Addressing Adaptation and Learning in the Context of MPC with MHE

David A. Copp and João P. Hespanha

Abstract This paper considers the estimation and control of systems with parametric uncertainty. An approach that combines moving horizon estimation and model predictive control into a single min-max optimization is employed to estimate past and current values of the state, compute a sequence of optimal future control inputs, predict future values of the state, as well as estimate current values of uncertain parameters. This is done by including the state, inputs, and uncertain parameters as optimization variables. Learning the true values of the uncertain parameters requires a sufficiently large number of past measurements and that the system is persistently excited. The true values of the uncertain parameters may change over time, and the optimization computes future control inputs that adapt to changing estimates of the uncertain parameters in order to better control the uncertain system. Several linear and nonlinear examples with parametric uncertainty are discussed and effectively controlled using this combined moving horizon estimation and model predictive control approach.

1 Introduction

Having an accurate model of a system to be controlled is often vital for effective control of that system. This is certainly true for a model predictive control (MPC) approach in which a finite-horizon online optimization

David A. Copp
University of California, Santa Barbara, CA 93106 USA, e-mail: dacopp@engr.ucsb.edu

João P. Hespanha
University of California, Santa Barbara, CA 93106 USA, e-mail: hespanha@ece.ucsb.edu

problem is solved in order to determine an optimal control input given the system's dynamics and a desired control objective [18]. However, in most practical applications, there are unknown parameters in the model of a system or, at least, uncertain parameters that are known only to be within some set of values. These uncertainties may include uncertain model parameters, input disturbances, and measurement noise. Because of this, much work on MPC approaches have involved investigating robustness to model parameter uncertainty, input disturbances, and measurement noise. This work is known as robust MPC [5, 13] which also includes worst-case, or min-max MPC [10].

An attractive, and perhaps less conservative, approach to controlling systems with parameter uncertainties is to update the model of the system with new estimates of the parameters as they become available, which is the underlying idea behind indirect adaptive control (see, e.g. [3, 9]). Very little work has been done on adaptive MPC, but there are a few proposed approaches. The authors of [14] propose an adaptive MPC scheme that uses a standard estimator and certainty equivalence to update the model with the current estimates of the parameters. The authors of [1] investigate nonlinear systems that are affine with respect to unknown parameters and perform adaptive control by combining a parameter adjustment mechanism with robust MPC algorithms such as min-max MPC. A cost function is minimized with respect to feedback control policies and maximized with respect to the unknown parameters so that the MPC approach is robust to the worst-case values of the unknown parameters. For both of these approaches, it is assumed that the full state is available for feedback. This is often the case for MPC approaches in order to alleviate issues that arise from uncertainties, noise, and disturbances.

Unfortunately, in most practical applications, the full state is not known or necessarily available for feedback. Because of this, output-feedback MPC should be considered, and an independent algorithm for estimating the state is needed. A convenient estimation algorithm for use with MPC is moving-horizon estimation (MHE). MHE can be used for estimating the state of constrained nonlinear systems and similarly involves the solution of a finite-horizon online optimization problem where a criterion based on a finite number of past output measurements is minimized in order to find the best estimate of the state [2, 17]. It is straightforward to also incorporate parameter estimation into the formulation of MHE, so the state and parameters can both be estimated using the same estimator [19].

In [6], a framework for solving the output-feedback MPC problem with MHE is presented that solves both the control and estimation problems as a single min-max optimization problem. This framework already incorporates input disturbances and measurement noise. In this chapter, we further incorporate uncertain model parameters into this framework and obtain parameter estimates by including the uncertain parameters as optimization variables. In this way, we solve simultaneously the control problem and the state and parameter estimation problems, resulting in effective control of uncertain systems. Our approach can be likened to an indirect model reference adaptive control approach as described in the adaptive control literature [3, 9] in that, at each time step, new estimates of the uncertain parameters are computed and used to update the model while a new sequence of future control inputs that minimize an objective criterion is also simultaneously computed.

Because MPC and MHE involve the solution of an online optimization problem, this approach lends itself to adapting to both constant and time-varying parameters because, at each time step, a new estimate is computed, and the model can be updated accordingly. We show in examples that when the system is sufficiently excited, this MPC with MHE approach is able to learn the true values of the uncertain parameters, but even if the system is not sufficiently excited to learn the true values of the parameters, this approach finds estimates that are consistent with the dynamics and often still enables effective control and disturbance rejection.

The main assumption for this work is that a saddle-point solution exists for the min-max optimization problem at each sampling time. This assumption presumes an appropriate form of observability for the closed-loop system and is a common requirement in game theoretical approaches to control design [4]. For controllability, we additionally require that there exists a terminal cost that is an ISS-control Lyapunov function with respect to a disturbance input, which is a common assumption in MPC [15].

The rest of the chapter is organized as follows. In Section 2 we formulate the adaptive MPC with MHE problem that we would like to solve. Stability results that can be used to prove state boundedness and reference tracking are given in Section 3. In Section 4 we discuss several linear and nonlinear systems with parameter uncertainty and show that using our MPC with MHE scheme we are able to not only stabilize the system, but also estimate the correct values of the uncertain parameters. Finally, we provide some conclusions and directions for future work in Section 5.

2 Problem Formulation

In the formulation of standard MPC and MHE problems, a time-varying nonlinear discrete-time process of the form

$$x_{t+1} = f_t(x_t, u_t, d_t), \qquad y_t = g_t(x_t) + n_t, \qquad \forall t \in \mathbb{Z}_{\geqslant 0} \tag{1}$$

is considered with *state* x_t taking values in a set $\mathcal{X} \subset \mathbb{R}^{n_x}$. The inputs to this system are the *control input* u_t that must be restricted to the set $\mathcal{U} \subset \mathbb{R}^{n_u}$, the *unmeasured disturbance* d_t that is assumed to belong to the set $\mathcal{D} \subset \mathbb{R}^{n_d}$, and the *measurement noise* $n_t \in \mathbb{R}^{n_n}$. The signal $y_t \in \mathbb{R}^{n_y}$ denotes the *measured output* that is available for feedback.

In this chapter, we investigate MPC and MHE of processes with uncertain model parameters. These uncertain parameters are denoted by the vector θ whose elements are known to belong to the set $\Theta \subset \mathbb{R}^{n_\theta}$. In this formulation, the process dynamics depend explicitly on the uncertain parameter θ, so we redefine the process dynamics in (1) to include the uncertain parameters as

$$x_{t+1} = f_t(x_t, \theta, u_t, d_t), \qquad y_t = g_t(x_t, \theta) + n_t, \qquad \forall t \in \mathbb{Z}_{\geqslant 0}. \tag{2}$$

We assume that θ is a constant parameter, i.e. $\theta = \theta_t$ for all $t \in \mathbb{Z}_{\geqslant 0}$, but as will be shown later, we are still able to adapt to and learn changing parameter values.

A block diagram depicting the process (2) is shown in Figure 1.

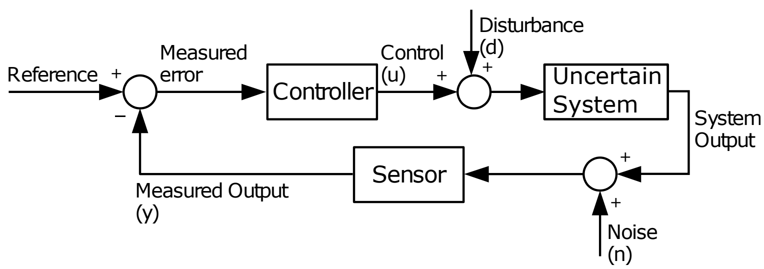

Fig. 1 Block diagram of the process given in (2).

2.1 Moving Horizon Estimation

In MHE, the current state of the system x_t at time t is estimated by solving a finite-horizon online optimization problem using a finite number of past measurements [17]. If we consider a finite horizon of L time steps, then the objective of the MHE problem is to find an estimate of the current state x_t so as to minimize a criterion of the form

$$\sum_{s=t-L}^{t} \eta_s\big(y_s - g_s(x_s)\big) + \sum_{s=t-L}^{t-1} \rho_s(d_s), \tag{3}$$

given the system dynamics (1). The functions $\eta_s(\cdot)$ and $\rho_s(\cdot)$ are assumed to take non-negative values. This is similar to the MHE criterion considered in [2, 17].

If the system dynamics also include uncertain model parameters, as in (2), the MHE problem can be formulated so as to estimate both the current state x_t and the uncertain parameter θ. Then the MHE problem can be written as

$$\min_{\hat{x}_{t-L}\in\mathcal{X},\hat{d}_{t-L:t-1}\in\mathcal{D},\hat{\theta}\in\Theta} \sum_{s=t-L}^{t} \eta_s\big(y_s - g_s(\hat{x}_s,\hat{\theta})\big) + \sum_{s=t-L}^{t-1} \rho_s(\hat{d}_s), \tag{4}$$

where the initial state x_{t-L} is constrained to belong to the set $\mathcal{X}$, each element of the input disturbance sequence $d_{t-L:t-1}$ is assumed to belong to the set $\mathcal{D}$, and the uncertain parameter θ is known to belong to the set Θ. Throughout the chapter, given two times t_1 and t_2 with $t_1 < t_2$, we use the notation $x_{t_1:t_2}$ to denote the time series $x_{t_1}, x_{t_1+1}, \ldots, x_{t_2-1}, x_{t_2}$. An estimate of the current state is then determined from the dynamics (2) given the known past control inputs applied $u_{t-L:t-1}$ and estimates of the initial state $\hat{x}_{t-L}$, the input disturbance sequence $\hat{d}_{t-L:t-1}$, and the uncertain parameter $\hat{\theta}$. The optimization (4) is re-solved at each time t in a receding horizon fashion.

2.2 Model Predictive Control

In MPC, a sequence of future control inputs that achieve a desired control objective is computed by solving a finite-horizon online optimization problem using an estimate of the current state $\hat{x}_t$ and the system dynamics [18]. If we consider a finite-horizon of T time steps, then the objective of the MPC problem is to find a sequence of future control inputs $u_{t:t+T-1}$ that minimizes a criterion of the form

$$\sum_{s=t}^{t+T-1} \big(c_s(x_s,u_s,d_s) - \rho_s(d_s)\big) + q_{t+T}(x_{t+T}), \tag{5}$$

given the system dynamics (1). The functions $c_s(\cdot)$, $\rho_s(\cdot)$, and $q_{t+T}(\cdot)$ are all assumed to take non-negative values. The negative sign in front of $\rho_t(\cdot)$ penalizes the maximizer for using large values of d_t. The function $q_{t+T}(x_{t+T})$ is a terminal cost that penalizes the "final" state x_{t+T} and is needed for proving stability (see Assumption 3 below). The criterion (5) is similar to the closed-loop min-max MPC criterion considered in [12, 16].

If the system dynamics also include uncertain model parameters, as in (2), the MPC criterion (5) can be reformulated in order to incorporate worst-case values of the uncertain parameters θ, and the MPC problem can be written as

$$\min_{\hat{u}_{t:t+T-1}\in\mathcal{U}} \max_{\hat{d}_{t:t+T-1}\in\mathcal{D},\hat{\theta}\in\Theta} \sum_{s=t}^{t+T-1} \big(c_s(x_s,\hat{u}_s,\hat{d}_s) - \rho_s(\hat{d}_s)\big) + q_{t+T}(x_{t+T}), \tag{6}$$

where each element of the future control input sequence $u_{t:t+T-1}$ is constrained to belong to the set $\mathcal{U}$, each element of the future disturbance sequence $d_{t:t+T-1}$ is assumed to belong to the set $\mathcal{D}$, and the uncertain parameter θ is known to belong to the set Θ. In order to overcome the conservativeness of open-loop control, at each time step t, the first element $\hat{u}_t^*$ of the future control input sequence $\hat{u}^*_{t:t+T-1}$ that is the solution to (6) is applied to the system, and the optimization (6) is solved again at each time step in a receding horizon fashion. This is similar to the adaptive MPC problem with exogenous inputs considered in [8].

2.3 Adaptive MPC combined with MHE

Next we show how both the MPC problem (6) and MHE problem (4) can be formulated and solved simultaneously as a single min-max optimization problem.

Taking the criterion (5) and subtracting the criterion (3) gives a criterion of the form

$$J_t \coloneqq \sum_{s=t}^{t+T-1} c_s(x_s, u_s, d_s) + q_{t+T}(x_{t+T}) - \sum_{s=t-L}^{t} \eta_s(n_s) - \sum_{s=t-L}^{t+T-1} \rho_s(d_s), \quad (7)$$

which contains $T \in \mathbb{Z}_{\geqslant 1}$ terms of the running cost $c_s(x_s, u_s, d_s)$, which recede as the current time t advances, $L + 1 \in \mathbb{Z}_{\geqslant 1}$ terms of the measurement cost $\eta_s(n_s)$, and $L + T \in \mathbb{Z}_{\geqslant 1}$ terms of the cost on the input disturbance $\rho_s(d_s)$. Again, the function $q_{t+T}(x_{t+T})$ acts as a terminal cost in order to penalize the "final" state at time $t + T$. The functions $c_t(\cdot)$, $q_{t+T}(\cdot)$, $\eta_t(\cdot)$, and $\rho_t(\cdot)$ in (7) are all assumed to take non-negative values. We use finite-horizons into the past and into the future in order to decrease the computational complexity of the optimization problem, and we use online optimization to generate closed-loop solutions.

The *control objective* is to select the control signal $u_t \in \mathcal{U}$, $\forall t \in \mathbb{Z}_{\geqslant 0}$, so as to minimize the criterion defined in (7) under worst-case assumptions on the *unknown* system's initial condition $x_{t-L} \in \mathcal{X}$, unmeasured disturbances $d_t \in \mathcal{D}$, measurement noise $n_t \in \mathbb{R}^{n_n}$, and uncertain parameter $\theta \in \Theta$, for all $t \in \mathbb{Z}_{\geqslant 0}$, subject to the constraints imposed by the system dynamics (2) and the measurements $y_{t-L:t}$ collected up to the current time t.

Because the objective is to optimize the criterion (7) at the current time t in order to compute control inputs u_s for times $s \geqslant t$, there is no reason to penalize other irrelevant terms. For instance, the first summation in (7) starts at time t because there is no reason to penalize the running cost $c_s(x_s, u_s, d_s)$ for past time instants $s < t$. There is also no reason to consider the values of future measurement noise at times $s > t$ as they will not affect choices made at time t. Thus, the second summation in (7) ends at time t. However, all values of the unmeasured disturbance d_s for $t - L \leqslant s \leqslant t + T - 1$ need to be considered because past values affect the (unknown) current state x_t, and future values affect the future values of the running cost.

Boundedness of (7) by a constant γ guarantees that

$$\sum_{s=t}^{t+T-1} c_s(x_s, u_s, d_s) + q_{t+T}(x_{t+T}) \leqslant \gamma + \sum_{s=t-L}^{t} \eta_s(n_s) + \sum_{s=t-L}^{t+T-1} \rho_s(d_s). \quad (8)$$

This shows that we can bound the running and final costs involving the future states $x_{t:t+T}$ in terms of bounds on the noise and disturbance.

Remark 1 (Quadratic case). While the results presented here are general, it may be easier to gain intuition on the results when considering a quadratic criterion for (7) such as $c_t(x_t, u_t, d_t) \coloneqq \|x_t\|^2 + \|u_t\|^2$, $\eta_t(n_t) \coloneqq \|n_t\|^2$, $\rho_t(d_t) \coloneqq \|d_t\|^2$. For this choice of the criterion, boundedness of (7) guarantees that the state x_t and input u_t are ℓ_2 provided that the disturbance d_t and noise n_t are also ℓ_2 [c.f. (8)]. □

With the objective of optimizing the criterion (7) at a given time $t \in \mathbb{Z}_{\geqslant 0}$ for the future control inputs $u_{t:t+T-1}$ and worst-case estimates of x_{t-L}, $d_{t-L:t+T-1}$, and θ, the combined adaptive MPC with MHE problem amounts to solving the following min-max optimization

$$J_t^* = \min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1} \in \mathcal{D}, \hat{\theta} \in \Theta}$$
$$\sum_{s=t}^{t+T-1} c_s(\hat{x}_s, \hat{u}_s, \hat{d}_s) + q_{t+T}(\hat{x}_{t+T}) - \sum_{s=t-L}^{t} \eta_s\big(y_s - g_s(\hat{x}_s, \hat{\theta})\big) - \sum_{s=t-L}^{t+T-1} \rho_s(\hat{d}_s), \quad (9)$$

with the understanding that

$$\hat{x}_{s+1} = \begin{cases} f_s(\hat{x}_s, \hat{\theta}, u_s, \hat{d}_s) & \text{for } t - L \leqslant s < t, \\ f_s(\hat{x}_s, \hat{\theta}, \hat{u}_s, \hat{d}_s) & \text{for } t \leqslant s < t + T. \end{cases}$$

We can view the optimization variables $\hat{x}_{t-L}$, $\hat{d}_{t-L:t+T-1}$, and $\hat{\theta}$ as (worst-case) estimates of the initial state, disturbances, and uncertain parameter, respectively, based on the past inputs $u_{t-L:t-1}$ and outputs $y_{t-L:t}$ available at time t.

Just as in the MPC problem, after solving this optimization problem at each time t, we use as the control input the first element of the sequence

$$\hat{u}^*_{t:t+T-1} = \{\hat{u}^*_t, \hat{u}^*_{t+1}, \hat{u}^*_{t+2}, \ldots, \hat{u}^*_{t+T-1}\} \in \mathcal{U}$$

that minimizes (9), leading to the control law

$$u_t = \hat{u}^*_t, \quad \forall t \geqslant 0. \quad (10)$$

A depiction of an example solution to the combined finite-horizon control and estimation problem is shown in Figure 2.

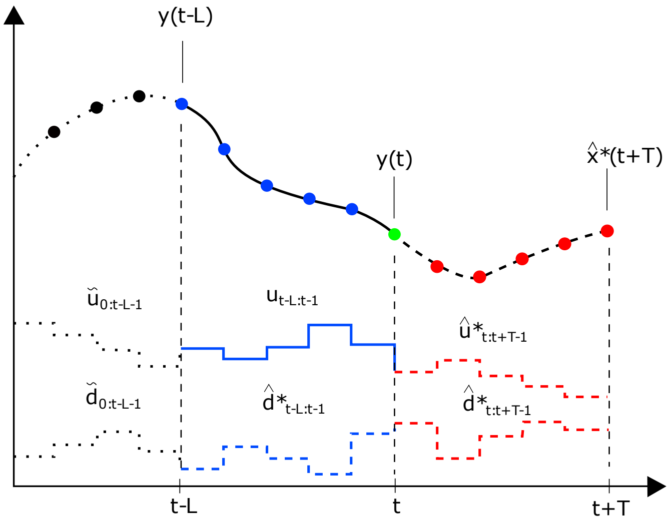

Fig. 2 Example solution to the combined finite-horizon control and estimation problem. The elements from $t-L$ to t correspond to the MHE problem, and the elements from t to $t+T$ correspond to the MPC problem. The variables denoted as $\hat{\cdot}$ are optimization variables, the variables denoted as $\tilde{\cdot}$ are not relevant for the optimization, and the other variables are known.

3 Stability Results

Next we discuss under what appropriate assumptions the control law (10) leads to boundedness of the state of the closed-loop system resulting from the finite-horizon optimization introduced in Section 2.3.

In order to implement the control law (10), the outer minimization in (9) must lead to a finite value for the optimum. For the stability results in this section, we require the existence of a finite-valued saddle-point solution to the min-max optimization in (9), which is a common requirement in game theoretical approaches to control design [4].

Assumption 1 (Saddle-point) *The min-max optimization* (9) *always has a finite-valued saddle-point solution for which the min and max commute. Specifically, for all* $t \in \mathbb{Z}_{\geqslant 0}$, $u_{t-L:t-1} \in \mathcal{U}$, $y_{t-L:t}$, *there exists* $J_t^* \in \mathbb{R}$, $\hat{x}_{t-L}^* \in \mathcal{X}$, $\hat{u}_{t:t+T-1}^* \in \mathcal{U}$, $\hat{d}_{t-L:t+T-1}^* \in \mathcal{D}$, *and* $\hat{\theta}^* \in \Theta$ *such that*

$$\begin{aligned} J_t^* &= \min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \quad \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1} \in \mathcal{D}, \hat{\theta} \in \Theta} \quad J_t \\ &= \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1} \in \mathcal{D}, \hat{\theta} \in \Theta} \quad \min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \quad J_t \\ &< \infty. \end{aligned}$$

□

Assumption 1 presumes an appropriate form of *observability/detectability* adapted to the criterion $\sum_{s=t}^{t+T-1} c_s(x_s, u_s, d_s)$. In particular, it implies that the size of the *current* state can be bounded using past outputs and past/future input disturbances, regardless of the value of $\theta \in \Theta$.

To ensure controllability and to establish state boundedness under the control (10) defined by the finite-horizon optimization (9), we require additional assumptions regarding the dynamics and the terminal cost $q_t(\cdot)$.

Assumption 2 (Reversible Dynamics) *For every $t \in \mathbb{Z}_{\geqslant 0}$, $x_{t+1} \in \mathcal{X}$, $\theta \in \Theta$, and $u_t \in \mathcal{U}$, there exists a state $\tilde{x}_t \in \mathcal{X}$ and a disturbance $\tilde{d}_t \in \mathcal{D}$ such that*

$$x_{t+1} = f_t(\tilde{x}_t, \theta, u_t, \tilde{d}_t). \tag{11}$$

□

Assumption 3 (ISS-control Lyapunov function) *The terminal cost $q_t(\cdot)$ is an* ISS-control Lyapunov function, *in the sense that, for every $t \in \mathbb{Z}_{\geqslant 0}$, $x \in \mathcal{X}$, $d \in \mathcal{D}$, and $\theta \in \Theta$, there exists a control $u \in \mathcal{U}$ such that*

$$q_{t+1}\big(f_t(x, \theta, u, d)\big) - q_t(x) \leqslant -c_t(x, u, d) + \rho_t(d). \tag{12}$$

□

The mild Assumption 2 essentially implies that the sets of disturbances $\mathcal{D}$ and past states $\mathcal{X}$ are sufficiently rich to allow for a jump to any future state in $\mathcal{X}$. In fact, for linear dynamics, Assumption 2 is satisfied if the state-space A matrix has no eigenvalues at the origin (e.g., if it results from the time-discretization of a continuous-time system).

Assumption 3 plays the role of the common assumption in MPC that the terminal cost must be a control Lyapunov function for the closed-loop [15]. Without the disturbance d_t, (12) would imply that $q_t(\cdot)$ could be viewed as a control Lyapunov function that decreases along system trajectories for an appropriate control input u_t [20]. With the disturbance d_t, $q_t(\cdot)$ should be viewed as an ISS-control Lyapunov function that satisfies an ISS stability condition for the disturbance input d_t and an appropriate control input u_t [11]. In the case of linear dynamics and a quadratic cost function, a terminal cost $q_t(\cdot)$ can typically be found by solving a system of linear matrix inequalities.

3.1 State Boundedness

The following theorem provides a bound that can be used to prove boundedness of the state when the control signal is computed by solving the finite-horizon optimization (9).

Theorem 1 (Finite-horizon cost-to-go bound). *Suppose that Assumptions 1, 2, and 3 hold. Then there exists a finite constant J_0^* and vectors $\tilde{d}_s \in \mathcal{D}$, $\tilde{n}_s \in \mathbb{R}^{n_n}$, $\forall s \in \{0, 1, \ldots, t-L-1\}$ for which*

$$\eta_s(\tilde{n}_s), \rho_s(\tilde{d}_s) < \infty, \quad \forall s \in \{0, 1, t-L-1\},$$

and the trajectories of the process (2) *with control* (10) *defined by the finite-horizon optimization* (9) *satisfy*

$$c_t(x_t, u_t, d_t) \leqslant J_0^* + \sum_{s=0}^{t-L-1} \eta_s(\tilde{n}_s) + \sum_{s=0}^{t-L-1} \rho_s(\tilde{d}_s) + \sum_{s=t-L}^{t} \eta_s(n_s) + \sum_{s=t-L}^{t} \rho_s(d_s), \quad \forall t \in \mathbb{Z}_{\geqslant 0} \tag{13}$$

□

Proof. This result is an extension of Theorem 1 presented in [6]. If the state is augmented such that $\bar{x}_t = [x_t \ \theta]^\top$, and the process is defined as $\bar{x}_{t+1} = [\bar{f}_t(\bar{x}_t, u_t, d_t) \ \theta]^\top$, then the same proof used for Theorem 1 in [6] can be applied here using $\bar{x}_t$ in place of x_t. □

We refer the reader to [6] for a discussion on how this result can be used to ensure state boundedness and reference tracking.

4 Simulation Study

In this section we consider several examples of systems with parametric uncertainty and present closed-loop simulations using the control approach described in Section 2. For all of the following examples, we use a cost function of the form

$$J_t = \sum_{s=t}^{t+T-1} \|h_s(x_s)\|_2^2 + \lambda_u \sum_{s=t}^{t+T-1} \|u_s\|_2^2 - \lambda_n \sum_{s=t-L}^{t} \|n_s\|_2^2 - \lambda_d \sum_{s=t-L}^{t+T-1} \|d_s\|_2^2. \tag{14}$$

where $h_s(x_s)$ is a function of the state x_s that is especially relevant for the example under consideration, and λ_u, λ_n, and λ_d are positive weighting constants.

Given the optimization criterion (14), the following examples involve optimizing this criterion with respect to the future control inputs $u_{t:t+T-1}$ under worst-case assumptions on x_{t-L}, $d_{t-L:t+T-1}$, and θ by solving the following min-max optimization problem:

$$\min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \quad \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1}, \hat{\theta} \in \Theta} J_t. \tag{15}$$

The first time this optimization is solved, guesses for the initial values of the uncertain parameter θ, initial state x_{t-L}, past control inputs $u_{t-L:t-1}$ and input disturbances $d_{t-L:t-1}$ need to be made. Then values for the past states $x_{t-L+1:t}$ that are consistent with the dynamics are picked. These states can be used to determine the output measurements $y_{t-L:t}$, and then the optimization (15) can be solved for the first time. At subsequent times, all of the variables from the solution of (15) at the previous time step (after moving away from the constraints) can be used as a "warm start" for solving (15) at the current time step.

In order to solve the optimization (15), we have developed a primal-dual-like interior-point method that finds the saddle-point solution. Details of this method can be found in [7]. Under appropriate convexity assumptions, this method is guaranteed to terminate at a global solution. However, the simulation results show that it also converges in problems that are severely non-convex, such as most of the examples below.

Example 1 (Linear System - uncertain gain and poles).

Consider a discrete-time linear system described by the transfer function

$$G(z) = \frac{b}{(z - p_1)(z - p_2)}, \tag{16}$$

where $p = [p_1 \; p_2]^\top$ denotes the uncertain pole locations p_1 and p_2 that are assumed to belong to the set $\mathcal{P} \coloneqq \{p \in \mathbb{R}^2 : 0 \leqslant p_i \leqslant 2, i = 1, 2\}$, so they may be stable or unstable. The parameter b is an uncertain gain assumed to belong in the interval $\mathcal{B} \coloneqq \{b \in \mathbb{R} : 1 \leqslant b \leqslant 5\}$.

The transfer function (16) can be rewritten in state space controllable canonical form as

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} p_1 + p_2 & -p_1 p_2 \\ 1 & 0 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} (u_t + d_t), \\ y_t &= \begin{bmatrix} 0 & b \end{bmatrix} x_t + n_t, \qquad\qquad \forall t \in \mathbb{Z}_{\geqslant 0}, \end{aligned} \tag{17}$$

where y_t is the measured output at time t with noise n_t, and d_t is an additive input disturbance. For all $t \in \mathbb{Z}_{\geqslant 0}$, the control input u_t is constrained to belong in the set $\mathcal{U} \coloneqq \{u_t \in \mathbb{R} : \|u_t\|_\infty \leqslant 8\}$, and the input disturbance d_t is assumed to belong to the set $\mathcal{D} \coloneqq \{d_t \in \mathbb{R} : \|d_t\|_\infty \leqslant 0.1\}$.

By defining $a_1 = p_1 + p_2$ and $a_2 = p_1 p_2$, the state space model (17) can be reparametrized as

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} a_1 & -a_2 \\ 1 & 0 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} (u_t + d_t), \\ y_t &= \begin{bmatrix} 0 & b \end{bmatrix} x_t + n_t, \qquad \forall t \in \mathbb{Z}_{\geqslant 0}. \end{aligned} \tag{18}$$

Letting $a = [a_1 \ a_2]^\top$, the uncertain parameter a is assumed to belong to $\mathcal{A} \coloneqq \{a \in \mathbb{R}^2 : 0 \leqslant a_i \leqslant 4, i = 1, 2\}$. This set $\mathcal{A}$ is conservative, and a tighter non-convex set could be used. Now the model (18) is linear in the uncertain parameters. This is a standard problem that can be solved using classical adaptive control techniques. We will show that our MPC with MHE approach can solve this problem, and in following examples, we will see that our approach does not require reparametrization such that the system is linear in the uncertain parameters.

The uncertain parameters (a and b) can be estimated by including them as optimization variables in the following problem

$$\min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \ \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1} \in \mathcal{D}, \hat{a} \in \mathcal{A}, \hat{b} \in \mathcal{B}} \\ \sum_{s=t}^{t+T} \|y_s - r_s\|_2^2 + \lambda_u \sum_{s=t}^{t+T-1} \|\hat{u}_s\|_2^2 - \lambda_n \sum_{s=t-L}^{t} \|n_s\|_2^2 - \lambda_d \sum_{s=t-L}^{t+T-1} \|\hat{d}_s\|_2^2, \tag{19}$$

where r_t is a desired reference signal for the output of the system to follow. Figures 3 and 4 show simulations of the resulting closed-loop system for a square-wave reference defined as $r_t = 10\,\mathrm{sgn}(\sin(0.4t))$ and the backward and forward horizon lengths chosen as $L = 10$, and $T = 10$, respectively. The weights in the cost function are chosen to be $\lambda_u = 0.1$, $\lambda_n = 1000$, and $\lambda_d = 1000$. In this simulation, the actual input disturbance d_t and measurement noise n_t are unmeasured Gaussian independently and identically distributed (i.i.d.) random variables with zero mean and standard deviations of 0.001 and 0.005, respectively.

Figure 3 shows the output of the system successfully following the given square-wave reference trajectory. The system is initialized with incorrect guesses for the initial values of the uncertain parameters, and zero control input (i.e. $u_t = 0$) is applied for the first $L = 10$ time steps. After that point, starting at time $t = 11$, the optimization problem (19) is solved at each time step, and the computed control input $\hat{u}_t^*$ is applied in a receding horizon fashion.

At several times throughout the simulation, the true model of the system (17) is altered by changing the value of the gain or the poles (which can be seen in Figure 4). The estimates of the gain and pole locations shown

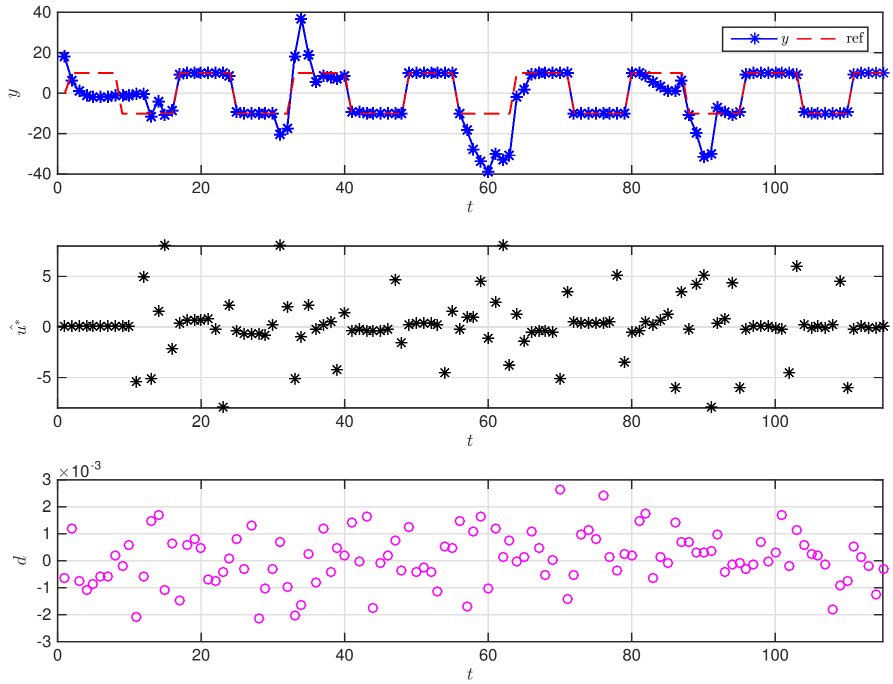

Fig. 3 Linear System: output and inputs. The top plot shows the measured output (denoted by o's) and the reference signal (denoted by -'s). The second plot shows the input $\hat{u}_t^*$ applied to the system, and the third plot shows the unmeasured disturbance input d_t that the system is subjected to.

in Figure 4 are obtained from the estimates of the uncertain parameters a and b from the solution of the optimization (19). Figure 4 shows that the estimates of the gain and pole locations converge to their true values. Even after the true values of the gain and pole locations are changed during the simulation, the combined MPC and MHE scheme is able to adapt to the changing system, effectively regulating the system to the reference trajectory and correctly learning the new parameters of the model. □

Example 2 (Inverted Pendulum - uncertain mass and friction).

Consider an inverted pendulum actuated by a torque at the base as shown in Figure 5 and described by the model

$$ml^2\ddot{\phi} = mgl\sin(\phi) - b\dot{\phi} + \tau,$$

where m is the mass at the end of the pendulum, l is the length of the link, ϕ is the angle from vertical, g is the gravitational constant, b is the coefficient of friction, and τ is the torque applied at the base.

We can rewrite this model in state space form as

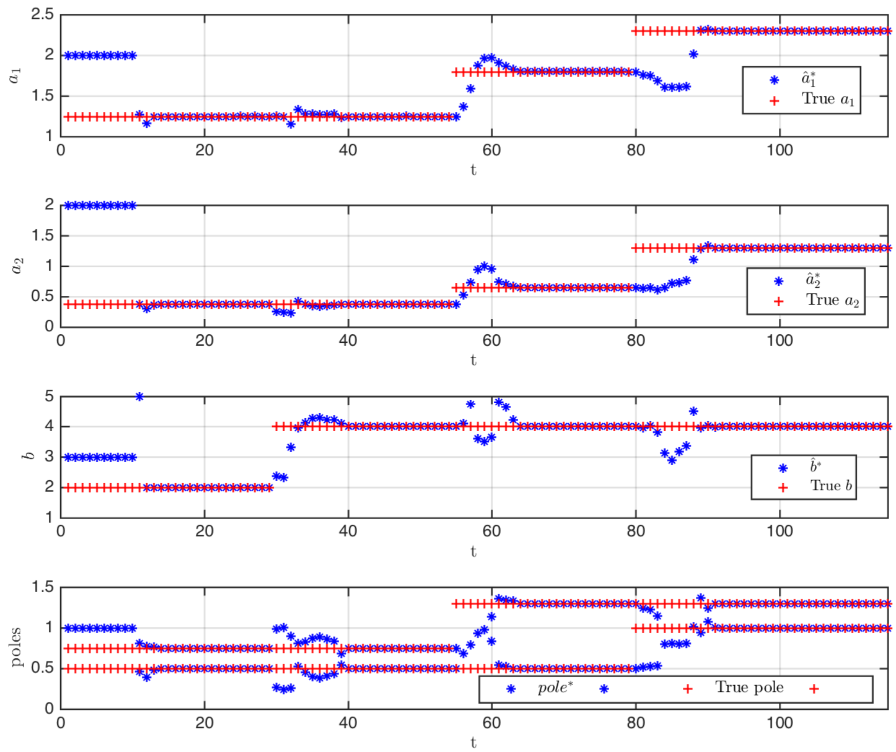

Fig. 4 Linear System: parameters. The top two plots show the true values of a_1 and a_2 (denoted by +'s) and their estimated values (denoted by *'s). The third plot shows the true value of the gain b (denoted by +'s) and its estimated value (denoted by *'s). The bottom plot shows the true values of the poles p_1 and p_2 (denoted by +'s) and their estimated values (denoted by *'s) computed from $\hat{a}_1^*$ and $\hat{a}_2^*$.

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \frac{g}{l}\sin(x_1) - \frac{b}{ml^2}x_2 + \frac{1}{ml^2}u,
\end{aligned} \tag{20}$$

where $x_1 = \phi$, $x_2 = \dot{\phi}$, and $u = \tau$.

By letting $l = 1$, $g = 9.81$, and $a = 1/m$, adding an input disturbance d, and discretizing using Euler's Method with time step Δt, the system (20) becomes

$$\begin{aligned}
x_{1,t+1} &= x_{1,t} + \Delta t\ x_{2,t}, \\
x_{2,t+1} &= x_{2,t} + \Delta t\big(9.81\sin(x_{1,t}) - abx_{2,t} + a(u_t + d_t)\big), \\
y_t &= x_{1,t} + n_t, \qquad\qquad \forall t \in \mathbb{Z}_{\geqslant 0},
\end{aligned} \tag{21}$$

where y_t is the measurement available at time t with noise n_t. According to this model, a noisy measurement of the angle x_1 is available at each time t.

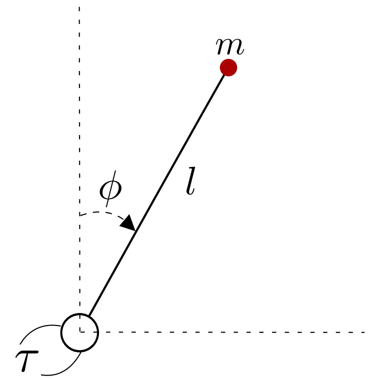

Fig. 5 Diagram of the pendulum considered in Examples 2 and 3.

The inverse of the mass and the coefficient of friction (a and b, respectively) are uncertain but assumed to belong to the sets $\mathcal{A} \coloneqq \{a \in \mathbb{R} : 1/2 \leqslant a \leqslant 1\}$ and $\mathcal{B} \coloneqq \{b \in \mathbb{R} : 0.2 \leqslant b \leqslant 0.7\}$, respectively. The control input u_t is constrained to the set $\mathcal{U} \coloneqq \{u_t \in \mathbb{R} : \|u_t\|_\infty \leqslant 5\}$, and the disturbance input d_t is assumed to belong to $\mathcal{D} \coloneqq \{d_t \in \mathbb{R} : \|d_t\|_\infty \leqslant 0.3\}$ for all $t \in \mathbb{Z}_{\geqslant 0}$.

The control objective is to regulate the output (the noisy measurement of the angle ϕ) to a desired reference. The uncertain mass and coefficient of friction can be determined using estimates of the parameters a and b in (21). These parameters can be estimated by including them as optimization variables in the following problem.

$$\min_{\hat{u}_{t:t+T-1} \in \mathcal{U}} \quad \max_{\hat{x}_{t-L} \in \mathcal{X}, \hat{d}_{t-L:t+T-1} \in \mathcal{D}, \hat{a} \in \mathcal{A}, \hat{b} \in \mathcal{B}} \sum_{s=t}^{t+T} \|y_s - r_s\|_2^2 + \lambda_u \sum_{s=t}^{t+T-1} \|\hat{u}_s\|_2^2 - \lambda_n \sum_{s=t-L}^{t} \|n_s\|_2^2 - \lambda_d \sum_{s=t-L}^{t+T-1} \|\hat{d}_s\|_2^2. \quad (22)$$

A noteworthy challenge in this nonlinear problem is that the unknown parameters a and b appear multiplied by the unmeasurable state x_2 in the system dynamics (21).

Figure 6 shows a simulation of the closed-loop system with the discrete time-step chosen as $\Delta t = 0.2$ and a square-wave reference given as $r_t = 5(\pi/180)\,\mathrm{sgn}\big(\sin(0.5t)\big)$. The backward and forward horizon lengths are chosen to be $L = 6$, and $T = 7$, respectively. The weights in the cost function are chosen to be $\lambda_u = 0$ (i.e. the control input is not penalized), $\lambda_n = 1000$, and $\lambda_d = 10$. In this simulation, the actual input disturbance d_t and measurement noise n_t are unmeasured Gaussian i.i.d. random variables with zero mean and standard deviations of 0.001 and 0.0001, respectively. The system is initialized with incorrect guesses for the initial values of the uncertain parameters, and zero control input (i.e. $u_t = 0$) is applied for the first $L = 6$ time steps. After that point, starting at time $t = 7$, the optimization problem

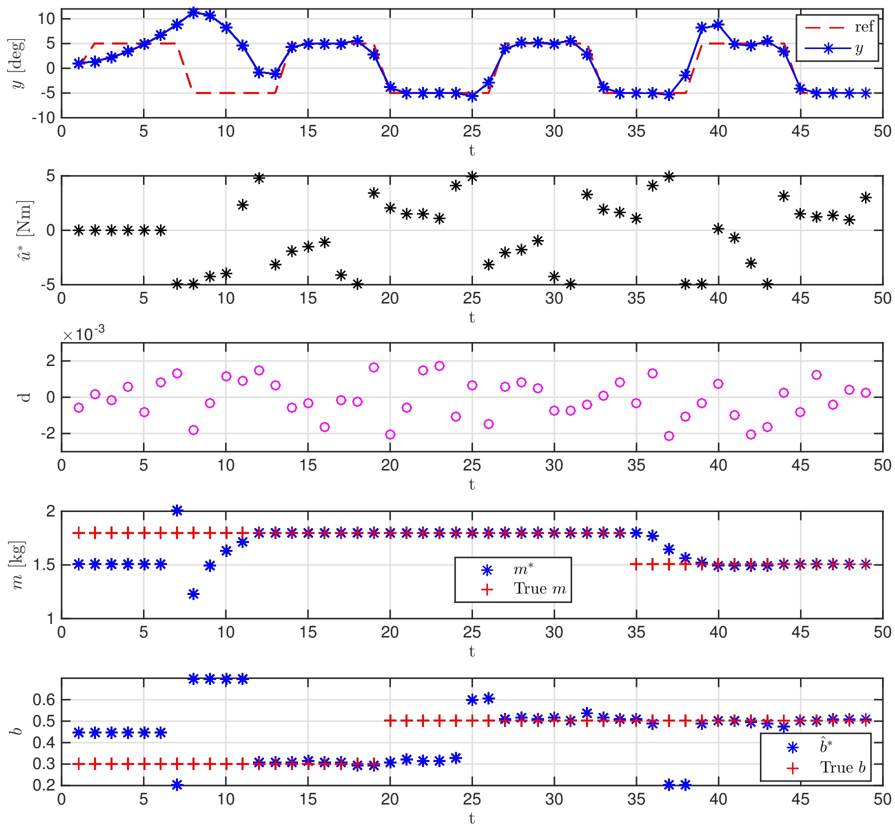

Fig. 6 Inverted Pendulum: uncertain mass and friction. The top plot shows the measured output (denoted by *'s) tracking the square reference signal (denoted by -'s). The second plot shows the control input $\hat{u}_t^*$ that is applied. The third plot shows the unmeasured input disturbance d_t that the system is subjected to. The bottom two plots show the true values of the mass and coefficient of friction (denoted by +'s) and the estimated values of the mass m^* (computed from $\hat{a}^*$) and coefficient of friction $\hat{b}^*$ (denoted by *'s).

(22) is solved at each time step, and the computed control input $\hat{u}_t^*$ is applied in a receding horizon fashion.

As in the previous example, the true values of the uncertain parameters (the mass and coefficient of friction) are changed several times throughout the simulation. Even as the true values of the mass and coefficient of friction change, the control input $\hat{u}_t^*$, computed by solving the optimization (22), is able to successfully regulate the output of the system to the reference trajectory, and the estimates of the uncertain mass m and coefficient of friction b converge to their true values. □

Example 3 (Inverted Pendulum - stabilization and disturbance rejection without learning true parameter values).

This example shows that this adaptive MPC with MHE approach can stabilize uncertain systems even when the systems are not persistently excited and the true values of the uncertain parameters are not learned. Furthermore, the results of this example show that this estimation and control approach is not only robust to the model uncertainty but is also able to reject large input disturbances.

Again we consider an inverted pendulum as depicted in Figure 5 and described using a discretized model of the form (21). This time, rather than following a reference trajectory, the control objective is to stabilize the system at the unstable equilibrium $x_1 = 0$. This means that the same optimization given in (22) is solved but with $r_t = 0$. Figure 7 shows a simulation of the resulting closed-loop system.

For this example, the parameters in the optimization (22) are chosen the same as in Example 2 except with respect to the input disturbance d_t. In this example, the unmeasured disturbance d_t is larger and assumed to belong to the set $\mathcal{D} \coloneqq \{d_t \in \mathbb{R} : \|d_t\|_\infty \leqslant 0.5\}$. The weight on the input disturbance in (22) is chosen as $\lambda_d = 1$. The actual input disturbance that the system is subjected to is a Gaussian i.i.d. random variable with zero mean and a standard deviation of 0.15. Again, the system is initialized with guesses for the initial values of the uncertain parameters, and zero control input (i.e. $u_t = 0$) is applied for the first $L = 6$ time steps. Starting at time $t = 7$, the optimization problem (22) is solved at each time step, and the computed control input $\hat{u}_t^*$ is applied in a receding horizon fashion.

Figure 7 shows that the system is not sufficiently excited in order to correctly learn the true values of the mass m and coefficient of friction b. However, the control input $\hat{u}_t^*$ computed by solving the optimization (22) is nonetheless able to stabilize the system (even as the true values of the uncertain parameters change) and reject the large unmeasured disturbance input. □

Example 4 (Nonlinear Pursuit-Evasion - uncertain wind).

In this example, we consider a two-player pursuit-evasion game where the pursuer is modeled as a unicycle vehicle, and the evader is modeled as a double-integrator. The pursuer is an aerial vehicle that is subject to wind disturbances, and the evader is a ground vehicle that is not susceptible to the wind. A nonlinear discrete-time model of the overall system is given as follows:

$$x_{t+1} = x_t + \begin{bmatrix} v\cos\phi_t + w_1 \\ v\sin\phi_t + w_2 \\ u_t \\ d_t \end{bmatrix}, \qquad y_t = x_t + n_t, \qquad \forall t \in \mathbb{Z}_{\geqslant 0}. \tag{23}$$

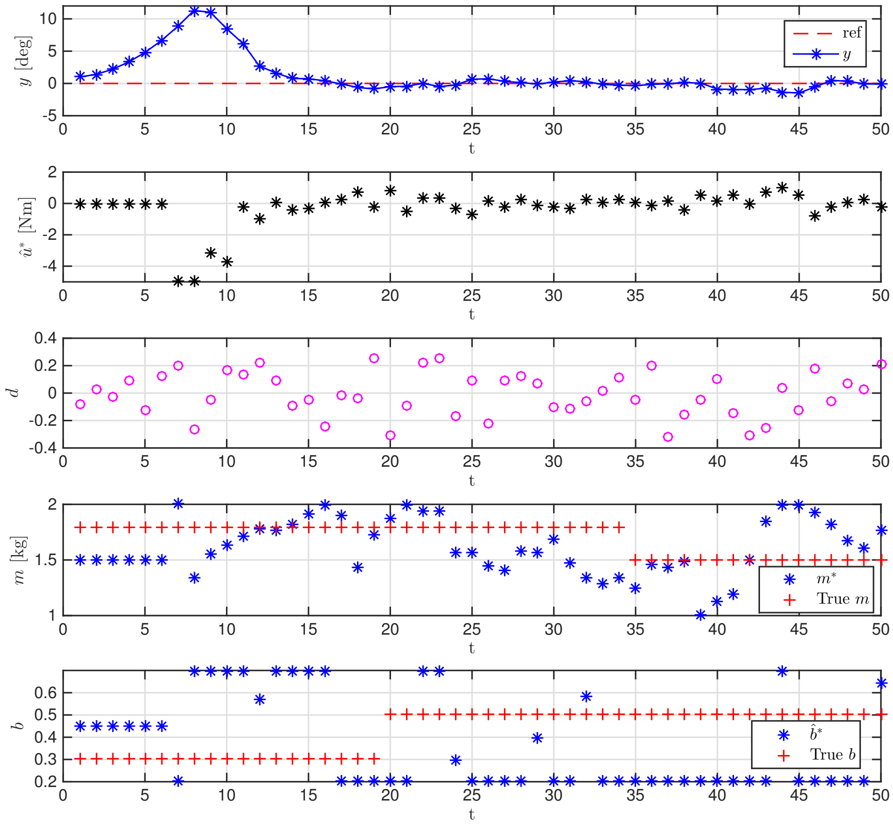

Fig. 7 Inverted Pendulum: stabilization and disturbance rejection. The top plot shows the output (denoted by *'s) converging to the unstable equilibrium $\phi = 0$ (denoted by -'s). The second plot shows the control input $\hat{u}_t^*$ that is applied. The third plot shows the large unknown input disturbance d_t that the system is subjected to. The fourth and fifth plots show the true values of the mass and coefficient of friction (denoted by +'s) and the estimated values of the mass m^* (computed from $\hat{a}^*$) and coefficient of friction $\hat{b}^*$ (denoted by *'s).

The state of the system is denoted by $x_t = [p_t\ \phi_t\ z_t]^\top$, where $p_t = [p_1\ p_2]_t^\top$ denotes the planar position of the pursuer, ϕ_t denotes the orientation of the pursuer, and $z_t = [z_1\ z_2]_t^\top$ denotes the planar position of the evader. The planar wind speed is denoted by $w = [w_1\ w_2]^\top$ where w_1 is the component of the wind speed in the x-direction, w_2 is the component of the wind speed in the y-direction, and w is known to belong to the set $\mathcal{W} \coloneqq \{w \in \mathbb{R}^2 : \|w\|_\infty \leqslant 0.05\}$. The control input u_t is constrained to belong in the set $\mathcal{U} \coloneqq \{u_t \in \mathbb{R} : \|u_t\|_\infty \leqslant 0.35\}$. The evader's velocity is given by $d_t = \begin{bmatrix} d_1\ d_2 \end{bmatrix}_t^\top$ and is constrained to the set $\mathcal{D} \coloneqq \{d_t \in \mathbb{R}^2 : \|d_t\|_\infty \leqslant 0.05\}$, and $n_t \in \mathbb{R}^{n_n}$ is measurement noise.

The pursuer's objective is to make the distance between its position p_t and the position of the evader z_t as small as possible, so the pursuer wants to minimize the value of $\|p_t - z_t\|$. The evader's objective is to do the opposite, namely, maximize the value of $\|p_t - z_t\|$. The pursuer and evader try to achieve these objectives by choosing appropriate values for u_t and d_t, respectively. The wind speed is unknown, but both the pursuer and evader would benefit from learning the wind speed. Therefore, the optimal solution will involve each player adapting his or her action (choice of u_t and d_t) to the current estimate of the wind speed. These considerations motivate solving the following problem

$$\min_{\hat{u}_{t:t+T-1}\in\mathcal{U}} \quad \max_{\hat{x}_{t-L}\in\mathcal{X},\hat{d}_{t-L:t+T-1}\in\mathcal{D},\hat{w}\in\mathcal{W}}$$
$$\sum_{s=t}^{t+T} \|p_s - z_s\|_2^2 + \lambda_u \sum_{s=t}^{t+T-1} \|\hat{u}_s\|_2^2 - \lambda_n \sum_{s=t-L}^{t} \|n_s\|_2^2 - \lambda_d \sum_{s=t-L}^{t+T-1} \|\hat{d}_s\|_2^2, \quad (24)$$

where the pursuer's future actions $u_{t:t+T-1}$, the unknown evader's actions $d_{t-L:t+T-1}$, the unknown initial state x_{t-L}, and the unknown wind speed w are included as optimization variables. A simulation of the resulting closed-loop system is shown in Figures 8, 9, and 10.

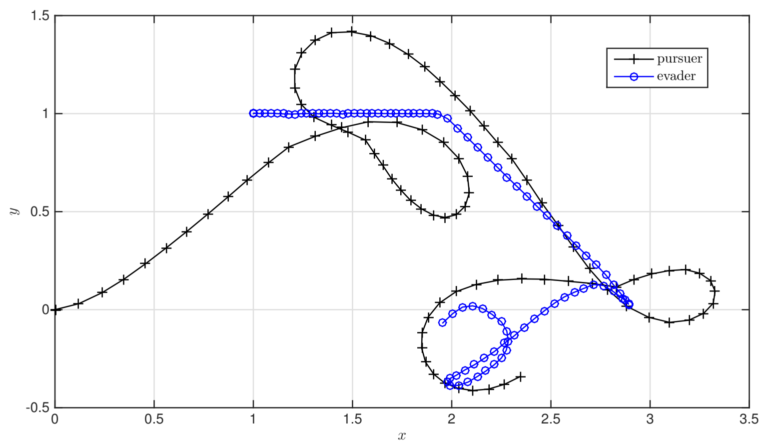

Fig. 8 Pursuit-evasion: trajectories. The pursuer's trajectory (denoted by +'s) begins at (0,0), and the evader's trajectory (denoted by o's) begins at (1,1).

Parameters chosen for the model (23) and the optimization (24) are given as follows. The pursuer moves with constant velocity $v = 0.1$. The backward and forward horizon lengths are chosen to be $L = 8$, and $T = 12$, respectively. The weights in the cost function in (24) are chosen to be $\lambda_u = 10$, $\lambda_n = 10000$, and $\lambda_d = 100$. The actual measurement noise n_t is an unmeasured Gaussian i.i.d. random variable with zero mean and a standard deviation of 0.001.

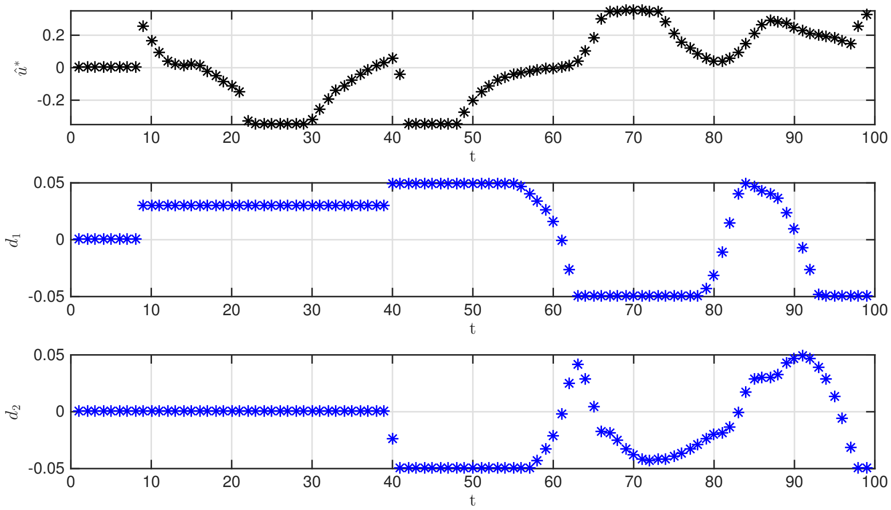

Fig. 9 Pursuit-evasion: inputs. The top plot shows the pursuer's input $\hat{u}_t^*$ that is applied. The lower two plots show the evader's input d_t that is applied. The evader applies constant velocity until time $t = 40$ at which time the optimal $\hat{d}_t^*$ is applied for the remainder of the simulation.

The trajectories that each player follow are shown in Figure 8. The evader moves with constant velocity in the positive x-direction until time $t = 40$ when the optimal $\hat{d}_t^*$ begins to be applied. The pursuer applies $\hat{u}_t^*$ throughout the entire simulation. Rapidly the pursuer catches up to the evader and is forced to make a loop due to its nonholonomic dynamics. The evader, on the other hand, is able to make sharp maneuvers due to its double-integrator dynamics. The inputs that each player applies are shown in Figure 9. Figure 10 shows that the estimates of the uncertain wind speed converge to their true values even as they change throughout the simulation. $\square$

5 Conclusions

In this chapter, we addressed adaptation and learning in the context of output-feedback MPC with MHE. Often the MPC and MHE problems are formulated with a known model of the dynamics. However, in this chapter, we investigated solving the MPC and MHE problems using a model with uncertain parameters. This was done by simultaneously solving the MPC and MHE problems as a single min-max optimization problem and including the uncertain model parameters as optimization variables to be estimated. Under appropriate assumptions ensuring controllability and observability, Theorem 1 provides bounds on the state of the system.

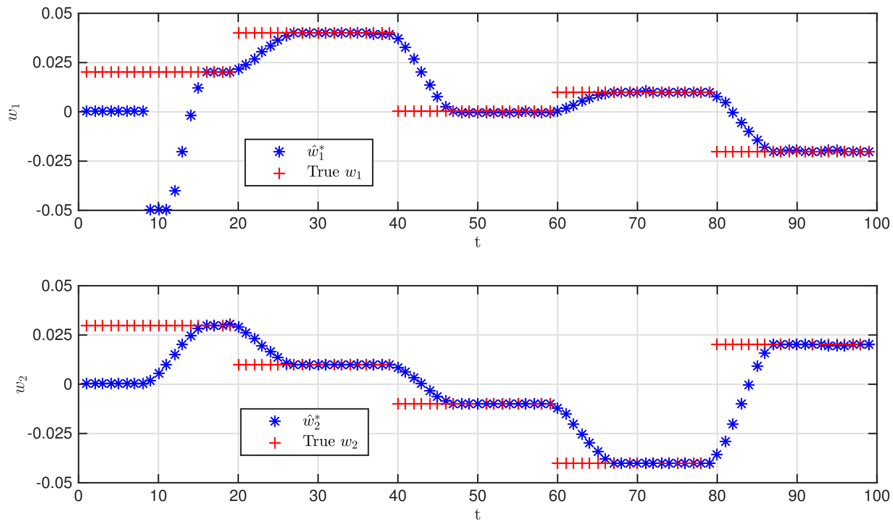

Fig. 10 Pursuit-evasion: wind. The top and bottom plots show the true values of the wind speed (denoted by +'s) and the estimated values $\hat{w}_1^*$ and $\hat{w}_2^*$ of the wind speed (denoted by *'s) in the x- and y-directions, respectively.

In a simulation study, we showed that the combined control and estimation approach effectively controls linear and nonlinear systems with model parameter uncertainty, adapts to changing model parameters, and also learns the uncertain model parameters when the system is sufficiently excited. However, even when the system is not sufficiently excited to learn the true values of the uncertain model parameters, the computed control law is still able to effectively reject disturbances and stabilize the system. Using a primal-dual-like interior point method, solutions to this MPC with MHE approach can be found even for severely non-convex examples.

Future work may involve investigating under what specific conditions the estimates of the uncertain parameters are guaranteed to converge to their true values.

References

[1] V. Adetola, D. DeHaan, and M. Guay. Adaptive model predictive control for constrained nonlinear systems. *Systems & Control Letters*, 58(5): 320–326, May 2009.

[2] A. Alessandri, M. Baglietto, and G. Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44(7):1753–1765, 2008.

[3] K. J. Åström and B. Wittenmark. *Adaptive Control*. Courier Corporation, 2013.

[4] T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press, London, 1995.

[5] A. Bemporad and M. Morari. Robust model predictive control: A survey. In *Robustness in identification and control*, pages 207–226. Springer, 1999.

[6] D. A. Copp and J. P. Hespanha. Nonlinear output-feedback model predictive control with moving horizon estimation. In *IEEE Conference on Decision and Control*, pages 3511–3517, December 2014.

[7] D. A. Copp and J. P. Hespanha. Nonlinear output-feedback model predictive control with moving horizon estimation. Technical report, Univ. California, Santa Barbara, 2014. `http://www.ece.ucsb.edu/~hespanha/techrep.html`.

[8] D. DeHaan and M. Guay. Adaptive robust mpc: A minimally-conservative approach. In *American Control Conference*, pages 3937–3942. IEEE, July 2007.

[9] P. A. Ioannou and J. Sun. *Robust Adaptive Control*. Courier Corporation, 2012.

[10] J. H. Lee and Z. Yu. Worst-case formulations of model predictive control for systems with bounded parameters. *Automatica*, 33(5):763–781, 1997.

[11] D. Liberzon, E. D. Sontag, and Y. Wang. Universal construction of feedback laws achieving ISS and integral-ISS disturbance attenuation. *Systems & Control Letters*, 46(2):111–127, 2002.

[12] L. Magni and R. Scattolini. Robustness and robust design of MPC for nonlinear discrete-time systems. In *Assessment and future directions of nonlinear model predictive control*, pages 239–254. Springer, 2007.

[13] L. Magni, G. De Nicolao, R. Scattolini, and F. Allgöwer. Robust model predictive control for nonlinear discrete-time systems. *International Journal of Robust and Nonlinear Control*, 13(3-4):229–246, 2003.

[14] D. Q. Mayne and H. Michalska. Adaptive receding horizon control for constrained nonlinear systems. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 1286–1291, December 1993.

[15] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6): 789–814, 2000.

[16] D. M. Raimondo, D. Limon, M. Lazar, L. Magni, and E. Camacho. Min-max model predictive control of nonlinear systems: A unifying overview on stability. *European Journal of Control*, 15(1):5–21, 2009.

[17] C. V. Rao, J. B. Rawlings, and D. Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *Automatic Control, IEEE Transactions on*, 48(2):246–258, 2003.

[18] J. B. Rawlings and D. Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2009.

[19] D. G. Robertson, J. H. Lee, and J. B. Rawlings. A moving horizon-based approach for least-squares estimation. *AIChE Journal*, 42(8):2209–2224, August 1996.

[20] E. D. Sontag. Control-lyapunov functions. In *Open problems in mathematical systems and control theory*, pages 211–216. Springer, 1999.