Learning-Based Model Predictive Control

Master Thesis

Author(s):
Ho, Kin Wa

Publication date:
2024-06-17

Permanent link:
https://doi.org/10.3929/ethz-b-000704865

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Automatic Control Laboratory

Master Thesis

Learning-Based Model Predictive Control

Kin Wa Ho
June 17, 2024

Advisors
Dr. Raffaele Soloperto
Dr. Anastasios Tsiamis
Prof. Dr. John Lygeros

Abstract

In this thesis, we investigate the convergence behavior of the dual adaptive model predictive controller (MPC). The dual adaptive MPC is a compelling area of research due to its ability to actively explore unknown parameters within the system. The current convergence result using Lyapunov stability analysis depends on two practical challenges: the difference between the current and the previous estimated parameter and the covariance matrix derived from the recursive least squares (RLS) estimation. Our research aims to achieve convergence independent of these factors. By applying logarithmic upper bounds to both the average squared Euclidean distance between consecutive parameter estimates and the terms in the average Lyapunov decrease function, we establish the average convergence of the dual adaptive MPC. We validate our results with detailed mathematical proofs and simulations, which also illustrate limitations. We contribute valuable insights into the convergence behavior of the dual adaptive MPC and provide some assurance of its operational reliability in practical applications.

Contents

List of Figures

Chapter 1

Introduction

Motivation: Adaptive control has been an attractive area of research in control theory due to its ability to adjust the controller based on real-time data. A typical adaptive control system includes a component that processes measured data using an estimation algorithm to refine unknown parameter estimates. These updated estimates are then fed into the controller, which computes the optimal tracking input for the plant. However, this architecture only explores the system's unknowns passively because there are no components actively encouraging the exploration of the system since the controller solely optimizes the tracking path.

Related Works: [1] addresses this by integrating exploration and exploitation into a single optimal control problem where the optimized control inputs meet both criteria simultaneously. Initially, the practical implementation was challenging due to computational limitations, but advancements in hardware and the development of efficient solvers have revitalized interest in this approach. Dual adaptive MPC, which integrates dual theory with model predictive control (MPC), has gained significant attention. MPC is favored in various industries due to its ability to handle constraints and its straightforward implementation with current tools.
In the work of [2], a stochastic approach to dual adaptive MPC for single-input single-output (SISO) systems was presented and converted into a deterministic quadratic constrained quadratic programming (QCQP) problem. This formulation enables the dual adaptive MPC to be solved using well-established solvers. However, stability proofs for dual adaptive MPC remained challenging and were not provided.
[3] presented a stability analysis where the dual adaptive MPC was considered within the reinforcement learning framework. The cost function was treated as a Bellman equation to explain stability. However, they did not address tracking, which is a common application of MPC.
A stability analysis of dual adaptive MPC for tracking was provided in [4] using Lyapunov stability analysis. They achieved convergence that depends on the difference between consecutive parameter estimates and the current covariance matrix. This dependence introduces practical challenges that could impact the operational safety of dual adaptive MPC.

Thesis Goal: In this thesis, we aim to enhance the convergence aspect compared to the results presented in [4]. We consider a discrete linear time-invariant system with a scalar output disturbed by additive bounded noise and an unknown parameter vector. The unknown parameter vector is estimated using recursive least squares (RLS) estimation. To improve the estimation accuracy, we project the RLS estimate onto a predefined set, that the true parameter vector is assumed to be contained within this set.
In contrast to the approach proposed in [4], which utilizes set-membership estimation to dynamically update the predefined set, our method maintains a constant predefined set throughout the

algorithm. This approach may lead to conservative constraints, but it simplifies the mathematical proof of stability. The objective of the dual adaptive MPC is to steer the output towards the desired value while concurrently refining the estimate of the unknown parameter.
In typical MPC, the control law ensures that the system remains within a positive invariant set, and a terminal controller is employed to guide the system towards the origin. We introduce a terminal equality constraint to define a positive invariant set that includes only a single point, thereby simplifying the mathematical proof. Our investigation reveals that, with these modifications and additional assumptions, both the squared Euclidean norm of the difference between consecutive parameter estimates and the tracking error converge on average. This implies average convergence of the dual adaptive MPC independent of the difference between the consecutive parameter estimates and the covariance matrix. This result is less strict than the classical convergence criteria, yet it offers meaningful observations for practical applications.

Contributions: This investigation ensures the proper operation of the dual adaptive MPC algorithm by ensuring recursive feasibility and achieving average convergence of the Lyapunov decrease function. Without these properties, the system may fail to reach the desired output or exhibit destabilizing behavior. By removing the dependence of variables obtained from the parameter estimation in the average convergence analysis, the dual adaptive MPC is ensured to operate correctly despite online updates of the parameter estimates. This research advances the practical usability and reliability of the dual adaptive MPC.
We validate our findings by providing a mathematical proof of the average convergence. The analysis involves demonstrating a logarithmic upper bound for the average squared Euclidean norm of the difference between consecutive parameter estimates and for the Lyapunov decrease function. To support our theoretical findings, we will simulate the MPC algorithm under several examples.

Structure: The thesis is structured as follows: We will begin by outlining the problem and detailing the necessary assumptions and design choices of the dual adaptive MPC for the average convergence analysis. Then, we will present our main result as a theorem which states that the change between the parameter estimates and the tracking error will converge on average. We highlight the meaning of the theorem for the average convergence of the dual adaptive MPC.
Then, we will provide the proof of the average convergence of the change between the parameter estimates first, since it will be utilized for the tracking error proof as well. This will be accomplished by identifying a logarithmic upper bound for the average term. Then we proceed with the proof of the average convergence of the tracking error, where we first determine a candidate solution for the dual adaptive MPC scheme. Following this, we validate that the candidate solution is recursively feasible and satisfies all constraints. Then we derive the upper bound for the Lyapunov decrease using the candidate solution and the last optimal solution and form the average of the consecutive Lyapunov decrease function. We demonstrate that this upper bound converges on average.
Finally, we illustrate our findings with simulations of various examples.

Chapter 2

Literature Review

Dual Theory: The concept of simultaneous control and exploration, called dual theory, was suggested in [1] in 1960. The innovative idea involved introducing controlled noise into the system plant to promote exploration. While the primary objective of the controller remained steering towards a reference value, this noise injection enabled the acquisition of information regardless of the specific control strategy employed. The information gained from steering towards a reference value is described as passive exploration in [1]. By deliberately introducing controlled noise to the system, the controller was driven to actively gather information. [1] demonstrated that the quality of the information depended on the characteristics of the noise sequence, demonstrating how different noise strategies could influence the exploration process. This important insight encouraged further exploration and adoption of dual control strategies in control systems.

Dual Adaptive MPC: With model predictive control widely adopted in industries for its ability to handle constraints and its straightforward implementation with current software, [2] presented dual adaptive MPC in 2017, integrating dual control strategies into model predictive control. They utilized a linear system output model $y(t) = \theta^\top \phi(t-1) + v(t)$, where θ represents an unknown parameter and $v(t)$ denotes additive noise. The unknown parameter was estimated using recursive least squares estimation, and the covariance matrix was incorporated into the cost function to promote active exploration. Their work involved formulating the dual adaptive MPC as a stochastic optimal-control problem and reformulating it as a quadratic constrained quadratic programming problem. This reformulation is crucial because it improves the computational efficiency of the dual adaptive MPC and enables its use in practical applications. To ensure safety and achieve the control goal, stability analysis is required for the dual adaptive MPC.

Stability Analysis using Reinforcement Learning: Prior research has explored the stability of the dual adaptive MPC. For instance, [3] provided a stability analysis without tracking. Inspired by reinforcement learning theories, they treated the MPC algorithm as an agent and formulated the cost function as a cost-to-go function, similar to a Bellman equation. In their analysis, they approximated the cost function by neglecting the exploration cost after a certain number of steps. This approximation allowed them to demonstrate two costs: a finite horizon cost with exploration and an infinite horizon cost without exploration, resembling a terminal cost. This approach provided insights into the stability properties of the dual adaptive MPC algorithm.

Stability Analysis using Lyapunov Stability Analysis: Another approach to stability analysis was done in [4]. They provided a comprehensive stability analysis with a tracking objective using Lyapunov stability analysis. With the introduction of set-membership estimation on top of the RLS estimation, the accuracy of the parameter estimate improved. Furthermore, it enabled the imposition of robust constraints on truncated Gaussian noise models to ensure recursive feasibility. As a result, a stability analysis of the dual adaptive MPC is possible using the Lyapunov decrease method. They demonstrated uniformly practically exponentially stability with respect to the covariance matrix of the estimated parameter and the change between consecutive parameter estimates.

Summary: In summary, the evolution and current state of research in dual adaptive MPC began with [1] and progressed through recent advancements in [2], [3], and [4]. The reviewed studies collectively highlight the potential of integrating dual control strategies with MPC frameworks. However, stability analysis remains a significant challenge, with current approaches showing promises but also leaving room for improvement.

Chapter 3

Problem Formulation

3.1 System Properties

In this thesis, we consider a discrete time-invariant linear system model, which is described by the following state transition equation:

$$x_{t+1} = Ax_t + Bu_t, \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$ represents the state transition matrix, $B \in \mathbb{R}^{n\times m}$ represents the input matrix, and $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ represent the state and the input, respectively, at time t. The output of the system is denoted by

$$y_t = \theta_{true}^\top x_t + w_t, \tag{3.2}$$

where $y_t \in \mathbb{R}$ and $w_t \in \mathbb{R}$ denote the system output and the additive noise, respectively, at time t and $\theta_{true} \in \mathbb{R}^n$ is a constant unknown parameter vector. The goal is to steer the output to a desired output y^d as close as possible, while simultaneously providing an estimation of the unknown parameter vector.

Systems potentially using this model are, for example, autonomous car driving systems, where the known internal system is the engine system, the road condition is constant but unknown, and the output is the velocity. The objective of the car is to drive at a desired target velocity. Another example is in an electric heating system, where the heating system of a plate is known but the heat capacity of the material in the pot is unknown. The output is the temperature of the material, and the goal is to reach a desired temperature.

We consider the following assumptions for our discrete linear time-invariant system:

Assumption 1. *The pair (A, B) is stabilizable.*

Assumption 2. *The unknown parameter θ_{true} lies inside a predefined hyperbox set Θ_0.*

Assumption 3. *If y^d is unreachable, then given $\mu \in (0, 1)$ and any $\theta \in \Theta_0$, the following optimal steady state problem is strictly convex:*

$$\begin{aligned} &\min_{x,u} (\theta^\top x - y^d)^2 \\ &s.t.\ x = Ax + Bu \\ &(x, u) \in \mu\mathbb{Z} \end{aligned} \tag{3.3}$$

Assumption 1 is standard for tracking problems to ensure non-trivial solutions for the steady state and input. Assumption 2 allows us to impose a robust constraint to deal with bounded noise while having an unknown parameter vector. Assumption 3 ensures, that if y^d is unreachable, then there exists a unique steady state and input pair, where the error of the output to the desired output is minimized.

Both the state and the input are constrained within a polytope described by:

$$\mathbf{Z} := \{(x, u) | x \in \mathbf{X}, u \in \mathbf{U}\}, \tag{3.4}$$

where

$$\mathbf{X} := \{x | H_x x \leq h_x\}, \tag{3.5}$$

and

$$\mathbf{U} := \{u | H_u u \leq h_u\}, \tag{3.6}$$

for some matrices H_x, H_u and vectors h_x, h_u of appropriate dimensions. The output of our system is constrained within the interval:

$$\mathbb{Y} := [y_{min}, y_{max}]. \tag{3.7}$$

We consider that the additive noise is bounded within the interval:

$$\mathbb{W} := [-\bar{w}, \bar{w}], \tag{3.8}$$

with zero mean and a variance of σ^2. We consider polytopic constraints, because they are widely used in systems with parameter uncertainties and in MPC controllers. They provide the advantage of representing the constraints as a set of linear inequalities. This results in increased computation efficiency when solving the optimization problem of the MPC.

3.2 Parameter Estimation

Since θ_{true} is unknown, an estimation is needed to achieve our control objective. From the current measured state x_t and output y_t and the last estimated parameter $\bar{\theta}_{t-1}$, we first employ the recursive least squares estimation algorithm to obtain the estimate $\hat{\theta}_t$ and its current covariance matrix Π_t^{-1}. The recursive least squares update algorithm is described by the following equations for all time $t \leq 1$ with a given initial parameter estimation $\bar{\theta}_0 \in \Theta_0$ and covariance matrix $\Pi_0^{-1} \preceq \frac{1}{\lambda} I$ with $\lambda > 0$:

$$\hat{\theta}_t = \bar{\theta}_{t-1} + \sigma^{-2} \Pi_t^{-1} x_t e_t \tag{3.9}$$

$$\Pi_t = \Pi_{t-1} + \sigma^{-2} x_t x_t^\top \tag{3.10}$$

The vector $e_t = y_t - \bar{\theta}_{t-1}^\top x_t$ in (3.9) describes the estimation error of the current measured output and the noise-free estimated output. Since we assumed that the predefined hyperbox set Θ_0 contains the unknown parameter vector θ_{true}, we project the RLS estimated parameter vector $\hat{\theta}_t$ into the hyperbox set Θ_0 to increase the accuracy of the estimate. It is possible to update the hyperbox with the set-membership estimation to further improve the accuracy, which is done in [4]. We decided to ignore the update to simplify the analysis. The resulting projection is the next estimated parameter vector, $\bar{\theta}_t$. The projection is done by using the following formula:

$$\bar{\theta}_t = \arg \min_{\theta \in \Theta_0} ||\theta - \hat{\theta}_t||_{\Pi_t}^2 \tag{3.11}$$

We define the difference between consecutive parameter estimates as:

$$\Delta \bar{\theta}_{t-1} = \bar{\theta}_t - \bar{\theta}_{t-1} \tag{3.12}$$

3.3 MPC Scheme

The control goal is to steer the output y_t to the desired output y^d as close as possible and to improve the parameter estimation concurrently while satisfying the constraints. Since y^d might be unreachable, we define at each time t the closest reachable output y_t^d with the correspond steady state x_t^d and input u_t^d under dependence on the current estimated parameter vector $\bar{\theta}_t$ as follows:

$$
\begin{aligned}
(x_t^d, u_t^d, y_t^d) = \arg\min \, &||y^d - y||_2 && \text{(3.13a)} \\
\text{s.t.} \quad & (x, u) = M_\Psi \Psi_t && \text{(3.13b)} \\
& y = \bar{\theta}_t^\top x && \text{(3.13c)} \\
& y_{min} + \bar{w} \leq \bar{\theta}_{i,0}^\top x \leq y_{max} - \bar{w} && \text{(3.13d)} \\
& i = 1, ..., 2^n
\end{aligned}
$$

It is important to recognize that due to the update of the estimated parameter vector at each time step, the estimated parameter vector $\bar{\theta}_{t+1}$ might differ from $\bar{\theta}_t$. This consequently leads to different solutions between x_{t+1}^d, u_{t+1}^d, and y_{t+1}^d and x_t^d, u_t^d, and y_t^d, respectively, when solving (3.13).

We solve the control problem with the dual adaptive MPC, which has the following scheme:

$$
\begin{aligned}
\min \, & J_t(x_t, \bar{\theta}_t, \Theta_0, \Pi_t, y^d) && \text{(3.14a)} \\
\text{s.t.} \quad & x_{k+1|t} = A x_{k|t} + B u_{k|t} && \text{(3.14b)} \\
& \Pi_{k+1|t} = \Pi_{k|t} + \sigma^{-2} x_{k|t} x_{k|t}^\top && \text{(3.14c)} \\
& y_{min} + \bar{w} \leq \bar{\theta}_{i,0}^\top x_{k|t} \leq y_{max} - \bar{w} && \text{(3.14d)} \\
& (x_t^s, u_t^s) = M_\Psi \Psi_t && \text{(3.14e)} \\
& y_t^s = \bar{\theta}_t^\top x_t^s && \text{(3.14f)} \\
& x_{N|t} = x_t^s && \text{(3.14g)} \\
& (x_{k|t}, u_{k|t}) \in \mathbf{Z} && \text{(3.14h)} \\
& x_{0|t} = x_t, \quad \Pi_{0|t} = \Pi_t && \text{(3.14i)} \\
& k = 0, ..., N-1 \quad i = 1, ..., 2^n,
\end{aligned}
$$

where x_t^s, u_t^s, and y_t^s are the optimal artificial steady state, input, and output, respectively. We use the artificial steady state and input pair (x_t^s, u_t^s) for tracking, as proposed in [5], to ensure recursive feasibility of the dual adaptive MPC with Assumption 1. Instead of directly tracking the desired state and input, which may pose some challenges due to the unknown parameter vector θ_{true}, this approach establishes a reference trajectory that the system can track with a feasibility guarantee in the presence of uncertainties.

We also introduce the parametrization of the artificial steady state and input with the matrix M_Ψ and the parameter vector Ψ_t to solve the optimization problem more efficiently. This parametrization is included as the constraint (3.14e), where the matrix M_Ψ is chosen by design and the vector Ψ_t is to be optimized by the dual adaptive MPC.

We refer to $x_{k|t}$ and $u_{k|t}$ as the optimal predicted state and input, respectively, after k steps at time t. Compared to the usual MPC scheme, where the terminal set is an invariant set, our MPC scheme has a terminal equality constraint (3.14g), such that the last predicted state equals the optimal artificial steady state, hence $x^*_{N|t} = x^{s,*}_t$. This can be conservative in practice, but it simplifies the analysis. With this change, we also no longer need to calculate the closed-loop control gain matrix since $x^*_{N|t} - x^{s,*}_t = 0_n$ and our terminal control law is $u = u^s_t$.

The constraint (3.14d) ensures the constraint satisfaction of (3.7) for all future outputs, for any future noise $w_{t+1} \in \mathbb{W}$ and parameter estimate $\theta_{t+1} \in \Theta_0$. The hyperbox set Θ_0 is described by its i vertices $\bar{\theta}_{i,0}$ for $i = 1, ..., 2^n$.
The cost function $J_t(x_t, \bar{\theta}_t, \Theta_0, \Pi_t, y^d)$ of the dual adaptive MPC scheme is defined as:

$$
\begin{aligned}
J_t(x_t, \bar{\theta}_t, \Theta_0, \Pi_t, y^d) &:= \sum_{k=0}^{N-1} (||\bar{\theta}_t^\top (x_{k|t} - x^s_t)||^2_q + ||x_{k|t}||^2_{\Pi^{-1}_{k|t} q} \\
&+ ||u_{k|t} - u^s_t||^2_R + ||x_{k|t} - x^s_t||^2_{Q_0}) + ||\bar{\theta}_t^\top x^s_t - y^d||^2_{T_c},
\end{aligned}
\tag{3.15}
$$

where $q > 0$ is the weight of the information matrix, R is the input cost matrix, Q_0 is the state cost matrix, and T_c is the terminal cost matrix. The first, third, and fourth, terms of (3.15) refer to the tracking cost penalty for the artificial steady output, input, and state, respectively. We refer to the second term of (3.15) as the learning cost, since this term encourages active exploration of the system by promoting the state to go to the origin. This learning cost corresponds to the mean cost of a stochastic process. The last term of (3.15) refer to the cost penalty between the artificial steady output and the desired output.

For the analysis of this thesis, we assume the following for Q_0:

Assumption 4. *The the following condition holds:*

$$
Q_0 \succeq \frac{4q}{\lambda} I \succeq 4q\Pi_0^{-1}. \tag{3.16}
$$

This assumption establishes a relation between the state cost matrix Q_0 and the initial covariance matrix Π_0^{-1}. It will be crucial when we upper bound the Lyapunov decrease function.

The overall algorithm is summarized in Algorithm 1.

Algorithm 1 Dual Adaptive MPC Scheme

Given the model, initial state x_0, hyperbox set Θ_0 ,prior estimate $\bar{\theta}_0 \in \Theta_0$, and covariance matrix Π_0^{-1}.
for $t \in \mathbb{N}_0$ **do**
  Measure the state x_t and the output y_t.
  Solve the MPC optimization problem (3.14).
  Update the covariance matrix Π_t^{-1} with (3.10).
  Update the estimate parameter vector $\bar{\theta}_t$ with (3.9) and 3.11.
  Apply the control input $u_t = u^*_{0|t}$ to the system (3.1).
end for

Chapter 4

Average Convergence Analysis of the Dual Adaptive MPC

We present the following theorem as the main result of this thesis.

Theorem 1. *Suppose that Assumptions 1, 2, 3, and 4 hold. Suppose that the MPC scheme (3.14) is feasible at* $t = 0$*, then the following hold:*

1. *The average squared Euclidean norm of the difference between consecutive parameter estimates converges to zero:*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 = 0, \tag{4.1}$$

2. *The average squared weighted norm of the tracking error converges to zero:*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} ||x_t - x_t^d||_{Q_{\bar{\theta}_t}}^2 = 0, \tag{4.2}$$

where $Q_{\bar{\theta}_t} = \frac{1}{2}Q_0 + q\bar{\theta}_t\bar{\theta}_t^\top$.

Theorem 1 refers to the average convergence of the dual adaptive MPC. Equation (4.1) indicates the average convergence of the estimated parameter vector. This also implies the average convergence of the solution of the optimization problem (3.13), since the constraints (3.13d) and (3.13b) doesn't change dynamically and the solution solely depends on the estimated parameter vector. Equation (4.2) suggests that the dual adaptive MPC will steer towards the closest reachable state x_t^d on average, satisfying (3.13). Since $Q_{\bar{\theta}_t}$ depends on $\bar{\theta}_t$, this also implies that the output will steer towards the closest reachable output y_t^d.
It is important to note that even if the closest reachable output from the solution of the optimization problem (3.13) converges, there is no guarantee that the converged value is the desired output due to the constraints in (3.13). Furthermore, we have no guarantee that the estimated parameter vector converges to the real one, as there might not be enough excitation in the system to ensure this. Still, Theorem 1 provides some guarantees for average convergence, which is essential for practical applications.
In order to prove Theorem 1, we will proceed as follows: First, we prove the average convergence of the parameter estimation. Then we analyze the stability of the MPC with standard procedures. We define a candidate solution and verify that the chosen candidate solution is feasible and satisfies all constraints. Then, we derive an upper bound for the Lyapunov decrease function and the corresponding average from the cost function (3.15) and show that this upper bound converges on average.

4.1 Proof of the Equation (4.1)

In this section, we provide the proof for Equation (4.1), which is the average convergence of the estimated parameter.

Proof (4.1): We first introduce the following Lemma:

Lemma 1. *With the constraints introduced in (3.5), (3.7), and under Assumption 2 and 4, the following inequality holds:*

$$\sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 \leq \frac{\sigma^{-2} n e_{max}}{\lambda} \log\left(\frac{(T+1)\sigma^{-2}(x_{max})^2 + \lambda}{\lambda}\right) \tag{4.3}$$

with $\Delta\bar{\theta}_t$ defined as in (3.12).

Lemma 1 enables us to establish a logarithmic upper bound of the cumulative sum of the difference between consecutive parameter estimates for the average analysis of Equation (4.1).

Proof Lemma 1: We begin by stating the following using Assumption 4 and (3.10):

$$\Delta\bar{\theta}_t^T \Pi_{t+1} \Delta\bar{\theta}_t \overset{(3.10)}{\geq} \Delta\bar{\theta}_t^T \Pi_0 \Delta\bar{\theta}_t \overset{(3.16)}{\geq} \Delta\bar{\theta}_t^T \lambda I \Delta\bar{\theta}_t, \tag{4.4}$$

The reasoning of this inequality can be found in Appendix A.5. We use (4.4) to reformulate the following:

$$\sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 = \Delta\bar{\theta}_t^T \Delta\bar{\theta}_t \overset{(4.4)}{\leq} \sum_{t=0}^{T} \frac{1}{\lambda} \Delta\bar{\theta}_t^T \Pi_{t+1} \Delta\bar{\theta}_t \leq \sum_{t=0}^{T} \frac{1}{\lambda} ||\bar{\theta}_{t+1} - \bar{\theta}_t||_{\Pi_{t+1}}^2. \tag{4.5}$$

Using the property of the non-expansiveness of the projection, which is proven in Appendix A.6 we can further reformulate (4.5) to:

$$\sum_{t=0}^{T} \frac{1}{\lambda} ||\bar{\theta}_{t+1} - \bar{\theta}_t||_{\Pi_{t+1}}^2 \overset{(A.17)}{\leq} \sum_{t=0}^{T} \frac{1}{\lambda} ||\hat{\theta}_{t+1} - \bar{\theta}_t||_{\Pi_{t+1}}^2 \tag{4.6}$$

Since $\bar{\theta}_t$ is already within the predefined hyperbox, the projection of $\bar{\theta}_t$ is $\bar{\theta}_t$ itself. We now utilize the RLS algorithm update (3.10) and Lemma 2 to further reformulate the term:

$$\begin{aligned}
&\sum_{t=0}^{T} \frac{1}{\lambda} ||\hat{\theta}_{t+1} - \bar{\theta}_t||_{\Pi_{t+1}}^2 = \sum_{t=0}^{T} \frac{1}{\lambda} ||\Pi_{t+1}^{\frac{1}{2}} (\hat{\theta}_{t+1} - \bar{\theta}_t)||_2^2 \\
&\overset{(3.9)}{=} \sum_{t=0}^{T} \frac{1}{\lambda} ||\Pi_{t+1}^{\frac{1}{2}} \sigma^{-2} \Pi_{t+1}^{-1} x_{t+1} e_{t+1}||_2^2 = \sum_{t=0}^{T} \frac{1}{\lambda} ||\sigma^{-2} \Pi_{t+1}^{-\frac{1}{2}} x_{t+1} e_{t+1}||_2^2 \\
&\overset{(A.9)}{\leq} \sum_{t=0}^{T} \frac{1}{\lambda} \sigma^{-4} ||\Pi_{t+1}^{-\frac{1}{2}} x_{t+1}||_2^2 ||e_{t+1}||_2^2 \\
&= \frac{\sigma^{-4}}{\lambda} \sum_{t=0}^{T} ||\Pi_{t+1}^{-\frac{1}{2}} x_{t+1}||_2^2 ||e_{t+1}||_2^2 \leq \frac{\sigma^{-4}}{\lambda} \sum_{t=0}^{T} ||\Pi_{t+1}^{-\frac{1}{2}} x_{t+1}||_2^2 e_{max} \\
&= \frac{\sigma^{-4} e_{max}}{\lambda} \sum_{t=1}^{T+1} ||x_t||_{\Pi_t^{-1}}^2
\end{aligned} \tag{4.7}$$

where $e_{max} = \max ||e||_2^2$. The error $e_{t+1} = y_{t+1} - \bar{\theta}_t^T x_{t+1}$ is bounded because of the state and output constraints, and Assumption 2. Thus every variable appearing in e_{t+1} is bounded, implying that e_{t+1} is bounded too.
Before proceeding, we provide another Lemma that is similar to the result in [6] to deal with the sum in Equation (4.7):

Lemma 2. *For $\Pi_0^{-1} \succeq \lambda I$, Π_t defined as in (3.10), and x_t is bounded for all t. Then the following relation holds:*

$$\sum_t^T ||x_t||^2_{\Pi_t^{-1}} \leq \sigma^2 n \log(\frac{T\sigma^{-2}(x_{max})^2 + \lambda}{(t-1)\sigma^{-2}(x_{max})^2 + \lambda}), \tag{4.8}$$

This lemma provides a logarithmic upper bound of the learning cost. We utilize this result both now and later in our analysis of the average convergence of the dual adaptive MPC.

Proof Lemma 2: We start by reformulating $\det(\Pi_{t-1})$ as follows:

$$\begin{aligned} &\det(\Pi_{t-1}) \overset{(3.10)}{=} \det(\Pi_t - \sigma^{-2} x_t x_t^\top) = \det(\Pi_t)\det(I - \sigma^{-2}\Pi_t^{-\frac{1}{2}} x_t x_t^\top \Pi_t^{-\frac{1}{2}}) \\ &= \det(\Pi_t)\det(I - \sigma^{-2} x_t^\top \Pi_t^{-1} x_t) = \det(\Pi_t)(1 - \sigma^{-2} x_t^\top \Pi_t^{-1} x_t). \end{aligned} \tag{4.9}$$

We can reformulate $||x_t||^2_{\Pi_t^{-1}}$ as follows:

$$\begin{aligned} &\frac{\det(\Pi_{t-1})}{\det(\Pi_t)} = 1 - \sigma^{-2} x_t^\top \Pi_t^{-1} x_t \\ &\sigma^{-2} x_t^\top \Pi_t^{-1} x_t = 1 - \frac{\det(\Pi_{t-1})}{\det(\Pi_t)} \\ &||x_t||^2_{\Pi_t^{-1}} = \sigma^2(1 - \frac{\det(\Pi_{t-1})}{\det(\Pi_t)}), \end{aligned} \tag{4.10}$$

From (A.13) in the Appendix, we have the following relation:

$$\frac{\det(\Pi_{t-1})}{\det(\Pi_t)} \leq 1. \tag{4.11}$$

Given the fact that $1 - x \leq \log(\frac{1}{x})$ for $x \leq 1$, we get from (4.10) to

$$\begin{aligned} &||x_t||^2_{\Pi_t^{-1}} = \sigma^2(1 - \frac{\det(\Pi_{t-1})}{\det(\Pi_t)}) \leq \sigma^2 \log(\frac{\det(\Pi_t)}{\det(\Pi_{t-1})}) \\ &= \sigma^2(\log\det(\Pi_t) - \log\det(\Pi_{t-1})). \end{aligned} \tag{4.12}$$

The next step is to reformulate $\sum_t^T ||x_t||^2_{\Pi_t^{-1}}$ as follows:

$$\begin{aligned} &\sum_t^T ||x_t||^2_{\Pi_t^{-1}} \overset{(4.12)}{\leq} \sum_t^T \sigma^2(\log\det(\Pi_t) - \log\det(\Pi_{t-1})) \\ &= \sigma^2(\log\det(\Pi_T) - \log\det(\Pi_{t-1})) \end{aligned} \tag{4.13}$$

With Π_t defined as in (3.10) for all t, $\log\det(\Pi_t)$ is rewritten as:

$$\begin{aligned} &\log\det(\Pi_t) \overset{(3.10)}{=} \log\det(\Pi_{t-1} + \sigma^{-2} x_t x_t^\top) \\ &\overset{(3.10)}{=} \log\det(\lambda I + \sigma^{-2}\sum_{k=0}^t x_k x_k^\top) \leq \log\det(\lambda I + \sigma^{-2}\sum_{k=0}^t (x_{max})^2 I) \\ &= \log\det(\lambda I + t\sigma^{-2}(x_{max})^2 I) = \log\det((\lambda + t\sigma^{-2}(x_{max})^2)I) \\ &= \log(\lambda + t\sigma^{-2}(x_{max})^2)^n = n\log(\lambda + t\sigma^{-2}(x_{max})^2), \end{aligned} \tag{4.14}$$

where $x_{max} = \max_{x \in \mathbf{X}} ||x_t||_2^2$, and n is the dimension of the state. The inequality holds because of the state being constrained. Thus, we can upper bound $\sum_t^T ||x_t||^2_{\Pi_t^{-1}}$ with the following logarithmic bound:

$$\begin{aligned}
\sum_t^T ||x_t||^2_{\Pi_t^{-1}} &\leq \sigma^2(\log\det(\Pi_T) - \log\det(\Pi_t)) \\
&\overset{(4.14)}{\leq} \sigma^2 n \log(\frac{T\sigma^{-2}(x_{max})^2 + \lambda}{(t-1)\sigma^{-2}(x_{max})^2\lambda})
\end{aligned} \tag{4.15}$$

□

Now we proceed with the proof of Lemma 1. By utilizing (4.8) in (4.7), we get:

$$\begin{aligned}
\sum_{t=0}^T ||\Delta\bar{\theta}_t||_2^2 &\leq \frac{\sigma^{-4}e_{max}}{\lambda} \sum_{t=1}^{T+1} ||x_t||^2_{\Pi_t^{-1}} \\
&\overset{(4.8)}{\leq} \frac{\sigma^{-4}e_{max}}{\lambda}\sigma^2 n \log(\frac{(T+1)\sigma^{-2}(x_{max})^2 + \lambda}{\lambda}) \\
&= \frac{\sigma^{-2}ne_{max}}{\lambda} \log(\frac{(T+1)\sigma^{-2}(x_{max})^2 + \lambda}{\lambda}).
\end{aligned} \tag{4.16}$$

□

With Lemma 1, we recognize that (4.1) holds:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^T ||\Delta\bar{\theta}_t||_2^2 \overset{(4.3)}{\leq} \lim_{T\to\infty} \frac{\sigma^{-2}ne_{max}}{T\lambda} \log(\frac{(T+1)\sigma^{-2}(x_{max})^2 + \lambda}{\lambda}) = 0, \tag{4.17}$$

with the reasoning behind the last equality being that as T approaches infinity, $\frac{\log(T)}{T}$ tends to zero. □

4.2 MPC Candidate Solution

Now, let's proceed with the proof of the Equation (4.2) of Theorem 1. We start by choosing a suitable candidate solution. Given the optimal solution at time t, denoted by *, we define the following candidate solution at time $t+1$:

$$\begin{aligned}
&\Psi_{t+1} := \Psi_t^*, \quad x_{t+1}^s := x_t^{s,*}, \quad u_{t+1}^s := u_t^{s,*} \\
&u_{k|t+1} := u_{k+1|t}^*, \quad k = 0, ..., N-2 \\
&x_{k|t+1} := x_{k+1|t}^*, \quad \Pi_{k|t+1} := \Pi_{k+1|t}^*, \quad k = 0, ..., N-1 \\
&\Pi_{N|t+1} := \Pi_{N+1|t}^* := \Pi_{N-1|t+1} + \sigma^{-2} x_{N|t+1} x_{N|t+1}^\top \\
&u_{N-1|t+1} := u_{N|t}^* := u_t^{s,*} \\
&x_{N|t+1} := Ax_{N-1|t+1} + Bu_{N-1|t+1}
\end{aligned} \tag{4.18}$$

The chosen candidate solution at time $t+1$, is the optimal solution at time t, shifted by one time instant. Since the state x and the input u are not influenced by uncertainty, the constraint (3.14h) from the MPC ensures constraint satisfaction of the candidate state and input. The constraint satisfaction of the candidate output is ensured by (3.14d), since we have $\bar{\theta}_{t+1} \in \Theta_0$ from our parameter estimation and (3.14d) takes account of all possible noise robustly. With the choice of the candidate solution and the terminal equality constraint (3.14g), recursive feasibility is ensured for this particular candidate solution.

4.3 Upper Bounding the Average Lypunov Decrease Function

In this section we present the upper bounded, decomposed equation of the average Lyapunov decrease function, for which we will analyze each term separately for convergence.

Lemma 3. *The Lyapunov decrease function for (3.15) is upper bounded by:*

$$\begin{aligned} J_{t+1}^* - J_t^* \leq J_{t+1} - J_t^* \leq ||x_t||^2_{\Pi_t^{-1}q} + \sum_{k=1}^{N-1} ||x^*_{k|t} - x_t^{s,*}||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} \\ + ||\Delta\bar{\theta}_t x_t^{s,*}||^2_{T_c} + 2\Delta\bar{\theta}_t^T x_t^{s,*} T_c(y_t^{s,*} - y^d) - (||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R), \end{aligned} \tag{4.19}$$

with $Q_{\bar{\theta}_t} = \frac{1}{2}Q_0 + q\bar{\theta}_t\bar{\theta}_t^\top$ *and* J_{t+1} *as the cost of a candidate solution defined as*

$$\begin{aligned} J_{t+1}(x_{t+1}, \bar{\theta}_{t+1}, \Theta_0, \Pi_{t+1}, y^d) := \sum_{k=0}^{N-1} (||\bar{\theta}_{t+1}^\top (x_{k|t+1} - x_{t+1}^s)||^2_q \\ + ||x_{k|t+1}||^2_{\Pi^{-1}_{k|t+1}q} + ||u_{k|t+1} - u_{t+1}^s||^2_R + ||x_{k|t+1} - x_{t+1}^s||^2_{Q_0}) \\ + ||\bar{\theta}_{t+1}^\top x_{t+1}^s - y^d||^2_{T_c} \end{aligned} \tag{4.20}$$

Proof Lemma 3: We first calculate the difference in cost between a candidate solution at time $t+1$ and the optimal solution at time t, which is

$$\begin{aligned} J_{t+1} - J_t^* \overset{(3.15)}{=} \sum_{k=0}^{N-1} (||\bar{\theta}_{t+1}^\top (x_{k|t+1} - x_{t+1}^s)||^2_q + ||x_{k|t+1}||^2_{\Pi^{-1}_{k|t+1}q} \\ + ||u_{k|t+1} - u_{t+1}^s||^2_R + ||x_{k|t+1} - x_{t+1}^s||^2_{Q_0}) + ||\bar{\theta}_{t+1}^\top x_{t+1}^s - y^d||^2_{T_c} \\ - (\sum_{k=0}^{N-1} (||\bar{\theta}_t^\top (x^*_{k|t} - x_t^{s,*})||^2_q + ||x^*_{k|t}||_{\Pi^{*-1}_{k|t}q} + ||u^*_{k|t} - u_t^{s,*}||^2_R + \\ ||x^*_{k|t} - x_t^{s,*}||^2_{Q_0}) + ||\bar{\theta}_t^\top x_t^{s,*} - y^d||^2_{T_c}). \end{aligned} \tag{4.21}$$

Next, we simplify (4.21) by substituting (4.18) into our chosen candidate solution. The equation then becomes:

$$\begin{aligned} J_{t+1} - J_t^* \overset{(4.21,4.18)}{=} \sum_{k=0}^{N-1} (||\bar{\theta}_{t+1}^\top (x^*_{k+1|t} - x_t^{s,*})||^2_q + ||x^*_{k+1|t}||^2_{\Pi^{*-1}_{k+1|t}q} \\ + ||u^*_{k+1|t} - u_t^{s,*}||^2_R + ||x^*_{k+1|t} - x_t^{s,*}||^2_{Q_0}) + ||\bar{\theta}_{t+1}^\top x_t^{s,*} - y^d||^2_{T_c} - \\ (\sum_{k=0}^{N-1} (||\bar{\theta}_t^\top (x^*_{k|t} - x_t^{s,*})||^2_q + ||x^*_{k|t}||^2_{\Pi^{*-1}_{k|t}q} + ||u^*_{k|t} - u_t^{s,*}||^2_R + ||x^*_{k|t} - x_t^{s,*}||^2_{Q_0}) \\ + ||\bar{\theta}_t^\top x_t^{s,*} - y^d||^2_{T_c}) \end{aligned} \tag{4.22}$$

Since the candidate solution at time $t+1$ that we choose is the optimal solution at time t, we result in similar terms in (4.22) that appear in the optimal and candidate solutions, which allows us to simplify them. We group up similar terms and simplify them through reformulation.
We begin with the reorganization of the learning cost terms, which we can write as:

$$\sum_{k=0}^{N-1} ||x^*_{k+1|t}||^2_{\Pi^{*-1}_{k+1|t}q} - \sum_{k=0}^{N-1} ||x^*_{k|t}||^2_{\Pi^{*-1}_{k|t}q} = ||x^*_{N|t}||^2_{\Pi^{*-1}_{N|t}q} - ||x_t||^2_{\Pi_t^{-1}q}. \tag{4.23}$$

Similarly, simplification is done for the difference in state tracking costs between the optimal solution and the candidate, which is then:

$$\begin{aligned}&\sum_{k=0}^{N-1} ||x^*_{k+1|t} - x^{s,*}_t||^2_{Q_0} - \sum_{k=0}^{N-1} ||x^*_{k|t} - x^{s,*}_t||^2_{Q_0} = \\ &||x^*_{N|t} - x^{s,*}_t||^2_{Q_0} - ||x_t - x^{s,*}_t||^2_{Q_0} \overset{(3.14g)}{=} -||x_t - x^{s,*}_t||^2_{Q_0},\end{aligned} \tag{4.24}$$

where we use the terminal equality constraint (3.14g).

An analogous reformulation can be made for the difference in the input tracking cost:

$$\begin{aligned}&\sum_{k=0}^{N-1} ||u^*_{k+1|t} - u^{s,*}_t||^2_{R} - \sum_{k=0}^{N-1} ||u^*_{k|t} - u^{s,*}_t||^2_{R} \\ &\overset{(4.18)}{=} ||u^*_{N|t} - u^{s,*}_t||^2_{R} - ||u_t - u^{s,*}_t||^2_{R} \overset{(3.14g)}{=} -||u_t - u^{s,*}_t||^2_{R}\end{aligned} \tag{4.25}$$

where $u^*_{N|t} = u^{s,*}_t$ comes from (4.18).

For the output tracking cost, we first reformulate the candidate solution as follows:

$$\begin{aligned}&\sum_{k=0}^{N-1} ||\bar{\theta}^\top_{t+1}(x^*_{k+1|t} - x^{s,*}_t)||^2_q \\ &= \sum_{k=0}^{N-2} ||\bar{\theta}^\top_{t+1}(x^*_{k+1|t} - x^{s,*}_t)||^2_q + ||\bar{\theta}^\top_{t+1}(x^*_{N|t} - x^{s,*}_t)||^2_q \\ &\overset{(3.12, 3.14g)}{=} \sum_{k=0}^{N-2} ||(\bar{\theta}_t + \Delta\bar{\theta}_t)^\top (x^*_{k+1|t} - x^{s,*}_t)||^2_q \\ &= \sum_{k=1}^{N-1} (x^*_{k|t} - x^{s,*}_t)^\top (\bar{\theta}_t + \Delta\bar{\theta}_t) q (\bar{\theta}_t + \Delta\bar{\theta}_t)^\top (x^*_{k|t} - x^{s,*}_t) \\ &= \sum_{k=1}^{N-1} (x^*_{k|t} - x^{s,*}_t)^\top q (\bar{\theta}_t\bar{\theta}^\top_t + 2\Delta\bar{\theta}_t\bar{\theta}^\top_t + \Delta\bar{\theta}_t\Delta\bar{\theta}^\top_t)(x^*_{k|t} - x^{s,*}_t)\end{aligned} \tag{4.26}$$

The second term of the output tracking cost can be rewritten as:

$$\begin{aligned}&\sum_{k=0}^{N-1} ||\bar{\theta}^\top_t (x^*_{k|t} - x^{s,*}_t)||^2_q = \sum_{k=1}^{N-1} ||\bar{\theta}^\top_t (x^*_{k|t} - x^{s,*}_t)||^2_q + ||\bar{\theta}^\top_t (x_t - x^{s,*}_t)||^2_q \\ &= \sum_{k=1}^{N-1} (x^*_{k|t} - x^{s,*}_t)^\top q\bar{\theta}_t\bar{\theta}^\top_t (x^*_{k|t} - x^{s,*}_t) + ||x_t - x^{s,*}_t||^2_{q\bar{\theta}_t\bar{\theta}^\top_t}\end{aligned} \tag{4.27}$$

Now, we combine (4.26) and (4.27) together and simplify the output tracking cost to:

$$
\begin{aligned}
&\sum_{k=0}^{N-1} ||\bar{\theta}_{t+1}^{\top}(x_{k+1|t}^{*} - x_t^{s,*})||_q^2 - \sum_{k=0}^{N-1} ||\bar{\theta}_t^{\top}(x_{k|t}^{*} - x_t^{s,*})||_q^2 \\
&\overset{(4.26,4.27)}{=} \sum_{k=1}^{N-1} (x_{k|t}^{*} - x_t^{s,*})^{\top} q(\bar{\theta}_t\bar{\theta}_t^{\top} + 2\Delta\bar{\theta}_t\bar{\theta}_t^{\top} + \Delta\bar{\theta}_t\Delta\bar{\theta}_t^{\top})(x_{k|t}^{*} - x_t^{s,*}) \\
&- \sum_{k=1}^{N-1} (x_{k|t}^{*} - x_t^{s,*})^{\top} q\bar{\theta}_t\bar{\theta}_t^{\top}(x_{k|t}^{*} - x_t^{s,*}) - ||x_t - x_t^{s,*}||_{q\bar{\theta}_t\bar{\theta}_t^{\top}}^2 \\
&= \sum_{k=1}^{N-1} (x_{k|t}^{*} - x_t^{s,*})^{\top} q(2\Delta\bar{\theta}_t\bar{\theta}_t^{\top} + \Delta\bar{\theta}_t\Delta\bar{\theta}_t^{\top})(x_{k|t}^{*} - x_t^{s,*}) - ||x_t - x_t^{s,*}||_{q\bar{\theta}_t\bar{\theta}_t^{\top}}^2 \\
&= \sum_{k=1}^{N-1} ||x_{k|t}^{*} - x_t^{s,*}||_{2q\Delta\bar{\theta}_t\bar{\theta}_t^{\top}}^2 + \sum_{k=1}^{N-1} ||x_{k|t}^{*} - x_t^{s,*}||_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^{\top}}^2 - ||x_t - x_t^{s,*}||_{q\bar{\theta}_t\bar{\theta}_t^{\top}}^2 .
\end{aligned}
\tag{4.28}
$$

The terms of the terminal cost can be expressed as:

$$
\begin{aligned}
&||\bar{\theta}_{t+1}^{\top} x_t^{s,*} - y^d||_{T_c}^2 - ||\bar{\theta}_t^{\top} x_t^{s,*} - y^d||_{T_c}^2 \\
&\overset{(3.12)}{=} ||(\Delta\bar{\theta}_t + \bar{\theta}_t)^{\top} x_t^{s,*} - y^d||_{T_c}^2 - ||\bar{\theta}_t^{\top} x_t^{s,*} - y^d||_{T_c}^2 \\
&= (\Delta\bar{\theta}_t^{\top} x_t^{s,*} + \bar{\theta}_t^{\top} x_t^{s,*} - y^d)^{\top} T_c(\Delta\bar{\theta}_t^{\top} x_t^{s,*} + \bar{\theta}_t^{\top} x_t^{s,*} - y^d) \\
&- (\bar{\theta}_t^{\top} x_t^{s,*} - y^d)^{\top} T_c(\bar{\theta}_t^{\top} x_t^{s,*} - y^d) \\
&= x_t^{s,*\top}\Delta\bar{\theta}_t T_c \Delta\bar{\theta}_t^{\top} x_t^{s,*} - 2\bar{\theta}_t^{\top} x_t^{s,*} T_c y^d + 2x_t^{s,*\top}\Delta\bar{\theta}_t T_c \bar{\theta}_t^{\top} x_t^{s,*} \\
&- 2x_t^{s,*\top}\Delta\bar{\theta}_t T_c y^d + x_t^{s,*\top}\bar{\theta}_t T_c \bar{\theta}_t^{\top} x_t^{s,*} + y^d T_c y^d - x_t^{s,*\top}\bar{\theta}_t T_c \bar{\theta}_t^{\top} x_t^{s,*} \\
&+ 2\bar{\theta}_t^{\top} x_t^{s,*} T_c y^d - y^d T_c y^d \\
&= x_t^{s,*\top}\Delta\bar{\theta}_t T_c \Delta\bar{\theta}_t^{\top} x_t^{s,*} + 2x_t^{s,*\top}\Delta\bar{\theta}_t T_c \bar{\theta}_t^{\top} x_t^{s,*} - 2x_t^{s,*\top}\Delta\bar{\theta}_t T_c y^d \\
&= ||\Delta\bar{\theta}_t^{\top} x_t^{s,*}||_{T_c}^2 + 2x_t^{s,*\top}\Delta\bar{\theta}_t T_c(\bar{\theta}_t^{\top} x_t^{s,*} - y^d)
\end{aligned}
\tag{4.29}
$$

Combine the terms back, the difference of the cost function transforms into:

$$
\begin{aligned}
&J_{t+1} - J_t^{*} \overset{(4.23,4.24,4.25,4.28,4.29)}{=} ||x_{N|t}^{*}||_{\Pi_{N|t}^{*-1} q}^2 - ||x_t||_{\Pi_t^{-1} q}^2 - ||x_t - x_t^{s,*}||_{Q_0}^2 \\
&- ||x_t - x_t^{s,*}||_{q\bar{\theta}_t\bar{\theta}_t^{\top}} - ||u_t - u_t^{s,*}||_R^2 + \sum_{k=1}^{N-1} ||x_{k|t}^{*} - x_t^{s,*}||_{2q\Delta\bar{\theta}_t\bar{\theta}_t^{\top}}^2 \\
&+ \sum_{k=1}^{N-1} ||x_{k|t}^{*} - x_t^{s,*}||_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^{\top}}^2 + ||\Delta\bar{\theta}_t^{\top} x_t^{s,*}||_{T_c}^2 + 2x_t^{s,*T}\Delta\bar{\theta}_t T_c(\bar{\theta}_t^{\top} x_t^{s,*} - y^d)
\end{aligned}
\tag{4.30}
$$

Now, we upper bound the term $||x_{N|t}^{*}||_{\Pi_{N|t}^{*-1} q}^2$ as follows:

$$
\begin{aligned}
&||x_{N|t}^{*}||_{\Pi_{N|t}^{*-1} q}^2 \overset{(3.14g)}{=} ||x_t^{s,*}||_{\Pi_{N|t}^{*-1} q}^2 \overset{(A.12)}{\leq} ||x_t^{s,*}||_{\Pi_t^{-1} q}^2 = ||x_t^{s,*} - x_t + x_t||_{\Pi_t^{-1} q}^2 \\
&\leq 2||x_t^{s,*} - x_t||_{\Pi_t^{-1} q}^2 + 2||x_t||_{\Pi_t^{-1} q}^2 \leq 2||x_t - x_t^{s,*}||_{\Pi_0^{-1} q}^2 + 2||x_t||_{\Pi_t^{-1} q}^2 .
\end{aligned}
\tag{4.31}
$$

Finally, we arrive to our upper bounded Lyapunov decrease function by using Assumption 4:

$$
\begin{aligned}
&J_{t+1}^* - J_t^* \leq J_{t+1} - J_t^* \overset{(4.31)}{\leq} 2||x_t - x_t^{s,*}||^2_{\Pi_0^{-1}q} + 2||x_t||^2_{\Pi_t^{-1}q} - ||x_t||^2_{\Pi_t^{-1}q} - ||x_t - x_t^{s,*}||^2_{Q_0} \\
&- ||x_t - x_t^{s,*}||_{q\bar{\theta}_t\bar{\theta}_t^\top} - ||u_t - u_t^{s,*}||^2_R + \sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&+ \sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} + ||\Delta\bar{\theta}_t x_t^{s,*}||^2_{T_c} + 2\Delta\bar{\theta}_t^T x_t^{s,*} T_c(y_t^{s,*} - y^d) \\
&\overset{(3.16)}{\leq} ||x_t||^2_{\Pi_t^{-1}q} + \sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&+ \sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} + ||\Delta\bar{\theta}_t x_t^{s,*}||^2_{T_c} + 2\Delta\bar{\theta}_t^T x_t^{s,*} T_c(y_t^{s,*} - y^d) \\
&- ||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}t}} - ||u_t - u_t^{s,*}||^2_R,
\end{aligned}
\tag{4.32}
$$

with $Q_{\bar{\theta}t} = \frac{1}{2}Q_0 + q\bar{\theta}_t\bar{\theta}_t^\top$. □

This implies, that the upper bound of the average Lyapunov decrease function is

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=0}^{T} J_{t+1}^* - J_t^* \leq \frac{1}{T}\sum_{t=0}^{T} J_{t+1} - J_t^* \\
&\leq \frac{1}{T}\sum_{t=0}^{T} ||x_t||^2_{\Pi_t^{-1}q} + \frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&+ \frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} + \frac{1}{T}\sum_{t=0}^{T} ||\Delta\bar{\theta}_t x_t^{s,*}||^2_{T_c} \\
&+ \frac{1}{T}\sum_{t=0}^{T} 2\Delta\bar{\theta}_t^T x_t^{s,*} T_c(y_t^{s,*} - y^d) - \frac{1}{T}\sum_{t=0}^{T}(||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} \\
&+ ||u_t - u_t^{s,*}||^2_R).
\end{aligned}
\tag{4.33}
$$

4.4 Proof of the Equation (4.2)

After establishing an upper bound for the average Lyapunov decrease function, the next step is to show the convergence of these terms.
For the proof of Equation (4.2), we use the following two lemmas.

Lemma 4. *The following terms converge to zero as time approaches infinity:*

$$
\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} ||x_t||^2_{\Pi_t^{-1}q} + \frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&+ \frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1} ||x_{k|t}^* - x_t^{s,*}||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} + \frac{1}{T}\sum_{t=0}^{T} ||\Delta\bar{\theta}_t x_t^{s,*}||^2_{T_c} \\
&+ \frac{1}{T}\sum_{t=0}^{T} 2\Delta\bar{\theta}_t^T x_t^{s,*} T_c(y_t^{s,*} - y^d) = 0.
\end{aligned}
\tag{4.34}
$$

Lemma 5. *There exists a constant $c > 0$ sufficiently small, such that the following holds:*

$$\frac{1}{T}(\sum_{t=0}^{T} ||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R) \geq \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}} \tag{4.35}$$

Lemma 4 indicates the average convergence of the terms appearing in (4.33), while Lemma 5 establishes a relation between x_t and x_t^d, as x_t^d does not appear in Equation (4.33). This is important, since the goal is not only to achieve convergence of (4.33) as time approaches infinity but also to show (4.2).
Compared to the solution of [4], where case distinction analysis is done based on the distance between x_t, $x_t^{s,*}$ and x_t^d, in this thesis, we assume:

Assumption 5. *There exists a constant c_0 such that the following inequality holds for all $t \geq 0$:*

$$||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R \geq c_0||x_t^{s,*} - x_t^d||^2_{Q_{\bar{\theta}_t}}, \tag{4.36}$$

with c_0 sufficiently small.

This assumption is motivated by the results in [7], where they have shown in a different but similar scenario that the case

$$||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R \leq c_0||x_t^{s,*} - x_t^d||^2_{Q_{\bar{\theta}_t}} \tag{4.37}$$

does not exist. We do not provide the mathematical verification for this assumption within this thesis and leave it for future works. However, due to the similarity between the scenarios, we utilize this assumption to prove Lemma 5.

Proof Lemma 4. We start with the convergence analysis of the first term in Lemma 4. Using Lemma 2, we can directly prove that:

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}||x_t||^2_{\Pi_t^{-1}q} = \lim_{T\to\infty}\frac{q}{T}(\sum_{t=1}^{T}||x_t||^2_{\Pi_t^{-1}} + ||x_0||^2_{\Pi_0^{-1}}) \\
&\overset{(4.8)}{\leq} \lim_{T\to\infty}\frac{q\sigma^2 n}{T}\log(\frac{T\sigma^{-2}(x_{max})^2+\lambda}{\lambda}) + \frac{1}{T}||x_0||^2_{\Pi_0^{-1}} = 0.
\end{aligned} \tag{4.38}$$

The reasoning behind the last equality is that the growth rate of T is bigger than $\log(T)$.
For the second term in (4.34), we first utilize the boundedness of the estimated parameter and state and reformulate it to:

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1}||x^*_{k|t} - x_t^{s,*}||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&= \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1}2q||(\Delta\bar{\theta}_t\bar{\theta}_t^\top)^{\frac{1}{2}}(x^*_{k|t} - x_t^{s,*})||^2_2 \\
&\overset{(A.7,A.9)}{\leq} \lim_{T\to\infty}\frac{2q}{T}\sum_{t=0}^{T}\sum_{k=1}^{N-1}||x^*_{k|t} - x_t^{s,*}||^2_2||\bar{\theta}_t||_2||\Delta\bar{\theta}_t||_2 \\
&= \lim_{T\to\infty}\frac{2q}{T}\sum_{t=0}^{T}||\bar{\theta}_t||_2||\Delta\bar{\theta}_t||_2\sum_{k=1}^{N-1}||x^*_{k|t} - x_t^{s,*}||^2_2 \\
&\leq \lim_{T\to\infty}\frac{2q}{T}\sum_{t=0}^{T}\theta_{max}||\Delta\bar{\theta}_t||_2(N-1)\Delta x_{max} \\
&= \lim_{T\to\infty}\frac{2q(N-1)\theta_{max}\Delta x_{max}}{T}\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2
\end{aligned} \tag{4.39}$$

where $\theta_{max} = \max_{\theta \in \Theta_0} ||\theta||_2$ and $\Delta x_{max} = \max_{x \in \mathbf{X}} ||x^*_{k|t} - x^{s,*}_t||_2^2$. Then average convergence of the second term in (4.34) is established with Lemma 1 and Appendix A.3:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k=1}^{N-1} ||x^*_{k|t} - x^{s,*}_t||^2_{2q\Delta\bar{\theta}_t\bar{\theta}_t^\top} \\
&\leq \lim_{T\to\infty} \frac{2q(N-1)\theta_{max}\Delta x_{max}}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2 ||1||_2 \\
&\overset{(A.6)}{\leq} \lim_{T\to\infty} \frac{2q(N-1)\theta_{max}\Delta x_{max}}{T} (\sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2)^{\frac{1}{2}} (\sum_{t=0}^{T} ||1||_2^2)^{\frac{1}{2}} \\
&= \lim_{T\to\infty} \frac{2q(N-1)\theta_{max}\Delta x_{max}}{T} (\sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2)^{\frac{1}{2}} (T+1)^{\frac{1}{2}} \\
&\overset{(4.3)}{\leq} \lim_{T\to\infty} \frac{2q(N-1)\sigma^{-1} n^{\frac{1}{2}} \theta_{max}\Delta x_{max} e^{\frac{1}{2}}_{max} (T+1)^{\frac{1}{2}}}{T\lambda^{\frac{1}{2}}} \\
&(\log(\frac{(T+1)\sigma^{-2}(x_{max})^2 + \lambda}{\lambda}))^{\frac{1}{2}} \\
&= 0,
\end{aligned} \tag{4.40}$$

The inequality holds as T grows faster than $(T+1)^{\frac{1}{2}} \log(T)$ when T approaches infinity. Utilizing the boundedness argument similar to the analysis of the second term and (4.1), we show average convergence for the third term in (4.34) by following:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k=1}^{N-1} ||x^*_{k|t} - x^{s,*}_t||^2_{q\Delta\bar{\theta}_t\Delta\bar{\theta}_t^\top} \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k=1}^{N-1} q||\Delta\bar{\theta}_t^\top (x^*_{k|t} - x^{s,*}_t)||_2^2 \\
&\overset{(A.7)}{\leq} \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k=1}^{N-1} q||x^*_{k|t} - x^{s,*}_t||_2^2 ||\Delta\bar{\theta}_t||_2^2 \\
&= \lim_{T\to\infty} \frac{q}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 \sum_{k=1}^{N-1} ||x^*_{k|t} - x^{s,*}_t||_2^2 \\
&\leq \lim_{T\to\infty} \frac{q(N-1)\Delta x_{max}}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 \\
&\overset{(4.1)}{=} 0,
\end{aligned} \tag{4.41}$$

where we utilize Equation (4.1) for the convergence criteria. The convergence of the fourth term in (4.34) is established using (4.1) as follows:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t x^{s,*}_t||^2_{T_c} \overset{(A.7)}{\leq} \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} T_c ||x^{s,*}_t||_2^2 ||\Delta\bar{\theta}_t||_2^2 \\
&\leq \lim_{T\to\infty} \frac{T_c x_{max}}{T} \sum_{t=0}^{T} ||\Delta\bar{\theta}_t||_2^2 \overset{(4.1)}{=} 0
\end{aligned} \tag{4.42}$$

Finally, the last term in (4.34) converges using a similar method as the second term:

$$
\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}2\Delta\bar{\theta}_t^\top x_t^{s,*}T_c(y_t^{s,*}-y^d)\\
&\le \lim_{T\to\infty}\frac{2T_c}{T}\sum_{t=0}^{T}||\Delta\bar{\theta}_t^\top x_t^{s,*}||_2||(y_t^{s,*}-y^d)||_2\\
&\le \lim_{T\to\infty}\frac{2T_c}{T}\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2||x_t^{s,*}||_2\Delta y_{max}\\
&\le \lim_{T\to\infty}\frac{2T_c}{T}\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2 x_{max}\Delta y_{max}\\
&= \lim_{T\to\infty}\frac{2T_c x_{max}\Delta y_{max}}{T}\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2||1||_2\\
&\overset{(A.7)}{\le} \lim_{T\to\infty}\frac{2T_c x_{max}\Delta y_{max}}{T}(\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2^2)^{\frac{1}{2}}(\sum_{t=0}^{T}||1||_2^2)\frac{1}{2}\\
&= \lim_{T\to\infty}\frac{2T_c x_{max}\Delta y_{max}}{T}(\sum_{t=0}^{T}||\Delta\bar{\theta}_t||_2^2)^{\frac{1}{2}}(T+1)^{\frac{1}{2}}\\
&\overset{(4.3)}{\le} \lim_{T\to\infty}\frac{2T_c\sigma^{-1}n^{\frac{1}{2}}x_{max}\Delta y_{max}e_{max}^{\frac{1}{2}}(T+1)^{\frac{1}{2}}}{T\lambda^{\frac{1}{2}}}\\
&(\log(\frac{(T+1)\sigma^{-2}(x_{max})^2+\lambda}{\lambda}))^{\frac{1}{2}}=0,
\end{aligned}
\tag{4.43}
$$

where $\Delta y_{max}=\max_{y\in\mathbf{Y}}||y_t^{s,*}-y^d||_2$ holds due to the boundedness of the output. The last inequality holds for the same reason as in (4.40). Thus, the convergence of all terms in (4.34) is established, and the proof of Lemma 4 is complete. □

With the remaining term, which has not yet been used in (4.33):

$$
-\frac{1}{T}(\sum_{t=0}^{T}||x_t-x_t^{s,*}||_{Q_{\bar{\theta}_t}}^2+||u_t-u_t^{s,*}||_R^2), \tag{4.44}
$$

we now provide the proof for Lemma 5.

Proof Lemma 5: With Assumption 5 we show the following reformulation:

$$
\begin{aligned}
&||x_t-x_t^{s,*}||_{Q_{\bar{\theta}_t}}^2+||u_t-u_t^{s,*}||_R^2\\
&\overset{(4.36)}{\ge}\frac{1}{2}(||x_t-x_t^{s,*}||_{Q_{\bar{\theta}_t}}^2+||u_t-u_t^{s,*}||_R^2)+\frac{c_0}{2}||x_t^{s,*}-x_t^d||_{Q_{\bar{\theta}_t}}^2\\
&\ge\frac{1}{2}||x_t-x_t^{s,*}||_{Q_{\bar{\theta}_t}}^2+\frac{c_0}{2}||x_t^{s,*}-x_t^d||_{Q_{\bar{\theta}_t}}^2\\
&\ge\min\{\frac{1}{4},\frac{c_0}{4}\}||x_t-x_t^d||_{Q_{\bar{\theta}_t}}^2,
\end{aligned}
\tag{4.45}
$$

where we used the fact that $\frac{1}{2}(a+b)^2 \leq (a^2+b^2)$ for a, $b \in \mathbb{R}$ and $c_0 > 0$. Then we have

$$\begin{aligned} &\frac{1}{T}(\sum_{t=0}^{T} ||x_t - x_t^{s,*}||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R) \geq \frac{1}{T}\sum_{t=0}^{T} \min\{\frac{1}{4}, \frac{c_t}{4}\}||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}} \\ &\geq \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}}, \end{aligned} \tag{4.46}$$

where $c = \min\{\frac{1}{4}, \frac{c_0}{4}, ..., \frac{c_T}{4}\}$. □

With the introduction of Lemma 4, Lemma 5, and (4.33), we now provide the proof of (4.2).

Proof 4.2: One can observe that as time approaches infinity, we have the following relation using Lemma 4 and Lemma 5:

$$\begin{aligned} &\lim_{T\to\infty} \frac{-J_0}{T} \leq \lim_{T\to\infty} \frac{J^*_{T+1} - J^*_0}{T} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} J^*_{t+1} - J^*_t \leq \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} J_{t+1} - J^*_t \\ &\overset{(4.34)}{\leq} (-1)\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T}(||x_t^{s,*} - x_t||^2_{Q_{\bar{\theta}_t}} + ||u_t - u_t^{s,*}||^2_R) \\ &\overset{(4.35)}{\leq} (-1)\lim_{T\to\infty} \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}} \end{aligned} \tag{4.47}$$

Especially, we have the relation

$$\lim_{T\to\infty} \frac{1}{T} J_0 \geq \lim_{T\to\infty} \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}}, \tag{4.48}$$

with J_0 finite, implies that the term

$$\lim_{T\to\infty} \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}} = 0 \tag{4.49}$$

converges on average. Furthermore, for our average Lyapunov decrease function, this means:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} J^*_{t+1} - J^*_t \leq (-1)\lim_{T\to\infty} \frac{c}{T}\sum_{t=0}^{T} ||x_t - x_t^d||^2_{Q_{\bar{\theta}_t}} = 0, \tag{4.50}$$

which concludes the proof of the Equation (4.2) and thus the proof for Theorem 1. □

Chapter 5

Simulation

In this chapter, we present different scenario examples of our proposed dual adaptive MPC with simulations. The simulations were implemented in MATLAB, and the optimization problem of the dual adaptive MPC was solved using the CasADi solver.

5.1 Simulation System Model

For all simulations, we use the following discrete linear time-invariant system:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0.5 \\ 1 & 0.5 \end{bmatrix},$$

The real parameter vector in the output system is:

$$\theta_{true} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Furthermore, the steady states and inputs are characterized by the vector $\Psi \in \mathbf{R}^2$, which is optimized by the dual adaptive MPC, and the matrix

$$M_\Psi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & -2 \end{bmatrix}.$$

The predefined hyperbox set is chosen to be $\Theta_0 = [1.4, 2.6] \times [0.7, 1.3]$. The MPC has a fixed horizon of $N = 20$. In all simulations, the initial covariance matrix is $\Pi_0^{-1} = \frac{1}{10} I_2$, and we choose the following cost matrices for the cost function: $Q_0 = I_2$, $R = I_2$, $q = 1$ and $T_c = 1$. By the choice of Q_0 and q Assumption 4 is satisfied.

5.2 Active Exploration Example without Noise

5.2.1 Example with Constant Desired Output

In this example we demonstrate the exploration and exploitation effects of our dual adaptive MPC scheme. The prior distribution of the unknown parameter is set to have a mean of $\bar{\theta}_0 = (2.4, 0.8)^\top$. The system is noisefree, and to prevent numerical issues in (3.10), we set $\sigma = \sqrt{\frac{2}{3}}$. The desired output is set as $y^d = 2$, and the initial state is $x_0 = (0.5, 1)^\top$. The constraints for the states, inputs and outputs are intentionally set to relatively loose standards, as they are not critical for the outcome of this simulation. We compare the performance of the dual adaptive MPC with the certainty-equivalent MPC (CE-MPC), which uses the same constraints and costs as the dual adaptive MPC except the learning cost is omitted.

As illustrated in Figure 5.1, even when the initial output matches the desired output and the initial estimated parameter matches the true parameter, the output still deviates from the desired output. This deviation arises from the cumulative learning cost $\sum_{k=0}^{N-1} ||x_{k|t}||_{\Pi_{k|t}^{-1}q}$, which encourages the system to explore actively and improves the parameter estimate as shown in Figure 5.2. In contrast, the CE-MPC maintains the system at the desired output. As the simulation time progresses, the dual adaptive MPC attempts to steer the output back to the desired output, demonstrating effective tracking. The optimal cost of the dual adaptive MPC converges, as illustrated in Figure 5.3.

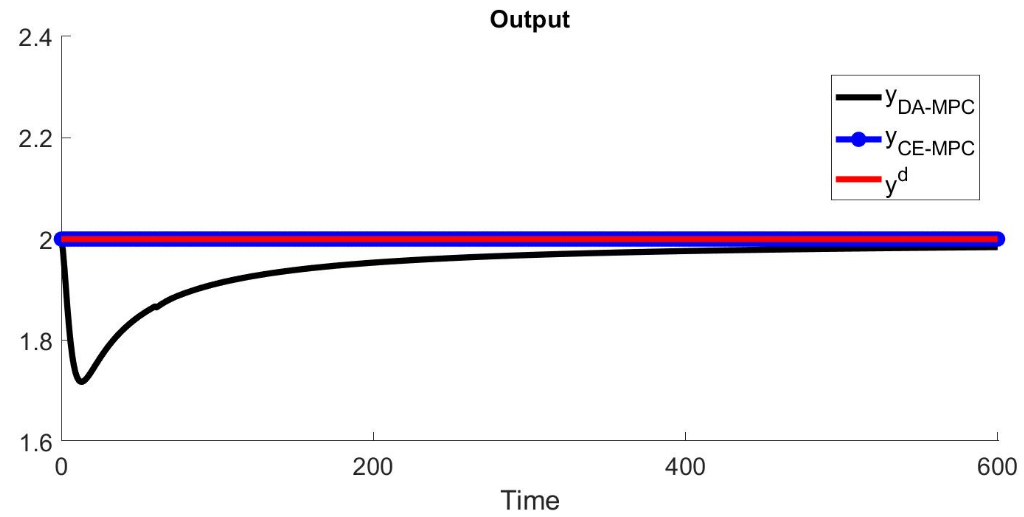

Figure 5.1: The output shows the exploration effect of the dual adaptive MPC compared to the certainty-equivalent MPC in Example 5.2.1.

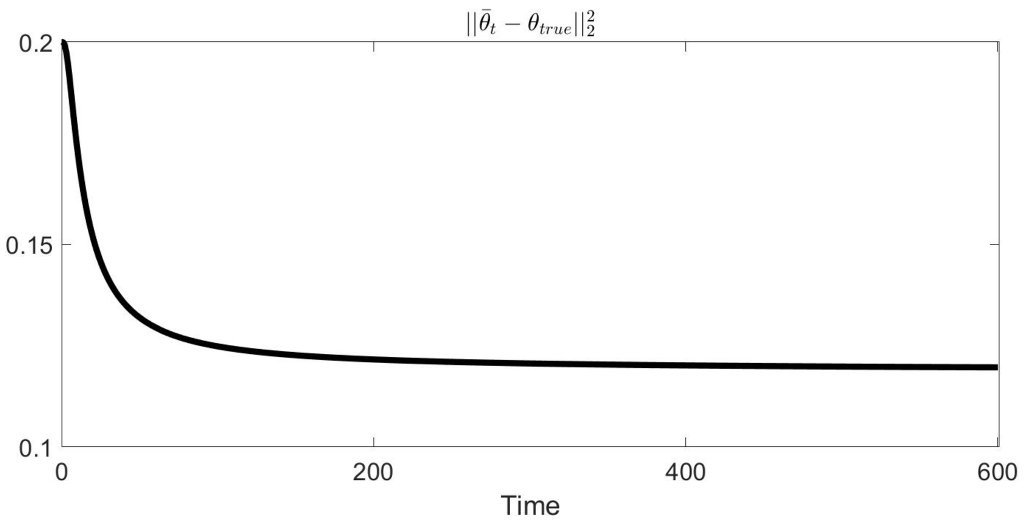

Figure 5.2: Error of the parameter estimation improves through exploration in Example 5.2.1.

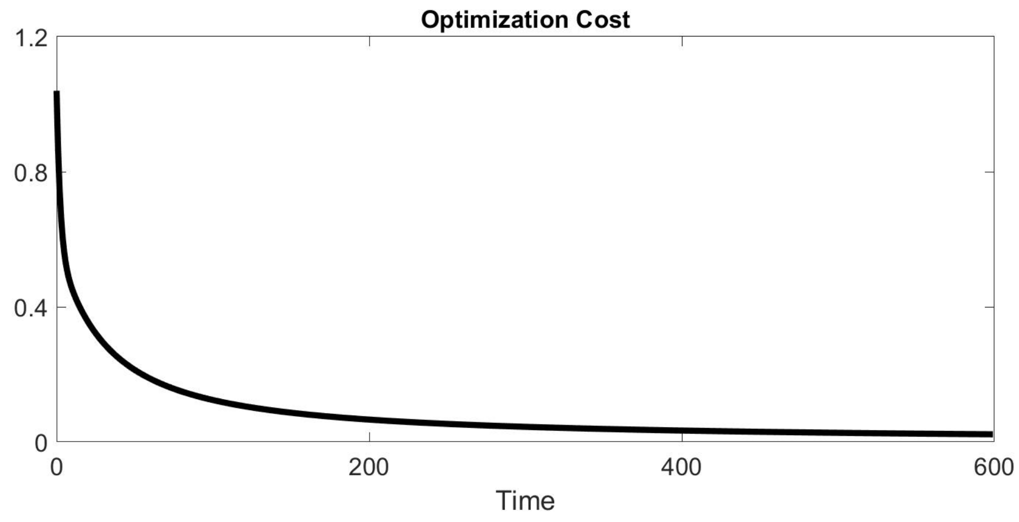

Figure 5.3: The cost of the dual adaptive MPC converges in Example 5.2.1

5.2.2 Example with Changing Desired Output

Next, we repeat the simulation with the same settings but the desired output is changed to $y^d = 0.5$ at $t = 200$. This allows us to observe how the dual adaptive MPC behaves after undergoing active exploration in response to the change in desired output compared to the CE-MPC.

Figure 5.4 shows that through exploration, the dual adaptive MPC adapts faster to the new desired output compared to the CE-MPC. Due to the change in desired output, the CE-MPC undergoes a larger correction in the estimation process as the experienced error is larger than the dual adaptive MPC. As a result, it achieves better overall parameter estimation than the dual adaptive MPC, as illustrated in Figure 5.5.

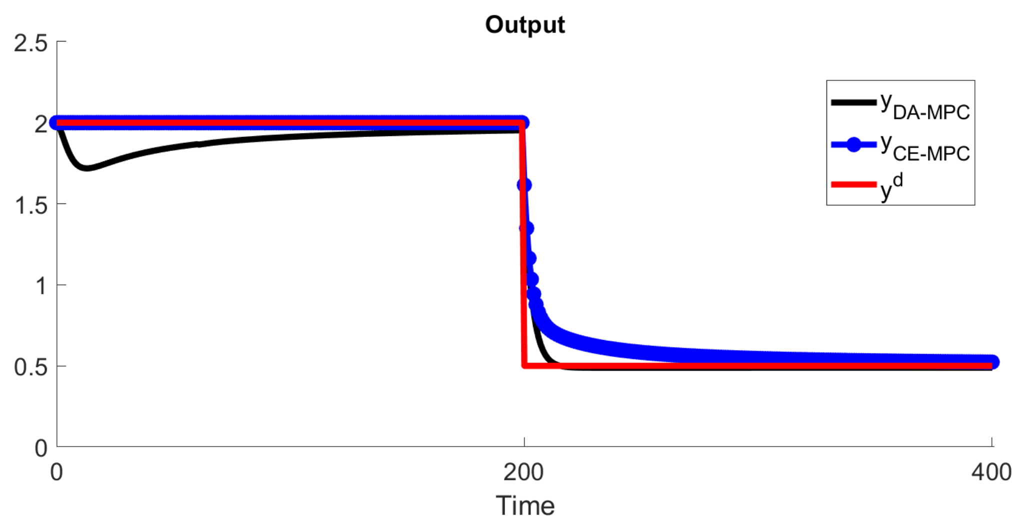

Figure 5.4: The change of the desired output demonstrates improved tracking with the dual adaptive MPC in Example 5.2.2.

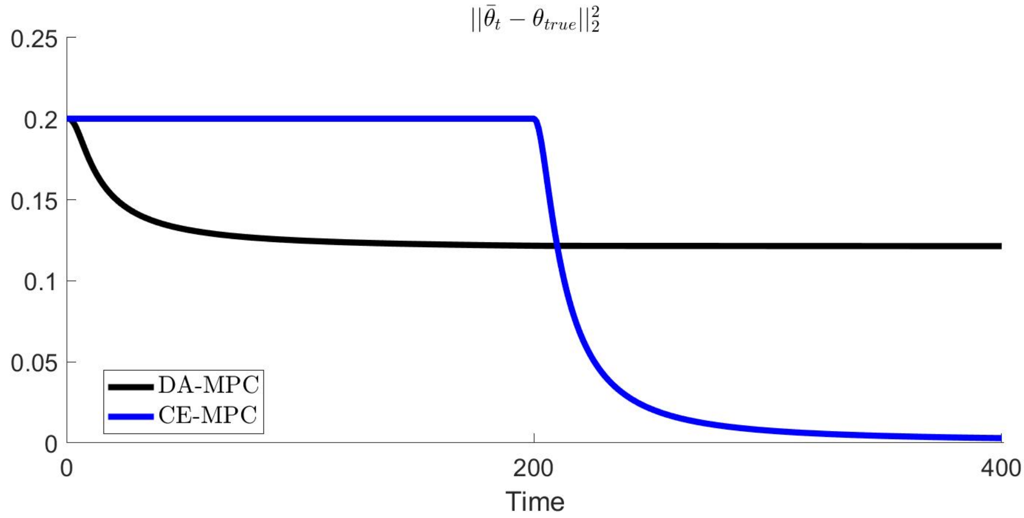

Figure 5.5: The CE-MPC achieves better parameter estimation compared to the dual adaptive MPC following the change in desired output in Example 5.2.2.

5.3 Convergence Example with Truncated Gaussian Noise

In this example, we simulate the system where the output is disturbed by a truncated Gaussian noise w_t within the range $[-0.2, 0.2]$ with a standard deviation of $\sigma = \frac{1}{\sqrt{5}}$ and zero mean. The desired output is set to $y^d = 2$, the initial state for this example is set to $x_0 = (0.5, 0.5)^\top$, and the initial estimate is set to $\bar{\theta}_0 = (2.4, 0.8)^\top$ This results in a initial output y_0 that differs from the desired output. The state is constrained to a two-dimensional hyperbox ranging from -2 to 2 along each dimension, while the input is restricted to a two-dimensional hyperbox ranging from -4 to 4. The output is constrained within the interval $[-3.5, 3.5]$. The simulation is repeated 10 times.

Figure 5.6 illustrates the convergence of the averaged output to the desired output. Moreover, the convergence of the parameter estimate despite noise in one of the simulations is illustrated in Figure 5.7. It is important to highlight that the converged value differs from the true value of the estimate due to the system not being sufficient excited, even in instances where the initial output differs from the desired output.

5.4 Example of Unreachable Desired Output

This simulation illustrates a scenario in which our dual adaptive MPC encounters difficulties to reach the desired output. For this purpose, we set the desired output to $y^d = 2.5$, the initial state to $x_0 = (0.5, 0.5)^\top$, and the initial estimate to $\bar{\theta}_0 = (2.4, 0.8)^\top$. The state is constrained within a two-dimensional hyperbox ranging from -2 to 2 along each dimension, while the input is restricted within a two-dimensional hyperbox ranging from -8 to 8. The output is constrained within the interval $[-3, 3]$. To clearly illustrate the result, we do not introduce noise into the output, but we chose $\sigma = 1$ to avoid numerical issues.

We observe that the dual adaptive MPC is unable to reach the desired output, as illustrated in Figure 5.8. Furthermore, the estimated parameter vector converges to the vector $\bar{\theta}_t \approx (2.13, 0.7)^\top$, as illustrated in Figure 5.9. Similarly, the state x_t converges to the vector $x_t \approx (0.98, 0.35)^\top$ in Figure 5.10. Despite the lack of tight input constraints, the state vector does not progress further towards another state, where the resulting output is closer to the desired output than the current output with the current estimated parameter vector. This issue arises from

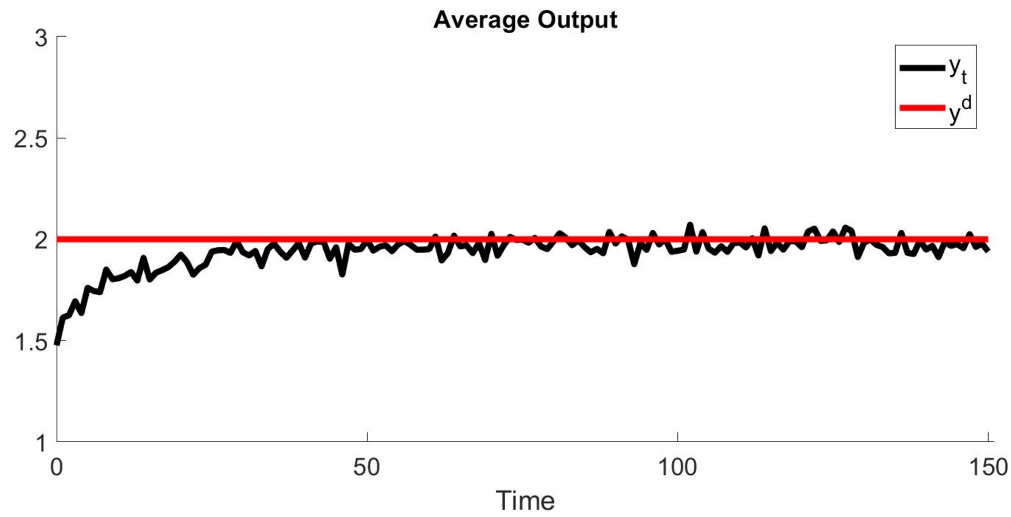

Figure 5.6: The average output converges to the desired output with respect to the current noise in Example 5.3.

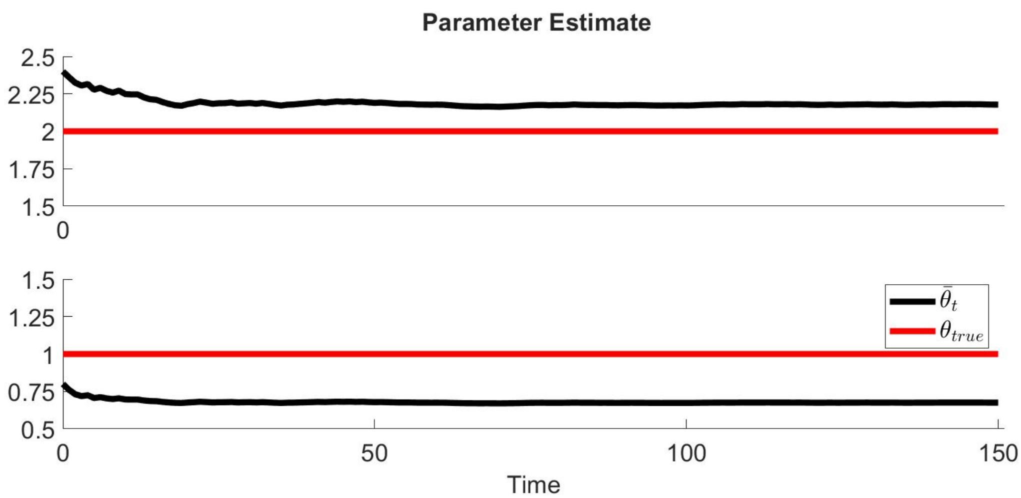

Figure 5.7: Parameter estimate converges to a different value than the real value in Example 5.3.

constraint (3.14d), where potential solutions violating (3.14d) are not considered. As a result, the dual adaptive MPC converges to the closest reachable output, but this closest reachable output does not converge to the desired output.
To illustrate this, we examine the vertex $\bar{\theta}_{1,0} = (2.6, 1.3)^\top$ of Θ_0 that $\bar{\theta}_{1,0}^\top x_t \approx 3$. Further increases in the state vector would result in violations of the constraint (3.14d). This issue can be addressed by introducing set membership estimation to the set Θ_0. This would allow for a decrease in this set over time and increase the number of feasible solutions.

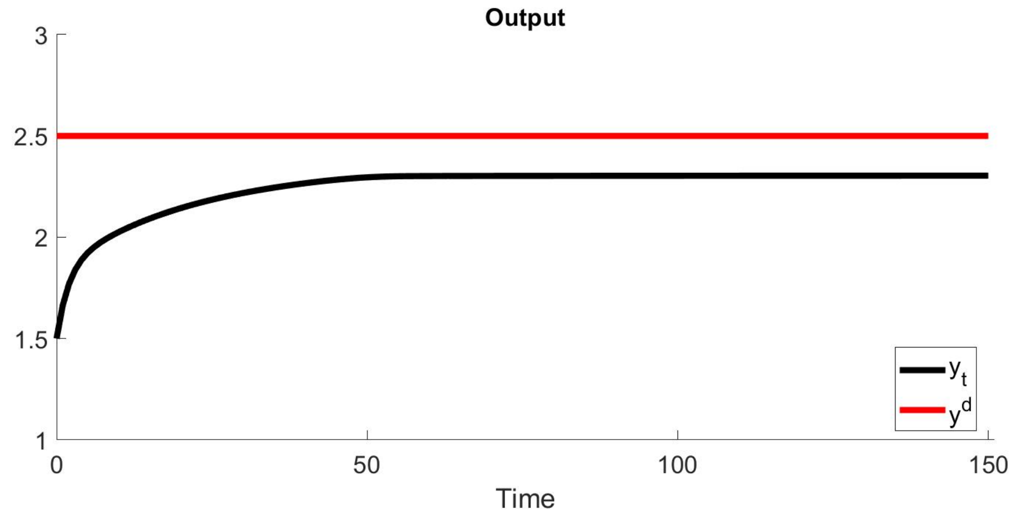

Figure 5.8: The output is unable to reach the desired output in Example 5.4.

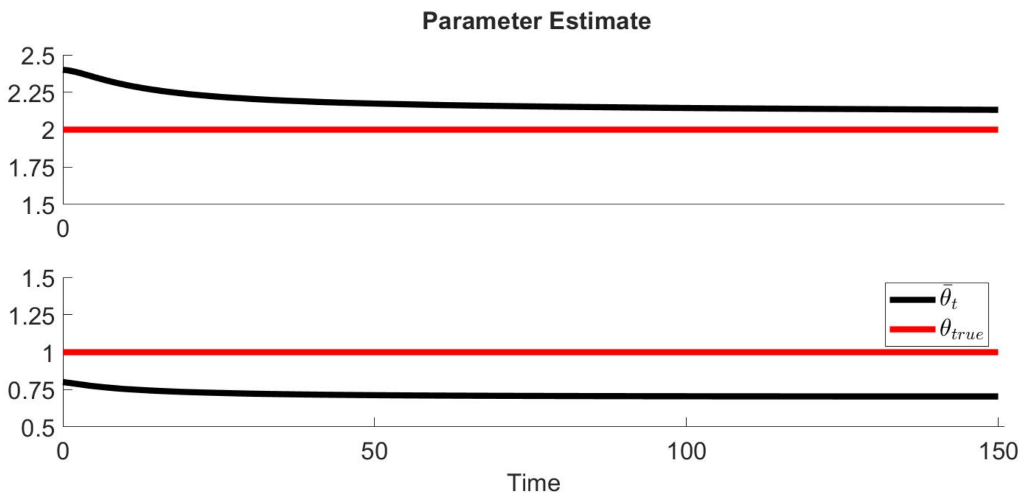

Figure 5.9: The estimated parameter converges to a an estimate, which should be able to reach the desired output in Example 5.4.

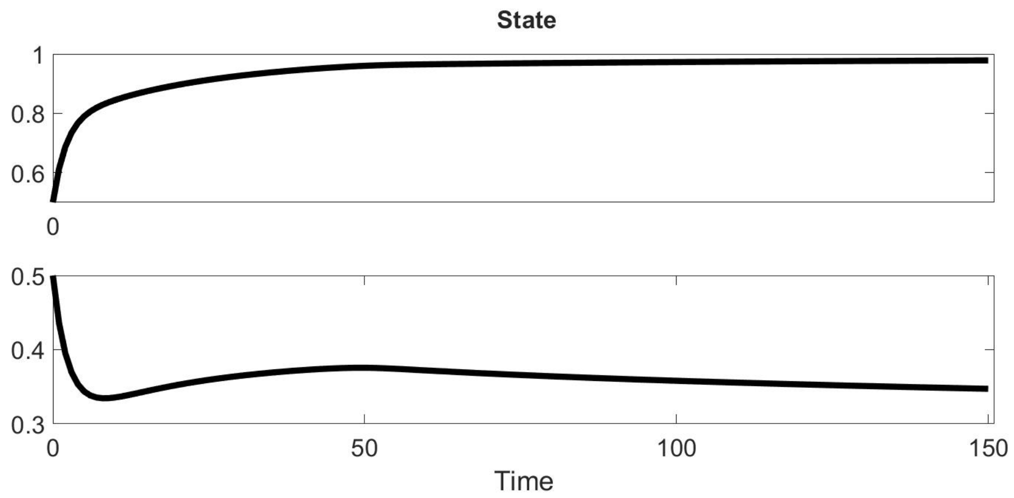

Figure 5.10: The state converges to a specific vector in Example 5.4.

Chapter 6

Conclusion

In this thesis, we investigated the convergence behavior of the dual adaptive MPC in a discrete time-invariant linear system. The goal was to improve the current state-of-the-art convergence result by analyzing the average convergence.

We introduced a theorem that ensures the average convergence of the dual adaptive MPC to the closest reachable output. This result is based on the average convergence of the parameter estimation and the average convergence of the tracking error. The average convergence of the parameter estimation implies the average convergence of the closest reachable output because the solution of the optimization problem (3.13) depends solely on the estimated parameter vector. Therefore, it is sufficient to prove the average convergence of the tracking error to the closest reachable output to confirm the overall average convergence of the dual adaptive MPC.

We have provided the mathematical proof of this theorem. The average convergence of the parameter estimation is proved by establishing a logarithmic upper bound for the cumulative sum of the differences between estimates at consecutive time steps. For the average convergence proof of the dual adaptive MPC, we defined the candidate solution as the optimal solution at the previous time instant and shifted it by one time step. Thus, we demonstrated an upper bound for the Lyapunov decrease function by considering the difference between the candidate solution and the optimal solution from the previous time instant. We first showed the average convergence of specific terms by showing their dependence on the estimated parameter. Then we established a relation between the current state and the steady state corresponding to the closest reachable output with the remaining term of the upper bound function, and proved its average convergence. To support the mathematical results, we provided simulation examples, that also illustrated the convergence behavior.

Our average convergence result provided valuable insight into the behavior of the dual adaptive MPC algorithm. In particular, we removed the convergence dependence of the covariance matrix and the estimated parameter from the result of [4]. By demonstrating the average convergence of the dual adaptive MPC, we confirm the stability and reliability of the dual adaptive MPC.

However, our result has limitations. We have not introduced the set-membership estimation in our parameter estimation. Therefore, the constraint to ensure the recursive feasibility of the dual adaptive MPC becomes too conservative, and the dual adaptive MPC may not reach the desired output. Future research should be done to develope the mathematical proof for the convergence of the dual adaptive MPC with set-membership estimation. Furthermore, the validation of Assumption 5, for which we did not provide formal proof, is another opportunity to research. Given that our learning cost is the mean of a stochastic process and does not guarantee exact parameter estimation, exploring alternative learning costs for the dual adaptive MPC could improve its overall tracking and exploration capabilities.

In conclusion, our work has successfully improved the convergence result of the dual adaptive MPC by showing average convergence without the dependence on the covariance matrix and the difference between the parameter estimates at consecutive time steps. Nevertheless, there remains potential for further investigation and enhancement of the dual adaptive MPC.

Bibliography

[1] Aleksandr Aronovich Feldbaum. "Dual control theory. I". In: *Avtomatika i Telemekhanika* 21.9 (1960), pp. 1240–1249.

[2] Tor Aksel N Heirung, B Erik Ydstie, and Bjarne Foss. "Dual adaptive model predictive control". In: *Automatica* 80 (2017), pp. 340–348.

[3] Juan E Morinelly and B Erik Ydstie. "Dual mpc with reinforcement learning". In: *IFAC-PapersOnLine* 49.7 (2016), pp. 266–271.

[4] Raffaele Soloperto, Johannes Köhler, Matthias A Müller, and Frank Allgöwer. "Dual adaptive MPC for output tracking of linear systems". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, pp. 1377–1382.

[5] Daniel Limón, Ignacio Alvarado, Teodoro Alamo, and Eduardo F Camacho. "MPC for tracking piecewise constant references for constrained linear systems". In: *Automatica* 44.9 (2008), pp. 2382–2387.

[6] Anastasios Tsiamis and George J Pappas. "Online learning of the kalman filter with logarithmic regret". In: *IEEE Transactions on Automatic Control* 68.5 (2022), pp. 2774–2789.

[7] Raffaele Soloperto, Johannes Köhler, and Frank Allgöwer. "A nonlinear MPC scheme for output tracking without terminal ingredients". In: *IEEE Transactions on Automatic Control* 68.4 (2022), pp. 2368–2375.

[8] Peter C Young. *Recursive estimation and time-series analysis: an introduction*. Springer science & business media, 2012.

[9] Ankit Goel, Adam L Bruce, and Dennis S Bernstein. "Recursive least squares with variable-direction forgetting: Compensating for the loss of persistency [lecture notes]". In: *IEEE Control Systems Magazine* 40.4 (2020), pp. 80–102.

[10] Zicheng Cai and B Erik Ydstie. "Dual adaptive model predictive control with disturbances". In: *IFAC-PapersOnLine* 54.3 (2021), pp. 206–211.

[11] MATLAB. *version 9.10.0.1602886 (R2021a)*. Natick, Massachusetts: The MathWorks Inc., 2021.

[12] Joel A E Andersson, Joris Gillis, Greg Horn, James B Rawlings, and Moritz Diehl. "CasADi – A software framework for nonlinear optimization and optimal control". In: *Mathematical Programming Computation* 11.1 (2019), pp. 1–36.

[13] Arcadii Z Grinshpan. "Weighted inequalities and negative binomials". In: *Advances in Applied Mathematics* 45.4 (2010), pp. 564–606.

[14] Eduardo H Zarantonello. "Projections on convex sets in hilbert space and spectral theory: Part i. projections on convex sets: Part ii. spectral theory". In: *Contributions to nonlinear functional analysis*. Elsevier, 1971, pp. 237–424.

Appendix A

Appendix

A.1 Inner Product in Euclidean Space

The inner product in Euclidean space of a vector $u \in \mathbb{R}^n$ is defined by:

$$\langle u, u \rangle := \sum_{k=1}^{n} u_k^2 = ||u||_2^2, \tag{A.1}$$

where u_k is the k-th entry of the vector u, and $||\cdot||_2$ is the Euclidean l_2 norm. The inner product in Euclidean space of two vectors u and $v \in \mathbb{R}^n$ is defined by:

$$\langle u, v \rangle := \sum_{k=1}^{n} u_k v_k. \tag{A.2}$$

The squared magnitude of the inner product of u and v is given by:

$$|\langle u, v \rangle|^2 := |\sum_{k=1}^{n} u_k v_k|^2 = (\sum_{k=1}^{n} u_k v_k)^2 \tag{A.3}$$

A.2 Hoelder's Inequality

Hoelder's inequality is stated as follows: Let (S, Σ, μ) be a measure space, and let $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then, for all measurable real- or complex-valued functions f and g on S,

$$||fg||_1 \leq ||f||_p ||g||_q. \tag{A.4}$$

With Hoelder's inequality, we state the following inequality for the inner product of two vectors $u, v \in \mathbb{R}^n$:

$$\left| \sum_{k=1}^{n} u_k v_k \right| \leq \left(\sum_{k=1}^{n} |u_k|^p \right)^{1/p} \left(\sum_{k=1}^{n} |v_k|^q \right)^{1/q}. \tag{A.5}$$

A.3 Cauchy-Schwartz Inequality

The Cauchy-Schwartz inequality can be derived from Hoelder's inequality with $p, q = 2$, and for two vectors u, $v \in \mathbb{R}^n$ is defined by:

$$|\langle u, v \rangle|^2 \overset{(A.3)}{=} (\sum_{k=1}^{n} u_k v_k)^2 \leq (\sum_{k=1}^{n} |u_k v_k|)^2 \overset{(A.5)}{\leq} ||u||_2^2 \cdot ||v||_2^2. \tag{A.6}$$

With (A.6), we write the following inequality:

$$||u^T v||_2^2 = \sum_{k=1}^{n}(u_k v_k)^2 \leq (\sum_{k=1}^{n} u_k v_k)^2 \overset{(A.6)}{\leq} ||u||_2^2 \cdot ||v||_2^2. \tag{A.7}$$

A.4 Sub-Multiplicative Property of the Spectral Norm

With the definition of the spectral norm:

$$||A||_2 := \max_{x \neq 0} \frac{||Ax||_2}{||x||_2} \geq \frac{||Ax||_2}{||x||_2}, \tag{A.8}$$

we have the following inequality:

$$||A||_2 ||x||_2 \geq ||Ax||_2. \tag{A.9}$$

A.5 Relation of the Covariance Matrix

By definition, the initial given covariance matrix Π_0^{-1} is positive-definite. Furthermore, the inverse is uniquely defined for a positive-definite matrix and is also positive-definite by the eigenvalue decomposition:

$$\Pi_t^{-1} = U\Sigma_t^{-1}U^\top \rightarrow \Pi_t = U\Sigma_t U^\top, \tag{A.10}$$

where Σ_t is the diagonal matrix with the eigenvalues of Π_t as the diagonal elements and U the matrix containing the eigenvector. From (3.10), we have:

$$\Pi_t \succeq \Pi_{t-1} \succeq \Pi_0 \quad \forall t \geq 1, \tag{A.11}$$

since the outer product $x_t x_t^T$ from (3.10) is positive semi-definite. And with Π_0 being positive-definite, resulting Π_t is also positive-definite for all $t \geq 0$. Hence, the covariance matrix has the following relation:

$$\Pi_0^{-1} \succeq \Pi_{t-1}^{-1} \succeq \Pi_t^{-1} \tag{A.12}$$

For the determinant of the covariance matrices, the following relation also holds:

$$\begin{aligned} &\det(\Pi_t) = \det(\Pi_t - \Pi_{t-1} + \Pi_{t-1}) \geq \det(\Pi_t - \Pi_{t-1}) + \det(\Pi_{t-1}) \\ &\geq \det(\Pi_{t-1}), \end{aligned} \tag{A.13}$$

where we used the sum rule of positive semi-definite matrices and that $\Pi_t - \Pi_{t-1}$ is positive semi-definite.

A.6 Proof of the Non-Expansiveness of the Projection

In this section, we exploit the non-expansiveness of the projection with weighted norm.

Proof. First, we define relation between the inner product and the squared weighted norm of a vector $x \in \mathbb{R}^n$ as follows:

$$\begin{aligned} \langle x, x \rangle &= x^\top \Pi x = ||x||_\Pi^2 \\ \langle x, y \rangle &= x^\top \Pi y \leq ||x||_\Pi ||y||_\Pi, \end{aligned} \tag{A.14}$$

with Π positive-definite. The weighted projection is defined as:

$$P(x) = \arg\min_{P(x) \in C} ||x - x_c||_\Pi^2, \tag{A.15}$$

with C as a convex set and x_c as a constant. Since the inner product and the projection use the same norm weight, with the Hilbert projection theorem, the following inequality holds for any $y \in C$:

$$\begin{aligned} \langle x_1 - P(x_1), y - P(x_1) \rangle \leq 0 \\ \langle x_2 - P(x_2), y - P(x_2) \rangle \leq 0 \end{aligned} \tag{A.16}$$

By replacing y with $P(x_2)$ in the first inequality and with $P(x_1)$ in the second inequality and sum these two inequality together results in:

$$\begin{aligned}
&\langle x_1 - P(x_1), P(x_2) - P(x_1) \rangle + \langle x_2 - P(x_2), P(x_1) - P(x_2) \rangle \leq 0 \\
&\langle x_1 - P(x_1), P(x_2) - P(x_1) \rangle - \langle x_2 - P(x_2), P(x_2) - P(x_1) \rangle \leq 0 \\
&\langle x_1 - P(x_1) - x_2 + P(x_2), P(x_2) - P(x_1) \rangle \leq 0 \\
&\langle P(x_2) - P(x_1), P(x_2) - P(x_1) \rangle - \langle x_2 - x_1, P(x_2) - P(x_1) \rangle \leq 0 \\
&\langle P(x_2) - P(x_1), P(x_2) - P(x_1) \rangle \leq \langle x_2 - x_1, P(x_2) - P(x_1) \rangle \\
&||P(x_2) - P(x_1)||_{\Pi}^2 \leq ||x_2 - x_1||_{\Pi} ||P(x_2) - P(x_1)||_{\Pi} \\
&||P(x_2) - P(x_1)||_{\Pi} \leq ||x_2 - x_1||_{\Pi},
\end{aligned} \tag{A.17}$$

which concludes the proof. □

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. In consultation with the supervisor, one of the following three options must be selected:

I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].

I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used and cited generative artificial intelligence technologies[2].

I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used generative artificial intelligence technologies[3]. In consultation with the supervisor, I did not cite them.

Title of paper or thesis:

Authored by:
If the work was compiled in a group, the names of all authors are required.

Last name(s): **First name(s):**

With my signature I confirm the following:

- I have adhered to the rules set out in the Citation Guide.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

Place, date **Signature(s)**

If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.

[1] E.g. ChatGPT, DALL E 2, Google Bard
[2] E.g. ChatGPT, DALL E 2, Google Bard
[3] E.g. ChatGPT, DALL E 2, Google Bard