

Closed-Loop Finite-Time Analysis of Suboptimal Online Control

Aren Karapetyan, Efe C. Balta, Andrea Iannelli, and John Lygeros

Abstract—Suboptimal methods in optimal control arise due to a limited computational budget, unknown system dynamics, or a short prediction window among other reasons. In this work, we study the transient closed-loop performance of such methods by providing finite-time suboptimality gap guarantees. We consider the control of discrete-time, nonlinear time-varying dynamical systems and establish sufficient conditions for such guarantees. These allow the control design to distribute a limited computational budget over a time horizon and estimate the on-the-go loss in performance due to suboptimality. We study exponential incremental input-to-state stabilizing policies and show that for nonlinear systems, under some mild conditions, this property is directly implied by exponential stability without further assumptions on global smoothness. The analysis is showcased on a suboptimal model predictive control use case.

Index Terms—Nonlinear Systems, Optimization Algorithms, Predictive Control

I. INTRODUCTION

Optimal control aims to compute an input signal to drive a dynamical system to a given target state, while optimizing a performance cost subject to constraints. In the absence of uncertainty, the problem has been studied using calculus of variations [1] and dynamic programming [2]. However, in many practical applications with limited computational power, it becomes difficult or infeasible to solve due to the curse of dimensionality [2]. This is further exacerbated if there are unknown system and/or cost parameters. As a result, control designers rely on approximate or suboptimal methods [3] to solve the problem. If there are adequate computational resources and an accurate simulator of the true system, the problem can be solved up to an arbitrary accuracy using approximate dynamic programming [4] or reinforcement learning [5] techniques. When this is not the case, e.g. the system has unpredictable dynamics or the cost to be optimized for is changing adversarially, offline methods alone are not

This work has been supported by the Swiss National Science Foundation under NCCR Automation (grant agreement 51NF40.225155) and by the German Research Foundation (DFG) under Germany's Excellence Strategy - EXC 2075 - 390740016.

A. Karapetyan, and J. Lygeros are with the Automatic Control Laboratory, Swiss Federal Institute of Technology (ETH Zürich), 8092 Zürich, Switzerland (E-mails: {akarapetyan, lygeros}@control.ee.ethz.ch).

E. C. Balta is with Inspire AG, 8005 Zürich, Switzerland (E-mail: efe.balta@inspire.ch).

A. Iannelli is with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Stuttgart 70569, Germany (E-mail: andrea.iannelli@ist.uni-stuttgart.de).

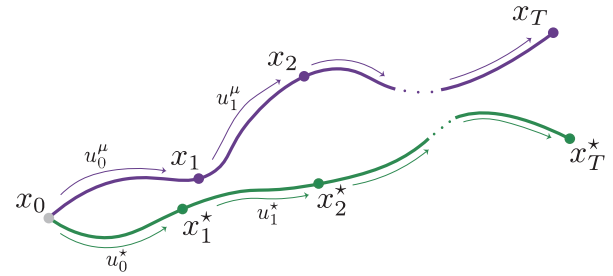


Fig. 1: Two separate closed-loop trajectories, generated by applying a suboptimal input signal u^μ , and a benchmark input signal u^* .

sufficient. In such cases, the policy is updated online or adaptively as more data becomes available. We consider a controller to be *online*, rather than *offline* or pre-fixed if it is capable of adapting to changes in the control problem during its single trajectory execution. For example, a model predictive controller (MPC) can utilize a finite window of predictions for possibly time-varying dynamics, disturbances, references, or costs to generate an input at each step. While such a definition partially overlaps with that of adaptive control [6], it defines a wider class of controllers, that also includes online optimization-based methods, like [7], [8].

The focus of this work is the closed-loop suboptimality analysis of online controllers in the finite-time or transient domain. Given their real-time implementation, online methods need to stay computationally efficient while stabilizing the system. Additionally, their performance is measured in terms of the accumulated cost that needs to be kept to a minimum. To quantify this, we fix a benchmark policy that we may deem to be close to the desired optimal one, visualized in Figure 1, and study the suboptimality gap of the given online algorithm in terms of the additional incurred cost due to its suboptimality with respect to the benchmark. Such an analysis provides a relative measure on the performance of the given algorithm, since, in general, the benchmark policy attains a non-zero cost. In this context, we pose the following questions.

- 1) How does the transient cost performance of an online algorithm scale with a measure of its suboptimality?
- 2) How should the benchmark policy be chosen to achieve meaningful finite-time bounds?

We consider nonlinear time-varying systems and derive conditions on the suboptimal and benchmark policies such that the suboptimality gap, defined as the difference of their

respective closed-loop costs, can be quantified. We show that it scales with the pathlength of the suboptimal trajectory and the rate of convergence of the suboptimal policy to the benchmark. The former can be computed online providing an on-the-go estimate of the suboptimality, and the latter usually depends on suboptimal algorithm parameters, allowing the control designer to tune these accordingly. We study the suboptimal, projected gradient method (PGM)-based linear-quadratic MPC [9] as an example of online control satisfying the proposed assumptions, and bound its suboptimality gap with respect to optimal MPC. Our contributions are summarized below:

- a) *We show that if the system dynamics in closed-loop with the benchmark policy are exponentially incrementally input-to-state stable (E - δ -ISS), and the suboptimal policy is linearly converging, then the suboptimality gap scales with the pathlength of the closed-loop suboptimal trajectory, and the convergence rate.*
- b) *We derive sufficient conditions under which exponential stability (ES) of a non-smooth nonlinear time-varying system implies E - δ -ISS, making the condition on the benchmark policy easier to verify.*
- c) *We study the suboptimal, PGM-based linear-quadratic MPC problem as an example satisfying these assumptions, and bound its suboptimality gap.*

The E - δ -ISS assumption on the benchmark policy is crucial for the validity of the results. Incremental input-to-state-stability is studied in [10], in the continuous-time and in [11], in the discrete-time settings, providing a condition on the deviation of two separate trajectories of the same system. If the dynamics are smooth, a sufficient condition for E - δ -ISS to hold is that of contraction [12]–[15]. However, this is often not the case when one considers closed-loop dynamics under optimal controllers, e.g. under a constrained MPC [16] policy. Since we are often interested in such benchmarks, we devote Section IV to the derivation of sufficient conditions for E - δ -ISS to hold when the dynamics are non-smooth in general.

Such an incremental stability analysis allows for the derivation of asymptotically tight bounds in the sense that they scale with the *level of suboptimality*, or the rate of convergence of the suboptimal policy to the optimal, converging to zero when the algorithm matches with the benchmark. The bounds also scale with the pathlength of the suboptimal trajectory, allowing an on-the-go calculation of the suboptimality gap that is independent of the benchmark states. Hence, for the suboptimal MPC example we achieve asymptotically tighter bounds under the same assumptions compared to [17]. Moreover, our result is independent of the asymptotic properties of the suboptimal algorithm, providing finite-time performance bounds even when the closed-loop is not exponentially stable.

Several notable examples of settings where such finite-time suboptimality analysis can be of use include adaptive control [18]–[20], with suboptimality due to unknown system parameters, online feedback optimization [21], [22] and online control [7], [8], with suboptimality due to unknown future costs, or real-time optimization-based control, such as real-time MPC [23]–[25], suboptimal due to finite computational resources.

While our analysis and results hold for any algorithm satisfying the outlined assumptions, optimization-based methods, including MPC, and their suboptimal variants are of particular interest given their wide applicability in practice. Suboptimal MPC has been studied extensively in the literature [23]–[30]. Real-time MPC with iterative optimization methods has been studied in [23]–[28]. Efficient warm-start methods are explored in [27], [28] with [27] also showing a lower bound on the number of gradient descent steps needed to achieve an open-loop suboptimality level, but no closed-loop analysis is performed. The closed-loop stability is analyzed in [26] using a robust MPC formulation, and also in later works in [24], [25] that extend the real-time MPC work of [23] and show asymptotic stability using Lyapunov methods and the small gain theorem of interconnected systems. Other sources of suboptimality for MPC have been studied in [29], analyzing the effect of the prediction horizon on the suboptimality of the MPC closed-loop cost, or in [30] in the context of distributed control, to name a few. However, in this line of work the suboptimality gap as we define here is not considered. This is a common performance metric in online learning, where it is referred to as *regret*. While there are several works [7], [31], [32] studying the regret of various model predictive controllers, their settings are different from ours. For instance, [31] considers Bayesian MPC, [7] studies the effect of predictions for linear time-invariant systems, and [32] studies optimal controllers with inexact predictions using perturbation analysis, thus leading to settings that are not considered under the assumptions in this paper. The earlier work [17] also analyzes the suboptimality of PGM-based MPC but does not utilize nonlinear incremental stability analysis achieving a weaker bound that is not asymptotically tight.

The article is structured as follows. In Section II we provide the preliminaries and the problem setup. In Section III, we conduct the suboptimality gap analysis. Sufficient conditions for E - δ -ISS are derived in Section IV, and in Section V, the suboptimal PGM-based linear MPC use case is studied with a numerical example.

Notation: The sets of positive real numbers, positive integers, and non-negative integers are denoted by \mathbb{R}_+ , \mathbb{N}_+ and \mathbb{N} , respectively. For some $k_0 \in \mathbb{N}$, the set of integers greater than or equal to k_0 is denoted by $\mathbb{N}_{\geq k_0}$. For a given vector x , its Euclidean norm is denoted by $\|x\|$, and the two-norm weighted by a $Q \succ 0$ by $\|x\|_Q = \sqrt{x^\top Q x}$. For a square matrix W , the spectral radius and the spectral norm are denoted by $\rho(W)$, and $\|W\|$, respectively. Given a symmetric, positive definite matrix $M \succ 0$, we define its unique square-root by $M^{\frac{1}{2}}$, such that $M = M^{\frac{1}{2}} M^{\frac{1}{2}}$ and $M^{\frac{1}{2}} \succ 0$. For a matrix W and $M \succ 0$, $\lambda_M^-(W)$ and $\lambda_M^+(W)$ denote the minimum and maximum eigenvalues of $M^{-\frac{1}{2}} W M^{-\frac{1}{2}}$; for any vector x , they satisfy $\lambda_M^-(W) \|x\|_M^2 \leq \|x\|_W^2 \leq \lambda_M^+(W) \|x\|_M^2$. The Euclidean point-to-set distance of a vector x from a nonempty, closed, convex set \mathcal{A} is denoted by $|x|_{\mathcal{A}} := \min_{y \in \mathcal{A}} \|x - y\|$, and the projection onto it by $\Pi_{\mathcal{A}}[x] = \arg \min_{y \in \mathcal{A}} \|x - y\|$. The n -dimensional closed ball of radius $r > 0$, centered at the origin is denoted by $\mathcal{B}(r) := \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$.

II. PRELIMINARIES AND PROBLEM SETUP

We consider the optimal control problem for discrete-time, nonlinear time-varying systems of the form

$$x_{k+1} = f_0(k, x_k) + g(k, x_k, u_k), \quad k \in \mathbb{N}_{\geq k_0}, \quad (1)$$

where $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^m$ denote the state and control input at time k , respectively, $f_0 : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ denotes the unforced nominal dynamics and $g : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ the controlled dynamics. The system evolution starts from some initial state $x_{k_0} \in \mathbb{R}^n$ at some initial time $k_0 \in \mathbb{N}$. The optimal control objective is to find the sequence of control inputs $\mathbf{u} = [u_{k_0}^\top \dots u_{T-1}^\top]^\top$ that minimizes the finite-time cost

$$J_T(x_{k_0}, \mathbf{u}) = F(T, x_T) + \sum_{k=k_0}^{T-1} c(k, x_k, u_k), \quad (2)$$

where $c : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is the stage cost at time k , $F : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathbb{R}$ is the terminal cost, and $T \in \mathbb{N}_{\geq k_0+1}$ is the control horizon. In addition, the control input has to satisfy $u_k \in \mathcal{U}$ for all k , for some bounded $\mathcal{U} \subset \mathbb{R}^m$.

An admissible policy $\pi(k, x) : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathcal{U}$ maps the state at time k to a control input, generating the control signal $\mathbf{u}^\pi = [u_{k_0}^{\pi^\top} \dots u_{T-1}^{\pi^\top}]^\top$ and the associated trajectory $\mathbf{x}^\pi = [x_{k_0}^{\pi^\top} \dots x_{T-1}^{\pi^\top}]^\top$. With a slight abuse of notation, its associated cost is denoted by $J_T(x_{k_0}, \pi)$. We consider time-varying systems for generality, and note that the analysis holds directly for time-invariant systems and policies as a special case. We showcase this in Section V.

We are interested in the relation of a policy μ , corresponding to a given suboptimal algorithm, with respect to another benchmark policy μ^* that is equipped with desirable characteristics, e.g. optimality. The two policies are defined as follows.

Benchmark dynamics: Consider a benchmark policy $\mu^* : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathcal{U}$. Given an initial state $x_{k_0}^* \in \mathbb{R}^n$, the benchmark dynamics are given by¹

$$x_{k+1}^* = f_0(k, x_k^*) + g(k, x_k^*, \mu_k^*(x_k^*)) := f(k, x_k^*), \quad (3)$$

for all $k \in \mathbb{N}_{\geq k_0}$. We assume that there exists a set $\mathcal{D}^* \subseteq \mathbb{R}^n$ that is forward invariant under the closed-loop dynamics (3), and restrict attention to $x_{k_0}^* \in \mathcal{D}^*$. Hence, $x_k^* \in \mathcal{D}^*$ for all $k \in \mathbb{N}_{\geq k_0}$.

Suboptimal dynamics: The suboptimal state evolution for a given policy $\mu : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathcal{U}$ can be represented in the following form² for any $x_{k_0} \in \mathbb{R}^n$

$$x_{k+1} = f(k, x_k) + \underbrace{g(k, x_k, \mu_k(x_k)) - g(k, x_k, \mu_k^*(x_k))}_{:=w_k(x(k))}, \quad (4)$$

for all $k \in \mathbb{N}_{\geq k_0}$. The mapping $w : \mathbb{N}_{\geq k_0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ can be thought of as a state-dependent disturbance acting on the benchmark state dynamics (3), introduced due to suboptimality. It is assumed to be such that the closed-loop suboptimal dynamics (4) evolve within a set $\mathcal{D}^\mu \subseteq \mathcal{D}^*$ that is forward invariant under (4). Restricting attention to initial states $x_{k_0} \in \mathcal{D}^\mu$, it then holds that $x_k \in \mathcal{D}^\mu$ for all $k \in \mathbb{N}_{\geq k_0}$.

¹For readability, we place the time k in the subscripts of μ , c , and F .

²We drop the explicit reference to μ from the superscript of x for readability.

Figure 1 shows the pictorial evolution of the two considered trajectories starting from the same initial state x_0 at time $k_0 = 0$. For each x_k^* , u_k^* denotes the control input generated by $\mu_k^*(x_k^*)$ and for each x_k , $u_k^\mu := \mu_k(x_k)$ the input generated by the suboptimal policy.

To quantify the relation between μ and μ^* , we define the suboptimality gap of the policy μ as the additional incurred cost compared to the benchmark

$$\mathcal{R}_T^\mu(x_{k_0}) := J_T(x_{k_0}, \mu) - J_T(x_{k_0}, \mu^*), \quad (5)$$

given some $x_{k_0} \in \mathcal{D}^\mu$. An informative finite-time bound on it, which depends on fundamental quantities of the online control problem, can provide a quantifiable tradeoff between the effort needed to compute the suboptimal policy μ and the additional cost incurred by using it instead of μ^* .

We assume the benchmark policy μ^* to have good performance, since otherwise, \mathcal{R}_T^μ can be uninformative. We characterize this performance in terms of E - δ -ISS.

Definition 1. A nonlinear dynamical system $x_{k+1} = f(k, x_k)$ is E - δ -ISS in some forward invariant $\mathcal{D} \subseteq \mathbb{R}^n$, if there exist constants $c_0, c_w, r_w \in \mathbb{R}_+$ and $\rho \in (0, 1)$, such that for any $(x_{k_0}, y_{k_0}) \in \mathcal{D} \times \mathcal{D}$, and $w_k \in \mathcal{B}(r_w), k \in \mathbb{N}_{\geq k_0}$, the perturbed dynamics $y_{k+1} = f(k, y_k) + w_k$ satisfy

$$\|x_k - y_k\| \leq c_0 \rho^{k-k_0} \|x_{k_0} - y_{k_0}\| + c_w \sum_{i=k_0}^{k-1} \rho^{k-i-1} \|w_i\|,$$

for all $k \in \mathbb{N}_{\geq k_0}$, where the disturbances w_k are such that $y_k \in \mathcal{D}$ for all $k \in \mathbb{N}_{\geq k_0}$. If $\mathcal{D} = \mathbb{R}^n$ the system is called globally E - δ -ISS.

E - δ -ISS for continuous-time systems has been introduced in [10]. For an in-depth discussion and analysis of incremental stability in discrete-time, and its relation to contraction [15] and convergent dynamics [33], we refer the interested reader to [11], and the references therein. We assume the following for the benchmark policy.

Assumption 1. (Benchmark Policy). Given the closed-loop system (3), the benchmark policy μ^* is such that

- i. it is uniformly L -Lipschitz continuous in \mathcal{D}^* , i.e. there exists a constant $L \in \mathbb{R}_+$, such that for all $(x, y) \in \mathcal{D}^* \times \mathcal{D}^*$, and $k \in \mathbb{N}_{\geq k_0}$

$$\|\mu_k^*(x) - \mu_k^*(y)\| \leq L \|x - y\|,$$

- ii. there exist $a_k \geq 0, k \in \mathbb{N}_{\geq k_0}$, such that for all $x \in \mathcal{D}^*$ and $k \in \mathbb{N}_{\geq k_0}$

$$\|\mu_{k+1}^*(x) - \mu_k^*(x)\| \leq a_k,$$

- iii. the closed-loop dynamics (3) are E - δ -ISS in \mathcal{D}^* with a rate $\rho \in (0, 1)$.

Although the Lipschitz continuity condition excludes benchmark policies with abrupt changes, such as discontinuities or jumps in the policy, it still holds for a relevant class of policies. In particular, when μ^* is defined as a solution to an optimization problem, as is, for example, the MPC policy, then Lipschitz continuity holds [34] if the optimal solution mapping is strongly regular [35]. This can be verified for a range

of problems by checking for strong second order sufficient conditions (SOSC), which is well studied in the literature, e.g. [25], [36]. Additionally, when state constraints are present, these can be relaxed by introducing soft constraints [37], [38], making the SOSC conditions easier to verify. Furthermore, we stress that the benchmark policy is not necessarily the optimal policy but any comparator policy as long as the conditions in Assumption 1 are satisfied. The second assumption limits how fast the benchmark policy changes given the same state between two timesteps. Since \mathcal{U} is bounded, such an a_k always exists, for all k , and can be set equal to the diameter of \mathcal{U} . However, it can also encode additional information, such as stationarity of the benchmark policy, in which case $a_k = 0$ for all k . The E - δ -ISS condition, on the other hand, is often harder to verify, unless one assumes the existence of an incremental Lyapunov equation [10], [11], or uses contraction to infer incremental stability [13], which in turn assumes a smooth policy. Since a large class of benchmark policies, including MPC, are non-smooth globally, we show in Section IV, that under further mild conditions on f , ES is enough to guarantee E - δ -ISS.

Next, we impose a linear convergence condition on the suboptimal policy μ .

Assumption 2. (*Suboptimal Policy*). Let $u_k^\mu = \mu_k(x_k)$ denote the suboptimal input evaluated on the suboptimal trajectory (4). There exist $\eta_k \in [0, 1]$, $k \in \mathbb{N}_{\geq k_0}$ and some $u_{k_0-1}^\mu = \nu \in \mathcal{U}$, such that for all $x_{k_0} \in \mathcal{D}^\mu$, $k \in \mathbb{N}_{\geq k_0}$

$$\|u_k^\mu - \mu_k^*(x_k)\| \leq \eta_k \|u_{k-1}^\mu - \mu_k^*(x_k)\|. \quad (6)$$

The assumption imposes at least a linear rate of convergence on the suboptimal input with respect to the benchmark given the suboptimal state. In some cases, the rate η_k can be thought of as a design parameter that can be tuned to control the desired level of suboptimality depending on the available computational budget. Notably, such rates appear in iterative optimization algorithms, such as PGM [39] or alternating direction method of multipliers (ADMM) [40], where the benchmark input is optimal with respect to some objective function, and the suboptimal input is obtained by the corresponding optimization method. When such methods are applied to optimization-based control, warm-started with the previous input, the exact linear convergence form of (6) appears in [9], [24], [25] for PGM and in [41] for ADMM. In Section V, we consider the PGM-based MPC setting of [9] and show how both assumptions are verified.

We restrict our attention to systems where the controlled dynamics g are Lipschitz continuous with respect to u , uniformly in x and k .

Assumption 3. There exists a constant $L_u \in \mathbb{R}_+$, such that for any $(u, v) \in \mathbb{R}^m \times \mathbb{R}^m$, for all $x \in \mathcal{D}^\mu$ and $k \in \mathbb{N}_{\geq k_0}$

$$\|g(k, x, u) - g(k, x, v)\| \leq L_u \|u - v\|.$$

This is satisfied, for instance, in linear time-invariant systems or control-affine nonlinear systems of the form $x_{k+1} = f(x_k) + g(x_k)u_k$, often studied in the context of feedback linearization (see [42], [43] for details). Finally, we restrict our analysis to local Lipschitz continuous stage costs.

Assumption 4. There exist constants $M_x, M_u \in \mathbb{R}_+$, such that for all $(x, y) \in \mathcal{D}^\mu \times \mathcal{D}^\mu$, $(u, z) \in \mathcal{U} \times \mathcal{U}$ and $k \in \mathbb{N}_{\geq k_0}$

$$\begin{aligned} \|c_k(x, u) - c_k(y, z)\| &\leq M_x \|x - y\| + M_u \|u - z\|, \\ \|F_k(x) - F_k(y)\| &\leq M_x \|x - y\|. \end{aligned}$$

III. SUBOPTIMALITY GAP ANALYSIS

In this section, we analyze the suboptimality gap for a given policy and show that \mathcal{R}_T scales with the product of the pathlength of the suboptimal dynamics and a vector dependent on the convergence rates. In the analysis we take $k_0 = 0$ without loss of generality. We define the backward difference path vector, $\Delta \in \mathbb{R}^{T-1}$, to be

$$\Delta := [\|\delta x_1\| \quad \|\delta x_2\| \quad \dots \quad \|\delta x_{T-1}\|]^\top,$$

where $\delta x_k = x_k - x_{k-1}$, $k \in \mathbb{N}_+$, where x_k is the state at time k for the suboptimal dynamics (4). The pathlength of the suboptimal trajectory is then defined as $\mathcal{S}_T = \|\Delta\|_1$ and the Euclidean pathlength as $\mathcal{S}_{T,2} := \|\Delta\|$.

The policy convergence rate vector, $\tilde{\eta} \in \mathbb{R}^{T-1}$ is defined as

$$\tilde{\eta} := [\tilde{\eta}_1 \quad \tilde{\eta}_2 \quad \dots \quad \tilde{\eta}_{T-1}]^\top,$$

where

$$\tilde{\eta}_k := \sum_{i=k}^{T-1} \prod_{j=k}^i \eta_j, \quad \forall k \in [0, T-1].$$

Note that $\tilde{\eta}_k = \mathcal{O}(\eta_k)$ and provides a weighting on the influence of the δx_k on \mathcal{R}_T . This is analyzed further in Section III-B. We denote the Euclidean norm of the suboptimality vector by $\tilde{\eta} := \|\tilde{\eta}\|$. The rate of change of the benchmark input $\mu^*(x^*)$ is captured by the vector $a \in \mathbb{R}_+^{T-1}$, defined as

$$a := [a_1 \quad a_2 \quad \dots \quad a_{T-1}]^\top.$$

A. Upper Bound

The bound in the following theorem captures the tradeoff between suboptimality and the additional cost in closed-loop.

Theorem 1. Let Assumptions 1, 3 and 4 hold, then the suboptimality gap of any policy, μ , fulfilling Assumption 2 satisfies

$$\mathcal{R}_T^\mu(x_0) = \mathcal{O}\left((a + \Delta)^\top \tilde{\eta}\right),$$

for all $x_0 \in \mathcal{D}^\mu$. Specifically, it is bounded by

$$\mathcal{R}_T^\mu(x_0) \leq \bar{M} \left(\tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta} \right),$$

for all $x_0 \in \mathcal{D}^\mu$, where $\delta u_0 := \nu - \mu_0^*(x_0)$ and $\bar{M} := \left(M_u + \frac{c_w L_u (M_u L + M_x)}{1 - \rho} \right)$.

The asymptotic variable in the $\mathcal{O}(\cdot)$ notation is the time horizon length T , with all terms in the product $(a + \Delta)^\top \tilde{\eta}$ scaling with T , as $T \rightarrow \infty$. The bound tends to zero as $\tilde{\eta}$ decreases. This is intuitive, as smaller $\tilde{\eta}$ suggests that the benchmark and suboptimal trajectories are closer to each other. Additionally, the suboptimality gap is a relative measure, but the bound is fully decoupled from the performance of μ^* and only depends on the performance of the suboptimal state evolution if $a = \mathbf{0}$. In this case, the bound can also be given

as $\mathcal{R}_T(x_0) = \mathcal{O}(\tilde{\eta}\mathcal{S}_{T,2})$, where $\mathcal{S}_{T,2}$ captures the transient behavior of the suboptimal system and is well-defined in the limit as $T \rightarrow \infty$ when, for example, (4) is exponentially stable.

Before we prove Theorem 1, we introduce several supporting lemmas, and the Cauchy Product inequality defined for two finite series $\{a_i\}_{i=1}^T$ and $\{b_i\}_{i=1}^T$ as follows

$$\sum_{i=0}^T \left| \sum_{j=0}^i a_j b_{i-j} \right| \leq \left(\sum_{i=0}^T |a_i| \right) \left(\sum_{j=0}^T |b_j| \right). \quad (7)$$

We let $u_k^\mu = \mu_k(x_k)$ denote the suboptimal input as in Assumption 2.

Lemma 1. *Let Assumption 1 hold, then for any policy, μ , fulfilling Assumption 2, and for all $x_0 \in \mathcal{D}^\mu$*

$$\sum_{k=0}^{T-1} \|d_k\| \leq \tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta}, \quad (8)$$

where we define $d_k := u_k^\mu - \mu_k^*(x_k)$, and $\delta u_0 := \nu - \mu_0^*(x_0)$.

Proof. Consider the bound on the per-step error in the input

$$\begin{aligned} \|d_k\| &\stackrel{(a)}{\leq} \eta_k \|u_{k-1}^\mu - \mu_{k-1}^*(x_k)\| \\ &\stackrel{(b)}{\leq} \eta_k \|u_{k-1}^\mu - \mu_{k-1}^*(x_{k-1})\| \\ &\quad + \eta_k \|\mu_{k-1}^*(x_{k-1}) - \mu_{k-1}^*(x_k)\| \\ &\stackrel{(c)}{\leq} \eta_k \|d_{k-1}\| + \eta_k \|\mu_{k-1}^*(x_{k-1}) - \mu_{k-1}^*(x_k)\| \\ &\quad + \eta_k \|\mu_{k-1}^*(x_k) - \mu_k^*(x_k)\| \\ &\stackrel{(d)}{\leq} \eta_k \|d_{k-1}\| + \eta_k L \|x_k - x_{k-1}\| + \eta_k a_k, \end{aligned}$$

the inequality (a) follows directly from Assumption 2, (b) and (c) follow from the triangle inequality for vector norms and (d) from the uniform Lipschitz condition in Assumption 1.i. and Assumption 1.ii.. Applying the above inequality recursively leads to

$$\|d_k\| \leq \|d_0\| \prod_{i=1}^k \eta_i + \sum_{j=1}^k (a_j + L\|d_{j-1}\|) \prod_{i=j}^k \eta_i,$$

for all $k \in \mathbb{N}_+$. Summing up for $k = 0, \dots, T-1$

$$\begin{aligned} \sum_{k=0}^{T-1} \|d_k\| &\leq \|\delta u_0\| \sum_{k=0}^{T-1} \prod_{i=0}^k \eta_i + \sum_{k=1}^{T-1} \sum_{j=1}^k (a_j + L\|d_{j-1}\|) \prod_{i=j}^k \eta_i \\ &= \tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta}, \end{aligned}$$

where the inequality follows from Assumption 2 and the equality from the definitions of $\tilde{\eta}$, a and Δ . \square

The following lemma provides an upper bound on the norm of the finite-time trajectory mismatch.

Lemma 2. *Let Assumptions 1 and 3 hold, and consider any policy μ fulfilling Assumption 2. Then, there exist constants $c_0, c_w \in \mathbb{R}_+$, such that for all $(x_0, x_0^*) \in (\mathcal{D}^\mu \times \mathcal{D}^\mu)$*

$$\begin{aligned} \sum_{k=0}^T \|x_k - x_k^*\| &\leq \|x_0 - x_0^*\| \left(\frac{c_0(1 - \rho^{T+1})}{1 - \rho} \right) \\ &\quad + \frac{c_w L u (1 - \rho^T)}{1 - \rho} \left(\tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta} \right), \end{aligned}$$

where x_k and x_k^* denote the states at time k under, respectively, the suboptimal and benchmark policies.

Proof. Given the boundedness of \mathcal{U} , the uniform Lipschitz continuity of g in u , and recalling the definition of w_k from (4), it follows that there exists a $r_w \in \mathbb{R}_+$, such that $w_k(x_k) \in \mathcal{B}(r_w), k \in \mathbb{N}$. Consider then the dynamics (4) as a perturbed version of the benchmark dynamics (3). Assumption 1 ensures that there exist $c_0, c_w \in \mathbb{R}_+$ and $\rho \in (0, 1)$, such that for all $k \in \mathbb{N}$ and $(x_0, x_0^*) \in \mathcal{D}^\mu \times \mathcal{D}^\mu$

$$\begin{aligned} \|x_k - x_k^*\| &\leq c_0 \rho^k \|x_0 - x_0^*\| + c_w \sum_{i=0}^{k-1} \rho^{k-i-1} \|w_i(x_i)\| \\ &\leq c_0 \rho^k \|x_0 - x_0^*\| + c_w L u \sum_{i=0}^{k-1} \rho^{k-i-1} \|d_i\|, \end{aligned}$$

where the second inequality follows from Assumption 3, where $d_k := u_k^\mu - \mu_k^*(x_k)$. Summing up over the whole trajectory and noting the resultant finite geometric series

$$\begin{aligned} \sum_{k=0}^T \|x_k - x_k^*\| &\leq \|x_0 - x_0^*\| \left(\frac{c_0(1 - \rho^{T+1})}{1 - \rho} \right) + c_w L u \sum_{k=0}^{T-1} \sum_{i=0}^k \rho^{k-i} \|d_i\| \\ &\leq \|x_0 - x_0^*\| \left(\frac{c_0(1 - \rho^{T+1})}{1 - \rho} \right) + c_w L u \sum_{k=0}^{T-1} \rho^k \sum_{i=0}^{T-1} \|d_i\|, \end{aligned}$$

where the last inequality follows from (7). Finally, the result follows by using the bound (8) in Lemma 1. \square

Similarly, the norm of the finite-time input signal mismatch due to the difference in applied inputs can be bounded in the following Lemma.

Lemma 3. *Let Assumptions 1 and 3 hold, and consider any policy μ fulfilling Assumption 2. Then, there exist constants $c_0, c_w \in \mathbb{R}_+$, such that for all $(x_0, x_0^*) \in (\mathcal{D}^\mu \times \mathcal{D}^\mu)$*

$$\begin{aligned} \sum_{k=0}^{T-1} \|u_k^\mu - \mu_k^*(x_k^*)\| &\leq \|x_0 - x_0^*\| \frac{L c_0 (1 - \rho^T)}{1 - \rho} \\ &\quad + \left(\tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta} \right) \left(1 + \frac{L c_w L u (1 - \rho^{T-1})}{1 - \rho} \right), \end{aligned}$$

where x_k and x_k^* are the states at time k under, respectively, the suboptimal and benchmark policies, μ and μ^* .

Proof. Using the triangle inequality for vector norms and defining $d_k := u_k^\mu - \mu_k^*(x_k)$

$$\begin{aligned} \|u_k^\mu - \mu_k^*(x_k^*)\| &\leq \|d_k\| + \|\mu_k^*(x_k) - \mu_k^*(x_k^*)\| \\ &\leq \|d_k\| + L \|x_k - x_k^*\|, \end{aligned}$$

where the last inequality follows from Lipschitz continuity of μ^* from Assumption 1.i.. Summing up over the trajectory horizon, using the bound (8) from Lemma 1 and the bound in Lemma 2 completes the proof. \square

Proof of Theorem 1. By Assumption 4, for all $(x_0, x_0^*) \in \mathcal{D}^\mu \times \mathcal{D}^\mu$

$$\begin{aligned} \mathcal{R}_T(x_0, x_0^*) &:= J_T(x_0, \mu) - J_T(x_0^*, \mu^*) \\ &\leq M_x \sum_{k=0}^T \|x_k - x_k^*\| + M_u \sum_{k=0}^{T-1} \|u_k^\mu - \mu_k^*(x_k^*)\|. \end{aligned}$$

Then, using Lemmas 2 and 3 for the two respective sums

$$\begin{aligned} \mathcal{R}_T(x_0, x_0^*) &< \|x_0 - x_0^*\| \frac{c_0 (M_u L + M_x)}{1 - \rho} \\ &\quad + \bar{M} \left(\tilde{\eta}_0 \|\delta u_0\| + (a + L\Delta)^\top \tilde{\eta} \right). \end{aligned}$$

The result follows by taking $x_0 = x_0^*$. \square

For the special case when the convergence rate of the suboptimal policy is constant, the following Corollary follows.

Corollary 1. *Let Assumptions 1, 3 and 4 hold, and consider a suboptimal policy, μ , fulfilling Assumption 2 with $\eta_k = \eta$, $k \in \mathbb{N}$. Then, its suboptimality gap is given by*

$$\mathcal{R}_T^\mu(x_0) = \mathcal{O} \left(\frac{\eta}{1 - \eta} (\mathcal{S}_T + \|a\|_1) \right),$$

for all $x_0 \in \mathcal{D}^\mu$.

Proof. If $\eta_k = \eta$, $k \in \mathbb{N}$, $\tilde{\eta}_k = \eta (1 - \eta^{T-k}) / (1 - \eta)$, and is bounded by

$$\tilde{\eta}_k \leq \frac{\eta}{1 - \eta}, \quad k \in [0, T].$$

The complexity term then satisfies

$$(a + \Delta)^\top \tilde{\eta} \leq \frac{\eta (\mathcal{S}_T + \|a\|_1)}{1 - \eta}.$$

The rest of the proof follows directly from Theorem 1 by replacing the complexity term with the new bound. \square

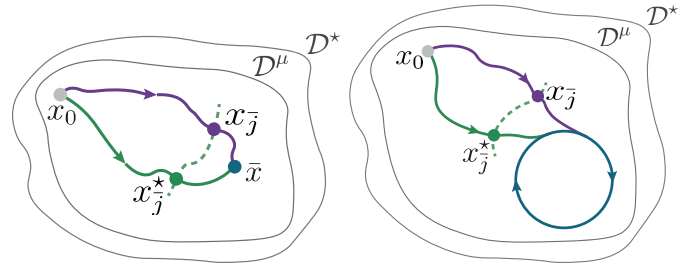
For bounded \mathcal{S}_T , the corollary shows the effect of η on the suboptimality gap. Crucially, the gap is 0 in the absence of suboptimality. On the other hand, given a fixed η , one can calculate an upper bound on the suboptimality gap of the policy μ by only considering its pathlength \mathcal{S}_T . The independence from the pathlength of the benchmark policy is particularly useful for online suboptimality estimation since it cannot be generally accessed by the control designer.

B. Interpretation of the Upper Bound

The term $\tilde{\eta}_0 \|\delta u_0\|$ in the bound of Theorem 1 captures the error due to the initial mismatch in the control input, δu_0 . This term in general cannot be avoided, unless the initial ‘‘guess’’ of the input ν is correct, or $\eta_0 = 0$, so that the suboptimal and benchmark policies match at the initial timestep.

The second term, $a^\top \tilde{\eta}$, scales with the magnitude of the rate of change of the time-varying benchmark policy μ^* , as defined in Assumption 1.ii.. It vanishes either when μ^* is stationary, or when the benchmark and suboptimal policies coincide.

The main complexity term of interest, $\Delta^\top \tilde{\eta}$ captures the suboptimality of the policy through the inner product of the path vector Δ and the suboptimality vector $\tilde{\eta}$. To study the interplay of these two quantities in more detail, let us



(a) Exponentially stable case. (b) Case of limit cycles. Suboptimal trajectories are captured by (9). Benchmark trajectories are captured by (10).

Fig. 2: The pictorial evolution of suboptimal and benchmark trajectories evolving in $\mathcal{D}^\mu \subseteq \mathcal{D}^*$.

consider the case when the benchmark dynamics (3) have an equilibrium at some $\bar{x} \in \mathbb{R}^n$, and $k_0 = 0$ without loss of generality. If the suboptimal closed-loop system (4) is also exponentially stable with $\eta_k \neq 0$, $\forall k \in \mathbb{N}$, then $\|\delta x_j\| \approx 0$, $\forall j \geq \bar{j}$, for some $\bar{j} \in \mathbb{N}$, as visualised in Figure 2a. Note that, in order to provide this interpretation it is assumed that the systems (3) and (4) have the same single equilibrium point \bar{x} . In such a setting, the complexity term captures the fact that the suboptimality gap is finite as follows

$$\Delta^\top \tilde{\eta} = \left[\|\delta x_1\| \dots \underbrace{\|\delta x_{\bar{j}}\| \dots \|\delta x_{T-1}\|}_{\approx 0} \right] \begin{Bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_{\bar{j}} \\ \vdots \\ \tilde{\eta}_{T-1} \end{Bmatrix} \neq 0 \quad (9)$$

This example coincides with the suboptimal MPC use case discussed in detail in Section V. When there are multiple equilibria, a similar interpretation of the bound holds. Among other possibilities, one can also consider the case when the benchmark dynamics (3) converge to a limit cycle. Since (3) is E - δ -ISS it follows from (4) that if at a given point in time $\bar{j} \in \mathbb{N}$, the suboptimal policy coincides with the considered benchmark, i.e. $\eta_j = 0$, $\forall j \geq \bar{j}$, then the trajectories will necessarily coincide, as visualized in Figure 2b. This is captured by the complexity term as

$$\Delta^\top \tilde{\eta} = \left[\|\delta x_1\| \dots \underbrace{\|\delta x_{\bar{j}}\| \dots \|\delta x_{T-1}\|}_{\neq 0} \right] \begin{Bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_{\bar{j}} \\ \vdots \\ \tilde{\eta}_{T-1} \end{Bmatrix} = 0 \quad (10)$$

Even though the pathlength keeps increasing, the norm of the suboptimality vector is finite, resulting in a finite suboptimality gap, containing only the additional cost due to suboptimality at the first \bar{j} timesteps.

IV. EXPONENTIALLY STABLE POLICIES AND E - δ -ISS

The analysis in the previous section is agnostic to the steady-state properties of (3) and, therefore, has general validity. In this section, we analyze a common case when the dynamics (3) have a unique equilibrium point at the origin. We derive sufficient conditions under which the exponential stability of globally non-smooth nonlinear dynamics implies E - δ -ISS, thus verifying the otherwise non-trivial Assumption 1.iii..

We treat (3) as a general nonlinear time-varying system of the form

$$x_{k+1} = f(k, x_k), \quad k \in \mathbb{N}_{\geq k_0}, \quad (11)$$

where $f : \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathcal{D}$ is continuous with respect to both arguments, $x_{k_0} = \xi \in \mathbb{R}^n$ for some $k_0 \in \mathbb{N}$, $\mathcal{D} \subset \mathbb{R}^n$. Moreover, the set \mathcal{D} is assumed to be compact and to contain the origin. The solution of the system (11) at time $k \in \mathbb{N}_{\geq k_0}$ is characterized by the function $\phi : \mathbb{N}_{\geq k_0} \times \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathcal{D}$ mapping the current time, initial time and the initial state to the current state, i.e. $\phi(k+1, k_0, \xi) = f(k, \phi(k, k_0, \xi))$ for all $k \in \mathbb{N}_{\geq k_0}$. We consider the origin to be an equilibrium point for (11), i.e. $f(k, \mathbf{0}) = \mathbf{0}$ for all $k \in \mathbb{N}_{\geq k_0}$. Although this restricts the attention to regulation problems, one can always recast a tracking problem into a regulation one by considering the nonlinear evolution of the error between the state and the reference. We impose the following assumption.

Assumption 5. (*Local Behavior*). *The dynamics f in (11) are*

- i. L_f -Lipschitz continuous in $\mathcal{D} \subset \mathbb{R}^n$,
- ii. continuously differentiable with respect to x , and the Jacobian matrix $[\partial f(k, x)/\partial x]$ is bounded and Lipschitz continuous uniformly in k within some region $\mathcal{D}_0 \subseteq \mathcal{D}$ containing the origin.

For completeness, we present a series of auxiliary definitions and theorems for (incremental) exponential stability before presenting the main results of this section.

A. Preliminaries on Exponential Stability

We define uniform exponential stability for discrete-time, nonlinear time-varying systems [43].

Definition 2. *Given the system (11), the equilibrium point $x = \mathbf{0}$ is uniformly exponentially stable in \mathcal{D} with a rate λ , if there exist constants $d \in \mathbb{R}_+$ and $\lambda \in (0, 1)$, such that for all $\xi \in \mathcal{D}$, and $k \in \mathbb{N}_{\geq k_0}$*

$$\|\phi(k, k_0, \xi)\| \leq d \|\xi\| \lambda^{k-k_0}. \quad (12)$$

If $\mathcal{D} = \mathbb{R}^n$, then the equilibrium is uniformly globally exponentially stable.

If the origin is exponentially stable, we also refer to the system (11) as such. Lyapunov theory provides necessary and sufficient conditions for the exponential stability of nonlinear systems. Below are the discrete-time Lyapunov theorems for exponential stability.

Theorem 2. [43, Thm. 13.11] *If there exists a continuous mapping $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathbb{R}_+$, and some constants $c_1, c_2 \in$*

\mathbb{R}_+ , $\beta \in (0, 1)$ and $p \geq 1$, such that for all $\xi \in \mathcal{D}$, and $k \in \mathbb{N}_{\geq k_0}$

$$\begin{aligned} c_1 \|\xi\|^p &\leq V(k, \xi) \leq c_2 \|\xi\|^p, \\ V(k+1, f(k, \xi)) &\leq \beta^p V(k, \xi), \end{aligned}$$

then the nonlinear system (11) is uniformly exponentially stable in \mathcal{D} , with a rate β .

The converse Lyapunov theorem for the discrete-time case shows the implication in the opposite direction.

Theorem 3. *If the nonlinear system (11) is uniformly exponentially stable in \mathcal{D} , then there exists a continuous function $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathbb{R}$ and constants $c_1, c_2 \in \mathbb{R}_+$ and $\beta \in (0, 1)$, such that for all $\xi \in \mathcal{D}$, and $k \in \mathbb{N}_{\geq k_0}$*

$$c_1 \|\xi\|^2 \leq V(k, \xi) \leq c_2 \|\xi\|^2, \quad (13)$$

$$V(k+1, f(k, \xi)) \leq \beta^2 V(k, \xi). \quad (14)$$

The proof for Theorem 3 follows directly from [44, Thm 2], by adapting the global result to the case with forward invariant subspace \mathcal{D} . The above theorems generalize to uniform global exponential stability if $\mathcal{D} = \mathbb{R}^n$ [42], [44]. In continuous-time, the rate of change of the Lyapunov function with respect to the state is bounded by the norm of the state [42, Thm 4.14]. The following Lemma is the discrete-time equivalent of this.

Lemma 4. *Let Assumption 5.i. hold, then if the system (11) is uniformly exponentially stable in \mathcal{D} , there exists a constant $c_3 \in \mathbb{R}_+$ and a continuous Lyapunov function $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathbb{R}$, that, in addition to (13) and (14), for all $(\xi, \zeta) \in \mathcal{D} \times \mathcal{D}$, and $k \in \mathbb{N}_{\geq k_0}$ also satisfies*

$$|V(k, \xi) - V(k, \zeta)| \leq c_3 \|\xi - \zeta\| (\|\xi\| + \|\zeta\|).$$

The proof of Lemma 4 is provided in the appendix.

B. Preliminaries on Exponential Incremental Stability

Exponential incremental stability shows the exponential convergence of two trajectories generated by the same system to each other [11], [13].

Definition 3. *The system (11) is uniformly exponentially incrementally stable in \mathcal{D} if there exist constants $d \in \mathbb{R}_+$ and $\lambda \in (0, 1)$, such that for all $(\xi, \zeta) \in \mathcal{D} \times \mathcal{D}$ and $k \in \mathbb{N}_{\geq k_0}$*

$$\|\phi(k, k_0, \xi) - \phi(k, k_0, \zeta)\| \leq d \|\xi - \zeta\| \lambda^{k-k_0}.$$

If this holds only for a subset of initial conditions in $\mathcal{D} \times \mathcal{D}$ the system is uniformly locally exponentially incrementally stable in this subset. If $\mathcal{D} = \mathbb{R}^n$ the definition is global.

The theory of exponential convergence of two trajectories of the same system has first been studied as *uniform convergence* by Demidovich [33], [45], and later extended through contraction theory [12], [13]. Contraction is a sufficient condition for exponential incremental stability, that when f is smooth, can be checked by the following condition.

Theorem 4. *Let Assumption 5.ii. hold, and suppose that there exists a sequence of positive-definite, bounded matrices $P(k) \in \mathbb{R}^{n \times n}$, $k \in \mathbb{N}_{\geq k_0}$, and some $\rho \in (0, 1)$ such that*

$$D(k, x) := \frac{\partial f}{\partial x}(k, x)^T P(k+1) \frac{\partial f}{\partial x}(k, x) - \rho^2 P(k) \prec 0 \quad (15)$$

for all $x \in \tilde{\mathcal{D}}_0 \subseteq \mathcal{D}_0$. Then, there exists a constant $r_{\tilde{\mathcal{D}}_0} \in \mathbb{R}_+$, such that (11) is uniformly locally exponentially incrementally stable in $\mathcal{B}(r_{\tilde{\mathcal{D}}_0}) \subseteq \tilde{\mathcal{D}}_0$ with a rate ρ .

We also state the following converse Lyapunov theorem.

Theorem 5. *If the system (11) is uniformly exponentially incrementally stable in \mathcal{D} , then there exists a function $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}_+$ and constants $c_1, c_2 \in \mathbb{R}_+$ and $\beta \in (0, 1)$, such that for all $(\xi, \zeta) \in \mathcal{D} \times \mathcal{D}$ and $k \in \mathbb{N}_{\geq k_0}$*

$$c_1 \|\xi - \zeta\|^2 \leq V(k, \xi, \zeta) \leq c_2 \|\xi - \zeta\|^2, \quad (16)$$

$$V(k+1, f(k, \xi), f(k, \zeta)) \leq \beta^2 V(k, \xi, \zeta). \quad (17)$$

Moreover, under Assumption 5.i., there exists a constant $c_3 \in \mathbb{R}_+$, such that for all $(\xi, \zeta, \tilde{\xi}, \tilde{\zeta}) \in \mathcal{D} \times \mathcal{D} \times \mathcal{D} \times \mathcal{D}$ and $k \in \mathbb{N}_{\geq k_0}$

$$\begin{aligned} |V(k, \xi, \zeta) - V(k, \tilde{\xi}, \tilde{\zeta})| &\leq \\ &\leq c_3 \left(\|\xi - \tilde{\xi}\| + \|\zeta - \tilde{\zeta}\| \right) \left(\|\xi - \zeta\| + \|\tilde{\xi} - \tilde{\zeta}\| \right). \end{aligned} \quad (18)$$

The proofs of Theorem 4, adapted from [13], and Theorem 5, extended from [10] and [11], are provided in the appendix.

C. Main Results

In this subsection, we show that if the nonlinear dynamics (11) are uniformly exponentially stable in \mathcal{D} and satisfy Assumption 5, then they are also E - δ -ISS in the same region. First, we show that under the local Lipschitz continuity assumption, exponential incremental stability implies E - δ -ISS.

Theorem 6. *Let Assumption 5 hold, then if the nonlinear system (11) is uniformly exponentially incrementally stable in \mathcal{D} , it is E - δ -ISS in the same region.*

For completeness, we provide the proof in the appendix, and note that the implication is similar to existing results in the literature, e.g. [46], that studies the asymptotic stability of the equilibrium, or [47, Lemma 2] that considers a non-autonomous setting.

Next, we show that if in addition to local Lipschitz continuity, the nonlinear dynamics are also locally differentiable in some arbitrarily small region \mathcal{D}_0 around the equilibrium, then uniform exponential stability implies uniform exponential incremental stability and E - δ -ISS. The results are formalized in Theorem 7 and Corollary 2.

Theorem 7. *Let Assumption 5 hold, then if the nonlinear system (11) is uniformly exponentially stable in \mathcal{D} , it is also uniformly exponentially incrementally stable in the same region.*

Proof. We start by showing that the exponential stability of f implies that the linearized dynamics around the origin are also stable by following similar arguments to [42]. Let

$$A(k) := \frac{\partial f(k, x)}{\partial x}(k, \mathbf{0}),$$

which is well-defined given Assumption 5. Moreover, there exists a $\bar{A} \in \mathbb{R}_+$, such that $\|A(k)\| \leq \bar{A}$, $k \in \mathbb{N}_{\geq k_0}$. It follows from Theorem 3, that there exists a continuous mapping $V : k \in \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathbb{R}$ satisfying (13)-(14). Let us consider V as a candidate Lyapunov function for $A(k)$. Then for all $x \in \mathcal{D}$,

$k \in \mathbb{N}_{\geq k_0}$ there exist constants $\beta \in (0, 1)$, and $c_4, d \in \mathbb{R}_+$ such that

$$\begin{aligned} V(k+1, A(k)x) &= \\ V(k+1, f(k, x)) + [V(k+1, A(k)x) - V(k+1, f(k, x))] &\leq \beta^2 V(k, x) + [V(k+1, A(k)x) - V(k+1, f(k, x))] \\ &\leq \beta^2 V(k, x) + c_4 \|f(k, x) - A(k)x\| \|f(k, x) + A(k)x\| \\ &\leq \beta^2 V(k, x) + c_4 \|x\| \cdot \|f(k, x) - A(k)x\| (d\beta + \|A(k)\|), \end{aligned}$$

where the first inequality follows from Theorem 3, the second from Lemma 4 and the last from Definition 2 and properties of induced norms. Denoting $h(k, x) := f(k, x) - A(k)x$, it follows from the Lipschitz continuity of the Jacobian of f [42, Chpt. 4.6] that there exists a $L_h \in \mathbb{R}_+$, such that $\|h(k, x)\| \leq L_h \|x\|^2$, for all $k \in \mathbb{N}_{\geq k_0}$ for all $x \in \mathcal{D}_0$. Using this

$$\begin{aligned} V(k+1, A(k)x) &\leq \beta^2 V(k, x) + c_4 L_h (d\beta + \bar{A}) \|x\|^3 \\ &\leq \left(\beta^2 + \frac{c_4 L_h (d\beta + \bar{A})}{c_1} \|x\| \right) V(k, x) := \gamma V(k, x), \end{aligned}$$

where the second inequality follows from the converse Lyapunov Theorem 3 for some $c_1 \in \mathbb{R}_+$. Defining $r_1 := \min \left\{ \frac{c_1(1-\beta^2)}{c_4 L_h (d\beta + \bar{A})}, \max\{r > 0 \mid \|x\| < r, x \in \mathcal{D}_0\} \right\}$, for all $\|x\| < r_1$, it holds that $\gamma < 1$. Hence, using Theorem 2 the linearized dynamics $x_{k+1} = A(k)x_k$ are uniformly exponentially stable. Then, from [48, Thm 23.3] there exists a sequence of uniformly positive definite $P(k)$ that solves the difference Lyapunov equation $A^\top(k)P(k+1)A(k) - P(k) \leq -cI$, for some $c \in \mathbb{R}_+$.

Considering now equation (15) for some $k \in \mathbb{N}_{\geq k_0}$, $\bar{x} \in \mathcal{D}_0$ and denoting $d_k(\bar{x}) := \frac{\partial f}{\partial x}(k, \bar{x}) - A(k)$, it holds that

$$\begin{aligned} &\frac{\partial f}{\partial x}(k, \bar{x})^\top P(k+1) \frac{\partial f}{\partial x}(k, \bar{x}) - P(k) \\ &= [A(k) + d_k(\bar{x})]^\top P(k+1) [A(k) + d_k(\bar{x})] - P(k) \\ &= A(k)^\top P(k+1)A(k) - P(k) \\ &\quad + A(k)^\top P(k+1)d_k(\bar{x}) + d_k(\bar{x})^\top P(k+1) [A(k) + d_k(\bar{x})] \\ &\leq -cI + A(k)^\top P(k+1)d_k(\bar{x}) \\ &\quad + d_k(\bar{x})^\top P(k+1) [A(k) + d_k(\bar{x})], \end{aligned}$$

where the first equality follows from the definition of $d_k(\bar{x})$ and the last one from [48, Thm 23.3].

Following the same arguments as in [42, Chpt. 4.6], there exists a constant $L_2 \in \mathbb{R}_+$, such that for all $k \in \mathbb{N}_{\geq k_0}$ and $\bar{x} \in \mathcal{D}_0$, $\|d_k(\bar{x})\| \leq L_2 \|\bar{x}\|$. Pre- and post-multiplying the above with some x^\top and x , respectively then yields

$$\begin{aligned} &x^\top \left[\frac{\partial f}{\partial x}(k, \bar{x})^\top P(k+1) \frac{\partial f}{\partial x}(k, \bar{x}) - P(k) \right] x \leq \\ &\quad - c \|x\|^2 + 2 \|x\|^2 \|A(k)^\top P(k+1)d_k(\bar{x})\| \\ &\quad \quad + \|x\|^2 \|d_k(\bar{x})^\top P(k+1)d_k(\bar{x})\| \\ &\leq - \|x\|^2 (c - 2\bar{A}\bar{P}L_2 \|x\| - \bar{P}r_d L_2^2 \|x\|), \end{aligned}$$

where $r_d = \max_{x \in \mathcal{D}} \|x\|$, and $\|P(k)\| \leq \bar{P}$, $k \in \mathbb{N}_{\geq k_0}$. Note that the rate of exponential stability of the linear system $A(k)$ is $\sqrt{1 - \frac{c}{\bar{P}}} \in (0, 1)$. Then, choosing $\rho^2 > 1 - \frac{c}{\bar{P}}$, adding

$x^\top (1 - \rho^2) P(k)x$ to both sides of the above inequality and defining $r_2 := \frac{c-(1-\rho^2)\bar{P}}{2\bar{A}\bar{P}L_2 + \bar{P}r_dL_2^2}$ ensures that

$$x^\top \left[\frac{\partial f}{\partial x}(k, x)^\top P(k+1) \frac{\partial f}{\partial x}(k, x) - \rho^2 P(k) \right] x < 0,$$

uniformly in k and x , for all $\|x\| < r := \min(r_1, r_2)$. This implies by Theorem 4 that there exists a positive constant $r_{\bar{\mathcal{D}}_0} \leq r$, such that for all $x \in \mathcal{B}(r_{\bar{\mathcal{D}}_0})$ the system (11) is uniformly exponentially incrementally stable with rate of ρ .

To show that the system is also uniformly exponentially incrementally stable in \mathcal{D} , consider any $\xi_1, \xi_2 \in \mathcal{D}$, then for all $k \in \mathbb{N}_{\geq k_0}$, and some $d \in \mathbb{R}_+$, the following two inequalities hold

$$\|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \leq \|\phi(k, k_0, \xi_1)\| + \|\phi(k, k_0, \xi_2)\| \leq 2r_d d \beta^{k-k_0}, \quad (19)$$

$$\|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \leq L_f^{k-k_0} \underbrace{\|\xi_1 - \xi_2\|}_{:=\Delta\xi}. \quad (20)$$

The bound in (19) follows from the exponential stability of f , and the one in (20) from its Lipschitz continuity. Combining the two

$$\|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \leq \min\{2r_d d \beta^{k-k_0}, L_f^{k-k_0} \|\Delta\xi\|\}.$$

Define $k' \in \mathbb{N}_{\geq k_0}$ such that both $\|\phi(k', k_0, \xi_1)\| < r_{\bar{\mathcal{D}}_0}$, $\|\phi(k', k_0, \xi_2)\| < r_{\bar{\mathcal{D}}_0}$. Then, from the above analysis, there exists a $d' \in \mathbb{R}_+$, such that for all $k \in \mathbb{N}_{\geq k'}$

$$\begin{aligned} & \|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \\ & \leq d' \rho^{k-k'} \cdot \|\phi(k', k_0, \xi_1) - \phi(k', k_0, \xi_2)\|. \end{aligned}$$

It then follows that

$$\begin{aligned} & \|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \\ & \leq d' \rho^{k-k'} \cdot \min\{r_d d \cdot \beta^{k-k_0}, L_f^{k-k_0} \|\Delta\xi\|\}. \end{aligned}$$

Note that for all $k = \{k_0, \dots, k'\}$

$$\|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \leq c_5 \|\Delta\xi\| \rho^{k-k_0},$$

where $c_5 := \max\{1, \frac{L_f^{k'}}{\rho^{k'}}\}$, makes it a constant independent of $\|\Delta\xi\|$, and since k' is finite, also independent of time. Combining the bounds, the following holds for all $k \in \mathbb{N}_{\geq k_0}$

$$\|\phi(k, k_0, \xi_1) - \phi(k, k_0, \xi_2)\| \leq c_5 d' \rho^{k-k_0} \|\Delta\xi\|,$$

which is the definition of uniform exponential incremental stability. \square

Combining Theorems 6 and 7 the following corollary follows directly.

Corollary 2. *Let Assumption 5 hold, then if the nonlinear system (11) is uniformly exponentially stable in \mathcal{D} , it is also E- δ -ISS in the same region.*

Remark 1. *The conditions on the Jacobian matrix of f in Assumption 5 are required only in the time-varying case, as these are satisfied trivially in the time-invariant setting; for further details, we refer the reader to [42, Theorems 4.7, 4.12].*

Remark 2. *It is worth highlighting that the implication in Corollary 2 is also claimed in [11, Thm. 12] without assuming local differentiability. While the proof in [11] holds for asymptotic stability, the same arguments break down in the exponential case since one cannot find constants $d \in \mathbb{R}_+$ and $\lambda \in (0, 1)$ as in Definition 3 that are independent of the initial deviation $\|\xi - \zeta\|$. Hence, the results in this section serve to fill this gap by deriving sufficient conditions in the form of Assumption 5 under which the implication holds.*

In the sequel, we use these insights to address the closed-loop dynamics under suboptimal and optimal linear MPC policies as a notable use case.

V. MODEL PREDICTIVE CONTROL - A USE CASE

In Section III, we showed that under certain assumptions on the suboptimal policy μ (Assumption 2) and the benchmark policy μ^* (Assumption 1), the suboptimality gap of μ can be bounded for a certain family of costs. We now exploit the results of Section IV to show that these assumptions are satisfied in the case of suboptimal, PGM-based linear-quadratic MPC, and derive an asymptotically tighter suboptimality gap bound compared to [17] under the same assumptions.

Consider the control of linear time-invariant dynamical systems, modeled by

$$x_{k+1} = Ax_k + Bu_k, \quad k \in \mathbb{N},$$

where $A \in \mathbb{R}^{n \times n}$, and $B \in \mathbb{R}^{n \times m}$ are known system matrices, and the state and input vectors are defined as before. We consider the finite horizon linear-quadratic regulator (LQR) problem with the objective of minimizing the finite-time cost

$$J_T(x_0, \mathbf{u}) = \|x_T\|_P^2 + \sum_{k=0}^{T-1} \|x_k\|_Q^2 + \|u_k\|_R^2, \quad (21)$$

where $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are design matrices and P is taken to be the solution of the discrete-time Algebraic Riccati Equation, $P = Q + K^\top RK + (A - BK)^\top (A - BK)$, with $K = (R + B^\top PB)^{-1} (B^\top PA)$. The control inputs must satisfy $u_k \in \mathcal{U}$ for all $k \in \mathbb{N}$, where $\mathcal{U} \subseteq \mathbb{R}^m$ is a constraint set. The following standard assumptions ensure a unique minimizer for (21) always exists [16], [49].

Assumption 6. (*Well-posed problem*)

- i. *The pair (A, B) is stabilizable, $Q \succ 0$, $R \succ 0$.*
- ii. *The input constraint set \mathcal{U} is closed, convex and contains the origin in its interior.*

The model predictive controller solves this problem in a receding horizon fashion, solving the following parametric optimal control problem (POCP) at each timestep k , having measured a state $x \in \mathbb{R}^n$

$$\begin{aligned} & \mu^*(x) := \arg \min_{\nu} J_N(\xi_0, \nu) \\ & \text{s.t. } \xi_{i+1} = A\xi_i + B\nu_i, \quad i=0, \dots, N-1, \\ & \xi_0 = x, \quad \nu_i \in \mathcal{U}, \quad i=0, \dots, N-1. \end{aligned} \quad (22)$$

Here N is the prediction horizon length, and $\nu = [\nu_0^\top \dots \nu_{N-1}^\top]^\top$ denotes the predicted input vector. We refer

to the minimiser of (22) for a given initial state (parameter) $x \in \mathbb{R}^n$, $\mu^*(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{Nm}$, as the optimal mapping. For this example, we take as the benchmark policy μ^* , the map solving the POCP (22). The optimal cost attained by this mapping is denoted by $J_N^*(x) := J_N(x, \mu^*(x))$, which serves as an approximate value function for the problem. For each k , the model predictive controller applies the first element of $\mu^*(x_k)$ to the system, and the process is repeated in a receding horizon fashion. The optimal state evolution under this optimal MPC policy is then given by

$$x_{k+1}^* = Ax_k^* + \bar{B}\mu^*(x_k^*) := f(x_k^*), \quad k \in \mathbb{N}, \quad (23)$$

where $x_0^* := x_0$, $\bar{B} := BS$, and $S := [I_{m \times m} \ \mathbf{0} \ \dots \ \mathbf{0}] \in \mathbb{R}^{m \times Nm}$ is a matrix selecting the first control input. Note that the optimal MPC policy, μ^* is time-invariant due to the structure of the problem.

Problem (22) is a parametric quadratic program and for a given parameter $x \in \mathbb{R}^n$ can be represented in an equivalent condensed form $J_N^*(x) = \min_{\nu \in \mathcal{N}} \|(x, \nu)\|_M^2$, where $\mathcal{N} = \mathcal{U}^N \subseteq \mathbb{R}^{Nm}$

$$M = \begin{bmatrix} W & G^\top \\ G & H \end{bmatrix}, \quad (24)$$

and the definitions of $H \in \mathbb{R}^{Nm \times Nm}$, $W \in \mathbb{R}^{n \times n}$ and, $G \in \mathbb{R}^{Nm \times n}$ can be found in [9].

As the optimal $\mu^*(x)$ may often be infeasible to compute exactly, suboptimal schemes are often considered. In our setting, a suboptimal policy is computed by performing only a finite number of optimization steps for (22). In particular, given $x \in \mathbb{R}^n$ and an input vector $\nu \in \mathbb{R}^{Nm}$, consider the operator that performs one step of the projected gradient method

$$\mathcal{T}(x, \nu) := \Pi_{\mathcal{N}}[\nu - \alpha \nabla \nu J_N(x, \nu)], \quad (25)$$

where $\alpha \in \mathbb{R}$ is a step size. Applying (25) iteratively $\ell_k \in \mathbb{N}_+$ times provides an approximation for the optimal input, and hence the optimal policy. The combined dynamics of the system and the approximate optimizer are then given by³

$$u_k^\mu = \mathcal{T}^{\ell_k}(x_k, u_{k-1}^\mu), \quad (26a)$$

$$x_{k+1} = Ax_k + \bar{B}u_k^\mu, \quad (26b)$$

where $u_{-1}^\mu \in \mathbb{R}^{Nm}$ is an initialization vector, and for some $l \in \mathbb{N}$, $x \in \mathbb{R}^n$ and $\nu \in \mathbb{R}^{Nm}$, we define

$$\mathcal{T}^l(x, \nu) = \mathcal{T}(x, \mathcal{T}^{l-1}(x, \nu)), \quad \mathcal{T}^0(x, \nu) = \nu. \quad (27)$$

The dynamics under the suboptimal policy are (26b) by taking $u_k^\mu = \mu_k(x_k)$ for all $k \in \mathbb{N}$, i.e.

$$x_{k+1} = \underbrace{Ax_k + \bar{B}\mu^*(x_k)}_{f(x_k)} + \underbrace{\bar{B}(u_k^\mu - \mu^*(x_k))}_{=d_k}, \quad k \in \mathbb{N}.$$

Note that μ_k is also a function of the previous input state u_{k-1}^μ . However, since the closed-loop evolution is noise-free, it can be uniquely determined given the initialization vector, the current time k , and the current state. Hence, the dependence on u_{k-1}^μ is encoded in the subscript of μ_k . The suboptimal policy can in general be defined as a function of the information vector $\mathcal{I}_k = \{x_k, u_{k-1}, \dots, u_0\}$; as long as Assumption 2 is satisfied, the results in this manuscript hold.

³The subscript of ℓ_k is dropped when it is taken to be a constant.

A. Optimal MPC

In this subsection, we review the properties of the optimal mapping $\mu^*(x)$. As shown in [9], [50], [51], the system (23) is exponentially stable with the forward invariant ROA estimate

$$\Gamma_N := \{x \in \mathbb{R}^n \mid \psi(x) \leq r_N\},$$

where $\psi(x) := \sqrt{J_N^*(x)}$, $d = c \cdot \lambda^-(Q)/\lambda^+(P)$, $r_N = \sqrt{Nd} + c$ and $c > 0$ is such that the following set

$$\Omega = \{x \in \mathbb{R}^n \mid \|x\|_P^2 \leq c\},$$

also satisfies $\Omega \subset \{x \in \mathbb{R}^n \mid -Kx \in \mathcal{U}\}$. The function $\psi(x)$ is a Lyapunov function for the optimal MPC algorithm, satisfying

$$\|x\|_P \leq \psi(x) \leq \|x\|_W \quad (28)$$

$$\psi(f(x)) \leq \beta\psi(x), \quad (29)$$

where $\beta \in (0, 1)$ is the exponential decay rate. The Lipschitz continuity of the optimal mapping is formalized in the following lemma.

Lemma 5. [9, Corollary 2] For any $(x, y) \in \Gamma_N \times \Gamma_N$, the optimal solution mapping, $\mu^*(x)$, satisfies

$$\|\mu^*(x) - \mu^*(y)\| \leq \|H^{-\frac{1}{2}}\| \|G(x - y)\|_{H^{-1}} \leq L\|x - y\|$$

with a Lipschitz constant $L := \|H^{-\frac{1}{2}}\| \cdot \|H^{-\frac{1}{2}}G\|$.

The proof follows from the parametric quadratic program structure of the MPC problem and is derived in [9] or [16] from an explicit MPC point of view.

B. Suboptimal MPC

The suboptimal policy in this setting is defined by (26a). The following well-established result shows the linear rate of convergence of the PGM method.

Theorem 8. [39, Theorem 3.1] For any $x \in \mathbb{R}^n$, $\nu \in \mathbb{R}^{Nm}$, $\ell \in \mathbb{N}_+$, and for $\alpha = \frac{1}{\lambda^+(H) + \lambda^-(H)}$

$$\|\mathcal{T}^\ell(x, \nu) - \mu^*(x)\| \leq \eta^\ell \|\nu - \mu^*(x)\|,$$

where $\eta = (\lambda^+(H) - \lambda^-(H))/(\lambda^+(H) + \lambda^-(H))$.

The suboptimal MPC scheme is treated by considering the combined evolution of the system-optimizer dynamics (26). The stability of such a scheme, also referred to as Time-Distributed MPC (TD-MPC) or as real-time implementation of MPC is shown in [9], [24], [25] for a fixed number of iterations ℓ and in [17] for a time-varying ℓ_k . In particular, if $\ell_k > \ell^*$ for all $k \in \mathbb{N}$, where

$$\ell^* = \frac{\log(1 - \beta) - \log(\sigma\kappa + \omega(1 - \beta))}{\log(\eta)},$$

where β is the same as in Section V-A, $\omega = 1 + \|H^{-\frac{1}{2}}\| \|H^{-\frac{1}{2}}G\bar{B}\|$, $\sigma = \|W^{\frac{1}{2}}\bar{B}\|$, and

$$\begin{aligned} \kappa &= \|H^{-\frac{1}{2}}\| \|H^{-\frac{1}{2}}G(A - I)P^{-\frac{1}{2}}\| \\ &+ \|H^{-\frac{1}{2}}\| \sqrt{\lambda_H^+(G\bar{B})(\lambda_P^+(W) - 1)}, \end{aligned}$$

Then, the combined dynamics (26) are exponentially stable in the following forward invariant ROA estimate

$$\Sigma_N = \left\{ (x, z) \in \Gamma_N \times \mathcal{N} \mid \psi(x) \leq r_N, \right. \\ \left. \|z - \mu^*(x)\| \leq \frac{(1 - \beta)r_N}{\sigma} \right\},$$

recalling r_N from Section V-A. For the rest of the section we take $u_{-1}^\mu = 0$, with the assumption that the resulting $u_0^\mu = \mathcal{T}^\ell(x_0, u_{-1}^\mu)$ is such that $(x_0, u_0^\mu) \in \Sigma_N$. For further analysis of the choice of the initial input and the forward invariant dynamics, we refer the interested reader to [50].

The exponential stability result from [17] is summarized in the following Lemma.

Lemma 6. [17, Lemma 5] *Given the dynamics (26), the following holds for all $(x_0, u_0^\mu) \in \Sigma_N$, $k \in \mathbb{N}$ and $\ell_k > \ell^*$*

$$\|x_k\| \leq h_0 \|P^{-\frac{1}{2}}\| \cdot \|x_0\|_W \prod_{i=-1}^k \varepsilon_i,$$

where $\varepsilon_k := \max\{\beta + \tau\kappa\eta^{\ell_k}, \frac{\sigma + \tau\eta^{\ell_k}\omega}{\tau}\} \in (0, 1)$, $\varepsilon_{-1} := 1$ and $h_0 = 1 + \tau\eta^{\ell_0}L\|W^{-\frac{1}{2}}\|$.

If the same number of optimization iterations are taken at all times the bound reduces to the following.

Corollary 3. [17, Corollary 3] *Given the dynamics (26), the following holds for all $\ell > \ell^*$, $(x_0, u_0^\mu) \in \Sigma_N$ and $k \in \mathbb{N}$*

$$\|x_k\| \leq h \|P^{-\frac{1}{2}}\| \cdot \|x_0\|_W \cdot \varepsilon^k,$$

where $\varepsilon := \max\{\beta + \tau\kappa\eta^\ell, \frac{\sigma + \tau\eta^\ell\omega}{\tau}\} \in (0, 1)$ and $h = 1 + \tau\eta^\ell L\|W^{-\frac{1}{2}}\|$.

C. Suboptimality Gap

We define the suboptimality vector, $\tilde{\eta}_\ell$, for this use case as

$$\tilde{\eta}_\ell := [\tilde{\eta}_{\ell,1} \quad \tilde{\eta}_{\ell,2} \quad \dots \quad \tilde{\eta}_{\ell,T-1}],$$

where $\tilde{\eta}_{\ell,k} := \eta^{\ell k}(1 + \tilde{\eta}_{\ell,k+1})$, $k \in [0, T-2]$, and $\tilde{\eta}_{\ell,T-1} := \eta^{\ell(T-1)}$. We denote the Euclidean norm of the suboptimality vector by $\|\tilde{\eta}_\ell\|$. The main result for a suboptimal MPC scheme is summarized in the following Theorem.

Theorem 9. *Let Assumption 6 hold, then if $\ell_k > \ell^*$, $\forall k \in \mathbb{N}$, the suboptimality gap of the suboptimal MPC (26) is given by*

$$\mathcal{R}_T(x_0) = \mathcal{O}(\|\tilde{\eta}_\ell\| \|x_0\|),$$

for all $(x_0, u_0^\mu) \in \Sigma_N$. Specifically, it is bounded by

$$\mathcal{R}_T(x_0) \leq \bar{M} (\tilde{\eta}_{\ell,0} \|\delta u_0\| + L\Delta^\top \tilde{\eta}_\ell),$$

with \bar{M}, Δ defined as in Theorem 1 and $\delta u_0 = u_{-1}^\mu - \mu_0^*(x_0)$. Moreover, if $\ell_k = \ell > \ell^*$, $\forall k \in \mathbb{N}$, then, for all $(x_0, u_0^\mu) \in \Sigma_N$

$$\mathcal{R}_T(x_0) = \mathcal{O}\left(\frac{\eta^\ell}{1 - \eta^\ell} \|x_0\|\right).$$

Proof. We start by showing that Assumptions 1-4 are satisfied in the MPC use-case. In this setting, the linear dynamics to be controlled are given by

$$x_{k+1} = Ax_k + \bar{B}\bar{u}_k, \quad k \in \mathbb{N}, \quad (30)$$

with the input $\bar{u}_k \in \mathbb{R}^{N^m}$. First, we note that Assumption 3 is satisfied trivially with $L_u = \|\bar{B}\|$. The benchmark controller is the optimal policy μ^* that solves the POCP (22), and is given in (23).

Taking \mathcal{D}^* to be the forward invariant ROA estimate $\mathcal{D}^* = \Gamma_N$, it follows from Lemma 5 that Assumption 1.i. is satisfied. Since the policy is time-invariant, Assumption 1.ii. is satisfied trivially with $a_k = 0$ for all k . To show that the Assumption 1.iii. is satisfied, we use the analysis from Section IV. Exploiting the structure of the optimization problem (22), it has been shown in [16] that under Assumption 6, the solution of the MPC problem is piecewise-affine in the state. Moreover, the set Ω , that defines a forward invariant set around the origin for the dynamics (23) with inactive input constraints, is non-empty [4]. This and Lemma 5 imply that Assumption 5 is satisfied. As discussed in Section V-A, the benchmark dynamics (23) are exponentially stable in \mathcal{D}^* , therefore, by Corollary 2 we conclude that they are also E - δ -ISS in the same region. The suboptimal policy is given by (26a). The results in [17] show that for all $\ell_k \geq \ell^*$, Σ_N is a forward invariant ROA estimate for the combined dynamics (26). Given this and an initial $u_{-1}^\mu = 0$, consider $\mathcal{D}^\mu = \{x \in \mathbb{R}^n \mid (x, \mathcal{T}^\ell(x, u_{-1}^\mu)) \in \Sigma_N\}$. Then Assumption 2 is satisfied directly from Theorem 8 and the suboptimal policy definition.

Finally, to reconcile the quadratic cost defined in (21) and the modified dynamics (30), we redefine the cost, as

$$J_T(x_0, \bar{u}) = \|x_T\|_P^2 + \sum_{k=0}^{T-1} \|x_k\|_Q^2 + \|\bar{u}_k\|_{\bar{R}}^2,$$

where $\bar{R} = S^\top R S$. For the quadratic costs in (21), and for any $(x, y) \in \mathcal{D}^\mu \times \mathcal{D}^\mu$, and $(u, z) \in (\mathcal{N} \times \mathcal{N})$

$$\begin{aligned} |x^\top Q x - y^\top Q y + u^\top R u - z^\top R z| \\ \leq 2\|x - y\| \|Q\| x_m + 2\|u - z\| \|R\| u_m, \end{aligned}$$

where x_m and u_m are such that, $\|x\| \leq x_m$ and $\|u\| \leq u_m$ for all $x \in \mathcal{D}^\mu, u \in \mathcal{N}$. Then the condition in Assumption 4 is satisfied with $M_x = 2x_m \max\{\|Q\|, \|P\|\}$ and $M_u = 2u_m \|\bar{R}\|$.

Since Assumption 6 implies Assumptions 1-4 are satisfied, we can invoke the bound in Theorem 1 for the suboptimality gap. As the suboptimal dynamics are exponentially stable, its pathlength is finite. In particular

$$\begin{aligned} \mathcal{S}_{T,2}^2 &= \sum_{k=1}^T \|x_k - x_{k-1}\|^2 \leq 4 \sum_{k=0}^T \|x_k\|^2 \\ &\leq c_0^2 \|x_0\|^2 \sum_{k=0}^T \prod_{i=0}^k \varepsilon_{i-1}^2 \leq \frac{c_0^2}{1 - \varepsilon^2} \|x_0\|^2, \end{aligned}$$

where the first inequality follows by the triangle inequality, the second from Lemma 6 and by denoting $c_0 := 2h_0 \|P^{-\frac{1}{2}}\| \|W^{\frac{1}{2}}\|$, and the last one by bounding the geometric series and denoting $\bar{\varepsilon} := \max\{\varepsilon_i\}_{i=0}^{T-1}$. Noting that $a_k = 0$ for all $k \in \mathbb{N}$ in this example, the suboptimality gap is bounded by $\mathcal{O}(\mathcal{S}_{T,2} \tilde{\eta}_\ell) = \mathcal{O}(\|\tilde{\eta}_\ell\| \|x_0\|)$.

In the case when $\ell_k = \ell > \ell^*$ for all $k \in \mathbb{N}$, one can use the bound derived in Corollary 1, with a simplified expression

for the pathlength

$$\mathcal{S}_T \leq \frac{c}{1-\varepsilon} \|x_0\|.$$

The above is obtained directly from the bound in Corollary 3 and by denoting $c := 2h\|P^{-\frac{1}{2}}\| \|W^{\frac{1}{2}}\|$. \square

The theorem shows that the suboptimality gap of the MPC suboptimal controller (26) scales with $\tilde{\eta}_\ell$ or η^ℓ , where the number of iterations ℓ_k are design parameters. Note that the higher ℓ_k the more computation is required at each timestep, but the lower the suboptimality gap; in the limit, as $\ell \rightarrow \infty$, the suboptimality gap is zero. This is a tighter result than the one derived in [17], as in the latter no incremental properties of the optimal controller are used. Specifically, when looking at the limit case of $\eta_k = 0$, $k \in \mathbb{N}$, the suboptimality gap in [17] is strictly positive, while the bound in Theorem 9 vanishes, reflecting the exact matching of the suboptimal and benchmark trajectories. The derived bounds can be used by control designers to give a quantifiable measure of the finite-time suboptimality of the controller. This can then be utilized to find the best sequence of ℓ_k to deliver a desired tradeoff between suboptimality and computational power.

In practice, the suboptimal MPC can be asymptotically stable even when the number of optimization iterations ℓ is less than ℓ^* . In this case, the existence of a forward invariant region of attraction \mathcal{D}^μ is not given, but the bounds in Theorem 1 and Corollary 1 still hold, as long as the closed-loop suboptimal system stays stable. This is shown in the following numerical example.

D. Numerical Example

The suboptimal TD-MPC scheme described in this section is implemented for the following linearized, continuous-time model of an inverted pendulum from [50], [17]

$$A_c = \begin{bmatrix} 0 & 1 \\ 14.7 & 0 \end{bmatrix}, B_c = \begin{bmatrix} 0 \\ 30 \end{bmatrix},$$

where the state is $x = [\theta, \dot{\theta}]^\top$, θ is the angle relative to the equilibrium position and the control input is the applied torque. We consider the control of the discretized model of the plant with a sampling time of $T_s = 0.1$. The input constraint set is taken to be $\mathcal{U} = [-1, 1]$, the cost matrices are $Q = I_2$, and $R = 1$ and the initial state is $x_0 = [-\pi/4 \ \pi/3]^\top$.

The left panel of Figure 3 shows the evolution of two trajectories in closed-loop with the TD-MPC policy with $\ell = 6$ and with an optimal MPC. For this example $\ell^* = 849$. However, even with the low value of $\ell = 6$ the closed-loop system stays stable, as also observed in [9], [50]. Although the asymptotic/exponential stability cannot be proven, the finite time bounds can still be computed online using only the suboptimal states as per Corollary 1. The order of this upper bound, as well as the empirically observed suboptimality gap, $\mathcal{R}_T(x_0)$, are plotted in the right panel of Figure 3 for $T = 30$ for a range of values of ℓ , increasing from 1 to 5000. The decrease of the suboptimality gap for increasing values of ℓ is juxtaposed with the increase of simulation/computational time in the same figure. The simulation time for each ℓ is

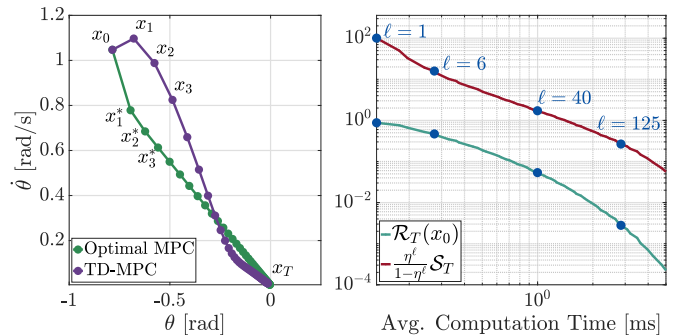


Fig. 3: The phase plot on the left shows two separate trajectories generated by applying respectively TD-MPC, with a constant $\ell = 6$ (in green) and optimal MPC policies (in purple). The logarithmic scale plot on the right shows the empirical suboptimality gap (in green), as well as the order of the upper bound decrease (in red) as ℓ , and therefore the computation time is increased. The two curves on the right are parameterized by ℓ , ranging from 1 to 5000.

calculated as the sum of the times it takes to solve the TD-MPC for each timestep, over the horizon T . To obtain an averaged value for this time, its average over 100 repeated independent runs from the same initial conditions is taken. The initial states are initialized in Σ_N following the procedure described in [50]. In the right panel of the figure, only the complexity $\frac{\eta^\ell}{1-\eta^\ell} \mathcal{S}_T$ of the suboptimality gap upper bound is plotted. The true constant multiplying the complexity term above is much larger; our aim here is not to compare the very conservative theoretical bound with the practical performance, but to give an estimate of whether the bound captures the order correctly. Indeed, it can be observed from the figure that the order of the true suboptimality gap is approximately captured by the upper bound with an underestimation.

Among other possible uses of the closed-loop suboptimality analysis, the insights in the figure can be used to design the allocation of finite computational resources. The right-side plot in the figure can be used to estimate the relative gain in computational time and loss in optimality for a given change in ℓ . For example, a change of $\ell = 6$ to $\ell = 40$, or equivalently $\eta^\ell = 0.92$ to $\eta^\ell = 0.56$ results in a 3.6 times increase in computational time and a 9.3 times decrease in the suboptimality gap bound. This provides an approximation for the variation of the true suboptimality gap that decreased 8.7 times in the same interval.

VI. CONCLUSIONS

We study the finite-time suboptimality gap of policies for discrete-time nonlinear, time-varying systems. We show that when the benchmark policy is chosen to be exponentially incrementally stable, then given a geometric improvement condition on the suboptimal policy, its suboptimality gap scales with its pathlength and improvement factor. We further show, that for non-smooth nonlinear systems E - δ -ISS is implied by exponential stability under certain conditions. The assumptions are verified on the suboptimal linear quadratic

MPC use case and on a numerical example. The generality of the provided analysis enables the study of other examples where the suboptimality is due to unknown system parameters or cost functions, for example in the fields of adaptive and online control. We leave the analysis of these use cases as well as the treatment of measurement and process noise to future work.

APPENDIX

Proof of Lemma 4

Proof. Given a state $\xi \in \mathcal{D}$ and some $N \in \mathbb{R}_+$, let

$$V(k_0, \xi) = \sum_{k=0}^{N-1} \phi(k+k_0, k_0, \xi)^\top \phi(k+k_0, k_0, \xi).$$

Then

$$V(k_0, \xi) = \xi^\top \xi + \sum_{k=1}^{N-1} \phi(k+k_0, k_0, \xi)^\top \phi(k+k_0, k_0, \xi) \geq \|\xi\|^2,$$

and, from Definition 2

$$V(k_0, \xi) \leq \sum_{k=0}^{N-1} d^2 \|\xi\|^2 \lambda^{2k} \leq \frac{d^2}{1-\lambda^2} \|\xi\|^2.$$

Thus, (13) is satisfied with $c_1 = 1$ and $c_2 = \frac{d^2}{1-\lambda^2}$. To show that (14) holds, consider

$$\begin{aligned} & V(k_0+1, f(k_0, \xi)) - V(k_0, \xi) \\ &= \sum_{k=0}^{N-1} (\|\phi(k+k_0+1, k_0, \xi)\|^2 - \|\phi(k+k_0, k_0, \xi)\|^2) \\ &= \|\phi(N+k_0, k_0, \xi)\|^2 - \|\xi\|^2 \leq d^2 \lambda^{2N} \|\xi\|^2 - \|\xi\|^2 \\ &= -(1-d^2 \lambda^{2N}) \|\xi\|^2. \end{aligned}$$

Choosing N large enough such that $d^2 \lambda^{2N} < 1$, ensures (14) holds with $\beta^2 = 1 - \frac{(1-d^2 \lambda^{2N})}{c_2} \in (0, 1)$ since $c_2 \geq c_1 = 1$ and $c_2 \in \mathbb{R}_+$. Finally, for some $(\xi, \zeta) \in \mathcal{D} \times \mathcal{D}$ and $k_0 \in \mathbb{N}$, denote $\Delta\phi(k, k_0, \xi, \zeta) := \phi(k+k_0, k_0, \xi) - \phi(k+k_0, k_0, \zeta)$ and consider

$$\begin{aligned} & |V(k_0, \xi) - V(k_0, \zeta)| \\ &= \left| \sum_{k=0}^{N-1} (\|\phi(k+k_0, k_0, \xi)\|^2 - \|\phi(k+k_0, k_0, \zeta)\|^2) \right| \\ &\leq \sum_{k=0}^{N-1} (\|\phi(k+k_0, k_0, \xi)\| + \|\phi(k+k_0, k_0, \zeta)\|) \\ &\quad \cdot \|\Delta\phi(k, k_0, \xi, \zeta)\| \\ &\leq \sum_{k=0}^{N-1} d\lambda^k (\|\xi\| + \|\zeta\|) \cdot L_f^k \|\xi - \zeta\| \\ &= (\|\xi\| + \|\zeta\|) \|\xi - \zeta\| \sum_{k=0}^{N-1} d\lambda^k L_f^k, \end{aligned}$$

where the last inequality follows from the exponential stability and L_f -Lipschitz continuity of the nonlinear mapping. Taking $c_3 = \sum_{k=0}^{N-1} d\lambda^k L_f^k$ completes the proof. \square

Proof of Theorem 4

Proof. The condition in (15) implies the contraction of the dynamics within the contraction region $\tilde{\mathcal{D}}_0$, as defined in [12]. We will first show that there exists a $r_{\tilde{\mathcal{D}}_0}$ such that for all initial states $x \in \mathcal{B}(r_{\tilde{\mathcal{D}}_0})$ the dynamics (11) evolve in a forward invariant set contained in $\tilde{\mathcal{D}}_0$. To see this, for some $r \in \mathbb{R}_+$, and $k \in \mathbb{N}_{\geq k_0}$ define $\mathcal{E}_0(k, r) := \{x \in \mathbb{R}^n \mid \|x\|_{P(k)} \leq r\}$, and $r^* := \max\{r > 0 \mid \mathcal{E}_0(k, r) \subseteq \tilde{\mathcal{D}}_0, \forall k \in \mathbb{N}_{\geq k_0}\}$. This implies that if a state satisfies $x \in \mathcal{E}_0(k, r^*)$, then $x \in \tilde{\mathcal{D}}_0$ necessarily. The largest ball contained in the intersection of the sets $\mathcal{E}_0(k, r^*)$, $k \in \mathbb{N}_{\geq k_0}$ is then given by $\mathcal{B}(\frac{r^*}{\bar{P}})$, where $P(k) \preceq \bar{P}I$ for all $k \in \mathbb{N}_{\geq k_0}$, and $\bar{P} > 0$ exists since $P(k)$ is uniformly bounded. Consider now the differential displacement δx for the dynamics (11), defining an infinitesimal displacement at a given time k . The shortest path integral with respect to the metric $P(k)$ between 0 and a given state x_k at time k is given by

$$V_\ell(\delta x_k, k) := \int_0^{x_k} \|\delta x\|_{P(k)} = \|x_k\|_{P(k)},$$

where the second equality follows from the fact that the shortest path integral corresponds to the Riemannian distance (see e.g. [11], [13], [52]). Using the same arguments as in [13, Thm 2.8] or [11, Thm. 15], one can then show that for any $x_k \in \mathcal{B}(\frac{r^*}{\bar{P}})$, it holds that $V_\ell(\delta x_{k+1}, k+1) \leq \rho V_\ell(\delta x_k, k) < r^*$. This, in turn, implies that $\|x_{k+1}\|_{P(k+1)} < r^* \implies x_{k+1} \in \mathcal{E}_0(k+1, r^*) \implies x_{k+1} \in \tilde{\mathcal{D}}_0$. Taking $r_{\tilde{\mathcal{D}}_0} = \frac{r^*}{\bar{P}}$ completes the claim that the dynamics are forward invariant within $\tilde{\mathcal{D}}_0$. For any initial state in $\mathcal{B}(r_{\tilde{\mathcal{D}}_0})$, uniform local exponential incremental stability with a rate ρ then follows from standard Lyapunov arguments, see e.g. [13, Thm 2.8]. \square

Proof of Theorem 5

The proof hinges on extending results from [44], [53], [10] and [11]. Before presenting the proof, we provide the following auxiliary definition and theorem.

Definition 4. [44], [53] *Given a closed, positively invariant set $\mathcal{A} \subset \mathbb{R}^n$, the system (11) is uniformly exponentially stable in \mathcal{D} with respect to \mathcal{A} , if there exist constants $d \in \mathbb{R}_+$ and $\lambda \in (0, 1)$, such that for all $\xi \in \mathcal{D}$ and $k \in \mathbb{N}_{\geq k_0}$*

$$\|\phi(k, k_0, \xi)\|_{\mathcal{A}} \leq d \|\xi\|_{\mathcal{A}} \lambda^{k-k_0}.$$

The extension of the converse Lyapunov results from [10] and [44] to the case of stability in a forward invariant region is included below for completeness.

Theorem 10. *If the system (11) is uniformly exponentially stable in \mathcal{D} with respect to some closed set \mathcal{A} , then there exists a function $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \rightarrow \mathbb{R}_+$ and constants $c_1, c_2 \in \mathbb{R}_+$ and $\beta \in (0, 1)$, such that for all $\xi \in \mathcal{D}$, and $k \in \mathbb{N}_{\geq k_0}$*

$$c_1 \|\xi\|_{\mathcal{A}}^2 \leq V(k, \xi) \leq c_2 \|\xi\|_{\mathcal{A}}^2, \quad (31)$$

$$V(k+1, f(k, \xi)) \leq \beta^2 V(k, \xi). \quad (32)$$

Proof. Following the same line of arguments from the proof of Lemma 4, consider a $x \in \mathcal{D}$ and some $N \in \mathbb{R}_+$ and let

$$V(k_0, \xi) = \sum_{k=0}^{N-1} |\phi(k+k_0, k_0, \xi)|_{\mathcal{A}}^2, \quad (33)$$

for some $k_0 \in \mathbb{N}$. Then,

$$V(k_0, \xi) = |\xi|_{\mathcal{A}}^2 + \sum_{k=1}^{N-1} |\phi(k+k_0, k_0, \xi)|_{\mathcal{A}}^2 \geq |\xi|_{\mathcal{A}}^2,$$

and, from Definition 4

$$V(k_0, \xi) \leq \frac{d^2}{1-\lambda^2} |\xi|_{\mathcal{A}}^2.$$

Next, consider

$$\begin{aligned} & V(k_0+1, f(k_0, \xi)) - V(k_0, \xi) \\ &= \sum_{k=0}^{N-1} (|\phi(k_0+k+1, k_0, \xi)|_{\mathcal{A}}^2 - |\phi(k_0+k, k_0, \xi)|_{\mathcal{A}}^2) \\ &= |\phi(k_0+N, \xi)|_{\mathcal{A}}^2 - |\xi|_{\mathcal{A}}^2 \leq d^2 \lambda^{2N} |\xi|_{\mathcal{A}}^2 - |\xi|_{\mathcal{A}}^2 \\ &= -(1-d^2 \lambda^{2N}) |\xi|_{\mathcal{A}}^2. \end{aligned}$$

Choosing N large enough such that $d^2 \lambda^{2N} < 1$ completes the proof with $c_1 = 1$, $c_2 = \frac{d^2}{1-\lambda^2}$, and $\beta^2 = 1 - \frac{(1-d^2 \lambda^{2N})}{c_2} \in (0, 1)$ since $c_2 \geq c_1 = 1$ and $c_2 \in \mathbb{R}_+$. \square

Proof of Theorem 5. Consider the augmented system

$$\begin{cases} \xi_{k+1} = f(k, \xi_k) \\ \zeta_{k+1} = f(k, \zeta_k). \end{cases}$$

We define the diagonal as the set $\Delta := \{[\xi^\top, \zeta^\top]^\top \in \mathcal{D} \times \mathcal{D} : \xi \in \mathcal{D}\}$. Let $z := [\xi^\top, \zeta^\top]^\top \in \mathcal{D} \times \mathcal{D}$, then it is shown in [10] that

$$|z|_{\Delta} = \frac{\sqrt{2}}{2} \|\xi - \zeta\|. \quad (34)$$

Then, considering the evolution of the combined system

$$z_{k+1} = \mathcal{F}(k, z_k) := \begin{bmatrix} f(k, \xi_k) \\ f(k, \zeta_k) \end{bmatrix}, \quad (35)$$

one can note that $z_k \in \mathcal{D} \times \mathcal{D}$ for all $k \in \mathbb{N}_{\geq k_0}$, for any $(\xi_{k_0}, \zeta_{k_0}) \in \mathcal{D} \times \mathcal{D}$, since \mathcal{D} is forward invariant. It follows that system (11) is exponentially incrementally stable in \mathcal{D} if and only if (35) is exponentially stable in \mathcal{D} with respect to the diagonal Δ .

Moreover, using Theorem 10, the combined system (35) admits a Lyapunov function satisfying (31)-(32). Given a $z \in \mathcal{D} \times \mathcal{D}$, it follows from the equivalence in (34), and the proof of Theorem 10, that (16)-(17) are satisfied with $c_1 = \frac{\sqrt{2}}{2}$, and $c_2 = \frac{d^2}{\sqrt{2}-\lambda^2\sqrt{2}}$ for the following Lyapunov function

$$V(k_0, z) = \sum_{k=0}^{N-1} |\phi_{\mathcal{F}}(k_0+k, k_0, z)|_{\Delta}^2,$$

where $\phi_{\mathcal{F}} : \mathbb{N}_{\geq k_0} \times \mathbb{N}_{\geq k_0} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ is the state transition matrix for the system (35).

We show the inequality (18) next. For any $(\xi, \zeta, \tilde{\xi}, \tilde{\zeta}) \in \mathcal{D} \times \mathcal{D} \times \mathcal{D} \times \mathcal{D}$, $k \in \mathbb{N}_{\geq k_0}$, $w := [\tilde{\xi}^\top, \tilde{\zeta}^\top]^\top$ and z defined as above

$$\begin{aligned} & |V(k_0, z) - V(k_0, w)| \\ &= \left| \sum_{k=0}^{N-1} (|\phi_{\mathcal{F}}(k_0+k, k_0, z)|_{\Delta}^2 - |\phi_{\mathcal{F}}(k_0+k, k_0, w)|_{\Delta}^2) \right| \\ &\leq \sum_{k=0}^{N-1} (|\phi_{\mathcal{F}}(k_0+k, k_0, z) - \phi_{\mathcal{F}}(k_0+k, k_0, w)|) \\ &\quad \cdot (|\phi_{\mathcal{F}}(k_0+k, k_0, z)|_{\Delta} + |\phi_{\mathcal{F}}(k_0+k, k_0, w)|_{\Delta}) \\ &\leq \sum_{k=0}^{N-1} \frac{d\lambda^k}{\sqrt{2}} \left\| \begin{bmatrix} \phi(k_0+k, k_0, \xi) - \phi(k_0+k, k_0, \tilde{\xi}) \\ \phi(k_0+k, k_0, \zeta) - \phi(k_0+k, k_0, \tilde{\zeta}) \end{bmatrix} \right\| \\ &\quad \cdot (\|\xi - \zeta\| + \|\tilde{\xi} - \tilde{\zeta}\|) \\ &\leq (\|\xi - \tilde{\xi}\| + \|\zeta - \tilde{\zeta}\|) (\|\xi - \zeta\| + \|\tilde{\xi} - \tilde{\zeta}\|) \sum_{k=0}^{N-1} \frac{d\lambda^k L^k}{\sqrt{2}}, \end{aligned}$$

where the last two inequalities follow from the Lipschitz continuity and exponential stability (with respect to Δ) of the solutions. The first inequality follows from the following. For any two vectors $\xi, \zeta \in \mathcal{D}$ and closed, convex set \mathcal{A}

$$\begin{aligned} |\xi|_{\mathcal{A}} - |\zeta|_{\mathcal{A}} &= |\xi|_{\mathcal{A}} - \|\zeta - \zeta_p\| \leq \|\xi - \zeta_p\| - \|\zeta - \zeta_p\| \\ &\leq \|\xi - \zeta_p\| - \|\zeta - \zeta_p\| \leq \|\xi - \zeta\|, \end{aligned}$$

where $\zeta_p \in \mathcal{A}$ is such that $|\zeta|_{\mathcal{A}} = \|\zeta - \zeta_p\|$, which exists and is unique, since \mathcal{A} is closed and convex. Finally, choosing $c_3 = \sum_{k=0}^{N-1} \frac{d\lambda^k L^k}{\sqrt{2}}$ completes the proof. \square

Proof of Theorem 6

Proof. Given the nonlinear system (11), consider the evolution of two, respectively unperturbed and perturbed trajectories

$$\begin{aligned} x_{k+1} &= f(k, x_k), \\ y_{k+1} &= f(k, y_k) + w_k, \end{aligned}$$

for some $(x_{k_0}, y_{k_0}) \in \mathcal{D} \times \mathcal{D}$, where $w_k \in \mathcal{B}(r_w)$, for some $r_w \in \mathbb{R}_+$ is such that $y_k \in \mathcal{D}$ for all $k \in \mathbb{N}_{\geq k_0}$. Given f is uniformly exponentially incrementally stable in \mathcal{D} , then from Theorem 5, there exists a Lyapunov function $V : \mathbb{N}_{\geq k_0} \times \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}_+$ satisfying (16)-(18). Then, for any $(x, y) \in \mathcal{D} \times \mathcal{D}$, $k \in \mathbb{N}_{\geq k_0}$, and $w \in \mathcal{B}(r_w)$

$$\begin{aligned} & V(k+1, f(k, x), f(k, y) + w) - V(k, x, y) = \\ &= V(k+1, f(k, x), f(k, y)) - V(k, x, y) \\ &+ V(k+1, f(k, x), f(k, y) + w) - V(k+1, f(k, x), f(k, y)) \\ &\leq -(1-\beta^2) c_1 \|x - y\|^2 \\ &\quad + c_3 \|w\| (\|f(k, x) - f(k, y) - w\| + \|f(k, x) - f(k, y)\|) \\ &\leq -(1-\beta^2) c_1 \|x - y\|^2 + c_3 \|w\|^2 + 2c_3 L_f \|w\| \|x - y\|, \end{aligned}$$

where the first inequality follows from Theorem 5, and the second from the triangle inequality. Completing the square,

and denoting $c_4 := (1 - \beta^2)c_1$ it follows from above that

$$\begin{aligned} V(k+1, f(k, x), f(k, y) + w) - V(k, x, y) &\leq \\ &\leq \frac{-3c_4}{2} \|x - y\|^2 \\ &\quad + \left(\frac{\sqrt{c_4}}{\sqrt{2}} \|x - y\| + \frac{c_3 L_f \sqrt{2} \|w\|}{\sqrt{c_4}} \right)^2 - \frac{2c_3^2 L_f^2 \|w\|^2}{c_4} \\ &\leq -\frac{c_4}{2} \|x - y\|^2 + \frac{2c_3^2 L_f^2 \|w\|^2}{c_4} \\ &\leq -\frac{c_4}{2c_2} V(k, x, y) + \frac{2c_3^2 L_f^2 \|w\|^2}{c_4}, \end{aligned}$$

where the second inequality follows by the fact that $(a + b)^2 \leq 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$, and the last inequality from Theorem 5. Finally, rearranging the Lyapunov equations it follows that

$$V(k+1, f(k, x), f(k, y) + w) \leq \rho^2 V(k, x, y) + c_5 \|w\|^2,$$

for all $k \in \mathbb{N}_{\geq k_0}$, with $\rho^2 := 1 - \frac{c_4}{2c_2} \in (0, 1)$, since $c_2 > c_4$, and $c_5 := \frac{2c_3^2 L_f^2}{c_4}$. Unrolling the recursion and using (16)

$$\begin{aligned} c_1 \|x_k - y_k\|^2 &\leq \rho^{2(k-k_0)} V(k_0, x_{k_0}, y_{k_0}) + c_5 \sum_{i=k_0}^{k-1} \rho^{2(k-i-1)} \|w_i\|^2 \\ &\leq c_2 \rho^{2(k-k_0)} \|x_{k_0} - y_{k_0}\|^2 + c_5 \sum_{i=k_0}^{k-1} \rho^{2(k-i-1)} \|w_i\|^2. \end{aligned}$$

Dividing by c_1 and taking the square root, completes the proof

$$\begin{aligned} \|x_k - y_k\| &\leq \sqrt{\frac{c_2}{c_1}} \rho^{(k-k_0)} \|x_{k_0} - y_{k_0}\| \\ &\quad + \sqrt{\frac{c_5}{c_1}} \sum_{i=k_0}^{k-1} \rho^{k-i-1} \|w_i\|. \end{aligned}$$

□

REFERENCES

- [1] L. S. Pontryagin, *Mathematical theory of optimal processes*. CRC press, 1987.
- [2] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.
- [3] D. Bertsekas, *Lessons from AlphaZero for optimal, model predictive, and adaptive control*. Athena Scientific, 2022.
- [4] D. Bertsekas, *Dynamic programming and optimal control: Volume I*, vol. 4. Athena scientific, 2012.
- [5] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [6] K. J. Åström and B. Wittenmark, *Adaptive control*. Courier Corporation, 2013.
- [7] Y. Li, X. Chen, and N. Li, "Online optimal control with linear dynamics and predictions: Algorithms and regret analysis," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [8] E. Hazan, S. Kakade, and K. Singh, "The nonstochastic control problem," in *Algorithmic Learning Theory*, pp. 408–421, PMLR, 2020.
- [9] D. Liao-McPherson, T. Skibik, J. Leung, I. Kolmanovsky, and M. M. Nicotra, "An analysis of closed-loop stability for linear model predictive control based on time-distributed optimization," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2618–2625, 2021.
- [10] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, 2002.
- [11] D. N. Tran, B. S. Rüffer, and C. M. Kellett, "Convergence properties for discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3415–3422, 2018.
- [12] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.
- [13] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview," *Annual Reviews in Control*, vol. 52, pp. 135–169, 2021.
- [14] A. Davydov, S. Jafarpour, and F. Bullo, "Non-euclidean contraction theory for robust nonlinear stability," *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6667–6681, 2022.
- [15] F. Bullo, *Contraction Theory for Dynamical Systems*. Kindle Direct Publishing, 1.1 ed., 2023.
- [16] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.
- [17] A. Karapetyan, E. C. Balta, A. Iannelli, and J. Lygeros, "On the finite-time behavior of suboptimal linear model predictive control," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 5053–5058, IEEE, 2023.
- [18] N. Hovakimyan and C. Cao, *L1 adaptive control theory: Guaranteed robustness with fast adaptation*. SIAM, 2010.
- [19] N. M. Boffi, S. Tu, and J.-J. E. Slotine, "Regret bounds for adaptive nonlinear control," in *Learning for Dynamics and Control*, pp. 471–483, PMLR, 2021.
- [20] A. Karapetyan, E. C. Balta, A. Tsiamis, A. Iannelli, and J. Lygeros, "On the regret of recursive methods for discrete-time adaptive control with matched uncertainty," *2024 63rd IEEE Conference on Decision and Control (CDC)*, arXiv preprint arXiv:2404.02023, 2024.
- [21] G. Belgioioso, D. Liao-McPherson, M. H. de Badyn, S. Bolognani, R. S. Smith, J. Lygeros, and F. Dörfler, "Online feedback equilibrium seeking," *IEEE Transactions on Automatic Control*, vol. 70, no. 1, pp. 203–218, 2025.
- [22] Z. He, S. Bolognani, J. He, F. Dörfler, and X. Guan, "Model-free nonlinear feedback optimization," *IEEE Transactions on Automatic Control*, 2023.
- [23] M. Diehl, R. Findeisen, F. Allgöwer, H. G. Bock, and J. P. Schlöder, "Nominal stability of real-time iteration scheme for nonlinear model predictive control," *IEE Proceedings-Control Theory and Applications*, vol. 152, no. 3, pp. 296–308, 2005.
- [24] A. Zanelli, Q. Dinh, and M. Diehl, "A lyapunov function for the combined system-optimizer dynamics in nonlinear model predictive control," *Automatica*, 2020.
- [25] D. Liao-McPherson, M. M. Nicotra, and I. Kolmanovsky, "Time-distributed optimization for real-time model predictive control: Stability, robustness, and constraint satisfaction," *Automatica*, vol. 117, p. 108973, 2020.
- [26] M. N. Zeilinger, C. N. Jones, D. M. Raimondo, and M. Morari, "Real-time MPC-stability through robust MPC design," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 3980–3986, IEEE, 2009.
- [27] S. Richter, C. N. Jones, and M. Morari, "Computational complexity certification for real-time MPC with input constraints based on the fast gradient method," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1391–1403, 2011.
- [28] M. N. Zeilinger, C. N. Jones, and M. Morari, "Real-time suboptimal model predictive control using a combination of explicit MPC and online optimization," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1524–1534, 2011.
- [29] L. Grüne, J. Pannek, M. Seehafer, and K. Worthmann, "Analysis of unconstrained nonlinear MPC schemes with time varying control horizon," *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 4938–4962, 2010.
- [30] A. Ferramosca, D. Limón, I. Alvarado, and E. F. Camacho, "Cooperative distributed MPC for tracking," *Automatica*, vol. 49, no. 4, pp. 906–914, 2013.
- [31] K. P. Wabersich and M. N. Zeilinger, "Cautious bayesian MPC: regret analysis and bounds on the number of unsafe learning episodes," *IEEE Transactions on Automatic Control*, vol. 68, no. 8, pp. 4896–4903, 2022.
- [32] Y. Lin, Y. Hu, G. Qu, T. Li, and A. Wierman, "Bounded-regret MPC via perturbation analysis: Prediction error, constraints, and nonlinearity," *Advances in Neural Information Processing Systems*, vol. 35, pp. 36174–36187, 2022.
- [33] B. P. Demidovich, "Lectures on stability theory (in Russian)," 1967.
- [34] A. L. Dontchev, M. Krastanov, R. T. Rockafellar, and V. M. Veliov, "An euler–newton continuation method for tracking solution trajectories

of parametric variational inequalities,” *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 1823–1840, 2013.

- [35] S. M. Robinson, “Strongly regular generalized equations,” *Mathematics of Operations Research*, vol. 5, no. 1, pp. 43–62, 1980.
- [36] A. L. Dontchev, I. Kolmanovsky, M. I. Krastanov, M. Nicotra, and V. M. Veliov, “Lipschitz stability in discretized optimal control with application to SQP,” *SIAM Journal on Control and Optimization*, vol. 57, no. 1, pp. 468–489, 2019.
- [37] M. N. Zeilinger, M. Morari, and C. N. Jones, “Soft constrained model predictive control with robust stability guarantees,” *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1190–1202, 2014.
- [38] N. Chatzikiriakos, K. P. Wabersich, F. Berkel, P. Pauli, and A. Iannelli, “Learning soft constrained MPC value functions: Efficient MPC design and implementation providing stability and safety guarantees,” in *6th Annual Learning for Dynamics & Control Conference*, pp. 387–398, PMLR, 2024.
- [39] A. B. Taylor, J. M. Hendrickx, and F. Glineur, “Exact worst-case convergence rates of the proximal gradient method for composite convex minimization,” *Journal of Optimization Theory and Applications*, vol. 178, no. 2, pp. 455–476, 2018.
- [40] R. Nishihara, L. Lessard, B. Recht, A. Packard, and M. Jordan, “A general analysis of the convergence of ADMM,” in *International conference on machine learning*, pp. 343–352, PMLR, 2015.
- [41] A. Srikanthan, A. Karapetyan, V. Kumar, and N. Matni, “Closed-loop analysis of admm-based suboptimal linear model predictive control,” *IEEE Control Systems Letters*, vol. 8, pp. 3195–3200, 2024.
- [42] H. Khalil, “Nonlinear systems, third edition, vol. 115,” *Upper Saddle River, NJ, USA: Prentice-Hall*, 2002.
- [43] W. M. Haddad and V. Chellaboina, *Nonlinear dynamical systems and control: a Lyapunov-based approach*. Princeton university press, 2008.
- [44] Z.-P. Jiang and Y. Wang, “A converse Lyapunov theorem for discrete-time systems with disturbances,” *Systems & control letters*, vol. 45, no. 1, pp. 49–58, 2002.
- [45] A. Pavlov, A. Pogromsky, N. van de Wouw, and H. Nijmeijer, “Convergent dynamics, a tribute to Boris Pavlovich Demidovich,” *Systems & Control Letters*, vol. 52, no. 3-4, pp. 257–261, 2004.
- [46] Z.-P. Jiang and Y. Wang, “Input-to-state stability for discrete-time nonlinear systems,” *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.
- [47] J. Köhler and M. N. Zeilinger, “Globally stable and locally optimal model predictive control using a softened initial state constraint-extended version,” *arXiv preprint arXiv:2207.10216*, 2022.
- [48] W. J. Rugh, *Linear system theory*. Prentice-Hall, Inc., 1996.
- [49] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert, “Constrained model predictive control: Stability and optimality,” *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.
- [50] J. Leung, D. Liao-McPherson, and I. V. Kolmanovsky, “A computable plant-optimizer region of attraction estimate for time-distributed linear model predictive control,” in *2021 American Control Conference (ACC)*, pp. 3384–3391, IEEE, 2021.
- [51] D. Limón, T. Alamo, F. Salas, and E. F. Camacho, “On the stability of constrained MPC without terminal constraint,” *IEEE transactions on automatic control*, vol. 51, no. 5, pp. 832–836, 2006.
- [52] S. Singh, A. Majumdar, J.-J. Slotine, and M. Pavone, “Robust online motion planning via contraction theory and convex optimization,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5883–5890, IEEE, 2017.
- [53] E. D. Sontag and Y. Wang, “New characterizations of input-to-state stability,” *IEEE transactions on automatic control*, vol. 41, no. 9, pp. 1283–1294, 1996.



Aren Karapetyan received his MEng Engineering Science degree from the University of Oxford, Oxford, United Kingdom in 2020, with the specialization in Information and Control Engineering. He worked on the Enhanced Cabin Monitoring project for electric vehicles at the Jaguar Land Rover Research Department in Coventry, United Kingdom in 2018 and 2019. Since September 2020, he is a doctoral student at the Automatic Control Laboratory (IfA), ETH Zürich, Zürich, Switzerland, working on sequential decision making under uncertainty and computational constraints.



Efe C. Balta received the B.S. degree in manufacturing engineering from the Faculty of Mechanical Engineering, Istanbul Technical University, in 2016, and the M.S. and Ph.D. degrees in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2018 and 2021, respectively. He was a Post-Doctoral Researcher with the Automatic Control Laboratory (IfA), ETH Zürich between 2021 and 2023. Since September 2023, he has been leading the Control and Automation research group at inspire AG. His research interests include control theory, optimization, statistical learning, robotics, cyber-physical systems, and additive manufacturing.



Andrea Iannelli is an Assistant Professor in the Institute for Systems Theory and Automatic Control at the University of Stuttgart (Germany). He completed his B.Sc. and M.Sc. degrees in Aerospace Engineering at the University of Pisa (Italy) and received his PhD from the University of Bristol (United Kingdom) on robust control and dynamical systems theory. He was a postdoctoral researcher in the Automatic Control Laboratory at ETH Zürich (Switzerland). His main research interests are at the intersection of control theory, optimization, and learning, with a particular focus on robust and adaptive optimization-based control, uncertainty quantification, and sequential decision-making problems. He serves the community as Associated Editor for the International Journal of Robust and Nonlinear Control and as IPC member of international conferences in the areas of control, optimization, and learning.



John Lygeros completed a B.Eng. degree in electrical engineering in 1990 and an M.Sc. degree in Systems Control in 1991, both at Imperial College of Science Technology and Medicine, London, U.K.. In 1996 he obtained a Ph.D. degree from the Electrical Engineering and Computer Sciences Department, University of California, Berkeley. During the period 1996–2000 he held a series of research appointments at the National Automated Highway Systems Consortium, Berkeley, the Laboratory for Computer Science, M.I.T., and the Electrical Engineering and Computer Sciences Department at U.C. Berkeley. Between 2000 and 2003 he was a University Lecturer at the Department of Engineering, University of Cambridge, U.K., and a Fellow of Churchill College. Between 2003 and 2006 he was an Assistant Professor at the Department of Electrical and Computer Engineering, University of Patras, Greece. In July 2006 he joined the Automatic Control Laboratory at ETH Zurich, first as an Associate Professor, and since January 2010 as a Full Professor. Between 2015 and 2018 he served as the Head of the Department of Information Technology and Electrical Engineering; since 2009 he is serving as the Head of the Automatic Control Laboratory. His research interests include modelling, analysis, and control of hierarchical, hybrid, and stochastic systems, with applications to biochemical networks, transportation systems, energy systems, and industrial processes. John Lygeros is a Fellow of the IEEE, and a member of the IET and the Technical Chamber of Greece; between 2013 and 2023 he served as the Vice President for Finances and a Council Member of the International Federation of Automatic Control (IFAC), as well as the Board of the IFAC Foundation. Since 2020 he serves as the Director of the National Centre of Competence in Research “Dependable Ubiquitous Automation” (NCCR Automation)