nabla V(x)\|\|F(x)\| \\ &\leq -c_1 c_2 \|F(x)\|. \end{aligned}$$

A similar argument can be used to show that the above inequality also holds when $V(x) \leq V^*$. Since ψ is increasing, $x^* \in U \cap \mathcal{K}$ is a local minimizer of $\tilde{V}$ if and only if x^* is a local minimizer of V. Hence, the result follows by applying Lemma A.2 with the Lyapunov function $\tilde{V}$. ■

[2]The online version [55] discusses classes of functions for which the Kurdyka-Łojasiewicz inequality holds.

Appendix B
Locally Bounded Set of Lagrange Multipliers

The proof of Theorem 5.7(ii) requires the following result, which establishes conditions under which $\Lambda_\alpha(x)$ is locally bounded.

Lemma B.1: (Local boundedness of $\Lambda_\alpha(x)$): Let $x^* \in X_{\text{KKT}}$ and suppose MFCQ is satisfied at x^*. Let U be a bounded, open set containing x^* on which (17) is well defined and $\Lambda_\alpha(x) \neq \emptyset$ for all $x \in U$. Then, there exists $B < \infty$ with

$$\sup_{x \in U} \left\{ \sup_{(u,v) \in \Lambda_\alpha(x)} \|(u, v)\|_\infty \right\} < B. \tag{34}$$

Proof: By [58, Corollary 4.3], the solution map of (17), $x \mapsto \{\mathcal{G}_\alpha(x)\} \times \Lambda_\alpha(x)$, satisfies the Lipschitz stability property that there exists $\ell > 0$ where

$$\|\mathcal{G}_\alpha(x)\| + \text{dist}_{\Lambda_\alpha(x^*)}(u, v) \leq \ell\|x - x^*\|, \tag{35}$$

for all $(u, v) \in \Lambda_\alpha(x)$ and all $x \in U$. By Proposition 5.1, $\Lambda_\alpha(x^*)$ is precisely the set of Lagrange multipliers of (4) at x^*, so MFCQ implies that $\Lambda_\alpha(x^*)$ is bounded [30]. Suppose by contradiction that (34) does not hold. Then there exists a sequence $\{x^\nu\}_{\nu=1}^\infty \subset U$ and $(u^\nu, v^\nu) \in \Lambda_\alpha(x^\nu)$ where $\|(u^\nu, v^\nu)\|_\infty \to \infty$ as $\nu \to \infty$. Since $\Lambda_\alpha(x^*)$ is bounded, $\|\mathcal{G}_\alpha(x^\nu)\| + \text{dist}_{\Lambda_\alpha(x^*)}((u^\nu, v^\nu)) \to \infty$, which contradicts (35) and the fact that U is bounded. ■

References

[1] A. Allibhoy and J. Cortés, "Anytime solution of constrained nonlinear programs via control barrier functions," in *IEEE Conf. on Decision and Control*, (Austin, Texas), pp. 6520–6525, Dec. 2021.

[2] K. Arrow, L. Hurwitz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford, CA: Stanford University Press, 1958.

[3] R. W. Brockett, "Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems," *Linear Algebra and its Applications*, vol. 146, pp. 79–91, 1991.

[4] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*. Springer, 1994.

[5] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Optimization algorithms as robust feedback controllers," *arXiv preprint arXiv:2103.11329*, 2021.

[6] A. Jokic, M. Lazar, and P. J. van den Bosch, "On constrained steady-state regulation: Dynamic KKT controllers," *IEEE Transactions on Automatic Control*, vol. 54, no. 9, pp. 2250–2254, 2009.

[7] M. Colombino, E. Dall'Anese, and A. Bernstein, "Online optimization as a feedback controller: Stability and tracking," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 422–432, 2020.

[8] L. S. P. Lawrence, J. W. Simpson-Porco, and E. Mallada, "Linear-convex optimal steady-state control," *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5377–5384, 2021.

[9] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Control Systems*, vol. 22, no. 1, pp. 28–43, 2002.

[10] G. Bianchin, J. Cortés, J. I. Poveda, and E. Dall'Anese, "Time-varying optimization of LTI systems via projected primal-dual gradient flows," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 1, pp. 474–486, 2022.

[11] K. Tanabe, "A geometric method in nonlinear programming," *Journal of Optimization Theory and Applications*, vol. 30, no. 2, pp. 181–210, 1980.

[12] H. Yamashita, "A differential equation approach to nonlinear programming," *Mathematical Programming*, vol. 18, no. 1, pp. 155–168, 1980.

[13] J. Schropp and I. Singer, "A dynamical systems approach to constrained minimization," *Numerical functional analysis and optimization*, vol. 21, no. 3-4, pp. 537–551, 2000.

[14] A. Nagurney and D. Zhang, *Projected Dynamical Systems and Variational Inequalities with Applications*, vol. 2 of *International Series in Operations Research and Management Science*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

[15] A. Hauswirth, S. Bolognani, and F. Dörfler, "Projected dynamical systems on irregular, non-Euclidean domains for nonlinear optimization," *SIAM Journal on Control and Optimization*, vol. 59, no. 1, pp. 635–668, 2021.

[16] J. Cortés, "Discontinuous dynamical systems - a tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Systems*, vol. 28, no. 3, pp. 36–73, 2008.

[17] T. L. Friesz, D. H. Bernstein, N. J. Mehta, R. L. Tobin, and S. Ganjlizadeh, "Day-to-day dynamic network disequilibria and idealized traveler information systems," *Operations Research*, vol. 42, no. 6, pp. 1120–1136, 1994.

[18] Y. S. Xia and J. Wang, "On the stability of globally projected dynamical systems," *Journal of Optimization Theory & Applications*, vol. 106, no. 1, pp. 129–150, 2000.

[19] M. Muehlebach and M. I. Jordan, "On constraints in first-order optimization: A view from non-smooth dynamical systems," *Journal of Machine Learning Research*, vol. 23, pp. 1–47, 2022.

[20] T. Kose, "Solutions of saddle value problems by differential equations," *Econometrica*, vol. 24, no. 1, pp. 59–70, 1956.

[21] T. Lin, C. Jin, and M. I. Jordan, "Near-optimal algorithms for minimax optimization," *arXiv preprint arXiv:2002.02417*, 2020.

[22] D. Feijer and F. Paganini, "Stability of primal-dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, pp. 1974–1981, 2010.

[23] A. Cherukuri, B. Gharesifard, and J. Cortés, "Saddle-point dynamics: conditions for asymptotic stability of saddle points," *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 486–511, 2017.

[24] A. Cherukuri, E. Mallada, S. H. Low, and J. Cortés, "The role of convexity in saddle-point dynamics: Lyapunov function and robustness," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2449–2464, 2018.

[25] W. Haddad and V. S. Chellaboina, *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton, NJ: Princeton University Press, 2008.

[26] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, vol. 317 of *Comprehensive Studies in Mathematics*. New York: Springer, 1998.

[27] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2017.

[28] 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), pp. 3420–3431, 2019.

[29] 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), pp. 151–158, 2015.

[30] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2nd ed., 1999.

[31] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[32] S. M. Robinson, "Stability theory for systems of inequalities. part i: Linear systems," *SIAM Journal on Numerical Analysis*, vol. 12, no. 5, pp. 754–769, 1975.

[33] M. Best and B. Ding, "On the continuity of the minimum in parametric quadratic programs," *Journal of Optimization Theory & Applications*, vol. 86, no. 1, pp. 245–250, 1995.

[34] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2009.

[35] J. W. Daniel, "The continuity of metric projections as functions of the data," *Journal of Approximation Theory*, vol. 12, no. 3, pp. 234–239, 1974.

[36] S. M. Robinson, "Strongly regular generalized equations," *Mathematics of Operations Research*, vol. 5, no. 1, pp. 43–62, 1980.

[37] G. Di Pillo and L. Grippo, "Exact penalty functions in constrained optimization," *SIAM Journal on Control and Optimization*, vol. 27, no. 6, pp. 1333–1360, 1989.

[38] J. Nocedal and S. Wright, *Numerical optimization*. Berlin, Heidelberg: Springer, 2006.

[39] A. V. Fiacco, "Sensitivity analysis for nonlinear programming using penalty methods," *Mathematical programming*, vol. 10, no. 1, pp. 287–311, 1976.

[40] K. Jittorntrum, "Solution point differentiability without strict complementarity in nonlinear programming," in *Sensitivity, Stability and Parametric Analysis*, pp. 127–138, Springer, 1984.

[41] P.-A. Absil and K. Kurdyka, "On the stable equilibrium points of gradient systems," *Systems & Control Letters*, vol. 55, pp. 573–577, 2006.

[42] S. Lojasiewicz, "Sur les trajectoires du gradient d'une fonction analytique," *Seminari di geometria*, vol. 1983, pp. 115–117, 1982.

[43] J. V. Burke and M. C. Ferris, "Weak sharp minima in mathematical programming," *SIAM Journal on Control and Optimization*, vol. 31, no. 5, pp. 1340–1359, 1993.

[44] M. Studniarski and D. E. Ward, "Weak sharp minima: characterizations and sufficient conditions," *SIAM Journal on Control and Optimization*, vol. 38, no. 1, pp. 219–236, 1999.

[45] A. Shapiro, "Second order sensitivity analysis and asymptotic theory of parametrized nonlinear programs," *Mathematical Programming*, vol. 33, no. 3, pp. 280–299, 1985.

[46] K. Kurdyka, "On gradients of functions definable in o-minimal structures," *Annales de l'institut Fourier*, vol. 48, no. 3, pp. 769–783, 1998.

[47] E. Bierstone and P. D. Milman, "Semianalytic and subanalytic sets," *Inst. Hautes Études Sci. Publ. Math.*, vol. 67, no. 1, pp. 5–42, 1988.

[48] L. Van den Dries and C. Miller, "Geometric categories and o-minimal structures," *Duke Mathematical Journal*, vol. 84, no. 2, pp. 497–540, 1996.

[49] A. Arsie and C. Ebenbauer, "Locating omega-limit sets using height functions," *Journal of Differential Equations*, vol. 248, no. 10, pp. 2458–2469, 2010.

[50] J. Bolte, A. Daniilidis, and A. Lewis, "A nonsmooth Morse–Sard theorem for subanalytic functions," *Journal of mathematical analysis and applications*, vol. 321, no. 2, pp. 729–740, 2006.

[51] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM review*, vol. 44, no. 4, pp. 525–597, 2002.

[52] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques*. SIAM, 1990.

[53] S. P. Bhat and D. S. Bernstein, "Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria," *SIAM Journal on Control and Optimization*, vol. 42, no. 5, pp. 1745–1775, 2003.

[54] S. P. Bhat and D. S. Bernstein, "Arc-length-based Lyapunov tests for convergence and stability with applications to systems having a continuum of equilibria," *Mathematics of Control, Signals and Systems*, vol. 22, no. 2, pp. 155–184, 2010.

[55] A. Allibhoy and J. Cortés, "Control barrier function based design of gradient flows for constrained nonlinear programming," `http://arxiv.org/abs/2204.01930`, 2022.

[56] P.-A. Absil, R. Mahony, and B. Andrews, "Convergence of the iterates of descent methods for analytic cost functions," *SIAM Journal on Optimization*, vol. 16, no. 2, pp. 531–547, 2005.

[57] C. Lageman, "Pointwise convergence of gradient-like systems," *Mathematische Nachrichten*, vol. 280, no. 13-14, pp. 1543–1558, 2007.

[58] S. M. Robinson, "Generalized equations and their solutions, part II: Applications to nonlinear programming," in *Optimality and Stability in Mathematical Programming*, pp. 200–221, Berlin, Heidelberg: Springer, 1982.

Ahmed Allibhoy received a B.S. in Math and a B.S. in Electrical Engineering from the University of California, Los Angeles in 2018, and an M.S. in Mechanical Engineering from the University of California, San Diego, in 2020. He is currently a Ph.D. student in the Department of Mechanical and Aerospace Engineering at UCSD. His main research interests include optimization, nonlinear control, and networked systems.

Jorge Cortés (M'02, SM'06, F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California Santa Cruz, CA, USA. He is a Professor in the Department of Mechanical and Aerospace Engineering, University of California San Diego, CA, USA. He is a Fellow of IEEE, SIAM, and IFAC. His research interests include distributed control and optimization, network science, nonsmooth analysis, reasoning and decision making under uncertainty, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.