Active exploration in adaptive model predictive control*

Anilkumar Parsi, Andrea Iannelli and Roy S. Smith

Abstract— A dual adaptive model predictive control (MPC) algorithm is presented for linear, time-invariant systems subject to bounded disturbances and parametric uncertainty in the state-space matrices. Online set-membership identification is performed to reduce the uncertainty and thus control affects both the informativity of identification and the system's performance. The main contribution of the paper is to include this dual effect in the MPC optimization problem using a predicted worst-case cost in the objective function. This allows the controller to perform active exploration, that is, the control input reduces the uncertainty in the regions of the parameter space that have most influence on the performance. Additionally, the MPC algorithm ensures robust constraint satisfaction of state and input constraints. Advantages of the proposed algorithm are shown by comparing it to a passive adaptive MPC algorithm from the literature.

I. INTRODUCTION

Adaptive control is a technique where the controller parameters are updated using measurement data. Conventional methods of adaptive control like gain scheduling are based on certainty equivalence, and neglect the model uncertainties [1]. For this reason, they cannot handle constraints on states and inputs of the system. Model predictive control (MPC) is a popular technique since it guarantees stability and constraint satisfaction under uncertainty [2]. The structure of MPC controllers facilitates easy integration of model adaptation into the controller. Utilizing this advantage, a variety of adaptive MPC schemes have been proposed in the recent past using different model structures (state-space, impulse response, etc.) and adaptation methods (set-membership identification, recursive least-squares, etc.) [3], [4], [5], [6].

An adaptive MPC algorithm using set-membership identification is proposed in [4], which ensures robust constraint satisfaction for systems affected by bounded measurement noise. An extension of this algorithm was proposed in [7], where a worst-case cost is used to improve robustness of the performance. The algorithm uses an impulse response model which depends on a large number of parameters. An alternative method has been proposed in [5] which uses uncertain state-space models subject to bounded disturbances. Using tube-MPC, the algorithm ensures robust constraint satisfaction while reducing the uncertainty online. However, in all these methods the adaptation is passive, that is, the MPC optimizer does not exploit the fact that identification and control are being simultaneously performed.

These disadvantages can be addressed using dual control [8], a technique which computes control inputs under deci-

sion relevant, reducible uncertainty. An optimal dual control problem can be formulated by modeling the dependence of uncertainty reduction on the control inputs. The solution to this problem is given by dynamic programming, whose computational complexity is high [9]. Instead, the existing dual control algorithms approximate the optimal control problem using heuristics to add a probing effect on the control input. In [10], an adaptive MPC algorithm is presented with a constraint on the control input to ensure persistent excitation. Though parameter convergence is guaranteed, this method could result in excessive probing, especially after the uncertainty is reduced. In [11], the parameter error covariance is penalized in the objective function of MPC. However, a lower covariance of the parameter uncertainty might not always translate into improved performance, and the cost function requires tuning the probing effect. These problems can be mitigated by using an application-oriented approach to dual control [12], [13]. Here the probing effect is induced by using a measure of the robust performance, such as worst-case cost, instead of geometric measures of uncertainty. This ensures that active exploration is performed, that is, the uncertainty is reduced in regions of parameter space to improve the control performance.

The main contribution of this paper is to formulate active exploration in a dual adaptive MPC framework. For this purpose, the regulation of a linear, time-invariant system with affine uncertainty in the state space matrices is considered. The system is subject to bounded disturbances and must satisfy state and input constraints. Using an approach similar to [14], the problems of feasibility and learning are decoupled by using two state tubes. A robust state tube is used to ensure feasibility for all model parameters in an initial parameter set, which can be large. A predicted parameter set is then defined as a function of the control input, and a predicted state tube is constructed such that it is robust to uncertainties in predicted parameter set. The cost function is defined as the worst-case cost over the predicted state tube. The algorithm requires a non-convex optimization problem to be solved online. The algorithm has the flexibility to trade-off between the computational complexity and quality of active exploration using the length of the predicted state tube. The performance of the algorithm with varying predicted state tube lengths is compared to a passive adaptive MPC algorithm using numerical simulations.

A. Notation

The sets of real numbers and non-negative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ respectively. The sequence of integers from n_1 to n_2 is represented by $\mathbb{N}_{n_1}^{n_2}$. For a vector

b, b^T represents its transpose, and $[b]_i$ refers to its ith element. The ith row of a matrix A is denoted by $[A]_i$. The dimensions of matrices and vectors are not explicitly specified when they can be inferred from the context. For any real scalar-valued function J, $\max\limits_{h \in \mathbb{H}} J(h)$ refers to the maximum value of J over the set $\mathbb{H}$. The Minkowski sum of two sets A and B is denoted by $A \oplus B$, and $\mathbf{1}$ denotes a column vector of appropriate length whose elements are equal to 1. The convex hull of the elements of a set S is represented by co$\{$S$\}$. The notation $a_{l|k}$ denotes the value of a at time step $k + l$ predicted at the time step k.

II. PROBLEM CONFIGURATION

A. System description

We consider a discrete time, linear time-invariant system with state $x_k \in \mathbb{R}^n$, control input $u_k \in \mathbb{R}^m$ and disturbance $w_k \in \mathbb{W} \subset \mathbb{R}^n$ at the time step k. The system dynamics can be described according to the parametric equation

$$x_{k+1} = A(\theta)x_k + B(\theta)u_k + w_k, \tag{1}$$

where $\theta \in \mathbb{R}^p$ is an unknown, constant parameter and θ^* is its true value. It is assumed that all the state variables are measurable. The state matrices are parameterized as

$$A(\theta) = A_0 + \sum_{i=1}^{p} A_i[\theta]_i, \quad B(\theta) = B_0 + \sum_{i=1}^{p} B_i[\theta]_i, \tag{2}$$

and θ belongs to the bounded polytope

$$\Theta := \{\theta \in \mathbb{R}^p | H_\theta \theta \le h_\theta\}, \tag{3}$$

such that $\theta^* \in \Theta$ and $H_\theta \in \mathbb{R}^{n_\theta \times p}$. The states and inputs of the system must satisfy the constraints

$$\mathbb{Z} = \left\{(x_k, u_k) \in \mathbb{R}^n \times \mathbb{R}^m \big| Fx_k + Gu_k \le \mathbf{1}\right\}, \tag{4}$$

where $\mathbb{Z}$ is a compact set and $F \in \mathbb{R}^{n_c \times n}$. The objective is to regulate the system state from the initial condition x_0 to the origin, while robustly satisfying the constraints in (4).

Assumption 1: The disturbance set $\mathbb{W}$ is a bounded polytope described by the n_w constraints in the set

$$\mathbb{W} = \{w \in \mathbb{R}^n | H_w w \le h_w\}. \tag{5}$$

B. Online set-membership identification

Set-membership identification is a technique used to identify systems affected by bounded noise with unknown statistical properties [15]. The identification procedure defines a feasible parameter set (FPS), which contains the set of all parameters to be robust against. The FPS is initialized with Θ and updated at each time step k to Θ_k. To perform the update, a set of non-falsified parameters is constructed using measurement data from the previous s time steps as

$$
\begin{aligned}
\Delta_k &:= \left\{ \theta \in \mathbb{R}^p \,\middle|\, x_{t+1} - A(\theta)x_t - B(\theta)u_t \in \mathbb{W}, \, \forall t \in \mathbb{N}_{k-s}^{k-1} \right\} \\
&= \left\{ \theta \in \mathbb{R}^p \,\middle|\, -H_w D_t \theta \le h_w + H_w d_{t+1}, \forall t \in \mathbb{N}_{k-s}^{k-1} \right\} \\
&= \{\theta \in \mathbb{R}^p \mid H_\Delta \theta \le h_\Delta\},
\end{aligned}
\tag{6}
$$

where $D_t \in \mathbb{R}^{n \times p}$ and $d_{k+1} \in \mathbb{R}^n$ are

$$
\begin{aligned}
D_t &:= D(x_t, u_t) = \left[A_1 x_t + B_1 u_t, \ \ldots, \ A_p x_t + B_p u_t\right], \\
d_{t+1} &:= A_0 x_t + B_0 u_t - x_{t+1}, \qquad \forall t \in \mathbb{N}_{k-s}^{k-1}.
\end{aligned}
\tag{7}
$$

Note that D_t and d_t are quantities which linearly depend on the measured state and input vectors, but the dependence is omitted for clarity. This notation is adopted so that Δ can be represented by hyperplane constraints in $\mathbb{R}^p$.

In (6), the non-falsified set Δ_k defines the set of all parameters that could have generated the measurement sequence $\{x_{k-s}, \ldots, x_k\}$. The set Θ_k is defined using a fixed number of polytopic constraints given by

$$\Theta_k := \{\theta \in \mathbb{R}^p | H_\theta \theta \le h_{\theta_k}\}. \tag{8}$$

The matrix H_θ is chosen offline and h_{θ_k} is updated online such that

$$\Theta_k \supseteq \Theta_{k-1} \cap \Delta_k \tag{9}$$

is satisfied. This is ensured by calculating h_{θ_k} as a solution to the following set of linear programs:

$$
\begin{aligned}
[h_{\theta_k}]_i &= \max_{\theta \in \mathbb{R}^p} \quad [H_\theta]_i \theta \\
&\text{s. t.} \quad \begin{bmatrix} H_\theta \\ H_\Delta \end{bmatrix} \theta \le \begin{bmatrix} h_{\theta_{k-1}} \\ h_\Delta \end{bmatrix}, \quad i = 1, 2, \ldots, n_\theta.
\end{aligned}
\tag{10}
$$

III. ROBUST STATE TUBE AND CONSTRAINTS

A. Tube MPC

To ensure robust constraint satisfaction, the tube MPC approach proposed in [16] is used. The prediction horizon of the MPC problem is N, and the control input is parameterized using a feedback gain K as

$$u_{l|k} = Kx_{l|k} + v_{l|k}, \tag{11}$$

where $\{v_{l|k}\}_{l=0}^{N-1}$ are decision variables in the MPC optimization problem.

Assumption 2: The feedback gain K is chosen such that $A_{\text{cl}}(\theta) = A(\theta) + B(\theta)K$ is asymptotically stable $\forall \theta \in \Theta$. The gain K can be computed using standard robust control techniques, for example, following the approach in [17]. A state tube is defined using the set-based dynamics

$$
\begin{aligned}
\mathbb{X}_{0|k} &\ni \{x_k\}, \tag{12a} \\
\mathbb{X}_{l+1|k} &\supseteq A(\theta)\mathbb{X}_{l|k} \oplus B(\theta)u_{l|k} \oplus \mathbb{W}, \tag{12b} \\
&\forall \theta \in \Theta_k, \quad l = 0, 1, \ldots, N-1,
\end{aligned}
$$

which ensures that $x_{l|k} \in \mathbb{X}_{l|k}$ for all the realizations of uncertainty and disturbance. The tube cross-section at each time step, $\mathbb{X}_{l|k}$, is parameterized by translation and scaling of the set

$$\mathbb{X}_0 := \{x | H_x x \le \mathbf{1}\} = \text{co}\{x^1, x^2, \ldots, x^v\}, \tag{13}$$

where the vertices $\{x^1, x^2, \ldots, x^v\}$ and the matrix $H_x \in \mathbb{R}^{n_x \times n}$ are computed offline. The variables $z_{l|k} \in \mathbb{R}^n$ and

$\alpha_{l|k} \in \mathbb{R}_{\geq 0}$ define the translation and scaling of $\mathbb{X}_0$ respectively, and are decision variables in the MPC optimization. Then, for $l = \mathbb{N}_0^N$, the state tube is parameterized as

$$\mathbb{X}_{l|k} = \{z_{l|k}\} \oplus \alpha_{l|k}\mathbb{X}_0 \quad = \{x|H_x(x - z_{l|k}) \leq \alpha_{l|k}\mathbf{1}\}$$
$$= \{z_{l|k}\} \oplus \alpha_{l|k}\text{co}\{x^1, x^2, \ldots, x^v\}. \tag{14}$$

B. Reformulation of constraints

The state and input constraints defined in (4) and the set dynamics proposed in (12) must be robustly satisfied for all $\theta \in \Theta_k$ and disturbances in $\mathbb{W}$. To reformulate these in a convex manner, the following notation is defined

$$x_{l|k}^j = z_{l|k} + \alpha_{l|k}x^j, \quad d_{l|k}^j = A_0 x_{l|k}^j + B_0 u_{l|k}^j - z_{l+1|k},$$
$$u_{l|k}^j = K x_{l|k}^j + v_{l|k}, \quad D_{l|k}^j = D(x_{l|k}^j, u_{l|k}^j), \tag{15}$$

where $j \in \mathbb{N}_1^v, l \in \mathbb{N}_0^{N-1}$. Note that unlike the definition in (7) where D_t, d_t are a function of known states and inputs, the quantities $D_{l|k}^j, d_{l|k}^j$ linearly depend on the decision variables of MPC. Additionally, the vectors $\bar{f}$ and $\bar{w}$ are computed offline such that for $i \in \mathbb{N}_1^{n_c}, j \in \mathbb{N}_1^{n_x}$

$$[\bar{f}]_i = \max_{x \in \mathbb{X}_0}[F + GK]_i x, \quad [\bar{w}]_j = \max_{w \in \mathbb{W}}[H_x]_j w. \tag{16}$$

The following proposition from [5] reformulates the robust constraints and set-dynamics as linear equality and inequality constraints.

Proposition 1: Let the state tube $\{\mathbb{X}_{l|k}\}_{l=0}^N$ be parameterized according to (14). Then, the constraints (4) and set-dynamics (12) are satisfied if and only if $\forall j \in \mathbb{N}_1^v, l \in \mathbb{N}_0^{N-1}$ there exists $\Lambda_{l|k}^j \in \mathbb{R}_{\geq 0}^{n_x \times n_\theta}$ such that

$$(F + GK)z_{l|k} + Gv_{l|k} + \alpha_{l|k}\bar{f} \leq \mathbf{1}, \tag{17a}$$
$$-H_x z_{0|k} - \alpha_{0|k}\mathbf{1} \leq -H_x x_k, \tag{17b}$$
$$\Lambda_{l|k}^j h_{\theta_k} + H_x d_{l|k}^j - \alpha_{l+1|k}\mathbf{1} \leq -\bar{w}, \tag{17c}$$
$$H_x D_{l|k}^j = \Lambda_{l|k}^j H_\theta. \tag{17d}$$

C. Terminal set

To obtain an MPC algorithm which ensures recursive feasibility, the state tube is directed to a terminal set. The terminal constraints are imposed on $z_{N|k}$ and $\alpha_{N|k}$ since they define the last cross section of the state tube.

Assumption 3: There exists a nonempty terminal set $\mathbb{X}_T = \{(z, \alpha) \in \mathbb{R}^n \times \mathbb{R}| z=0, \alpha \in [0, \bar{\alpha}]\}$, such that for all $\theta \in \Theta$ it holds that

$$\alpha \in [0, \bar{\alpha}] \implies \exists \alpha^+ \in [0, \bar{\alpha}] \text{ s.t.}$$
$$A_{\text{cl}}(\theta)(\alpha\mathbb{X}_0) \oplus \mathbb{W} \subseteq \alpha^+\mathbb{X}_0,$$
$$\alpha \in [0, \bar{\alpha}] \implies (x, Kx) \in \mathbb{Z} \quad \forall x \in \alpha\mathbb{X}_0.$$

Assumption 3 implies that the set $\mathbb{X}_T$ is a robust positively invariant (RPI) set for the set-dynamics in (z, α), with an additional constraint that the set $\mathbb{X}_{N|k}$ remains centered at origin. Note that Assumption 2 is a necessary condition for Assumption 3 to be satisfied, but they are stated separately to emphasize that the stronger assumption is only needed to implement the terminal condition. Thus, the terminal constraint for the MPC algorithm is

$$z_{N|k} = 0; \quad \alpha_{N|k} \leq \bar{\alpha}. \tag{18}$$

IV. PREDICTED STATE TUBE FOR EXPLORATION

In this section predicted variables and sets are defined, which are analogous to the variables and sets defined in the previous section. Each of the predicted quantities is denoted with a hat (^) on the top. Using the predicted sets and variables, the dual effect of the input is captured by the MPC optimization problem. The control input u_k and the future parameter set Θ_{k+1} are connected by the identification step (9). By predicting the next state measurement $\hat{x}_{1|k}$, a predicted parameter set $\hat{\Theta}_k$ is defined as a function of u_k. A predicted state tube is then constructed to contain all the state trajectories generated by the predicted parameter set.

A. Predicted parameter set

To predict the next measurement, an estimate of the parameters is required. For this purpose, a least mean squares filter is used to calculate $\hat{\theta}_k$ as an estimate of θ^* [10]. Alternative filters such as recursive least squares can also be used for this purpose. Using $\hat{\theta}_k$, the predicted state measurement at the next time step can be written as

$$\hat{x}_{1|k} = A(\hat{\theta}_k)x_k + B(\hat{\theta}_k)u_k. \tag{19}$$

Using the predicted state $\hat{x}_{1|k}$, the future constraints on the parameter set Θ_k are given by

$$\hat{\Delta}_k := \{\theta \in \mathbb{R}^p | \hat{x}_{1|k} - A(\theta)x_k - B(\theta)u_k \in \mathbb{W}\},$$
$$= \{\theta \in \mathbb{R}^p | -H_w D_k \theta \leq h_w - H_w D_k \hat{\theta}_k\}, \tag{20}$$

where $u_k = Kx_k + v_{0|k}$ is the first control input calculated by the MPC controller. Since u_k is the only input applied in closed loop, the predicted constraints from other inputs are not considered. A predicted parameter set $\hat{\Theta}_k \subseteq \Theta_k$ can now be defined as

$$\hat{\Theta}_k := \Theta_k \cap \hat{\Delta}_k = \left\{\theta \in \mathbb{R}^p \middle| \begin{array}{c} H_\theta\theta \leq h_{\theta_k} \\ -H_w D_k\theta \leq h_w - H_w D_k\hat{\theta}_k \end{array}\right\}$$
$$= \{\theta \in \mathbb{R}^p | \hat{H}_\theta\theta \leq \hat{h}_{\theta_k}\}. \tag{21}$$

Note that $\hat{\Theta}_k$ is dependent on the control input through the definition of D_k (7), but this dependence is omitted for clarity. This captures the effect of the control input on the identification. Figure 1a shows the parameter estimate $\hat{\theta}_k$, the parameter sets Θ_k and $\hat{\Theta}_k$, and the predicted constraints.

B. Predicted state tube

Using the set $\hat{\Theta}_k$, a predicted state tube of length $\hat{N} \leq N$ is constructed such that $\forall\theta \in \hat{\Theta}_k$ and $l = \mathbb{N}_0^{\hat{N}-1}$

$$\hat{\mathbb{X}}_{0|k} \ni \{x_k\}, \quad \hat{\mathbb{X}}_{l+1|k} \supseteq A(\theta)\hat{\mathbb{X}}_{l|k} \oplus B(\theta)u_{l|k} \oplus \mathbb{W}. \tag{22}$$

The control input applied is the same for the set-dynamics (12) and (22), while the parameter sets used are different. The evolution of the predicted state tube and the robust state tube is shown in Figure 1b. Since the parameter sets satisfy $\hat{\Theta}_k \subseteq \Theta_k$, the predicted state tube lies within the robust state tube $\left(\hat{\mathbb{X}}_{l|k} \subseteq \mathbb{X}_{l|k}\right)$. Each set in the predicted state tube
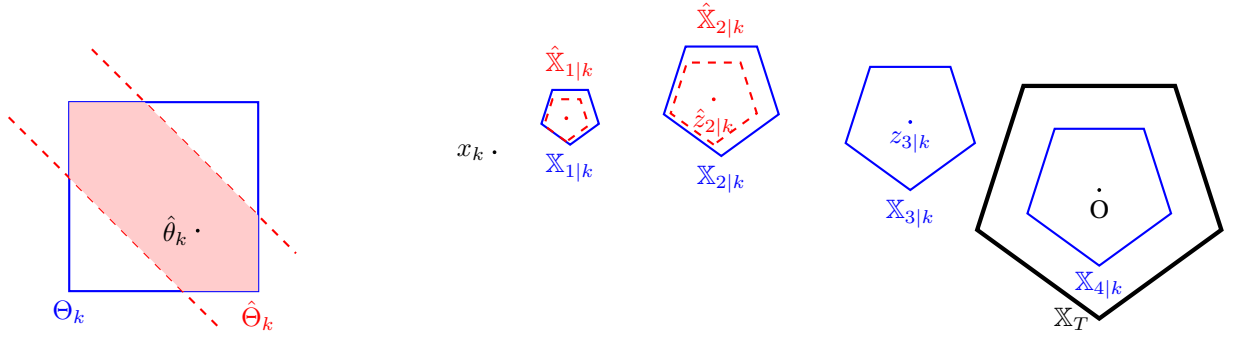
(a) The estimated parameter is $\hat{\theta}_k$. The parameter set Θ_k is bounded by the blue constraints and the dashed lines represent the predicted constraints. The shaded region shows the predicted parameter set $\hat{\Theta}_k$.

(b) The state tube $\{\mathbb{X}_{l|k}\}_{l=1}^4$ is shown in blue and predicted state tube $\{\hat{\mathbb{X}}_{l|k}\}_{l=1}^2$ is shown in red (dashed). The values of N and $\hat{N}$ are 4 and 2 respectively. The depiction of the terminal set $\mathbb{X}_T$ in $\mathbb{R}^n$ is shown in black and contains the set $\mathbb{X}_{N|k}$ centered at origin.

Fig. 1: Depiction of parameter set, predicted parameter set, state tube and predicted state tube

is parameterized as $\hat{\mathbb{X}}_{l|k} = \{\hat{z}_{l|k}\} \oplus \hat{\alpha}_{l|k}\mathbb{X}_0$ where $\hat{z}_{l|k} \in \mathbb{R}^n, \hat{\alpha}_{l|k} \in \mathbb{R}_{\geq 0}$, for $l = \mathbb{N}_0^{\hat{N}}$ are decision variables in the MPC optimization problem. To reformulate the predicted set-dynamics in (22), the following definitions are used for $j \in \mathbb{N}_1^v, l \in \mathbb{N}_0^{\hat{N}-1}$

$$\hat{x}_{l|k}^j = \hat{z}_{l|k} + \hat{\alpha}_{l|k}x^j, \quad \hat{d}_{l|k}^j = A_0 x_{l|k}^j + B_0 u_{l|k}^j - \hat{z}_{l+1|k},$$
$$\hat{u}_{l|k}^j = K\hat{x}_{l|k}^j + v_{l|k}, \quad \hat{D}_{l|k}^j = D(\hat{x}_{l|k}^j, \hat{u}_{l|k}^j). \tag{23}$$

The next proposition formulates the dynamics of the predicted state tube as constraints. The proof is similar to Proposition 1 and is omitted.

Proposition 2: The predicted state tube $\{\hat{\mathbb{X}}_{l|k}\}_{l=0}^{\hat{N}-1}$ satisfies the set-dynamics (22) if and only if for all $j \in \mathbb{N}_1^v$ and $l \in \mathbb{N}_0^{\hat{N}-1}$ there exists $\hat{\Lambda}_{l|k}^j \in \mathbb{R}_{\geq 0}^{n_x \times (n_\theta + n_w)}$ such that

$$-H_x \hat{z}_{0|k} - \hat{\alpha}_{0|k}\mathbf{1} \leq -H_x x_k, \tag{24a}$$
$$\hat{\Lambda}_{l|k}^j \hat{h}_{\theta_k} + H_x \hat{d}_{l|k}^j - \hat{\alpha}_{l+1|k}\mathbf{1} \leq -\bar{w}, \tag{24b}$$
$$H_x \hat{D}_{l|k}^j = \hat{\Lambda}_{l|k}^j \hat{H}_\theta. \tag{24c}$$

The constraints are bilinear in the variables since $\hat{H}_\theta, \hat{h}_{\theta_k}$ are linearly dependent on the control input u_k as seen in (21).

C. Predicted worst-case cost

The cost function to be minimized is

$$J(\mathbf{v}, \hat{N}, N) = \sum_{i=0}^{\hat{N}} l(\hat{\mathbb{X}}_{i|k}, v_{i|k}) + \sum_{i=\hat{N}+1}^{N} l(\mathbb{X}_{i|k}, v_{i|k}), \tag{25}$$

where $l(\mathbb{X}, v) = \max_{x \in \mathbb{X}} ||Qx||_\infty + ||R(Kx + v)||_\infty$, and Q, R are positive definite matrices. A linear cost is chosen so that it can be reformulated using linear inequalities. The cost function is the sum of the predicted worst-case cost over the horizon $\hat{N}$ and the worst-case cost over the remaining prediction horizon. This combination is used because propagating the predicted state tube is computationally expensive due to the bilinear constraints (24), and results in a non-convex optimization problem. The parameter $\hat{N}$ offers a trade-off between the computational complexity and active

---

Algorithm 1 Adaptive MPC with active exploration

Offline Choose K, $\bar{\alpha}$ and $\mathbb{X}_0$. Initialize h_{θ_k} and $\hat{\theta}_k$.
Online
1: $k \leftarrow 1$
2: **repeat**
3:      Obtain the measurement x_k
4:      Construct Δ_k according to (6)
5:      Update h_{θ_k} using (10) and compute $\hat{\theta}_k$
6:      Solve optimization problem (27)
7:      Apply the control input $u_k = Kx_k + v_{0|k}$
8:      $k \leftarrow k + 1$
9: **until**

---

exploration. A higher value of $\hat{N}$ increases the effect of a smaller parameter set $\hat{\Theta}_k$ and thus promotes exploration. However, it also increases the number of bilinear constraints and the computational complexity. This trade-off will be exemplified in the numerical tests shown in Section V.

D. MPC algorithm

The MPC optimization can now be defined using all the elements described above. The decision variables are

$$q_k = \begin{Bmatrix} \{z_{l|k}, \alpha_{l|k}, \{\Lambda_{l|k}^j\}_{j=1}^v\}_{l=0}^N, \{v_{l|k}\}_{l=0}^{N-1}, \\ \{\hat{z}_{l|k}, \hat{\alpha}_{l|k}, \{\hat{\Lambda}_{l|k}^j\}_{j=1}^v\}_{l=0}^{\hat{N}} \end{Bmatrix}, \tag{26}$$

and the optimization problem can be written as

$$\begin{aligned} & \text{minimize } J(25) \\ & \text{s.t. } q_k \in \mathcal{Q}_k := \{q_k | (17), (18), (24)\}, \end{aligned} \tag{27}$$

where $\mathcal{Q}_k$ represents the feasible region at time step k. The adaptive MPC algorithm with active exploration is described in Algorithm 1. The values of the prestabilizing gain K, the state tube shape $\mathbb{X}_0$, terminal set bound $\bar{\alpha}$ must be computed offline. The value of h_{θ_k} is initialized according to (3), and an initial guess is used for $\hat{\theta}_k$. The following proposition establishes the control theoretic properties of the algorithm

Proposition 3: Let the assumptions 1,2 and 3 be satisfied and an initial feasible solution exist, that is, $\mathcal{Q}_0 \neq \{\varnothing\}$.

Then, the closed loop system using Algorithm 1 satisfies the following properties for all $k > 0$:

(i) $\theta^* \in \Theta_k$

(ii) $\mathcal{Q}_k \neq \{\varnothing\}$

(iii) $(x_k, u_k) \in \mathbb{Z}$.

Proof: The properties (i) and (ii) are proven by induction. Let $\theta^* \in \Theta_k$ for some $k \geq 0$ which implies $\theta^* \in \Delta_k$ according to (9). Since $w_k \in \mathbb{W}$, the definition of the non-falsified set Δ_{k+1} in (6) implies $\theta^* \in \Delta_{k+1}$. This is because Δ_{k+1} is constructed using the $s - 1$ measurements used in Δ_k. Applying (9) for the time step $k+1$ proves $\theta^* \in \Theta_{k+1}$.

Property (ii) implies recursive feasibility, i.e., if the MPC problem is feasible at the first time step, it remains feasible. For the proof, assume there exists a feasible solution at time step $k \geq 0$. It is sufficient to find a feasible solution for the control variables $\{v_{l|k}\}_{l=0}^{N-1}$ and state tube variables $\{z_{l|k}, \alpha_{l|k}, \{\Lambda_{l|k}^j\}_{j=1}^v\}_{l=0}^{N-1}$ at the next time step to prove that the optimization problem is feasible. This is because the state tube satisfies the set-dynamics of the predicted state tube (22). Consider the state tube $\{\mathbb{X}_{l|k}\}_{l=0}^N$ computed at time step k. Assumption 3 implies that the feedback controller $u = Kx$ maps the set $\mathbb{X}_{N|k}$ to a set $\alpha^+ \mathbb{X}_0$, where $\alpha^+ \in [0, \bar{\alpha}]$. Since the relation $\Theta_k \supseteq \Theta_{k+1}$ holds according to (9), a feasible sequence of control inputs at the next time step is

$$\{v_{l|k+1}\}_{l=0}^{N-2} = \{v_{l+1|k}\}_{l=0}^{N-2}, \quad v_{N-1|k+1} = 0.$$

and a feasible sequence of sets defining the state tube is

$$\{\mathbb{X}_{l|k+1}\}_{l=0}^{N-1} = \{\mathbb{X}_{l+1|k}\}_{l=0}^{N-1}, \quad \mathbb{X}_{N|k+1} = \alpha^+ \mathbb{X}_0.$$

Property (iii) is a direct result of Proposition 1 and recursive feasibility. ∎

V. NUMERICAL RESULTS

In this section, the performance of the dual adaptive MPC (DAMPC) algorithm presented in this paper is compared to a passive adaptive MPC (PAMPC) algorithm from [5]. The DAMPC algorithm performs active exploration using a predicted state tube, and the dependence of the performance its length $\hat{N}$ is studied. The PAMPC algorithm uses the MPC cost function with $\hat{N}$ set to 0. The system matrices used in the simulation are given by

$$A_0 = \begin{bmatrix} 0.85 & 0.5 \\ 0.2 & 0.6 \end{bmatrix}, A_1 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$
$$B_0 = \begin{bmatrix} 1 & 0.4 \\ 0.2 & 0.4 \end{bmatrix}, B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0.5 \\ 0 & 0.4 \end{bmatrix}.$$

The uncertainty in the parameters is described by $\Theta = \{\theta \in \mathbb{R}^2 \mid \|\theta\|_\infty \leq 1\}$, with $\theta^* = [0.95, 0.3]^\top$. The disturbance set is $\mathbb{W} = \{w \in \mathbb{R}^2 \mid \|w\|_\infty \leq 0.1\}$ and the state and input constraints are described by

$$\mathbb{Z} = \left\{ (x, u) \in \mathbb{R}^{2 \times 2} \,\middle|\, \begin{array}{c} \|x\|_\infty \leq 10 \\ -0.5 \leq [u]_1 \leq 1, \, -2 \leq [u]_2 \leq 2 \end{array} \right\}.$$

The initial state of the system is $x_0 = [1, 1.5]^\top$. In both PAMPC and DAMPC, the state tube is constructed by translating and scaling the set $\mathbb{X}_0 = \{x \in \mathbb{R}^2 \mid \|x\|_\infty \leq 1\}$. The bounded complexity update of Θ_k is performed using
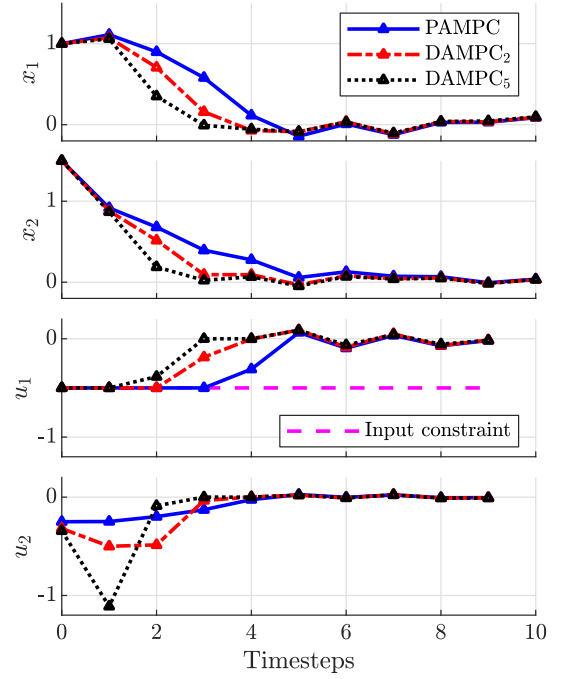

Fig. 2: Closed loop trajectories achieved under PAMPC and two DAMPC schemes with predicted state tube length $\hat{N} = 2, 5$.
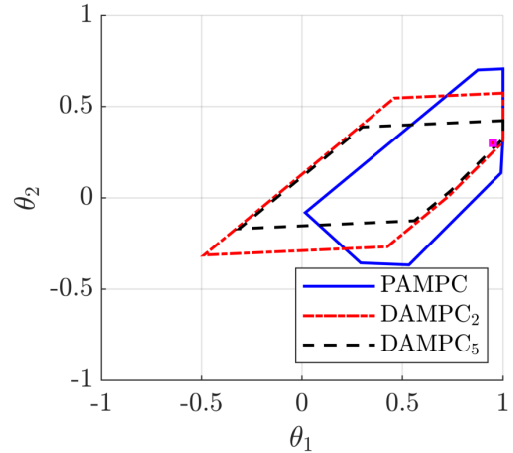

Fig. 3: Parameter sets obtained after running the adaptive MPC schemes for 10 timesteps. The square represents the true parameter.

$n_\theta = 58$ hyperplanes which are initially chosen as outer bounds of the set Θ. The cost matrices are given as $Q = R = \mathbb{I}_{2 \times 2}$, the prestabilizing gain used is

$$K = \begin{bmatrix} -0.5625 & 0 \\ 0 & 0 \end{bmatrix},$$

and the corresponding terminal set bound $\bar{\alpha}$ is 0.89. The prediction horizon chosen is $N = 8$ time steps for all the algorithms. Two different values of the $\hat{N}$ are used, and the corresponding adaptive MPC schemes are referred to as DAMPC$_2$ and DAMPC$_5$ for $\hat{N} = 2, 5$ respectively. The DAMPC schemes are initialized at $\hat{\theta}_0 = [0.5, 0.5]^\top$.

The closed loop trajectories using each of the controllers are shown in Figure 2. It can be seen that active exploration
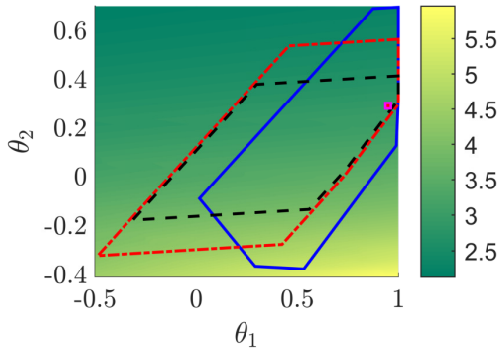
Fig. 4: Colormap showing the distribution of closed loop costs of MPC controllers as a function of the parameters. The legend is same as in Figure 3, and is omitted for readability.

improves the regulation performance. The PAMPC scheme achieves a closed loop cost of 6.04, while DAMPC$_2$ and DAMPC$_5$ achieve 4.49 (25% lower) and 4.21 (30% lower) respectively. The coefficients of control input $[u]_2$ have high uncertainty, and the PAMPC algorithm does not excite this input since the MPC optimizer within does not explicitly include the benefit of online identification. However, the DAMPC algorithms use a higher value of $[u]_2$ which improves the identification and reduces the closed loop cost. The updated uncertainty set Θ_k of each scheme after 10 time steps is shown in Figure 3. Even though the uncertainty sets for DAMPC$_2$ and PAMPC have similar size, the DAMPC$_2$ algorithm has a lower uncertainty in the parameter $[\theta]_2$ which has a stronger influence on the performance. This can be interpreted using Figure 4, which shows the closed loop cost of an MPC controller specifically designed for each plant in the uncertainty set, plotted as a function of the corresponding parameters. Since the goal is to investigate the exploratory actions of the three aforementioned controllers, only the region around the uncertainty sets depicted in Figure 3 is considered. Figure 4 reveals the relative difficulty in controlling the systems, and thus motivates why some regions of parameter space are removed from the uncertainty set Θ_k rather than the others. The figure shows that compared to the DAMPC$_2$ controller, the PAMPC controller results in a parameter set associated with worse performance. This is because the predicted worst-case cost function induces exploration so as to remove the systems difficult to control from the future parameter set. Additionally, it can be seen that using a larger $\hat{N}$ improves exploration. The DAMPC$_5$ algorithm has the smallest uncertainty set, while also having the least closed loop cost. The cost-reduction offered by DAMPC schemes is achieved at the price of computational complexity. The simulations were performed on a laptop using Intel i7-8550U 1.8 GHz processor, and the optimization problems were setup using YALMIP [18] and solved using IPOPT [19]. The average solver time for the optimization problem in PAMPC was 0.042s, while that of DAMPC$_2$ and DAMPC$_5$ were 0.89s and 1s respectively. A similar trend was observed for the performance and solver times with different values of the predicted state tube length.

VI. CONCLUSION

A dual adaptive MPC scheme was presented for systems with parametric uncertainty in state-space matrices. The algorithm uses online set-membership identification to reduce the uncertainty in the parameters and a tube MPC approach to ensure robust constraint satisfaction. A predicted state-tube is used to capture the effect of the future control inputs on identification, and a predicted worst-case cost is optimized. The resulting optimization problem in the MPC is non-convex, but offers the flexibility to trade-off the computational complexity with performance. The algorithm ensures recursive feasibility and consistency of the parameter set, and performs better compared to a passive adaptive MPC approach from literature while regulating a system.

REFERENCES

[1] S. Sastry and M. Bodson, *Adaptive control: stability, convergence and robustness*. Courier Corporation, 2011.
[2] J. B. Rawlings and D. Q. Mayne, *Model predictive control: Theory and design*. Nob Hill Pub., 2009.
[3] T.-H. Kim and T. Sugie, "Adaptive receding horizon predictive control for constrained discrete-time linear systems with parameter uncertainties," *International Journal of Control*, vol. 81, no. 1, pp. 62–73, 2008.
[4] M. Tanaskovic, L. Fagiano, R. Smith, and M. Morari, "Adaptive receding horizon control for constrained MIMO systems," *Automatica*, vol. 50, no. 12, pp. 3019–3029, 2014.
[5] M. Lorenzen, F. Allgöwer, and M. Cannon, "Adaptive model predictive control with robust constraint satisfaction," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 3313–3318, 2017.
[6] T. A. N. Heirung, B. E. Ydstie, and B. Foss, "Dual adaptive model predictive control," *Automatica*, vol. 80, pp. 340–348, 2017.
[7] A. Parsi, A. Iannelli, M. Yin, M. Khosravi, and R. S. Smith, "Robust adaptive model predictive control with worst-case cost," in *21st IFAC World Congress*, 2020.
[8] B. Wittenmark, "Adaptive dual control methods: An overview," in *In 5th IFAC symposium on Adaptive Systems in Control and Signal Processing*, pp. 67–72, 1995.
[9] A. Mesbah, "Stochastic model predictive control with active uncertainty learning: a survey on dual control," *Annual Reviews in Control*, vol. 45, pp. 107–117, 2018.
[10] M. Lorenzen, M. Cannon, and F. Allgöwer, "Robust MPC with recursive model update," *Automatica*, vol. 103, pp. 461–471, 2019.
[11] A. Weiss and S. Di Cairano, "Robust dual control MPC with guaranteed constraint satisfaction," in *53rd IEEE Conference on Decision and Control*, pp. 6713–6718, IEEE, 2014.
[12] C. A. Larsson, A. Ebadat, C. R. Rojas, X. Bombois, and H. Hjalmarsson, "An application-oriented approach to dual control with excitation for closed-loop identification," *European Journal of Control*, vol. 29, pp. 1–16, 2016.
[13] A. Iannelli and R. S. Smith, "A multiobjective lqr synthesis approach to dual control for uncertain plants," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 952–957, 2020.
[14] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.
[15] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: an overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.
[16] W. Langson, I. Chryssochoos, S. Raković, and D. Q. Mayne, "Robust model predictive control using tubes," *Automatica*, vol. 40, no. 1, pp. 125–133, 2004.
[17] M. V. Kothare, V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361–1379, 1996.
[18] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *In Proceedings of the CACSD Conference*, (Taipei, Taiwan), 2004.
[19] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.