Near-Optimal 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 near-optimal feedback control of input-affine nonlinear systems. An approximate value function for the unconstrained control problem is computed offline by solving the Hamilton–Jacobi–Bellman equation. Online, a quadratic program is solved that minimizes the associated approximate Hamiltonian subject to safety constraints imposed via control barrier functions. Our proposed architecture decouples performance from constraint enforcement, allowing constraints to be modified online without recomputing the value function. Validation on a linear 2-state 1D hovercraft and a nonlinear 9-state spacecraft attitude control problem demonstrates near-optimal performance relative to open-loop optimal control benchmarks and superior performance compared to control Lyapunov function-based controllers.**

I. Introduction

The optimal feedback control problem for general nonlinear systems remains a central challenge in control theory. The Hamilton–Jacobi–Bellman (HJB) equation provides the theoretical foundation for optimal control; however, its nonlinear partial differential equation (PDE) structure renders analytical solutions generally unavailable, while the curse of dimensionality makes numerical solutions intractable for high-dimensional systems [1]. In practice, control systems must also satisfy safety-critical constraints, such as actuator limits, state bounds, and geometric exclusion zones, further complicating the design of optimal controllers.

Model predictive control (MPC) addresses both optimality and safety via receding-horizon optimization, but solving a nonlinear program online limits real-time applicability for fast or high-dimensional systems.

Alternatively, one can address optimality and safety separately. On the optimization side, approximate dynamic programming methods solve relaxed forms of the HJB equation: the Successive Galerkin Approximation (SGA) [2] projects the Generalized HJB (GHJB) PDE onto polynomial bases, while Sum-of-Squares (SOS) programming [3], [4] replaces intractable non-negativity conditions with semidefinite constraints. Policy iteration schemes [5], [6] then produce monotonically improving value function approximations with global stability guarantees. However, these methods typically address unconstrained problems; embedding constraints into these methods is nontrivial, as it can significantly increase the problem size, require higher-order approximation to maintain feasibility, or render the problem infeasible altogether.

On the constraint enforcement side, Control Barrier Functions (CBFs) [7] and their high-order extensions (HOCBFs) [8] translate state constraints into linear inequalities on the control input, enabling real-time enforcement via quadratic programs (QPs). CBFs have been combined with Control Lyapunov Functions (CLFs) in CLF-CBF-QP architectures [9] for simultaneous stability and safety. However, as observed in our prior work on spacecraft attitude control [10], CLF-based approaches have several limitations: (i) CLF-QP objectives guarantee stability but not optimality, and typically require hand-designed Lyapunov functions via input-output linearization [10]; (ii) fixed decay rates cause input chattering at low sampling frequencies [10], [11]; and (iii) the prescribed decay rate introduces a comparison function design problem [12], where the choice of decay shape governs a feasibility/convergence trade-off further complicated by actuator bounds. All three issues stem from enforcing a pointwise decay rate and require nontrivial design choices.

Several recent works couple safety and performance by incorporating CBF conditions directly into the value function computation [13]–[15], but some guarantee only ultimate boundedness rather than asymptotic stability [13], and all require a full problem re-solve if the safe set changes, eliminating the flexibility to modify constraints online.

This work bridges the gap between the two described lines of work. In contrast to [13]–[15], we propose a two-stage architecture that decouples performance optimization from safety constraint enforcement. The key idea is to replace the hand-designed CLF with a *near-optimal value function* computed offline via policy iteration, and to enforce constraints online through a CBF-QP that requires only the value function gradient and is agnostic to how it was obtained. The contributions of this paper are as follows:

1. We propose the *GHJB-CBF-QP*, a convex QP that uses the gradient of a precomputed value function to drive the system toward *near-optimal* trajectories, while CBF constraints enforce state safety and input bounds. Unlike CLF-CBF-QP [9], [10], the objective reflects optimality rather than mere stability, and the controller is inherently chatter-free: it avoids input-output linearization entirely and imposes no CLF decay rate constraint, eliminating all three CLF limitations discussed above.
2. We characterize the optimality gap, decomposing it into three terms: value function approximation error, constraint projection error, and CBF conservatism, each admitting a separate mitigation strategy, and show that the CBF

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

conservatism term vanishes for integrator-state constraints (relative degree one with zero drift).

3. We validate the framework on a 1D hovercraft (linear, 2 states) and a spacecraft attitude control problem (nonlinear, 9 states), demonstrating near-optimal performance relative to open-loop optimal control and superior performance compared to CLF-CBF-QP controllers with real-time implementability and formal safety guarantees.

The remainder of this paper is organized as follows. Section II formulates the problem and presents the offline value function approximation via policy iteration. Section III develops the online GHJB-CBF-QP controller, including its optimality and safety properties. Section IV validates the framework on linear and nonlinear examples, and Section V concludes with future research directions.

II. Problem Formulation

A. System Dynamics and Objectives

Consider a nonlinear *input-affine* control system

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

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the control input, and $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are assumed to be locally Lipschitz continuous with $f(0) = 0$. The system is subject to state constraints $h_i(x) \le 0$, $i = 1, \dots, n_h$ and input constraints $u \in \mathcal{U} := \{u \mid u_{\min} \le u \le u_{\max}\}$.

We aim to minimize the infinite-horizon cost

$$J(x_0, u(\cdot)) = \int_0^\infty \big(q(x) + u^\mathsf{T} R\, u\big)\,\mathrm{d}t \quad \text{s.t.} \quad (1), \tag{2}$$

where $q : \mathbb{R}^n \to \mathbb{R}_{\ge 0}$ is positive semidefinite with $q(0) = 0$ (here we consider $q(x) := x^\mathsf{T} Q x$ with $Q \succeq 0$), and $R \succ 0$, while satisfying input and state constraints.

Our approach decouples this problem into two layers. First, the *unconstrained* optimal control problem is approximately solved offline via the HJB equation. Then, constraints are enforced online via a CBF-QP. We begin by reviewing the unconstrained problem.

B. Hamilton–Jacobi–Bellman (HJB) Equation

Ignoring constraints for the moment, the optimal value function $V^*(x_0) = \min_{u(\cdot)} J(x_0, u(\cdot))$ satisfies the HJB equation $\mathcal{H}(V^*) = 0$, where the Hamiltonian is

$$\mathcal{H}(V) := \nabla V^\mathsf{T} f(x) + q(x) - \tfrac{1}{4}\nabla V^\mathsf{T} g(x) R^{-1} g(x)^\mathsf{T} \nabla V, \tag{3}$$

and the optimal feedback control is

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

Solving $\mathcal{H}(V^*) = 0$ exactly is generally intractable since (3) is a nonlinear PDE due to the quadratic dependence on ∇V. For a given admissible policy $u(x)$, this nonlinearity is removed: the associated value function V satisfies the Generalized HJB (GHJB) equation, a linear PDE in ∇V:

$$\nabla V^\mathsf{T}\big(f(x) + g(x)u\big) + q(x) + u^\mathsf{T} R u = 0. \tag{5}$$

C. Relaxed HJB Formulation

Following [5], we relax the HJB equality to an inequality. Let $\Omega \subset \mathbb{R}^n$ be a compact set containing the origin, and $\mathcal{P}$ denote the set of positive definite, proper, $\mathcal{C}^1$ functions. The relaxed problem is:

$$\min_V \int_\Omega V(x)\,\mathrm{d}x \quad \text{s.t.} \quad \mathcal{H}(V) \le 0, \quad V \in \mathcal{P}. \tag{6}$$

***Assumption** 1:* There exist $V_0 \in \mathcal{P}$ and a feedback policy u_0 such that $\mathcal{L}(V_0, u_0) \ge 0$ for all $x \in \mathbb{R}^n$, where $\mathcal{L}(V, u) := -\nabla V^\mathsf{T}(f + g\,u) - q(x) - u^\mathsf{T} R u$ is the Bellman operator.

***Assumption** 2:* There exists $V^* \in \mathcal{P}$ satisfying the HJB equation $\mathcal{H}(V^*) = 0$.

***Theorem** 1 (Relaxed Solution Properties [5]):* Under the above assumptions: (i) Problem (6) has a nonempty feasible set. (ii) For any feasible V, the control $\bar{u}(x) = -\frac{1}{2}R^{-1}g^\mathsf{T}\nabla V$ is globally asymptotically stabilizing. (iii) V upper-bounds the cost: $V(x_0) \ge J(x_0, \bar{u})$ for all x_0. (iv) $V^*(x) \le V(x)$ for all feasible V. (v) V^* is the global optimum of (6).

Property (ii) is the key insight: *any* feasible V simultaneously provides a stabilizing controller and a Lyapunov certificate, since $\dot{V} = \mathcal{H}(V) - q(x) - \bar{u}^\mathsf{T} R\,\bar{u} \le -q(x)$. Minimizing $\int V\,\mathrm{d}x$ tightens this bound, driving V toward V^*.

D. Value Function Approximation

Our proposed online control framework requires a precomputed approximate value function $\hat{V}$. The quality of $\hat{V}$ affects performance but not safety (Proposition 2).

For our numerical examples, we used two different approximation methods, which we now briefly review. In all cases, we parameterized V as a polynomial $\hat{V}(x) = \sum_i c_i\, m_i(x)$ in even-degree monomials and solved for the coefficients via *policy iteration* [5]: alternating between evaluating the current policy (solving for $V^{(k)}$ with $\mathcal{L}(V^{(k)}, u^{(k)}) \ge 0$) and improving via $u^{(k+1)} = -\frac{1}{2}R^{-1}g^\mathsf{T}\nabla V^{(k)}$. Under Assumptions 1 and 2, the iterates $\{V^{(k)}\}$ decrease monotonically toward V^*, and each policy is globally stabilizing [5]. The policy evaluation step requires verifying $\mathcal{L}(V^{(k)}, u^{(k)}) \ge 0$. For our linear example, we used Successive Galerkin Approximation (SGA) [2], which reduces the computation to solving a system of linear equations with analytically computable entries, converging to the exact LQR solution as basis richness increases. For our second example, a nonlinear polynomial system, we used Sum-of-Squares (SOS) programming, which replaced the (NP-hard) non-negativity check with a semidefinite program, inheriting the same convergence guarantees [5].

E. Safety via Control Barrier Functions (CBFs)

We now review the (high-order) CBF tools used in the online layer.

***Definition** 1 (Relative Degree [8]):* A (sufficiently) differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ has relative degree r with respect to (1) if $L_g L_f^k h(x) = 0$ for all $k < r - 1$ and $L_g L_f^{r-1} h(x) \ne 0$.

***Definition** 2 (HOCBF [7], [8]):* Given $B(x) \geq 0$ with relative degree r and $\alpha_1, \ldots, \alpha_r > 0$, define

$$\psi_0(x) := B(x), \quad \psi_k(x) := \dot{\psi}_{k-1}(x) + \alpha_k \, \psi_{k-1}(x), \quad (7)$$

for $k = 1, \ldots, r$, with $\mathcal{C}_k = \{x \mid \psi_k(x) \geq 0\}$. Then B is a HOCBF if $\sup_{u \in \mathcal{U}} \psi_r(x, u) \geq 0$ for all $x \in \mathcal{C}_0 \cap \cdots \cap \mathcal{C}_{r-1}$.

***Theorem** 2 (HOCBF Forward Invariance [7], [8]):* If B is a HOCBF and $u(x)$ satisfies $\psi_r(x, u) \geq 0$ on $\mathcal{C}_0 \cap \cdots \cap \mathcal{C}_{r-1}$, then this set is forward invariant.

For $r = 1$, the HOCBF condition reduces to the standard CBF constraint [7]: $L_f B + L_g B \, u \geq -\alpha_1 \, B$, and for $r = 2$, expanding (7) yields

$$L_f^2 B + L_g L_f B \cdot u + (\alpha_1 + \alpha_2) \, L_f B + \alpha_1 \alpha_2 \, B \geq 0,$$

which is affine in u and directly incorporable into a QP.

III. GHJB-CBF-QP Controller

We now present our main contribution, GHJB-CBF-QP, which combines the offline value function from Section II-D with the CBF machinery from Section II-E into a single QP.

A. Formulation

Given the precomputed $\hat{V}(x)$, the online QP objective is derived from the GHJB equation. Recall from (5) that the Hamiltonian evaluated at a control u (before minimization) is $\nabla V(x)^\mathsf{T}(f(x) + g(x) \, u) + q(x) + u^\mathsf{T} R \, u$. Since the drift terms $\nabla V^\mathsf{T} f$ and q do not depend on u, minimizing over u is equivalent to minimizing

$$J_{\text{GHJB}}(u) = \nabla \hat{V}(x)^\mathsf{T} g(x) \, u + u^\mathsf{T} R \, u, \quad (8)$$

which is convex quadratic in u. The unconstrained minimizer is $u^* = -\frac{1}{2} R^{-1} g^\mathsf{T} \nabla \hat{V}$, recovering the offline policy (4) with $\hat{V}$ in place of V^*. The GHJB-CBF-QP augments this objective with CBFs/HOCBFs and input constraints:

$$\begin{aligned} \min_u \quad & \nabla \hat{V}(x)^\mathsf{T} g(x) \, u + u^\mathsf{T} R \, u \\ \text{s.t.} \quad & \psi_r(x, u) \geq 0, && \text{(CBFs/HOCBFs)} \\ & u_{\min} \leq u \leq u_{\max}. && \text{(Input bounds)} \end{aligned} \quad (9)$$

By Theorem 2, any feasible solution guarantees forward invariance of all safe sets, regardless of $\hat{V}$.

***Remark** 1 (No comparison function):* Unlike CLF-CBF-QPs, which require a decay constraint $\dot{V} \leq -\gamma(V)$ with a carefully chosen comparison function γ [12] and a slack variable to ensure feasibility under actuator bounds, the GHJB-CBF-QP (9) imposes no decay rate constraint and requires no slack variable. The effective decay is determined implicitly by $\nabla \hat{V}$, eliminating both design choices.

B. Theoretical Properties

***Proposition** 1 (Optimality Recovery):* In (9), if all constraints are inactive, the solution recovers the unconstrained offline policy: $u_{\text{GHJB-CBF}} = -\frac{1}{2} R^{-1} g^\mathsf{T} \nabla \hat{V}$.

Proof: With no active constraints, (9) reduces to unconstrained minimization of (8). Setting $\partial J_{\text{GHJB}} / \partial u = g^\mathsf{T} \nabla \hat{V} + 2Ru = 0$ yields $u^* = -\frac{1}{2} R^{-1} g^\mathsf{T} \nabla \hat{V}$. ■

***Proposition** 2 (Safety-Performance Decoupling):* For any value function approximation $\hat{V}$, the online policy (9): (i) is guaranteed to be safe and (ii) recovers optimal performance as $\hat{V} \to V^*$ when the constraints are inactive.

Proof: (i) The CBF and HOCBF constraints in (9) are enforced as hard constraints regardless of the objective. By Theorem 2, any feasible u renders the safe sets forward invariant, independent of $\hat{V}$. (ii) As $\hat{V} \to V^*$, we have $\nabla \hat{V} \to \nabla V^*$, so the QP objective approaches the true Hamiltonian minimization. By Proposition 1, whenever constraints are inactive the controller recovers $u^* = -\frac{1}{2} R^{-1} g^\mathsf{T} \nabla V^*$. ■

***Remark** 2 (Optimality Gap Decomposition):* Let $J^{\text{c}}(x_0)$ denote the cost achieved by the constrained optimal controller (solving (2) subject to all state and input constraints), and let $J_{\text{GHJB-CBF}}(x_0)$ denote the cost achieved by the GHJB-CBF-QP controller (9). The total suboptimality $J_{\text{GHJB-CBF}}(x_0) - J^{\text{c}}(x_0)$ can be attributed to three sources: (i) Δ_{approx}: error from approximating V^* with $\hat{V}$; (ii) Δ_{proj}: error from projecting the unconstrained policy onto the feasible set, which is zero when constraints are inactive; and (iii) Δ_{CBF}: conservatism of the CBF formulation relative to hard state constraints, which depends on the constraint structure and the parameter α (CBF decay rate). These sources are conceptually distinct and admit separate mitigation strategies: Δ_{approx} by improving the value function approximation, Δ_{proj} by loosening constraints, and Δ_{CBF} by increasing α or exploiting favorable constraint structure (see Proposition 3).

To our knowledge, the following observation has not been formally stated in prior work, though it follows directly from the CBF construction.

***Proposition** 3 (Integrator-State Constraints):* Consider a box constraint $x_{i,\min} \leq x_i \leq x_{i,\max}$ on a state component whose dynamics satisfy $\dot{x}_i = f_i(x) + g_i(x)^\mathsf{T} u$ with $f_i(x) = 0$ for all x and t. Define $\bar{B}(x) = x_{i,\max} - x_i$ and $\underline{B}(x) = x_i - x_{i,\min}$. Then each CBF has relative degree one with $L_f B = 0$, and the CBF constraints recover the hard state constraints exactly as $\alpha \to \infty$, so $\Delta_{\text{CBF}} = 0$.

Proof: Since $f_i(x) = 0$, the Lie derivatives are $L_f \bar{B} = -f_i = 0$ and $L_g \bar{B} = -g_i(x)^\mathsf{T}$, so the CBF conditions reduce to $-g_i^\mathsf{T} u \geq -\alpha(x_{i,\max} - x_i)$ and $g_i^\mathsf{T} u \geq -\alpha(x_i - x_{i,\min})$.

In the interior ($x_{i,\min} < x_i < x_{i,\max}$), the right-hand sides tend to $-\infty$ as $\alpha \to \infty$, so the constraints become vacuous and impose no restriction on u. At the boundary $x_i = x_{i,\max}$, the upper constraint becomes $g_i^\mathsf{T} u \leq 0$, which exactly prevents x_i from increasing past the boundary (since $\dot{x}_i = g_i^\mathsf{T} u$). The lower boundary is analogous. Thus, in the limit $\alpha \to \infty$, the CBF constraints are equivalent to the hard state constraints, and no conservatism is introduced. ■

In practice, a moderately large α (e.g., $\alpha = 10$) suffices to make the CBF conservatism practically vanish, so that $\Delta_{\text{CBF}} \approx 0$ without requiring $\alpha \to \infty$ (see Section IV).

IV. Application Examples

We validate the GHJB-CBF-QP on two examples: Example 1 on a linear system where the exact optimum is known, and Example 2 on a nonlinear system with constraints of

relative degree one (Case 1) and two (Case 2). Together, the examples demonstrate (i) near-zero CBF conservatism for integrator-state constraints, (ii) the flexibility to handle different constraint types without modifying the offline value function, and (iii) the inherent suboptimality of pointwise QP-based controllers relative to trajectory-aware optimal control for higher-order constraints.

All simulations run in Julia on an Apple M2 Pro chip with 32 GB of RAM at a sampling rate of 10 Hz. The constrained Optimal Control Problem (OCP) (2) is solved via direct transcription in JuMP/Ipopt; all QP-based controllers use JuMP/HiGHS. MPC and constrained Linear Quadratic Regulator (LQR) serve as additional benchmarks for Example 1. The Optimal-Decay (OD) and Rapidly-Exponentially-Stabilizing (RES) CLF-CBF-QP controllers for Example 2 are from [10]. All shared parameters are identical across controllers.

A. Example 1 (Linear): 1D Hovercraft

The 1D hovercraft has state $x = [p,\, v]^\mathsf{T} \in \mathbb{R}^2$ (position and velocity). The input $u \in \mathbb{R}$ is the thrust force. The system is a double integrator in input-affine form (1) with

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

The SGA uses basis functions $\{p^2, v^2, pv\}$ over $\Omega = [-1, 1]^2$ and the initial policy $u_0 = -p - v$. The SGA policy iteration took approximately 0.4 seconds and converged after 4 iterations. The velocity is constrained by $v_{\min} \le v \le v_{\max}$ and the input by $u_{\min} \le u \le u_{\max}$. Since $\dot{v} = u$, velocity is an integrator-state with relative degree one. Defining $\bar{B}(x) = v_{\max} - v$ and $\underline{B}(x) = v - v_{\min}$, the CBF conditions (Proposition 3) reduce to $-\alpha(v - v_{\min}) \le u \le \alpha(v_{\max} - v)$.

The parameters used in the simulations are $x_0 = [10,\, 0]^\mathsf{T}$, $Q = I_2$, $R = 1$, $\alpha = 10$, $|v| \le 1$ m/s, and $|u| \le 1$ N.

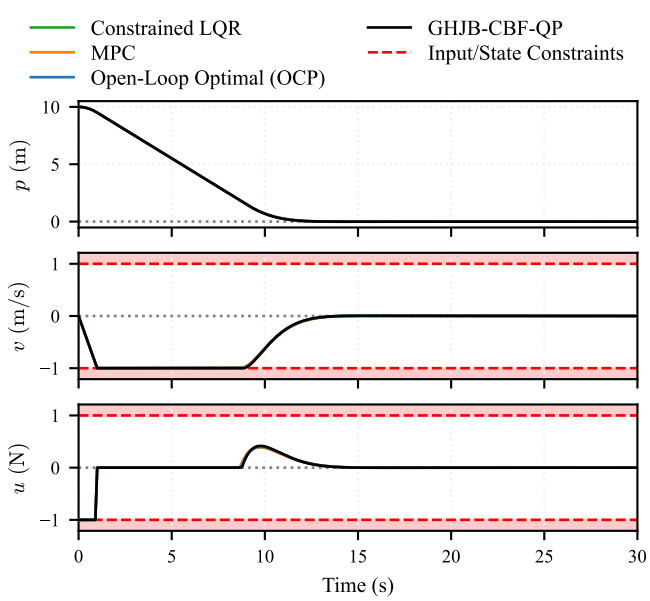

Fig. 1: Simulated trajectories (position, velocity, input) for the 1D hovercraft for various controllers. All trajectories are the same for this example.

TABLE I: Cost and wall clock time comparison for different controllers. For OCP, this is the offline computation time. For the other control methods, this is the total online computation time (all iterations). The corresponding state and input trajectories are shown in Fig. 1.

Control Method	Cost J	Time (s)
Constrained LQR	398.3448	0.07
MPC	398.3475	1.07
Open-Loop Optimal (OCP)	398.3440	0.03
GHJB-CBF-QP	**398.3448**	**0.28**

The simulation results in Fig. 1 and Table I confirm that the closed-loop GHJB-CBF-QP achieves near-identical performance (all trajectories are superimposed) to all three benchmarks for linear systems. All reported times reflect the total wall clock time to generate the full trajectory; for QP-based controllers this is the cumulative cost of all per-step solves, confirming real-time implementability. The close agreement with the open-loop OCP is consistent with Proposition 3, which predicts $\Delta_{\mathrm{CBF}} \approx 0$ for the velocity constraint since it is an integrator-state with zero drift. Moreover, the relative-degree-one structure means the constraints reduce to instantaneous bounds on u, for which pointwise projection incurs little cost penalty ($\Delta_{\mathrm{proj}} \approx 0$).

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 state is $x = [\sigma^\mathsf{T}, \omega^\mathsf{T}, h_w^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^9$, where $\sigma \in \mathbb{R}^3$ denotes the Rodrigues Parameters (RPs) attitude representation, $\omega \in \mathbb{R}^3$ is the angular velocity in the body-fixed frame, 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 assumed to be aligned with the spacecraft's principal axes. The system takes the input-affine form (1) with [10], [16]

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

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.

The SOS approximation uses 147 even-degree monomials (degrees 2 and 4) in $(\sigma, \omega) \in \mathbb{R}^6$ with domain $\Omega = [-1, 1]^6$ enforced via the S-procedure. The initial stabilizing policy is $u_0 = -\sigma - 3\omega$. The SOS policy iteration using MOSEK took approximately 7.6 seconds and converged after 7 iterations.

We consider a *rest-to-rest maneuver* to the origin from $x_0 = \begin{bmatrix}\sigma_0^\mathsf{T} & 0^\mathsf{T} & 0^\mathsf{T}\end{bmatrix}^\mathsf{T}$, where $\sigma_0 = \begin{bmatrix}0.312 & -0.666 & 0.606\end{bmatrix}^\mathsf{T}$ corresponds to initial Euler angles $\begin{bmatrix}80^\circ & -30^\circ & 60^\circ\end{bmatrix}^\mathsf{T}$. The remaining parameters are $Q = I_6$, $R = I_3$, inertia matrix $J = \begin{bmatrix} 1.8140 & -0.1185 & 0.0275 \\ -0.1185 & 1.7350 & 0.0169 \\ 0.0275 & 0.0169 & 3.4320 \end{bmatrix}$ kg m^2, and $|u_i| \le 0.123$ N m.

Case 1 (CBF): Reaction wheel momentum constraints: The momentum h_w satisfies $\dot{h}_w = -u$, an integrator-state with zero drift under box constraints $|h_{w,i}| \le 0.4$ N m s.

Notably, h_w enters the system exclusively through the CBF layer and is excluded from the SOS approximation. Extending the basis to all nine states would require 540 monomials, so this separation is crucial for tractability. Defining vector-valued CBFs $\bar{B}(x) = h_{w,\max} - h_w$ and $\underline{B}(x) = h_w - h_{w,\min}$, the CBF conditions (Proposition 3) reduce to $-\alpha(h_w - h_{w,\min}) \leq u \leq \alpha(h_{w,\max} - h_w)$ [10], with $\alpha = 10$.

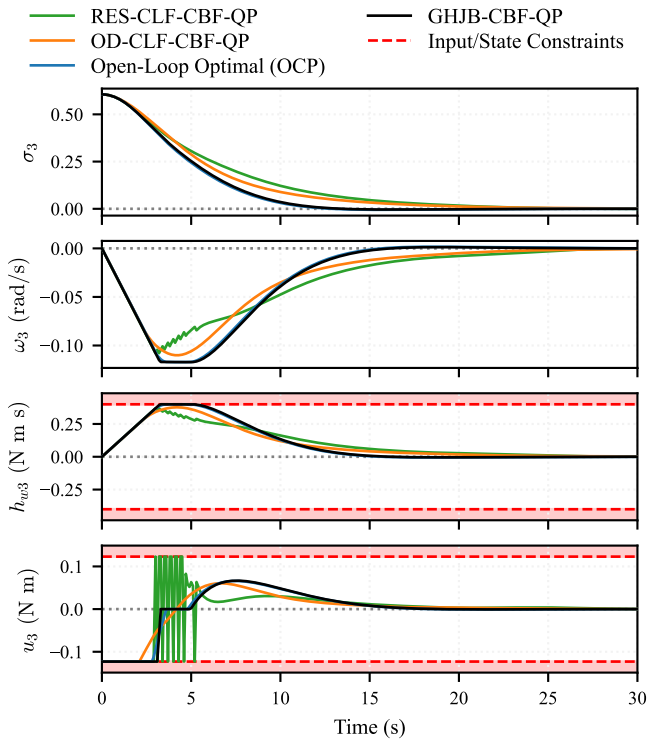

Fig. 2: Representative attitude RPs, angular velocity, reaction wheel angular momentum, and control torque simulation responses for various controllers. Only the 3rd components of σ, ω, h_w, and u are shown.

Fig. 2 and Table II present the spacecraft results with reaction wheel momentum constraints active. The GHJB-CBF-QP closely tracks the OCP across all channels, while both CLF-CBF-QP controllers exhibit larger overshoots and slower convergence. The GHJB-CBF-QP also runs significantly faster than the OCP, as shown in Table II. Consistent with Proposition 3, the momentum constraints introduce almost no CBF conservatism ($\Delta_{\mathrm{CBF}} \approx 0$). The near-optimal cost further suggests that Δ_{proj} is small in this case, as the relative-degree-one structure reduces the constraints to instantaneous bounds on u, for which pointwise projection suffices without trajectory-level lookahead.

Case 2 (HOCBF): Forbidden pointing constraint: Consider an instrument with boresight direction $b \in \mathbb{R}^3$ (body frame) that must avoid a bright celestial object in inertial direction $n \in \mathbb{R}^3$ by at least a half-cone angle θ. The pointing constraint is $B(\sigma) = \cos\theta - b^{\mathsf{T}} R(\sigma)\, n \geq 0$ [17], where the rotation matrix from 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 $\dot{\sigma} = G(\sigma)\omega$, the first derivative $\dot{B} = \frac{\partial B}{\partial \sigma} G(\sigma)\omega$ does not contain u, so $L_g B = 0$ and the constraint has relative degree two. The HOCBF construction (Definition 2) applies with

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

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 $\psi_2 \geq 0$ is affine in u and enters the QP (9) directly. We use boresight $b = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathsf{T}}$ (instrument along the body $+z$ axis), bright object direction $n = \begin{bmatrix} -0.47 & -0.19 & 0.86 \end{bmatrix}^{\mathsf{T}}$ (normalized), half-cone angle $\theta = 15°$, and $\alpha_1 = \alpha_2 = 1$.

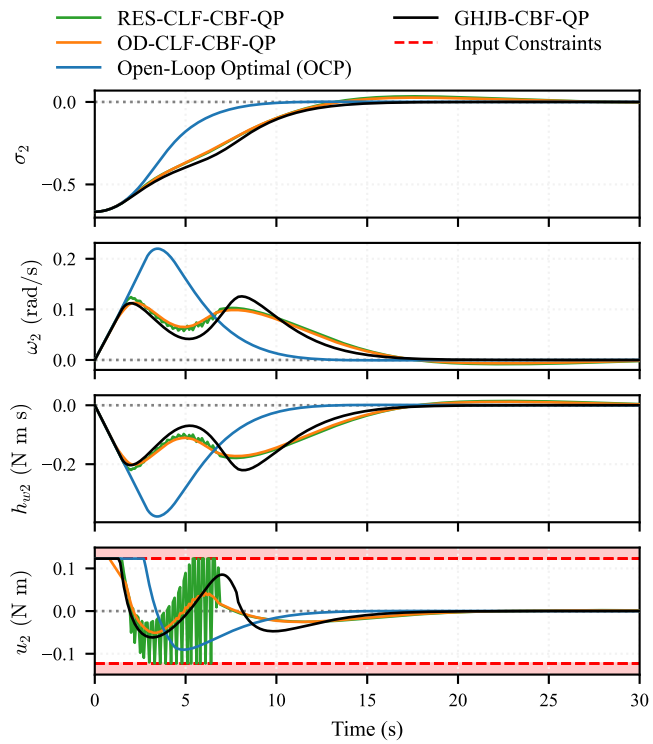

Fig. 3: Representative attitude RPs, angular velocity, reaction wheel angular momentum, and control torque simulation responses for various controllers. Only the 2nd components of σ, ω, h_w, and u are shown.

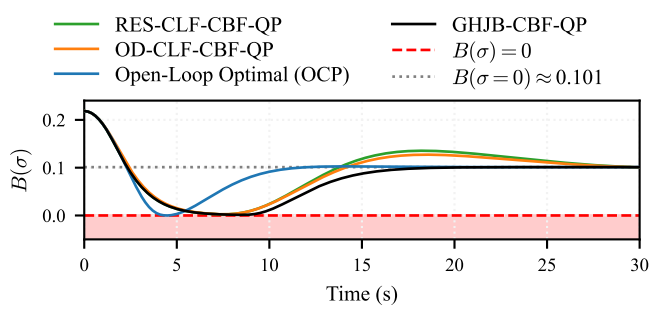

Fig. 4: Pointing constraint $B(\sigma)$ simulation response for various controllers. Since $B(\sigma) > 0$, we are assured that the pointing constraint is satisfied during the trajectory.

figs. 3 to 5 and Table II present the spacecraft results with the forbidden pointing constraint active. The GHJB-CBF-QP closely tracks the OCP (Fig. 3), while the CLF-CBF-QP controllers exhibit larger transients, with the RES-CLF-CBF-QP additionally showing notable input chattering

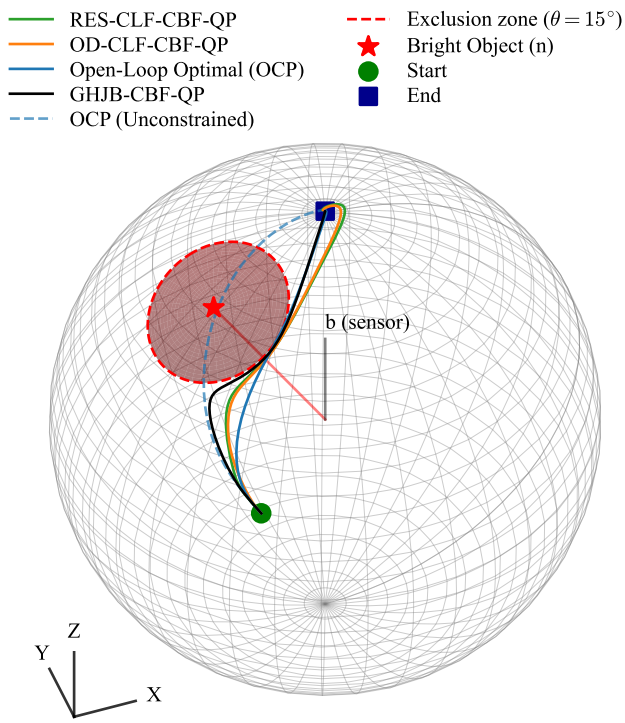

Fig. 5: Sensor boresight trajectory on the celestial sphere. The red region denotes the exclusion zone of half-angle $\theta = 15°$ around the bright object. All constrained controllers maintain $B(\sigma) \geq 0$ throughout the maneuver.

TABLE II: Cost and wall clock time comparison for the spacecraft attitude control problem under two constraint cases. For OCP, this is the offline computation time. For the other control methods, this is the total online computation time (all iterations). Case 1: Reaction wheel momentum constraints (Fig. 2). Case 2: Forbidden pointing constraint (Fig. 3).

	Case 1		Case 2	
Control Method	Cost J	Time (s)	Cost J	Time (s)
RES-CLF-CBF-QP	3.5841	0.35	4.1937	0.35
OD-CLF-CBF-QP	3.4755	0.35	4.0626	0.36
Open-Loop Optimal (OCP)	3.2211	3.66	3.3768	5.49
GHJB-CBF-QP	**3.2249**	**0.51**	**3.9842**	**0.53**

due to the CLF decay rate enforcement. Fig. 4 confirms that $B(\sigma) \geq 0$ holds at all times for all controllers. In Fig. 5, the open-loop OCP takes the shortest feasible path along the exclusion zone boundary. The GHJB-CBF-QP initially follows the unconstrained OCP path toward the exclusion zone, then converges back toward the constrained OCP path after passing it, remaining closer to the OCP throughout than the CLF-CBF-QP controllers. Unlike Case 1, the HOCBF auxiliary sets $\mathcal{C}_1 \cap \mathcal{C}_2$ are a strict subset of $\mathcal{C}_0$, so $\Delta_{\text{CBF}} > 0$. The larger cost gap relative to Case 1 (Table II) reflects both this HOCBF conservatism and the relative-degree-two structure, which means optimal constraint satisfaction requires trajectory-level lookahead that a pointwise QP cannot replicate, making Δ_{proj} a more significant source of suboptimality. The source code reproducing all figures is available at `https://github.com/QCGroup/ghjbcbf`.

V. Conclusion and Future Work

We presented a two-stage framework for constrained optimal control of input-affine nonlinear systems, in which an approximate value function is computed offline via policy iteration, and constraints are enforced online via a GHJB-based CBF-QP that is solvable in real time. The key architectural advantage is the clean decoupling of performance optimization from constraint enforcement: constraints can be added, removed, or modified without recomputing the value function. We established conditions under which the optimality gap is small, showing that for integrator-state constraints, i.e., box constraints of relative degree one with zero drift, the CBF introduces zero additional conservatism.

Future work includes extension to the finite-horizon setting with time-varying value functions, a more rigorous study of the optimality gap, and deriving tighter gap upper bounds.

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, J. L. Junkins *et al.*, "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.