Multi-Rate Planning and Control of Uncertain Nonlinear Systems: Model Predictive Control and Control Lyapunov Functions

Noel Csomay-Shanklin*, Andrew J. Taylor*, Ugo Rosolia, Aaron D. Ames

Abstract—**Modern control systems must operate in increasingly complex environments subject to safety constraints and input limits, and are often implemented in a hierarchical fashion with different controllers running at multiple time scales. Yet traditional constructive methods for nonlinear controller synthesis typically "flatten" this hierarchy, focusing on a single time scale, and thereby limited the ability to make rigorous guarantees on constraint satisfaction that hold for the entire system. In this work we seek to address the stabilization of constrained nonlinear systems through a *multi-rate* control architecture. This is accomplished by iteratively planning continuous reference trajectories for a nonlinear system using a linearized model and Model Predictive Control (MPC), and tracking said trajectories using the full-order nonlinear model and Control Lyapunov Functions (CLFs). Connecting these two levels of control design in a way that ensures constraint satisfaction is achieved through the use of *Bézier curves*, which enable planning continuous trajectories respecting constraints by planning a sequence of discrete points. Our framework is encoded via convex optimization problems which may be efficiently solved, as demonstrated in simulation.**

I. Introduction

The study and design of nonlinear control systems has long been framed through the lens of stabilization, often in an optimal sense. This is coupled with the fact that one typically considers a single model, implicitly representing a single time scale. However in most modern engineering settings, especially in the context of autonomous and robotic systems, the task of stabilization is complicated by the need to meet safety-critical constraints on the system's state while respecting input limitations. To address this need, implementations often utilize a hierarchical approach that spans multiple time-scales, from the planning layer—which typically leverages discrete-time models—to the real-time controller layer which often considers continuous-time representations. Thus it is necessary to develop efficient control synthesis techniques that provide rigorous guarantees of stability, even in the presence of such constraints, and across multiple time scales.

At the level of real-time control design, a rich catalog of methods have been developed for stabilizing nonlinear systems in the presence of unknown disturbances by utilizing underlying structural properties of the system [1]–[4]. In particular, the tools of Control Lyapunov Functions (CLFs) [5], [6] and Input-to-State Stability [7] have enabled the joint synthesis of stabilizing controllers and Lyapunov certificates of stability

* Authors contributed equally. N. Csomay-Shanklin, A. J. Taylor, U. Rosolia, and A. D. Ames are with the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA, {noelcs, ajtaylor, urosolia, ames}@caltech.edu.
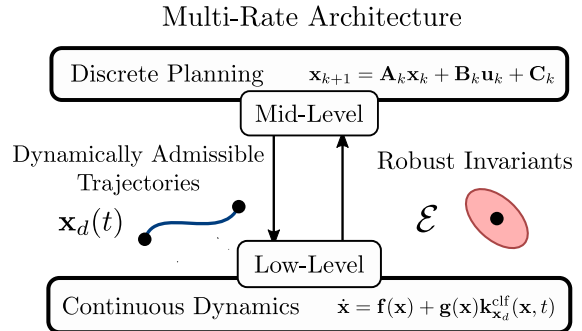
Multi-Rate Architecture

Fig. 1. Overview of Multi-Rate Architecture, with discrete planning producing reference trajectories at a mid-level and continuous controllers producing invariant sets at a low-level.

in the presence of disturbances, including through convex optimization [8]–[10]. These methods for stabilization yield highly structured controllers, and modifying these designs to accommodate state and input constraints may destroy the stability properties guaranteed by the controller. This issue is often circumnavigated theoretically by limiting the domain on which stability is guaranteed, effectively ignoring constraints.

In contrast, Model Predictive Control (MPC) provides an effective method for addressing constraints [11]–[13]. This is achieved by directly incorporating constraints a into controller that iteratively plans a finite sequence of states and inputs that are related through a discrete model of the system dynamics and satisfy required constraints. Although MPC has been successfully demonstrated in several challenging control settings [14]–[24], it is rarely implemented in real-time using the full-order continuous time nonlinear dynamics while accounting for unknown disturbances acting on the system. Thus, MPC implementation for nonlinear systems usually lack strong theoretical guarantees on constraint satisfaction in the presence of disturbances. This is because (i) it is typically difficult to find a closed-form expression for the exact temporal discretization of continuous time nonlinear dynamics [25], (ii) approximating the exact discretization through numerical integration typically yields a non-convex relationship between planned states and inputs, and (iii) exactly propagating disturbances through high-dimensional nonlinear dynamics is often computationally intractable [26]. These challenges often preclude achieving the computational efficiency needed for real-time implementation.

The difficulty in realizing MPC based controllers at a fast enough rate to allow for real-time implementation is often resolved by using an approximate model of the system dynamics that is amenable to efficient planning, typically through reduced-order models or via linearization and temporal discretization of the continuous time nonlinear system dynamics

[13], [22]–[24], [27]. The use of such approximations creates a gap between the system which is being planned for and the actual evolution of the nonlinear system, requiring an additional measure of robustness to ensure constraint satisfaction. This robustness is often achieved by tightening the constraint sets by the maximum deviation between the approximate model and the continuous time nonlinear system dynamics [28]–[35]. Approximating worse-case deviations is typically done using properties of the dynamics which may be difficult to compute, such as Lipschitz constants for which over-approximations yield conservativeness, or by solving computationally intensive optimization programs. More recently, hierarchical control frameworks have been proposed that plan with an approximate model, but directly address nonlinear dynamics with a low-level controller [36], [37]. However, this work does not address if the low-level controller respects state and input constraints as it follows the planned trajectory under disturbances.

In this work we propose a novel multi-rate control architecture that unifies the planning capabilities of Model Predictive Control with the ability to directly address nonlinear dynamics provided by Control Lyapunov Functions. The fundamental tool that allows our framework to explicitly address the relationship between a planner and controller operating at different time scales are *Bézier curves* [38], [39]. By directly planning over the *control points* that parameterize Bézier curves, we capitalize on a critical convex hull property to ensure that state and input constraints are met by the nonlinear system evolving under an optimization-based CLF controller. While Bézier curves have been used in motion planning, or to verify constraint satisfaction after solving an MPC problem [40], this is to the best of our knowledge the first result directly planning over Bézier control points in an MPC formulation, and using the resulting continuous trajectories to ensure constraint satisfaction for a nonlinear system with disturbances.

We begin in Section II by reviewing nonlinear dynamics, and how structural properties can be used to synthesize CLFs and optimization-based controllers for stabilizing a class of *dynamically admissible* reference trajectories. These controllers yield a description of how accurately a reference trajectory is tracked in the presence of disturbances that is amenable to being incorporated into planning. Next, in Section III we provide a review of Bézier curves, and show how they may be used to synthesize reference trajectories for the disturbed nonlinear system such that state constraints are satisfied. Section IV uses the properties of Bézier curves in conjunction with the structure of the low-level controller to formulate constraints on Bézier control points that ensure the low-level controller satisfies input constraints. In Section V we integrate the preceding constructions into an MPC formulation that plans over Bézier control points and synthesizes continuous reference trajectories using a locally linearized and discretized model while ensuring recursive feasibility. We conclude in Section VI with simulation results. We note that proofs may be found in the appendix.

II. LOW-LEVEL CONTROLLER DESIGN

In this section we review nonlinear dynamical systems and discuss the design of nonlinear feedback controllers that provide a measure of disturbance rejection. Importantly, these controllers will yield a quantitative description of reference trajectory tracking that is amenable to being directly incorporated into the synthesis of the reference trajectory itself.

Consider the nonlinear control-affine system:

$$\dot{\mathbf{x}} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ 0 & \mathbf{0}^\top \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{0} \\ f(\mathbf{x}) \end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} \mathbf{0} \\ g(\mathbf{x}) \end{bmatrix}}_{\mathbf{g}(\mathbf{x})} u + \mathbf{w}(t), \qquad (1)$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $u \in \mathbb{R}$, piecewise continuous[1] disturbance signal $\mathbf{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, and functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, assumed to be continuously differentiable on $\mathbb{R}^n$. Furthermore, we make the following assumption:

Assumption 1. The function f satisfies $f(\mathbf{0}) = 0$ and the function g satisfies $g(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

The first assumption takes the origin to be an unforced equilibrium point of the undisturbed system. The second assumption amounts to the system (1) possessing a relative degree [1]. We note that while we consider a single-input, single-output system, this is purely to simplify the presentation of our contributions, and the subsequent developments may be easily extended to the multiple-input, multiple-output setting under an equivalent assumption of a vector relative degree.

Let $\underline{t}, \overline{t} \in \mathbb{R}_{\geq 0}$ with $\underline{t} < \overline{t}$, and let $k : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}$ be a feedback controller that is locally Lipschitz continuous with respect to its first argument[2] and piecewise continuous with respect to its second argument on $\mathbb{R}^n \times [\underline{t}, \overline{t}]$. This controller yields the closed-loop system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})k(\mathbf{x}, t) + \mathbf{w}(t). \qquad (2)$$

As the functions $\mathbf{f}$, $\mathbf{g}$, and k are locally Lipschitz continuous with respect to $\mathbf{x}$ and k is piecewise continuous with respect to t, for any initial condition $\mathbf{x}_0 \in \mathbb{R}^n$ and any piecewise continuous disturbance $\mathbf{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, there exists an interval $I(\underline{t}, \mathbf{x}_0, \mathbf{w}) \triangleq [\underline{t}, \underline{t} + \delta(\mathbf{x}_0, \mathbf{w}))$ with $\delta(\mathbf{x}_0, \mathbf{w}) \in \mathbb{R}_{>0}$ such that the system (2) has a unique piecewise continuously differentiable[3] solution $\boldsymbol{\varphi} : I(\underline{t}, \mathbf{x}_0, \mathbf{w}) \to \mathbb{R}^n$ satisfying:

$$\dot{\boldsymbol{\varphi}}(t) = \mathbf{f}(\boldsymbol{\varphi}(t)) + \mathbf{g}(\boldsymbol{\varphi}(t))k(\boldsymbol{\varphi}(t), t) + \mathbf{w}(t), \qquad (3)$$
$$\boldsymbol{\varphi}(\underline{t}) = \mathbf{x}_0, \qquad (4)$$

for almost all $t \in I(\underline{t}, \mathbf{x}_0, \mathbf{w})$ [3].

With a view towards controller design, the system (1) may also be used to define a class of reference trajectories:

---

[1] This definition is taken as in [3], with piecewise continuity requiring the existence of one-sided limits at points of discontinuity.

[2] This definition is taken as in [3], with local Lipschitz continuity holding with a Lipschitz constant that is uniform in the function's second argument.

[3] Piecewise continuous differentiability is taken to mean a continuous function with a derivative defined on the open intervals of a finite partition with one-sided limits.

Definition 1 (*Dynamically Admissible Trajectory*). A piecewise continuously differentiable function $\mathbf{x}_d : [\underline{t}, \overline{t}] \to \mathbb{R}^n$ is a *dynamically admissible trajectory* for the system (1) if there is a piecewise continuous function $u_d : [\underline{t}, \overline{t}] \to \mathbb{R}$ such that:

$$\dot{\mathbf{x}}_d(t) = \mathbf{f}(\mathbf{x}_d(t)) + \mathbf{g}(\mathbf{x}_d(t))u_d(t), \qquad (5)$$

for almost all $t \in [\underline{t}, \overline{t}]$.

Given a dynamically admissible trajectory $\mathbf{x}_d : [\underline{t}, \overline{t}] \to \mathbb{R}^n$ for (1), let us denote: $\dot{\mathbf{x}}_d(t) = \begin{bmatrix} \dot{x}_d^1(t) & \cdots & \dot{x}_d^n(t) \end{bmatrix}^\top$, and define a error function $\mathbf{e}_{\mathbf{x}_d} : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}^n$:

$$\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t) = \mathbf{x} - \mathbf{x}_d(t), \qquad (6)$$

and its derivative $\dot{\mathbf{e}}_{\mathbf{x}_d} : \mathbb{R}^n \times [\underline{t}, \overline{t}] \times \mathbb{R} \to \mathbb{R}^n$ as:

$$\dot{\mathbf{e}}_{\mathbf{x}_d}(\mathbf{x}, t, u) = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{w}(t) - \dot{\mathbf{x}}_d(t). \qquad (7)$$

Denoting:

$$\mathcal{F}_{\mathbf{x}_d}(\mathbf{x}, t) = f(\mathbf{x}) - \dot{x}_d^n(t), \qquad (8)$$

the structure of the system (1) implies that:

$$\dot{\mathbf{e}}_{\mathbf{x}_d}(\mathbf{x}, t, u) = \overbrace{\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ 0 & \mathbf{0}^\top \end{bmatrix} \mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t) + \begin{bmatrix} \mathbf{0} \\ \mathcal{F}_{\mathbf{x}_d}(\mathbf{x}, t) \end{bmatrix}}^{\mathbf{f}_{\mathbf{x}_d}(\mathbf{x}, t)} \qquad (9)$$
$$+ \mathbf{g}(\mathbf{x})u + \mathbf{w}(t).$$

This structure in conjunction with the assumption that $\mathbf{g}(\mathbf{x}) \neq 0$ for any $\mathbf{x} \in \mathbb{R}^n$ enables a controller $k_{\mathbf{x}_d}^{\text{fbl}} : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}$:

$$k_{\mathbf{x}_d}^{\text{fbl}}(\mathbf{x}, t) = g(\mathbf{x})^{-1} \left(-\mathcal{F}_{\mathbf{x}_d}(\mathbf{x}, t) - \mathbf{K}^\top \mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t) \right), \qquad (10)$$

where $\mathbf{K} \in \mathbb{R}^n$ is selected to yield the relationship:

$$\dot{\mathbf{e}}_{\mathbf{x}_d}(\mathbf{x}, t, k_{\text{fbl}}(\mathbf{x}, t)) = \mathbf{F}\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t) + \mathbf{w}(t), \qquad (11)$$

with $\mathbf{F} \in \mathbb{R}^{n \times n}$ a Hurwitz matrix. For any $\mathbf{Q} \in \mathbb{S}_{\succ 0}^n$ (symmetric positive definite matrices) there exists a unique $\mathbf{P} \in \mathbb{S}_{\succ 0}^n$ solving the Continuous Time Lyapunov Equation:

$$\mathbf{F}^\top \mathbf{P} + \mathbf{P}\mathbf{F} = -\mathbf{Q}. \qquad (12)$$

For a particular $\mathbf{Q}$, the corresponding solution $\mathbf{P}$ may be used to define the following function $V_{\mathbf{x}_d} : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}_{\geq 0}$:

$$V_{\mathbf{x}_d}(\mathbf{x}, t) = \mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)^\top \mathbf{P}\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t). \qquad (13)$$

Denoting $\nabla V_{\mathbf{x}_d}(\mathbf{x}, t) = 2\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)^\top \mathbf{P}$, we have that:

$$\lambda_{\min}(\mathbf{P})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2 \leq V_{\mathbf{x}_d}(\mathbf{x}, t) \leq \lambda_{\max}(\mathbf{P})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2, \quad (14)$$
$$\nabla V_{\mathbf{x}_d}(\mathbf{x}, t)(\mathbf{f}_{\mathbf{x}_d}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x})k_{\mathbf{x}_d}^{\text{fbl}}(\mathbf{x}, t)) \qquad (15)$$
$$\leq -\lambda_{\min}(\mathbf{Q})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2,$$

for all $\mathbf{x} \in \mathbb{R}^n$ and $t \in [\underline{t}, \overline{t}]$. Let $\gamma = 4\lambda_{\max}(\mathbf{P})^3/\lambda_{\min}(\mathbf{Q})^2$ and for a given disturbance signal $\mathbf{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ define $\|\mathbf{w}\|_\infty = \sup_{t \geq 0} \|\mathbf{w}(t)\|_2$. The preceding construction yields the following result:

Lemma 1. *Let $\overline{w} \in \mathbb{R}_{\geq 0}$, and for $t \in [\underline{t}, \overline{t}]$ define the set:*

$$\Omega_{\mathbf{x}_d}(t, \overline{w}) = \{\mathbf{x} \in \mathbb{R}^n \mid V_{\mathbf{x}_d}(\mathbf{x}, t) \leq \gamma \overline{w}^2\}. \qquad (16)$$

Let the controller $k : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}$ satisfy:

$$\nabla V_{\mathbf{x}_d}(\mathbf{x}, t)(\mathbf{f}_{\mathbf{x}_d}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x})k(\mathbf{x}, t)) \qquad (17)$$
$$\leq -\lambda_{\min}(\mathbf{Q})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2,$$

for almost all $t \in [\underline{t}, \overline{t}]$ and all $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$. Then for initial time $\underline{t}$, any initial condition $\mathbf{x}_0 \in \Omega_{\mathbf{x}_d}(\underline{t}, \overline{w})$, and any disturbance signal $\mathbf{w}$ satisfying $\|\mathbf{w}\|_\infty \leq \overline{w}$, we have that $I(\underline{t}, \mathbf{x}_0, \mathbf{w}) = [\underline{t}, \overline{t})$, and $\boldsymbol{\varphi}(t) \in \Omega_{\mathbf{x}_d}(t, \overline{w})$ for all $t \in [\underline{t}, \overline{t})$, and $\lim_{t \to \overline{t}} \boldsymbol{\varphi}(t)$ exists and satisfies $\lim_{t \to \overline{t}} \boldsymbol{\varphi}(t) \in \Omega_{\mathbf{x}_d}(\overline{t}, \overline{w})$.

The preceding result follows by a standard input-to-state stability argument [7]. For any $t \in [\underline{t}, \overline{t}]$, the set $\Omega_{\mathbf{x}_d}(t, \overline{w})$ captures how accurately the nonlinear closed-loop system (2) tracks $\mathbf{x}_d$ with disturbances. Importantly, for a given $t \in [\underline{t}, \overline{t}]$ the set $\Omega_{\mathbf{x}_d}(t, \overline{w})$ is convex – as we will see later, this property will allow us to efficiently synthesize a dynamically admissible trajectory $\mathbf{x}_d$ while knowing how accurately it will be tracked and ensuring state and input constraint satisfaction.

In contrast to cancelling the nonlinear dynamics to achieve linear dynamics as in (11), which may be unnecessary and inefficient [8], Control Lyapunov Functions (CLFs) provide an alternative method for synthesizing stabilizing controllers via convex optimization. In particular, we have that (15) implies:

$$\inf_{u \in \mathbb{R}} \nabla V_{\mathbf{x}_d}(\mathbf{x}, t)(\mathbf{f}_{\mathbf{x}_d}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x})u) \qquad (18)$$
$$\leq -\lambda_{\min}(\mathbf{Q})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2.$$

for all $\mathbf{x} \in \mathbb{R}^n$ and $t \in [\underline{t}, \overline{t}]$. Define a feed-forward controller $k_{\mathbf{x}_d}^{\text{ff}} : \mathbb{R}^n \times [\underline{t}, \overline{t}] \to \mathbb{R}$ as:

$$k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t) = -g(\mathbf{x})^{-1}\mathcal{F}_{\mathbf{x}_d}(\mathbf{x}, t). \qquad (19)$$

This feed-forward controller is incorporated into the following controller specified via a convex quadratic program (QP):

$$k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t) = \operatorname*{argmin}_{u \in \mathbb{R}} \frac{1}{2}\|u - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2^2 \qquad \text{(CLF-QP)}$$

$$\text{s.t. } \nabla V_{\mathbf{x}_d}(\mathbf{x}, t)(\mathbf{f}_{\mathbf{x}_d}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x})u) \leq -\lambda_{\min}(\mathbf{Q})\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2^2.$$

Note that the constraint in this controller ensures that $k_{\mathbf{x}_d}^{\text{clf}}$ satisfies the condition in (17).

III. BÉZIER CURVES & STATE CONSTRAINTS

In this section we present the first main contribution of this work by addressing how the properties of the low-level tracking controller can be used to place requirements on a dynamically admissible trajectory $\mathbf{x}_d$ that ensure state constraint satisfaction by the closed-loop nonlinear system (2) evolving under controllers such as $k_{\mathbf{x}_d}^{\text{fbl}}$ or $k_{\mathbf{x}_d}^{\text{clf}}$.

We first make the following assumption regarding the state constraints for the system:

Assumption 2. The state constraint set $\mathcal{X} \subset \mathbb{R}^n$ is a compact, convex polytope, with the existence of $\mathbf{L}_j \in \mathbb{R}^n$ and $\ell_j \in \mathbb{R}$ for $j = 1, \ldots, q$ such that $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid \forall j, \mathbf{L}_j^\top \mathbf{x} \leq \ell_j\}$. Furthermore, we have that $\mathbf{0} \in \text{Int}(\mathcal{X})$.

Given the above state constraints, it is not the case – even for a dynamically admissible trajectory satisfying $\mathbf{x}_d(t) \in \mathcal{X}$

for all $t \in [\underline{t}, \overline{t}]$ – that the state will remain inside the set $\mathcal{X}$, as we may have that $\Omega_{\mathbf{x}_d}(t, \overline{w}) \not\subseteq \mathcal{X}$ for some $t \in [\underline{t}, \overline{t}]$. To ensure these constraints are met by the closed-loop system without directly modifying the low-level control design, we will incorporate information about the low-level controller when constructing $\mathbf{x}_d$. The core tool that will enable incorporating this information is *Bézier curves* [38].

Let $T \in \mathbb{R}_{>0}$. A Bézier curve $r : [0, T] \to \mathbb{R}$ of order p is defined as:

$$r(\tau) = \boldsymbol{\xi}_0^\top \mathbf{z}(\tau), \qquad (20)$$

where $\boldsymbol{\xi}_0 = \begin{bmatrix} \xi_{0,0} & \dots & \xi_{0,p} \end{bmatrix}^\top \in \mathbb{R}^{p+1}$ is a vector with elements consisting of the $p + 1$ *control points*, $\xi_{0,i} \in \mathbb{R}$, of the curve and $\mathbf{z} : [0, T] \to \mathbb{R}^{p+1}$ is a Bernstein polynomial defined elementwise as:

$$z_i(\tau) = \binom{p}{i} \left(\frac{\tau}{T} \right)^i \left(1 - \frac{\tau}{T} \right)^{p-i}, \quad i = 0, \dots, p. \qquad (21)$$

The curve r is smooth, and there exists a matrix[4] $\mathbf{H} \in \mathbb{R}^{p+1 \times p+1}$ such that the jth derivative of r is given by:

$$r^{(j)}(\tau) = \frac{1}{T^j} \boldsymbol{\xi}_0^\top \mathbf{H}^j \mathbf{z}(\tau) \triangleq \boldsymbol{\xi}_j^\top \mathbf{z}(\tau). \qquad (22)$$

Consequently, $r^{(j)} : [0, T] \to \mathbb{R}$ is a Bézier curve of order p with the elements of $\boldsymbol{\xi}_j$ (which are uniquely and linearly defined by $\boldsymbol{\xi}_0$) as control points. Define the function $\mathbf{r} : [0, T] \to \mathbb{R}^n$:

$$\mathbf{r}(\tau) = \begin{bmatrix} r(\tau) & r^{(1)}(\tau) & \cdots & r^{(n-1)}(\tau) \end{bmatrix}^\top. \qquad (23)$$

There exists a matrix[4] $\mathbf{D} \in \mathbb{R}^{2n \times 2n}$ such that for any two vectors $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^n$, the unique Bézier curve r of order $2n - 1$ satisfying $\mathbf{r}(0) = \mathbf{x}_0$ and $\mathbf{r}(T) = \mathbf{x}_1$ with a vector of control points $\boldsymbol{\xi}_0 \in \mathbb{R}^{2n}$ is given by:

$$\boldsymbol{\xi}_0^\top = \begin{bmatrix} \mathbf{x}_0^\top & \mathbf{x}_1^\top \end{bmatrix} \mathbf{D}^{-1}. \qquad (24)$$

The following result shows how a sequence of points may be used to construct a set of Bézier curves that constitute a dynamically admissible trajectory for (1):

Lemma 2. *Let $N \in \mathbb{N}$, $\underline{t} \in \mathbb{R}_{\geq 0}$, and define $\overline{t} = \underline{t} + NT$. For $k = 0, \dots, N$, consider a collection of points $\{\mathbf{x}_k\}$ with $\mathbf{x}_k \in \mathbb{R}^n$ and define $t_k \in \mathbb{R}_{\geq 0}$ as $t_k = \underline{t} + kT$. For $k = 0, \dots, N-1$, let $r_k : [0, T] \to \mathbb{R}$ be a Bézier curve of order $2n - 1$ with control points $(\boldsymbol{\xi}_k)_0 = \begin{bmatrix} (\xi_k)_{0,0} & \dots & (\xi_k)_{0,2n-1} \end{bmatrix}^\top \in \mathbb{R}^{2n}$ given by:*

$$(\boldsymbol{\xi}_k)_0^\top = \begin{bmatrix} \mathbf{x}_k^\top & \mathbf{x}_{k+1}^\top \end{bmatrix} \mathbf{D}^{-1}. \qquad (25)$$

Defining the functions $\mathbf{r}_k : [0, T] \to \mathbb{R}^n$ as in (23), we have that the function $\mathbf{x}_d : [\underline{t}, \overline{t}] \to \mathbb{R}^n$ defined as:

$$\begin{aligned} \mathbf{x}_d(t) &= \mathbf{r}_k(t - t_k), \, t \in [t_k, t_{k+1}), \\ \mathbf{x}_d(\overline{t}) &= \mathbf{x}_N, \end{aligned} \qquad (26)$$

is a dynamically admissible trajectory for the system (1).

---

[4]The matrices $\mathbf{H}$ and $\mathbf{D}$ are uniquely defined by the order of the Bézier curve p and can be constructed as shown in appendix.
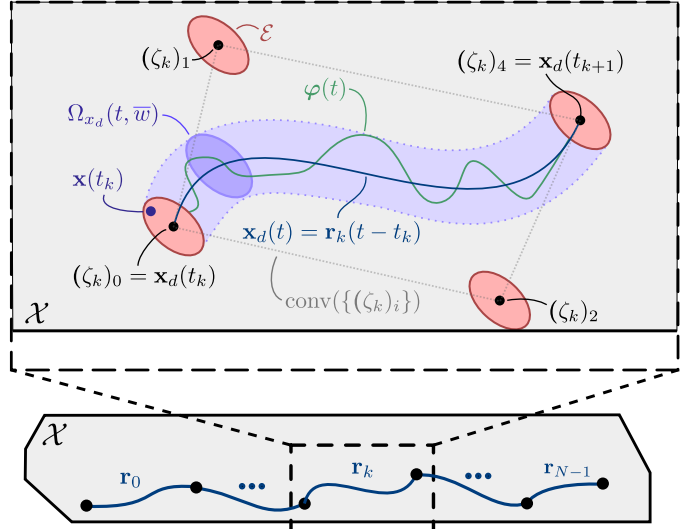

Fig. 2. A depiction of the proposed method, where the control points of the Bézier curve are constraint tightened by the size of the robust invariant tube coming from the low-level controller.

We note that the preceding result reduces planning of an (infinite dimensional) continuous time trajectory to planning a finite sequence of points. This aligns with planning dynamically admissible trajectories online in a multi-rate approach. While other classes of functions (such as general polynomials) may similarly be used to construct dynamically admissible trajectories for (1), the motivation for using Bézier curves lies in the convex hull relationship between the curve $\mathbf{r}_k$ and the control points $(\boldsymbol{\xi}_k)_0, \dots, (\boldsymbol{\xi}_k)_{n-1}$. More precisely, for $i = 0, \dots, 2n - 1$ denote:

$$(\boldsymbol{\zeta}_k)_i \triangleq \begin{bmatrix} (\xi_k)_{0,i} & \cdots & (\xi_k)_{n-1,i} \end{bmatrix}^\top \in \mathbb{R}^n. \qquad (27)$$

The points $(\boldsymbol{\xi}_k)_j$ can be viewed as the control points in time for the curve $r_k^{(j)}$, while $(\boldsymbol{\zeta}_k)_i$ reflects the control points for the curve $\mathbf{r}_k$ realized in state space. This enables the following:

Fact 1 ([38] §4). *We have that $\mathbf{r}_k(\tau) \in \mathrm{conv}(\{(\boldsymbol{\zeta}_k)_i\})$ for all $\tau \in [0, T]$.*

We may immediately use this property to establish the following result regarding state constraints:

Lemma 3. *Define the convex, compact set $\mathcal{E} \subseteq \mathbb{R}^n$ as:*

$$\mathcal{E} = \{ \mathbf{v} \in \mathbb{R}^n \mid \mathbf{v}^T \mathbf{P} \mathbf{v} \leq \gamma \overline{w}^2 \}. \qquad (28)$$

If $(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}$ for $i = 0, \dots, 2n-1$ and $k = 0, \dots, N-1$, then we have that $\Omega_{\mathbf{x}_d}(t, \overline{w}) \subseteq \mathcal{X}$ for all $t \in [\underline{t}, \overline{t}]$.

This result states that by constraining the Bézier curve control points, we can ensure the evolution of the system under the low-level controller satisfies state constraints. The requirement that $(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}$ can be expressed as an affine inequality constraint as in the following result:

Lemma 4. *We have that for $j = 1, \dots, q$:*

$$(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E} \Leftrightarrow \mathbf{L}_j^\top (\boldsymbol{\zeta}_k)_i \leq \ell_j - \sqrt{\gamma \overline{w}^2 \mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j}. \qquad (29)$$

IV. Input Constraints

In this section, we present the second main contribution of this work. We show how the structure of a low-level tracking controller can be used to place requirements on a dynamically admissible trajectory $\mathbf{x}_d$ to ensure input constraint satisfaction.

We will make the following assumption regarding input constraints for the system:

Assumption 3. The input constraint set $\mathcal{U} \subset \mathbb{R}$ is given by $\mathcal{U} = [-u_{\max}, u_{\max}]$ for some $u_{\max} \in \mathbb{R}_{>0}$.

Neither of the controllers $k^{\text{fbl}}_{\mathbf{x}_d}$ or $k^{\text{clf}}_{\mathbf{x}_d}$ are necessarily required to take values in the set $\mathcal{U}$. Thus, satisfying input constraints may require violating the inequality constraint in (17), potentially invalidating the claim that $\boldsymbol{\varphi}(t) \in \Omega_{\mathbf{x}_d}(t, \overline{w})$ for all $t \in [\underline{t}, \overline{t}]$. To address this limitation, knowledge of how much control action is required by the controller to track the reference trajectory under disturbances should be incorporated when synthesizing $\mathbf{x}_d$.

To this end, we state the following definitions. For $\alpha, \beta \in \mathbb{R}_{\geq 0}$, define the matrix $\mathbf{M}_{\alpha,\beta} \in \mathbb{S}^2_{\succeq 0}$ and the functions $\mathbf{N}_{\alpha,\beta} : \mathcal{X} \to \mathbb{R}^2_{\geq 0}$ and $\Gamma_{\alpha,\beta} : \mathcal{X} \to \mathbb{R}_{\geq 0}$ as:

$$\mathbf{M}_{\alpha,\beta} = \boldsymbol{\pi}_{\text{PSD}}\left(\begin{bmatrix} 2\alpha\beta & \beta \\ \beta & 0 \end{bmatrix}\right), \tag{30}$$

$$\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}) = \begin{bmatrix} 2\alpha\beta\overline{e} + \alpha|g(\overline{\mathbf{x}})^{-1}| + \beta\|\mathbf{K}\|_2\overline{e} \\ |g(\overline{\mathbf{x}})^{-1}| + \beta\overline{e} \end{bmatrix}, \tag{31}$$

$$\Gamma_{\alpha,\beta}(\overline{\mathbf{x}}) = \overline{e}(\beta\overline{e} + |g(\overline{\mathbf{x}})^{-1}|)(\alpha + \|\mathbf{K}\|_2), \tag{32}$$

where $\boldsymbol{\pi}_{\text{PSD}} : \mathbb{S}^2 \to \mathbb{S}^2_{\succeq 0}$ denotes the projection from symmetric matrices to symmetric positive semidefinite matrices, $\overline{e} \triangleq \sqrt{\gamma\overline{w}^2/\lambda_{\min}(\mathbf{P})}$, and $\mathbf{K}$ is defined in (10). Given these definitions, we state one of our main results:

Theorem 1. There exists constants $\underline{\alpha}, \underline{\beta} \in \mathbb{R}_{\geq 0}$ such that if $\alpha \geq \underline{\alpha}$, $\beta \geq \underline{\beta}$, and $\Omega_{\mathbf{x}_d}(t, \overline{w}) \subseteq \mathcal{X}$ for all $t \in [\underline{t}, \overline{t}]$, then for any collection of points $\{\overline{\mathbf{x}}_k\}$ with $\overline{\mathbf{x}}_k \in \mathcal{X}$ for $k = 0, \dots, N-1$, we have that for all $t \in [t_k, t_{k+1})$:

$$\|k^{\text{fbl}}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2 \leq \frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta}\boldsymbol{\sigma}_{\mathbf{x}_d}(t) \tag{33}$$
$$+ \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k),$$

for all $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$, where $\boldsymbol{\sigma}_{\mathbf{x}_d} : [\underline{t}, \overline{t}] \to \mathbb{R}^2_{\geq 0}$ is defined as:

$$\boldsymbol{\sigma}_{\mathbf{x}_d}(t) = \begin{bmatrix} \|\mathbf{x}_d(t) - \overline{\mathbf{x}}_k\|_2 \\ \|\dot{x}^n_d(t) - f(\overline{\mathbf{x}}_k)\|_2 \end{bmatrix}, \quad t \in [t_k, t_{k+1}). \tag{34}$$

This result is motivated by the key observation that the upper bound achieved in (33) is convex in the quantity $\boldsymbol{\sigma}_{\mathbf{x}_d}(t)$ for each $t \in [\underline{t}, \overline{t}]$, such that the constraint:

$$\frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta}\boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) \leq u_{\max}. \tag{35}$$

is a convex quadratic inequality constraint in the quantity $\boldsymbol{\sigma}_{\mathbf{x}_d}(t)$. Given that the function $\boldsymbol{\sigma}_{\mathbf{x}_d}$ is defined by Bézier control points, we seek to translate this constraint into one on control points.

Remark 1. We note that the proof of Theorem 1 establishes the existence of values of $\underline{\alpha}$ and $\underline{\beta}$ through Lipschitz properties of the dynamics. In practice, it may be difficult to compute these values, and they may not necessarily be the minimum values for which this result holds. Moreover, choosing very large values of α and β may lead to conservative behavior, as the constraint in (35) will constrain the dynamically admissible trajectory $\mathbf{x}_d$ to a small neighborhood of $\overline{\mathbf{x}}_k$. These issues are not unexpected, as the challenge of input constraint satisfaction for general nonlinear systems is known to be quite difficult. Instead, with this result we seek to highlight an important monotonic structural property of the system that permits a well-posed and practical approach for achieving input constraint satisfaction. In particular, one may begin with small values of α and β and increase them until the closed-loop nonlinear system meets input constraints. We will demonstrate this type of procedure in Section VI.

Before relating Theorem 1 to the Bézier control points defining $\mathbf{x}_d$, we state the following lemma:

Lemma 5. *For any $\mathbf{x} \in \mathbb{R}^n$, we have that:*

$$\|\mathbf{r}_k(\tau) - \mathbf{x}\|_2 \leq \sup_i \|(\boldsymbol{\zeta}_k)_i - \mathbf{x}\|_2, \tag{36}$$

$$\|r^{(n)}_k(\tau) - f(\mathbf{x})\|_2 \leq \sup_i \|(\xi_k)_{n,i} - f(\mathbf{x})\|_2, \tag{37}$$

for all $\tau \in [0, T]$.

With this result, we now state one of our main results for tractably enforcing input bounds:

Lemma 6. *If given a collection of points $\{\overline{\mathbf{x}}_k\}$ with $\overline{\mathbf{x}}_k \in \mathcal{X}$ for $k = 0, \dots, N-1$, there exists $\mathbf{s}_k \in \mathbb{R}^2_{\geq 0}$ such that:*

$$\begin{bmatrix} \|(\boldsymbol{\zeta}_k)_i - \overline{\mathbf{x}}_k\|_2 \\ \|(\xi_k)_{n,i} - f(\overline{\mathbf{x}}_k)\|_2 \end{bmatrix} \leq \mathbf{s}_k, \tag{38}$$

$$\frac{1}{2}\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta}\mathbf{s}_k + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) \leq u_{\max}, \tag{39}$$

for $i = 0, \dots, 2n-1$ and $k = 0, \dots, N-1$, then we have that the inequality (35) is satisfied with $\boldsymbol{\sigma}_{\mathbf{x}_d}(t)$ defined as in (34) for all $t \in [\underline{t}, \overline{t}]$.

A consequence of this result is that for sufficiently high values of α and β, meeting the conditions of Lemma 6 implies $\|k^{\text{fbl}}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2 \leq u_{\max}$ for all $t \in [\underline{t}, \overline{t}]$ and $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$. Moreover, the constraint (38) is a second-order cone constraint and the constraint (39) is a convex quadratic constraint (which may be reformulated as a second-order cone constraint, see the appendix), and thus they may be incorporated into a convex program for determining Bézier control points. Lastly, we state the following corollary relating bounds on $k^{\text{fbl}}_{\mathbf{x}_d}$ and $k^{\text{clf}}_{\mathbf{x}_d}$:

Corollary 1. *If the function $k^{\text{fbl}}_{\mathbf{x}_d}$ is bounded as in (33) for all $t \in [t_k, t_{k+1})$ and $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$, then we have that:*

$$\|k^{\text{clf}}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2 \leq \frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta}\boldsymbol{\sigma}_{\mathbf{x}_d}(t) \tag{40}$$
$$+ \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k),$$

for all $t \in [t_k, t_{k+1})$ and all $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$.

V. Multi-Rate Control Architecture

Utilizing the developments presented in the previous sections, we now construct a multi-rate control architecture which iteratively produces dynamically admissible trajectories for the system (1) and tracks them with the low-level controller designed in Section II. Importantly, by achieving robustness to disturbances with the low-level controller, the trajectory planning algorithm can reason about a disturbance-free system.

A. Model Predictive Control

In this section we establish how to compute the collection of points $\{\mathbf{x}_k\}$ used to define $\mathbf{x}_d$ in Lemma 2 while meeting the desired constraints on the Bézier control points. Consider a collection of points $\{\overline{\mathbf{x}}_k\}$ with $\overline{\mathbf{x}}_k \in \mathcal{X}$ and $\{\overline{u}_k\}$ with $\overline{u}_k \in \mathbb{R}$ for $k = 0, \ldots N-1$. To incorporate information about the system dynamics when synthesizing $\mathbf{x}_d$ as in Lemma 2, we will use linearizations of the system dynamics (1) around these collections of points. This approximation of the dynamics will provide constraints on sequential state points (and the corresponding Bézier control points as defined by (25)) defining $\mathbf{x}_d$. We neglect the disturbances $\mathbf{w}$ in this approximation as the low-level controller rejects these disturbances and provides a robust invariant set around $\mathbf{x}_d$. More precisely, consider a linear, temporal discretization of (1):

$$\mathbf{x}_{k+1} = \mathbf{A}(\overline{\mathbf{x}}_k, \overline{u}_k)\mathbf{x}_k + \mathbf{B}(\overline{\mathbf{x}}_k)u_k + \mathbf{C}(\overline{\mathbf{x}}_k, \overline{u}_k), \qquad (41)$$

where $\mathbf{A} : \mathcal{X} \times \mathbb{R} \to \mathbb{R}^{n \times n}, \mathbf{B} : \mathcal{X} \to \mathbb{R}^n$, and $\mathbf{C} : \mathcal{X} \times \mathbb{R} \to \mathbb{R}^n$ come from linearizing and taking the exact temporal discretization[5] (with sample period T) of the dynamics in (1). For notational simplicity let us define:

$$\mathbf{A}_k \triangleq \mathbf{A}(\overline{\mathbf{x}}_k, \overline{u}_k), \quad \mathbf{B}_k \triangleq \mathbf{B}(\overline{\mathbf{x}}_k), \quad \mathbf{C}_k \triangleq \mathbf{C}(\overline{\mathbf{x}}_k, \overline{u}_k). \quad (42)$$

Given these, let us denote the state at a time $t \in \mathbb{R}_{\geq 0}$ by $\mathbf{x}(t)$. Building upon the previous two sections, we propose a Finite Time Optimal Control Problem (FTOCP):

$$\min_{\substack{u_k, \mathbf{x}_k \\ \mathbf{s}_k, \boldsymbol{\xi}_k}} \sum_{k=0}^{N-1} h(\mathbf{x}_k, u_k) + J(\mathbf{x}_N) \qquad \text{(FTOCP)}$$

$$\text{s.t.} \quad \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k u_k + \mathbf{C}_k, \qquad (43a)$$

$$\mathbf{x}_0 \in \mathbf{x}(t) \oplus \mathcal{E}, \qquad (43b)$$

$$\mathbf{x}_N = \mathbf{0}, \qquad (43c)$$

$$(\boldsymbol{\xi}_k) = \begin{bmatrix} \mathbf{x}_k^\top & \mathbf{x}_{k+1}^\top \end{bmatrix} \mathbf{D}^{-1}, \qquad (43d)$$

$$(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}, \qquad \forall i \in \mathcal{I} \qquad (43e)$$

$$\begin{bmatrix} \|(\boldsymbol{\zeta}_k)_i - \overline{\mathbf{x}}_k\|_2 \\ \|(\xi_k)_{n,i} - f(\overline{\mathbf{x}}_k)\|_2 \end{bmatrix} \leq \mathbf{s}_k, \qquad \forall i \in \mathcal{I} \qquad (43f)$$

$$\frac{1}{2}\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta}\mathbf{s}_k + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k$$
$$\qquad\qquad + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) \leq u_{\max}, \quad (43g)$$

where $h : \mathcal{X} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a convex stage cost, $J : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a convex terminal cost, and $\mathcal{I} = \{0, \ldots, 2n-1\}$. The constraint in (43a) requires that the sequence of discrete points

---

[5]See the appendix for a formula for these linearizations and discretizations.

defining $\mathbf{x}_d$ satisfy a linear, discrete time approximation of the system dynamics. The constraint in (43b) requires that the beginning of $\mathbf{x}_d$ is close to the current state $\mathbf{x}(t)$, such that $\mathbf{x}(t) \in \Omega(t, \overline{w})$ as required by Lemma 1. The constraint in (43c) requires the end of $\mathbf{x}_d$ to be placed at origin. The constraints in (43d)-(43g) relate the discrete points $\mathbf{x}_k$ to Bézier control points, and consequently the continuous trajectory $\mathbf{x}_d$ tracked by the low-level controller. Note that as in Fact 1, the coefficients $(\boldsymbol{\xi})_k$ and $(\boldsymbol{\zeta}_k)_i$ are linearly related for $i = 0, \ldots, 2n-1$, a constraint implicitly assumed in (FTOCP). If h and J are positive definite quadratic functions, (FTOCP) is a second-order cone program (SOCP), which can be efficiently solved via standard solvers [41].

Remark 2. Note that we do not explicitly enforce input constraints on the decision variables u_k. Instead, constraints are induced on these decision variables through the linear dynamics constraint (43a) and the constraints on the Bézier coefficients in (43d) and (43f)-(43g). Moreover, these constraints ensure that the low-level controller will satisfy input constraints as desired.

B. The Multi-Rate Architecture

We now present the multi-rate architecture that integrates the low-level controller design posed in Section II with the preceding trajectory planner encoded in (FTOCP).

We first recall the role T plays in dynamically admissible trajectories synthesized through Bézier curves as in Lemma 2, as well as its role as a sampling period for the temporal discretization established in (41). Let us denote $\mathcal{T} = \cup_{i=0}^{\infty}\{iT\}$. This set serves to index the discrete points in time (separated by T) at which a dynamically admissible trajectory for the system will be replanned by solving the (FTOCP). The multi-rate architecture is initialized at time $t = 0$ with collections of points $\{\overline{\mathbf{x}}_{k|0}\}$ and $\{\overline{u}_{k|0}\}$ with $\overline{\mathbf{x}}_{k|0} \in \mathcal{X}$ and $\overline{u}_{k|0} \in \mathbb{R}$ for $k = 0, \ldots, N-1$. Let us denote the linearized and discretized dynamics computed around these collections by $\{\mathbf{Lin}_{k|0}\} = \{(\mathbf{A}_{k|0}, \mathbf{B}_{k|0}, \mathbf{C}_{k|0})\}$.

Assumption 4. Given an initial condition $\mathbf{x}(0) \in \mathcal{X}$, (FTOCP) is feasible using $\{\mathbf{Lin}_{k|0}\}$.

Algorithm 1 $u = \text{C-MPC}(\mathbf{x}, t)$
1: **if** $t \in \mathcal{T} = \cup_{i=0}^{\infty}\{iT\}$ **then**
2:      Compute $\{\mathbf{Lin}_{k|i}\}$ in (41) about $\{\overline{\mathbf{x}}_{k|i}\}$ and $\{\overline{u}_{k|i}\}$;
3:      Solve (FTOCP) with $\{\mathbf{Lin}_{k|i}\}$;
4:      **if** (FTOCP) is infeasible **then**
5:          $\{\mathbf{Lin}_{k|i}\} \leftarrow \{\mathbf{Lin}_{1|i-1}, \ldots, \mathbf{Lin}_{N-1|i-1}, \mathbf{Lin}_O\}$;
6:          Solve (FTOCP) with $\{\mathbf{Lin}_{k|i}\}$;
7:      **end if**
8:      $\{\overline{\mathbf{x}}_{k|i+1}\} \leftarrow \{\mathbf{x}_{1|i}^*, \ldots, \mathbf{x}_{N-1|i}^*, \mathbf{x}_{N|i}^*\}$;
9:      $\{\overline{u}_{k|i+1}\} \leftarrow \{u_{1|i}^*, \ldots, u_{N-1|i}^*, 0\}$;
10: **end if**
11: Calculate $\mathbf{x}_d|i$ from $\{\mathbf{x}_{k|i}^*\}$, as in (25)–(26);
12: **return** $u = k_{\mathbf{x}_d|i}^{\text{clf}}(\mathbf{x}, t)$;

We now describe our multi-rate framework as summarized in Algorithm 1. As in Line 1, let $t \in \mathcal{T}$ such that $t = iT$ for some $i \in \mathbb{Z}$. In Line 2, the linearized and discretized dynamics are computed around the collections $\{\overline{\mathbf{x}}_{k|i}\}$ and $\{\overline{u}_{k|i}\}$, and are denoted by $\{\mathbf{Lin}_{k|i}\} = \{(\mathbf{A}_{k|i}, \mathbf{B}_{k|i}, \mathbf{C}_{k|i})\}$. In Line 3 these dynamics are used to solve the (FTOCP) using the state at the current time, $\mathbf{x}(t)$, in (43b). If the (FTOCP) is feasible, it returns collections of points $\{\mathbf{x}_{k|i}^*\}$ with $\mathbf{x}_{k|i}^* \in \mathcal{X}$ for $k = 0, \ldots, N$ and $\{u_{k|i}^*\}$ with $u_{k|i}^* \in \mathbb{R}$ for $k = 0, \ldots, N-1$. If (FTOCP) is infeasible, in Line 5 we set the linearized and discretized dynamics $\{\mathbf{Lin}_{k|i}\}$ to the previous linearization shifted by one and appending the linearization and discretization around the origin, denoted $\mathbf{Lin}_O = (\mathbf{A}(\mathbf{0}, 0), \mathbf{B}(\mathbf{0}), \mathbf{C}(\mathbf{0}, 0))$. In Line 6 we solve (FTOCP) and similarly return collections of points $\{\mathbf{x}_{k|i}^*\}$ and $\{u_{k|i}^*\}$. As we will show in Theorem 2, our assumption about feasibility at time $t = 0$ will ensure that switching to this set of linearizations will always ensure (FTOCP) is feasible. In Line 8–9 the collection $\{\mathbf{x}_{k|i}^*\}$ is shifted and the collection $\{u_{k|i}^*\}$ is shifted and appended with 0 to define collections $\{\overline{\mathbf{x}}_{k|i+1}\}$ and $\{\overline{u}_{k|i+1}\}$ used for linearization and discretization in the next iteration. In Line 11 the collection $\{\mathbf{x}_{k|i}^*\}$ is then used to define a dynamically admissible trajectory $\mathbf{x}_d|i$ as in Lemma 2, which yields a corresponding low-level controller $k_{\mathbf{x}_d|i}^{\mathrm{clf}}$ that defines the output of our algorithm. We may view our algorithm as a time-varying controller that yields a closed-loop system (2). Importantly, our algorithm ensures state and input constraints are satisfied as the continuous time system evolves under this controller, as stated in the following theorem:

Theorem 2. *Suppose that* $\alpha \geq \underline{\alpha}$ *and* $\beta \geq \underline{\beta}$ *are such that* $\Gamma_{\alpha,\beta}(\mathbf{0}) \leq u_{\max}$. *Let* (FTOCP) *be defined with* α *and* β, *and consider the closed-loop system* (2) *with a feedback controller given by C-MPC in Algorithm 1 and a disturbance signal satisfying* $\|\mathbf{w}\|_\infty \leq \overline{w}$. *If* $\mathbf{0} \in \mathcal{X} \ominus \mathcal{E}$ *and* (FTOCP) *is feasible at* $t_0 = 0$ *with initial condition* $\mathbf{x}(0) \in \mathcal{X}$, *then C-MPC is well-defined for all time, and the closed-loop system* (2) *satisfies state and input constraints.*

VI. SIMULATION

We consider the following nonlinear system in simulation:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \sin(x_1) + x_2^3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}.$$

The goal is to drive the system to the origin while satisfying state and input constraints for all time. Fig. 3 demonstrates that at different time scales, both with and without added disturbances, using only either a low-level or mid-level controller results in state and/or input violation, whereas the proposed combined approach is able to satisfy both for all time. Fig. 4 shows the behavior of the system for increasing values of α and β. As the parameter values increase, the planned MPC points become closer to reduce deviation from the linearization points, and in doing so the deviation of the low-level controller from the planned input u_k decreases as the system evolves from $\mathbf{x}_k^*$ to $\mathbf{x}_{k+1}^*$. Simulation code is provided at [42].
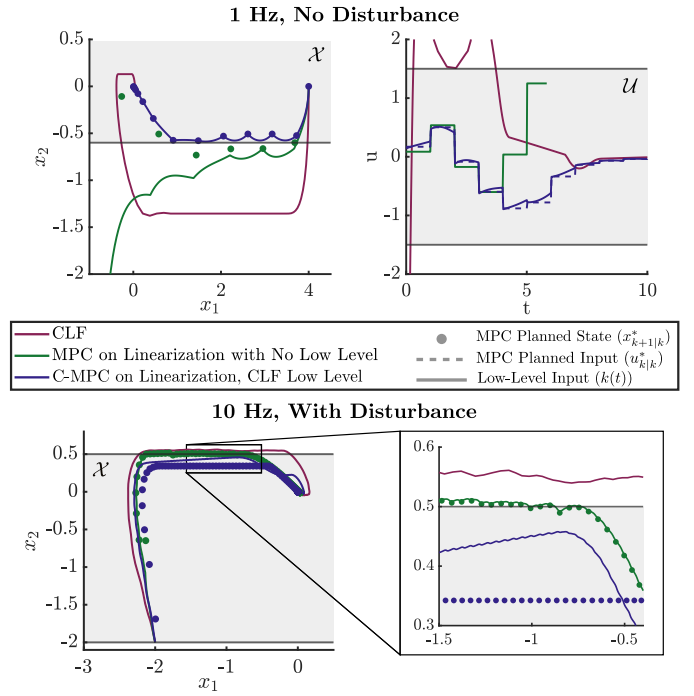

Fig. 3. Comparison of three control methods: only using a low level controller (CLF), applying MPC with no low-level controller, and applying the proposed C-MPC with a CLF at the low-level. In both scenarios, just using the low-level or mid-level controller separately yields both state and input violation.
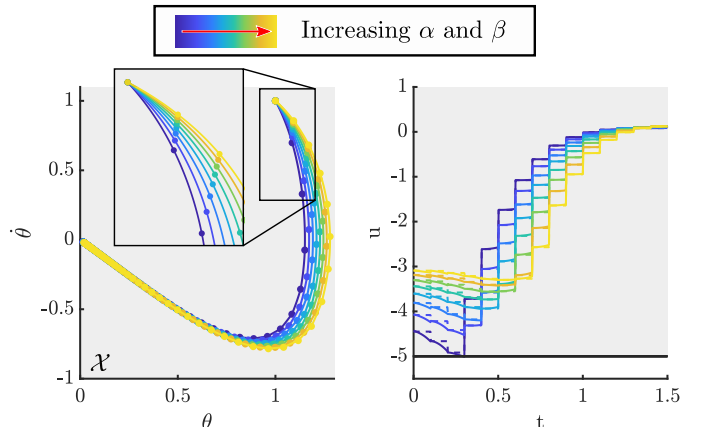

Fig. 4. The proposed C-MPC for increasing user parameter values α and β. Notice that as the parameters increase, the planned MPC points become spatially closer so as to reduce the linearization error, and in doing so the deviation of the low-level controller from the planned control input decreases.

VII. CONCLUSION AND FUTURE WORK

In conclusion, we have presented a multi-rate control architecture for nonlinear systems that utilizes MPC in conjunction with Bézier curves to iteratively plan continuous time trajectories that are tracked using Control Lyapunov Function based controllers. Our approach allows us to ensure that the low-level controller satisfies state and input constraints as it tracks the desired trajectory. We believe there are a number of meaningful directions for future work. First, in the pursuit of a truly multi-*rate* scheme, the low-level CLF control design could be adapted to the sampled-data setting [43]. Next, our

work uses the origin as the terminal set, but developing constructive approaches to synthesizing terminal sets using the ideas in [44] could greatly improve the feasible domain of our method. Lastly, we believe that the challenge of underactuation and unstable *zero-dynamics* may be best approached through a joint planning and low-level control mindset, and believe our work serves as a first step in this direction [45].

REFERENCES

[1] A. Isidori, *Nonlinear control systems*. Springer-Verlag London, 1995, vol. 3.

[2] P. Kokotović and M. Arcak, "Constructive nonlinear control: a historical perspective," *Automatica*, vol. 37, no. 5, pp. 637–662, 2001.

[3] H. K. Khalil and J. W. Grizzle, *Nonlinear Systems*. Upper Saddle River, NJ: Prentice Hall, 2002, vol. 3.

[4] R. Sepulchre, M. Jankovic, and P. V. Kokotovic, *Constructive nonlinear control*. Springer Science & Business Media, 2012.

[5] Z. Artstein, "Stabilization with relaxed controls," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 7, no. 11, pp. 1163–1173, 1983.

[6] E. D. Sontag, "A 'universal' construction of artstein's theorem on nonlinear stabilization," *Systems & Control Letters*, vol. 13, no. 2, pp. 117–123, 1989.

[7] ——, "Smooth stabilization implies coprime factorization," *Transactions on Automatic Control*, vol. 34, no. 4, pp. 435–443, 1989.

[8] R. Freeman and P. V. Kokotovic, *Robust nonlinear control design: state-space and Lyapunov techniques*. Birkhauser Basel, 1996.

[9] A. D. Ames and M. Powell, "Towards the unification of locomotion and manipulation through control lyapunov functions and quadratic programs," in *Control of Cyber-Physical Systems*. Springer, 2013, pp. 219–240.

[10] S. Kolathaya, J. Reher, A. Hereid, and A. D. Ames, "Input to state stabilizing control lyapunov functions for robust bipedal robotic locomotion," in *American Control Conference (ACC)*. IEEE, 2018, pp. 2224–2230.

[11] F. Allgower, R. Findeisen, Z. K. Nagy, *et al.*, "Nonlinear model predictive control: From theory to application," *Journal-Chinese Institute Of Chemical Engineers*, vol. 35, no. 3, pp. 299–316, 2004.

[12] F. Allgöwer and A. Zheng, *Nonlinear model predictive control*. Birkhäuser, 2012, vol. 26.

[13] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[14] J. Di Carlo, P. M. Wensing, B. Katz, G. Bledt, and S. Kim, "Dynamic locomotion in the mit cheetah 3 through convex model-predictive control," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE/RSJ, 2018, pp. 1–9.

[15] J.-P. Sleiman, F. Farshidian, M. V. Minniti, and M. Hutter, "A unified mpc framework for whole-body dynamic locomotion and manipulation," *Robotics and Automation Letters*, vol. 6, no. 3, pp. 4688–4695, 2021.

[16] P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, and D. Hrovat, "Predictive active steering control for autonomous vehicle systems," *Transactions on Control Systems Technology*, vol. 15, no. 3, pp. 566–580, 2007.

[17] D. Hrovat, S. Di Cairano, H. E. Tseng, and I. V. Kolmanovsky, "The development of model predictive control in automotive industry: A survey," in *International Conference on Control Applications*. IEEE, 2012, pp. 295–302.

[18] P. F. Lima, G. C. Pereira, J. Mårtensson, and B. Wahlberg, "Experimental validation of model predictive control stability for autonomous driving," *Control Engineering Practice*, vol. 81, pp. 244–255, 2018.

[19] S. Bengea, A. Kelman, F. Borrelli, R. Taylor, and S. Narayanan, "Model predictive control for mid-size commercial building hvac: Implementation, results and energy savings," in *International Conference on Building Energy and Environment*, 2012, pp. 979–986.

[20] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, "Model predictive control (mpc) for enhancing building and hvac system energy efficiency: Problem formulation, applications and opportunities," *Energies*, vol. 11, no. 3, p. 631, 2018.

[21] E. T. Maddalena, Y. Lian, and C. N. Jones, "Data-driven methods for building control—a review and promising future directions," *Control Engineering Practice*, vol. 95, p. 104211, 2020.

[22] U. Rosolia and F. Borrelli, "Learning how to autonomously race a car: a predictive control approach," *Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2713–2719, 2019.

[23] A. Liniger, A. Domahidi, and M. Morari, "Optimization-based autonomous racing of 1: 43 scale rc cars," *Optimal Control Applications and Methods*, vol. 36, no. 5, pp. 628–647, 2015.

[24] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using gaussian process regression," *Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2736–2743, 2019.

[25] D. Nešić, A. R. Teel, and P. V. Kokotović, "Sufficient conditions for stabilization of sampled-data nonlinear systems via discrete-time approximations," *Systems & Control Letters*, vol. 38, no. 4-5, pp. 259–270, 1999.

[26] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-jacobi reachability: A brief overview and recent advances," in *Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 2242–2253.

[27] A. Carvalho, Y. Gao, A. Gray, H. E. Tseng, and F. Borrelli, "Predictive control of an autonomous ground vehicle using an iterative linearization approach," in *International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2013, pp. 2335–2340.

[28] Y. Gao, A. Gray, H. E. Tseng, and F. Borrelli, "A tube-based robust nonlinear predictive control approach to semiautonomous ground vehicles," *Vehicle System Dynamics*, vol. 52, no. 6, pp. 802–823, 2014.

[29] M. Kögel and R. Findeisen, "Discrete-time robust model predictive control for continuous-time nonlinear systems," in *American Control Conference (ACC)*. IEEE, 2015, pp. 924–930.

[30] S. Yu, C. Maier, H. Chen, and F. Allgöwer, "Tube mpc scheme based on robust control invariant set with application to lipschitz nonlinear systems," *Systems & Control Letters*, vol. 62, no. 2, pp. 194–200, 2013.

[31] S. Singh, A. Majumdar, J.-J. Slotine, and M. Pavone, "Robust online motion planning via contraction theory and convex optimization," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 5883–5890.

[32] J. Köhler, R. Soloperto, M. A. Müller, and F. Allgöwer, "A computationally efficient robust model predictive control framework for uncertain nonlinear systems," *Transactions on Automatic Control*, vol. 66, no. 2, pp. 794–801, 2020.

[33] S. L. Herbert, M. Chen, S. Han, S. Bansal, J. F. Fisac, and C. J. Tomlin, "Fastrack: A modular framework for fast and guaranteed safe motion planning," in *Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 1517–1522.

[34] S. Singh, M. Chen, S. L. Herbert, C. J. Tomlin, and M. Pavone, "Robust tracking with model mismatch for fast and safe planning: an sos optimization approach," in *International Workshop on the Algorithmic Foundations of Robotics (WAFR)*. Springer, 2018, pp. 545–564.

[35] H. Yin, M. Bujarbaruah, M. Arcak, and A. Packard, "Optimization based planner–tracker design for safety guarantees," in *American Control Conference (ACC)*. IEEE, 2020, pp. 5194–5200.

[36] U. Rosolia, A. Singletary, and A. D. Ames, "Unified multi-rate control: from low level actuation to high level planning," *arXiv preprint arXiv:2012.06558*, 2020.

[37] U. Rosolia and A. D. Ames, "Multi-rate control design leveraging control barrier functions and model predictive control policies," *Control Systems Letters*, vol. 5, no. 3, pp. 1007–1012, 2021.

[38] M. Kamermans, "A primer on bézier curves," *(online book)*, 2020.

[39] G. E. Farin, *Curves and surfaces for CAGD: a practical guide*. Morgan Kaufmann, 2002.

[40] M. Muehlebach, C. Sferrazza, and R. D'Andrea, "Implementation of a parametrized infinite-horizon model predictive control scheme with stability guarantees," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 2723–2730.

[41] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.3.*, 2022.

[42] "C-mpc matlab code," *https://github.com/noelc-s/C-MPC*, 2022.

[43] A. J. Taylor, V. D. Dorobantu, Y. Yue, P. Tabuada, and A. D. Ames, "Sampled-data stabilization with control lyapunov functions via quadratically constrained quadratic programs," *Control Systems Letters*, vol. 6, pp. 680–685, 2022.

[44] T. Marcucci, J. Umenberger, P. A. Parrilo, and R. Tedrake, "Shortest paths in graphs of convex sets," *arXiv preprint arXiv:2101.11565*, 2021.

[45] J. Koehler, M. A. Muller, and F. Allgower, "Constrained nonlinear output regulation using model predictive control," *Transactions on Automatic Control*, 2021.

VIII. APPENDIX

A. Construction of $\mathbf{H}$:

Let r be a Bézier curve of order p defined as in (20):

$$r(\tau) = \sum_{i=0}^{p} \xi_i z_i(\tau),$$

with control points given by the elements of $\boldsymbol{\xi} = \begin{bmatrix} \xi_0 & \cdots & \xi_p \end{bmatrix}^\top$. The derivative of this curve is given by:

$$\dot{r}(\tau) = \sum_{i=0}^{p} \xi_i \dot{z}_i(\tau),$$

which we may equivalently express as [38, §13]:

$$\dot{r}(\tau) = \frac{1}{T}\sum_{i=0}^{p-1} p(\xi_{i+1} - \xi_i)z_i(\tau) \triangleq \frac{1}{T}\sum_{i=0}^{p-1}\vartheta_i z_i(\tau),$$

with $\vartheta_i \triangleq p(\xi_{i+1} - \xi_i)$ for $i = 0,\ldots,p-1$. We observe that $\dot{r}$ is a Bézier curve of order $p-1$ with control points ϑ_i/T. We may increase the order of $\dot{r}$ by one (making it a Bézier curve of order p) by the following transformation [38, §12]:

$$\dot{r}(\tau) = \frac{1}{T}\sum_{i=0}^{p}\left(\frac{(p-i)\vartheta_i + i\vartheta_{i-1}}{p}\right)z_i(\tau) \triangleq \frac{1}{T}\sum_{i=0}^{p}\omega_i z_i(\tau),$$

where $\vartheta_{-1} = \vartheta_p \triangleq 0$. Thus $\dot{r}$ is a Bézier curve of order p with control points ω_i/T. Noting that ω_i is a linear function of ξ_{i-1}, ξ_i, and ξ_{i+1}, we may rewrite $\dot{r}$ as:

$$\dot{r}(\tau) = \frac{1}{T}\boldsymbol{\xi}^\top \underbrace{(\mathbf{S}^\top \mathbf{R}^\top)}_{\mathbf{H}}\mathbf{z}(\tau),$$

where $\mathbf{S} \in \mathbb{R}^{p \times p+1}$ and $\mathbf{R} \in \mathbb{R}^{p+1 \times p}$ are defined by:

$$\mathbf{S}_{ii} = -p, \quad \mathbf{S}_{i,i+1} = p,$$
$$\mathbf{R}_{ii} = \frac{p+1-i}{p}, \quad \mathbf{R}_{i+1,i} = \frac{i}{p},$$

for $i = 1,\ldots,p$ with all other entries zero [38]. The matrix $\mathbf{S}$ corresponds to the differentiation of r, and the matrix $\mathbf{R}$ corresponds to increasing the order of the curve by one. Furthermore, we may reapply this transformation an arbitrary number of times to produce higher-order derivatives:

$$r^{(j)}(\tau) = \frac{1}{T^j}\boldsymbol{\xi}^\top \mathbf{H}^j \mathbf{z}(\tau).$$

B. Construction of $\mathbf{D}$

Let us denote:

$$\mathbf{x}_0 = \begin{bmatrix} x_{0,0} & \cdots & x_{0,n-1} \end{bmatrix}^\top,$$
$$\mathbf{x}_1 = \begin{bmatrix} x_{1,0} & \cdots & x_{1,n-1} \end{bmatrix}^\top.$$

Consider the set of equality constraints on the boundary of the Bézier curve:

$$r^{(j)}(0) = x_{0,j}, \quad j = 0,\ldots,n-1$$
$$r^{(j)}(T) = x_{1,j}, \quad j = 0,\ldots,n-1.$$

Substituting in the definition of a Bézier curve results in:

$$\boldsymbol{\xi}_j^\top \mathbf{z}(0) = x_{0,j}, \quad j = 0,\ldots,n-1$$
$$\boldsymbol{\xi}_j^\top \mathbf{z}(T) = x_{1,j}, \quad j = 0,\ldots,n-1.$$

As all of the control points are linear in $\boldsymbol{\xi}_0$, we can again reformulate this as:

$$\boldsymbol{\xi}_0^\top \frac{1}{T^j}\mathbf{H}^j \mathbf{z}(0) = x_{0,j}, \quad j = 0,\ldots,n-1$$
$$\boldsymbol{\xi}_0^\top \frac{1}{T^j}\mathbf{H}^j \mathbf{z}(T) = x_{1,j}, \quad j = 0,\ldots,n-1.$$

From this, we can construct a collection of linear equality constraints:

$$\boldsymbol{\xi}_0^\top \underbrace{\begin{bmatrix} \mathbf{D}_0 & \mathbf{D}_1 \end{bmatrix}}_{\mathbf{D}} = \begin{bmatrix} \mathbf{x}_0^\top & \mathbf{x}_1^\top \end{bmatrix},$$

with the matrices $\mathbf{D}_0 \in \mathbb{R}^{2n \times n}$ and $\mathbf{D}_1 \in \mathbb{R}^{2n \times n}$ defined as:

$$\mathbf{D}_0 = \begin{bmatrix} \frac{1}{T^0}\mathbf{H}^0\mathbf{z}(0) & \cdots & \frac{1}{T^{n-1}}\mathbf{H}^{n-1}\mathbf{z}(0) \end{bmatrix},$$
$$\mathbf{D}_1 = \begin{bmatrix} \frac{1}{T^0}\mathbf{H}^0\mathbf{z}(T) & \cdots & \frac{1}{T^{n-1}}\mathbf{H}^{n-1}\mathbf{z}(T) \end{bmatrix}.$$

C. Proof of Lemma 2:

Proof. Let $k \in \{0,\ldots,N-1\}$. As each function $r_k^{(j)}$, $j = 0,\ldots,n-1$, is a Bézier polynomial, the function $\mathbf{r}_k$ is continuously differentiable on the interval $(0,T)$ and the respective one-sided limits of the derivative exist at 0 and T. The definition of the Bézier control points in (25) implies that:

$$\mathbf{r}_k(T) = \begin{cases} \mathbf{r}_{k+1}(0) & \text{if } k \in \{0,\ldots,N-2\}, \\ \mathbf{x}_N & \text{if } k = N-1. \end{cases}$$

Thus the function $\mathbf{x}_d$ is continuous on $[\underline{t},\overline{t}]$, which with the previous differentiability properties, implies it is piecewise continuously differentiable on $[\underline{t},\overline{t}]$. Next, observe that:

$$\dot{\mathbf{x}}_d(t) = \dot{\mathbf{r}}_k(t - t_k), \, t \in (t_k, t_{k+1}),$$

where $\dot{\mathbf{r}}_k : (0,T) \to \mathbb{R}^n$ is given by:

$$\dot{\mathbf{r}}_k(\tau) = \begin{bmatrix} r_k^{(1)}(\tau) & \cdots & r_k^{(n)}(\tau) \end{bmatrix}^\top,$$

with $r_k^{(j)}$ defined as in (22). This may be rewritten as:

$$\dot{\mathbf{r}}_k(\tau) = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ 0 & \mathbf{0}^\top \end{bmatrix}\mathbf{r}_k(\tau) + \begin{bmatrix} \mathbf{0} \\ r_k^{(n)}(\tau) \end{bmatrix}.$$

Thus we have that:

$$\dot{\mathbf{x}}_d(t) = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ 0 & \mathbf{0}^\top \end{bmatrix}\mathbf{x}_d(t) + \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}r_k^{(n)}(t - t_k),$$

for $t \in (t_k, t_{k+1})$. Defining the function $u_d : [\underline{t},\overline{t}] \to \mathbb{R}$ as:

$$u_d(t) = g(\mathbf{x}_d(t))^{-1}(-f(\mathbf{x}_d(t)) + r_k^{(n)}(t - t_k)),$$

for $t \in [t_k, t_{k+1})$, the continuity of f, g on $\mathbb{R}^n$, the continuity of $r_k^{(n)}$ on $[0,T]$, and the fact $g(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, implies u_d is piecewise continuous. Moreover, we have that:

$$\dot{\mathbf{x}}_d(t) = \mathbf{f}(\mathbf{x}_d(t)) + \mathbf{g}(\mathbf{x}_d(t))u_d(t),$$

for almost all $t \in [\underline{t}, \overline{t}]$. Thus $\mathbf{x}_d$ is a dynamically admissible trajectory for the system (1). $\square$

D. Proof of Lemma 3:

Proof. Because $\mathcal{X}$ and $\mathcal{E}$ are convex, their Minkowski difference, $\mathcal{X} \ominus \mathcal{E}$, is also convex. As such, if $(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}$ for $i = 0, \ldots, 2n-1$, then $\operatorname{conv}(\{(\boldsymbol{\zeta}_k)_i\}) \subseteq \mathcal{X} \ominus \mathcal{E}$. As in Fact 1, the convex hull property of Bèzier curves implies that $\mathbf{r}_k(\tau) \in \operatorname{conv}(\{(\boldsymbol{\zeta}_k)_i\})$ for all $\tau \in [0, T]$. Thus we have that $\mathbf{r}_k(\tau) \in \mathcal{X} \ominus \mathcal{E}$ for all $\tau \in [0, T]$, implying that for any time $t \in [\underline{t}, \overline{t}]$ we have that $\mathbf{x}_d(t)$ as defined in (26) is contained in $\mathcal{X} \ominus \mathcal{E}$. Therefore, $\mathbf{x}_d(t) \oplus \mathcal{E} = \Omega_{\mathbf{x}_d}(t, \overline{w}) \subseteq \mathcal{X}$, as desired. $\square$

E. Proof of Lemma 4:

Proof. First, suppose that $(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}$ and let $j \in \{1, \ldots, q\}$. We then have that:

$$\mathbf{L}_j^\top ((\boldsymbol{\zeta}_k)_i + \mathbf{v}) \leq \ell_j, \quad \forall \mathbf{v} \in \mathcal{E}.$$

Equivalently, we have that:

$$\mathbf{L}_j^\top (\boldsymbol{\zeta}_k)_i \leq \ell_j - \sup_{\mathbf{v} \in \mathcal{E}} \mathbf{L}_j^\top \mathbf{v}.$$

Noting that $\mathcal{E} = \{\mathbf{v} \mid \mathbf{v}^\top \mathbf{P} \mathbf{v} \leq \gamma \overline{w}^2\}$ is a convex set as $\mathbf{P}$ is positive definite, taking the Lagrangian yields:

$$\sup_{\mathbf{v} \in \mathcal{E}} \mathbf{L}_j^\top \mathbf{v} = \inf_{\lambda \in \mathbb{R}_+} \sup_{\mathbf{v} \in \mathbb{R}^n} \mathbf{L}_j^\top \mathbf{v} - \lambda(\mathbf{v}^\top \mathbf{P} \mathbf{v} - \gamma \overline{w}^2),$$
$$\triangleq \inf_{\lambda \in \mathbb{R}_+} \sup_{\mathbf{v} \in \mathbb{R}^n} \mathcal{L}(\mathbf{v}, \lambda).$$

The stationarity conditions implies that:

$$\nabla_{\mathbf{v}} \mathcal{L}(\mathbf{v}^*, \lambda^*) = \mathbf{L}_j - 2\lambda^* \mathbf{P} \mathbf{v}^* = 0 \implies \mathbf{v}^* = \frac{1}{2\lambda^*} \mathbf{P}^{-1} \mathbf{L}_j.$$

Substituting this expression for $\mathbf{v}^*$ into the Lagrangian yields the dual problem:

$$\sup_{\mathbf{v} \in \mathcal{E}} \mathbf{L}_j^\top \mathbf{v} = \inf_{\lambda \in \mathbb{R}_+} \frac{1}{4\lambda} \mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j + \lambda \gamma \overline{w}^2.$$

The stationarity condition yields:

$$\lambda^* = \frac{1}{2} \sqrt{\frac{\mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j}{\gamma \overline{w}^2}},$$

From this we arrive at:

$$\sup_{\mathbf{v} \in \mathcal{E}} \mathbf{L}_j^\top \mathbf{v} = \sqrt{\gamma \overline{w}^2 \mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j},$$

whereby we then know that:

$$\mathbf{L}_j^\top (\boldsymbol{\zeta}_k)_i \leq \ell_j - \sqrt{\gamma \overline{w}^2 \mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j},$$

as desired.

Second, let $j \in \{1, \ldots, N\}$, and suppose that:

$$\mathbf{L}_j^\top (\boldsymbol{\zeta}_k)_i \leq \ell_j - \sqrt{\gamma \overline{w}^2 \mathbf{L}_j^\top \mathbf{P}^{-1} \mathbf{L}_j},$$

Then for $\mathbf{w} \in \mathcal{E}$, we have that:

$$\mathbf{L}_j^\top ((\boldsymbol{\zeta}_k)_i + \mathbf{w}) \leq \mathbf{L}_j^\top (\boldsymbol{\zeta}_k)_i + \sup_{\mathbf{v} \in \mathcal{E}} \mathbf{L}_j^\top \mathbf{v} \leq \ell_j,$$

following from our previous evaluation of the supremum. Thus we have $(\boldsymbol{\zeta}_k)_i + \mathbf{w} \in \mathcal{X}$, and since $\mathbf{w}$ and j were arbitrary, we have $(\boldsymbol{\zeta}_k)_i \in \mathcal{X} \ominus \mathcal{E}$. $\square$

F. Proof of Theorem 1

Proof. Let $k \in \{0, \ldots, N-1\}$, let $t \in [t_k, t_{k+1})$ and let $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$. For notational simplicity, let $g^\dagger : \mathbb{R}^n \to \mathbb{R}$ be defined as $g^\dagger(\mathbf{x}) = g(\mathbf{x})^{-1}$. We first bound the feed-forward input defined in (19) as follows:

$$\|k_{\mathbf{x}_d}^{\mathrm{ff}}(\mathbf{x}, t)\|_2 = \left\| g^\dagger(\mathbf{x}) \mathcal{F}_{\mathbf{x}_d}(\mathbf{x}, t) \right\|_2,$$
$$\leq \left\| g^\dagger(\mathbf{x}) \right\|_2 \| f(\mathbf{x}) - \dot{x}_d^n(t) \|_2.$$

with $\mathcal{F}_{\mathbf{x}_d}$ defined in (8). Given this, we have that:

$$\|k_{\mathbf{x}_d}^{\mathrm{fbl}}(\mathbf{x}, t)\|_2 \leq \|k_{\mathbf{x}_d}^{\mathrm{fbl}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\mathrm{ff}}(\mathbf{x}, t)\|_2 + \|k_{\mathbf{x}_d}^{\mathrm{ff}}(\mathbf{x}, t)\|_2, \quad (44)$$
$$\leq \|g^\dagger(\mathbf{x}) \mathbf{K}^\top \mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2 + \|k_{\mathbf{x}_d}^{\mathrm{ff}}(\mathbf{x}, t)\|_2,$$
$$\leq \left\| g^\dagger(\mathbf{x}) \right\|_2 (\|\mathbf{K}\|_2 \overline{e} + \|f(\mathbf{x}) - \dot{x}_d^n(t)\|_2).$$

As the set $\mathcal{X}$ is compact, we have that f and $g^\dagger$ are Lipschitz continuous on $\mathcal{X}$ (as g is Lipschitz continuous and non-zero) with Lipschitz constants $L_f, L_{g^\dagger} \in \mathbb{R}_{\geq 0}$, respectively. We continue by observing that:

$$\|g^\dagger(\mathbf{x})\|_2 \leq \|g^\dagger(\mathbf{x}) - g^\dagger(\mathbf{x}_d(t))\|_2 + \|g^\dagger(\mathbf{x}_d(t)) - g^\dagger(\overline{\mathbf{x}}_k)\|_2$$
$$+ \|g^\dagger(\overline{\mathbf{x}}_k)\|_2,$$
$$\leq L_{g^\dagger}(\|\mathbf{e}_{\mathbf{x}_d}(\mathbf{x}, t)\|_2 + \|\mathbf{x}_d(t) - \overline{\mathbf{x}}_k\|_2) + \|g^\dagger(\overline{\mathbf{x}}_k)\|_2,$$
$$\leq L_{g^\dagger}(\overline{e} + \|\mathbf{x}_d(t) - \overline{\mathbf{x}}_k\|_2) + \|g^\dagger(\overline{\mathbf{x}}_k)\|_2.$$

Similarly, we have that:

$$\|f(\mathbf{x}) - \dot{x}_d^n(t)\|_2 \leq L_f(\overline{e} + \|\mathbf{x}_d(t) - \overline{\mathbf{x}}_k\|_2)$$
$$+ \|f(\overline{\mathbf{x}}_k) - \dot{x}_d^n(t)\|_2.$$

The previous bounds allow us to construct a matrix $\mathbf{M} \in \mathbb{S}^2$ and functions $\mathbf{N} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}^2$, and $\Gamma : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ defined as:

$$\mathbf{M} = \begin{bmatrix} 2L_{g^\dagger} L_f & L_{g^\dagger} \\ L_{g^\dagger} & 0 \end{bmatrix},$$
$$\mathbf{N}(\overline{\mathbf{x}}_k) = \begin{bmatrix} 2L_{g^\dagger} L_f \overline{e} + L_f \|g^\dagger(\overline{\mathbf{x}}_k)\|_2 + L_{g^\dagger} \|\mathbf{K}\|_2 \overline{e} \\ \|g^\dagger(\overline{\mathbf{x}}_k)\|_2 + L_{g^\dagger} \overline{e} \end{bmatrix},$$
$$\Gamma(\overline{\mathbf{x}}_k) = \overline{e}(L_{g^\dagger} \overline{e} + \|g^\dagger(\overline{\mathbf{x}}_k)\|_2)(L_f + \|\mathbf{K}\|_2),$$

such that:

$$\|k_{\mathbf{x}_d}^{\mathrm{fbl}}(\mathbf{x}, t)\|_2 \leq \frac{1}{2} \boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M} \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \mathbf{N}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma(\overline{\mathbf{x}}_k).$$

where:

$$\boldsymbol{\sigma}_{\mathbf{x}_d}(t) = \begin{bmatrix} \|\mathbf{x}_d(t) - \overline{\mathbf{x}}_k\|_2 \\ \|\dot{x}_d^n(t) - f(\overline{\mathbf{x}}_k)\|_2 \end{bmatrix}, \quad t \in [t_k, t_{k+1}).$$

Let $\underline{\alpha} = L_f$ and $\underline{\beta} = L_{g^\dagger}$. We can then see that if both $\alpha \geq \underline{\alpha}$ and $\beta \geq \underline{\beta}$, then:

$$\mathbf{N}_{\alpha, \beta}(\overline{\mathbf{x}}_k) \geq \mathbf{N}(\overline{\mathbf{x}}_k),$$

where the inequality is element-wise, and:

$$\Gamma_{\alpha, \beta}(\overline{\mathbf{x}}_k) \geq \Gamma(\overline{\mathbf{x}}_k).$$

As the elements of $\mathbf{N}$ and $\mathbf{N}_{\alpha,\beta}$ are non-negative, we have:

$$\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{v} \geq \mathbf{N}(\overline{\mathbf{x}}_k)^\top \mathbf{v},$$

for any $\mathbf{v} \in \mathbb{R}^2_{\geq 0}$. Given the definition of $\boldsymbol{\sigma}_{\mathbf{x}_d}$ (with non-negative elements by definition of a norm), we thus have:

$$\|k^{\text{fbl}}_{\mathbf{x}_d}(\mathbf{x},t)\|_2 \leq \frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}\boldsymbol{\sigma}_{\mathbf{x}_d}(t) \\ +\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k).$$

We next observe that if both $\alpha \geq \underline{\alpha}$ and $\beta \geq \underline{\beta}$, then:

$$\frac{1}{2}\mathbf{v}^\top \mathbf{M}\mathbf{v} = L_{g^\dagger}L_f v_1^2 + L_{g^\dagger}v_1 v_2$$
$$\leq \alpha\beta v_1^2 + \beta v_1 v_2 \triangleq \frac{1}{2}\mathbf{v}^\top \widetilde{\mathbf{M}}_{\alpha,\beta}\mathbf{v},$$

for all $\mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^\top \in \mathbb{R}^2_{\geq 0}$, where:

$$\widetilde{\mathbf{M}}_{\alpha,\beta} = \begin{bmatrix} 2\alpha\beta & \beta \\ \beta & 0 \end{bmatrix}.$$

It can be seen that the matrix $\widetilde{\mathbf{M}}_{\alpha,\beta}$ will have both a positive and a negative eigenvalue for any $\alpha, \beta \in \mathbb{R}_{\geq 0}$, and thus using $\widetilde{\mathbf{M}}_{\alpha,\beta}$ directly in (35) will yield a non-convex constraint for an optimization program. To resolve this, we will project $\widetilde{\mathbf{M}}_{\alpha,\beta}$ onto the positive semidefinite cone to get $\mathbf{M}_{\alpha,\beta}$, such that:

$$\mathbf{M}_{\alpha,\beta} \triangleq \pi_{\text{PSD}}(\widetilde{\mathbf{M}}_{\alpha,\beta}), = \lambda_1(\widetilde{\mathbf{M}}_{\alpha,\beta})\mathbf{v}_1(\widetilde{\mathbf{M}}_{\alpha,\beta})\mathbf{v}_1(\widetilde{\mathbf{M}}_{\alpha,\beta})^\top,$$

where $\lambda_1(\widetilde{\mathbf{M}}_{\alpha,\beta})$ is the positive eigenvalue of $\widetilde{\mathbf{M}}_{\alpha,\beta}$, and:

$$\mathbf{v}_1(\widetilde{\mathbf{M}}_{\alpha,\beta}) = \frac{1}{\sqrt{1 + \lambda_1(\widetilde{\mathbf{M}}_{\alpha,\beta})^2}} \begin{bmatrix} \lambda_1(\widetilde{\mathbf{M}}_{\alpha,\beta}) \\ 1 \end{bmatrix}, \quad (45)$$

is the corresponding unit eigenvector. By construction, we have that $\mathbf{M}_{\alpha,\beta} \succeq \widetilde{\mathbf{M}}_{\alpha,\beta}$, and thus we may conclude that:

$$\frac{1}{2}\mathbf{v}^\top \mathbf{M}\mathbf{v} \leq \frac{1}{2}\mathbf{v}^\top \widetilde{\mathbf{M}}_{\alpha,\beta}\mathbf{v} \leq \frac{1}{2}\mathbf{v}^\top \mathbf{M}_{\alpha,\beta}\mathbf{v},$$

for all $\mathbf{v} \in \mathbb{R}^2_{\geq 0}$. Thus we can conclude that:

$$\|k^{\text{fbl}}_{\mathbf{x}_d}(\mathbf{x},t)\|_2 \leq \frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta}\boldsymbol{\sigma}_{\mathbf{x}_d}(t) \\ +\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k). \quad \square$$

Remark 3. The projection of the matrix $\widetilde{\mathbf{M}}_{\alpha,\beta}$ onto the positive semidefinite cone is a relaxation of the problem in that it will shrink the set of feasible dynamically admissible trajectories. In doing so, it provides a tractable way for guaranteeing that input bounds are met. Importantly, this is a type of "minimal" relaxation as the projection onto the positive semidefinite cone is the closest matrix that yields a convex inequality constraint.

G. Proof of Lemma 5:

Proof. Let $\mathbf{x} \in \mathbb{R}^n$. The convex hull property of Bézier curves implies that for any $\tau \in [0, T]$, we may write:

$$\mathbf{r}_k(\tau) = \sum_{i=0}^{2n-1} \lambda_i(\tau)(\boldsymbol{\zeta}_k)_i,$$

where $\lambda_i(\tau) \geq 0$ and $\sum_{i=0}^{2n-1}\lambda_i(\tau) = 1$. Thus we have that:

$$\|\mathbf{r}_k(\tau) - \mathbf{x}\|_2 = \left\| \sum_{i=0}^{2n-1} \lambda_i(\tau)\left((\boldsymbol{\zeta}_k)_i - \mathbf{x}\right) \right\|_2,$$

using the fact that $\sum_{i=0}^{2n-1}\lambda_i(\tau) = 1$. Moving the norm inside the sum, we have that:

$$\|\mathbf{r}_k(\tau) - \mathbf{x}\|_2 \leq \sum_{i=0}^{2n-1} \left\| \lambda_i(\tau)\left((\boldsymbol{\zeta}_k)_i - \mathbf{x}\right) \right\|_2,$$
$$= \sum_{i=0}^{2n-1} \lambda_i(\tau) \left\| (\boldsymbol{\zeta}_k)_i - \mathbf{x} \right\|_2,$$

as $\lambda_i(\tau) \geq 0$. We may further conclude that:

$$\|\mathbf{r}_k(\tau) - \mathbf{x}\|_2 \leq \sum_{i=0}^{2n-1} \lambda_i(\tau)\sup_i \left\| (\boldsymbol{\zeta}_k)_i - \mathbf{x} \right\|_2,$$
$$= \sup_i \left\| (\boldsymbol{\zeta}_k)_i - \mathbf{x} \right\|_2,$$

as desired. To establish the second part, we begin by noting that $r^{(n)}$ is a Bézier curve of order $2n-1$, such that:

$$\|r^{(n)}(\tau) - f(\mathbf{x})\|_2 = \left\| \sum_{i=0}^{2n-1} (\xi_k)_{n,i}z_i(\tau) - f(\mathbf{x}) \right\|,$$

Noting that:

$$\sum_{i=0}^{2n-1} z_i(\tau) = 1$$

for all $\tau \in [0, T]$ (see [39]), we have that:

$$\|r^{(n)}(\tau) - f(\mathbf{x})\|_2 = \left\| \sum_{i=0}^{2n-1} ((\xi_k)_{n,i} - f(\mathbf{x}))z_i(\tau) \right\|,$$
$$\leq \sum_{i=0}^{2n-1} \|((\xi_k)_{n,i} - f(\mathbf{x}))z_i(\tau)\|,$$
$$\leq \sum_{i=0}^{2n-1} |(\xi_k)_{n,i} - f(\mathbf{x})||z_i(\tau)|,$$

Noting that $z_i(\tau) \geq 0$ for all $\tau \in [0, T]$ and $i \in \{0, \ldots, 2n-1\}$, we then have that:

$$\|r^{(n)}(\tau) - f(\mathbf{x})\|_2 \leq \sum_{i=0}^{2n-1} |(\xi_k)_{n,i} - f(\mathbf{x})|z_i(\tau),$$
$$\leq \sum_{i=0}^{2n-1} \sup_i |(\xi_k)_{n,i} - f(\mathbf{x})|z_i(\tau),$$
$$\leq \sup_i |(\xi_k)_{n,i} - f(\mathbf{x})| \sum_{i=0}^{2n-1} z_i(\tau),$$
$$= \sup_i |(\xi_k)_{n,i} - f(\mathbf{x})|,$$

yielding the desired result. $\square$

H. Proof of Lemma 6

Proof. Assume that the inequalities in (38) and (39) hold for some $k \in \{0, \dots, N-1\}$. From Lemma 5, we have that:

$$\begin{bmatrix} \|\mathbf{r}_k(\tau) - \overline{\mathbf{x}}_k\|_2 \\ \|r_k^{(n)}(\tau) - f(\overline{\mathbf{x}}_k)\|_2 \end{bmatrix} \leq \mathbf{s}_k,$$

for all $\tau \in [0, T]$, where the inequality is element-wise. Given the definition of the dynamically admissible trajectory $\mathbf{x}_d$ in (26), we then have that:

$$\boldsymbol{\sigma}_{\mathbf{x}_d}(t) \leq \mathbf{s}_k$$

for all $t \in [t_k, t_{k+1})$. As the elements of $\mathbf{M}_{\alpha,\beta}$ are positive (as the elements of $\mathbf{v}_1(\widetilde{\mathbf{M}}_{\alpha,\beta})$ in (45) are both positive), and the elements of $\boldsymbol{\sigma}_{\mathbf{x}_d}(t)$ and $\mathbf{s}_k$ are non-negative, we have that:

$$\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta} \boldsymbol{\sigma}_{\mathbf{x}_d}(t) \leq \mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k.$$

Furthermore, as the elements of $\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)$ are non-negative, we have that:

$$\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) \leq \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k.$$

Therefore, we can conclude that:

$$\frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta} \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k)$$
$$\leq \frac{1}{2}\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k)$$
$$\leq u_{\max},$$

as enforced via (39). $\qquad\square$

I. Reformulation to a SOCP

Consider a positive semidefinite matrix $\mathbf{M}_{\alpha,\beta} \in \mathbb{S}_{\succeq 0}^2$. We may take its Cholesky decomposition, yielding:

$$\mathbf{M}_{\alpha,\beta} = \mathbf{L}_{\alpha,\beta} \mathbf{L}_{\alpha,\beta}^\top,$$

for some $\mathbf{L}_{\alpha,\beta} \in \mathbb{R}^{2\times 2}$. Let $\mathbf{s}_k \in \mathbb{R}^2$. We have that:

$$\frac{1}{2}\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)\mathbf{s}_k + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) \leq u_{\max}, \qquad (46)$$

if and only if there exists a $\sigma_k \in \mathbb{R}$ such that:

$$\left\| \begin{bmatrix} \mathbf{L}_{\alpha,\beta}^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{s}_k \\ \sigma_k \end{bmatrix} \right\|_2 \leq \sigma_k + \frac{1}{2}, \qquad (47)$$

and:

$$\sigma_k + \frac{1}{4} \leq -\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k - \Gamma_{\alpha,\beta} + u_{max}. \qquad (48)$$

To see the if direction, assume there exists a $\sigma_k \in \mathbb{R}$ such that (47) and (48) hold. We then have that:

$$\left\| \begin{bmatrix} \mathbf{L}_{\alpha,\beta}^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{s}_k \\ \sigma_k \end{bmatrix} \right\|_2^2 \leq \sigma_k^2 + \sigma_k + \frac{1}{4},$$

which may be rewritten as:

$$\mathbf{s}_k^\top \mathbf{L}_{\alpha,\beta} \mathbf{L}_{\alpha,\beta}^\top \mathbf{s}_k + \sigma_k^2 \leq \sigma_k^2 + \sigma_k + \frac{1}{4}.$$

Using the definition of $\mathbf{L}_{\alpha,\beta}$, we arrive at:

$$\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k \leq -\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top s_k - \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) + u_{max},$$

and thus have:

$$\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k + \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k) \leq u_{max}$$

For the only if direction, suppose that (46) is satisfied, and let:

$$\sigma_k = -\mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \mathbf{s}_k - \Gamma_{\alpha,\beta} + u_{max} - \frac{1}{4},$$

such that (48) is satisfied. Substituting this into (46) yields:

$$\mathbf{s}_k^\top \mathbf{M}_{\alpha,\beta} \mathbf{s}_k \leq \sigma_k + \frac{1}{4}.$$

Adding σ_k^2 to each side and using the definition of $\mathbf{L}_{\alpha,\beta}$ yields:

$$\mathbf{s}_k^\top \mathbf{L}_{\alpha,\beta} \mathbf{L}_{\alpha,\beta}^\top \mathbf{s}_k + \sigma_k^2 \leq \sigma_k^2 + \sigma_k + \frac{1}{4}.$$

This may be rewritten as:

$$\left\| \begin{bmatrix} \mathbf{L}_{\alpha,\beta}^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{s}_k \\ \sigma_k \end{bmatrix} \right\|_2^2 \leq \left(\sigma_k + \frac{1}{2} \right)^2.$$

Taking the square root of each side yields (47) as desired.

J. Proof of Corollary 1

Proof. Let $k \in \{0, \dots, N-1\}$, let $t \in [t_k, t_{k+1})$, and let $\mathbf{x} \in \Omega_{\mathbf{x}_d}(t, \overline{w})$. From (15) we know that $k_{\mathbf{x}_d}^{\text{fbl}}(\mathbf{x})$ is a feasible solution to the optimization problem defining $k_{\mathbf{x}_d}^{\text{clf}}$, and thus we may conclude:

$$\frac{1}{2}\|k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2^2 \leq \frac{1}{2}\|k_{\mathbf{x}_d}^{\text{fbl}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2^2.$$

From this we have that:

$$\|k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2 + \|k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2$$
$$\leq \|k_{\mathbf{x}_d}^{\text{fbl}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2 + \|k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2.$$

From the triangle inequality, we have that:

$$\|k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t)\|_2 \leq \|k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t) - k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2 + \|k_{\mathbf{x}_d}^{\text{ff}}(\mathbf{x}, t)\|_2.$$

Using this to replace the left-hand side of the inequality in (44), we may proceed as in the proof of Theorem 1 to arrive at:

$$\|k_{\mathbf{x}_d}^{\text{clf}}(\mathbf{x}, t)\|_2 \leq \frac{1}{2}\boldsymbol{\sigma}_{\mathbf{x}_d}(t)^\top \mathbf{M}_{\alpha,\beta} \boldsymbol{\sigma}_{\mathbf{x}_d}(t)$$
$$+ \mathbf{N}_{\alpha,\beta}(\overline{\mathbf{x}}_k)^\top \boldsymbol{\sigma}_{\mathbf{x}_d}(t) + \Gamma_{\alpha,\beta}(\overline{\mathbf{x}}_k). \quad \square$$

K. Linearization and Discretization:

We can linearize the dynamics of (1) to generate a linear, continuous time representation about the point $(\overline{\mathbf{x}}_k, \overline{u}_k)$ as:

$$\mathbf{A}_c(\overline{\mathbf{x}}_k, \overline{u}_k) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\overline{\mathbf{x}}_k) + \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\overline{\mathbf{x}}_k)\overline{u}_k,$$
$$\mathbf{B}_c(\overline{\mathbf{x}}_k) = \mathbf{g}(\overline{\mathbf{x}}_k),$$
$$\mathbf{C}_c(\overline{\mathbf{x}}_k, \overline{u}_k) = \mathbf{f}(\overline{\mathbf{x}}_k) + \mathbf{g}(\overline{\mathbf{x}}_k)\overline{u}_k - \mathbf{A}_c(\overline{\mathbf{x}}_k, \overline{u}_k)\overline{\mathbf{x}}_k - \mathbf{B}_c(\overline{\mathbf{x}}_k)\overline{u}_k.$$

We can then employ exact temporal discretization over a time interval T to obtain:

$$\mathbf{A}(\overline{\mathbf{x}}_k, \overline{u}_k) = e^{\mathbf{A}_c(\overline{\mathbf{x}}_k, \overline{u}_k)T},$$

$$\mathbf{B}(\overline{\mathbf{x}}_k) = \int_0^T e^{\mathbf{A}_c(\overline{\mathbf{x}}_k, \overline{u}_k)(T-\tau)} \mathbf{B}_c(\overline{\mathbf{x}}_k) d\tau,$$

$$\mathbf{C}(\overline{\mathbf{x}}_k, \overline{u}_k) = \int_0^T e^{\mathbf{A}_c(\overline{\mathbf{x}}_k, \overline{u}_k)(T-\tau)} \mathbf{C}_c(\overline{\mathbf{x}}_k, \overline{u}_k) d\tau.$$

L. Proof of Theorem 2

Proof. Let $i \in \mathbb{Z}_{\geq 0}$. Suppose that at time $t_i = iT$ and state $\mathbf{x}(t_i) \in \mathcal{X}$, we have that (FTOCP) is feasible using the collections of points $\{\overline{\mathbf{x}}_{k|i}\}$ and $\{\overline{u}_{k|i}\}$ for $k = 0, \dots, N-1$ and the corresponding linearizations $\{\mathbf{Lin}_{k|i}\}$. Let $\{\mathbf{x}^*_{k|i}\}$ for $k = 0, \dots, N$ and $\{u^*_{k|i}\}$, $\{(\boldsymbol{\xi}^*_{k|i})\}$, and $\{\mathbf{s}^*_{k|i}\}$ for $k = 0, \dots, N-1$ be the collection of points composing the solution to (FTOCP), and let $\mathbf{x}_d|i : [t_i, t_{i+1}] \to \mathcal{X}$ be the continuous reference trajectory defined as in Lemma 2. Given that $\|\mathbf{w}\|_\infty \leq \overline{w}$, Lemma 1 implies that $\boldsymbol{\varphi}(t) \in \Omega(t, \overline{w})$ for all $t \in [t_i, t_{i+1}]$. We have from Lemma 3 that $\Omega(t, \overline{w}) \subseteq \mathcal{X}$ for all $t \in [t_i, t_{i+1}]$, implying that $\boldsymbol{\varphi}(t) \in \mathcal{X}$ for all $t \in [t_i, t_{i+1}]$. Given this, we may further conclude from Theorem 1, Corollary 1, and Lemma 6 that:

$$\|k^{\text{clf}}_{\mathbf{x}_d}(\boldsymbol{\varphi}(t), t)\|_2 \leq u_{max} \implies k^{\text{clf}}_{\mathbf{x}_d}(\boldsymbol{\varphi}(t), t) \in \mathcal{U}.$$

for all $t \in [t_i, t_{i+1}]$.

To see that our algorithm is recursively feasible (i.e, feasible at time t_{i+1} given feasibility at time t_i), it is sufficient for us to show that (FTOCP) is feasible at the time t_{i+1} with $\mathbf{x}(t_{i+1}) = \boldsymbol{\varphi}(t_{i+1})$ and linearizations:

$$\{\mathbf{Lin}_{k|i+1}\} = \{\mathbf{Lin}_{1|i}, \dots, \mathbf{Lin}_{N-1|i}, \mathbf{Lin}_O\},$$

for $k = 0, \dots, N-1$, i.e., those calculated at time t_i shifted by one index and appended with the linearization at the origin. This reflects the case in which solving (FTOCP) with the linearizations about the previous optimal solution is infeasible, so it is sufficient to check feasibility only in this case.

To show that under such these conditions a feasible solution for (FTOCP) is given by:

$$\{\widehat{\mathbf{x}}_{k|i+1}\} = \left\{\mathbf{x}^*_{1|i}, \dots, \mathbf{x}^*_{N|i}, \mathbf{0}\right\}, \qquad k = 0, \dots, N,$$

$$\{\widehat{u}_{k|i+1}\} = \left\{u^*_{1|i}, \dots, u^*_{N-1|i}, 0\right\}, \qquad k = 0, \dots, N-1,$$

$$\{\widehat{\mathbf{s}}_{k|i+1}\} = \left\{\mathbf{s}^*_{1|i}, \dots, \mathbf{s}^*_{N-1|i}, \mathbf{0}\right\}, \qquad k = 0, \dots, N-1,$$

$$\{(\widehat{\boldsymbol{\xi}}_{k|i+1})\} = \left\{(\boldsymbol{\xi}^*_{1|i}), \dots, (\boldsymbol{\xi}^*_{N-1|i}), \mathbf{0}\right\}, \quad k = 0, \dots, N-1,$$

First, we observe that:

$$\mathbf{x}^*_{k+1|i} = \mathbf{A}_{k|i}\mathbf{x}^*_{k|i} + \mathbf{B}_{k|i}u^*_{k|i} + \mathbf{C}_{k|i} \qquad (49)$$

for $k = 1, \dots, N-1$ by the previous solution. Thus:

$$\widehat{\mathbf{x}}_{k+1|i+1} = \mathbf{A}_{k|i+1}\widehat{\mathbf{x}}_{k|i+1} + \mathbf{B}_{k|i+1}\widehat{u}_{k|i+1} + \mathbf{C}_{k|i+1},$$

for $k = 0, \dots, N-2$. Noting that $\widehat{\mathbf{x}}_{N-1|i+1} = \mathbf{x}^*_{N|i} = \mathbf{0}$ as required by the previous solution, and the fact that $\widehat{\mathbf{x}}_{N|i+1} = \mathbf{0}$

and $\widehat{u}_{N-1|i+1} = 0$, and $\mathbf{C}_{N-1|i} = \mathbf{C}(\mathbf{0}, 0) = \mathbf{0}$ we have that:

$$\widehat{\mathbf{x}}_{N|i+1} = \mathbf{A}_{N-1|i+1}\widehat{\mathbf{x}}_{N-1|i+1} + \mathbf{B}_{N-1|i+1}\widehat{u}_{N-1|i+1}$$
$$+ \mathbf{C}_{N-1|i+1}.$$

Thus we have:

$$\widehat{\mathbf{x}}_{k+1|i+1} = \mathbf{A}_{k|i+1}\widehat{\mathbf{x}}_{k|i+1} + \mathbf{B}_{k|i+1}\widehat{u}_{k|i+1} + \mathbf{C}_{k|i+1},$$

$k = 0, \dots, N-1$ as required in (43a).

Next, we observe that by Lemma 1, we must have that $\boldsymbol{\varphi}(t_{i+1}) \in \Omega_{\mathbf{x}_d|i}(t_{i+1}, \overline{w})$, such that $\boldsymbol{\varphi}(t_{i+1}) \in \mathbf{x}_d|i(t_{i+1}) \oplus \mathcal{E} = \mathbf{x}^*_{1|i} \oplus \mathcal{E}$. Noting that if $\mathbf{v} \in \mathcal{E}$, we must have that $-\mathbf{v} \in \mathcal{E}$, and thus $\mathbf{x}^*_{1|i} \in \boldsymbol{\varphi}(t_{i+1}) \oplus \mathcal{E}$. As $\mathbf{x}(t_{i+1}) = \boldsymbol{\varphi}(t_{i+1})$, we have $\widehat{\mathbf{x}}_{k|i+1} = \mathbf{x}^*_{1|i} \in \mathbf{x}(t_{i+1}) \oplus \mathcal{E}$ as required by (43b). As $\widehat{\mathbf{x}}_{N|i+1} = \mathbf{0}$, (43c) is satisfied.

By virtue of the previous solution, we can see that (43d)-(43g) are satisfied for $k = 0, \dots, N-2$ by our proposed solution. We need only show they hold for $k = N-1$. As $\widehat{\mathbf{x}}_{N-1|i+1} = \widehat{\mathbf{x}}_{N|i+1} = \mathbf{0}$ and $(\widehat{\boldsymbol{\xi}}_{N-1|i+1}) = \mathbf{0}$, we have that (43d) is satisfied. Furthermore, we have that $(\widehat{\boldsymbol{\xi}}_{N-1|i+1}) = \mathbf{0}$ implies the corresponding $(\widehat{\boldsymbol{\zeta}}_{N-1|i+1})_j = \mathbf{0}$ for $j = 0, \dots, 2n-1$. By assumption we have that $\mathbf{0} \in \mathcal{X} \ominus \mathcal{E}$, and thus we have that (43e) is satisfied.

Lastly, we note that $\overline{\mathbf{x}}_{N-1|i+1}$ used to define the linearization $\mathbf{Lin}_{N-1|i+1}$ is the origin, i.e., $\overline{\mathbf{x}}_{N-1|i+1} = \mathbf{0}$. Thus we have that the left-hand side of (43f) is $\mathbf{0}$, and thus because $\Gamma_{\alpha, \beta}(\mathbf{0}) \leq u_{\max}$, we see that $\widehat{\mathbf{s}}_{N-1|i+1} = \mathbf{0}$ satisfies (43f) and (43g), such that our proposed solution is feasible.

$\square$

Remark 4. Note that recomputing the linearizations about the previous trajectory is not strictly necessary to ensure feasibility – trajectories generated from *any* linearization of the system will be feasible for a full-state feedback linearizeable system. However, what is payed is performance – keeping the trajectory close to its linearization will reduce the conservativeness that the MPC program exhibits. As such, there is a conditional statement in Algorithm 1: if a feasible trajectory about the new linearizations can be found, use it, and if not, use the previous linearizations as a contingency plan to ensure feasibility.