Bounded-Regret MPC via Perturbation Analysis: Prediction Error, Constraints, and Nonlinearity

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study Model Predictive Control (MPC) and propose a general analysis pipeline to bound its dynamic regret. The pipeline first requires deriving a perturbation bound for a finite-time optimal control problem. Then, the perturbation bound is used to bound the per-step error of MPC, which leads to a bound on the dynamic regret. Thus, our pipeline reduces the study of MPC to the well-studied problem of perturbation analysis, enabling the derivation of regret bounds of MPC under a variety of settings. To demonstrate the power of our pipeline, we use it to generalize existing regret bounds on MPC in linear time-varying (LTV) systems to incorporate prediction errors on costs, dynamics, and disturbances. Further, our pipeline leads to regret bounds on MPC in systems with nonlinear dynamics and constraints.

1 Introduction

Model Predictive Control (MPC) is an optimal control approach that solves a Finite-Time Optimal Control Problem (FTOCP) using future predictions in a receding horizon manner [1]. It is a flexible approach that is able to accommodate nonlinear and time-varying dynamics, state and actuation constraints, and general cost functions [2–5]. As a result, it is broadly applied in a wide spectrum of control problems, including robotics [6–10], autonomous vehicles [11–17], power systems [18–24], process control [25–27], etc.

Despite the popularity of MPC, its theoretic analysis has been quite challenging. Early works along this line focused on the stability and recursive feasibility of MPC [28–31]. More recently, there has been tremendous interest in providing finite-time learning-theoretic performance guarantees for MPC, such as regret and/or competitive ratio bounds [32, 33]. For example, progress has recently been made toward (i) regret analysis of MPC in linear time-invariant (LTI) systems with prediction errors on the trajectory to track [34], (ii) the dynamic regret and competitive ratio bounds of MPC under linear time-varying (LTV) dynamics with exact predictions [35], and (iii) exponentially decaying perturbation bounds of the finite-time optimal control problem in time-varying, constrained, and non-linear systems [36, 37]. Beyond MPC, providing regret and/or competitive ratio guarantees for a variety of (predictive) control policies has been a focus in recent years. Examples include RHGC [38, 39] and AFHC [20, 40] for online control/optimization with prediction horizons, OCO-based controllers [41, 42] for no-regret online control, and variations of ROBD for competitive online control without predictions [43, 44] or with delayed observations [45]. In addition, regret lower bounds have been studied in known LTI systems [46] and unknown LTV systems [47].

A promising analysis approach that has emerged from the literature studying MPC and, more generally, predictive control, is the use of perturbation analysis techniques, or more particularly, the

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

use of so-called exponential decaying perturbation bounds. Such techniques underlie the results in [34–37]. This research direction is particularly promising since perturbation bounds exist for FTOCP in many dynamical systems, e.g., [48–52], and thus it potentially allows the derivation of regret and/or competitive ratio bounds in a variety of settings. However, to this point the approach has only yielded results in unconstrained linear systems with no prediction errors (e.g., [35]), and often requires adjusting MPC to include a counter-intuitively large re-planning window due to technical challenges in the analysis (e.g., [48, 49]).

Thus, though perturbation analysis techniques might seem promising, many important questions about applying them for the study of predictive control remain open. Firstly, one of the major reasons for the extensive application of MPC is its flexibility in incorporating constraints and nonlinear dynamics [53]. However, none of the existing results and approaches can analyze the performance of MPC under constraints and/or nonlinear dynamics. In fact, the anlyasis of MPC under constraints or nonlinearity has long been known to be challenging because of the intractable form of cost-to-go functions and optimal solutions. Secondly, prediction error is inevitable for real-world implementations of MPC due to unpredictable noise and model mismatch, yet the analysis of MPC subject to prediction errors is limited. Thirdly, existing approaches analyze MPC in a case-by-case manner and, in most cases, the analysis framework is specific to the assumptions of the particular case (e.g. quadratic costs, perfect predictions, etc) in a way that does not generalize to other settings [33–35, 48, 49].

Contributions. In this paper, we propose a general analysis pipeline (Section 3) that converts perturbation bounds for an FTOCP into dynamic regret bounds for MPC across a variety of settings. More specifically, the pipeline consists of three steps (see Figure 1). In Step 1, we obtain the required perturbation bounds for the specific setting. In Step 2, as shown in Lemma 3.1, the perturbation bounds are used to bound the *per-step error*, which is defined to be the error of the MPC action against the clairvoyant optimal action (see Definition 3.1). In Step 3, the per-step error bound is converted to a dynamic regret bound for MPC, as shown in Lemma 3.2. The full pipeline is summarized into a *Pipeline Theorm* (Theorem 3.3), which directly converts perturbation bounds into bounds on the dynamic regret of MPC in general settings, including those with time-variation, prediction error, constraints, and nonlinearities. The key technical insight that enables the pipeline is the following recursive relationship between Step 2 and Step 3 (Lemma 3.1 and Lemma 3.2): Step 2 guarantees a "small" per-step error e_t once the current state x_t of MPC is "near" the offline optimal trajectory (OPT), while Step 3 guarantees the next state x_{t+1} of MPC will be near OPT if all previous per-step errors ($\{e_\tau\}_{\tau \leq t}$) are small. Thus Step 2 and Step 3 work together to guarantee MPC states are always near OPT and thus MPC per-step errors are always small (Theorem 3.3).

To demonstrate the power of the proposed pipeline, we apply it to a range of settings, as summarized in Table 1. Our first applications are to two settings with linear time-varying (LTV) dynamics and prediction errors on (i) disturbances, Section 4.1, and (ii) the dynamical matrices and cost functions, Section 4.2. The state-of-the-art results in the LTV setting are [35], which requires exact knowledge of the disturbances and of the dynamics. To the best of our knowledge, our work provides the first regret result for MPC with prediction error on the dynamics (see Theorem 4.2), a result that enables the bounds in settings where MPC is applied to learned dynamics [54].

Our second application is to a setting with nonlinear dynamics and constraints (Section 5). We show the first dynamic regret bound for MPC under state and actuation constraints in nonlinear systems with general costs (Theorem 5.1). Very few prior results exist for MPC in this setting, even with nonlinear dynamics or constraints individually. The most related works are [48], which studies constrained MPC, and [49], which studies nonlinear MPC. In both cases, a counter-intuitive re-planning window is added to MPC to facilitate the analysis, a downside that our pipeline could avoid. Besides, [48] and [49] require exact predictions of the cost functions, dynamics, and constraints for the exponential convergence property of MPC to hold, while our result can apply to more general noisy predictions.

2 Preliminaries

We consider a general, finite-horizon, discrete-time optimal control problem with *time-varying costs, dynamics and constraints*, namely

$$\begin{aligned}
\min_{x_{0:T}, u_{0:T-1}} \; & \sum_{t=0}^{T-1} f_t(x_t, u_t; \xi_t^*) + F_T(x_T; \xi_T^*) \\
\text{s.t. } & x_{t+1} = g_t(x_t, u_t; \xi_t^*), && \forall 0 \le t < T, \\
& s_t(x_t, u_t; \xi_t^*) \le 0, && \forall 0 \le t < T, \\
& x_0 = x(0).
\end{aligned} \tag{1}$$

Here, $x_t \in \mathbb{R}^n$ is the *state*, $u_t \in \mathbb{R}^m$ is the *control input* or *action*; f_t is a time-varying *stage cost* function, g_t is a time-varying *dynamical* function, and s_t is a time-varying *constraint* function, all parameterized by a ground-truth parameter ξ_t^* (unknown to an online controller); and F_T is a terminal cost function parameterized by ξ_T^* that regularizes the terminal state.

The offline optimal trajectory OPT is obtained by solving (1) with the full knowledge of the true parameters $\xi_{0:T}^*$. In contrast, an online controller can only observe noisy estimations of the parameters in a fixed prediction horizon to decide its current action u_t at each time step t. The objective is to design an online controller that can compete against the offline optimal trajectory OPT. We use *dynamic regret* as the performance metric, which is widely used to evaluate the performance of online controllers/algorithms in the literature of online control [32, 34, 35] and online optimization [38, 43, 55]. Specifically, for a concrete problem instance $(x(0), \xi_{0:T}^*)$, let $\mathrm{cost}(\mathsf{OPT})$ denote the total cost incurred by OPT, and $\mathrm{cost}(\mathsf{ALG})$ denote the total cost incurred by an online controller ALG. The *dynamic regret* is defined as the worst-case additional cost incurred by ALG against OPT, i.e., $\sup_{x(0), \xi_{0:T}^*} (\mathrm{cost}(\mathsf{ALG}) - \mathrm{cost}(\mathsf{OPT}))$.

The formulation in (1) is general enough to include a variety of challenging settings. In this paper, we consider three important settings to illustrate how to apply our analysis pipeline. The settings differ in (a) the form of costs, dynamics, and constraints, and (b) the quantities in the system to be predicted (i.e., parameterized by ξ_t^*), and the prediction error allowed. An overview of the settings is presented in Table 1 below.

Table 1: Overview of the settings considered in this paper

Section	Costs	Dynamics	Constraints	Prediction $\boldsymbol{\xi_t}$	Prediction error
4.1	decomposable	LTV	none	disturbance: w_t	arbitrary
4.2	quadratic	LTV	none	cost: $Q_t, R_t, \bar{x}_t$ dynamics: A_t, B_t	sufficiently small
5	general	non-linear time-varying	non-linear stage constraint	cost: f_t dynamics: g_t constraints: s_t	sufficiently small

In each setting, we impose different assumptions on cost functions, dynamical systems, constraints, and properties of the predicted quantities as functions of parameter ξ_t. In general, we require well-defined costs, Lipschitz and uniformly controllable dynamics, and Lipschitzness of the predicted quantities with regard to ξ_t. For constraints, additional assumptions characterizing the active constraints along and near the optimal trajectory are imposed. Detailed definitions and statements are deferred to Appendix B and Sections 3, 4, and 5. To facilitate the statement of the pipeline, we assume the following *universal properties* hold throughout the paper:

- *Stability of* OPT*:* there exists a constant D_{x^*} such that $\|x_t^*\| \le D_{x^*}$ for every state x_t^* on the offline optimal trajectory OPT.
- *Lipschitz dynamics:* the ground-truth dynamical function $g_t(\cdot, \cdot; \xi_t^*)$ is Lipschitz in action; i.e., for any feasible x_t, u_t, u_t', g_t satisfies $\|g_t(x_t, u_t; \xi_t^*) - g_t(x_t, u_t'; \xi_t^*)\| \le L_g \|u_t - u_t'\|$.
- *Well-conditioned costs:* every stage cost $f_t(\cdot, \cdot; \xi_t^*)$ and the terminal cost $F_T(\cdot; \xi_T^*)$ are nonnegative, convex, and ℓ-smooth in (x_t, u_t) and x_T, respectively.

2.1 Predictive Online Control

While Step 3 (Lemma 3.2) in our pipeline can be generally applied to all online controllers, in the subsequent applications we focus on *Model Predictive Control (MPC)*, a popular classical controller. In this subsection, we first define the available information (predictions) as well as its quality (prediction power), and how general predictive online controllers make decisions. Then, we define a useful optimization problem called FTOCP, and introduce MPC as a predictive online controller.

We represent the uncertainties in cost functions, dynamics, constraints, and terminal costs as function families parameterized by ξ_t: $\mathcal{F}_t := \{f_t(x_t, u_t; \xi_t) \mid \xi_t \in \Xi_t\}, \mathcal{G}_t := \{g_t(x_t, u_t; \xi_t) \mid \xi_t \in \Xi_t\}$, $\mathcal{S}_t := \{s_t(x_t, u_t; \xi_t) \mid \xi_t \in \Xi_t\}$, and $\mathcal{F}_T := \{F_T(x_T; \xi_T) \mid \xi_T \in \Xi_T\}$. The online controller knows the function families $\mathcal{F}_{0:T}$, $\mathcal{G}_{0:T-1}$, and $\mathcal{S}_{0:T-1}$ as prior knowledge, but it does not know the true parameters $\xi^*_{0:T}$. Instead, at time step t, the online controller has access to noisy predictions of these parameters for the future k time steps (where k is called the *prediction horizon*), represented by $\xi_{t:t+k|t} \in \prod_{\tau=t}^{t+k} \Xi_\tau$. The parameter space Ξ_t at each time step t may have different dimensions.

We formally define the quality of predictions by introducing the following notion of prediction error.

Definition 2.1. *The prediction error is defined as* $\rho_{t,\tau} := \left\|\xi_{t+\tau|t} - \xi^*_{t+\tau}\right\|$ *for an integer* $\tau \geq 0$. *The power of* τ*-step-away predictions (for parameter* ξ*) is defined as* $P(\tau) := \sum_{t=0}^{T-\tau} \rho^2_{t,\tau}$.

Under this noisy prediction model, a general predictive online controller ALG decides the control action based on the current state and the latest available predictions of future parameters. We formally define the class of predictive online controllers considered in this paper in Definition 2.2, which includes MPC as a special case.

Definition 2.2. *A predictive online controller* ALG *is a function that takes the current state* x_t *and the available predictions* $\xi_{t:t+k|t}$ *as inputs at time t and outputs the current control action* u_t*, i.e.,* $u_t = \mathsf{ALG}(x_t, \xi_{t:t+k|t})$. *We use* $x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} \cdots \xrightarrow{u_{T-1}} u_T$ *to denote the trajectory achieved by* ALG*, and use* $x_0 \xrightarrow{u^*_0} x^*_1 \xrightarrow{u^*_1} \cdots \xrightarrow{u^*_{T-1}} u^*_T$ *to denote the offline optimal trajectory* OPT.

A core component of both the design of online controllers and our analysis is the following *finite-time optimal control problem* (FTOCP). Given a time interval $[t_1, t_2]$, the FTOCP solves the optimal sub-trajectory subjected to the given initial state z, terminal cost F, and a sequence of (potentially noisy) parameters $\xi_{t_1:t_2-1}, \zeta_{t_2}$, as formalized in the following definition.

Definition 2.3. *The finite-time optimal control problem (FTOCP) over the horizon* $[t_1, t_2]$*, with initial state z, parameters* $\xi_{t_1:t_2-1}$ *and* ζ_{t_2}*, and terminal cost* $F(\cdot;\cdot)$*, is defined as*

$$
\begin{aligned}
\iota_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \zeta_{t_2}; F) := \min_{y_{t_1:t_2}, v_{t_1:t_2-1}} & \sum_{t=t_1}^{t_2-1} f_t(y_t, v_t; \xi_t) + F(y_{t_2}; \zeta_{t_2}) \\
\textit{s.t. } & y_{t+1} = g_t(y_t, v_t; \xi_t), && \forall t_1 \leq t < t_2, \\
& s_t(y_t, v_t; \xi_t) \leq 0, && \forall t_1 \leq t < t_2, \\
& y_{t_1} = z,
\end{aligned}
\tag{2}
$$

and a corresponding optimal solution as $\psi_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \zeta_{t_2}; F)$. *We shall use the shorthand notation* $\psi_{t_1}^{t_2}(z, \xi_{t_1:t_2}; F) := \psi_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \xi_{t_2}; F)$ *when the context is clear.*

Note that the formulation of the FTOCP in Definition 2.3 does not include a terminal constraint set. To compensate for this, we allow the terminal cost $F(\cdot; \zeta_{t_2})$ to take value $+\infty$ in some subset of $\mathbb{R}^n$, and ζ_{t_2} is not necessarily an element in Ξ_{t_2}. For example, a terminal cost function that we frequently use later is the indicator function of the terminal parameter ζ_{t_2}, where $\zeta_{t_2} \in \mathbb{R}^n$. We use $\mathbb{I}$ to denote such indicator terminal cost (i.e., $\mathbb{I}(y_{t_2}; \zeta_{t_2}) = 0$ if $y_{t_2} = \zeta_{t_2}$ and $\mathbb{I}(y_{t_2}; \zeta_{t_2}) = +\infty$ otherwise).

Finally, given the definition of the FTOCP, we are ready to formally introduce MPC. The pseudocode of this online controller is given in Algorithm 1. Basically, at time step t, MPC_k solves a k-step predictive FTOCP using the latest available parameter predictions, and commits the first control action in the solution. When there are only fewer than k steps left, MPC_k directly solves a $(T-t)$-step FTOCP at time t until the end of the horizon, using the predicted real terminal cost $F_T(\cdot;\xi_{T|t})$. This MPC controller (and its variants) has a wide range of real-world applications.

Algorithm 1 Model Predictive Control (MPC_k)

Require: Specify the terminal costs F_t for $k \le t < T$.
1: **for** $t = 0, 1, \ldots, T-1$ **do**
2: $\quad t' \leftarrow \min\{t+k, T\}$
3: $\quad$ Observe current state x_t and obtain predictions $\xi_{t:t'|t}$.
4: $\quad$ Solve and commit control action $u_t := \psi_t^{t'}(x_t, \xi_{t:t'|t}; F_{t'})_{v_t}$.

3 The Pipeline: Bounded Regret via Perturbation Analysis

The goal of this section is to give an overview of a novel analysis pipeline that converts a perturbation bound into a bound on the dynamic regret. We begin by highlighting the form of perturbation bounds required in the pipeline, and then describe the 3-step process of applying the pipeline. In subsequent sections, we apply this pipeline to obtain new regret bounds for MPC in different settings.

3.1 Per-Step Error and Perturbation Bounds

A key challenge when comparing the performance of an online controller against the offline optimal trajectory is that the online controller's state x_t is different from the offline optimal state x_t^* at time step t. Due to such discrepancy in states, we cannot simply evaluate the online controller's action u_t via comparison against the offline optimal action u_t^*. To address this challenge, our pipeline uses the notion of per-step error (Definition 3.1) inspired by the performance difference lemma and its proofs in reinforcement learning (RL) [35]. Specifically, we compare u_t to the clairvoyant optimal action one may adopt at the same state x_t if all true future parameters $\xi_{t:T}^*$ are known, which leads to the definition of *per-step error* as follows.

Definition 3.1. *The per-step error e_t incurred by a predictive online controller* ALG *at time step t is defined as the distance between its actual action u_t and the clairvoyant optimal action, i.e.,*

$$e_t := \left\| u_t - \psi_t^T(x_t, \xi_{t:T}^*; F_T)_{v_t} \right\|, \text{ where } u_t = \mathsf{ALG}(x_t, \xi_{t:t+k|t}).$$

The clairvoyant optimal trajectory starting from x_t is defined as $x_{t:T|t}^* := \psi_t^T(x_t, \xi_{t:T}^*; F_T)_{y_{t:T}}$.

Note that the clairvoyant optimal trajectory can be viewed as being generated by an MPC controller with long enough prediction horizon and exact predictions. This notion highlights the reason why MPC can compete against the clairvoyant optimal trajectory, since the per-step error in a system controlled by MPC_k becomes $e_t = \left\| \psi_t^{t+k}(x_t, \xi_{t:t+k|t}; F_{t+k})_{v_t} - \psi_t^T(x_t, \xi_{t:T}^*; F_T)_{v_t} \right\|$. Intuitively, the per-step error converges to zero as the prediction horizon k increases and the quality of predictions improves (i.e. $\left\| \xi_{t:t+k|t} - \xi_{t:t+k}^* \right\| \to 0$).

This intuition highlights the important role of perturbation bounds in comparing online controllers against (offline) clairvoyant optimal trajectories. As we have discussed in Section 1, many previous works [36, 37, 48, 49] have established (local) decaying sensitivity/perturbation bounds for different instances of the FTOCP (2). These bounds may take different forms, but for the application of our pipeline we require two types of perturbation bounds that are both common in the literature:

(a) *Perturbations of the parameters $\xi_{t_1:t_2}$ given a fixed initial state z:*

$$\left\| \psi_{t_1}^{t_2}\left(z, \xi_{t_1:t_2}; F\right)_{v_{t_1}} - \psi_{t_1}^{t_2}\left(z, \xi'_{t_1:t_2}; F\right)_{v_{t_1}} \right\| \le \left(\sum_{t=t_1}^{t_2} q_1(t-t_1)\delta_t \right) \|z\| + \sum_{t=t_1}^{t_2} q_2(t-t_1)\delta_t, \tag{3}$$

where $\delta_t := \|\xi_t - \xi'_t\|$ for $t \in [t_1, t_2]$, and scalar functions q_1 and q_2 satisfy $\lim_{t\to\infty} q_i(t) = 0$, $\sum_{t=0}^{\infty} q_i(t) \le C_i$ for constants $C_i \ge 1, i = 1, 2$. This perturbation bound is useful in bounding the per-step error e_t, as we will discuss in Lemma 3.1.

(b) *Perturbation of the initial state z given fixed parameters $\xi_{t_1:t_2}$:*

$$\left\| \psi_{t_1}^{t_2}\left(z, \xi_{t_1:t_2}; F\right)_{y_t/v_t} - \psi_{t_1}^{t_2}\left(z', \xi_{t_1:t_2}; F\right)_{y_t/v_t} \right\| \le q_3(t - t_1) \left\| z - z' \right\|, \text{ for } t \in [t_1, t_2], \quad (4)$$

where the scalar function q_3 satisfies $\sum_{t=0}^{\infty} q_3(t) \le C_3$ for some constant $C_3 \ge 1$. This bound is useful in preventing the accumulation of per-step errors e_t throughout the horizon (see Lemma 3.2). Compared with (3), the right hand side of (4) has a simpler form.

Existing perturbation bounds usually combine the above two types ((3) and (4)) into a single equation that characterizes perturbations on z and $\xi_{t_1:t_2}$ simultaneously, e.g., [35, 37]. Here, we decompose them into two separate types because they are used in different parts of our pipeline.

3.2 A 3-Step Pipeline from Perturbation Bounds to Regret

An overview of the pipeline is given in Figure 1, which illustrates the high-level ideas of the pipeline that starts by obtaining perturbation bounds, proceeds to bound the per-step error using perturbation bounds, and finally combines the per-step error and perturbation bounds to bound the dynamic regret. In the following we describe each step in detail.

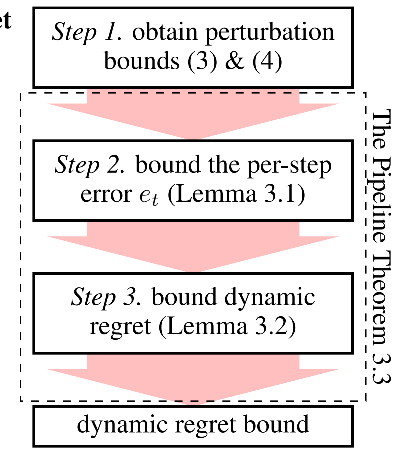

Figure 1: Illustrative diagram of the 3-step pipeline from perturbation analysis to bounded regret.

Step 1: Obtain the perturbation bounds given in (3) and (4). The form of the perturbation bounds depends heavily on the specific form of the FTOCP, and thus the derivation requires case-by-case study (e.g., see Section 4 and Section 5). However, off-the-shelf bounds are available in most cases, as there has been a rich literature on perturbation analysis of control systems (e.g., [35–37, 48, 49] and the references therein). The following property summarizes precisely what is expected to be derived for bounds (3) and (4) in Steps 2 and 3.

Property 3.1. *Suppose there exists a positive constant R such that the perturbation bound* (3) *holds for the following specifications: with $t_1 = t$ and $t_2 = t + k$ for $t < T - k$,* (3) *holds for*

$$z \in \mathcal{B}(x_t^*, R);\ \xi_{t:t+k-1} \in \Xi_{t:t+k-1}, \xi'_{t:t+k-1} = \xi^*_{t:t+k-1};\ \xi_{t+k}, \xi'_{t+k} \in \mathcal{B}(x^*_{t+k}, R);\ F = \mathbb{I};$$

with $t_1 = t$ and $t_2 = T$ for $t \ge T - k$, (3) *holds for $z \in \mathcal{B}(x_t^*, R);\ \xi_{t:T} \in \Xi_{t:T}, \xi_{t:T} = \xi^*_{t:T};\ F = F_T$. Further, perturbation bound* (4) *holds for any $z, z' \in \mathcal{B}(x_t^*, R)$.*

Recall that the indicator terminal cost function $\mathbb{I}$ is defined below Definition 2.3. Intuitively, Property 3.1 states that perturbation bounds (3) and (4) hold in a small neighborhood (specifically, a ball with radius R) around the offline optimal trajectory OPT, which is much weaker than the global exponentially decaying perturbation bounds required by previous work (e.g., [35]) in the following sense: (i) in the general settings where the dynamical function g_t is non-linear, or where there are constraints on states and actions, one cannot hope the perturbation bound to hold globally for all possible parameters [37, 49, 50]; (ii) the decay functions $\{q_i\}_{i=1,2,3}$ are only required to converge to zero and satisfy $\sum_{\tau=0}^{\infty} q_i(\tau) \le C_i$, which means the exponential decay rate as in [35] is not necessary — in fact, polynomial decay rates can also satisfy these properties, which greatly broadens the applicability of our pipeline.

Step 2: Bound the per-step error e_t. The core of the analysis is to apply the perturbation bounds to bound the per-step error. For MPC_k, under Property 3.1, this step can be done in a universal way, as summarized in Lemma 3.1 below. A complete proof of Lemma 3.1 can be found in Appendix C.

Lemma 3.1. *Let Property 3.1 hold. Suppose the current state x_t satisfies $x_t \in \mathcal{B}(x_t^*, R/C_3)$ and the terminal cost F_{t+k} of MPC_k is set to be the indicator function of some state $\bar{y}(\xi_{t+k|t})$ that satisfies $\bar{y}(\xi_{t+k|t}) \in \mathcal{B}(x^*_{t+k}, R)$ for $t < T - k$. Then, the per-step error of MPC_k is bounded by*

$$e_t \le \sum_{\tau=0}^{k} \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(\tau) + q_2(\tau) \right) \rho_{t,\tau} + 2R \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(k) + q_2(k) \right). \quad (5)$$

Lemma 3.1 is a straight-forward implication of perturbation bound (3) specified in Property 3.1. To see this, for $t < T - k$, note that the per-step error e_t can be bounded by

$$\begin{aligned} e_t &= \left\| \psi_t^{t+k}(x_t, \xi_{t:t+k-1|t}, \bar{y}(\xi_{t+k|t}); \mathbb{I})_{v_t} - \psi_t^T(x_t, \xi_{t:T}^*; F_T)_{v_t} \right\| && \text{(6a)} \\ &= \left\| \psi_t^{t+k}(x_t, \xi_{t:t+k-1|t}, \bar{y}(\xi_{t+k|t}); \mathbb{I})_{v_t} - \psi_t^{t+k}(x_t, \xi_{t:t+k-1}^*, x_{t+k|t}^*; \mathbb{I})_{v_t} \right\| && \text{(6b)} \\ &\leq \sum_{\tau=0}^{k-1} \left(\|x_t\| \cdot q_1(\tau) + q_2(\tau) \right) \rho_{t,\tau} + \left(\|x_t\| \cdot q_1(k) + q_2(k) \right) \left\| \bar{y}(\xi_{t+k|t}) - x_{t+k|t}^* \right\|. && \text{(6c)} \end{aligned}$$

Here, we apply the principle of optimality to conclude that the optimal trajectory from x_t to $x_{t+k|t}^*$ (i.e., $\psi_t^{t+k}(x_t, \xi_{t:t+k-1}^*, x_{t+k|t}^*; \mathbb{I})$ in (6b)) is a sub-trajectory of the clairvoyant optimal trajectory from x_t (i.e., $\psi_t^T(x_t, \xi_{t:T}^*; F_T)$ in (6a)), and (6c) is obtained by directly applying perturbation bound (3). Note that $\|x_t\| \leq \frac{R}{C_3} + D_{x^*}$, and that both $\bar{y}(\xi_{t+k|t})$ and $x_{t+k|t}^*$ are in $\mathcal{B}(x_{t+k}^*; R)$ by assumption and by perturbation bound (4) specified in Property 3.1, we conclude that (5) hold for $t < T - k$. The case $t \geq T - k$ can be shown similarly. We defer the detailed proof to Appendix C.

Step 3: Bound the dynamic regret by $\sum_{t=0}^{T-1} e_t^2$. This final step builds upon perturbation bound (4), and aims at deriving dynamic regret bounds in a universal way, as stated in Lemma 3.2 below. Specifically, under the assumption that a local decaying perturbation bound in the form of (4) holds around the offline optimal trajectory OPT, and the property that per-step errors e_t are sufficiently small, we can show that the online controller will not leave the "safe region" near the offline optimal trajectory as specified in Property 3.1, and thus the dynamic regret of ALG is bounded as in (7) (note that ALG is not confined to MPC, but is allowed to be any algorithm with bounded per-step errors). A complete proof of Lemma 3.2 can be found in Appendix D.

Lemma 3.2. *Let Property 3.1 hold. If the per-step errors of* ALG *satisfy* $e_\tau \leq R/(C_3^2 L_g)$ *for all time steps* $\tau < t$*, the trajectory of* ALG *will remain close to* OPT *at time* t*, i.e.* $x_t \in \mathcal{B}(x_t^*, R/C_3)$*. Further, if* $e_t \leq R/(C_3^2 L_g)$ *for all* $t < T$*, the dynamic regret of* ALG *is upper bounded by*

$$\mathrm{cost}(\mathsf{ALG}) - \mathrm{cost}(\mathsf{OPT}) = O\left(\sqrt{\mathrm{cost}(\mathsf{OPT}) \cdot \sum_{t=0}^{T-1} e_t^2} + \sum_{t=0}^{T-1} e_t^2 \right). \quad (7)$$

Summary. Combining Steps 2 and 3 of the pipeline yields the following *Pipeline Theorem* for MPC_k (see Theorem 3.3). Basically it states that, when the prediction horizon k is sufficiently large and the prediction errors $\rho_{t,\tau}$ are sufficiently small, Lemma 3.1 and Lemma 3.2 can work together to make sure that MPC_k never leaves a (R/C_3)-ball around the offline optimal trajectory OPT; thus we obtain a dynamic regret bound.

Theorem 3.3 (The Pipeline Theorem)**.** *Let Property 3.1 hold. Suppose the terminal cost* F_{t+k} *of* MPC_k *is set to be the indicator function of some state* $\bar{y}(\xi_{t+k|t})$ *that satisfies* $\bar{y}(\xi_{t+k|t}) \in \mathcal{B}(x_{t+k}^*, R)$ *for all time steps* $t < T - k$*. Further, suppose the prediction errors* $\rho_{t,\tau}$ *are sufficiently small and the prediction horizon* k *is sufficiently large, such that*

$$\sum_{\tau=0}^{k} \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(\tau) + q_2(\tau) \right) \rho_{t,\tau} + 2R \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(k) + q_2(k) \right) \leq \frac{R}{C_3^2 L_g}.$$

Then, the trajectory of MPC_k *will remain close to* OPT*, i.e.* $x_t \in \mathcal{B}(x_t^*, R/C_3)$ *for all time steps* t*, and the dynamic regret of* MPC_k *is upper bounded by*

$$\mathrm{cost}(\mathsf{MPC}_k) - \mathrm{cost}(\mathsf{OPT}) = O\left(\sqrt{\mathrm{cost}(\mathsf{OPT}) \cdot E} + E \right), \quad (8)$$

where $E := \sum_{\tau=0}^{k-1} (q_1(\tau) + q_2(\tau)) P(\tau) + \left(q_1(k)^2 + q_2(k)^2 \right) T$*.*

The proof of Theorem 3.3 can be found in Appendix E. To interpret the dynamic regret bound in (8), note that we have $\mathrm{cost}(\mathsf{OPT}) = O(T)$ as a result of our model assumptions. Thus, the dynamic regret of ALG is in the order of $\sqrt{TE} + E$. When there is no prediction error, the regret bound $O((q_1(k) + q_2(k)) \cdot T)$ reproduces the result in [35], and the bound will degrade as the prediction error increases. It is also worth noticing that, when the prediction power improves over time as the online controller learns the system better and $k = \Omega(\ln T)$, the dynamic regret can be $o(T)$.

4 Unconstrained LTV Systems

We now illustrate the use of the Pipeline Theorem by applying it in the context of (unconstrained) LTV systems with prediction errors, either on disturbances or the dynamical matrices.

4.1 Prediction Errors on Disturbances

In this section, we consider the following special case of problem (1), where the dynamics is LTV and the prediction error can only occur on the disturbances w_t:

$$\begin{aligned} \min_{x_{0:T}, u_{0:T-1}} & \sum_{t=0}^{T-1} \left(f_t^x(x_t) + f_t^u(u_t)\right) + F_T(x_T) \\ \text{s.t. } & x_{t+1} = A_t x_t + B_t u_t + w_t(\xi_t^*), \qquad \forall 0 \le t < T, \\ & x_0 = x(0). \end{aligned} \tag{9}$$

All necessary assumptions on the system are summarized below in Assumption 4.1.

Assumption 4.1. *Assume the following holds for the online control problem instance* (9)*:*

- *Cost functions: $\{f_t^x\}_{t=0}^{T-1}, \{f_t^u\}_{t=0}^{T-1}, F_T$ are nonnegative μ-strongly convex and ℓ-smooth. And we assume $f_t^x(0) = f_t^u(0) = F_T(0) = 0$ without the loss of generality.*
- *Dynamical systems: the LTV system $\{A_t, B_t\}$ is σ-uniform controllable with controllability index d, and $\|A_t\| \le a,\ \|B_t\| \le b$, and $\|B_t^\dagger\| \le b'$ hold for all t, where $B_t^\dagger$ denotes the Moore–Penrose inverse of matrix B_t.. The detailed definitions can be found in Assumption F.1 in Appendix F.*
- *Predicted quantities: $\|w_t(\xi_t)\| \le D_w$ holds for all $\xi_t \in \Xi_t$ and all t. For every time step t, $w_t(\xi_t)$ is a L_w-Lipschitz function in ξ_t, i.e., $\|w_t(\xi_t) - w_t(\xi_t')\| \le L_w \|\xi_t - \xi_t'\|, \forall \xi_t, \xi_t' \in \Xi_t$.*

Under Assumption 4.1, we can again apply the perturbation bounds shown in [35] to show Property 3.1. In particular, we already know that for some constants $H_1 \ge 1$ and $\lambda_1 \in (0,1)$, perturbation bounds (3) and (4) hold globally for $q_1(t) = 0$, $q_2(t) = H_1 \lambda_1^t$, and $q_3(t) = H_1 \lambda_1^t$. Since both of these perturbation bounds hold globally, radius R in Property 3.1 can be set arbitrarily, and we shall take $R \coloneqq \max\left\{D_{x^*}, \frac{2L_g H_1^3}{(1-\lambda_1)^3}\right\}$ so that Theorem 3.3 can be applied to MPC_k with terminal cost $F_{t+k}(\cdot; \xi_{t|t+k}) \equiv \mathbb{I}(\cdot; 0)$. This leads to the following dynamic regret bound:

Theorem 4.1. *In the unconstrained LTV setting* (9)*, under Assumption 4.1, when the prediction horizon k is sufficiently large such that $k \ge \ln\left(\frac{4H_1^3 L_g}{(1-\lambda_1)^2}\right) / \ln(1/\lambda_1)$, the dynamic regret of MPC_k (Algorithm 1) with terminal cost $F_{t+k}(\cdot; \xi_{t|t+k}) \equiv \mathbb{I}(\cdot; 0)$ is bounded by* $\mathrm{cost}(\mathsf{MPC}_k) - \mathrm{cost}(\mathsf{OPT}) \le O\left(\sqrt{T \cdot \sum_{\tau=0}^{k-1} \lambda_1^\tau P(\tau) + \lambda_1^{2k} T^2} + \sum_{\tau=0}^{k-1} \lambda_1^\tau P(\tau)\right)$.

A complete proof of Theorem 4.1 can be found in Appendix F. When there are no prediction errors, the bound in Theorem 4.1 reduces to $O(\lambda_1^k T)$, which reproduces the result of [35]. Further, it is also worth noticing that due to the form of discounted sum $\sum_{\tau=0}^{k-1} \lambda_1^\tau P(\tau)$, prediction errors for the near future matter more than those for the far future.

4.2 Prediction Error on Costs and Dynamical Matrices

We now consider prediction errors on cost functions and dynamics, rather than disturbances. Specifically, we consider the following instance of problem (1):

$$\begin{aligned} \min_{x_{0:T}, u_{0:T-1}} & \sum_{t=0}^{T-1} \left((x_t - \bar{x}_t(\xi_t^*))^\top Q_t(\xi_t^*)(x_t - \bar{x}_t(\xi_t^*)) + u_t^\top R_t(\xi_t^*) u_t\right) + F_T(x_T; \xi_t^*) \\ \text{s.t. } & x_{t+1} = A_t(\xi_t^*) \cdot x_t + B_t(\xi_t^*) \cdot u_t + w_t(\xi_t^*), \qquad \forall 0 \le t < T, \\ & x_0 = x(0), \end{aligned} \tag{10}$$

where the terminal cost is given by $F_T(x_T; \xi_T^*) \coloneqq (x_T - \bar{x}_T(\xi_T^*))^\top P_T(\xi_T^*)(x_T - \bar{x}(\xi_T^*))$.

All necessary assumptions on the system are summarized below in Assumption 4.2.

Assumption 4.2. *Assume the following holds for the online control problem instance* (10)*:*

- *Cost:* $\mu I \preceq Q_t(\xi_t) \preceq \ell I, \mu I \preceq R_t(\xi_t) \preceq \ell I$, *and* $\mu I \preceq P_T(\xi_T) \preceq \ell I$, $\forall \xi_t \in \Xi_t, \forall t$.
- *Dynamical systems: both the ground-truth LTV system* $\{A_t(\xi_t^*), B_t(\xi_t^*)\}_{t=0}^{T-1}$ *and any predicted LTV system* $\{A_t(\xi_{t+\tau|t}), B_t(\xi_{t+\tau|t})\}_{\tau=0}^{k-1}$ *(for all* $\xi_t \in \Xi_t$ *and all* t*) satisfy the controllability assumptions in Assumption G.1 in Appendix G.*
- *Predicted quantities: bounds* $\|w_t(\xi_t)\| \leq D_w, \|\bar{x}_t(\xi_t)\| \leq D_{\bar{x}}, \|A_t(\xi_t)\| \leq a, \|B_t(\xi_t)\| \leq b$ *hold for all* $\xi_t \in \Xi_t$ *and all* t. L_A *is a uniform Lipschitz constant such that* $\|A_t(\xi_t) - A_t(\xi_t')\| \leq L_A \|\xi_t - \xi_t'\|, \forall \xi_t, \xi_t' \in \Xi_t$ *holds for all* t*, and* $L_B, L_Q, L_R, L_{\bar{x}}, L_w$ *are defined similarly.*

Under Assumption 4.2, we can show that for some constants $H_2 \geq 1$ and $\lambda_2 \in (0, 1)$, perturbation bounds (3) and (4) hold globally for $q_1(t) = H_2\lambda_2^{2t}$, $q_2(t) = H_2\lambda_2^t$, and $q_3(t) = H_2\lambda_2^t$ under the specifications of Property 3.1. Thus, Property 3.1 holds for arbitrary R, and we can set $R = D_x^* + D_{\bar{x}}$ so that Theorem 3.3 can be applied to MPC_k with terminal cost $F_{t+k}(\cdot; \xi_{t|t+k}) = \mathbb{I}(\cdot; \bar{x}(\xi_{t|t+k}))$, which leads to the following dynamic regret bound:

Theorem 4.2. *In the unconstrained LTV setting* (10)*, under Assumption 4.2, when the prediction horizon* $k \geq O(1)$ *and the prediction errors satisfy* $\sum_{\tau=0}^{k} \lambda_2^{2\tau} \rho_{t,\tau} \leq \Omega(1)$*, the dynamic regret of* MPC_k *(Algorithm 1) with terminal cost* $F_{t+k}(\cdot; \xi_{t|t+k}) = \mathbb{I}(\cdot; \bar{x}(\xi_{t|t+k}))$ *is bounded by*

$$\mathrm{cost}(\mathsf{MPC}_k) - \mathrm{cost}(\mathsf{OPT}) \leq O\left(\sqrt{T \cdot \sum_{\tau=0}^{k-1} \lambda_2^\tau P(\tau) + \lambda_2^{2k} T^2} + \sum_{\tau=0}^{k-1} \lambda_2^\tau P(\tau)\right).$$

The exact constants and a complete proof of Theorem 4.2 can be found in Appendix G. Compared with Theorem 4.1, Theorem 4.2 additionally requires the discounted total prediction errors $\sum_{\tau=0}^{k} \lambda_2^{2\tau} \rho_{t,\tau}$ to be less than or equal to some constant. This is actually expected, and emphasizes the critical difference between the prediction errors on dynamical matrices (A_t, B_t) and the prediction errors on w_t, since an online controller cannot even stabilize the system when the predictions on (A_t, B_t) can be arbitrarily bad.

5 General Dynamical Systems

We now move beyond unconstrained linear systems to constrained nonlinear systems given by the general online control problem (1) in Section 2. All necessary assumptions are summarized in Assumption H.1 in Appendix H. Perhaps surprisingly, decaying perturbation bounds can hold even in this case. In particular, using Theorem 4.5 in [50], we can show that there exists a small constant R such that, for some constants $H_3 \geq 1$ and $\lambda_3 \in (0, 1)$, perturbation bounds (3) and (4) hold for $q_1(t) = 0$, $q_2(t) = H_3\lambda_3^t$, and $q_3(t) = H_3\lambda_3^t$. Thus, Property 3.1 holds (see Appendix H for formal statements) and we can apply Theorem 3.3 to obtain the following dynamic regret bound:

Theorem 5.1. *In the general system* (1)*, under Assumption H.1 in Appendix H, Property 3.1 holds for some positive constant* R *and* $q_1(t) = 0$, $q_2(t) = H_3\lambda_3^t$*, and* $q_3(t) = H_3\lambda_3^t$*. Suppose the terminal cost* F_{t+k} *of* MPC_k *is set to be the indicator function of some state* $\bar{y}(\xi_{t+k|t})$ *that satisfies* $\bar{y}(\xi_{t+k|t}) \in \mathcal{B}(x_{t+k}^*, R)$ *for* $t < T - k$*. Suppose the prediction errors* $\rho_{t,\tau}$ *are sufficiently small and the prediction horizon* k *is sufficiently large such that* $H_3 \sum_{\tau=0}^{k-1} \lambda_3^\tau \rho_{t,\tau} + 2RH_3\lambda_3^k \leq \frac{(1-\lambda_3)^2 R}{H_3^2 L_g}$*. Then, the dynamic regret of* MPC_k *is upper bounded by* $\mathrm{cost}(\mathsf{MPC}_k) - \mathrm{cost}(\mathsf{OPT}) \leq$
$$O\left(\sqrt{T \cdot \sum_{\tau=0}^{k-1} \lambda_3^\tau P(\tau) + \lambda_3^{2k} T^2} + \sum_{\tau=0}^{k-1} \lambda_3^\tau P(\tau)\right).$$

A complete proof of Theorem 5.1 can be found in Appendix H. An assumption in Theorem 5.1 that is difficult to satisfy in general is that the reference terminal states $\bar{y}(\xi_{t+k|t})$ of MPC_k must be close enough to the offline optimal state x_{t+k}^*, i.e., $\bar{y}(\xi_{t+k|t}) \in \mathcal{B}(x_{t+k}^*, R)$, while the offline optimal state x_{t+k}^* is generally unknown. This can be achieved in some special cases, for example, when we know $\|\xi_t^*\|$ is sufficiently small. In this case, one can first solve FTOCP $\psi_0^T(x_0, \mathbf{0}; F_T)$ and use it as a reference to set the terminal states of MPC_k. This intuition is formally shown in Appendix H. Another limitation is that Theorem 5.1 is only a bound on the cost of MPC, not its feasibility. There are many ways to guarantee recursive feasibility of MPC [53], which we leave as future work.

References

[1] Carlos E. Garcia, David M. Prett, and Manfred Morari. Model predictive control: theory and practice —- a survey. *Automatica*, 25(3):335–348, 1989.

[2] Ugo Rosolia and Francesco Borrelli. Learning model predictive control for iterative tasks: a data-driven control framework. *IEEE Transactions on Automatic Control*, 63(7):1883–1896, 2017.

[3] Milan Korda and Igor Mezić. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93:149–160, 2018.

[4] Frank Allgöwer and Alex Zheng. *Nonlinear model predictive control*, volume 26. Birkhäuser, 2012.

[5] Paolo Falcone, Manuela Tufo, Francesco Borrelli, Jahan Asgari, and H. Eric Tseng. A linear time varying model predictive control approach to the integrated vehicle dynamics control problem in autonomous systems. In *2007 46th IEEE Conference on Decision and Control*, pages 2980–2985. IEEE, 2007.

[6] Pierre-brice Wieber. Trajectory free linear model predictive control for stable walking in the presence of strong perturbations. In *2006 6th IEEE-RAS International Conference on Humanoid Robots*, pages 137–142, 2006.

[7] Dongbing Gu and Huosheng Hu. Receding horizon tracking control of wheeled mobile robots. *IEEE Transactions on Control Systems Technology*, 14(4):743–749, 2006.

[8] David H. Shim, H. Jin Kim, and Shankar Sastry. Decentralized nonlinear model predictive control of multiple flying robots. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, volume 4, pages 3621–3626, 2003.

[9] Holger Diedam, Dimitar Dimitrov, Pierre-Brice Wieber, Katja Mombaur, and Moritz Diehl. Online walking gait generation with adaptive foot positioning through linear model predictive control. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1121–1126, 2008.

[10] Michael Neunert, Cédric de Crousaz, Fadri Furrer, Mina Kamel, Farbod Farshidian, Roland Siegwart, and Jonas Buchli. Fast nonlinear model predictive control for unified trajectory optimization and tracking. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1398–1404, 2016.

[11] Daniel Morgan, Soon-Jo Chung, and Fred Y. Hadaegh. Model predictive control of swarms of spacecraft using sequential convex programming. *Journal of Guidance, Control, and Dynamics*, 37(6):1725–1740, 2014.

[12] Arthur Richards and Jonathan P. How. Robust variable horizon model predictive control for vehicle maneuvering. *International Journal of Robust and Nonlinear Control*, 16(7):333–351, 2006.

[13] Stefano Di Cairano, Diana Yanakiev, Alberto Bemporad, Ilya V. Kolmanovsky, and Davor Hrovat. An MPC design flow for automotive control and applications to idle speed regulation. In *2008 47th IEEE Conference on Decision and Control*, pages 5686–5691, 2008.

[14] Gregory Stewart and Francesco Borrelli. A model predictive control framework for industrial turbodiesel engine control. In *2008 47th IEEE Conference on Decision and Control*, pages 5704–5711, 2008.

[15] Rachid Amari, Mazen Alamir, and Paolino Tona. Unified MPC strategy for idle-speed control, vehicle start-up and gearing applied to an automated manual transmission. *IFAC Proceedings Volumes*, 41(2):7079–7085, 2008.

[16] Takeshi Hatanaka, Teruki Yamada, Masayuki Fujita, Shigeru Morimoto, and Masayuki Okamoto. *Explicit Receding Horizon Control of Automobiles with Continuously Variable Transmissions*, pages 561–569. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[17] Stefano Di Cairano, H. Eric Tseng, Daniele Bernardini, and Alberto Bemporad. Vehicle yaw stability control by coordinated active front steering and differential braking in the tire sideslip angles domain. *IEEE Transactions on Control Systems Technology*, 21(4):1236–1248, 2013.

[18] Patricia Hidalgo-Gonzalez, Rodrigo Henriquez-Auba, Duncan Callaway, and Claire Tomlin. Frequency regulation using data-driven controllers in power grids with variable inertia due to renewable energy. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5, 2019.

[19] Madasthu Santhosh, Chintham Venkaiah, and D. M. Vinod Kumar. Current advances and approaches in wind speed and wind power forecasting for improved renewable energy integration: A review. *Engineering Reports*, 2(6):e12178, 2020.

[20] Minghong Lin, Zhenhua Liu, Adam Wierman, and Lachlan L. H. Andrew. Online algorithms for geographical load balancing. In *Proceedings of the International Green Computing Conference (IGCC)*, pages 1–10, 2012.

[21] Xiaoyan Li, Chenghui Zhang, Alian Chen, Xiangyang Xing, and Guangxian Zhang. Model predictive direct current control strategy for three-level T-type rectifier under unbalanced grid voltage conditions. In *2018 IEEE Applied Power Electronics Conference and Exposition (APEC)*, pages 1514–1519, 2018.

[22] Ahmed A. Shetaya, Amr M. A. Amin, and Omar H. Abdalla. Model predictive control based power system operation planning of grid connected high share renewable energy. In *2017 Nineteenth International Middle East Power Systems Conference (MEPCON)*, pages 1500–1504, 2017.

[23] Alessandra Parisio, Evangelos Rikos, and Luigi Glielmo. A model predictive control approach to microgrid operation optimization. *IEEE Transactions on Control Systems Technology*, 22(5):1813–1827, 2014.

[24] Venkata Yaramasu and Bin Wu. Predictive control of a three-level boost converter and an NPC inverter for high-power PMSG-based medium voltage wind energy conversion systems. *IEEE Transactions on Power Electronics*, 29(10):5308–5322, 2014.

[25] Tong Wang, Huijun Gao, and Jianbin Qiu. A combined adaptive neural network and nonlinear model predictive control for multirate networked industrial process control. *IEEE Transactions on Neural Networks and Learning Systems*, 27(2):416–425, 2016.

[26] D. W. Clarke. Application of generalized predictive control to industrial processes. *IEEE Control Systems Magazine*, 8(2):49–55, 1988.

[27] Matthew Ellis, Helen Durand, and Panagiotis D. Christofides. A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24(8):1156–1178, 2014.

[28] Moritz Diehl, Rishi Amrit, and James B. Rawlings. A Lyapunov function for economic optimizing model predictive control. *IEEE Transactions on Automatic Control*, 56(3):703–707, 2010.

[29] David Angeli, Rishi Amrit, and James B. Rawlings. On average performance and stability of economic model predictive control. *IEEE transactions on automatic control*, 57(7):1615–1626, 2011.

[30] David Angeli, Alessandro Casavola, and Francesco Tedesco. Theoretical advances on economic model predictive control with time-varying costs. *Annual Reviews in Control*, 41:218–224, 2016.

[31] Lars Grüne and Simon Pirkelmann. Economic model predictive control for time-varying system: Performance and stability results. *Optimal Control Applications and Methods*, 41(1):42–64, 2020.

[32] Chenkai Yu, Guanya Shi, Soon-Jo Chung, Yisong Yue, and Adam Wierman. Competitive control with delayed imperfect information. *arXiv preprint arXiv:2010.11637*, 2020.

[33] Chenkai Yu, Guanya Shi, Soon-Jo Chung, Yisong Yue, and Adam Wierman. The power of predictions in online control. *Advances in Neural Information Processing Systems*, 33, 2020.

[34] Runyu Zhang, Yingying Li, and Na Li. On the regret analysis of online LQR control with predictions. *arXiv preprint arXiv:2102.01309*, 2021.

[35] Yiheng Lin, Yang Hu, Haoyuan Sun, Guanya Shi, Guannan Qu, and Adam Wierman. Perturbation-based regret analysis of predictive control in linear time varying systems. *arXiv preprint arXiv:2106.10497*, 2021.

[36] Sungho Shin, Victor M. Zavala, and Mihai Anitescu. Decentralized schemes with overlap for solving graph-structured optimization problems. *IEEE Transactions on Control of Network Systems*, 7(3):1225–1236, 2020.

[37] Sungho Shin and Victor M. Zavala. Controllability and observability imply exponential decay of sensitivity in dynamic optimization. *IFAC-PapersOnLine*, 54(6):179–184, 2021.

[38] Yingying Li, Guannan Qu, and Na Li. Online optimization with predictions and switching costs: Fast algorithms and the fundamental limit. *IEEE Transactions on Automatic Control*, 66(10):4761–4768, 2020.

[39] Yingying Li, Xin Chen, and Na Li. Online optimal control with linear dynamics and predictions: Algorithms and regret analysis. *Advances in Neural Information Processing Systems*, 32, 2019.

[40] Niangjun Chen, Anish Agarwal, Adam Wierman, Siddharth Barman, and Lachlan L. H. Andrew. Online convex optimization using predictions. *ACM SIGMETRICS Performance Evaluation Review*, 43(1):191–204, 2015.

[41] Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, and Karan Singh. Online control with adversarial disturbances. In *International Conference on Machine Learning*, pages 111–119. PMLR, 2019.

[42] Naman Agarwal, Elad Hazan, and Karan Singh. Logarithmic regret for online control. *Advances in Neural Information Processing Systems*, 32, 2019.

[43] Gautam Goel, Yiheng Lin, Haoyuan Sun, and Adam Wierman. Beyond online balanced descent: An optimal algorithm for smoothed online optimization. *Advances in Neural Information Processing Systems*, 32:1875–1885, 2019.

[44] Guanya Shi, Yiheng Lin, Soon-Jo Chung, Yisong Yue, and Adam Wierman. Online optimization with memory and competitive control. In *Advances in Neural Information Processing Systems*, pages 20636–20647, 2020.

[45] Weici Pan, Guanya Shi, Yiheng Lin, and Adam Wierman. Online optimization with feedback delay and nonlinear switching cost. *arXiv preprint arXiv:2111.00095*, 2021.

[46] Gautam Goel and Babak Hassibi. The power of linear controllers in LQR control. *Proceedings of Machine Learning Research vol TBD*, 1:13, 2020.

[47] Edgar Minasyan, Paula Gradu, Max Simchowitz, and Elad Hazan. Online control of unknown time-varying dynamical systems. *Advances in Neural Information Processing Systems*, 34, 2021.

[48] Wanting Xu and Mihai Anitescu. Exponentially convergent receding horizon strategy for constrained optimal control. *Vietnam Journal of Mathematics*, 47(4):897–929, 2019.

[49] Sen Na and Mihai Anitescu. Superconvergence of online optimization for model predictive control. *arXiv preprint arXiv:2001.03707*, 2020.

[50] Sungho Shin, Mihai Anitescu, and Victor M. Zavala. Exponential decay of sensitivity in graph-structured nonlinear programs. *arXiv preprint arXiv:2101.03067*, 2021.

[51] G. M. Lee, N. N. Tam, and N. D. Yen. Continuity of the solution map in quadratic programs under linear perturbations. *Journal of optimization theory and applications*, 129(3):415–423, 2006.

[52] Anthony V. Fiacco and Yo Ishizuka. Sensitivity and stability analysis for nonlinear programming. *Annals of Operations Research*, 27(1):215–235, 1990.

[53] Francesco Borrelli, Alberto Bemporad, and Manfred Morari. *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[54] Dimitris Papadimitriou, Ugo Rosolia, and Francesco Borrelli. Control of unknown nonlinear systems with linear time-varying MPC. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 2258–2263. IEEE, 2020.

[55] Yiheng Lin, Gautam Goel, and Adam Wierman. Online optimization with predictions and non-convex losses. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 4(1):1–32, 2020.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] See Contributions in Section 1 and the references within.
   (b) Did you describe the limitations of your work? [Yes] See the last paragraph of Section 5.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A] The goal of our work is to advance the theoretical understanding for predictive online controllers. We did not see any potential negative societal impacts of our work.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes] We did not find any aspect of our work that may lead to ethics violations.
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] See Assumption 4.1 and 4.2 in the main body, Appendix B, and Assumption F.1, G.1, and H.1 in the appendix.
   (b) Did you include complete proofs of all theoretical results? [Yes] See Appendices for complete proofs of our results. We also add references to the relevant sections of the Appendix when we introduced each theorem or lemma.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Notation Summary

In this paper, we use $\alpha_{t_1:t_2}$ ($t_2 \geq t_1$) to denote a sequence of vectors $(\alpha_{t_1}, \alpha_{t_1+1}, \ldots, \alpha_{t_2})$. For ease of reference, we summarize in the following table all the notations used in the paper.

Notation	Meaning
ξ_t	The uncertainty parameter of the system, used to parameterize costs, dynamics, and constraints.
ξ_t^*	The ground-truth parameter of the system, unknown to the controller.
$\xi_{\tau\|t}$	The prediction of ξ_τ^* revealed to the controller at time step t ($\tau \geq t$).
Ξ_t	The space of uncertainty parameters. ξ_t^* and $\xi_{t\|\tau}, \tau \leq t$ are in Ξ_t. We assume the diameter of Ξ_t is less than or equal to 1 without the loss of generality, i.e., $\\|\xi_t - \xi_t'\\| \leq 1$ for all $\xi_t, \xi_t' \in \Xi_t$.
k	The prediction horizon. At time t, the controller observes predictions $\xi_{t:t'\|t}$, where $t' := \min\{t+k, T\}$.
$\rho_{t,\tau}$	The error of predicting the system parameter after τ steps at time t, i.e., $\rho_{t,\tau} = \left\\|\xi_{t+\tau}^* - \xi_{t+\tau\|t}\right\\|$. We adopt the convention that $\rho_{t,\tau} := 0$ if $t + \tau > T$.
$P(\tau)$	The total error of predicting the system parameter after τ steps (the power of τ-step-away predictions), i.e., $P(\tau) := \sum_{t=0}^{T-\tau} \rho_{t,\tau}^2$.
$f_t(x_t, u_t; \xi_t)$	The stage cost of FTOCP at time step t, parameterized by $\xi_t \in \Xi_t$. The true stage cost is $f_t(x_t, u_t; \xi_t^*)$.
$g_t(x_t, u_t; \xi_t)$	The dynamical function at time step t, parameterized by $\xi_t \in \Xi_t$. The true dynamics is $x_{t+1} = g_t(x_t, u_t; \xi_t^*)$.
$s_t(x_t, u_t; \xi_t)$	The constraint function at time step t, parameterized by $\xi_t \in \Xi_t$. The true constraint is $s_t(x_t, u_t; \xi_t^*) \leq 0$.
F_T and $\{F_{t+k}\}_{t=0}^{T-k-1}$	F_T is the true terminal cost function defined by the original online control problem (1), while F_{t+k} for $t < T-k$ is the terminal cost function used by MPC_k at time t.
$\iota_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \zeta_{t_2}; F)$	The FTOCP defined on the time interval $[t_1, t_2]$, where z is the initial state at time t_1, and F is some terminal cost function at time t_2. $\xi_{t_1:t_2-1}$ are the parameters for the cost, dynamics, and constraints at time $[t_1, t_2-1]$, while ζ_{t_2} is the parameter for the terminal cost F.
$\psi_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \zeta_{t_2}; F)$	An optimal solution to the FTOCP $\iota_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \zeta_{t_2}; F)$. The entries are indexed by $y_{t_1:t_2}$ (for states) and $v_{t_1:t_2-1}$ (for actions).
$\psi_{t_1}^{t_2}(z, \xi_{t_1:t_2}; F)$	The shorthand notation of $\psi_{t_1}^{t_2}(z, \xi_{t_1:t_2-1}, \xi_{t_2}; F)$.

B Assumptions Overview

In this section, we give a more detailed overview of the assumptions that the online control problem (1) should satisfy in general so that our pipeline in Section 3.2 works. Specific assumptions in each specific setting will be presented separately in Assumption F.1, G.1, and H.1.

Cost functions. In general, we require the stage cost functions f_t and the terminal cost F_T to be *well-conditioned*, which includes non-negativity, strong convexity, smoothness (Lipschitz continuous gradient), and twice continuous differentiability. Note that these assumptions are equivalent to bounded Hessian ($\mu I \preceq \nabla^2 f_t \preceq \ell I$) and non-negative minimizer of the cost functions. Specifically, for quadratic costs $\nabla^2 f_t$ are constant, and the assumptions are further equivalent to bounded spectra of the cost matrices.

Dynamical systems. A basic requirement of the dynamical function g_t is *Lipschitzness* in u_t, i.e.,

$$\|g_t(x_t, u_t; \xi_t^*) - g_t(x_t, u_t'; \xi_t^*)\| \leq L_g \|u_t - u_t'\|.$$

We point out that only Lipschitzness in control action u_t is needed for the Pipeline Theorem to hold, which guarantees that an error on a control action u_t has a bounded impact on the next state x_{t+1}.

A more non-trivial assumption on dynamics is that the dynamical system should be *(uniformly) controllable*. Intuitively, this means the online controller should be able to steer the system to some target state in a finite number of time steps with some bounded control actions.

Definition B.1 (uniform controllability)**.** *Consider a general dynamics $x_{t+1} = g_t(x_t, u_t; \xi_t)$. For any time steps $t_2 \geq t_1$ and fixed (x_t, u_t), define $A_t := \nabla_{x_t}^\top g_t(x_t, u_t; \xi_t)$ and $B_t := \nabla_{u_t}^\top g_t(x_t, u_t; \xi_t)$, and we further define* ***transition matrix*** *$\Phi(t_2, t_1) \in \mathbb{R}^{n\times n}$ at (x_t, u_t) as*

$$\Phi(t_2, t_1) := \begin{cases} A_{t_2-1}A_{t_2-2}\cdots A_{t_1} & \textit{if } t_2 > t_1, \\ I & \textit{otherwise.} \end{cases}.$$

For any time t and time interval $p \geq 0$, define ***controllability matrix*** *$M(t, p; x_{t:t+p}, u_{t:t+p}) \in \mathbb{R}^{n\times(mp)}$ as*

$$M(t, p; x_{t:t+p}, u_{t:t+p}) := [\Phi(t + p, t + 1)B_t, \Phi(t + p, t + 2)B_{t+1}, \ldots, \Phi(t + p, t + p)B_{t+p}].$$

We say the system is ***controllable*** *if there exists a positive integer d, such that the controllability matrix $M(t, d; x_t, u_t)$ is of full row rank for any t and any (x_t, u_t). The smallest such constant d is called the* ***controllability index*** *of the system. Further, we say the system is* ***σ-uniformly controllable*** *if exists a positive constant σ such that $\sigma_{\min}(M(t, d)) \geq \sigma$ holds for all $t = 0, \ldots, T - d$.*

The definition has a clear control-theoretic interpretation for linear dynamics (where A_t and B_t are independent of (x_t, u_t)), but might seem trickier for non-linear dynamics (where A_t and B_t are functions of (x_t, u_t)). For the latter case, uniform controllability may be assumed for the offline optimal trajectory only, or for state-action pairs in a small neighborhood around it.

Constraints. Recursive feasibility is a well-known challenge for the design of online controllers in constrained systems [53]: at some time t, the controller may encounter an absence of feasible trajectories to continue from the current state x_t. Many solutions have been proposed for different controllers in a variety of systems. Since the purpose of this work is to establish dynamic regret guarantees for an online controller, and for the purpose of this paper, we would expect that there is a solution, potentially via a combination of proper controller design (e.g., setting the terminal cost/constraint of MPC) and some additional assumptions on the system (e.g., the SSOSC, strong second-order sufficient conditions, and LICQ, linear independent constraint qualification, which will be introduced in Section 5), so that we could focus on the sub-optimality of the online controller against the offline optimal trajectory.

We also need to point out that, although the additional assumptions on system that involve constraints might seem tricky, sometimes they are exactly the implications of previous assumptions on costs and dynamics that is actually needed in the proof. For example, Lemma 12 in [37] shows that Lipshitzness of dynamics and uniform controllability together imply uniform LICQ property of the system. For the clarity of exposition, these implications might be directly assumed in place of the low-level ones.

Parameter $\boldsymbol{\xi_t}$. In general, we require that all predicted quantities, which might include cost functions, dynamical functions, and constraints, should be *Lipschitz* in ξ_t, so that these quantities get closer to their ground truth value in a linearly-bounded way as the prediction error on the parameter ξ_t^* decreases. For a specific example of parameterized linear dynamics $x_{t+1} = A_t(\xi_t)x_t + B_t(\xi_t)u_t + w_t(\xi_t)$, the requirement is realized by assuming Lipschitzness of $A_t(\cdot), B_t(\cdot), w_t(\cdot)$ in ξ_t.

Offline optimal trajectory. We require the offline optimal trajectory OPT to be *stable*; i.e., there exists a constant D_{x^*} such that $\|x_t^*\| \leq D_{x^*}$ for any state x_t^* visited by OPT. While this can be shown under some assumptions in unconstrained LTV systems (see [35]), we introduce this assumption to simplify and unify the presentation for more complex systems.

C Proof of Lemma 3.1

We have already shown (5) holds for all time step $t < T - k$ in the main body. For $t \geq T - k$, we see that

$$e_t = \left\|\psi_t^T\left(x_t, \xi_{t:T|t}; F_T\right) - \psi_t^T\left(x_t, \xi_{t:T}^*; F_T\right)\right\| \tag{11a}$$

$$\le \sum_{\tau=0}^{k} \big(\|x_t\| \cdot q_1(\tau) + q_2(\tau)\big)\rho_{t,\tau} \tag{11b}$$

$$\le \sum_{\tau=0}^{k} \Big(\left(\frac{R}{C_3} + D_{x^*}\right) \cdot q_1(\tau) + q_2(\tau)\Big)\rho_{t,\tau}, \tag{11c}$$

where we used the definition of per-step error e_t in (11a); we used the perturbation bound (3) specified by Property 3.1 in (11b); we used the assumption $x_t \in \mathcal{B}\left(x_t^*, \frac{R}{C_3}\right)$, $\|x_t^*\| \le D_{x^*}$, and the convention $\rho_{t,\tau} := 0$ if $t + \tau > T$ in (11c). Thus e_t also satisfies (5) for $t \ge T - k$.

D Proof of Lemma 3.2

To simplify the notation, we will use $\psi_t^T(z)$ as a shorthand notation of $\psi_t^T(z, \xi_{t:T}^*; F_T)$ in the proof of Lemma 3.2, since the proof only relies on the perturbation bound (4).

Note that for any time step $t+1$, by Lipschitzness of the dynamics we have

$$\begin{aligned}\left\|x_{t+1} - \psi_t^T\left(x_t\right)_{y_{t+1}}\right\| &= \left\|g_t(x_t, u_t, w_t) - g_t\left(x_t, \psi_t^T(x_t)_{v_t}, w_t\right)\right\| \\ &\le L_g \left\|u_t - \psi_t^T(x_t)_{v_t}\right\| \\ &\le L_g e_t. \end{aligned} \tag{12}$$

Therefore, we can show the statement that $x_t \in \mathcal{B}\left(x_t^*, \frac{R}{C_3}\right)$ holds if $e_\tau \le R/(C_3^2 L_g), \forall \tau < t$ by induction. Note that this statement clearly holds for $t = 0$ since $x_0^* = x_0$. Suppose it holds for $0, 1, \ldots, t-1$. Then, we see that

$$\begin{aligned}\|x_t - x_t^*\| =&\ \left\|x_t - \psi_0^T(x_0)_{y_t}\right\| \\ \le&\ \left\|x_t - \psi_{t-1}^T(x_{t-1})_{y_t}\right\| + \sum_{i=1}^{t-1}\left\|\psi_{t-i}^T(x_{t-i})_{y_t} - \psi_{t-i-1}^T(x_{t-i-1})_{y_t}\right\| \\ \le&\ \left\|x_t - \psi_{t-1}^T(x_{t-1})_{y_t}\right\| + \sum_{i=1}^{t-1} q_3(i)\left\|x_{t-i} - \psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}}\right\| &\text{(13a)} \\ \le&\ \sum_{i=0}^{t-1} q_3(i)\left\|x_{t-i} - \psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}}\right\| &\text{(13b)} \\ \le&\ L_g \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}, &\text{(13c)}\end{aligned}$$

where in (13a), we apply the perturbation bound (4) specified by Property 3.1. To see why it can be applied, note that for $i \in [1, t-1]$, x_{t-i-1} satisfies $x_{t-i-1} \in \mathcal{B}\left(x_{t-i-1}^*, \frac{R}{C_3}\right)$ by the induction assumption, thus we have $\psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}} \in \mathcal{B}\left(x_{t-i}^*, R\right)$ because $q_3(1) \le \sum_{\tau=0}^{\infty} q_3(\tau) \le C_3$. Therefore, we can apply the perturbation bound (4) specified by Property 3.1 to compare the optimization solution vectors $\psi_{t-i}^T(x_{t-i})$ and $\psi_{t-i}^T\left(\psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}}\right)$, and by the principle of optimality, we see that

$$\psi_{t-i}^T\left(\psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}}\right)_{y_t} = \psi_{t-i-1}^T(x_{t-i-1})_{y_t}.$$

We also used $q_3(0) \ge 1$ in (13b) and (12) in (13c). Recall that we assume $e_{t-i} \le \frac{R}{C_3^2 L_g}$. Substituting this into (13) gives that

$$\|x_t - x_t^*\| \le L_g \cdot \frac{R}{C_3^2 L_g} \sum_{i=0}^{t-1} q_3(i) \le \frac{R}{C_3}.$$

Hence we have shown $x_t \in \mathcal{B}\left(x_t^*, \frac{R}{C_3}\right)$ holds if $e_\tau \le R/(C_3^2 L_g), \forall \tau < t$ by induction. An implication of this result is that $x_t \in \mathcal{B}\left(x_t^*, \frac{R}{C_3}\right)$ holds for all $t \le T$ if $e_t \le R/(C_3^2 L_g)$ holds for all $t < T$.

Similar with (13), we see the following inequality holds for all $t < T$ if $e_t \le R/(C_3^2 L_g), \forall t < T$:

$$\begin{aligned}
\|u_t - u_t^*\| &= \left\|u_t - \psi_0^T(x_0)_{v_t}\right\| \\
&\le \left\|u_t - \psi_t^T(x_t)_{v_t}\right\| + \sum_{i=0}^{t-1} \left\|\psi_{t-i}^T(x_{t-i})_{v_t} - \psi_{t-i-1}^T(x_{t-i-1})_{v_t}\right\| \\
&\le \left\|u_t - \psi_t^T(x_t)_{v_t}\right\| + \sum_{i=0}^{t-1} q_3(i) \left\|x_{t-i} - \psi_{t-i-1}^T(x_{t-i-1})_{y_{t-i}}\right\| \\
&\le e_t + L_g \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}, \qquad (14)
\end{aligned}$$

where the second inequality holds for the same reason as (13a).

By (13), we see that

$$\begin{aligned}
\|x_t - x_t^*\|^2 &\le L_g^2 \left(\sum_{i=0}^{t-1} q_3(i) e_{t-i-1}\right)^2 \\
&\le L_g^2 \left(\sum_{i=0}^{t-1} q_3(i)\right) \cdot \left(\sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right) \qquad (15a) \\
&\le C_3 L_g^2 \left(\sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right), \qquad (15b)
\end{aligned}$$

where we use the Cauchy-Schwarz inequality in (15a), and $\sum_{i=0}^{t-1} q_3(i) \le C_3$ in (15b).

Similarly, by (14), we see that

$$\begin{aligned}
\|u_t - u_t^*\|^2 &\le \left(e_t + L_g \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}\right)^2 \\
&\le \left(1 + L_g^2 \sum_{i=0}^{t-1} q_3(i)\right) \cdot \left(e_t^2 + \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right) \qquad (16a) \\
&\le \left(1 + C_3 L_g^2\right) \cdot \left(e_t^2 + \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right), \qquad (16b)
\end{aligned}$$

where we use the Cauchy-Schwarz inequality in (16a), and we use $\sum_{i=0}^{t-1} q_3(i) \le C_3$ in (16b).

Summing (15) and (16) over time steps t gives that

$$\begin{aligned}
&\sum_{t=1}^{T} \|x_t - x_t^*\|^2 + \sum_{t=0}^{T-1} \|u_t - u_t^*\|^2 \\
&\le C_3 L_g^2 \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right) + \left(1 + C_3 L_g^2\right) \cdot \sum_{t=0}^{T-1} \left(e_t^2 + \sum_{i=0}^{t-1} q_3(i) e_{t-i-1}^2\right) \\
&\le \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2, \qquad (17)
\end{aligned}$$

where we rearrange the terms and use $\sum_{j=0}^{\infty} q_3(j) \le C_3$ in the last inequality.

Since the cost function $f_t(\cdot, \cdot; \xi_t^*)$ and $F_T(\cdot; \xi_T^*)$ are nonnegative, convex, and ℓ-smooth in their inputs, by Lemma F.2 in [35], we see that the following inequality holds for arbitrary $\eta > 0$:

$$\begin{aligned}
&\mathrm{cost}(\mathsf{ALG}) - \mathrm{cost}(\mathsf{OPT}) \\
&\le \left(\sum_{t=0}^{T-1} f_t(x_t, u_t; \xi_t^*) + F_T(x_T; \xi_T^*)\right) - \left(\sum_{t=0}^{T-1} f_t(x_t^*, u_t^*; \xi_t^*) + F_T(x_T^*; \xi_T^*)\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \eta \left(\sum_{t=0}^{T-1} f_t(x_t^*, u_t^*; \xi_t^*) + F_T(x_T^*; \xi_T^*) \right) \\
&\quad + \frac{\ell}{2}\left(1 + \frac{1}{\eta}\right)\left(\sum_{t=1}^{T} \|x_t - x_t^*\|^2 + \sum_{t=0}^{T-1} \|u_t - u_t^*\|^2\right) \qquad (18a) \\
&\leq \eta \cdot \mathrm{cost}(\mathsf{OPT}) + \left(1 + \frac{1}{\eta}\right) \cdot \frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2 \qquad (18b) \\
&= \eta \cdot \mathrm{cost}(\mathsf{OPT}) + \frac{1}{\eta} \cdot \frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2 \\
&\quad + \frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2, \qquad (18c)
\end{aligned}
$$

where we apply Lemma F.2 in [35] in (18a), and we use (17) in (18b). Setting the tunable weight η in (18c) to be

$$
\eta = \left(\frac{\frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2}{\mathrm{cost}(\mathsf{OPT})} \right)^{\frac{1}{2}}
$$

gives that

$$
\begin{aligned}
&\mathrm{cost}(\mathsf{ALG}) - \mathrm{cost}(\mathsf{OPT}) \\
&\leq \sqrt{\left(\frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3)\right) \cdot \mathrm{cost}(\mathsf{OPT}) \cdot \sum_{t=0}^{T-1} e_t^2} \\
&\quad + \frac{\ell}{2} \cdot \left(1 + 2C_3 L_g^2\right) \cdot (1 + C_3) \cdot \sum_{t=0}^{T-1} e_t^2. \qquad (19)
\end{aligned}
$$

This finishes the proof of Lemma 3.2.

E Proof of Theorem 3.3

We first use induction to show that the following two conditions holds for all time steps $t < T$:

$$
\begin{aligned}
&x_t \in \mathcal{B}\left(x_t^*, \frac{R}{C_3}\right), \qquad (20a) \\
&e_t \leq \sum_{\tau=0}^{k} \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(\tau) + q_2(\tau) \right) \rho_{t,\tau} + 2R \left(\left(\frac{R}{C_3} + D_{x^*} \right) \cdot q_1(k) + q_2(k) \right). \qquad (20b)
\end{aligned}
$$

At time step 0, (20a) holds because $x_0 = x_0^*$, and (20b) holds by Lemma 3.1 and the assumption on the terminal cost F_k of MPC_k.

Suppose (20a) and (20b) hold for all time steps $\tau < t$. For time step t, by the assumption on the prediction errors $\rho_{t,\tau}$ and prediction horizon k in Theorem 3.3, we know that $e_\tau \leq \frac{R}{C_3^2 L_g}$ holds for all $\tau < t$ because (20b) holds for all $\tau < t$. Thus, we know that (20a) holds for time step t by Lemma 3.2. Then, since (20a) holds for time step t, and the terminal cost F_{t+k} of MPC_k is set to be the indicator function of some state $\bar{y}(\xi_{t+k|t})$ that satisfies $\bar{y}(\xi_{t+k|t}) \in \mathcal{B}(x_{t+k}^*, R)$ if $t < T - k$, we know (20b) also holds for time step t by Lemma 3.1. This finishes the induction proof of (20).

To simplify the notation, let $R_0 := \frac{R}{C_3} + D_{x^*}$. Note that (20b) implies that

$$
e_t^2 \leq \left(\sum_{\tau=0}^{k} \left(R_0 \cdot q_1(\tau) + q_2(\tau)\right) + 2R\left(R_0 + 1\right) \right)
$$

$$\cdot \left(\sum_{\tau=0}^{k} \left(R_0 \cdot q_1(\tau) + q_2(\tau)\right) \rho_{t,\tau}^2 + 2R \left(R_0 \cdot q_1(k)^2 + q_2(k)^2\right) \right) \tag{21a}$$

$$\leq \left(R_0 C_1 + C_2 + 2R(R_0+1)\right)$$

$$\cdot \left(\sum_{\tau=0}^{k-1} \left(R_0 \cdot q_1(\tau) + q_2(\tau)\right) \rho_{t,\tau}^2 + (2R+1) \left(R_0 \cdot q_1(k)^2 + q_2(k)^2\right) \right), \tag{21b}$$

where we use the Cauchy-Schwarz inequality in (21a); we use the bounds $\sum_{\tau=0}^{k} q_1(\tau) \leq C_1$, $\sum_{\tau=0}^{k} q_2(\tau) \leq C_2$, and $\rho_{t,\tau} \leq 1$ in (21b).

Since (20) and (21) holds for all time steps $t < T$, we can apply Lemma 3.2 to obtain that

$$\operatorname{cost}(\mathsf{MPC}_k) - \operatorname{cost}(\mathsf{OPT}) \leq \sqrt{\operatorname{cost}(\mathsf{OPT}) \cdot E_0} + E_0,$$

where

$$E_0 \coloneqq \left(R_0 C_1 + C_2 + 2R(R_0+1)\right) \cdot \left(\sum_{\tau=0}^{k-1} \left(R_0 \cdot q_1(\tau) + q_2(\tau)\right) P(\tau) + (2R+1) \left(R_0 \cdot q_1(k)^2 + q_2(k)^2\right) T \right).$$

This finishes the proof of Theorem 3.3.

F Assumptions and Proofs of Section 4.1

The formal definition of the controllability index d and σ-uniform controllable are given in [35]. For completeness, we restate them for LTV dynamics in Assumption F.1 below.

Assumption F.1. *For time steps $t_2 \geq t_1$, we define the transition matrix $\Phi(t_2, t_1) \in \mathbb{R}^{n\times n}$ as*

$$\Phi(t_2, t_1) \coloneqq \begin{cases} A_{t_2-1} A_{t_2-2} \cdots A_{t_1} & \text{if } t_2 > t_1 \\ I & \text{otherwise.} \end{cases},$$

For any positive integer p, we define the controllability matrix $M(t,p) \in \mathbb{R}^{n\times(mp)}$ as

$$M(t,p) \coloneqq \left[\Phi(t+p, t+1)B_t, \Phi(t+p, t+2)B_{t+1}, \ldots, \Phi(t+p, t+p)B_{t+p}\right].$$

We assume the LTV system $\{A_t, B_t\}$ is σ-uniform controllable with controllability index d, i.e., d is the smallest positive integer such that $\sigma_{min}\left(M(t,d)\right) > 0$ holds for all $t \in [0, T-d]$, and $\sigma_{min}\left(M(t,d)\right) \geq \sigma$ holds for all $t \in [0, T-d]$.

As a remark, the Assumption F.1 is a special case of Definition B.1 in unconstrained LTV systems. [35] has established a perturbation bound for the LTV system in (9) which implies the our requirements in Property 3.1. Thus we can use Theorem 3.3 to show Theorem 4.1.

Proof of Theorem 4.1. By Theorem 3.3 in [35], we know Property 3.1 holds under Assumption 4.1 for arbitrary R and $q_1(t) = 0, q_2(t) = H_1\lambda_1^t$, and $q_3(t) = H_1\lambda_1^t$, where $H_1 = H_1(\mu, \ell, d, \sigma, a, b, b', L_w) > 0$ is some constant, and $\lambda_1 = \lambda_1(\mu, \ell, d, \sigma, a, b, b', L_w) \in (0,1)$ is the decay rate. Here, H_1 corresponds to C and λ_1 corresponds to λ in Theorem 3.3 in [35].

By setting $R \coloneqq \max\left\{D_{x^*}, \frac{2L_g H_1^3}{(1-\lambda_1)^3}\right\}$, we guarantee that the terminal state 0 of MPC_k is always in the closed ball $\mathcal{B}(x_{t+k}^*, R)$, and the condition

$$\sum_{\tau=0}^{k} \left(\left(\frac{R}{C_3} + D_{x^*}\right) \cdot q_1(\tau) + q_2(\tau) \right) \rho_{t,\tau} + 2R \left(\left(\frac{R}{C_3} + D_{x^*}\right) \cdot q_1(k) + q_2(k) \right) \leq \frac{R}{C_3^2 L_g}$$

holds once $k \geq \ln\left(\frac{4H_1^3 L_g}{(1-\lambda_1)^2}\right) / \ln(1/\lambda_1)$ because $\rho_{t,\tau} \leq 1$. Therefore, we can apply Theorem 3.3 to finish the proof of Theorem 4.1. $\square$

G Assumptions and Proofs of Section 4.2

In this section, we give the detailed assumptions and proofs of the results in Section 4.2. Before we present the assumption on the uncertain LTV systems in (10), we first define several quantities that we will use heavily in the rest of this section:

For time steps $t_1 \leq t_2$ and $\xi_{t_1:t_2} \in \Xi_{t_1:t_2}$, define

$$N_{t_1}^{t_2}(\xi_{t_1:t_2}) \coloneqq \begin{bmatrix} I & & & & & \\ -A_{t_1}(\xi_{t_1}) & -B_{t_1}(\xi_{t_1}) & I & & & \\ & & \ddots & & & \\ & & & -A_{t_2}(\xi_{t_2}) & -B_{t_2}(\xi_{t_2}) & I \end{bmatrix}. \tag{22}$$

This matrix is closely related to the stability of the LTV system in the time interval $[t_1, t_2 + 1]$. To see this, note that $N_{t_1}^{t_2}(\xi_{t_1:t_2})$ always has full row rank, i.e., given any disturbance vector $w = (x_{t_1}, w_{t_1}, w_{t_1+1}, \ldots, w_{t_2})^\top$, one can always find a feasible sub-trajectory $z = (x_{t_1}, u_{t_1}, x_{t_1+1}, \ldots, u_{t_2}, x_{t_2+1})^\top$ that satisfies $N_{t_1}^{t_2}(\xi_{t_1:t_2})z = w$. If for any vector w, there exists a feasible sub-trajectory z such that $\|z\| \leq (1/\sigma) \cdot \|w\|$ for some positive constant σ, then the smallest singular value of $N_{t_1}^{t_2}(\xi_{t_1:t_2})$ is lower bounded by σ.

Similar with (22), we define matrix

$$\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2}) \coloneqq \begin{bmatrix} I & & & & \\ -A_{t_1}(\xi_{t_1}) & -B_{t_1}(\xi_{t_1}) & I & & \\ & & \ddots & & \\ & & & -A_{t_2}(\xi_{t_2}) & -B_{t_2}(\xi_{t_2}) \end{bmatrix} \tag{23}$$

for any time steps $t_1 \leq t_2$ and $\xi_{t_1:t_2} \in \Xi_{t_1:t_2}$, which removes the last column of (22). The matrix $\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2})$ is closely related to the controllability of the LTV system in the time interval $[t_1, t_2 + 1]$. To see this, given any disturbance vector $\hat{w} = (x_{t_1}, w_{t_1}, w_{t_1+1}, \ldots, w_{t_2} - x_{t_2+1})^\top$ whose first/last entry depends on the initial/terminal state, a feasible sub-trajectory $\hat{z} = (x_{t_1}, u_{t_1}, x_{t_1+1}, \ldots, u_{t_2})^\top$ must satisfy that $\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2})\hat{z} = \hat{w}$. Different from $N_{t_1}^{t_2}(\xi_{t_1:t_2})$, $\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2})$ is not guaranteed to have full row rank. If for any vector $\hat{w}$, there exists a feasible sub-trajectory $\hat{z}$ such that $\|\hat{z}\| \leq (1/\sigma) \cdot \|\hat{w}\|$ for some positive constant σ, then the smallest singular value of $\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2})$ is lower bounded by σ.

We make the following assumption on the smallest singular values of matrices $N_{t_1}^{t_2}(\xi_{t_1:t_2})$ and $\hat{N}_{t_1}^{t_2}(\xi_{t_1:t_2})$ so that the LTV system possesses uniform stability and controllability properties under any uncertainty parameters:

Assumption G.1. *There exists some universal constant $\sigma > 0$ such that $\sigma_{min}\left(N_t^{T-1}(\xi_{t:T-1})\right) \geq \sigma$ for any $t < T$, and $\sigma_{min}\left(\hat{N}_t^{t+k}(\xi_{t:t+k})\right) \geq \sigma$ for any $t < T - k$.*

While Assumption G.1 may seem more restricted than the uniform controllability defined in Definition B.1, it can actually be derived from Definition B.1 by Lemma 12 in [37].

In order to formulate (10) as a quadratic programming problem with equality constraints, we also need to define the matrix for cost functions:

$$\begin{aligned} M_t^T(\xi_{t:T}) &\coloneqq \operatorname{diag}\left(Q_t(\xi_t), R_t(\xi_t), Q_{t+1}(\xi_{t+1}), \ldots, R_{T-1}(\xi_{T-1}), P_T(\xi_T)\right), \forall t < T, \\ \hat{M}_t^{t+k}(\xi_{t:t+k}) &\coloneqq \operatorname{diag}\left(Q_t(\xi_t), R_t(\xi_t), Q_{t+1}(\xi_{t+1}), \ldots, R_{t+k-1}(\xi_{t+k-1})\right), \forall t < T - k. \end{aligned} \tag{24}$$

To write down the KKT condition of the equality constrained quadratic programming problem, we also need to define

$$\begin{aligned} H_t^T(\xi_{t:T}) &\coloneqq \begin{bmatrix} M_t^T(\xi_{t:T}) & N_t^{T-1}(\xi_{t:T-1})^\top \\ N_t^{T-1}(\xi_{t:T-1}) & 0 \end{bmatrix}, \\ \hat{H}_t^{t+k}(\xi_{t:t+k}) &\coloneqq \begin{bmatrix} \hat{M}_t^{t+k}(\xi_{t:t+k}) & \hat{N}_t^{t+k-1}(\xi_{t:t+k-1})^\top \\ \hat{N}_t^{t+k-1}(\xi_{t:t+k-1}) & 0 \end{bmatrix}, \end{aligned}$$

$$
\begin{aligned}
b_t^T(z,\xi_{t:T}) &\coloneqq \left(Q_t(\xi_t)\bar{x}_t(\xi_t), 0, \ldots, P(\xi_T)\bar{x}_T(\xi_T), z, w_t(\xi_t), \ldots, w_{T-1}(\xi_{T-1})\right)^\top, \\
\hat{b}_t^{t+k}(z,\xi_{t:t+k}) &\coloneqq \left(Q_t(\xi_t)\bar{x}_t(\xi_t), 0, \ldots, 0, z, w_t(\xi_t), \ldots, w_{t+k-1}(\xi_{t+k-1}) - \xi_{t+k}\right)^\top, \\
\chi_t^T &= \left(y_t, v_t, y_{t+1}, \ldots, v_{T-1}, y_T, \eta_t, \eta_{t+1}, \ldots, \eta_T\right)^\top, \\
\hat{\chi}_t^{t+k} &= \left(y_t, v_t, y_{t+1}, \ldots, v_{t+k-1}, \eta_t, \eta_{t+1}, \ldots, \eta_{t+k}\right)^\top.
\end{aligned}
$$

According to the KKT condition, the optimal primal-dual solution to $\iota_t^T(z,\xi_{t:T};F_T)$ ($t < T$) is the unique solution χ_t^T to the linear equation $H_t^T(\xi_{t:T})\chi_t^T = b_t^T(z,\xi_{t:T})$. Similarly, the optimal primal-dual solution to $\iota_t^{t+k}(z,\xi_{t:t+k};\mathbb{I})$ ($t < T-k$) is the unique solution χ_t^{t+k} to the linear equation $\hat{H}_t^{t+k}(\xi_{t:t+k})\hat{\chi}_t^{t+k} = \hat{b}_t^{t+k}(z,\xi_{t:t+k})$. We provide an illustrative example for χ_t^T with $(t,T) = (0,3)$ below:

$$
\left[\begin{array}{ccccccc|cccc}
Q_0 & & & & & & & I & -A_0^\top & & \\
& R_0 & & & & & & & -B_0^\top & & \\
& & Q_1 & & & & & & I & -A_1^\top & \\
& & & R_1 & & & & & & -B_1^\top & \\
& & & & Q_2 & & & & & I & -A_2^\top \\
& & & & & R_2 & & & & & -B_2^\top \\
& & & & & & P_3 & & & & I \\
\hline
I & & & & & & & & & & \\
-A_0 & -B_0 & I & & & & & & & & \\
& & -A_1 & -B_1 & I & & & & & & \\
& & & & -A_2 & -B_2 & I & & & &
\end{array}\right]
\left[\begin{array}{c}
y_0 \\ v_0 \\ y_1 \\ v_1 \\ y_2 \\ v_2 \\ y_3 \\ \hline \eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_3
\end{array}\right]
=
\left[\begin{array}{c}
Q_0\bar{x}_0 \\ 0 \\ Q_1\bar{x}_1 \\ 0 \\ Q_2\bar{x}_2 \\ 0 \\ P_3\bar{x}_3 \\ \hline z \\ w_0 \\ w_1 \\ w_2
\end{array}\right],
$$

where we omit the parameters $\xi_{0:3}$ to simplify the notations. Rearranging the rows and columns of the matrix on the left hand side gives the equation:

$$
\left[\begin{array}{ccc|ccc|ccc|cc}
Q_0 & & I & & & -A_0^\top & & & & & \\
& R_0 & & & & -B_0^\top & & & & & \\
I & & & & & & & & & & \\
\hline
& & & Q_1 & & I & & & -A_1^\top & & \\
& & & & R_1 & & & & -B_1^\top & & \\
-A_0 & -B_0 & & I & & & & & & & \\
\hline
& & & & & & Q_2 & & I & & -A_2^\top \\
& & & & & & & R_2 & & & -B_2^\top \\
& & & -A_1 & -B_1 & & I & & & & \\
\hline
& & & & & & & & & P & I \\
& & & & & & -A_2 & -B_2 & & I &
\end{array}\right]
\left[\begin{array}{c}
y_0 \\ v_0 \\ \eta_0 \\ \hline y_1 \\ v_1 \\ \eta_1 \\ \hline y_2 \\ v_2 \\ \eta_2 \\ \hline y_3 \\ \eta_3
\end{array}\right]
=
\left[\begin{array}{c}
Q_0\bar{x}_0 \\ 0 \\ z \\ \hline Q_1\bar{x}_1 \\ 0 \\ w_0 \\ \hline Q_2\bar{x}_2 \\ 0 \\ w_1 \\ \hline P\bar{x}_3 \\ w_2
\end{array}\right].
$$

Let Φ_t^T denote the permutation matrix that permute $(y_t, v_t, y_{t+1}, \ldots, v_{T-1}, y_T; \eta_t, \ldots, \eta_T)^\top$ to $(y_t, v_t, \eta_t, y_{t+1}, v_{t+1}, \eta_{t+1}, \ldots, y_T, \eta_T)^\top$. We use $\Upsilon_t^T(\xi_{t:T}) \coloneqq (\Phi_t^T)H_t^T(\xi_{t:T})(\Phi_t^T)^\top$ to denote the rearrangement of $H_t^T(\xi_{t:T})$ as illustrated in the above equation, and use $\beta_t^T(z,\xi_{t:T}) \coloneqq (\Phi_t^T)b_t^T(z,\xi_{t:T})$ to denote the corresponding rearrangement of $b_t^T(z,\xi_{t:T})$.

We also provide an illustrative example for $\hat{\chi}_t^{t+k}$ with $(t,k) = (0,3)$ below:

$$
\left[\begin{array}{cccccc|cccc}
Q_0 & & & & & & I & -A_0^\top & & \\
& R_0 & & & & & & -B_0^\top & & \\
& & Q_1 & & & & & I & -A_1^\top & \\
& & & R_1 & & & & & -B_1^\top & \\
& & & & Q_2 & & & & I & -A_2^\top \\
& & & & & R_2 & & & & -B_2^\top \\
\hline
I & & & & & & & & & \\
-A_0 & -B_0 & I & & & & & & & \\
& & -A_1 & -B_1 & I & & & & & \\
& & & & -A_2 & -B_2 & & & &
\end{array}\right]
\left[\begin{array}{c}
y_0 \\ v_0 \\ y_1 \\ v_1 \\ y_2 \\ v_2 \\ y_3 \\ \hline \eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_3
\end{array}\right]
=
\left[\begin{array}{c}
Q_0\bar{x}_0 \\ 0 \\ Q_1\bar{x}_1 \\ 0 \\ Q_2\bar{x}_2 \\ 0 \\ P_3\bar{x}_3 \\ \hline z \\ w_0 \\ w_1 \\ w_2
\end{array}\right],
$$

where we omit the parameters $\xi_{0:3}$ to simplify the notations. Rearranging the rows and columns of the matrix on the left hand side gives the equation:

$$
\left[\begin{array}{ccc|ccc|ccc|c}
Q_0 & & I & & & -A_0^\top & & & & \\
 & R_0 & & & & -B_0^\top & & & & \\
I & & & & & & & & & \\
\hline
 & & & Q_1 & & I & & & -A_1^\top & \\
 & & & & R_1 & & & & -B_1^\top & \\
-A_0 & -B_0 & & I & & & & & & \\
\hline
 & & & & & & Q_2 & & I & -A_2^\top \\
 & & & & & & & R_2 & & -B_2^\top \\
 & & & -A_1 & -B_1 & & I & & & \\
\hline
 & & & & & & -A_2 & -B_2 & &
\end{array}\right]
\left[\begin{array}{c}
y_0 \\ v_0 \\ \eta_0 \\ \hline y_1 \\ v_1 \\ \eta_1 \\ \hline y_2 \\ v_2 \\ \eta_2 \\ \hline \eta_3
\end{array}\right]
=
\left[\begin{array}{c}
Q_0\bar{x}_0 \\ 0 \\ z_0 \\ \hline Q_1\bar{x}_1 \\ 0 \\ w_0 \\ \hline Q_2\bar{x}_2 \\ 0 \\ w_1 \\ \hline w_2 - z_3
\end{array}\right].
$$

Let $\hat{\Phi}_t^{t+k}$ denote the permutation matrix that permute $(y_t, v_t, y_{t+1}, \ldots, v_{t+k-1}; \eta_t, \ldots, \eta_{t+k})^\top$ to $(y_t, v_t, \eta_t, y_{t+1}, v_{t+1}, \eta_{t+1}, \ldots, y_{t+k-1}, v_{t+k-1}, \eta_{t+k-1}, \eta_{t+k})^\top$. We use $\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k}) := (\hat{\Phi}_t^{t+k})\hat{H}_t^{t+k}(\xi_{t:t+k})(\hat{\Phi}_t^{t+k})^\top$ to denote the rearrangement of $\hat{H}_t^{t+k}(\xi_{t:t+k})$ as illustrated in the above equation, and use $\hat{\beta}_t^{t+k}(z, \xi_{t:t+k}) := (\hat{\Phi}_t^{t+k})\hat{b}_t^{t+k}(z, \xi_{t:t+k})$ to denote the corresponding rearrangement of $\hat{b}_t^{t+k}(z, \xi_{t:t+k})$.

Before showing the main result about the per-step error, we first show a technical lemma about the singular values of a block matrix in Lemma G.1.

Lemma G.1. *Consider a block matrix*

$$H = \begin{bmatrix} M & N^\top \\ N & 0 \end{bmatrix}.$$

Here $M \in \mathbb{R}^{n_0 \times n_0}$ is a symmetric positive definite matrix that satisfies $\underline{\sigma}_M I \preceq M \leq \overline{\sigma}_M I$ with $\underline{\sigma}_M > 0$, and $N \in \mathbb{R}^{n_1 \times n_0}$ with $n_1 \leq n_0$ satisfies that $\underline{\sigma}_N \leq \sigma(N) \leq \overline{\sigma}_N$ with $\underline{\sigma}_N > 0$. Then H satisfies that

$$\min(\underline{\sigma}_M, 1) \cdot \overline{\sigma}_N \cdot \sqrt{\frac{\overline{\sigma}_M}{2\underline{\sigma}_M\overline{\sigma}_M + \underline{\sigma}_M(\underline{\sigma}_N)^2}} \leq \sigma(H) \leq \sqrt{2}(\overline{\sigma}_M + \overline{\sigma}_N).$$

Proof of Lemma G.1. We first establish the lower bound on $\sigma(H)$: Suppose the singular value decomposition of $NM^{-\frac{1}{2}}$ is given by

$$NM^{-\frac{1}{2}} = U\Sigma V^\top,$$

where $U \in \mathbb{R}^{n_1 \times n_1}$ and $V \in \mathbb{R}^{n_0 \times n_0}$ are unitary matrices and $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_{n_1}) \in \mathbb{R}^{n_1 \times n_0}$ with $\frac{\overline{\sigma}_N}{\sqrt{\underline{\sigma}_M}} \geq \sigma_1 \geq \cdots \geq \sigma_m \geq \frac{\underline{\sigma}_N}{\sqrt{\overline{\sigma}_M}}$. We can decompose H as

$$\begin{aligned}
H &= \begin{bmatrix} M^{\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} I & M^{-\frac{1}{2}}F^\top \\ FM^{-\frac{1}{2}} & 0 \end{bmatrix} \cdot \begin{bmatrix} M^{\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} M^{\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \cdot \begin{bmatrix} I & \Sigma^\top \\ \Sigma & 0 \end{bmatrix} \cdot \begin{bmatrix} V^\top & 0 \\ 0 & U^\top \end{bmatrix} \cdot \begin{bmatrix} M^{\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix}. \qquad (25)
\end{aligned}$$

Note that for any $\alpha \in \mathbb{R}^{n_0}$ and $\beta \in \mathbb{R}^{n_1}$, we have

$$\begin{aligned}
\left\| \begin{bmatrix} I & \Sigma^\top \\ \Sigma & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|^2 &= \sum_{i=1}^{n_1} (\alpha_i + \sigma_i\beta_i)^2 + \sum_{i=n_1+1}^{n_0} \alpha_i^2 + \sum_{i=1}^{n_1} \sigma_i^2 \alpha_i^2 \\
&= \sum_{i=1}^{n_1} \left(\left(1 + \frac{\sigma_i^2}{2}\right) \left(\alpha_i + \frac{2\sigma_i}{2+\sigma_i^2}\beta_i\right)^2 + \frac{\sigma_i^2}{2}\alpha_i^2 + \frac{\sigma_i^2}{2+\sigma_i^2}\beta_i^2 \right) + \sum_{i=n_1+1}^{n_0} \alpha_i^2 \\
&\geq \left(\min_i \frac{\sigma_i^2}{2+\sigma_i^2}\right) \left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|^2 \\
&\geq \frac{\underline{\sigma}_M(\underline{\sigma}_N)^2}{2\underline{\sigma}_M\overline{\sigma}_M + \overline{\sigma}_M(\overline{\sigma}_N)^2} \left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|^2.
\end{aligned}$$

Therefore, by (25), we see that

$$\left\|H\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\right\| \geq \min(\underline{\sigma}_M, 1)\cdot\underline{\sigma}_N\cdot\sqrt{\frac{\underline{\sigma}_M}{2\underline{\sigma}_M\overline{\sigma}_M+\overline{\sigma}_M(\overline{\sigma}_N)^2}}\cdot\left\|\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\right\|.$$

This finishes the proof of the lower bound.

For the upper bound, note that

$$\begin{aligned}\left\|H\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\right\| &\leq \left\|M\alpha+N^\top\beta\right\|+\|N\alpha\|\\ &\leq \|M\alpha\|+\left\|N^\top\beta\right\|+\|N\alpha\|\\ &\leq (\overline{\sigma}_M+\overline{\sigma}_N)(\|x\|+\|y\|)\\ &\leq \sqrt{2}(\overline{\sigma}_M+\overline{\sigma}_N)\left\|\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\right\|.\end{aligned}$$

□

Since the primal-dual optimal solution to $\iota_t^T(z,\xi_{t:T};F_T)$ and $\iota_t^{t+k}(z,\xi_{t:t+k};\mathbb{I})$ are given by $\left(\Upsilon_t^T(\xi_{t:T})\right)^{-1}\beta_t^T(z,\xi_{t:T})$ and $\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})\right)^{-1}\hat{\beta}_t^{t+k}(z,\xi_{t:t+k})$ respectively, it is critical to establish the exponentially decaying bounds for the matrices $\left(\Upsilon_t^T(\xi_{t:T})\right)^{-1}$ and $\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})\right)^{-1}$. Note that after the rearrangement, $\Upsilon_t^T(\xi_{t:T})$ is a block matrix with $(T-t+1)\times(T-t+1)$ blocks, indexed by $(i,j)\in[t,T]^2$; $\Upsilon_t^{t+k}(\xi_{t:t+k})$ is a block matrix with $(k+1)\times(k+1)$ blocks, indexed by $(i,j)\in[t,t+k]^2$.

Lemma G.2. *Under Assumption 4.2, the following inequalities hold for the norm of the block entries of $\left(\Upsilon_t^T(\xi_{t:T})\right)^{-1}$ and $\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})\right)^{-1}$:*

$$\begin{aligned}\left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{ij}\right\| &\leq C_2\lambda_2^{|i-j|},\forall(i,j)\in[t,T]^2,\forall\xi_{t:T}\in\Xi_{t:T},\\ \left\|\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1}\right)_{ij}\right\| &\leq C_2\lambda_2^{|i-j|},\forall(i,j)\in[t,t+k]^2,\forall\xi_{t:t+k}\in\Xi_{t:t+k}.\end{aligned}\tag{26}$$

Further, the following inequalities hold for the norm of differences between the block entries of $\left(\Upsilon_t^T(\xi_{t:T})\right)^{-1}$ and $\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})\right)^{-1}$: For all $(i,j)\in[t,T]^2$ and $\xi_{t:T}\in\Xi_{t:T}$, we have

$$\left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}-\Upsilon_t^T(\xi'_{t:T})^{-1}\right)_{ij}\right\|\leq C_2'\sum_{\tau=t}^{T}\lambda_2^{|\tau-i|+|\tau-j|}\cdot\|\xi_\tau-\xi'_\tau\|,\tag{27}$$

For all $(i,j)\in[t,t+k]^2$ and $\xi_{t:T}\in\Xi_{t:t+k}$ with $t<T-k$, we have

$$\left\|\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1}-\hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1}\right)_{ij}\right\|\leq C_2'\sum_{\tau=t}^{t+k}\lambda_2^{|\tau-i|+|\tau-j|}\cdot\|\xi_\tau-\xi'_\tau\|,\tag{28}$$

where the constants C_2, C_2', and λ_2 are given by

$$\begin{aligned}\lambda_2&=\left(\frac{\overline{\sigma}_H-\underline{\sigma}_H}{\overline{\sigma}_H+\underline{\sigma}_H}\right)^{\frac{1}{2}},C_2=\frac{4(\ell+1+a+b)}{\underline{\sigma}_H^2\cdot\lambda_2},\\ C_2'&=C_2^2\left(\max\{L_Q+L_R,L_P\}+\frac{2}{\lambda_2}(L_A+L_B)\right).\end{aligned}$$

where $\underline{\sigma}_H$ and $\overline{\sigma}_H$ are defined as

$$\underline{\sigma}_H:=\min(\mu,1)\cdot(a+b+1)\cdot\sqrt{\frac{\ell}{2\mu\ell+\mu\sigma^2}},\ \textit{and}\ \overline{\sigma}_H:=\sqrt{2}(\ell+a+b+1).$$

Proof of Lemma G.2. In the proof, we only show the results for Υ_t^T. The results for $\hat{\Upsilon}_t^{t+k}$ can be shown using the same method.

We first show (26) holds. By Lemma G.1, we know that Note that $\Upsilon_t^T(\xi_{t:T})^2$ is a positive definite matrix that has band width 4 and satisfies

$$\underline{\sigma}_H^2 I \preceq \Upsilon_t^T(\xi_{t:T})^2 \preceq \overline{\sigma}_H^2 I.$$

Using the same method as the proof of Lemma B.1 in [35], one can show that for any $(i,j) \in [t,T]^2$,

$$\left(\left(\Upsilon_t^T(\xi_{t:T})^2\right)^{-1}\right)_{ij} \leq \frac{2}{\underline{\sigma}_H^2} \cdot \lambda_2^{|i-j|}. \tag{29}$$

Note that $\Upsilon_t^T(\xi_{t:T})^{-1} := \Upsilon_t^T(\xi_{t:T}) \cdot \left(\Upsilon_t^T(\xi_{t:T})^2\right)^{-1}$. Thus we see that

$$\begin{aligned}
\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{ij} &= \left(\Upsilon_t^T(\xi_{t:T}) \cdot \left(\Upsilon_t^T(\xi_{t:T})^2\right)^{-1}\right)_{ij} \\
&= \sum_{k=t}^{T} \Upsilon_t^T(\xi_{t:T})_{ik} \cdot \left(\left(\Upsilon_t^T(\xi_{t:T})^2\right)^{-1}\right)_{kj} \\
&= \sum_{k=i-1}^{i+1} \Upsilon_t^T(\xi_{t:T})_{ik} \cdot \left(\left(\Upsilon_t^T(\xi_{t:T})^2\right)^{-1}\right)_{kj}.
\end{aligned}$$

Therefore, by (29), we see that

$$\left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{ij}\right\| \leq \frac{4(\ell+1+a+b)}{\underline{\sigma}_H^2 \cdot \lambda_2} \cdot \lambda_2^{|i-j|}, \forall (i,j) \in [t,T]^2. \tag{30}$$

Note that we have

$$\Upsilon_t^T(\xi_{t:T})^{-1} - \Upsilon_t^T(\xi'_{t:T})^{-1} = -(\Upsilon_t^T(\xi'_{t:T}))^{-1}\left(\Upsilon_t^T(\xi_{t:T}) - \Upsilon_t^T(\xi'_{t:T})\right)\Upsilon_t^T(\xi_{t:T})^{-1}. \tag{31}$$

To simplify the notation, we define

$$E_\tau := \left(\Upsilon_t^T(\xi_{t:T}) - \Upsilon_t^T(\xi'_{t:T})\right)_{\tau\tau}, \forall \tau \in [t,T],$$

and

$$E'_\tau := \left(\Upsilon_t^T(\xi_{t:T}) - \Upsilon_t^T(\xi'_{t:T})\right)_{(\tau+1)\tau}, \forall \tau \in [t,T-1].$$

The right hand side of (31) is a linear equation. Thus we can study the following 3 equations separately and sum them up:

$$\Phi_\tau := (\Upsilon_t^T(\xi'_{t:T}))^{-1} \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & E_\tau & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} (\Upsilon_t^T(\xi_{t:T}))^{-1}, \forall \tau \in [t,T], \tag{32a}$$

$$\Phi_\tau^L := (\Upsilon_t^T(\xi'_{t:T}))^{-1} \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & E'_\tau & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} (\Upsilon_t^T(\xi_{t:T}))^{-1}, \forall \tau \in [t,T-1], \tag{32b}$$

$$\Phi_\tau^U := (\Upsilon_t^T(\xi'_{t:T}))^{-1} \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & (E'_\tau)^\top & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} (\Upsilon_t^T(\xi_{t:T}))^{-1}, \forall \tau \in [t,T-1]. \tag{32c}$$

By (32a), (32b), and (32c), we see that

$$\begin{aligned}\|(\Phi_\tau)_{ij}\| &\le \left\|\left((\Upsilon_t^T(\xi'_{t:T}))^{-1}\right)_{i\tau}\right\| \cdot \|E_\tau\| \cdot \left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{\tau j}\right\| \\ &\le C_2^2 \cdot \lambda_2^{|\tau-i|+|\tau-j|} \cdot \|E_\tau\|, \quad (33a)\\ \left\|(\Phi_\tau^L)_{ij}\right\| &\le \left\|\left((\Upsilon_t^T(\xi'_{t:T}))^{-1}\right)_{i\tau}\right\| \cdot \|E'_\tau\| \cdot \left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{(\tau+1)j}\right\| \\ &\le C_2^2 \cdot \lambda_2^{|\tau-i|+|\tau+1-j|} \cdot \|E'_\tau\|, \quad (33b)\\ \left\|(\Phi_\tau^U)_{ij}\right\| &\le \left\|\left((\Upsilon_t^T(\xi'_{t:T}))^{-1}\right)_{i(\tau+1)}\right\| \cdot \|E'_\tau\| \cdot \left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1}\right)_{\tau j}\right\| \\ &\le C_2^2 \cdot \lambda_2^{|\tau+1-i|+|\tau-j|} \cdot \|E'_\tau\|, \quad (33c)\end{aligned}$$

where we use the bound (30) on the norm of individual block entries in (33a), (33b), and (33c). Summing these inequalities up over τ, we see that

$$\begin{aligned}&\left\|\left(\Upsilon_t^T(\xi_{t:T})^{-1} - \Upsilon_t^T(\xi'_{t:T})^{-1}\right)_{ij}\right\| \\ =\ &\left\|\sum_{\tau=t}^{T}(\Phi_\tau)_{ij} + \sum_{\tau=t}^{T-1}(\Phi_\tau^L)_{ij} + \sum_{\tau=t}^{T-1}(\Phi_\tau^U)_{ij}\right\| \quad (34a)\\ \le\ &\sum_{\tau=t}^{T}\|(\Phi_\tau)_{ij}\| + \sum_{\tau=t}^{T-1}\left\|(\Phi_\tau^L)_{ij}\right\| + \sum_{\tau=t}^{T-1}\left\|(\Phi_\tau^U)_{ij}\right\| \quad (34b)\\ \le\ &C_2^2\left(\sum_{\tau=t}^{T}\lambda_2^{|\tau-i|+|\tau-j|} \cdot \|E_\tau\| + \frac{2}{\lambda_2}\sum_{\tau=t}^{T-1}\lambda_2^{|\tau-i|+|\tau-j|} \cdot \|E'_\tau\|\right) \quad (34c)\\ \le\ &C_2^2\left(\max\{L_Q+L_R, L_P\} + \frac{2}{\lambda_2}(L_A+L_B)\right)\sum_{\tau=t}^{T}\lambda_2^{|\tau-i|+|\tau-j|} \cdot \|\xi_\tau - \xi'_\tau\|, \quad (34d)\end{aligned}$$

where we use (31) in (34a); we use the triangle inequality in (34b); we use (33) in (34c); we use the Lipschitzness of dynamical and cost matrices in ξ (Assumption 4.2) in (34d). □

With Lemma G.2, we can derive the perturbation bounds specified by Property 3.1.

Theorem G.3. *Under Assumption 4.2, Property 3.1 holds for arbitrary positive constant R and $q_1(t) = H_2\lambda_2^{2t}$, $q_2(t) = H_2\lambda_2^t$, and $q_3(t) = H_2\lambda_2^t$, where λ_2 is defined in Lemma G.2, and H_2 is given by*

$$H_2 = C'_2\left(\frac{2(\ell D_{\bar{x}} + D_w)}{1-\lambda_2} + R + D_{x^*} + 1\right) + C_2\left(L_w + \ell L_{\bar{x}} + D_{\bar{x}}L_Q + 1\right).$$

Proof of Theorem G.3. For $t < T-k$, under the specification of Property 3.1, we see that

$$\begin{aligned}&\left\|\psi_t^{t+k}\left(z, \xi_{t:t+k}; \mathbb{I}\right)_{v_t} - \psi_t^{t+k}\left(z, \xi'_{t:t+k}; \mathbb{I}\right)_{v_t}\right\| \\ \le\ &\left\|\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1}\hat{\beta}_t^{t+k}(z, \xi_{t:t+k}) - \hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1}\hat{\beta}_t^{t+k}(z, \xi'_{t:t+k})\right)_{v_t}\right\| \quad (35a)\\ \le\ &\left\|\left(\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1} - \hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1}\right)\hat{\beta}_t^{t+k}(z, \xi_{t:t+k})\right)_{v_t}\right\| \\ &+ \left\|\left(\hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1}\left(\hat{\beta}_t^{t+k}(z, \xi_{t:t+k}) - \hat{\beta}_t^{t+k}(z, \xi'_{t:t+k})\right)\right)_{v_t}\right\|, \quad (35b)\end{aligned}$$

where we used the KKT condition in (35a) and the triangle inequality in (35b).

For the first term in (35b), we see that

$$\left\|\left(\left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1} - \hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1}\right) \cdot \hat{\beta}_t^{t+k}(z, \xi_{t:t+k})\right)_{v_t}\right\|$$

$$\leq \sum_{\tau=t}^{t+k} \left\| \left(\hat{\Upsilon}_t^{t+k}(\xi_{t:t+k})^{-1} - \hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1} \right)_{t\tau} \right\| \cdot \left\| \hat{\beta}_t^{t+k}(z, \xi_{t:t+k})_\tau \right\| \tag{36a}$$

$$\leq C_2' \left(\sum_{\tau=0}^{k} \lambda_2^{2\tau} \left\| \xi_{t+\tau} - \xi'_{t+\tau} \right\| \right) \cdot \|z\| + \sum_{\tau=t}^{t+k} C_2' \left(\sum_{i=t}^{t+k} \lambda_2^{i-t+|i-\tau|} \left\| \xi_i - \xi'_i \right\| \right) \cdot (\ell D_{\bar{x}} + D_w)$$

$$+ C_2' \lambda_2^k \cdot \left(\sum_{\tau=0}^{k} \left\| \xi_{t+\tau} - \xi'_{t+\tau} \right\| \right) \cdot \|\xi_{t+k}\| \tag{36b}$$

$$\leq C_2' \sum_{\tau=0}^{k} \lambda_2^{2\tau} \delta_{t+\tau} \cdot \|z\| + C_2' \left(\frac{2(\ell D_{\bar{x}} + D_w)}{1 - \lambda_2} + R + D_{x^*} \right) \sum_{\tau=0}^{k} \lambda_2^{\tau} \delta_{t+\tau}, \tag{36c}$$

where we use the triangle inequality in (36a); we use Lemma G.2 and the bounds on each entry of $\hat{\beta}_t^{t+k}(z, \xi_{t:t+k})$ in (36b); we rearrange the terms and use $\xi_{t+k} \in \mathcal{B}(x^*_{t+k}, R)$ in (36c). For the second error term (35b), we see that

$$\left\| \left(\hat{\Upsilon}_t^{t+k}(\xi'_{t:t+k})^{-1} \left(\hat{\beta}_t^{t+k}(z, \xi_{t:t+k}) - \hat{\beta}_t^{t+k}(z, \xi'_{t:t+k}) \right) \right)_{v_t} \right\|$$

$$\leq C_2 \sum_{\tau=t}^{t+k} \lambda_2^{\tau-t} (L_w + \ell L_{\bar{x}} + D_{\bar{x}} L_Q) \delta_\tau + C_2 \lambda_2^k \delta_{t+k}, \tag{37}$$

where we use the following inequality to bound the difference between $\hat{\beta}_t^{t+k}(z, \xi_{t:t+k})$ and $\hat{\beta}_t^{t+k}(z, \xi'_{t:t+k})$:

$$\begin{aligned}
&\|Q_\tau(\xi_\tau)\bar{x}_\tau(\xi_\tau) - Q_\tau(\xi'_\tau)\bar{x}_\tau(\xi'_\tau)\| \\
\leq\ & \|Q_\tau(\xi_\tau)\bar{x}_\tau(\xi_\tau) - Q_\tau(\xi'_\tau)\bar{x}_\tau(\xi_\tau)\| + \|Q_\tau(\xi'_\tau)\bar{x}_\tau(\xi_\tau) - Q_\tau(\xi'_\tau)\bar{x}_\tau(\xi'_\tau)\| \\
\leq\ & \|Q_\tau(\xi_\tau) - Q_\tau(\xi'_\tau)\| \cdot \|\bar{x}_\tau(\xi_\tau)\| + \|Q_\tau(\xi'_\tau)\| \cdot \|\bar{x}_\tau(\xi_\tau) - \bar{x}_\tau(\xi'_\tau)\| \\
\leq\ & (L_Q D_{\bar{x}} + \ell L_{\bar{x}})\delta_\tau.
\end{aligned}$$

Substituting (36) and (37) into (35) gives that for any $t < T - k$,

$$\left\| \psi_t^{t+k}\left(z, \xi_{t:t+k}; \mathbb{I}\right)_{v_t} - \psi_t^{t+k}\left(z, \xi'_{t:t+k}; \mathbb{I}\right)_{v_t} \right\| \leq \left(\sum_{\tau=0}^{k} q_1(\tau) \delta_{t+\tau} \right) \|z\| + \sum_{\tau=0}^{k} q_2(\tau) \delta_{t+\tau}$$

under the specification that $\xi_{t:t+k-1} \in \Xi_{t:t+k-1}, \xi'_{t:t+k-1} = \xi^*_{t:t+k-1};\ \xi_{t+k}, \xi'_{t+k} \in \mathcal{B}(x^*_{t+k}, R)$. We can use a similar methods to show that for any $t \geq T - k$,

$$\left\| \psi_t^T\left(z, \xi_{t:T}; F_T\right)_{v_t} - \psi_t^T\left(z, \xi'_{t:T}; F_T\right)_{v_t} \right\| \leq \left(\sum_{\tau=0}^{T-t} q_1(\tau) \delta_{t+\tau} \right) \|z\| + \sum_{\tau=0}^{T-t} q_2(\tau) \delta_{t+\tau}$$

under the specification that $\xi_{t:T} \in \Xi_{t:T}, \xi'_{t:T} = \xi^*_{t:T}$.

For any $t < T$, we see that

$$\begin{aligned}
&\left\| \psi_t^T(z, \xi^*_{t:T}; F_T)_{y_\tau/v_\tau} - \psi_t^T(z', \xi^*_{t:T}; F_T)_{y_\tau/v_\tau} \right\| \\
\leq\ & \left\| \left(\Upsilon_t^T(\xi^*_{t:T})^{-1} \left(\beta_t^T(z, \xi^*_{t:T}) - \beta_t^T(z', \xi^*_{t:T}) \right) \right)_{y_\tau/v_\tau} \right\| && \text{(38a)} \\
\leq\ & \left\| \left(\Upsilon_t^T(\xi^*_{t:T})^{-1} \right)_{\tau t} \right\| \cdot \|z - z'\| \\
\leq\ & C_2 \lambda_2^{\tau-t} \|z - z'\|, && \text{(38b)}
\end{aligned}$$

where we use the KKT condition in (38a); we use Lemma G.2 in (38b). □

Now we come back to the proof of Theorem 4.2.

Proof of Theorem 4.2. By Theorem G.3, Property 3.1 holds for arbitrary positive constant R and $q_1(t) = H_2 \lambda_2^{2t}$, $q_2(t) = H_2 \lambda_2^t$, and $q_3(t) = H_2 \lambda_2^t$, where the decay rate $\lambda_2 \in (0, 1)$ and constant

H_2 depends on R. We set $R := D_{x^*} + D_{\bar{x}}$ so that MPC_k with terminal state $\bar{x}_{t+k}(\xi_{t+k|t})$ satisfies the assumption of Theorem 3.3. The constant H_2 is given by

$$H_2 = C_2' \left(\frac{2(\ell D_{\bar{x}} + D_w)}{1 - \lambda_2} + 2D_{x^*} + D_{\bar{x}} + 1 \right) + C_2 \left(L_w + \ell L_{\bar{x}} + D_{\bar{x}} L_Q + 1 \right).$$

By Theorem 3.3, in order to achieve the claimed dynamic regret bound in Theorem 4.2, a sufficient condition is that the prediction errors $\rho_{t,\tau}$ satisfy

$$\sum_{\tau=0}^{k} \lambda_2^{\tau} \rho_{t,\tau} \leq \frac{(1-\lambda_2)^2 (D_{x^*} + D_{\bar{x}})}{2H_2^2 L_g \left((1-\lambda_2)(D_{x^*} + D_{\bar{x}}) + H_2(D_{x^*} + 1)\right)},$$

and the prediction horizon k satisfies that

$$\lambda_2^k \leq \frac{(1-\lambda_2)^2}{4H_2^2 L_g \left((1-\lambda_2)(D_{x^*} + D_{\bar{x}}) + H_2(D_{x^*} + 1)\right)}.$$

□

H Assumptions and Proofs of Section 5

To introduce the SSOSC assumption, we first define the *reduced Hessian* of the Lagrangian.

Definition H.1 (reduced Hessian)**.** *For a constrained optimization problem with primal variable z and dual variable η, let $H = \nabla_{zz}^2 \mathcal{L}$ denote the Hessian of the Lagrangian $\mathcal{L}(z, \eta; \xi)$. Let G denote the **active constraints Jacobian**, i.e. Jacobian of all equality constraints and active inequality constraints, and let Z be the null-space matrix of G (i.e., the column vectors of Z form an orthonormal basis of the null space of G). Then the **reduced Hessian** is defined as $H_{\mathrm{re}}(z, \eta; \xi) := Z^\top H Z$.*

We define the concept of *singular spectrum bounds* for a specific instance of FTOCP:

Definition H.2 (singular spectrum bounds)**.** *Consider the FTOCP $\iota_{t_1}^{t_2}(z, \xi_{t_1:t_2}; F)$. The positive real numbers $\overline{\sigma}_H, \overline{\sigma}_R, \underline{\sigma}_H$ are called **singular spectrum bounds** for this specific instance of FTOCP*[1] *if they satisfy that*

$$\overline{\sigma}_H \geq \overline{\overline{\sigma}}_H(z, \xi_{t_1:t_2}), \overline{\sigma}_R \geq \overline{\overline{\sigma}}_R(z, \xi_{t_1:t_2}), \text{ and } 0 < \underline{\sigma}_H \leq \underline{\underline{\sigma}}_H(z, \xi_{t_1:t_2}),$$

where $\overline{\overline{\sigma}}_H, \overline{\overline{\sigma}}_R$, and $\underline{\underline{\sigma}}_H$ are defined in (4.16a-c) in [50].

Assumption H.1. *We make the following assumptions on the costs, dynamics, and constraints of an FTOCP $\iota_{t_1}^{t_2}(z, \xi_{t_1:t_2}; F)$:*

1. *All cost functions, dynamical functions, and constraint functions are twice continuously differentiable in (x_t, u_t) and ξ_t* [2].
2. *(SSOSC) The reduced Hessian at the optimal primal-dual solution is positive-definite.*
3. *(LICQ) The active constraints Jacobian G at the optimal primal-dual solution has full row rank, i.e. $\sigma_{\min}(G) > 0$.*
4. *(Uniform singular spectrum bounds) There exist positive singular spectrum bounds $\overline{\sigma}_H, \overline{\sigma}_R, \underline{\sigma}_H$ for all FTOCP specifications below:*
   (a) *$t_1 = t, t_2 = t + k$ for $t < T - k$:*
   $$z \in \mathcal{B}(x_t^*, R), \xi_{t:t+k-1} \in \Xi_{t:t+k-1}, \xi_{t+k} \in \mathcal{B}(x_{t+k}^*, R), F = \mathbb{I}.$$
   (b) *$t_1 = t, t_2 = T$ for $t < T$:*
   $$z \in \mathcal{B}(x_t^*, R), \xi_{t:T} \in \Xi_{t:T}, F = F_T.$$

[1] We remind the reader that the functions $\overline{\overline{\sigma}}_H, \overline{\overline{\sigma}}_R$, and $\underline{\underline{\sigma}}_H$ depend on the form of FTOCP, i.e., different horizon $[t_1, t_2]$ and different terminal cost function F.

[2] If the terminal function F is an indicator function of some state, we view it as a constraint instead of cost.

We remind the readers that Lemma 12 in [37] shows that Lipshitzness of dynamics and uniform controllability together imply uniform LICQ property of the system.

Under Assumption H.1, we know Property 3.1 holds for $q_1(t) = 0$, $q_2(t) = H_3 \lambda_3^t$, and $q_3(t) = H_3 \lambda_3^t$ for some $H_3 > 0$ and $\lambda_3 \in (0, 1)$ by Theorem 4.5 in [50].

Theorem H.1. *Under Assumption H.1 that holds for some $R > 0$, Property 3.1 holds for $q_1(t) = 0$, $q_2(t) = H_3 \lambda_3^t$, and $q_3(t) = H_3 \lambda_3^t$ with the same R. The coefficient H_3 and decay factor λ_3 are given by*

$$H_3 := \left(\frac{\overline{\sigma}_H \overline{\sigma}_R}{\underline{\sigma}_H^2} \right)^{\frac{1}{2}}, \text{ and } \lambda_3 := \left(\frac{\overline{\sigma}_H^2 - \underline{\sigma}_H^2}{\overline{\sigma}_H^2 + \underline{\sigma}_H^2} \right)^{\frac{1}{8}}.$$

Combining Theorem H.1 with the Pipeline Theorem (Theorem 3.3) finishes the proof of Theorem 5.1. Note that when $\xi_t^* \leq \frac{(1-\lambda_3)R}{H_3}$, we know that $\left\| \psi_0^T \left(x_0, \mathbf{0}; F_T\right)_{y_{t+k}} - x_{t+k}^* \right\| \leq R$. Thus using $\psi_0^T \left(x_0, \mathbf{0}; F_T\right)_{y_{t+k}}$ as the terminal state of MPC_k at time step t can satisfy the requirement of Theorem 5.1.