Constrained Feedback Control of Nonlinear Systems via Approximate HJB and Control Barrier Functions

Milad Alipour Shahraki and Laurent Lessard

***Abstract*— This paper presents a two-stage framework for constrained feedback control of input-affine nonlinear systems. Offline, an approximate value function for the unconstrained problem is computed, for example using Hamilton–Jacobi–Bellman (HJB)-based policy iteration. Online, the proposed quadratic program (QP) minimizes the pre-Hamiltonian evaluated using the approximate value-function gradient subject to safety constraints enforced by control barrier functions (CBFs). This architecture decouples performance optimization from constraint enforcement, allowing constraints to be modified without recomputing the value function. As in CBF-QP architectures based on control Lyapunov functions (CLFs), safety is enforced as a hard constraint; however, the performance objective targets approximate optimality rather than a prescribed Lyapunov decay. Numerical results on a linear 2-state hovercraft and a nonlinear 9-state spacecraft attitude-control problem show agreement with the constrained open-loop optimal control problem (OCP) benchmark in the linear case, and performance close to the OCP benchmark, improving on CLF-based controllers, in the nonlinear case.**

I. Introduction

The optimal feedback control of general nonlinear systems remains a central challenge in control theory. The Hamilton–Jacobi–Bellman (HJB) equation provides a theoretical characterization of the optimal solution, but its nonlinear partial differential equation (PDE) structure generally precludes analytical solutions, while the curse of dimensionality makes numerical approximation computationally prohibitive for high-dimensional systems [1]. Practical systems must also satisfy safety-critical constraints, such as actuator limits, state bounds, and geometric exclusion zones, further complicating optimal controller design. Model predictive control (MPC) can address both performance and constraints, but the need to solve a nonlinear optimization problem online can limit real-time applicability for fast or high-dimensional systems.

One alternative is to address optimality and safety separately. On the optimization side, approximate dynamic programming methods seek tractable approximations to the HJB equation. Successive Galerkin Approximation (SGA) [2] projects the generalized HJB equation onto a finite-dimensional basis, while sum-of-squares (SOS) methods [3], [4] replace polynomial nonnegativity conditions with tractable semidefinite relaxations. These approaches are often combined with policy iteration [5], [6] to successively improve the value-function approximation and associated feedback policy. However, they typically address unconstrained control problems; incorporating state and input constraints directly into the offline value-function computation can substantially increase computational complexity.

On the constraint enforcement side, control barrier functions (CBFs) [7] and their high-order extensions (HOCBFs) [8] convert suitable state constraints into affine inequalities in the control input, enabling real-time enforcement through quadratic programs (QPs). CBFs are commonly combined with control Lyapunov functions (CLFs) in CLF-CBF-QP architectures [9] to jointly address stability and safety. However, as observed in our prior work on CLF-CBF-QPs [10], CLF-based formulations introduce several design considerations. First, the CLF condition enforces stability rather than optimal performance and requires a suitable Lyapunov function, often hand-designed via input-output linearization [10]. Second, fixed CLF decay rates can induce input chattering at low sampling frequencies [10], [11]. Third, selecting the comparison function governing the prescribed decay introduces a feasibility-convergence trade-off that is further complicated by input constraints [12]. These considerations motivate replacing the prescribed CLF decay objective with one derived from optimal control.

Another line of work couples safety and performance by incorporating CBF conditions directly into the value-function computation [13]–[15]. However, because the constraints are embedded in the offline optimization, changing the safe set generally requires recomputing the value function, reducing the ability to adapt to real-time changes in constraints.

Our approach uses a two-stage architecture that separates performance optimization from safety enforcement. Offline, an approximate value function for the unconstrained problem is computed. Online, its gradient defines the objective of a CBF-QP that enforces safety and input constraints. The controller therefore depends only on the value-function gradient, not on how the approximation is obtained. Our main contributions are as follows.

1. We propose the *H-CBF-QP*, a convex QP that minimizes the approximate pre-Hamiltonian induced by a precomputed value function while enforcing CBF and input constraints. Unlike CLF-CBF-QP architectures [9], [10], which impose a prescribed Lyapunov decay condition that may be relaxed for safety, the H-CBF-QP favors the unconstrained value-function policy whenever feasible. This removes the need for a prescribed CLF decay rate and slack variable, as well as the input-output-linearization-based CLF construction.

M. Alipour Shahraki is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA alipourshahraki.m@northeastern.edu

L. Lessard is with the Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA l.lessard@northeastern.edu

2. We characterize the safety, approximate optimality, and local stability properties of the resulting closed loop. The CBF constraints provide forward-invariance guarantees, while the controller recovers the unconstrained value-function policy whenever the constraints are inactive. Under a regional HJB inequality, asymptotic stability is guaranteed on forward-invariant sublevel sets of the approximate value function contained in the inactive-constraint region.
3. We validate our proposed controller on a linear 2-state hovercraft and a nonlinear 9-state spacecraft attitude-control problem with reaction-wheel momentum and pointing constraints. The results agree with the constrained OCP benchmark in the linear example, achieve near-OCP performance under momentum constraints, and outperform CLF-CBF-QP benchmarks under both spacecraft constraints while maintaining CBF-based safety and real-time QP implementation.

The remainder of the paper is organized as follows. Section II formulates the problem and reviews HJB and CBF concepts. Section III develops the online H-CBF-QP and establishes its safety, approximate optimality, and stability properties. Section IV reviews offline value-function approximation methods based on policy iteration. Section V validates the framework on linear and nonlinear examples, and Section VI concludes with future research directions.

II. Problem Formulation

A. System Dynamics and Objectives

We consider a nonlinear *input-affine* control system

$$\dot{x} = f(x) + g(x)u, \qquad x(0) = x_0, \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the control input, and $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are locally Lipschitz continuous, with $f(0) = 0$. We aim to minimize the infinite-horizon cost

$$J(x_0, u(\cdot)) = \int_0^\infty \big(q(x) + u^\mathsf{T} R\, u\big)\, \mathrm{d}t \tag{2}$$

subject to the dynamics (1), state constraint $x(t) \in \mathcal{X} \subseteq \mathbb{R}^n$, and input constraint $u(t) \in \mathcal{U} \subseteq \mathbb{R}^m$, where $\mathcal{U}$ is a nonempty convex polyhedron that includes the origin. Here, $q : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ satisfies $q(0) = 0$ and $R \succ 0$.

B. Hamilton–Jacobi–Bellman (HJB) Equation

We first consider the *unconstrained* problem. Let $\mathcal{P}$ denote the set of positive definite, proper, C^1 functions. Define the optimal value function $V^*(x_0) := \inf_{u(\cdot)} J(x_0, u(\cdot))$. Assuming $V^* \in \mathcal{P}$, it satisfies the HJB equation

$$H^*(x, \nabla V^*(x)) := \min_{u\in\mathbb{R}^m} H(x, \nabla V^*(x), u) = 0, \tag{3}$$

where H^* denotes the Hamiltonian and the pre-Hamiltonian H is defined as:

$$H(x, \nabla V(x), u) := \nabla V(x)^\mathsf{T}\big(f(x)+g(x)u\big)+q(x)+u^\mathsf{T}Ru.$$

The optimal (unconstrained) feedback law is given by

$$u^*(x) := -\tfrac{1}{2}R^{-1}g(x)^\mathsf{T}\nabla V^*(x). \tag{4}$$

C. Safety via Control Barrier Functions (CBFs)

We briefly review the CBF and HOCBF tools used for online constraint enforcement. For a scalar function B, define the Lie derivatives $L_fB := \nabla B^\mathsf{T} f$ and $L_gB := \nabla B^\mathsf{T} g$, so that $\dot{B} = L_fB + L_gB\, u$ along (1).

***Definition** 1 (Relative Degree [8]):* A sufficiently differentiable function $B : \mathbb{R}^n \to \mathbb{R}$ has relative degree r with respect to (1) if $L_gL_f^kB = 0$ for $k = 0, \dots, r-2$ and $L_gL_f^{r-1}B \neq 0$.

***Definition** 2 (HOCBF [8]):* Let $B : \mathbb{R}^n \to \mathbb{R}$ have relative degree r. Using linear class-$\mathcal{K}$ functions with gains $\alpha_k > 0$, define

$$\psi_0 := B, \quad \psi_k := \dot{\psi}_{k-1} + \alpha_k\psi_{k-1}, \quad k = 1, \dots, r, \tag{5}$$

and $\mathcal{C}_k := \{x \mid \psi_k(x) \geq 0\}$. Then B is a HOCBF if $\sup_{u\in\mathcal{U}} \psi_r(x, u) \geq 0$ for all $x \in \bigcap_{k=0}^{r-1} \mathcal{C}_k$.

***Theorem** 1 (HOCBF Forward Invariance [8]):* If B is a HOCBF and $u(x) \in \mathcal{U}$ satisfies $\psi_r(x, u) \geq 0$ on $\bigcap_{k=0}^{r-1} \mathcal{C}_k$, then this set is forward invariant.

For a relative-degree-r function B, the HOCBF condition $\psi_r(x, u) \geq 0$ is affine in u and can therefore be imposed directly in a QP. In particular, for $r = 1$ it reduces to $L_fB + L_gB\, u \geq -\alpha_1 B$, while for $r = 2$ it becomes $L_f^2B + L_gL_fB\, u + (\alpha_1 + \alpha_2)L_fB + \alpha_1\alpha_2 B \geq 0$.

III. Controller Design

We now present the H-CBF-QP, which combines an approximate value function for the *unconstrained* problem (computed offline) with the CBF constraints of Section II-C. Any approximation method may be used; specific examples are given in Section IV.

Given the precomputed $\nabla\hat{V}$ and the current state x, we solve an online QP that minimizes the pre-Hamiltonian $H(x, \nabla\hat{V}(x), u)$ over u, subject to the state constraints (via the CBF/HOCBF) and input constraints:

$$\begin{aligned} \bar{u}(x) \;:=\; \operatorname*{argmin}_{u}\; & H(x, \nabla\hat{V}(x), u) \\ \text{s.t. }\; & \psi_r(x, u) \geq 0, \qquad \text{(CBF/HOCBF)} \\ & u \in \mathcal{U}. \qquad\quad \text{(Input constraint)} \end{aligned} \tag{6}$$

This is a pointwise optimization. Since $\nabla\hat{V}(x)^\mathsf{T} f(x)$ and $q(x)$ do not depend on u, we may equivalently minimize

$$J_{\mathrm{H}}(u) := \nabla\hat{V}(x)^\mathsf{T} g(x)\, u + u^\mathsf{T} R\, u. \tag{7}$$

To mirror the notation defined in (4), we let $\hat{u}$ denote the unconstrained minimizer of the pre-Hamiltonian $H(x, \nabla\hat{V}(x), u)$:

$$\hat{u}(x) := -\tfrac{1}{2}R^{-1}g(x)^\mathsf{T}\nabla\hat{V}(x). \tag{8}$$

A. Theoretical Properties

Under suitable assumptions, we can prove safety, approximate optimality, and asymptotic stability properties of our proposed controller. Our first result follows directly from strict convexity of (7).

***Proposition** 1 (Unconstrained Policy Recovery):* If $\hat{u}(x)$ is feasible for (6), then the QP solution satisfies $\bar{u}(x) = \hat{u}(x)$.

Thus, wherever the constraints are inactive, the online controller matches the unconstrained value-function policy. In this case, the deviation from the optimal policy depends on the deviation of the approximate value function gradient from the optimal one.

***Proposition** 2 (Safety and Approximate Optimality):* Suppose B is a relative degree-r HOCBF and $\mathcal{C}_k$ is given by Definition 2. If (6) is feasible for $x \in \bigcap_{k=0}^{r-1} \mathcal{C}_k$, then: (i) $\bigcap_{k=0}^{r-1} \mathcal{C}_k$ is forward invariant under the resulting policy; and (ii) on any compact set $\Omega \subset \mathbb{R}^n$ where $\hat{u}$ is feasible, there exists a constant $C > 0$ such that

$$\|\bar{u} - u^*\|_\Omega \leq C \left\|\nabla \hat{V} - \nabla V^*\right\|_\Omega,$$

where $\|\phi\|_\Omega := \sup_{x\in\Omega} \|\phi(x)\|$. Thus, $\|\bar{u} - u^*\|_\Omega \to 0$ as $\|\nabla \hat{V} - \nabla V^*\|_\Omega \to 0$.

Proof: Part (i) follows directly from Theorem 1, since feasibility of (6) ensures satisfaction of the HOCBF constraint. For part (ii), Proposition 1 gives $\bar{u} = \hat{u}$ on Ω. Using (4) and (8), $\bar{u} - u^* = -\frac{1}{2} R^{-1} g(x)^\mathsf{T} (\nabla \hat{V} - \nabla V^*)$. Taking the supremum norm over Ω gives the stated bound with $C = \frac{1}{2}\|R^{-1} g^\mathsf{T}\|_\Omega$. Since g is continuous and Ω is compact, C is finite, proving the convergence claim. ∎

Finally, if the approximate value function satisfies a local HJB inequality, the unconstrained policy also provides a local Lyapunov decrease. This yields an asymptotic stability guarantee for sublevel sets contained in the region where the constraints are inactive.

***Theorem** 2 (Local Asymptotic Stability):* Suppose $\hat{V}$ is C^1 and positive definite on a compact set Ω containing the origin, and satisfies $H(x, \nabla \hat{V}(x), \hat{u}(x)) \leq 0$ on Ω. Let q be positive definite. Define $\mathcal{I} \subseteq \mathbb{R}^n$ as the set of states where $\hat{u}(x)$ is strictly feasible for (6), and assume $\mathcal{I} \cap \Omega$ contains a neighborhood of the origin. Then, for any $c > 0$ such that $\mathcal{S} := \{x \mid \hat{V}(x) \leq c\} \subseteq \mathcal{I} \cap \Omega$, the set $\mathcal{S}$ is forward invariant and contained in the region of attraction of the origin.

Proof: On $\mathcal{I}$, Proposition 1 gives $\bar{u} = \hat{u}$. Therefore, $\dot{\hat{V}} = H(x, \nabla \hat{V}, \hat{u}) - q(x) - \hat{u}^\mathsf{T} R \hat{u} \leq -q(x)$ for $x \in \mathcal{I} \cap \Omega$, since $H(x, \nabla \hat{V}, \hat{u}) \leq 0$ and $R \succ 0$. Since $\mathcal{S} \subseteq \mathcal{I} \cap \Omega$ is a sublevel set of $\hat{V}$ and $\dot{\hat{V}} \leq 0$ throughout $\mathcal{S}$, the set $\mathcal{S}$ is forward invariant. Because $\hat{V}$ and q are positive definite, $\dot{\hat{V}} < 0$ for all $x \in \mathcal{S} \setminus \{0\}$. Hence, the origin is asymptotically stable and $\mathcal{S}$ is contained in its region of attraction. ∎

IV. Value Function Approximation

The H-CBF-QP requires only the gradient $\nabla \hat{V}$ of a precomputed value-function approximation. The safety guarantee is independent of the approximation quality. The stability result of Theorem 2 additionally requires $\hat{V}$ to be C^1 and positive definite on a compact approximation domain Ω containing the origin, with $H(x, \nabla \hat{V}(x), \hat{u}(x)) \leq 0$ on Ω.

We review two popular approaches: Successive Galerkin Approximation (SGA) [2], and sum-of-squares (SOS) [5]. Both approaches parameterize $\hat{V}(x) = \sum_i c_i m_i(x)$ using polynomial basis functions m_i and solve for the coefficients via *policy iteration*: alternating between policy evaluation and policy improvement steps. Both use the same policy improvement step: $\hat{u}^{(k+1)} = -\frac{1}{2} R^{-1} g^\mathsf{T} \nabla \hat{V}^{(k)}$. For policy evaluation, both methods work with the negative pre-Hamiltonian

$$L^{(k)}(x) := -H(x, \nabla \hat{V}^{(k)}(x), \hat{u}^{(k)}(x)),$$

over a compact domain Ω.

A. Successive Galerkin Approximation

SGA performs policy evaluation by seeking $L^{(k)} \approx 0$. This is accomplished by solving a system of orthogonality equations, which are linear in the coefficients c_i: $\langle L^{(k)}, m_i \rangle_\Omega = 0$, where the inner product is an integral over $x \in \Omega$. SGA is scalable to larger systems because each step only involves linear solves and the integrals may be pre-computed for the monomial bases. Under suitable completeness conditions, increasing the number of basis functions yields convergence to V^*. For a finite basis, intermediate iterates do not come with stability guarantees, unlike the SOS formulation below.

B. Sum-of-squares

SOS performs policy evaluation by minimizing $\int_\Omega \hat{V}^{(k)}(x)\,\mathrm{d}x$ subject to $L^{(k)} \geq 0$ and $\hat{V}^{(k)}$ positive definite on Ω. For a semialgebraic domain Ω, the regional (restricted to Ω) inequalities are certified using SOS multipliers. These conditions yield a semidefinite program in the coefficients c_i. Since $\hat{u}^{(k+1)}$ minimizes the pre-Hamiltonian,

$$H(x, \nabla \hat{V}^{(k)}, \hat{u}^{(k+1)}) \leq H(x, \nabla \hat{V}^{(k)}, \hat{u}^{(k)}) = -L^{(k)}(x) \leq 0$$

on Ω. Thus, each feasible iterate satisfies the regional HJB inequality and is positive definite on Ω, providing the Lyapunov conditions used in Theorem 2 on sublevel sets contained in Ω and guaranteeing local stability.

V. Application Examples

We validate the H-CBF-QP on two examples: a linear system with a known constrained-optimal benchmark (Example 1) and a nonlinear spacecraft system with relative-degree-one and relative-degree-two constraints (Example 2). Together, they demonstrate (i) the ability to modify constraints without recomputing the offline value function, (ii) cases in which the H-CBF-QP closely reproduces the constrained-optimal trajectory, and (iii) the performance gap that can arise when the constrained optimum requires anticipation not encoded in the unconstrained value function. In both examples, the running state cost is a positive definite quadratic, $q(\xi) := \xi^\mathsf{T} Q \xi$, where ξ denotes the state used in the corresponding offline value-function problem.

The closed-loop simulations are implemented in Julia on an Apple M2 Pro with 32 GB of RAM and use a sampling rate of 10 Hz. The constrained open-loop optimal control problem (OCP) is solved by direct transcription using JuMP/Ipopt, while all QP-based controllers use JuMP/HiGHS. Constrained LQR and MPC provide additional benchmarks for Example 1. For Example 2, the Optimal-Decay (OD) and Rapidly-Exponentially-Stabilizing (RES) CLF-CBF-QP controllers from [10] are tuned for best performance. All methods use the same shared model and cost parameters.

A. Example 1 (Linear): 1D Hovercraft

As a linear benchmark, we apply the H-CBF-QP to a unit-mass 1D hovercraft with state $x = [p,\, v]^\mathsf{T} \in \mathbb{R}^2$ (position and velocity) and input $u \in \mathbb{R}$ (thrust force). The system is a double integrator of the form (1), with

$$f(x) := \begin{bmatrix} v \\ 0 \end{bmatrix}, \qquad g(x) := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For this example, we used SGA as a proof of concept, but since this is a standard linear-quadratic problem, the optimal value function may be obtained directly by solving the appropriate algebraic Riccati equation. We used the quadratic basis $\{p^2, v^2, pv\}$ on $\Omega = [-1, 1]^2$ with initial policy $u_0 = -p - v$. Since this basis spans the exact unconstrained LQR value function, SGA recovers $\hat{V} = V^*$ up to numerical tolerance, and the resulting value function satisfies the unconstrained HJB equation globally. The policy iteration terminates after 4 iterations in approximately 0.4 s.

We used $x_0 = [10,\, 0]^\mathsf{T}$, $Q = I_2$, $R = 1$, $\alpha = 10$, with constraints $|v| \le 1\,\mathrm{m\,s^{-1}}$, and $|u| \le 1\,\mathrm{N}$. Since $\dot{v} = u$, the velocity constraints have relative degree one. Using the CBFs $\bar{B} = v_{\max} - v$ and $\underline{B} = v - v_{\min}$, we obtain the input bound $-\alpha(v - v_{\min}) \le u \le \alpha(v_{\max} - v)$.

For comparison, the constrained LQR benchmark uses the infinite-horizon LQR value function $V_\infty(x) = x^\mathsf{T} P_\infty x$ as the terminal cost of a one-step QP with $\Delta t = 0.1\,\mathrm{s}$, enforcing input and next-step velocity bounds. The resulting scalar-input parametric QP is solved offline by active-set enumeration and KKT conditions, yielding an exact piecewise-affine feedback law evaluated online by region lookup. The MPC uses the same stage cost and constraints with horizon $N = 20$, $\Delta t = 0.1\,\mathrm{s}$, and terminal cost $x_N^\mathsf{T} P_\infty x_N$. No terminal constraint is imposed; a condensed QP is solved at each sample and only the first input is applied.

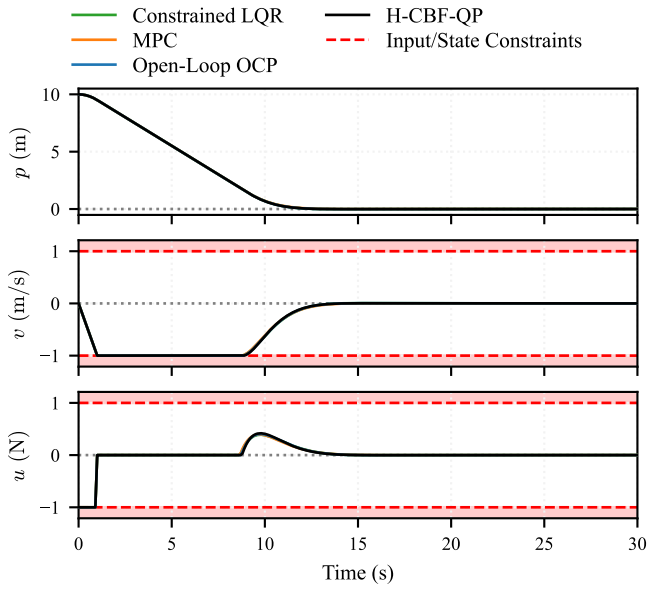

Fig. 1: Position, velocity, and input trajectories for the 1D hovercraft. The trajectories of all controllers are superimposed.

As shown in Fig. 1, the constrained LQR, MPC, open-loop OCP, and H-CBF-QP trajectories all coincide, with costs agreeing to within 10^{-3} ($J \approx 398.34$). Thus, for this example, the H-CBF-QP reproduces the numerical constrained-optimal benchmark. The trajectory consists of an input-saturated phase, a velocity-saturated coasting phase with $\dot{v} = u = 0$, and a terminal unconstrained LQR phase. The H-CBF-QP reproduces this structure through the input bound and relative-degree-one CBF constraints, while in the final unconstrained phase it recovers the LQR feedback by Proposition 1.

B. Example 2 (Nonlinear): Spacecraft Attitude Control

We apply the framework to spacecraft attitude control with reaction wheels, previously studied in [10] using CLF-CBF-QP. The online constrained model has state $x = [\sigma^\mathsf{T}, \omega^\mathsf{T}, h_w^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^9$, where $\sigma \in \mathbb{R}^3$ denotes the Rodrigues Parameters (RPs), $\omega \in \mathbb{R}^3$ is the body-frame angular velocity, and $h_w \in \mathbb{R}^3$ is the reaction-wheel angular momentum. The input $u \in \mathbb{R}^3$ is the control torque. Three identical, axially symmetric reaction wheels are aligned with the spacecraft body-frame axes. The dynamics take the input-affine form (1), with [10], [16]

$$f(x) := \begin{bmatrix} G(\sigma)\omega \\ -J^{-1}[\omega]_\times (J\omega + h_w) \\ 0_{3\times 1} \end{bmatrix}, \qquad g(x) := \begin{bmatrix} 0_{3\times 3} \\ J^{-1} \\ -I_3 \end{bmatrix},$$

where $G(\sigma) = \frac{1}{2}(I_3 + [\sigma]_\times + \sigma\sigma^\mathsf{T})$, $J \in \mathbb{R}^{3\times 3}$ is the spacecraft inertia matrix, and $[\cdot]_\times$ denotes the skew-symmetric cross-product matrix.

For this example, we used SOS policy iteration, as described in Section IV-B. For the offline value-function approximation, we use the reduced state $\xi = [\sigma^\mathsf{T}, \omega^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^6$ and corresponding unconstrained attitude dynamics, introducing h_w only in the online constrained model. The SOS approximation uses all even-degree monomials of degree 2 and 4 (147 monomials) in ξ over $\Omega = [-1, 1]^6$, with regional inequalities enforced via the S-procedure. Starting from $u_0 = -\sigma - 3\omega$, the MOSEK-based policy iteration terminates after 7 iterations in approximately 7.6 s.

We consider a *rest-to-rest maneuver* to the origin from $x_0 = [\sigma_0^\mathsf{T}, 0^\mathsf{T}, 0^\mathsf{T}]^\mathsf{T}$, where $\sigma_0 = [0.312,\, -0.666,\, 0.606]^\mathsf{T}$ corresponds to initial Euler angles[1] $[80°,\, -30°,\, 60°]^\mathsf{T}$. The remaining parameters are $Q = I_6$, $R = I_3$,

$$J = \begin{bmatrix} 1.8140 & -0.1185 & 0.0275 \\ -0.1185 & 1.7350 & 0.0169 \\ 0.0275 & 0.0169 & 3.4320 \end{bmatrix} \mathrm{kg\,m^2},$$

and $|u_i| \le 0.123\,\mathrm{N\,m}$.

Case 1 (CBF): Reaction wheel momentum constraints: The reaction-wheel momentum satisfies $\dot{h}_w = -u$ and is constrained by $|h_{w,i}| \le 0.4\,\mathrm{N\,m\,s}$. Each momentum constraint therefore has relative degree one. Although h_w is excluded from the offline SOS approximation, it is included in the online constrained dynamics and CBF-QP. Including

[1]Using the ZYX extrinsic Euler angles convention.

all nine states on the SOS basis would significantly increase the number of monomials, making the offline formulation substantially less tractable. Defining the componentwise CBFs $\bar{B}_i = h_{w,\max,i} - h_{w,i}$ and $\underline{B}_i = h_{w,i} - h_{w,\min,i}$ yields $-\alpha(h_{w,\max} - h_w) \le u \le \alpha(h_w - h_{w,\min})$ [10], with $\alpha = 10$.

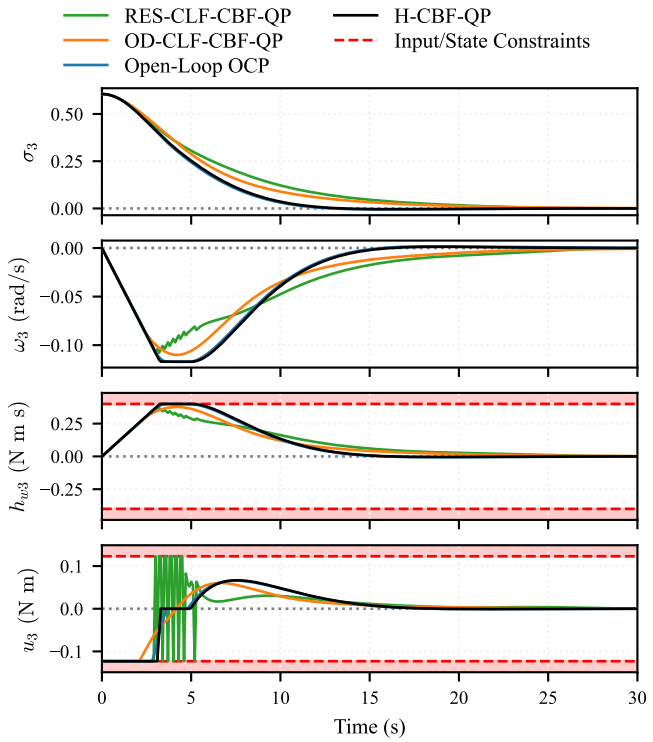

Fig. 2: Representative attitude RPs, angular velocity, reaction-wheel momentum, and control torque responses for the spacecraft under momentum constraints. Only the 3rd components of σ, ω, h_w, and u are shown.

Fig. 2 and Table I show the results with active reaction-wheel momentum constraints. The H-CBF-QP closely tracks the open-loop OCP across the state and input trajectories, while both CLF-CBF-QP controllers exhibit larger transients and slower convergence. In this example, the relative-degree-one momentum constraint reduces to an instantaneous affine bound on u, and the H-CBF-QP closely reproduces the constraint-active behavior of the OCP. Once the momentum constraint is no longer active, Proposition 1 gives $\bar{u} = \hat{u}$. This close trajectory agreement is reflected in the near-OCP cost reported in Table I.

Case 2 (HOCBF): Forbidden pointing constraint: Consider an instrument with body-frame boresight direction $b \in \mathbb{R}^3$ that must avoid a bright celestial object with inertial direction $n \in \mathbb{R}^3$ by at least a half-cone angle θ. The pointing constraint is $B(\sigma) := \cos\theta - b^\mathsf{T} R(\sigma) n \ge 0$ [17], where the rotation matrix from the inertial to body frame is $R(\sigma) = \frac{1}{1+\sigma^\mathsf{T}\sigma}\left[(1-\sigma^\mathsf{T}\sigma)I_3 + 2\sigma\sigma^\mathsf{T} - 2[\sigma]_\times\right]$ [16].

Since B depends only on σ and $\dot{\sigma} = G(\sigma)\omega$, $L_g B = 0$. Moreover, $L_g L_f B \neq 0$, so B has relative degree two. The HOCBF construction of Definition 2 gives

$$\psi_0 = B, \qquad \psi_1 = L_f B + \alpha_1 B,$$
$$\psi_2 = L_f^2 B + L_g L_f B\, u + (\alpha_1 + \alpha_2) L_f B + \alpha_1 \alpha_2 B,$$

where $L_f B = \frac{\partial B}{\partial \sigma} G(\sigma)\omega$ and $L_g L_f B = \frac{\partial B}{\partial \sigma} G(\sigma) J^{-1}$. The constraint $\psi_2 \ge 0$ is affine in u and is imposed directly in (6). We use $b = [0, 0, 1]^\mathsf{T}$, $n = [-0.47, -0.19, 0.86]^\mathsf{T}$ (normalized), $\theta = 15°$, and $\alpha_1 = \alpha_2 = 1$.

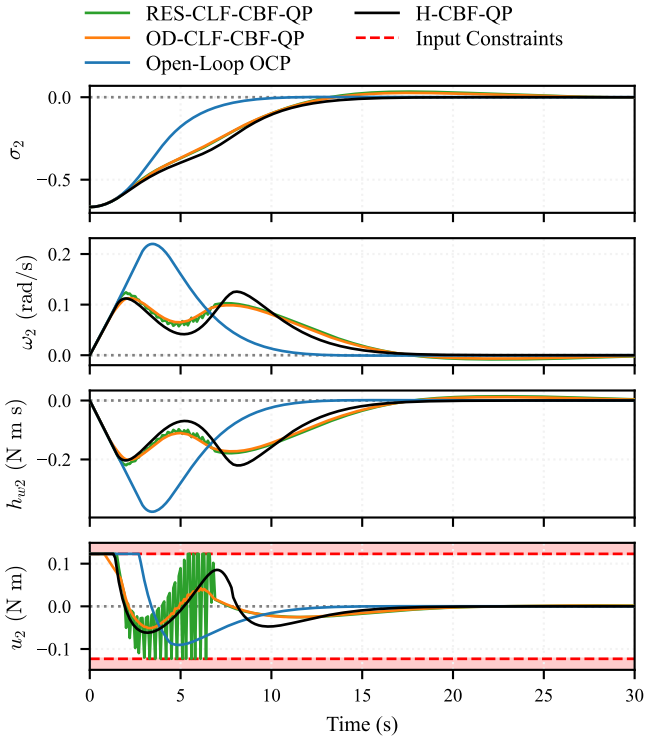

Fig. 3: Representative attitude RPs, angular velocity, reaction-wheel momentum, and control torque responses under the pointing constraint. Only the 2nd components of σ, ω, h_w, and u are shown.

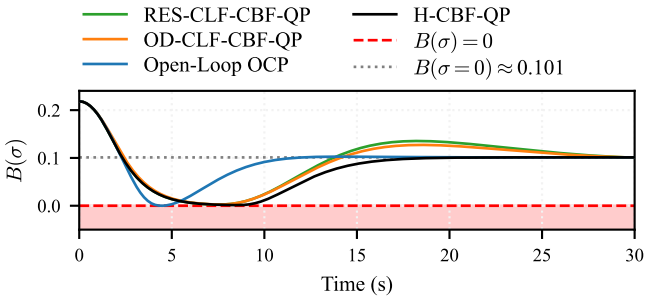

Fig. 4: Pointing-constraint function $B(\sigma)$ for the constrained controllers. All trajectories satisfy $B(\sigma) \ge 0$.

Figs. 3 to 5 and Table I summarize the results with the forbidden pointing constraint. The H-CBF-QP remains qualitatively closer to the OCP trajectory than the CLF-CBF-QP benchmarks, which exhibit larger transients; the RES-CLF-CBF-QP also exhibits noticeable input chattering associated with its prescribed CLF decay. Figs. 4 and 5 confirm satisfaction of the pointing constraint throughout the maneuver.

Unlike Case 1, the constrained OCP begins deviating from the unconstrained trajectory before the HOCBF constraint becomes active, thereby anticipating the future exclusion-zone encounter. The H-CBF-QP, by contrast, has a zero-

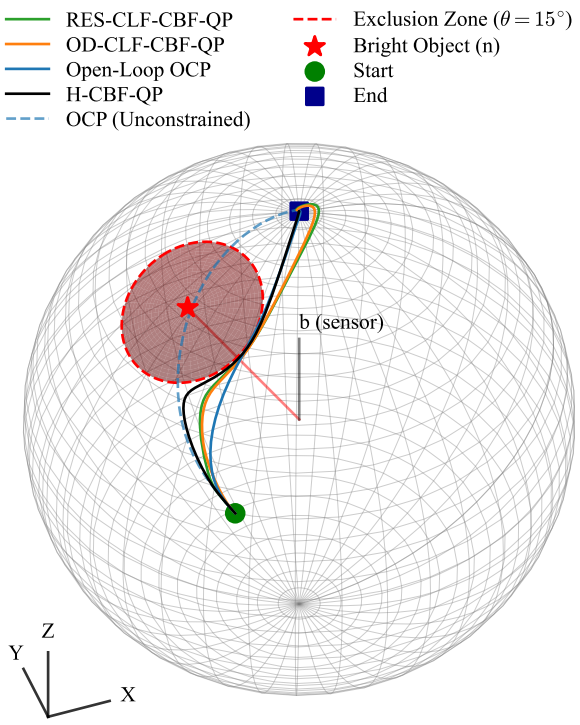

Fig. 5: Sensor boresight trajectories on the celestial sphere. The red region denotes the exclusion zone of half-angle $\theta = 15°$ around the bright object. All constrained trajectories remain outside the exclusion zone.

TABLE I: Cost and computation time for the spacecraft attitude-control problem. The reported times are the total computation times for the online feedback controllers. Case 1: reaction-wheel momentum constraints. Case 2: forbidden pointing constraint.

	Case 1		Case 2	
Control Method	Cost J	Time (s)	Cost J	Time (s)
RES-CLF-CBF-QP	3.5841	0.23	4.1937	0.23
OD-CLF-CBF-QP	3.4755	0.24	4.0626	0.24
H-CBF-QP	**3.2249**	**0.22**	**3.9842**	**0.22**
Open-Loop OCP	3.2211	–	3.3768	–

horizon online structure and uses an unconstrained value function: it follows the unconstrained value-function policy whenever feasible and modifies it only through the instantaneous HOCBF constraint. Consequently, future constraint encounters are not explicitly anticipated, which is consistent with the observed cost gap relative to the OCP in Table I. The HOCBF gains were not optimized; since performance depends on α_1 and α_2, further tuning may reduce this gap.

The source code reproducing all figures is available at `https://github.com/QCGroup/hcbf`.

VI. Conclusion and Future Work

We presented a two-stage framework for constrained feedback control of input-affine nonlinear systems in which an approximate value function is computed offline via HJB-based policy iteration and constraints are enforced online through the H-CBF-QP. The method decouples performance optimization from safety enforcement, allowing constraints to be modified without recomputing the value function. Under suitable conditions, we established safety, approximate optimality, and local asymptotic stability. The examples show near-OCP performance when the constrained-optimal trajectory aligns with the unconstrained value-function policy, while future constraint encounters can introduce a performance gap due to the zero-horizon CBF-QP structure and the use of the unconstrained offline value function.

Future work includes finite-horizon extensions, characterizing when constrained-optimal behavior is recovered, and incorporating limited lookahead to anticipate future constraints.

References

[1] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.

[2] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.

[3] T. H. Summers, K. Kunz, N. Kariotoglou, M. Kamgarpour, S. Summers, and J. Lygeros, "Approximate dynamic programming via sum of squares programming," in *Proc. European Control Conf.*, 2013, pp. 191–197.

[4] L. Yang, H. Dai, A. Amice, and R. Tedrake, "Approximate optimal controller synthesis for cart-poles and quadrotors via sums-of-squares," *IEEE Robot. Autom. Lett.*, vol. 8, no. 11, pp. 7376–7383, 2023.

[5] Y. Jiang and Z.-P. Jiang, "Global adaptive dynamic programming for continuous-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2917–2929, 2015.

[6] Y. Wang, B. O'Donoghue, and S. Boyd, "Approximate dynamic programming via iterated bellman inequalities," *Int. J. Robust Nonlinear Control*, vol. 25, no. 10, pp. 1472–1496, 2015.

[7] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *Proc. European Control Conf.*, 2019, pp. 3420–3431.

[8] W. Xiao and C. Belta, "High-order control barrier functions," *IEEE Trans. Autom. Control*, vol. 67, no. 7, pp. 3655–3662, 2021.

[9] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3861–3876, 2016.

[10] M. A. Shahraki and L. Lessard, "Spacecraft attitude control under reaction wheel constraints using control Lyapunov and control barrier functions," in *Proc. American Control Conf.*, 2025, pp. 947–952.

[11] L. Dihel and N.-S. P. Hyun, "Flexible convergence rate for quadratic programming-based control Lyapunov functions," in *Proc. Conf. Decision and Control*, 2024, pp. 8845–8851.

[12] S. Fan, G. Pan, and H. Werner, "Concave comparison functions for accelerating constrained Lyapunov decay," *arXiv preprint arXiv:2511.14626*, 2025.

[13] M. H. Cohen and C. Belta, "Approximate optimal control for safety-critical systems with control barrier functions," in *Proc. Conf. Decision and Control*, 2020, pp. 2062–2067.

[14] N. M. Yazdani, R. K. Moghaddam, B. Kiumarsi, and H. Modares, "A safety-certified policy iteration algorithm for control of constrained nonlinear systems," *IEEE Control Syst. Lett.*, vol. 4, no. 3, pp. 686–691, 2020.

[15] H. Almubarak, E. A. Theodorou, and N. Sadegh, "HJB based optimal safe control using control barrier functions," in *Proc. Conf. Decision and Control*, 2021, pp. 6829–6834.

[16] H. Schaub and J. L. Junkins, "Stereographic orientation parameters for attitude dynamics: A generalization of the Rodrigues parameters," *J. Astronaut. Sci.*, vol. 44, no. 1, pp. 1–19, 1996.

[17] X.-Z. Wang, F. Wu, X.-N. Shi, Z.-G. Zhou, and D. Zhou, "Control barrier function-based fixed-time spacecraft attitude control under pointing constraints," *Int. J. Control*, vol. 98, no. 11, pp. 2765–2773, 2025.