

Anytime-Feasible First-Order Optimization via Safe Sequential QCQP

Jiarui Wang and Mahyar Fazlyab

Abstract—This paper presents the Safe Sequential Quadratically Constrained Quadratic Programming (SS-QCQP) algorithm, a first-order method for smooth inequality-constrained nonconvex optimization that guarantees feasibility at every iteration. The method is derived from a continuous-time dynamical system whose vector field is obtained by solving a convex QCQP that enforces monotonic descent of the objective and forward invariance of the feasible set. The resulting continuous-time dynamics achieve an $O(1/t)$ convergence rate to first-order stationary points under standard constraint qualification conditions. We then propose a safeguarded Euler discretization with adaptive step-size selection that preserves this convergence rate while maintaining both descent and feasibility in discrete time. To enhance scalability, we develop an active-set variant (SS-QCQP-AS) that selectively enforces constraints near the boundary, substantially reducing computational cost without compromising theoretical guarantees. Numerical experiments on a multi-agent nonlinear optimal control problem demonstrate that SS-QCQP and SS-QCQP-AS maintain feasibility, exhibit the predicted convergence behavior, and deliver solution quality comparable to second-order solvers such as SQP and IPOPT.

Index Terms—Constrained optimization, Nonconvex optimization, First-order methods, Anytime algorithms, Quadratically constrained quadratic programming (QCQP), Safe optimization, Control barrier functions.

I. INTRODUCTION

Constrained optimization forms the foundation of numerous engineering disciplines, including optimal power flow [1], optimal control [2], and robotics [3]. In many practical settings, particularly when algorithms are deployed on high-frequency physical systems such as quadrotors, optimization must be executed in real time [4] [5]. This necessitates producing feasible solutions prior to full convergence while operating under strict computational limitations. Meeting these requirements poses fundamental challenges for the design of optimization methods that are simultaneously efficient, reliable, and capable of enforcing constraints at all intermediate iterates.

A key requirement in this setting is *anytime feasibility*. This property guarantees that the algorithm can be safely interrupted at any time while still returning a feasible, though possibly suboptimal, solution. Such a property is essential in real-world deployments, where constraints frequently encode

safety conditions; violating them may induce instability, equipment damage, or safety hazards.

Another desirable property is *monotonic descent*, meaning that the objective value decreases at every iteration. Since premature termination is common in real-time settings, monotonic descent ensures consistent improvement over time and provides a quantifiable performance guarantee for any intermediate solution.

A third desirable property is *scalability*. Modern machine learning, control, and robotics applications increasingly involve high-dimensional decision variables and large numbers of constraints, for example, model predictive control with long horizons, multi-agent coordination tasks, optimization problems arising in real-time perception and planning, or data-driven verification [6]. In such settings, each iteration must be computationally inexpensive, ideally relying only on first-order information, so that the algorithm can meet strict real-time deadlines and be deployed on resource-limited hardware.

Despite the importance of these properties, most classical algorithms do not satisfy them simultaneously. Interior-point methods [7]–[9] enforce feasibility only asymptotically and generally allow infeasible iterates until convergence, which is incompatible with safety-critical applications. Sequential Quadratic Programming (SQP) methods [10]–[12] may violate constraints even when initialized at a feasible point due to the mismatch between linearized and nonlinear constraints. Penalty and augmented Lagrangian methods deliberately relax constraints and therefore cannot guarantee feasibility during intermediate iterations. Moreover, many of these classical methods require second-order information or the solution of large linear systems at each iteration, which limits scalability and hinders deployment in real-time environments. These limitations highlight the need for first-order optimization algorithms that simultaneously ensure anytime feasibility, monotonic descent, and scalability.

Our Contribution: To address the limitations of existing constrained optimization methods, this paper proposes a new first-order framework for solving general inequality-constrained nonconvex problems. Our main contributions are as follows: (1) Adopting a control-theoretic perspective, we design a continuous-time dynamical system whose vector field is obtained by solving a convex quadratically constrained quadratic program (QCQP) that uses only first-order information. The dynamics guarantee forward invariance of the feasible set, monotonic descent, and global convergence to Karush–Kuhn–Tucker (KKT) points. We further show that the continuous-time system achieves an $O(1/t)$ ergodic conver-

The authors are with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: {jwang486, mahyarfazlyab}@jhu.edu).

gence rate in terms of a first-order stationary measure. (2) We discretize the continuous-time dynamics using a safeguarded Armijo-type line search. The discrete-time method preserves anytime feasibility and monotonic decrease of the objective and matches the continuous-time $O(1/t)$ rate. (3) To enhance scalability in problems with many constraints, we introduce an active-set variant that enforces only a subset of nearly active constraints at each iteration, substantially reducing per-iteration computational cost. (4) We establish convergence guarantees for both the full and active-set variants, showing that each achieves an $O(1/k)$ ergodic convergence rate in discrete time under standard Mangasarian-Fromovitz constraint qualification (MFCQ) and Lipschitz assumptions, matching the rate of classical gradient descent for unconstrained problems. (5) We provide detailed implementation guidelines, including conic reformulations, solver considerations, and adaptive hyperparameter selection, to ensure strong empirical performance in practical deployments.

A. Related Work

1) *SQP Methods*: A widely used approach for solving nonconvex constrained optimization problems is to approximate them by a sequence of convex subproblems. The most prominent example is Sequential Quadratic Programming (SQP), which computes search directions by solving a convex quadratic program (QP) obtained from a second-order expansion of the Lagrangian and a first-order linearization of the constraints. Viewed analytically, SQP corresponds to applying Newton's method to the KKT conditions. However, because linearized constraints only approximate the feasible region locally, the QP step can be overly optimistic—improving the objective while violating the true nonlinear constraints. Line search on an exact penalty function often forces very small step sizes, a limitation known as the Maratos effect [13]. Classical remedies, such as augmented Lagrangian merit functions [14] or second-order correction steps [15], [16], improve robustness but still do not guarantee anytime feasibility or monotonic descent. In model predictive control (MPC), SQP variants have been adapted to return feasible iterates, for instance through double-loop schemes [17] or feasibility perturbation techniques [18]. Nevertheless, these methods generally lack monotonic decrease guarantees and may still exhibit constraint violations during intermediate iterations.

2) *QCQP Methods*: Sequential quadratically constrained quadratic programming (SQCQP) methods follow a structure similar to SQP, but replace linearized constraints with quadratic approximations, leading to a QCQP subproblem at each iteration. Classical approaches construct second-order Taylor expansions of the constraints, which requires computing Hessians and generally yields nonconvex QCQPs. To avoid these difficulties, Fukushima [19] proposed computing a search direction via a QCQP but, due to the lack of efficient QCQP solvers at the time, implemented this idea through two successive convex QPs. Variants for convex problems have also been explored [20], though these methods still rely on line search with an ℓ_1 penalty merit function and therefore do not guarantee anytime feasibility or monotonic descent.

A related line of work is the phase-I–phase-II SQCQP scheme of [21], which uses quadratic approximations of both the objective and the constraints sharing a common Hessian, chosen as a positive multiple of the identity. Once feasibility is achieved, this approach ensures both anytime feasibility and monotonic decrease, with convergence established under standard assumptions including twice differentiability, MFCQ, second-order sufficiency, and strict complementarity. However, the method requires including *all* constraints in every QCQP subproblem, a significant limitation for large-scale problems with many constraints.

3) *First-Order Methods*: A growing body of work develops first-order algorithms for constrained optimization motivated by control theory [22]–[24] or Lipschitz-based surrogate modeling. One line of research designs continuous-time dynamics that guarantee invariance of the feasible set. In the safe gradient flow framework [25], each constraint is treated as a Control Barrier Function (CBF) [26], yielding a vector field that follows the unconstrained gradient flow as closely as possible while enforcing feasibility in continuous time. However, when discretized, these flows generally lose their invariance properties—an issue closely related to the Maratos effect—so that anytime feasibility cannot be guaranteed in discrete time. A related formulation in [27] provides a discrete-time update with convergence guarantees, but feasibility is not preserved at intermediate iterates and the stepsize depends on problem-dependent constants that are difficult to estimate in practice.

A complementary line of first-order methods constructs conservative local approximations of the objective and constraints using Lipschitz bounds. Representative examples include Successive Convex Approximation (SCA) [28] and the Moving Ball Approximation (MBA) method [29]. While these approaches guarantee feasibility of the surrogate subproblem, they require knowledge of the Lipschitz constants of the gradients, which are often unavailable or hard to estimate. Lu et al. [30] proposed a line-search-based Sequential Convex Programming (SCP) algorithm that avoids explicit Lipschitz constants, but each line-search step involves solving a convex QCQP, resulting in substantially higher computational cost. Overall, existing first-order methods either lack anytime feasibility after discretization or rely on problem-dependent constants, limiting their applicability in real-time or large-scale settings.

B. Notation

We denote the space of n -dimensional real vectors by \mathbb{R}^n . For a vector $x \in \mathbb{R}^n$, $\|x\|_2$ denotes the Euclidean norm, and $\|x\|_P = \sqrt{x^\top P x}$ denotes the weighted P -norm, where $P \in \mathbb{R}^{n \times n}$ is symmetric positive definite ($P \succ 0$). For a natural number m , we use $[m] := \{1, 2, \dots, m\}$ to denote the index set of the first m integers. For $x \in \mathbb{R}^n$ and $r > 0$, $\mathcal{B}(x, r)$ denotes the open ball centered at x with radius r .

II. BACKGROUND AND PROBLEM STATEMENT

This section introduces the constrained optimization problem studied in the paper, reviews the assumptions and optimal-

ity conditions used throughout, and summarizes key methods that motivate the development of our algorithm.

A. Problem Setup and Assumptions

We consider inequality-constrained, possibly nonconvex, optimization problems of the form

$$\begin{aligned} \min_x f(x) \\ \text{s.t. } g_i(x) \leq 0, \quad i \in [m], \end{aligned} \quad (\text{OPT})$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are continuously differentiable functions. We denote the feasible set as

$$\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \forall i \in [m]\}.$$

For $x \in \mathcal{F}$, we denote the set of active constraints at x as $\mathcal{I}(x) = \{i \in [m] : g_i(x) = 0\}$.

Definition 1: A continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is L -smooth if its gradient is L -Lipschitz, i.e.,

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2 \quad \forall x, y.$$

The L -smoothness of a function f implies the following quadratic upper bound (see, e.g., [31]):

$$f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2}\|y - x\|_2^2 \quad \forall x, y. \quad (1)$$

Definition 2 (MFCQ): We say the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at x for (OPT) if $\exists d \in \mathbb{R}^n$ such that $\nabla g_i(x)^\top d < 0 \forall i \in \mathcal{I}(x)$.

Assumption 1: Throughout the paper, the following assumptions are imposed on (OPT).

1. f is L_f -smooth and g_i is L_i -smooth $\forall i \in [m]$.
2. (OPT) is feasible and bounded below, and the optimal value is achieved at x^* .
3. The Mangasarian-Fromovitz constraint qualification (MFCQ) holds for all $x \in \mathcal{F}$.

Definition 3: A point $x \in \mathbb{R}^n$ is said to be *stationary* for (OPT) if there exists $\lambda \in \mathbb{R}^m$ such that the pair (x, λ) satisfies the Karush–Kuhn–Tucker (KKT) conditions:

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) = 0, \quad (2a)$$

$$g_i(x) \leq 0 \quad \forall i \in [m], \quad (2b)$$

$$\lambda \geq 0, \quad (2c)$$

$$\lambda_i g_i(x) = 0 \quad \forall i \in [m]. \quad (2d)$$

B. Background

1) Safe Gradient Flow: Leveraging the theory of Control Barrier Functions (CBFs) [32], Allibhoy *et al.* [22] introduced the *safe gradient flow* dynamics for solving constrained optimization problems of the form (OPT). Specifically, the safe gradient flow is defined as

$$\dot{x} = \tilde{u}(x), \quad (3)$$

where the velocity field $\tilde{u}(x)$ is obtained as the solution of the following convex quadratic program (QP):

$$\begin{aligned} \tilde{u}(x) := \arg \min_u \frac{1}{2}\|u + \nabla f(x)\|_2^2 \\ \text{s.t. } \nabla g_i(x)^\top u \leq -\alpha g_i(x), \quad \forall i \in [m], \end{aligned} \quad (4)$$

with $\alpha > 0$ a prescribed parameter. This QP can be interpreted in two complementary ways.

From a *control-theoretic* perspective, the QP computes the smallest correction to the unconstrained gradient descent direction that ensures satisfaction of the differential inequalities

$$\dot{g}_i(x) + \alpha g_i(x) \leq 0, \quad \forall i \in [m], \quad (5)$$

which serve as control barrier conditions. These inequalities guarantee both *forward invariance* and *exponential stability* of the feasible set. In particular, if the initial condition is feasible, $g_i(x(0)) \leq 0$, then $g_i(x(t)) \leq 0$ for all $t \geq 0$; moreover, if a constraint is initially violated, $g_i(x(0)) > 0$, its violation decays exponentially, i.e.,

$$g_i(x(t)) \leq g_i(x(0))e^{-\alpha t}.$$

From an *optimization* standpoint, when $\alpha = 1$, the QP in (4) projects the negative gradient $-\nabla f(x)$ onto the linearized feasible set

$$\{u : \nabla g_i(x)^\top u \leq -g_i(x), \quad \forall i \in [m]\}, \quad (6)$$

This coincides with the SQP search direction under an identity Hessian approximation of the Lagrangian and has also appeared in the MPC literature as a means of generating feasible descent directions [33].

Under the Mangasarian–Fromovitz constraint qualification (MFCQ) and the constant rank condition (CRC) [34], the mapping $x \mapsto \tilde{u}(x)$ is locally Lipschitz [22]. This guarantees existence and uniqueness of solutions of (3). Moreover, if x^* satisfies $\tilde{u}(x^*) = 0$, then (x^*, λ^*) satisfies the KKT conditions of (OPT), and conversely. If x^* is an isolated stationary point, it is asymptotically stable; exponential stability follows under LICQ, strict complementarity, and the second-order sufficient condition [22, Thm. 5.7].

Although these results provide strong guarantees for the continuous-time system, the analysis in [22] does not address the discretization required for numerical implementation. In particular, the anytime feasibility property—which holds exactly in continuous time—may be lost upon discretization.

2) Constrained Gradient Flow (Descent): In closely related work, Muehlebach *et al.* [27] proposed the *constrained gradient flow*, in which the invariance condition (5) is enforced only for active or violated constraints. This modification reduces the number of constraints appearing in the QP (4), but it renders the mapping $x \mapsto \tilde{u}(x)$ nonsmooth. The authors further developed a constant-step size Euler discretization of the resulting dynamics and established that, under convexity assumptions on problem (OPT), the squared distance to an optimal solution decreases at a rate of $\mathcal{O}(1/k)$. However, the discrete-time scheme in [27] does not guarantee feasibility of the iterates, and its admissible step size depends on problem-dependent constants that are often difficult to estimate in practice.

In the next section, we address these issues by developing a new continuous-time dynamical system and an associated safeguarded discretization procedure that provably preserves feasibility at every iteration while ensuring monotonic descent of the objective.

III. SAFE SEQUENTIAL QCQP

In this section, we develop a continuous-time dynamical system $\dot{x} = u(x)$ and an associated Euler discretization for solving problem (OPT). Given an iterate $x^{(k)}$, the update takes the form

$$x^{(k+1)} = x^{(k)} + t^{(k)}u(x^{(k)}), \quad (7)$$

where the vector field (or search direction) $u(x^{(k)})$ and step size $t^{(k)} > 0$ are designed to guarantee two fundamental properties: (1) *Anytime feasibility*: If the initial iterate $x^{(0)}$ satisfies $g_i(x^{(0)}) \leq 0$ for all $i \in [m]$, then feasibility is maintained at every iteration, i.e.,

$$g_i(x^{(k)}) \leq 0, \quad \forall k \geq 0, i \in [m]. \quad (8)$$

(2) *Monotonic descent*: The sequence of objective values is strictly decreasing,

$$f(x^{(k+1)}) < f(x^{(k)}), \quad \forall k \geq 0. \quad (9)$$

The proposed method, referred to as the *Safe Sequential QCQP (SS-QCQP) algorithm*, constructs $u(x^{(k)})$ by solving a convex quadratically constrained quadratic program (QCQP) that locally approximates (OPT) while enforcing feasibility (Section III-A). The step size t_k is then selected to guarantee both a sufficient decrease in the objective and preservation of safety along the resulting feasible direction (Section III-D).

A. ODE Construction

Discretizing the safe gradient flow (3) into an iterative algorithm introduces a fundamental difficulty: the linearization of the constraints in (4) neglects curvature, which can cause infeasibility for any positive step size. To illustrate this, consider a point x where $g_i(x) = 0$ for some $i \in [m]$ (an active constraint). In this case, the feasible set of (4) permits

$$\nabla g_i(x)^\top \tilde{u}(x) = 0, \quad (10)$$

meaning that $\tilde{u}(x)$ can be tangent to the zero-level set of g_i . Consequently, taking any positive step in the direction of $\tilde{u}(x)$ may violate feasibility, i.e.,

$$g_i(x + t\tilde{u}(x)) > 0, \quad \forall t > 0. \quad (11)$$

This phenomenon, similar to the *Maratos effect*, arises from the mismatch between the local linear approximation of the constraints and their true nonlinear geometry.

To mitigate this limitation, we incorporate a quadratic correction term into the feasibility constraints of the QP (4), leading to the following convex quadratically constrained quadratic program (QCQP):

$$\begin{aligned} u(x) := \arg \min_u & \quad \frac{1}{2} \|u + \nabla f(x)\|_2^2 \\ \text{s.t.} & \quad \nabla g_i(x)^\top u \leq -\alpha g_i(x) - w_i \|u\|_2^2, \quad \forall i \in [m], \end{aligned} \quad (\text{SS-QCQP}(x))$$

where $\alpha > 0$ and $w_i > 0$ are prescribed parameters. When $g_i(x) = 0$, i.e., when x lies on the boundary of the feasible region, the additional quadratic term $w_i \|u\|_2^2$ ensures that

$$\nabla g_i(x)^\top u(x) < 0, \quad (12)$$

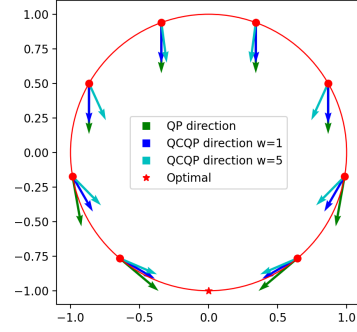


Fig. 1: The search directions computed by solving the QP (4) and the QCQP (SS-QCQP(x)) where $x = [x_1, x_2]^\top$, the objective $f(x) = x_2$ is a linear function and the constraint $g(x) = \|x\|_2^2 - 1 \leq 0$ is a unit ball constraint. Here we scale the search directions differently so that they do not overlap. When x is on the boundary of the feasible region, i.e., $g(x) = 0$, (4) might generate a direction tangent to the boundary, while (SS-QCQP(x)) tilts the direction towards the interior of the feasible region.

thereby preventing the vector field from being tangent to the constraint boundary. Geometrically, the penalty $w_i \|u\|_2^2$ tilts the direction towards the interior of the feasible set (see Figure 1), thereby providing a curvature-aware safeguard absent in the original QP formulation.

In the next subsection, we establish key properties of the direction field $u(x)$ defined by (SS-QCQP(x)), including feasibility preservation, regularity, and its correspondence with KKT points of (OPT).

B. Properties of the Search Direction

We begin by showing that (SS-QCQP(x)) always admits a strictly feasible point whenever the original problem (OPT) satisfies MFCQ. Because the QCQP is convex, strict feasibility guarantees that the optimizer $u(x)$ exists and is unique.

Lemma 1 (Strict Feasibility of (SS-QCQP(x))): Suppose the MFCQ condition holds for (OPT) and that x is feasible. Then (SS-QCQP(x)) is strictly feasible.

Proof: Under MFCQ, there exists a vector u such that

$$\nabla g_i(x)^\top u < 0, \quad \forall i \in \mathcal{I}(x),$$

where $\mathcal{I}(x) := \{i \in [m] : g_i(x) = 0\}$ denotes the set of active constraints. We consider two cases.

Case 1: $i \in \mathcal{I}(x)$ (active constraints). Since $g_i(x) = 0$ and $\nabla g_i(x)^\top u < 0$, choose any

$$0 < \eta_1 < \min_{i \in \mathcal{I}(x)} \frac{-\nabla g_i(x)^\top u}{w_i \|u\|_2^2}.$$

Then $\hat{u}_1 := \eta_1 u$ satisfies

$$\nabla g_i(x)^\top \hat{u}_1 = \eta_1 \nabla g_i(x)^\top u < -w_i \|\hat{u}_1\|_2^2 = -\alpha g_i(x) - w_i \|\hat{u}_1\|_2^2,$$

for all $i \in \mathcal{I}(x)$.

Case 2: $i \notin \mathcal{I}(x)$ (inactive constraints). For these indices, $g_i(x) < 0$. The inequality

$$\nabla g_i(x)^\top (\eta u) \leq -\alpha g_i(x) - w_i \|\eta u\|_2^2$$

is quadratic in η . It holds strictly for any

$$0 < \eta_2 < \min_{i \notin \mathcal{I}(x)} \frac{-\nabla g_i(x)^\top u + \sqrt{(\nabla g_i(x)^\top u)^2 - 4\alpha w_i \|u\|_2^2 g_i(x)}}{2w_i \|u\|_2^2}.$$

Finally, let $\eta := \min\{\eta_1, \eta_2\}$ and set $\hat{u} := \eta u$. Then \hat{u} satisfies

$$\nabla g_i(x)^\top \hat{u} < -\alpha g_i(x) - w_i \|\hat{u}\|_2^2, \quad \forall i \in [m],$$

showing that (SS-QCQP(x)) is strictly feasible. ■

Theorem 1: Assume that problem (OPT) satisfies Assumption 1, and let x be a feasible point of (OPT). Then the solution $u(x)$ to (SS-QCQP(x)) satisfies:

- 1) $u(x) = 0$ if and only if x is a stationary point of (OPT).
- 2) The search direction is a descent direction for f , i.e.,

$$\nabla f(x)^\top u(x) \leq -\|u(x)\|_2^2. \quad (13)$$

Proof: Since (SS-QCQP(x)) is convex and strictly feasible (by the preceding lemma), strong duality holds and the KKT conditions are necessary and sufficient [35]. Define

$$c_i(u; x) := \nabla g_i(x)^\top u + w_i \|u\|_2^2 + \alpha g_i(x), \quad i \in [m].$$

The KKT conditions of (SS-QCQP(x)) state that there exists $\lambda(x) \in \mathbb{R}_{\geq 0}^m$ such that

$$u(x) + \nabla f(x) + \sum_{i=1}^m \lambda_i(x) (\nabla g_i(x) + 2w_i u(x)) = 0, \quad (14a)$$

$$c_i(u(x); x) \leq 0, \quad i \in [m], \quad (14b)$$

$$\lambda_i(x) \geq 0, \quad i \in [m], \quad (14c)$$

$$\lambda_i(x) c_i(u(x); x) = 0, \quad i \in [m]. \quad (14d)$$

Part (1): (\implies) Setting $u(x) = 0$ in (14a) gives

$$\nabla f(x) + \sum_{i=1}^m \lambda_i(x) \nabla g_i(x) = 0.$$

Complementarity (14d) with $u(x) = 0$ yields $\alpha \lambda_i(x) g_i(x) = 0$, so $\lambda_i(x) = 0$ whenever $g_i(x) < 0$, while $\lambda_i(x) \geq 0$ is allowed if $g_i(x) = 0$. Thus, x together with multipliers satisfies the KKT conditions of (OPT), proving stationarity.

(\impliedby) Let x be a stationary point of (OPT). By Definition 3, there exists λ such that (x, λ) satisfies (2). Letting $u = 0$, (2a) gives

$$u + \nabla f(x) + \sum_{i=1}^m \lambda_i(x) (\nabla g_i(x) + 2w_i u) = 0.$$

Complementarity (2d) yields $\lambda_i c_i(0; x) = 0$, and (2b) yields $c_i(0; x) = \alpha g_i(x) \leq 0$. Thus, the pair $(0, \lambda)$ satisfies the KKT conditions (14). Since (SS-QCQP(x)) is a convex problem with a strongly convex objective, $u = 0$ is the unique solution to (SS-QCQP(x)).

Part (2): Taking the inner product of (14a) with $u(x)$ gives

$$\begin{aligned} & \nabla f(x)^\top u(x) \\ &= -\left(1 + 2 \sum_{i=1}^m \lambda_i(x) w_i\right) \|u(x)\|_2^2 - \sum_{i=1}^m \lambda_i(x) \nabla g_i(x)^\top u(x) \\ &= -\left(1 + 2 \sum_{i=1}^m \lambda_i(x) w_i\right) \|u(x)\|_2^2 + \sum_{i=1}^m \lambda_i(x) (\alpha g_i(x) + w_i \|u(x)\|_2^2) \\ &= -\left(1 + \sum_{i=1}^m \lambda_i(x) w_i\right) \|u(x)\|_2^2 + \sum_{i=1}^m \lambda_i(x) \alpha g_i(x) \\ &\leq -\|u(x)\|_2^2. \end{aligned}$$

Here we use the complementarity (14d) to get the second equality. ■

C. Convergence Analysis of Continuous-Time Dynamics

With the properties of the search direction $u(x)$ established in Theorem 1, we define the continuous-time system

$$\dot{x}(t) = u(x(t)) \quad \text{with } x(0) \text{ given} \quad (15)$$

to solve (OPT), where $u(x)$ is the unique solution to (SS-QCQP(x)). For this ODE to be well posed, we must ensure that the mapping $x \mapsto u(x)$ is continuous on a neighborhood of the feasible set \mathcal{F} . This property guarantees the existence of solutions to the ODE. In the following lemma, we will first extend the strict feasibility result of Lemma 1 from feasible points to an entire open neighborhood of \mathcal{F} . Then, we will show that every solution trajectory converges to the set of first-order stationary points of (OPT).

Lemma 2: There exists an open set U containing \mathcal{F} such that the QCQP (SS-QCQP(x)) is strictly feasible for all $x \in U$.

Proof: For any $x \in \mathcal{F}$, Lemma 1 establishes that (SS-QCQP(x)) is strictly feasible; that is, there exist u and $\delta_i < 0$ such that

$$\alpha g_i(x) + \nabla g_i(x)^\top u + w_i \|u\|_2^2 = \delta_i, \quad \forall i \in [m].$$

Since each g_i is L_i -smooth, there exists $\epsilon_i(x) > 0$ such that

$$\alpha g_i(y) + \nabla g_i(y)^\top u + w_i \|u\|_2^2 \leq \frac{\delta_i}{2} < 0,$$

for all $y \in \mathcal{B}(x, \epsilon_i(x))$. Hence, (SS-QCQP(y)) remains strictly feasible for all $y \in \mathcal{B}(x, \min_i \epsilon_i(x))$. Defining

$$U := \bigcup_{x \in \mathcal{F}} \mathcal{B}(x, \min_i \epsilon_i(x)),$$

we conclude that (SS-QCQP(x)) is strictly feasible for all $x \in U$. ■

Existence of Trajectories: The continuity of $u(\cdot)$ established above guarantees existence of solutions to the ODE (15), but it does not ensure uniqueness. Uniqueness would require local Lipschitz continuity of $u(\cdot)$, which in turn depends on the stability of the active-set structure of (SS-QCQP(x)). Such stability is typically ensured by the *Constant Rank Constraint Qualification* (CRC) applied directly to the QCQP. However, CRC is considerably stronger than MFCQ and is rarely imposed in nonlinear programming, where regularity assumptions are made for the *original* problem (OPT) rather than for subproblems such as (SS-QCQP(x)).

Moreover, CRC for (OPT) does not imply CRC for (SS-QCQP(x)). If the original problem satisfies CRC at a feasible point x , then the gradients of its active constraints, $\{\nabla g_i(x) : g_i(x) = 0\}$, are linearly independent. But when x is strictly feasible, this set is empty, and CRC provides no information. In contrast, the QCQP (SS-QCQP(x)) may still have active constraints,

$$\nabla g_i(x)^\top u(x) + \alpha g_i(x) + w_i \|u(x)\|_2^2 = 0.$$

The gradients associated with these QCQP-active constraints, $\{\nabla g_i(x) + 2w_i u(x) : \nabla g_i(x)^\top u(x) + \alpha g_i(x) + w_i \|u(x)\|_2^2 = 0\}$, need not be linearly independent, even if the original problem satisfies CRC. Thus, the QCQP subproblem may fail to satisfy CRC regardless of any regularity properties possessed by (OPT).

For these reasons, we do not assume uniqueness of trajectories, and all convergence guarantees below apply to *every* solution of the ODE.

Theorem 2: Suppose problem (OPT) satisfies Assumption 1, and $x(t)$ is any solution to the ODE (15) with a feasible initial condition $x(0) \in \mathcal{F}$. Then $x(t)$ satisfies the following properties:

1) *Anytime feasibility:*

$$g_i(x(t)) \leq 0, \quad \forall t > 0, i \in [m].$$

2) *Convergence:*

$$\frac{1}{2} \int_0^t \|u(x(\tau))\|_2^2 d\tau \leq f(x(0)) - f^*, \quad \forall t > 0.$$

where f^* is the optimal value of (OPT).

Proof: $\forall i \in [m]$, when $g_i(x) = 0$, we have

$$\dot{g}_i(x(t)) = \nabla g_i(x)^\top u(x(t)) \leq -\alpha g_i(x) - w_i \|u(x(t))\|_2^2 \leq 0.$$

According to Nagumo's theorem [36] [37], the 0-sublevel set $\{x : g_i(x) \leq 0\}$ is forward-invariant.

Let f^* be the optimal value of (OPT) and define the Lyapunov function $V(x(t)) = f(x(t)) - f^* + \frac{1}{2} \int_0^t \|u(x(\tau))\|_2^2 d\tau$. Using the result $\nabla f(x(t))^\top u(x(t)) \leq -\|u(x(t))\|_2^2$ from Theorem 1, we have

$$\begin{aligned} \dot{V}(x(t)) &= \nabla f(x(t))^\top u(x(t)) + \frac{1}{2} \|u(x(t))\|_2^2 \\ &\leq -\frac{1}{2} \|u(x(t))\|_2^2 \end{aligned}$$

Thus $V(x(t)) \leq V(x(0)) \forall t > 0$. Since $f(x(t)) - f^* \geq 0$, we have

$$\frac{1}{2} \int_0^t \|u(x(\tau))\|_2^2 d\tau \leq V(x(t)) \leq V(x(0)) = f(x(0)) - f^*.$$

This implies that

$$\min_{\tau \in [0, t]} \|u(x(\tau))\|_2^2 \leq \frac{1}{t} \int_0^t \|u(x(\tau))\|_2^2 d\tau \leq \frac{2(f(x(0)) - f^*)}{t},$$

Remark 1 (Non-ergodic vs. ergodic rates and stationarity): The estimate in the theorem yields the *non-ergodic* rate

$$\min_{\tau \in [0, t]} \|u(x(\tau))\|_2^2 \leq \frac{2(f(x(0)) - f^*)}{t}.$$

By averaging, we also obtain the *ergodic* bound

$$\frac{1}{t} \int_0^t \|u(x(\tau))\|_2^2 d\tau \leq \frac{2(f(x(0)) - f^*)}{t}.$$

Consequently $\int_0^\infty \|u(x(t))\|_2^2 dt < \infty$, so there exists a sequence $t_k \rightarrow \infty$ with $\|u(x(t_k))\| \rightarrow 0$.

D. Safeguarded Adaptive Discretization

The proposed formulation (SS-QCQP(x)) enables *automatic step size selection* via a safeguarded Armijo-type line search that guarantees both feasibility and descent at every iteration. Starting from a feasible point x_0 , we discretize the ODE $\dot{x}(t) = u(x(t))$ using the forward Euler method, where $u(x_k)$ is obtained by solving (SS-QCQP(x)) and $t_k > 0$ is chosen via a backtracking line search so that the next iterate

$$x^{(k+1)} = x^{(k)} + t^{(k)} u(x^{(k)})$$

satisfies

$$\begin{aligned} f(x^{(k+1)}) &\leq f(x^{(k)}) + \gamma \nabla f(x^{(k)})^\top (x^{(k+1)} - x^{(k)}), \\ g_i(x^{(k+1)}) &\leq 0, \quad \forall i \in [m], \end{aligned} \quad (16)$$

where $\gamma \in (0, 1)$ is fixed. The first inequality in (16) is the classical Armijo condition, ensuring sufficient descent in the objective. The rest guarantees *anytime feasibility*, guaranteeing that all iterates remain in the feasible set of (OPT).

The next lemma shows that, under Assumption 1, the backtracking procedure in (16) terminates in finitely many steps and returns a step size bounded away from zero.

Lemma 3: Suppose x_k is feasible for (OPT), and that $w_i \geq \underline{w}$ for all $i \in [m]$, for some $\underline{w} > 0$. Then, the backtracking conditions (16) are satisfied for all $t \in [0, \underline{t}]$, where

$$\underline{t} = \min \left\{ \frac{1}{\alpha}, \frac{2(1-\gamma)}{L_f}, \frac{2\underline{w}}{L_1}, \dots, \frac{2\underline{w}}{L_m} \right\}. \quad (17)$$

Proof: We first establish feasibility. For any $t \in [0, \underline{t}]$, it follows that $t \leq \min\{\frac{1}{\alpha}, \frac{2w_i}{L_i}\}$ for all $i = 1, \dots, m$. Using the L_i -smoothness of g_i and the feasibility constraint in (SS-QCQP(x)), we obtain

$$\begin{aligned} g_i(x + tu(x)) &\leq g_i(x) + t \nabla g_i(x)^\top u(x) + \frac{L_i t^2}{2} \|u(x)\|_2^2 \\ &\leq g_i(x) + t[-\alpha g_i(x) - w_i \|u(x)\|_2^2] + \frac{L_i t^2}{2} \|u(x)\|_2^2 \\ &= (1 - \alpha t) g_i(x) + \left(\frac{L_i t^2}{2} - w_i t \right) \|u(x)\|_2^2 \\ &\leq (1 - \alpha t) g_i(x) \leq 0, \end{aligned}$$

which establishes that $x + tu(x)$ is feasible for all $t \in [0, \underline{t}]$.

We now verify the Armijo condition. By the L_f -smoothness of f , we have

$$f(x + tu(x)) \leq f(x) + t \nabla f(x)^\top u(x) + \frac{L_f t^2}{2} \|u(x)\|_2^2.$$

From Theorem 1, the search direction satisfies $\nabla f(x)^\top u(x) \leq -\|u(x)\|_2^2$. Substituting this bound yields

$$\begin{aligned} f(x + tu(x)) &\leq f(x) + t \nabla f(x)^\top u(x) - \frac{L_f t^2}{2} \|u(x)\|_2^2 \\ &\leq f(x) + \left(\frac{L_f t^2}{2} - t \right) (-\nabla f(x)^\top u(x)). \end{aligned}$$

For $t \leq \frac{2(1-\gamma)}{L_f}$, the coefficient $(\frac{L_f t^2}{2} - t)$ is upper-bounded by γt , giving

$$f(x + tu(x)) \leq f(x) + \gamma t \nabla f(x)^\top u(x),$$

which confirms the Armijo condition. Thus, both conditions in (16) hold for all $t \in [0, \underline{t}]$. ■

In Lemma 3, we only require that the parameters w_i are lower bounded by some $\underline{w} > 0$. This condition guarantees a uniformly positive lower bound on the step size, unlike the pure QP case, where the step size can vanish, thus preventing infeasibility and stagnation during backtracking line search. At the same time, it allows for the adaptive selection of w_i , as discussed in Section V-B, enabling the algorithm to adjust the strength of the quadratic correction dynamically while preserving the theoretical convergence guarantees. The overall algorithm is summarized in Algorithm 1.

Algorithm 1 Safe Sequential QCQP (SS-QCQP) for solving (OPT)

- 1: **Input:** initial feasible point $x^{(0)}$ such that $g_i(x^{(0)}) \leq 0$ for all $i \in [m]$; parameters $\gamma \in (0, 1)$, $\underline{w} > 0$, $\alpha > 0$, and tolerance $\epsilon > 0$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: Choose adaptive parameters $\{w_i^{(k)} \geq \underline{w} : i \in [m]\}$.
 - 4: Compute the search direction $u = u(x^{(k)})$ by solving (SS-QCQP(x)) with $w_i = w_i^{(k)}$.
 - 5: **if** $\|u\|_2 \leq \epsilon$ **then**
 - 6: **return** $x^{(k)}$ ▷ Convergence achieved
 - 7: **end if**
 - 8: Initialize step size $t^{(k)} = 1$.
 - 9: **while** $\exists i \in [m]$ such that $g_i(x^{(k)} + t^{(k)}u) > 0$ **or** $f(x^{(k)} + t^{(k)}u) > f(x^{(k)}) + \gamma t^{(k)} \nabla f(x^{(k)})^\top u$ **do**
 - 10: $t^{(k)} := \frac{1}{2} t^{(k)}$ ▷ Backtracking line search
 - 11: **end while**
 - 12: Update $x^{(k+1)} := x^{(k)} + t^{(k)}u$.
 - 13: **end for**
-

In the next result, we show that the convergence rate of the discrete-time algorithm matches that of its continuous-time counterpart.

Theorem 3 (Convergence Guarantee of SS-QCQP): Let Assumption 1 hold for (OPT). Then, for all iterations $k = 0, 1, \dots$ generated by Algorithm 1, the following bound holds:

$$\min_{i=0, \dots, k} \|u(x^{(i)})\|_2^2 \leq \frac{f(x^{(0)}) - f^*}{\gamma \underline{t} (k+1)}, \quad (18)$$

where \underline{t} is the uniform lower bound on the step size established in Lemma 3.

Proof: Using Lemma 3, at each iteration k , the backtracking line search admits a step size satisfying $t^{(k)} \geq \underline{t}$. By the Armijo condition in (16) and the descent property $\nabla f(x^{(k)})^\top u(x^{(k)}) \leq -\|u(x^{(k)})\|_2^2$ from Theorem 1, we

obtain

$$\begin{aligned} f(x^{(k+1)}) &\leq f(x^{(k)}) + \gamma t^{(k)} \nabla f(x^{(k)})^\top u(x^{(k)}) \\ &\leq f(x^{(k)}) - \gamma t^{(k)} \|u(x^{(k)})\|_2^2 \\ &\leq f(x^{(k)}) - \gamma \underline{t} \|u(x^{(k)})\|_2^2. \end{aligned}$$

Summing over $i = 0, \dots, k$ and telescoping the inequality yields

$$\sum_{i=0}^k \gamma \underline{t} \|u(x^{(i)})\|_2^2 \leq f(x^{(0)}) - f(x^{(k+1)}) \leq f(x^{(0)}) - f^*.$$

Dividing both sides by $\gamma \underline{t} (k+1)$ gives

$$\begin{aligned} \min_{i=0, \dots, k} \|u(x^{(i)})\|_2^2 &\leq \frac{1}{k+1} \sum_{i=0}^k \|u(x^{(i)})\|_2^2 \\ &\leq \frac{f(x^{(0)}) - f^*}{\gamma \underline{t} (k+1)}, \end{aligned}$$

which establishes the result. ■

The bound in (18) implies a non-ergodic $O(1/k)$ rate of convergence, which directly parallels the continuous-time result in Theorem 3. By averaging over the iterates, one also obtains the corresponding *ergodic* bound

$$\frac{1}{k+1} \sum_{i=0}^k \|u(x^{(i)})\|_2^2 \leq \frac{f(x^{(0)}) - f^*}{\gamma \underline{t} (k+1)},$$

Consequently, $\sum_{k=0}^\infty \|u(x^{(k)})\|_2^2 < \infty$, which implies $\|u(x^{(k)})\| \rightarrow 0$ along a subsequence.

Remark 2 (Comparison with the MBA method [29]): The design of the proposed SS-QCQP algorithm is conceptually inspired by the Moving Balls Approximation (MBA) method [29], which constructs locally feasible quadratic models for nonlinear constrained problems. MBA iteratively solves

$$\begin{aligned} \min_y & f(x) + \nabla f(x)^\top (y - x) + \frac{L_f}{2} \|y - x\|_2^2 \\ \text{s.t. } & g_i(x) + \nabla g_i(x)^\top (y - x) + \frac{L_i}{2} \|y - x\|_2^2 \leq 0, \quad \forall i \in [m]. \end{aligned} \quad (19)$$

The key distinction lies in how feasibility is enforced. MBA builds a conservative local approximation using Lipschitz upper bounds (see (1)) so that the solution to (19) is guaranteed to be feasible and can be directly adopted as the next iterate. In contrast, (SS-QCQP(x)) computes a feasible *search direction* and updates the iterate through a safeguarded line search. Thus, unlike MBA, the SS-QCQP framework does not require prior knowledge of the Lipschitz constants L_f and L_i , which are often difficult to estimate in practice. We note that extensions of MBA incorporating line search have been proposed to relax this requirement [30], but each line-search step requires solving a convex QCQP, making these variants computationally more expensive than both SS-QCQP and the standard MBA method.

IV. SCALABILITY TO MANY CONSTRAINTS

The main computational bottleneck of the SS-QCQP algorithm lies in solving the QCQP subproblem (SS-QCQP(x)) at each iteration. When the original problem (OPT) contains

a large number of inequality constraints ($m \gg 1$), solving this QCQP with all constraints active can be prohibitively expensive. However, in many practical applications, only a small subset of constraints is active or nearly active at optimality [38]. Thus, enforcing all m constraints in every subproblem is overly conservative.

To improve scalability, we introduce a *specialized active-set strategy* that enforces forward invariance only with respect to constraints that are nearly active. This reduces the size of the QCQP without compromising theoretical guarantees. In particular, we prove that the active-set variant preserves the $O(1/k)$ convergence rate established in Theorem 3 while significantly lowering the per-iteration computational cost.

A. Active-Set Strategy

We define the δ -active set at a point x as

$$\mathcal{A}_\delta(x) = \{i \in [m] \mid g_i(x) \geq -\delta\} \quad (20)$$

where $\delta > 0$ is a prescribed threshold. We also define $\mathcal{A}(x)$ to be any set such that $\mathcal{A}_\delta(x) \subseteq \mathcal{A}(x) \subseteq [m]$. The particular choice of $\mathcal{A}(x)$ will be discussed in Section V-C.

At each iteration, instead of considering all m constraints, we propose to compute the search direction by solving a reduced QCQP involving only the nearly active constraints:

$$\begin{aligned} \hat{u}(x) &:= \arg \min_u \frac{1}{2} \|u + \nabla f(x)\|_2^2 \\ \text{s.t. } \nabla g_i(x)^\top u &\leq -\alpha g_i(x) - w_i \|u\|_2^2, \quad i \in \mathcal{A}(x). \end{aligned} \quad (21)$$

Next, we will show that the direction $\hat{u}(x)$ inherits the key properties established in Theorem 1. Hence, the active-set SS-QCQP maintains the same theoretical guarantees as the full SS-QCQP while achieving substantially improved scalability.

Theorem 4: Let Assumption 1 hold, and let x be a feasible point of (OPT). Then,

- 1) If $\hat{u}(x) = 0$, then x is a stationary point of (OPT).
- 2) The direction $\hat{u}(x)$ satisfies the descent condition

$$\nabla f(x)^\top \hat{u}(x) \leq -\|\hat{u}(x)\|_2^2. \quad (22)$$

Proof: Since problem (SS-QCQP(x)) is strictly feasible, it follows that the reduced subproblem (21) is also strictly feasible. Hence, by the KKT optimality conditions [35], there exists a multiplier vector $\lambda(x) \geq 0$ satisfying

$$\hat{u}(x) + \nabla f(x) + \sum_{i \in \mathcal{A}(x)} \lambda_i(x) (\nabla g_i(x) + 2w_i \hat{u}(x)) = 0, \quad (23a)$$

$$\nabla g_i(x)^\top \hat{u}(x) \leq -\alpha g_i(x) - w_i \|\hat{u}(x)\|_2^2, \quad \forall i \in \mathcal{A}(x), \quad (23b)$$

$$\lambda_i(x) \geq 0, \quad \forall i \in \mathcal{A}(x), \quad (23c)$$

$$\lambda_i(x) (\nabla g_i(x)^\top \hat{u}(x) + \alpha g_i(x) + w_i \|\hat{u}(x)\|_2^2) = 0, \quad \forall i \in \mathcal{A}(x). \quad (23d)$$

Comparing the KKT conditions in (23) with those of (2), we observe that when $\hat{u}(x) = 0$, the pair (x, λ) satisfies the original KKT system (2), where

$$\lambda_i = \begin{cases} \lambda_i(x), & \text{if } i \in \mathcal{A}(x), \\ 0, & \text{otherwise.} \end{cases} \quad (24)$$

Therefore, x is a stationary point of (OPT). The proof of the descent property (22) follows identically to the argument presented in Theorem 1. ■

Having defined $\hat{u}(x)$ in (21), we now consider the corresponding iterative scheme

$$x^{(k+1)} = x^{(k)} + t^{(k)} \hat{u}(x^{(k)}), \quad (25)$$

where the step size $t^{(k)} > 0$ is selected using the same backtracking conditions as in (16), ensuring that each new iterate $x^{(k+1)}$ remains feasible and achieves sufficient descent. To guarantee the existence of a uniform lower bound on the accepted step sizes, we impose the following mild assumption.

Assumption 2: The feasible set \mathcal{F} is compact.

Under this assumption, the reduced QCQP directions remain uniformly bounded as we show next.

Lemma 4: Suppose (OPT) satisfies Assumption 1 and Assumption 2, and $w_i \geq \underline{w}$ for some $\underline{w} > 0$ and all $i \in [m]$. Then, the set $\{\hat{u}(x) : x \in \mathcal{F}\}$ is bounded.

Proof: Fix any $x \in \mathcal{F}$ and define, for each $i \in [m]$,

$$\Omega_i(x) \triangleq \{u \in \mathbb{R}^n : w_i \|u\|_2^2 + \nabla g_i(x)^\top u + \alpha g_i(x) \leq 0\}.$$

Dividing both sides by w_i and completing the square, $\Omega_i(x)$ can be represented as

$$\{u \in \mathbb{R}^n : \|u + \frac{1}{2w_i} \nabla g_i(x)\|_2^2 \leq \frac{-\alpha}{w_i} g_i(x) + \|\frac{1}{2w_i} \nabla g_i(x)\|_2^2\}.$$

Since g_i is continuously differentiable and \mathcal{F} is compact, both $g_i(x)$ and $\nabla g_i(x)$ are bounded on \mathcal{F} . Hence, there exist finite constants

$$M_i \triangleq \max_{x \in \mathcal{F}} |g_i(x)|, \quad G_i \triangleq \max_{x \in \mathcal{F}} \|\nabla g_i(x)\|_2.$$

Using $w_i \geq \underline{w}$ and upper bounds on $g_i(x)$ and $\nabla g_i(x)$, we can write

$$\|u + \frac{1}{2w_i} \nabla g_i(x)\|_2^2 \leq R_i^2$$

where $R_i^2 \triangleq \frac{\alpha M_i}{\underline{w}} + \left(\frac{G_i}{2\underline{w}}\right)^2$ is independent of x . Therefore, for all $x \in \mathcal{F}$,

$$\Omega_i(x) \subseteq \bar{\mathcal{B}}\left(-\frac{1}{2\underline{w}} \nabla g_i(x), R_i\right),$$

where $\bar{\mathcal{B}}(c, r)$ denotes the closed Euclidean ball centered at c with radius r . Because \mathcal{F} is compact and ∇g_i is continuous, the union $\Omega_i \triangleq \bigcup_{x \in \mathcal{F}} \Omega_i(x)$ is bounded. Finally, since $\hat{u}(x)$ satisfies all active constraints,

$$\hat{u}(x) \in \bigcap_{i \in \mathcal{A}(x)} \Omega_i(x) \subseteq \bigcap_{i \in \mathcal{A}(x)} \Omega_i,$$

implying that $\{\hat{u}(x) : x \in \mathcal{F}\}$ is bounded. ■

Lemma 5: Suppose (OPT) satisfies Assumption 1, and let x be a feasible point of (OPT). Assume $\gamma \in (0, 1)$ and $w_i \geq \underline{w}$ for all $i \in [m]$, for some $\underline{w} > 0$. Then there exists a constant $M > 0$ such that the backtracking conditions in Equation (16) hold for all $t \in [0, t]$, where

$$t = \min \left\{ \frac{2(1-\gamma)}{L_f}, 1, \frac{\delta}{M}, \frac{1}{\alpha}, \frac{2\underline{w}}{\max_{i \in [m]} L_i} \right\}. \quad (26)$$

Proof: From Theorem 4, we have $\nabla f(x)^\top \hat{u}(x) \leq -\|\hat{u}(x)\|_2^2$. Hence, for all $t \in [0, \frac{2(1-\gamma)}{L_f}]$, the Armijo condition is satisfied, i.e.,

$$f(x + t\hat{u}(x)) \leq f(x) + t\nabla f(x)^\top \hat{u}(x) + \frac{L_f t^2}{2} \|\hat{u}(x)\|_2^2 \quad (27)$$

$$\leq f(x) + \gamma t \nabla f(x)^\top \hat{u}(x).$$

The proof follows the same steps as in Lemma 3 for the full QCQP case.

For feasibility, consider two cases.

Case 1: $i \in \mathcal{A}(x)$. Then, by the constraint in (21), $g_i(x + t\hat{u}(x)) \leq 0$ for all $t \in [0, \min\{\frac{1}{\alpha}, \frac{2w}{L_i}\}]$, as in Lemma 3.

Case 2: $i \notin \mathcal{A}(x)$. Since $\mathcal{A}_\delta(x) \subseteq \mathcal{A}(x)$, we have $g_i(x) \leq -\delta$. Define

$$s_i(t) \triangleq g_i(x) + t\nabla g_i(x)^\top \hat{u}(x) + \frac{L_i t^2}{2} \|\hat{u}(x)\|_2^2,$$

which is a quadratic upper bound on $g_i(x + t\hat{u}(x))$. The largest nonnegative root η_i of $s_i(t) = 0$ is given by

$$\eta_i = \frac{-\nabla g_i(x)^\top \hat{u}(x) + \sqrt{(\nabla g_i(x)^\top \hat{u}(x))^2 - 2L_i g_i(x) \|\hat{u}(x)\|_2^2}}{L_i \|\hat{u}(x)\|_2^2}.$$

By construction, $s_i(0) = g_i(x) \leq -\delta < 0$, $s_i(\eta_i) = 0$, and for all $t \in [0, \eta_i]$,

$$g_i(x + t\hat{u}(x)) \leq s_i(t) \leq 0.$$

Applying the mean value theorem, there exists $\theta_i \in [0, \eta_i]$ such that

$$\begin{aligned} -g_i(x) &= s_i(\eta_i) - s_i(0) = s'_i(\theta_i)\eta_i \\ &= \eta_i [\nabla g_i(x)^\top \hat{u}(x) + L_i \theta_i \|\hat{u}(x)\|_2^2] \\ &\leq \eta_i [|\nabla g_i(x)^\top \hat{u}(x)| + L_i \theta_i \|\hat{u}(x)\|_2^2]. \end{aligned}$$

If $\eta_i \leq 1$, then since $\theta_i \leq \eta_i$, we can write

$$-g_i(x) \leq \eta_i [|\nabla g_i(x)^\top \hat{u}(x)| + L_i \|\hat{u}(x)\|_2^2].$$

Since $-g_i(x) \geq \delta$, we have

$$\eta_i \geq \frac{\delta}{|\nabla g_i(x)^\top \hat{u}(x)| + L_i \|\hat{u}(x)\|_2^2}.$$

Since $\hat{u}(x)$ and $\nabla g_i(x)$ are bounded over the compact feasible set \mathcal{F} , there exists a constant $M > 0$ such that

$$|\nabla g_i(x)^\top \hat{u}(x)| + L_i \|\hat{u}(x)\|_2^2 \leq M, \quad \forall x \in \mathcal{F}, i \in [m].$$

Hence, $\eta_i \geq \frac{\delta}{M}$. Eliminating the temporary assumption $\eta_i \leq 1$, we obtain

$$\eta_i \geq \min \left\{ 1, \frac{\delta}{M} \right\}.$$

Combining all cases, we conclude that:

- The Armijo condition holds for all $t \in [0, \frac{2(1-\gamma)}{L_f}]$;
- For all $i \in \mathcal{A}(x)$, $g_i(x + t\hat{u}(x)) \leq 0$ for $t \in [0, \min\{\frac{1}{\alpha}, \frac{2w}{L_i}\}]$;
- For all $i \notin \mathcal{A}(x)$, $g_i(x + t\hat{u}(x)) \leq 0$ for $t \in [0, \min\{1, \frac{\delta}{M}\}]$.

Therefore, the backtracking conditions (16) hold for all $t \in [0, \underline{t}]$, where \underline{t} is given by (26). ■

Theorem 4 and Lemma 5 prove similar results to Theorem 1 and Lemma 3, respectively, enabling us to extend Algorithm 1

with active set strategy while preserving theoretical convergence guarantees. The algorithm is summarized in Algorithm 2.

Algorithm 2 SS-QCQP with Active Set Strategy (SS-QCQP-AS) for solving (OPT)

- 1: **Input:** initial feasible point $x^{(0)}$ such that $g_i(x^{(0)}) \leq 0$ for all $i \in [m]$; parameters $\delta > 0$, $\gamma \in (0, 1)$, $w > 0$, $\alpha > 0$, and tolerance $\epsilon > 0$.
- 2: **for** $k = 0, 1, 2, \dots$ **do**
- 3: Choose adaptive parameters $\{w_i^{(k)} \geq w : i \in [m]\}$.
- 4: Choose active set such that $\mathcal{A}_\delta(x) \subseteq \mathcal{A}(x)$ \triangleright Active set strategy
- 5: Compute the search direction $u = \hat{u}(x^{(k)})$ by solving (21) with $w_i = w_i^{(k)}$.
- 6: **if** $\|u\|_2 \leq \epsilon$ **then**
- 7: **return** $x^{(k)}$ \triangleright Convergence achieved
- 8: **end if**
- 9: Initialize step size $t^{(k)} = 1$.
- 10: **while** $\exists i \in [m]$ such that $g_i(x^{(k)} + t^{(k)}u) > 0$ **or** $f(x^{(k)} + t^{(k)}u) > f(x^{(k)}) + \gamma t^{(k)} \nabla f(x^{(k)})^\top u$ **do**
- 11: $t^{(k)} := \frac{1}{2} t^{(k)}$ \triangleright Backtracking line search
- 12: **end while**
- 13: Update $x^{(k+1)} := x^{(k)} + t^{(k)}u$.
- 14: **end for**

Theorem 5: Let Assumption 1 and Assumption 2 hold for (OPT). Then, there exists a constant $M > 0$ such that, for all iterations $k = 0, 1, 2, \dots$ generated by Algorithm 1, the following bound holds:

$$\min_{i=0, \dots, k} \|u(x^{(i)})\|_2^2 \leq \frac{f(x^{(0)}) - f^*}{\gamma \underline{t}(k+1)}, \quad (28)$$

where

$$\underline{t} = \min \left\{ \frac{2(1-\gamma)}{L_f}, 1, \frac{\delta}{M}, \frac{1}{\alpha}, \frac{2w}{\max_{i \in [m]} L_i} \right\}. \quad (29)$$

Proof: Given a lower bound on step size from Lemma 5, the proof is the same as in Theorem 3. ■

V. PRACTICAL IMPLEMENTATION

A. Efficient Computation of the Search Direction

The optimization problem (SS-QCQP(x)) (or its reduced form (21)) is convex and can be solved efficiently using modern conic solvers such as ECOS [39] and Clarabel [40], after reformulating the quadratic constraints as second-order cone (SOC) constraints. While these solvers are robust and well optimized, the efficiency of each iteration in SS-QCQP depends critically on how the problem is expressed in conic form. To improve practical performance, we exploit a structural feature of (SS-QCQP(x)), where both the objective and the constraints are quadratic functions with identity Hessians. For simplicity, we describe the formulation when all m constraints are included in the subproblem; the same approach applies directly to the reduced QCQP (21).

Introducing an auxiliary variable s , problem (SS-QCQP(x)) can be equivalently written as

$$\begin{aligned} \hat{u}(x), \hat{s}(x) = \arg \min_{u, s} \quad & \frac{1}{2} \|u + \nabla f(x)\|_2^2 \\ \text{s.t.} \quad & \nabla g_i(x)^\top u + w_i s + \alpha g_i(x) \leq 0, \quad \forall i \in [m], \\ & \|u\|_2^2 \leq s. \end{aligned} \quad (30)$$

See Lemma 6 below for proof of equivalence. Although the transformation (30) is conceptually straightforward, it significantly simplifies the problem representation for off-the-shelf conic solvers. For example, Clarabel [40] solves generic conic programs of the form

$$\begin{aligned} \min_{x, s} \quad & \frac{1}{2} x^\top P x + q^\top x \\ \text{s.t.} \quad & A x + s = b, \\ & s \in \mathcal{K}, \end{aligned} \quad (31)$$

where \mathcal{K} denotes a convex cone (e.g., a second-order cone). When modeling (SS-QCQP(x)) through high-level interfaces such as CVXPY [41], the auxiliary-variable form (30) leads to a sparser coefficient matrix A and a more compact representation, yielding consistent reductions in solver runtime across iterations.

Lemma 6: Problems (SS-QCQP(x)) and (30) are equivalent in the sense that, if $(\hat{u}(x), \hat{s}(x))$ is an optimal solution of (30), then $\hat{u}(x)$ is the unique optimal solution of (SS-QCQP(x)).

Proof: Let p_1^* and p_2^* denote the optimal values of (SS-QCQP(x)) and (30), respectively. For any (u, s) feasible to (30), the first constraint implies that u is feasible for (SS-QCQP(x)). Since both problems share the same objective, it follows that $p_1^* \leq p_2^*$. Conversely, for any u feasible to (SS-QCQP(x)), the pair $(u, \|u\|_2^2)$ is feasible for (30) and yields the same objective value, implying $p_2^* \leq p_1^*$. Thus, $p_1^* = p_2^*$.

Consequently, if $(\hat{u}(x), \hat{s}(x))$ is an optimal solution of (30), then $\hat{u}(x)$ attains the same optimal value in (SS-QCQP(x)), and hence is optimal for that problem. Since the objective of (SS-QCQP(x)) is strongly convex, the solution $\hat{u}(x)$ is unique. ■

B. Hyperparameter Choice

Another important practical aspect of the proposed SS-QCQP algorithm is the selection of the parameters α and $\{w_i\}_{i=1}^m$ in (SS-QCQP(x)). As illustrated in Figure 1, when w_i is small, the direction $\hat{u}(x)$ becomes nearly tangent to the constraint boundary, leading to a small step size required to maintain feasibility. Conversely, when w_i is large, $\hat{u}(x)$ deviates significantly from the negative gradient direction, resulting in a smaller decrease in the objective value. Therefore, the choice of w_i critically affects the empirical performance of Algorithm 1.

Inspired by the MBA method [29], we fix $\alpha = 1$ and set $w_i \approx \frac{L_i}{2}$. However, since L_i is typically unknown in practice, we adaptively estimate it during iterations. We initialize each $w_i^{(0)} = \underline{w}$ with a small positive value (e.g., 10^{-3}), and update

it according to

$$w_i^{(k+1)} = \max \left\{ w_i^{(k)}, \frac{\|\nabla g_i(x^{(k+1)}) - \nabla g_i(x^{(k)})\|_2}{2\|x^{(k+1)} - x^{(k)}\|_2} \right\}, \quad \forall i \in [m]. \quad (32)$$

Since the sequence $\{w_i^{(k)}\}$ is nondecreasing, the convergence guarantees established in Theorem 3 and Theorem 5 remain valid.

C. Active Set Choice

In Lemma 5 and Theorem 5, we only required that $\mathcal{A}_\delta(x) \subseteq \mathcal{A}(x)$, allowing flexibility in the choice of \mathcal{A} . A natural and straightforward option is to set $\mathcal{A}(x) = \mathcal{A}_\delta(x)$. However, with this choice, even small perturbations in x can cause significant changes in $\mathcal{A}_\delta(x)$, which in turn may induce large variations in the direction $\hat{u}(x)$. Such discontinuities can produce zigzagging trajectories and degrade the empirical performance of the algorithm. To improve stability, we propose to use $\mathcal{A}(x) = \mathcal{A}_\delta(x) \cup \mathcal{T}_q(x)$, where

$$\mathcal{T}_q(x) = \{i : g_i(x) \text{ is among top } q\% \text{ of } \{g_i(x) : i \in [m]\}\}. \quad (33)$$

This construction smooths transitions in $\mathcal{A}(x)$ by including a small fraction of nearly active constraints, effectively balancing scalability and stability.

VI. NUMERICAL EXPERIMENTS

In this section, we demonstrate the effectiveness of the proposed SS-QCQP algorithm on an optimal control problem involving multi-agent navigation. The objective is to steer four vehicles from their initial positions to their designated goal locations while avoiding collisions with both static obstacles and other vehicles. The problem setup is similar to that in [42].

The problem is summarized as follows,

$$\begin{aligned} \min_{X, U} \quad & \sum_{t=0}^{N-1} l(X(t), U(t)) + V(X(N)) \\ \text{s.t.} \quad & U_i(t) \in \mathcal{U} \quad \forall t \in [N-1], i \in [4], \\ & X_i(t) \in \mathcal{X} \quad \forall t \in [N], i \in [4], \\ & X_i(t+1) = f(X_i(t), U_i(t)) \quad \forall t \in [N-1], i \in [4], \\ & X_i(0) = X_i^s \quad \forall i \in [4] \\ & g_j(X_i(t)) \leq 0 \quad \forall i \in [4], j \in [3], t \in [N], \end{aligned} \quad (34)$$

where $X_i(t) = [x_i(t), y_i(t), \theta_i(t)]^\top$ is the state of the i -th car at time step t , consisting of the x -position, y -position, and orientation, and $U_i(t) = [v_i(t), w_i(t)]^\top$ is its control input, consisting of the linear and angular velocities. The discrete-time dynamics, shared by all cars, are given by

$$f(X_i(t), U_i(t)) = \begin{bmatrix} x_i(t) + v_i(t)T \cos \theta_i(t) - \Delta x \\ y_i(t) + v_i(t)T \sin \theta_i(t) \\ \theta_i(t) + w_i(t)T \end{bmatrix}.$$

where $T = \Delta x = 0.03$. Here, $X(t) = [X_i(t)^\top]_{i \in [4]}^\top$ stacks the states of all four cars at time t , and $X = [X(t)^\top]_{t \in [N]}$ and $U = [U(t)^\top]_{t \in [N-1]}$ denote the stacked states and control inputs over the entire horizon.

The initial and target states are specified as follows. The starting state of the i -th vehicle is denoted by X_i^s , with

$$\begin{aligned} X_1^s &= [-2, -2, 0]^\top, & X_2^s &= [-3, -1, 0]^\top, \\ X_3^s &= [-3, -3, 0]^\top, & X_4^s &= [-1, -3, 0]^\top, \end{aligned}$$

and the corresponding target (desired) states are

$$\begin{aligned} X_1^d &= [2, 3, 0]^\top, & X_2^d &= [3, 2, 0]^\top, \\ X_3^d &= [2, 1, 0]^\top, & X_4^d &= [1, 2, 0]^\top. \end{aligned}$$

The desired steady-state input is $U_i^d = [1, 0]^\top$ for all $i \in [4]$. The stage cost is defined as

$$l(X(t), U(t)) = \sum_{i \in [4]} (\|X_i(t) - X_i^d\|_2^2 + 0.01 \|U_i(t) - U_i^d\|_2^2),$$

and the terminal cost is

$$V(X(N)) = \sum_{i \in [4]} \|X_i(N) - X_i^d\|_{P_i}^2,$$

where each P_i is obtained from the discrete-time algebraic Riccati Equation

$$A_i^\top P_i A_i - P_i - (A_i^\top P_i B_i)(R + B_i^\top P_i B_i)^{-1} (B_i^\top P_i A_i) + Q = 0$$

with $A_i = \nabla_1 f(X_i^d, U_i^d)$, $B_i = \nabla_2 f(X_i^d, U_i^d)$, $Q = I$ and $R = 0.01I$. The state and input constraints are defined as

$$\begin{aligned} \mathcal{X} &= \{(x, y, \theta) : |x| \leq 3.7, |y| \leq 3.7, |\theta| \leq \pi\}, \\ \mathcal{U} &= \{(v, w) : -5 \leq v \leq 12, |w| \leq \tfrac{3}{2}\pi\}. \end{aligned}$$

Obstacle avoidance is enforced through the constraint

$$g_j(X_i(t)) = r_j^2 - \|X_i(t) - c_j\|_2^2 \leq 0,$$

which models circle obstacles centered at $c_1 = [-1, -1]^\top$, $c_2 = [1, 0]^\top$, $c_3 = [0, 1]^\top$ with radii $r_1 = 1$, $r_2 = 0.5$, $r_3 = 0.5$.

By unrolling the dynamics from the starting state and applying the control inputs, (34) can be transformed into an optimization problem in form of (OPT). We set the planning horizon to $N = 40$, resulting in a total number of 320 variables and 2320 constraints.

Algorithm 1 and Algorithm 2 were implemented in CasADi [43], and the corresponding subproblems (SS-QCQP(x)) and (21) were solved using Clarabel [40] through the CVXPY [41] interface. In both methods, the parameters were initialized as $w_i = 10^{-3}$ and adaptively updated according to (32). For the active-set strategy, we used $\mathcal{A}(x) = \mathcal{A}_{0.5}(x) \cup \mathcal{T}_5(x)$, where $\mathcal{A}_\delta(x)$ and $\mathcal{T}_q(x)$ are defined in (20) and (33), respectively. The experiments are conducted on a 2020 MacBook Pro equipped with an Apple M1 CPU and 16 GB of memory.

We compared SS-QCQP and SS-QCQP-AS against IPOPT [44] (as called from CasADi) and CasADi's built-in SQP solver. Figure 2 illustrates the position trajectories of the four agents obtained using SS-QCQP, SS-QCQP-AS, IPOPT, and SQP. The dotted lines, stars, and triangles represent the trajectories, initial positions, and target positions of the agents, respectively, while the gray circles denote obstacles. It can be observed that Algorithm 1, Algorithm 2, and SQP yield similar near-optimal trajectories. Figure 3a further depicts the

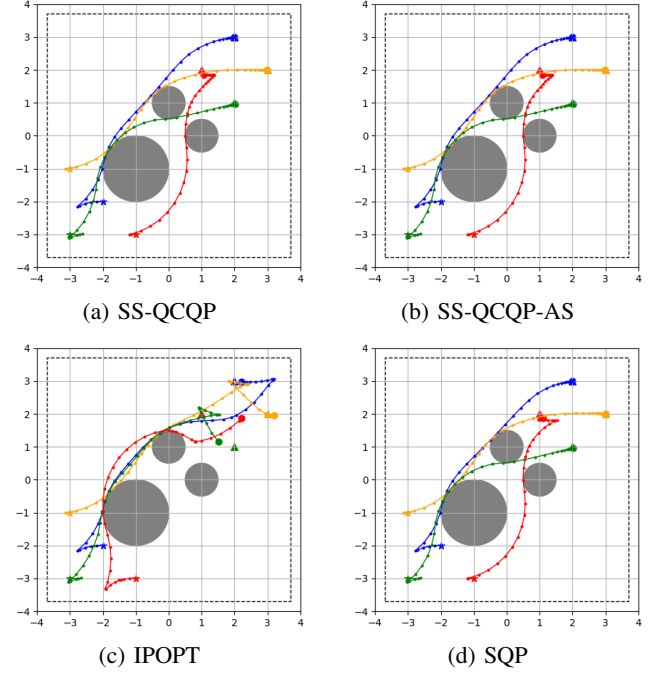


Fig. 2: Position trajectories of four agents obtained using SS-QCQP, SS-QCQP-AS, IPOPT, and SQP.

pairwise distances among the four agents over the planning horizon obtained using SS-QCQP.

Figure 3b and Figure 3c plot, for each algorithm, the objective value and the maximum constraint violation across iterations, respectively. Furthermore, Figure 3d shows the minimum squared norm of the search directions across iterations, which converges at the theoretical rate of $O(1/k)$ established in Theorems 3 and 5. Finally, Figure 4 shows the number of nearly active constraints at each iteration, illustrating that fewer than 10% of all constraints are nearly active and enter the QCQP subproblem. The runtime of Algorithm 1 is 186.53 seconds, whereas Algorithm 2 achieves a comparable solution in 49.47 seconds.

It can be observed that while the SQP method decreases the objective value more rapidly than both SS-QCQP and SS-QCQP-AS, it exhibits small constraint violations before convergence. In contrast, both SS-QCQP and SS-QCQP-AS remain feasible throughout all iterations, albeit with a slower decrease in the objective value. This property enables early termination with feasibility guarantees. It is worth noting that both SQP and IPOPT are second-order methods (in this experiment, CasADi's built-in SQP employs an L-BFGS Hessian approximation), whereas SS-QCQP and SS-QCQP-AS are first-order algorithms that achieve comparable solution quality while ensuring feasibility at every iteration.

VII. CONCLUSION

This paper introduced the Safe Sequential QCQP (SS-QCQP) algorithm, a first-order framework for solving inequality-constrained nonlinear optimization problems with guaranteed feasibility at every iteration. The method is derived from a continuous-time dynamical system whose vector field

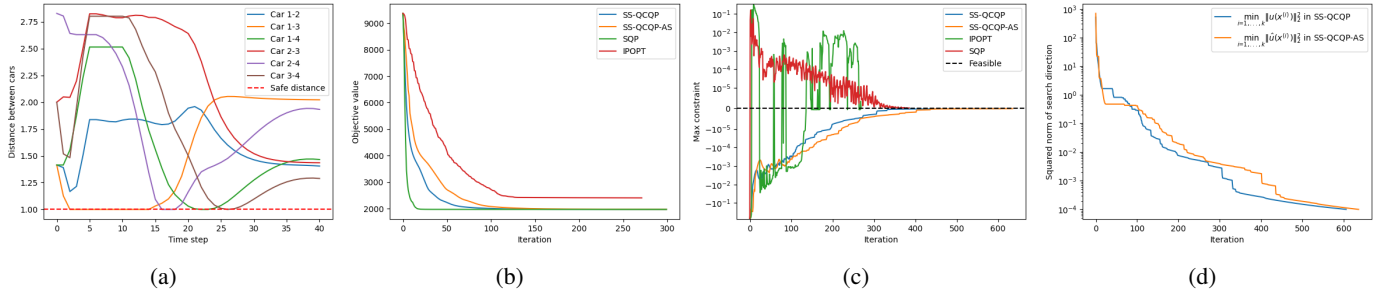


Fig. 3: (a): Pairwise distances among the four agents over the planning horizon solved by SS-QCQP. (b): Objective of SS-QCQP, SS-QCQP-AS, IPOPT, and SQP along iterations. (c): Maximum constraint of SS-QCQP, SS-QCQP-AS, IPOPT, and SQP along iterations. (d): $\min_{i=1,\dots,k} \|u(x^{(i)})\|_2^2$ in SS-QCQP and $\min_{i=1,\dots,k} \|\hat{u}(x^{(i)})\|_2^2$ in SS-QCQP-AS along iterations.

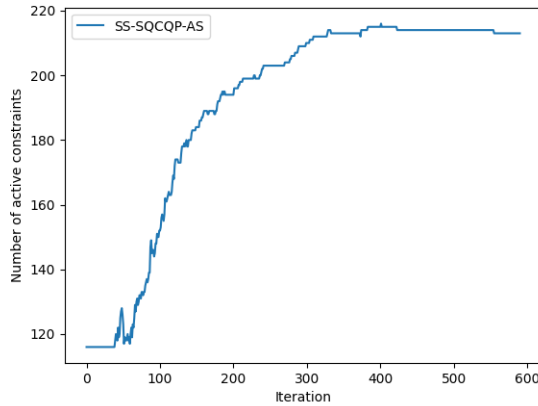


Fig. 4: Number of active constraints in Algorithm 2 along the iterations. The total number of constraints is 2320.

is obtained by solving a convex quadratically constrained quadratic program (QCQP) that enforces strict descent and forward invariance of the feasible set. A safeguarded discretization with adaptive step-size selection yields a discrete-time algorithm that maintains feasibility and achieves a provable $O(1/k)$ convergence rate.

To improve scalability, we developed an active-set variant (SS-QCQP-AS) that selectively enforces invariance for nearly active constraints, significantly reducing computational cost without compromising convergence guarantees. Extensive numerical experiments on a multi-agent optimal control problem confirmed that both SS-QCQP and SS-QCQP-AS maintain feasibility, exhibit the predicted convergence behavior, and deliver performance comparable to second-order solvers such as SQP and IPOPT.

Future work will explore the integration with model predictive control (MPC) schemes, as well as acceleration strategies using higher-order derivatives.

- [1] T. Erseghe, "Distributed optimal power flow using admm," *IEEE transactions on power systems*, vol. 29, no. 5, pp. 2370–2380, 2014.
- [2] M. Bagherzadeh, S. Savehshemshaki, and W. Lucia, "Guaranteed collision-free reference tracking in constrained multi unmanned vehicle systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 6, pp. 3083–3089, 2021.

- [3] P. M. Wensing, M. Posa, Y. Hu, A. Escande, N. Mansard, and A. Del Prete, "Optimization-based control for dynamic legged robots," *IEEE Transactions on Robotics*, vol. 40, pp. 43–63, 2023.
- [4] K. Nguyen, S. Schoedel, A. Alavilli, B. Plancher, and Z. Manchester, "Tinymc: Model-predictive control on resource-constrained microcontrollers," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 1–7.
- [5] A. L. Bishop, J. Z. Zhang, S. Gurumurthy, K. Tracy, and Z. Manchester, "Relu-qp: A gpu-accelerated quadratic programming solver for model-predictive control," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 13 285–13 292.
- [6] E. Dietrich, A. Devonport, and M. Arcak, "Nonconvex scenario optimization for data-driven reachability," in *6th Annual Learning for Dynamics & Control Conference*. PMLR, 2024, pp. 514–527.
- [7] A. S. Nemirovski and M. J. Todd, "Interior-point methods for optimization," *Acta Numerica*, vol. 17, pp. 191–234, 2008.
- [8] J. Nocedal, A. Wächter, and R. A. Waltz, "Adaptive barrier update strategies for nonlinear interior methods," *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1674–1693, 2009.
- [9] A. Wächter, *An interior point algorithm for large-scale nonlinear optimization with applications in process engineering*. Carnegie Mellon University, 2002.
- [10] E. R. Panier and A. L. Tits, "A superlinearly convergent feasible method for the solution of inequality constrained optimization problems," *SIAM Journal on control and Optimization*, vol. 25, no. 4, pp. 934–950, 1987.
- [11] K. Schittkowski, "On the convergence of a sequential quadratic programming method with an augmented lagrangian line search function," *Mathematische Operationsforschung und Statistik. Series Optimization*, vol. 14, no. 2, pp. 197–216, 1983.
- [12] E. R. Panier and A. L. Tits, "On combining feasibility, descent and superlinear convergence in inequality constrained optimization," *Mathematical programming*, vol. 59, no. 1, pp. 261–276, 1993.
- [13] N. Maratos, "Exact penalty function algorithms for finite dimensional and control optimization problems," Ph.D. dissertation, Imperial College London (University of London), 1978.
- [14] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Some theoretical properties of an augmented lagrangian merit function." Tech. Rep., 1986.
- [15] D. Q. Mayne and E. Polak, "A superlinearly convergent algorithm for constrained optimization problems," in *Algorithms for constrained minimization of smooth nonlinear functions*. Springer, 2009, pp. 45–61.
- [16] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "A trust region algorithm for nonlinearly constrained optimization," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1152–1170, 1987.
- [17] L. Numerow, A. Zanelli, A. Carron, and M. N. Zeilinger, "Inherently robust suboptimal mpc for autonomous racing with anytime feasible sqp," *IEEE Robotics and Automation Letters*, vol. 9, no. 7, pp. 6616–6623, 2024.
- [18] M. J. Tenny, S. J. Wright, and J. B. Rawlings, "Nonlinear model predictive control via feasibility-perturbed sequential quadratic programming," *Computational Optimization and Applications*, vol. 28, no. 1, pp. 87–121, 2004.
- [19] M. Fukushima, "A successive quadratic programming algorithm with global and superlinear convergence properties," *Mathematical Programming*, vol. 35, no. 3, pp. 253–264, 1986.

- [20] M. Fukushima, Z.-Q. Luo, and P. Tseng, "A sequential quadratically constrained quadratic programming method for differentiable convex minimization," *SIAM Journal on Optimization*, vol. 13, no. 4, pp. 1098–1119, 2003.
- [21] E. West and E. Polak, "A generalized quadratic programming-based phase i-phase ii method for inequality-constrained optimization," *Applied Mathematics and Optimization*, vol. 26, no. 3, pp. 223–252, 1992.
- [22] A. Allibhoy and J. Cortés, "Control-barrier-function-based design of gradient flows for constrained nonlinear programming," *IEEE Transactions on Automatic Control*, vol. 69, no. 6, pp. 3499–3514, 2023.
- [23] M. Muehlebach and M. I. Jordan, "On constraints in first-order optimization: a view from non-smooth dynamical systems," *J. Mach. Learn. Res.*, vol. 23, no. 1, Jan. 2022.
- [24] A. Raghunathan, J. Shamma, N. Li et al., "Constrained optimization from a control perspective via feedback linearization," *arXiv preprint arXiv:2503.12665*, 2025.
- [25] A. Allibhoy and J. Cortés, "Anytime solution of constrained nonlinear programs via control barrier functions," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 6527–6532.
- [26] C. Dawson, S. Gao, and C. Fan, "Safe control with learned certificates: A survey of neural lyapunov, barrier, and contraction methods for robotics and control," *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 1749–1767, 2023.
- [27] M. Muehlebach and M. I. Jordan, "On constraints in first-order optimization: A view from non-smooth dynamical systems," *arXiv preprint arXiv:2107.08225*, 2021.
- [28] G. Scutari, F. Facchinei, and L. Lampariello, "Parallel and distributed methods for constrained nonconvex optimization—part i: Theory," *IEEE Transactions on Signal Processing*, vol. 65, no. 8, pp. 1929–1944, 2016.
- [29] A. Auslender, R. Shefi, and M. Teboulle, "A moving balls approximation method for a class of smooth constrained minimization problems," *SIAM Journal on Optimization*, vol. 20, no. 6, pp. 3232–3259, 2010.
- [30] Z. Lu, "Sequential convex programming methods for a class of structured nonlinear programming," *arXiv preprint arXiv:1210.3039*, 2012.
- [31] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 2006.
- [32] 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, 2016.
- [33] G. Torrisi, S. Grammatico, R. S. Smith, and M. Morari, "A projected gradient and constraint linearization method for nonlinear model predictive control," *SIAM Journal on Control and Optimization*, vol. 56, no. 3, pp. 1968–1999, 2018.
- [34] J. Liu, "Sensitivity analysis in nonlinear programs and variational inequalities via continuous selections," *SIAM Journal on Control and Optimization*, vol. 33, no. 4, pp. 1040–1060, 1995.
- [35] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [36] F. Blanchini, S. Miani et al., *Set-theoretic methods in control*. Springer, 2008, vol. 78.
- [37] F. Blanchini, "Set invariance in control," *Automatica*, vol. 35, no. 11, pp. 1747–1767, 1999.
- [38] S. Misra, L. Roald, and Y. Ng, "Learning for constrained optimization: Identifying optimal active constraint sets," *INFORMS Journal on Computing*, vol. 34, no. 1, pp. 463–480, 2022.
- [39] A. Domahidi, E. Chu, and S. Boyd, "Ecos: An socp solver for embedded systems," in *2013 European control conference (ECC)*. IEEE, 2013, pp. 3071–3076.
- [40] P. J. Goulart and Y. Chen, "Clarabel: An interior-point solver for conic programs with quadratic objectives," *arXiv preprint arXiv:2405.12762*, 2024.
- [41] S. Diamond and S. Boyd, "Cvxpy: A python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.
- [42] J. Wu, L. Dai, S. Dou, and Y. Xia, "Accelerated successive convex approximation for nonlinear optimization-based control," *IEEE Transactions on Automatic Control*, 2025.
- [43] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.
- [44] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.