

Adaptive Learning-based Model Predictive Control for Uncertain Interconnected Systems: A Set Membership Identification Approach [★]

Ahmed Aboudonia ^a, John Lygeros ^a

^a*Automatic Control Laboratory, ETH Zurich*

Abstract

We propose a novel adaptive learning-based model predictive control (MPC) scheme for interconnected systems which can be decomposed into several smaller dynamically coupled subsystems with uncertain coupling. The proposed scheme is mainly divided into two main online phases; a learning phase and an adaptation phase. Set membership identification is used in the learning phase to learn an uncertainty set that contains the coupling strength using online data. In the adaptation phase, rigid tube-based robust MPC is used to compute the optimal predicted states and inputs. Besides computing the optimal trajectories, the MPC ingredients are adapted in the adaptation phase taking the learnt uncertainty set into account. These MPC ingredients include the prestabilizing controller, the rigid tube, the tightened constraints and the terminal ingredients. The recursive feasibility of the proposed scheme as well as the stability of the corresponding closed-loop system are discussed. The developed scheme is compared in simulations to existing schemes including robust, adaptive and learning-based MPC.

Key words: Model predictive Control; Optimization-based Control; Learning-based Control.

1 Introduction

Model Predictive Control (MPC) is an advanced control scheme that optimally controls a given system by repeatedly solving an online optimal control problem (OCP) (Kouvaritakis & Cannon (2016)). MPC was successfully used in various applications such as interconnected systems (e.g. see Müller & Allgöwer (2017) and the references therein). MPC, however, heavily relies on the availability of an accurate model which is, in many cases, difficult to find and/or complex to use. Various efforts were devoted to finding MPC variants that do not require an accurate model. These variants include robust, stochastic and learning-based MPC (Saltik et al. (2018), Mesbah (2016), Ren et al. (2022)). One drawback of these schemes is that the model is usually not updated online leading to potentially conservative performance. Recently, various adaptive and safe learning-based MPC schemes (Hewing et al. (2020)) have been proposed to ensure safety while enhancing closed-loop performance by improving the model using

data collected online. This goal is achieved by combining physics-based knowledge with machine learning and system identification.

In this paper, we propose a novel adaptive learning-based MPC scheme for interconnected systems which can be decomposed into several smaller subsystems. Each subsystem is dynamically coupled with a set of other subsystems called the set of neighbours. The source of uncertainty in each subsystem is the strength of the coupling with its neighbours. The proposed scheme learns the uncertainty set online by applying distributed set membership identification on collected data and adapts the MPC ingredients accordingly. In particular, the developed control scheme adapts the rigid tube, the prestabilizing controller, the tightened constraint sets and the terminal ingredients by considering them as decision variables in the OCP. The developed scheme is compared in simulation to existing schemes including robust MPC, adaptive MPC and learning-based MPC using a network of dynamically coupled double integrators and found to lead to a better trade-off between closed-loop performance and computation cost. The main contributions and challenges of this work can be listed as follows:

- Unlike many MPC schemes which usually only compute optimal state and input trajectories online, we

[★] Funded by NCCR Automation, the Swiss National Science Foundation (Corresponding Author: Ahmed Aboudonia)

Email addresses: ahmedab@ethz.ch (Ahmed Aboudonia), lygeros@ethz.ch (John Lygeros).

formulate an MPC problem which can also compute various MPC ingredients in the adaptation phase. The main challenge here is to maintain the convexity of the OCP since these ingredients should generally satisfy non-convex infinite dimensional conditions.

- Although set membership identification is extensively used for monolithic systems, we extend this approach here to interconnected systems with uncertain couplings where parameter estimation is not handled by a centralized unit.
- We show that the closed-loop system under the developed controller is input-to-state stable. The main challenges here are that the closed-loop system is switched due to updating the MPC ingredients online and that artificial equilibrium points are used. Hence, standard MPC stability arguments do not apply.
- We compare the proposed scheme to other off-the-shelf MPC schemes which prove to be successful in many applications.

In Section 2, we formulate the considered control problem. In Section 3, we develop the novel adaptive learning-based MPC scheme. In Section 4, we prove the recursive feasibility of the proposed scheme and the input-to-state stability of the corresponding closed-loop system. In Section 5, we validate the efficacy of the developed scheme by comparing it to other existing schemes in the literature. Finally, we draw some conclusions in Section 6. In the sequel, we refer to symmetric, symmetric positive semidefinite and symmetric positive definite matrices of size n as $\mathbb{S}^n$, $\mathbb{S}_+^n$ and $\mathbb{S}_{++}^n$, respectively. Similarly, we denote diagonal, diagonal positive semidefinite and diagonal positive definite matrices of size n by $\mathbb{D}^n$, $\mathbb{D}_+^n$ and $\mathbb{D}_{++}^n$, respectively. Additionally, we use $\oplus$ and $\ominus$ to refer to the Minkowski set addition and the Pontryagin set difference, respectively.

2 Problem Formulation

We consider discrete-time linear time-invariant interconnected systems with a sparse structure. We assume that such systems can be decomposed into M dynamically-coupled subsystems and that the i -th subsystem has a set of in-neighbours $\mathcal{N}_i^-$ including all subsystems which affect its dynamics and a set of out-neighbours $\mathcal{N}_i^+$ including all subsystems whose dynamics are affected by the i -th subsystem. We assume that the dynamics of the i -th subsystem is given by,

$$z_i(t+1) = A_i z_i(t) + B_i v_i(t) + E_i \sum_{j \in \mathcal{N}_i^-} a_{ij} C_j z_j(t) + d_i(t), \quad (1)$$

where $z_i \in \mathbb{R}^{n_i}$, $v_i \in \mathbb{R}^{p_i}$ and $d_i \in \mathbb{R}^{n_i}$ are the state, input and noise vectors of the i -th subsystem, $A_i \in \mathbb{R}^{n_i \times n_i}$, $B_i \in \mathbb{R}^{n_i \times p_i}$, $C_i \in \mathbb{R}^{m_0 \times n_i}$ and $E_i \in \mathbb{R}^{n_i \times m_0}$ are known matrices and $a_{ij} \in \mathbb{R}$ are unknown fixed parameters that represent the coupling strengths. We assume that the scalars a_{ij} satisfy $0 \leq a_{ij}^{min} \leq a_{ij} \leq a_{ij}^{max}$ where the bounds a_{ij}^{min} and a_{ij}^{max} are known. We also assume that

each subsystem is subject to the state and input constraints $z_i(t) \in \mathcal{Z}_i = \{z_i : G_i z_i \leq 1\}$ and $v_i(t) \in \mathcal{V}_i = \{v_i : H_i v_i \leq 1\}$ where $G_i \in \mathbb{R}^{q_i \times n_i}$, $H_i \in \mathbb{R}^{r_i \times p_i}$, $\mathcal{Z}_i$ and $\mathcal{V}_i$ are compact and include the origin in their interior. The noise vector $d_i(t)$ is assumed to lie in a known compact polytopic set $\mathcal{D}_i = \{d_i : \Xi_i d_i \leq 1\}$ that includes the origin in its interior. Following Rivero et al. (2013), we express (1) as,

$$z_i(t+1) = A_i z_i(t) + B_i v_i(t) + E_i w_i(t) + d_i(t), \quad (2)$$

$$w_i(t) = \sum_{j \in \mathcal{N}_i^-} a_{ij} C_j z_j(t) \in \mathcal{W}_i = \bigoplus_{j \in \mathcal{N}_i^-} a_{ij}^{max} C_j Z_j.$$

By denoting the nominal state and input vectors of the i -th subsystem by $x_i \in \mathbb{R}^{n_i}$ and $u_i \in \mathbb{R}^{p_i}$, we also define the nominal dynamics of the i -th subsystem to be, $x_i(t+1) = A_i x_i(t) + B_i u_i(t)$. Assuming that $v_i(t) = u_i(t) + K_i(z_i(t) - x_i(t))$ where $K_i \in \mathbb{R}^{p_i \times n_i}$ is a stabilizing control gain and defining $e_i(t) = z_i(t) - x_i(t)$, we introduce the error dynamics,

$$e_i(t+1) = (A_i + B_i K_i) e_i(t) + E_i w_i(t) + d_i(t). \quad (3)$$

We consider here a tracking problem where the state is required to follow the target point x_{r_i} which is an equilibrium point of the nominal system. The target point x_{r_i} and the corresponding input u_{r_i} are assumed to satisfy the state and input constraints. For this purpose, we define the artificial equilibrium point (x_{e_i}, u_{e_i}) satisfying $x_{e_i} = A_i x_{e_i} + B_i u_{e_i}$ and the cost function $J_i = \sum_{k=0}^{N-1} (\|x_i(k|t) - x_{e_i}\|_{Q_i}^2 + \|u_i(k|t) - u_{e_i}\|_{R_i}^2) + \|x_{e_i} - x_{r_i}\|_{S_i}^2$ where $x_i(k|t)$ and $u_i(k|t)$ are the k -step ahead predicted nominal state and input vectors of the i -th subsystem at the t -th time instant. Note that (x_{e_i}, u_{e_i}) is used to enlarge the feasible region in tracking problems (Limón et al. (2008)) and is considered as an artificial reference that converges asymptotically to the target point. Although the main aim is to drive the actual state to the target point, we use the nominal state in J_i since we can only predict the behavior of the nominal dynamics. Note, however, that the closed-loop stability guarantees derived in Section 4 apply to the actual states. In summary, an MPC problem can be formulated for the i -th subsystem with the local OCP,

$$\min J_i \text{ s.t. } \begin{cases} \text{for all } k \in \{0, \dots, N-1\}, \\ z_i(t) - x_i(0|t) \in \mathcal{Z}_{K_i}, \\ x_i(k+1|t) = A_i x_i(k|t) + B_i u_i(k|t), \\ x_i(k|t) \in \mathcal{X}_i, \quad u_i(k|t) \in \mathcal{U}_i, \\ x_i(N|t) = x_{e_i} = A_i x_{e_i} + B_i u_{e_i}, \\ x_{e_i} \in \xi \mathcal{X}_i, \quad u_{e_i} \in \xi \mathcal{U}_i, \end{cases} \quad (4)$$

where $x_i(k|t)|_{k \in \{0, \dots, N\}}$, $u_i(k|t)|_{k \in \{0, \dots, N-1\}}$, x_{e_i} and u_{e_i} are the decision variables, $z_i(t)$ is the measured state of the i -th subsystem, $\mathcal{Z}_{K_i}$ is a robust positive invariant (RPI) set for its error dynamics (3) under the disturbances $w_i(t)$ and $d_i(t)$, $\xi \in (0, 1)$, $\mathcal{X}_i = \{x_i : G_i x_i \leq g_i\}$ and $\mathcal{U}_i = \{u_i : H_i u_i \leq h_i\}$ are its tightened constraint sets with g_i and h_i positive scalars. The last three constraints in (4) indicate that $x_i(N|t)$ reaches an invariant

terminal set satisfying the state and input constraints. The scalar ξ is used to ensure that the artificial equilibrium lies in the interior of the constraint sets. This requirement is needed later in the stability proof. Note that the global prestabilizing control gain is denoted by $K = \text{diag}(K_1, \dots, K_M)$. Similarly, the global system matrices A and B can be obtained in the obvious way. Following Rivero et al. (2013), the control input $v_i(t) = u_i(0|t) + K_i(z_i(t) - x_i(0|t))$ is applied to (1) to ensure robust constraint satisfaction.

Interconnected systems with uncertain coupling (1) can be exploited in various scenarios with varying-topology networks and large-scale systems. For example, varying-topology networks comprise subsystems which can join and leave the network. While accurate models can be identified offline for the new subsystems, the coupling strength between the network and these subsystems can only be identified online after the plug-in operation. Practical applications comprising such systems include microgrids, power systems and vehicle platoons. Note that considering the whole coupling term as a disturbance might be conservative. This, however, leads to decentralized OCPs rather than distributed ones. Decentralized OCPs are usually more robust against failures and requires less communication and computation. This is more appealing for safety-critical systems with fast dynamics. Moreover, the inclusion of parametric uncertainties within an additive term might be also conservative. An alternative consideration involves employing adaptive MPC tailored for parametric uncertainties, as applied in monolithic systems (Lorenzen et al. (2019)). However, the complexity arises in interconnected systems when uncertainties are embedded within coupling terms, necessitating the formulation of a distributed OCP, which imposes substantial communication and computational burdens. Decentralized OCPs have been widely used with interconnected systems (Rivero et al. (2013), Betti et al. (2013)). Finally, note that coupling constraints among neighbours could in principle be incorporated in decentralized OCPs by finding a decentralized inner-approximation of the constraint sets. This, however, adds to the conservatism of decentralized OCPs.

In this paper, we make use of the OCP (4) to develop the proposed scheme which, at each time instant, performs two phases: a learning phase to learn the model uncertainty and an adaptation phase to compute the optimal control action. We assume that every subsystem can communicate with its neighbours at each time instant. While this assumption is not required for the adaptation phase which relies on the decentralized OCP (4) that requires no communication, it is still required for the learning phase. This is because the model uncertainty in (2) relies on the states of the neighbours. Hence, communication is required among neighbours only once at each time instant in the learning phase. Since communication is allowed by assumption, the decentralized

OCP (4) can in principle be replaced by a distributed OCP. While this may alleviate the conservatism of the developed scheme, the computation and communication complexity will increase significantly since distributed optimization is required.

In the learning phase, we aim to reduce the volume of the uncertainty set $\mathcal{W}_i$ in (2) by learning the bounds a_{ij}^{min} and a_{ij}^{max} of the uncertain parameter a_{ij} as more data is collected online. Hence, we develop a distributed set membership identification approach which learns these bounds by sharing information among the neighbours only once at each time instant. While the set $\mathcal{W}_i$ is a function of a_{ij}^{max} only, a_{ij}^{min} is also updated since it can speed up the shrinkage of $\mathcal{W}_i$ by finding smaller upper bounds. In the adaptation phase, we aim to adapt all MPC ingredients depending on the learnt uncertainty set and compute the local optimal control inputs taking the updated uncertainty set into consideration. The adapted MPC ingredients include the rigid tube represented by $\mathcal{Z}_{K_i}$, the prestabilizing controller K_i , the tightened state constraint set $\mathcal{X}_i$ and the tightened input constraint set $\mathcal{U}_i$. Note that these ingredients are updated at each time instant t based on the learnt uncertainty set $\mathcal{W}_i$ which is a function of the updated parameters a_{ij}^{max} and a_{ij}^{min} at this time instant. Hence, these MPC ingredients together with the uncertainty set are now functions of time.

In the sequel, we use the ellipsoidal rigid tube $\mathcal{Z}_{K_i}(t) = \{e_i : e_i^\top Z_i e_i \leq \alpha_i^2(t)\}$ and the linear prestabilizing controller $\kappa_i(x_i) = K_i(t)x_i$. We assume that $K_i(t) = K_{o_i}T_i(t)$ where K_{o_i} is a full-rank stabilizing control gain and $T_i(t) \in \mathbb{S}_{++}^{n_i}$. Recall also that the tightened polytopic state and control constraint sets are given, respectively, by $\mathcal{X}_i(t) = \{x_i : G_i x_i \leq g_i(t)\}$ and $\mathcal{U}_i(t) = \{u_i : H_i u_i \leq h_i(t)\}$. In addition to the decision variables in (4), $\alpha_i(t)$, $T_i(t)$, $g_i(t)$ and $h_i(t)$ are also considered as decision variables in the developed scheme in order to update the MPC ingredients online; K_{o_i} , on the other hand, is assumed to be known in advance, computed, for example, with methods as those in Aboudonia et al. (2021). When updating these ingredients at each time instant, one should ensure the following conditions,

- I. The set $\mathcal{Z}_{K_i}(t)$ is an RPI set for the dynamics (3).
- II. The local matrices $A_i + B_i K_i(t)$ are Schur.
- III. The global matrix $A + BK(t)$ is Schur.
- IV. The condition $\mathcal{X}_i(t) \oplus \mathcal{Z}_{K_i}(t) \subseteq \mathcal{Z}_i$ holds.
- V. The condition $\mathcal{U}_i(t) \oplus K_i(t)\mathcal{Z}_{K_i}(t) \subseteq \mathcal{V}_i$ holds.

The aforementioned conditions are ensured by adding extra constraints in the OCP of the adaptation phase while preserving its convexity. These additional constraints are derived in the next section. Although the constraint sets are polytopic, the rigid-tube is chosen to be ellipsoidal. As shown in the next section, this leads in the adaptation phase to an OCP which considers the prestabilizing controller as a decision variable while maintaining the convexity of the OCP by making use of

the S-lemma (Boyd et al. (1994)). Recall also the scalability properties of ellipsoidal sets in previous related works when controlling networked dynamical systems (Conte et al. (2016), Parsi et al. (2022)).

The system (1) can be described by a graph $\mathcal{G}(\mathcal{M}, \mathcal{E}, \mathcal{A})$ where $\mathcal{M} = \{1, \dots, M\}$ represents the set of nodes, $\mathcal{E} \subseteq (\mathcal{M} \times \mathcal{M})$ the set of edges and $\mathcal{A} = \{a_{ij} \in \mathbb{R}, (i, j) \in \mathcal{E}\}$ the set of weights. Each subsystem is represented by a node. An edge exists from the j -th node to the i -th node if the j -th subsystem belongs to the set of in-neighbours of the i -th subsystem. We define the adjacency matrix $A_g \in \mathbb{R}^{M \times M}$ of the graph $\mathcal{G}$ such that the element in its i -th row and j -th column is given by $A_{g_{ij}} = a_{ij}$ if $i \in \mathcal{N}_j^+$ and zero otherwise. The degree matrix $D_g \in \mathbb{D}_{++}^M$ of the graph $\mathcal{G}$ is defined such that the diagonal element in the i -th row is given by $\sum_{j \in \mathcal{N}_i^-} a_{ij}$ and the Laplacian matrix is given by $L_g = D_g - A_g \in \mathbb{S}_+^M$. These graph theory concepts are required to locally ensure Condition III that the global matrix $A+BK(t)$ is Schur. The global dynamics is then given by,

$$z(t+1) = Az(t) + Bv(t) + d(t), \quad (5)$$

where $z(t) = [z_1^\top(t), \dots, z_M^\top(t)]^\top$, $v(t) = [v_1^\top(t), \dots, v_M^\top(t)]^\top$, $d(t) = [d_1^\top(t), \dots, d_M^\top(t)]^\top$, $A_d = \text{diag}(A_1, \dots, A_M)$, $\tilde{A}_g = A_g \otimes I_{m_0}$, $A = A_d + E\tilde{A}_gC$, $B = \text{diag}(B_1, \dots, B_M)$, $C = \text{diag}(C_1, \dots, C_M)$ and $E = \text{diag}(E_1, \dots, E_M)$. Note that the proposed method is mainly developed for sparse graphs where only neighbours can communicate. While this method can be used with fully connected graphs where all subsystems can share information with each other, other efficient methods (e.g. Lorenzen et al. (2019)) can be used where all subsystems share information with a central entity that does all the computations instead. In the sequel, we make use of the following definitions.

Definition 1 (Kouvaritakis & Cannon (2016))

A set $Z_K \subset \mathbb{R}^n$ is robustly positively invariant under $z^+ = Az + Bv$, $v = Kz$ and $z \in Z$, $v \in \mathcal{V}$, if and only if $z^+ \in Z_K$, $z \in Z$, $v \in \mathcal{V}$ for all $z \in Z_K$.

Definition 2 (Aliyu (2017)) The system $z^+ = Az + Bv$, $y = Cz + Dv$ is strictly passive if there exist functions $V_p : Z \rightarrow \mathbb{R}$ and $\gamma_p : Z \rightarrow \mathbb{R}_+$ so that $V_p(0) = 0$ and $V_p(z^+) - V_p(z) \leq y^\top v - \gamma_p(z)$ for all $v \in \mathcal{V}$.

3 Adaptive Learning-based MPC

In this section, we start by discussing the distributed set membership identification approach used to learn the uncertainty set online in the learning phase. We then derive the constraints added to the local online OCP of each subsystem in the adaptation phase. First, we discuss the learning phase where we aim to learn the uncertainty set $\mathcal{W}_i(t)$ in real time, as more data is collected online. According to (2), this set is a function of the bounds

of the uncertain parameters. Hence, a better estimation of these bounds leads to a better description of this set. Thus, we develop here a distributed set-membership identification technique to improve the estimates of these bounds. Although such techniques are widely used with monolithic systems (e.g. see Lorenzen et al. (2019)), we extend this technique here to interconnected systems with uncertain couplings where the identification is not carried out by a central entity.

Assume that the bounds of the uncertain parameter a_{ij} at the $(t-1)$ -th time instant are $a_{ij}^{min}(t-1)$ and $a_{ij}^{max}(t-1)$. Our goal is to improve the estimates of these bounds at the t -th time instant. For this purpose, we construct for every a_{ij} the feasible parameter set defined by the inequality $F_{ij}a_{ij} \leq f_{ij}(t-1)$ where $F_{ij} = [1 \ -1]^\top$ and $f_{ij}(t-1) = [a_{ij}^{max}(t-1) \ a_{ij}^{min}(t-1)]^\top$. We then define the vector θ_i which comprises all the uncertain parameters to be identified by the i -th subsystem (i.e. a_{ij} , $j \in \mathcal{N}_i^-$). In this case, the feasible parameter set for the i -th subsystem is given by $F_i\theta_i \leq f_i(t-1)$ where F_i is a block diagonal matrix with the matrices F_{ij} for all $j \in \mathcal{N}_i^-$ on the diagonal and f_i is a vector vertically concatenating the subvectors f_{ij} for all $j \in \mathcal{N}_i^-$.

Recall that each subsystem can communicate with its neighbours in the learning phase. Hence, the states of the neighbours of the i -th subsystem at the $(t-1)$ -th time instant are available for the i -th subsystem at the t -th time instant. Furthermore, the states and inputs of the i -th subsystem at the $(t-1)$ -th time instant are available. Finally, the states of the i -th subsystem at the t -th time instant can be measured. Recall also that $\mathcal{D}_i = \{d_i : \Xi_i d_i \leq 1\}$. Defining $\chi_i(t-1)$ to be a matrix horizontally concatenating the vectors $C_j z_j(t-1)$ for all $j \in \mathcal{N}_i^-$, $\Omega_i(t-1) = -\Xi_i E_i \chi_i(t-1)$ and $\omega_i(t-1) = 1 + \Xi_i A_i z_i(t-1) + \Xi_i B_i v_i(t-1) - \Xi_i z_i(t)$, we can construct for the i -th subsystem the non-falsified parameter set $\Delta_i = \{\theta_i : \Omega_i(t-1)\theta_i \leq \omega_i(t-1)\}$.

In order to update the bounds of the uncertain parameters, the i -th subsystem should solve $2|\mathcal{N}_i^-|$ optimization problems at the t -th time instant where $|\cdot|$ denotes set cardinality. Each optimization problem updates one of these bounds. For each $j \in \mathcal{N}_i^-$, the lower and upper bounds of the parameter a_{ij} are given respectively by,

$$a_{ij}^{min}(t) = \min_{a_{ij}} e_{ij}^\top \theta_i \text{ s.t. } \begin{bmatrix} F_i \\ \Omega_i(t-1) \end{bmatrix} \theta_i \leq \begin{bmatrix} f_i(t-1) \\ \omega_i(t-1) \end{bmatrix}, \quad (6a)$$

$$a_{ij}^{max}(t) = \max_{a_{ij}} e_{ij}^\top \theta_i \text{ s.t. } \begin{bmatrix} F_i \\ \Omega_i(t-1) \end{bmatrix} \theta_i \leq \begin{bmatrix} f_i(t-1) \\ \omega_i(t-1) \end{bmatrix}, \quad (6b)$$

where e_{ij} is a unit vector which selects the parameter a_{ij} in the vector θ_i . Note that computing the new lower (upper) bound is a minimization (maximization) problem because we aim to find the smallest (largest) value of the corresponding parameter that lies inside the feasible

parameter set and the non-falsified set. In conclusion, the i -th subsystem learns the uncertainty set $\mathcal{W}_i(t)$ in (2) by solving at the t -th time instant the optimization problems (6) which identify the bounds of the uncertain parameter. These bounds are communicated among the in-neighbours and out-neighbours to be used in the adaptation phase. We also highlight the dependency of the upper bound $a_{ij}^{max}(t)$ on the lower bound $a_{ij}^{min}(t-1)$ which is part of the vector $f_i(t-1)$ in (6b) to support the fact that updating the lower bound might improve the updates of the upper bound. Note that using asymmetric duality of linear programs, (6) can be written compactly as a single optimization problem. We guarantee that the true parameters always lie within the updated bounds in the following theorem.

Theorem 3 *Assume that $a_{ij}^{min}(0) \leq a_{ij} \leq a_{ij}^{max}(0)$. Then, $a_{ij}^{min}(t) \leq a_{ij} \leq a_{ij}^{max}(t)$ for all $t \geq 0$ where $a_{ij}^{min}(t)$ and $a_{ij}^{max}(t)$ are computed using (6).*

PROOF. Note that $z_i(1) = A_i z_i(0) + B_i v_i(0) + E_i \sum_{j \in \mathcal{N}_i^-} a_{ij} C_j z_j(0) + d_i(0)$ and recall that $d_i(0) \in \mathcal{D}_i$ for all $i \in \mathcal{M}$. Thus, the true parameters satisfy $\Xi z_i(1) - \Xi A_i z_i(0) + \Xi B_i v_i(0) + \Xi E_i \sum_{j \in \mathcal{N}_i^-} a_{ij} C_j z_j(0) \leq 1$ and hence, $\Omega_i(0)\theta_i \leq \omega(0)$ for all $i \in \mathcal{M}$. Since $a_{ij}^{min}(0) \leq a_{ij} \leq a_{ij}^{max}(0)$, the true parameters also satisfy $F_i \theta_i \leq f_i(0)$ for all $i \in \mathcal{M}$ by assumption. Since $a_{ij}^{min}(1)$ and $a_{ij}^{max}(1)$ are computed using (6), then these are the smallest and largest values respectively satisfying $\Omega_i(0)\theta_i \leq \omega(0)$ and $F_i \theta_i \leq f_i(0)$ that the true parameters also satisfy and thus, $a_{ij}^{min}(1) \leq a_{ij} \leq a_{ij}^{max}(1)$. By induction, the same results hold for all t . $\square$

As a consequence of the above theorem, it is guaranteed that $a_{ij}^{min}(t) \geq a_{ij}^{min}(t-1)$ because $a_{ij}^{min}(t)$ is computed by minimizing over all possible values satisfying $a_{ij} \geq a_{ij}^{min}(t-1)$. The same arguments apply to the upper bound and hence, $\mathcal{W}_i(t) \subseteq \mathcal{W}_i(t-1)$ for all t (i.e. the volume of the uncertainty set can never increase). Note that the proposed scheme is not necessarily restricted to set membership identification. While this technique is used here to update uncertain parameters in linear coupling terms, the proposed scheme paves the way to develop other schemes which can learn more complex uncertainties with nonlinear couplings and other disturbances. In this case, one can use other machine learning methods which have shown success in learning uncertainty sets. We now move to discussing the adaptation phase. As mentioned in Section 2, the online OCP is solved in the adaptation phase to adapt the MPC ingredients based on the learnt uncertainty set and compute the optimal control inputs. To adapt the MPC ingredients online while ensuring robust constraint satisfaction, the five conditions mentioned in Section 2 should be satisfied. We derive here additional constraints which ensure the satisfaction of these conditions while maintaining the convexity of the OCP (4).

We start with the condition which ensures the robust positive invariance of the set $\mathcal{Z}_{K_i}(t)$. First, we define for the i -th subsystem the output vector $y_i(t) = C_i z_i(t)$ and the constraint set $\mathcal{Y}_i$, obtained from $\mathcal{Z}_i$ in the obvious way, such that $y_i \in \mathcal{Y}_i$. We then compute a minimum volume ellipsoidal set $\hat{\mathcal{Y}}_i = \{y_i : \|y_i - c_{y_i}\|_{Y_i}^2 \leq 1\}$ that contains the set $\mathcal{Y}_i$. For this purpose, the vertex description of $\mathcal{Y}_i$ is needed. Note that this is in general combinatorial in the linear inequalities needed to define $\mathcal{Y}_i$. However, the sets $\mathcal{Y}_i$ of all subsystems do not change over time and hence, their minimum volume ellipsoidal sets $\hat{\mathcal{Y}}_i$ can be computed once offline by solving a convex optimization problem for each. Finally, we construct the outer approximation $\hat{\mathcal{W}}_i(t) = \bigoplus_{j \in \mathcal{N}_i^-} a_{ij}^{max}(t) \hat{\mathcal{Y}}_j$ of the uncertainty set $\mathcal{W}_i(t) = \bigoplus_{j \in \mathcal{N}_i^-} a_{ij}^{max}(t) \mathcal{Y}_j$. Note that $\hat{\mathcal{W}}_i(t)$ is not necessarily an ellipsoidal set, but it is the Minkowski sum of the ellipsoidal sets $\mathcal{W}_{ij}(t) = a_{ij}^{max}(t) \hat{\mathcal{Y}}_j = \{w_{ij} : \|w_{ij} - a_{ij}^{max}(t) c_{y_j}\|_{Y_j}^2 \leq a_{ij}^{max^2}(t)\}$. Similarly, we also compute a minimum volume ellipsoidal set $\hat{\mathcal{D}}_i = \{d_i : \|d_i\|_{\Xi_i}^2 \leq 1\}$ offline for the set $\mathcal{D}_i$ of the i -th subsystem for all $i \in \{1, \dots, M\}$.

In Proposition 4, we consider Condition I in Section 2 and derive constraints which ensure the robust positive invariance of the set $\mathcal{Z}_{K_i}(t)$ for the error dynamics (3) under the disturbances $w_i \in \hat{\mathcal{W}}_i(t)$ and $d_i \in \hat{\mathcal{D}}_i$. Although this might be more conservative than considering the disturbances $w_i \in \mathcal{W}_i(t)$ and $d_i \in \mathcal{D}_i$, the convexity of the resulting OCP is preserved in this case. Another source of conservatism in Proposition 4 is using the approximate S-lemma to reach local convex constraints. Recall here that $K_i(t) = K_{o_i} T_i(t)$. The dependence on time in (7) is omitted in the interest in space

Proposition 4 *The set $\mathcal{Z}_{K_i}(t)$ is robustly positively invariant for the error dynamics (3) with the uncertainty sets $\hat{\mathcal{W}}_i(t)$ and $\hat{\mathcal{D}}_i$ if there exist non-negative scalars $\lambda_i(t)$, $\lambda_{d_i}(t)$ and $\lambda_{ij}(t)$ for all $j \in \mathcal{N}_i^-$ such that (7) holds.*

PROOF. According to Definition 1, the set $\mathcal{Z}_{K_i}(t)$ is robustly positively invariant under the considered dynamics and uncertainty sets if and only if $\|e_i^+\|_{Z_i}^2 \leq \alpha_i^2(t)$ for all e_i satisfying $\|e_i\|_{Z_i}^2 \leq \alpha_i^2(t)$, w_{ij} satisfying $\|w_{ij} - a_{ij}^{max}(t) c_{y_j}\|_{Y_j}^2 \leq a_{ij}^{max^2}(t)$ for all $j \in \mathcal{N}_i^-$ and d_i satisfying $\|d_i\|_{\Xi_i}^2 \leq 1$. By defining s_i such that $e_i = \alpha_i(t) s_i$, the invariance condition becomes $\|(A_i \alpha_i(t) + B_i K_{o_i} T_i(t) \alpha_i(t)) s_i + E_i w_i + d_i\|_{Z_i}^2 \leq \alpha_i^2(t)$ for all s_i satisfying $\|s_i\|_{Z_i}^2 \leq 1$, w_{ij} satisfying $\|w_{ij} - a_{ij}^{max}(t) c_{y_j}\|_{Y_j}^2 \leq a_{ij}^{max^2}(t)$ for all $j \in \mathcal{N}_i^-$ and d_i satisfying $\|d_i\|_{\Xi_i}^2 \leq 1$. We now define the block diagonal matrix $Y_{\lambda_i(t)}$ with the submatrices $\lambda_{ij}(t) Y_j$ for all $j \in \mathcal{N}_i^-$ on the diagonal and the vector $\tau_i(t)$ which vertically concatenates the subvectors $a_{ij}^{max}(t) c_{y_j}$. Using the S-lemma and applying the Schur complement yields (7) where I_{n_i} is an identity matrix of

$$\begin{bmatrix} \lambda_i Z_i & 0 & 0 & 0 & A_i \alpha_i + B_i K_{o_i} T_i \alpha_i^\top \\ * & Y_{\lambda_i} & 0 & -Y_{\lambda_i} \tau_i & \mathcal{E}_i^\top \\ * & * & \lambda_{d_i} \hat{\Xi}_i & 0 & I_{n_i} \\ * & * & * & \alpha_i - \lambda_i - \lambda_{d_i} - \sum_{j \in \mathcal{N}_i^-} \lambda_{ij} (1 - \|c_{y_j}\|_{Y_j}^2) a_{ij}^{max^2} & 0 \\ * & * & * & * & \alpha_i Z_i^{-1} \end{bmatrix} \geq 0. \quad (7)$$

size n_i and $\mathcal{E}_i$ is a matrix horizontally concatenating the matrix E_i for $|\mathcal{N}_i^-|$ times. $\square$

$\alpha_i(t) > 0$ such that,

$$\begin{bmatrix} (P_i - \Gamma_i) \alpha_i & 0.5 C_i^\top \alpha_i & (A_i \alpha_i + B_i K_{o_i} T_i \alpha_i)^\top \\ * & D_i \alpha_i & E_i^\top \alpha_i \\ * & * & P_i^{-1} \alpha_i \end{bmatrix} \geq 0. \quad (8)$$

Note that the dimension of the constraint in (7) increases linearly with respect to the number of neighbours. Rigid tubes are known for their little computation burden in comparison to homothetic tubes (Raković et al. (2012)). Since we use time-varying rigid tubes, one might think that rigid tubes lose their main computational advantage and could be better replaced by homothetic tubes. This is, however, not the case because when using homothetic tubes, the nominal dynamics would be replaced by tube dynamics. Since the prestabilizing controller is a decision variable and the tube is ellipsoidal, the tube dynamics is represented by one matrix inequality for each prediction step. In this case, the problem becomes computationally challenging as the prediction horizon gets longer.

Next, we consider Conditions II and III in Section 2 by deriving constraints ensuring that $A_i + B_i K_i(t)$ for all $i \in \{1, \dots, M\}$ and $A + BK(t)$ are Schur. For this purpose, we rely on passivity arguments for discrete-time interconnected systems. First, we derive a constraint ensuring that the noise-free dynamics of the i -th subsystem (i.e. (1) with $d_i(t) = 0$ for all t) under the controller $u_i(t) = K_{o_i} T_i(t) x_i(t)$ is strictly passive. We show that this implies that $A_i + B_i K_i(t)$ is Schur. We then derive other constraints ensuring that the strict passivity of all subsystems implies the asymptotic stability of the whole system ensuring that $A + BK(t)$ is Schur. To ensure strict passivity of the i -th subsystem, we define the virtual output $\tilde{y}_i(t) = C_i z_i(t) + D_i w_i(t)$ where $D_i \in \mathbb{D}_{++}^{m_o}$ is a decision variable. We also assume that a matrix $P_i \in \mathbb{S}_{++}^{n_i}$ is computed offline together with K_{o_i} , for example as in Aboudonia et al. (2021), so that $V_{p_i}(z_i) = \|z_i\|_{P_i}^2$ is a storage function over the set $\mathcal{Z}_i$. We finally define a minimum volume ellipsoidal set $\hat{\mathcal{Z}}_i = \{z_i : \|z_i - c_{z_i}\|_{\hat{\mathcal{Z}}_i} \leq 1\}$ for the state constraint set $\mathcal{Z}_i$. The dependence on time in (8) is omitted in the interest in space.

Proposition 5 *The noise-free dynamics of the i -th subsystem $z_i^+ = (A_i + B_i K_{o_i} T_i(t)) z_i + E_i w_i$ is strictly passive with respect to $w_i \in \hat{\mathcal{W}}_i(t)$ and $\tilde{y}_i \in C_i \hat{\mathcal{Z}}_i \oplus D_i \hat{\mathcal{W}}_i(t)$ if there exist $\Gamma_i(t) \in \mathbb{D}_{++}^{n_i}$, $D_i(t) \in \mathbb{D}_{++}^{n_i}$, $T_i(t) \in \mathbb{S}_{++}^{n_i}$,*

PROOF. To ensure strict passivity, we have to satisfy the condition in Definition 2 for all $w_i \in \hat{\mathcal{W}}_i(t)$. Recall also that the storage function $V_{p_i}(z_i)$ is defined over the set $\hat{\mathcal{Z}}_i$. Hence, the noise-free dynamics of the i -th subsystem under the controller $u_i(t) = K_{o_i} T_i(t) x_i(t)$ is strictly passive according to Definition 2 if $\|z_i^+\|_{P_i}^2 - \|z_i\|_{P_i}^2 \leq \tilde{y}_i^\top w_i - \|z_i\|_{\Gamma_i}^2$ for all z_i satisfying $\|z_i - c_{z_i}\|_{\hat{\mathcal{Z}}_i}^2 \leq 1$ and w_{ij} satisfying $\|w_{ij} - a_{ij}^{max}(t) c_{y_j}\|_{Y_j}^2 \leq a_{ij}^{max^2}(t)$ for all $j \in \mathcal{N}_i^-$. By substituting for the dynamics, the strict passivity condition is given by $\|z_i\|_{L_{1i}}^2 + 2w_i^\top L_{2i} z_i + \|w_i\|_{L_{3i}}^2 \geq 0$ for all z_i satisfying $\|z_i - c_{z_i}\|_{\hat{\mathcal{Z}}_i}^2 \leq 1$ and w_{ij} satisfying $\|w_{ij} - a_{ij}^{max}(t) c_{y_j}\|_{Y_j}^2 \leq a_{ij}^{max^2}(t)$ for all $j \in \mathcal{N}_i^-$ where $L_{1i} = P_i - \Gamma_i - (A_i + B_i K_{o_i} T_i(t))^\top P_i (A_i + B_i K_{o_i} T_i(t))$, $L_{2i} = 0.5 C_i - E_i^\top P_i (A_i + B_i K_{o_i} T_i(t))$ and $L_{3i} = D_i - E_i^\top P_i E_i$. We now define the block diagonal matrix $Y_{\sigma_i}(t)$ with the submatrices $\sigma_{ij}(t) Y_j$ for all $j \in \mathcal{N}_i^-$ on the diagonal and the vector $\tau_i(t)$ which vertically concatenates the subvectors $a_{ij}^{max}(t) c_{y_j}$ for all $j \in \mathcal{N}_i^-$. We also recall that the ellipsoidal set $\hat{\mathcal{Y}}_i$ contains the origin in its interior and hence, $\|c_{y_i}\|_{Y_i}^2 \leq 1$. Using the S-lemma, applying the Schur complement and multiplying the resulting matrix inequality by $\alpha_i(t)$ thus leads to (8). $\square$

Note that (8) implies that $P_i - (A_i + B_i K_{o_i} T_i(t))^\top P_i (A_i + B_i K_{o_i} T_i(t)) \geq \Gamma_i(t)$ where $\Gamma_i(t) \in \mathbb{D}_{++}^{n_i}$ and hence, $A_i + B_i K_i(t)$ is Schur. One source of conservatism in Proposition 5 is that the local controller is designed to not only stabilize the local dynamics, but also to passivize it. The benefit of passivizing the local dynamics will become obvious in the next proposition. We now consider Condition III in Section 2 by deriving constraints which ensure that the strict passivity of all noise-free subsystems implies the asymptotic stability of the overall noise-free system. For this purpose, we consider the noise-free global dynamics $z(t+1) = A_{cl}(t)z(t) + Ew(t)$ where $A_{cl}(t) = \text{diag}(A_1 + B_1 K_1(t), \dots, A_M + B_M K_M(t))$, $E =$

$\text{diag}(E_1, \dots, E_M)$ and $w(t) = [w_1^\top(t), \dots, w_M^\top(t)]^\top$. We also define the Lyapunov function $V(z) = \|z\|_P^2$ where $P = \text{diag}(P_1, \dots, P_M)$. Furthermore, we define the matrices $\Gamma(t) = \text{diag}(\Gamma_1(t), \dots, \Gamma_M(t))$, $C = \text{diag}(C_1, \dots, C_M)$ and $D(t) = \text{diag}(D_1(t), \dots, D_M(t))$. Recall that A_g and D_g are the adjacency and degree matrices, respectively. The global virtual output vector is given by $\tilde{y}(t) = Cz(t) + Dw(t)$ and the global coupling vector is given by $w(t) = \tilde{A}_g Cz(t)$.

Finally, we construct $A_g^{max}(t)$ and $D_g^{max}(t)$ by replacing all uncertain parameters in A_g and D_g with their upper bounds. Consequently, we define $U = \text{diag}(U_1, \dots, U_M) = C^\top D_g C$, $V = [V_1^\top, \dots, V_M^\top]^\top = C^\top \tilde{A}_g^\top$ and $W = [W_1^\top, \dots, W_M^\top]^\top = \tilde{A}_g C$ where $U_i \in \mathbb{R}^{n_i \times n_i}$, $V_i \in \mathbb{R}^{n_i \times m_0}$ and $W_i \in \mathbb{R}^{m_0 \times n_i}$. We also construct $U^{max}(t)$, $V^{max}(t)$ and $W^{max}(t)$ by replacing all uncertain parameters in U , V and W with their upper bounds. In the sequel, we denote the diagonal element in the j -th row of a matrix by $[\cdot]_j$ and the summation of the absolute values of the entries in the j -th row of a matrix by $|\cdot|_j$. The dependence on time in (9) is omitted in the interest in space.

Proposition 6 *The strict passivity of the noise-free dynamics of all subsystems under the corresponding controller $u_i(t) = K_{o_i} T_i(t) x_i(t)$ (i.e. $z_i(t+1) = (A_i + B_i K_{o_i} T_i(t)) z_i(t) + E_i w_i(t)$) implies the asymptotic stability of the noise-free dynamics of the overall system (i.e. $z(t+1) = A_{cl}(t) z(t) + E w(t)$) if,*

$$([\Gamma_i]_j - \epsilon) \alpha_i \geq |U_i^{max}|_j \alpha_i + |V_i^{max}|_j \alpha_i, \forall j \in \{1, \dots, n_i\}, \\ [D_i]_j \alpha_i \leq \alpha_i / |W_i^{max}|_j, \forall j \in \{1, \dots, m_0\} : |W_i^{max}|_j \neq 0. \quad (9)$$

PROOF. The passivity of the noise-free dynamics of each subsystem implies that $\|z_i^+\|_{P_i}^2 - \|z_i\|_{P_i}^2 \leq \tilde{y}_i^\top w_i - \|z_i\|_{\Gamma_i(t)}^2$. Summing up these inequalities for all subsystems yields $V(z^+) - V(z) \leq \tilde{y}^\top w - \|z\|_{\Gamma(t)}^2$. To ensure asymptotic stability, it suffices to guarantee that $\tilde{y}^\top w - \|z\|_{\Gamma(t)}^2 < 0$. Using the expressions of $\tilde{y}$ and w , this condition can be written as $\|z\|_{\Gamma(t) - C^\top \tilde{A}_g C - C^\top \tilde{A}_g^\top D(t) \tilde{A}_g C}^2 \geq \epsilon \|z\|^2$ where $\epsilon > 0$. Recall that $L_g = D_g - A_g$ and hence, $\|z\|_{\Gamma(t) + C^\top L_g C - C^\top D_g C - C^\top \tilde{A}_g^\top D(t) \tilde{A}_g C}^2 \geq \epsilon \|z\|^2$. Note also that L_g is positive semidefinite by definition and hence, it is sufficient to ensure that $\|z\|_{\Gamma(t) - C^\top D_g C - C^\top \tilde{A}_g^\top D(t) \tilde{A}_g C}^2 \geq \epsilon \|z\|^2$. This condition can be reformulated using the Schur Complement as,

$$\left[\begin{array}{c|c} \Gamma(t) - C^\top D_g C - \epsilon I_n & C^\top \tilde{A}_g^\top \\ \hline * & D(T)^{-1} \end{array} \right] \geq 0. \quad (10)$$

Using diagonal dominance, the constraint (10) is approximated for each subsystem by $[\Gamma_i(t)]_j - \epsilon \geq |U_i|_j + |V_i|_j$

for all $j \in \{1, \dots, n_i\}$ and $[D_i(t)^{-1}]_j \geq |W_i|_j$ for all $j \in \{1, \dots, m_0\}$. Multiplying the former by $\alpha_i(t)$ and the latter by $\alpha_i(t)^{-1}$ and inverting, these inequalities become $([\Gamma_i(t)]_j - \epsilon) \alpha_i(t) \geq |U_i|_j \alpha_i(t) + |V_i|_j \alpha_i(t)$ for all $j \in \{1, \dots, n_i\}$ and $[D_i(t)]_j \alpha_i(t) \leq \alpha_i(t) / |W_i|_j$ for all $j \in \{1, \dots, m_0\}$ such that $|W_i|_j \neq 0$. Note that the resulting inequalities are functions of the uncertain parameters. More conservative inequalities can be obtained by replacing the uncertain parameters with their upper bounds, leading to (9). $\square$

In Proposition 6, (9) ensures that the Lyapunov condition is satisfied for the noise-free global dynamics $z(t+1) = A_{cl}(t) z(t) + E w(t)$ or equivalently, $z(t+1) = (A + BK(t)) z(t)$ and hence, $A + BK(t)$ is Schur. Sources of conservatism in this proposition include implying $\tilde{y}^\top w - \|z\|_{\Gamma(t)}^2 < 0$ instead of $V(z^+) - V(z) < 0$, ignoring the term $C^\top L_g C$, using diagonal dominance and replacing the true values of the parameters with their upper bounds. These sources are utilized to eventually reach local convex constraints. We now move to Conditions IV and V in Section 2. In the sequel, we denote the j -th row of the matrices G_i and H_i by $G_i^{(j)}$ and $H_i^{(j)}$ respectively. The dependence on time is omitted in (13) and (14) in the interest in space.

Proposition 7 *The Minkowski sum of $\mathcal{X}_i(t)$ and $\mathcal{Z}_{K_i}(t)$ is contained in $\mathcal{Z}_i$ if there exist $\mu_{ij}(t) \geq 0$ for all $j \in \{1, \dots, q_i\}$ such that,*

$$\left[\begin{array}{c|c} \mu_{ij}(t) Z_i & -0.5 G_i^{(j)\top} \alpha_i(t) \\ \hline * & 1 - \mu_{ij}(t) - g_i(t) \end{array} \right] \geq 0. \quad (11)$$

PROOF. The Minkowski sum of $\mathcal{X}_i(t)$ and $\mathcal{Z}_{K_i}(t)$ is contained in $\mathcal{Z}_i$ if and only if for all $j \in \{1, \dots, q_i\}$, $G_i^{(j)}(x_{e_i} + x_{s_i}) \leq 1$ for all x_{e_i} such that $\|x_{e_i}\|_{Z_i}^2 \leq \alpha_i^2(t)$ and for all x_{s_i} such that $G_i x_{s_i} \leq g_i(t)$. By defining s_i and t_i such that $x_{e_i} = \alpha_i(t) s_i$ and $x_{s_i} = g_i(t) t_i$, this condition becomes $G_i^{(j)} \alpha_i(t) s_i + G_i^{(j)} g_i(t) t_i \leq 1$ for all $j \in \{1, \dots, q_i\}$, s_i satisfying $s_i^\top Z_i s_i \leq 1$ and t_i satisfying $G_i t_i \leq 1$. Using the S-lemma, this condition is given for all $j \in \{1, \dots, q_i\}$ by,

$$\left[\begin{array}{c|c} \mu_{ij}(t) Z_i & 0 & -0.5 \alpha_i(t) G_i^{(j)\top} \\ \hline * & 0 & \sum_{l=1}^{q_i} 0.5 \mu_{ijl}(t) G_i^{(l)\top} - 0.5 G_i^{(j)} g_i(t) \\ \hline * & * & 1 - \mu_{ij}(t) - \sum_{l=1}^{q_i} \mu_{ijl}(t) \end{array} \right] \geq 0. \quad (12)$$

Note that (12) holds if $\sum_{l=1}^{q_i} 0.5 \mu_{ijl}(t) G_i^{(l)} = 0.5 G_i^{(j)} g_i(t)$. By stacking this equality for all $j \in \{1, \dots, q_i\}$, we reach $g_i(t) G_i = \mu_i(t) G_i$ where $\mu_{ijl}(t)$ is the entry of $\mu_i(t)$ in the j -th row and l -th column. To satisfy this equality, we choose the diagonal elements of $\mu_i(t)$ to be equal to

$g_i(t)$ and the off-diagonal elements to be equal to zero and hence, (12) reduces to (11). $\square$

Proposition 8 *The Minkowski sum of $\mathcal{U}_i(t)$ and $K_{o_i}T_i(t)\mathcal{Z}_{K_i}(t)$ is contained in $\mathcal{V}_i$ if there exist $\nu_{ij}(t) > 0$ for all $j \in \{1, \dots, r_i\}$ and $M_i(t) \in \mathbb{S}_{++}^{p_i}$ such that,*

$$\begin{bmatrix} M_i^{-1} & K_i^\top \alpha_i \\ * & \nu_{ij} Z_i \end{bmatrix} \geq 0, \begin{bmatrix} M_i^{-1} & 0.5 M_i^{-1} H_i^{(j)\top} \\ * & 1 - \nu_{ij} - h_i \end{bmatrix} \geq 0. \quad (13)$$

PROOF. First, we define $M_i(t) \in \mathbb{S}_{++}^{p_i}$ and $\nu_{ij}(t) > 0$ such that $M_i(t) \leq \nu_{ij}(K_{o_i}T_i(t)\alpha_i(t)Z_i^{-1}\alpha_i(t)T_i^\top(t)K_{o_i}^\top)^{-1}$ for all $j \in \{1, \dots, r_i\}$. Since $T_i(t) \in \mathbb{S}_{++}^{n_i}$, $\alpha_i(t) > 0$ and K_{o_i} is full-rank, this inequality is invertible and hence, $M_i(t)^{-1} - \nu_{ij}(t)^{-1}K_{o_i}T_i(t)\alpha_i(t)Z_i^{-1}\alpha_i(t)T_i^\top(t)K_{o_i}^\top \geq 0$ for all $j \in \{1, \dots, r_i\}$. Applying the Schur Complement yields the first condition in (13). The Minkowski sum of $\mathcal{U}_i(t)$ and $K_{o_i}T_i(t)\mathcal{Z}_{K_i}(t)$ is contained in $\mathcal{V}_i$ if and only if for all $j \in \{1, \dots, r_i\}$, $H_i^{(j)}(u_{e_i} + u_{s_i}) \leq 1$ for all u_{e_i} satisfying $\|u_{e_i}\|_{(K_{o_i}T_i(t)Z_i^{-1}T_i^\top(t)K_{o_i}^\top)^{-1}} \leq \alpha_i^2(t)$ and u_{s_i} satisfying $H_i u_{s_i} \leq h_i(t)$. By defining s_i and t_i such that $u_{e_i} = s_i$ and $u_{s_i} = h_i(t)t_i$, we reach the condition $H_i^{(j)}s_i + H_i^{(j)}h_i(t)t_i \leq 1$ for all $j \in \{1, \dots, r_i\}$, s_i and t_i satisfying $s_i^\top(K_{o_i}T_i(t)\alpha_i(t)Z_i^{-1}\alpha_i(t)T_i^\top(t)K_{o_i}^\top)^{-1}s_i \leq 1$ and $H_i t_i \leq 1$. Using the S-lemma, we reach for all $j \in \{1, \dots, r_i\}$,

$$\begin{bmatrix} \nu_{ij}\psi_i^{-1} & 0 \\ * & \sum_{l=1}^{r_i} 0.5\nu_{ijl}H_i^{(l)\top} - 0.5H_i^{(j)\top}h_i(t) \\ * & 1 - \nu_{ij} - \sum_{l=1}^{r_i} \nu_{ijl} \end{bmatrix} \geq 0, \quad (14)$$

where $\psi_i(t) = K_{o_i}T_i(t)\alpha_i(t)Z_i^{-1}T_i^\top(t)K_{o_i}^\top\alpha_i(t)$. Note that (14) holds if $\sum_{l=1}^{r_i} 0.5\nu_{ijl}(t)H_i^{(l)\top} = 0.5H_i^{(j)\top}h_i(t)$. By stacking this equality for all $j \in \{1, \dots, r_i\}$, we reach $h_i(t)H_i = \nu_i(t)H_i$ where $\nu_{ijl}(t)$ is the entry of $\nu_i(t)$ in the j -th row and l -th column. To satisfy this equality, we choose the diagonal elements of $\nu_i(t)$ to equal $h_i(t)$ and the off-diagonal elements to equal zero. Recalling that $M_i(t) \leq \nu_{ij}(t)(K_i(t)\alpha_i(t)Z_i^{-1}K_i^\top(t)\alpha_i(t))^{-1}$, (14) reduces to a matrix, that if pre and post-multiplied by $\text{diag}(M_i(t)^{-1}, 1)$, yields the second condition in (13). $\square$

One source of conservatism in Proposition 7 is choosing the diagonal elements of $\mu_i(t)$ to equal $g_i(t)$ and the off-diagonal elements to equal zero. The same applies to $\nu_i(t)$ in Proposition 8. Additionally, introducing $M_i(t)$ adds one more degree of conservatism in Proposition 8. In

conclusion, the OCP of the i -th subsystem becomes,

$$\min J_i \text{ s.t. } \begin{cases} \forall k \in \{0, \dots, N-1\}, \\ z_i(k+1|t) - x_i(k|t) \in \mathcal{Z}_{K_i}(t), \\ x_i(k+1|t) = A_i x_i(k|t) + B_i u_i(k|t), \\ x_i(k|t) \in \mathcal{X}_i(t), \quad u_i(k|t) \in \mathcal{U}_i(t), \\ x_i(N|t) = x_{e_i}(t) = A_i x_{e_i}(t) + B_i u_{e_i}(t), \\ x_{e_i} \in \xi \mathcal{X}_i(t), \quad u_{e_i} \in \xi \mathcal{U}_i(t), \\ (7), \lambda_i(t), \lambda_{d_i}(t), \lambda_{ij}(t) \geq 0, \forall j \in \mathcal{N}_i^-, \\ (8), (9), (11), \mu_{ij}(t) \geq 0 \forall j \in \{1, \dots, q_i\}, \\ (13), \nu_{ij}(t) > 0 \forall j \in \{1, \dots, r_i\}, \\ M_i(t) \in \mathbb{S}_{++}^{n_i}, T_i(t)\alpha_i(t) \in \mathbb{S}_{++}^{n_i}, \\ \Gamma_i(t)\alpha_i(t) \in \mathbb{D}_{++}^{n_i}, D_i(t)\alpha_i(t) \in \mathbb{D}_{++}^{n_i}. \end{cases} \quad (15)$$

The decision variables are $x_i(k|t)$ and $u_i(k|t)$ for all $k \in \{0, \dots, N\}$, $x_{e_i}(t)$, $u_{e_i}(t)$, $g_i(t)$, $h_i(t)$, $\alpha_i(t)$, $T_i(t)\alpha_i(t)$, $\Gamma_i(t)\alpha_i(t)$, $D_i(t)\alpha_i(t)$, $M_i(t)^{-1}$, $\lambda_i(t)$, $\lambda_{d_i}(t)$, $\lambda_{ij}(t)$ for all $j \in \mathcal{N}_i^-$, $\mu(t)_{ij}$ for all $j \in \{1, \dots, q_i\}$ and $\nu(t)_{ij}$ for all $j \in \{1, \dots, r_i\}$. Although the OCP might seem nonconvex using these decision variables, one can turn it into an SDP by redefining the decision variables and, if desired, reversing the transformation to recover the original variables from the solution. As the number of neighbours of each subsystem increases, the dimensions of all constraints remain unchanged except for (7) which increases linearly and hence, the computation time is not significantly altered. Similar to the OCP (4), the online OCP (15) is solved at each time instant and the optimal control input $v_i(t) = u_i(0|t) + K_{o_i}T_i(t)(z_i(t) - x_i(0|t))$ is applied to the i -th subsystem.

4 Feasibility and Stability

In this section, we prove that the online OCP (15) is recursively feasible and the resulting closed-loop system is input-to-state stable. In the sequel, we use the superscript $(\cdot)^*$ to refer to the optimal solution. The proof is inspired from Kouvaritakis & Cannon (2016), Limón et al. (2008) and is divided into three main steps. First, we discuss the convergence of the nominal state $x^*(t) = [x_1^*(t), \dots, x_M^*(t)]^\top$ to the artificial equilibrium $x_e^*(t) = [x_{e_1}^*(t), \dots, x_{e_M}^*(t)]^\top$. Then, we discuss the convergence of the artificial equilibrium $x_e^*(t)$ to the target point $x_r = [x_{r_1}, \dots, x_{r_M}]^\top$. Finally, we discuss the stability of the error dynamics between the actual state $z^*(t) = [z_1^*(t), \dots, z_M^*(t)]^\top$ and the nominal state $x^*(t)$. We start with the first step in Lemma 9.

Lemma 9 *The nominal state trajectory converges asymptotically to the artificial equilibrium trajectory.*

The proof uses standard MPC arguments (Kouvaritakis & Cannon (2016)) by comparing the optimal cost at time t , the tail sequence cost at time t and the optimal cost at time $t+1$. Second, we discuss the convergence of $x_e^*(t)$ to the

target point x_r . For this purpose, we define $(\tilde{x}_r, \tilde{u}_r)$ to be the projection of (x_r, u_r) onto the set of equilibrium points contained in the set $(\xi\mathcal{X}_i(t), \xi\mathcal{U}_i(t))$. We also define the matrix $\tilde{P}$ such that $\tilde{P} - (A + BK(t))^\top \tilde{P} (A + BK(t)) - Q - K^\top(t)RK(t) \geq 0$. Note that such $\tilde{P}$ exists since $K(t)$ asymptotically stabilizes the actual global dynamics as shown in Proposition 6. Finally, we define $E(c, s) = \{x : \|x - c\|_{\tilde{P}}^2 \leq s\}$. For brevity, we omit the dependence on t in the proof of Lemma 10. This proof is inspired from Limón et al. (2008), however, we prove that $\lim_{t \rightarrow \infty} (x^*(t) - x_e^*(t)) = 0$ implies $\lim_{t \rightarrow \infty} (x_e^*(t) - x_r) = 0$ instead of $x^*(t) - x_e^*(t) = 0$ implying $x_e^*(t) - x_r = 0$.

Lemma 10 *If $x^*(t)$ and $x_e^*(t)$ are such that $\lim_{t \rightarrow \infty} (x^*(t) - x_e^*(t)) = 0$, then $\lim_{t \rightarrow \infty} (x_e^*(t) - \tilde{x}_r) = 0$.*

PROOF. According to the OCP (15), $(x_e^*, u_e^*) \in \text{int}(\mathcal{X} \times \mathcal{U})$. Hence, there always exists $E(x_e^*, \beta_1)$ such that $(x, K(x - x_e^*) + u_e^*) \in \gamma\mathcal{X} \times \gamma\mathcal{U}$ for all $x \in E(x_e^*, \beta_1)$ where $\gamma \in (0, 1)$. Note that $E(x_e^*, \beta_1)$ is invariant since it is defined using the level sets of a Lyapunov function. We now define an equilibrium point $(\tilde{x}_e, \tilde{u}_e) \in \text{int}(\mathcal{X} \times \mathcal{U})$ such that $\|\tilde{x}_e - \tilde{x}_r\|_S^2 = \beta_2\|x_e^* - \tilde{x}_r\|_S^2$ and $\|\tilde{x}_e - x_e^*\|_{\tilde{P}}^2 = (1 - \lambda)\|\tilde{x}_r - x_e^*\|_{\tilde{P}}^2$ where $\beta_2 \in (0, 1)$ and $\lambda \in (0, 1)$ is chosen sufficiently large such that $(0, K(\tilde{x}_e - x_e^*) + \tilde{u}_e - u_e^*) \in (1 - \gamma)\mathcal{X} \times (1 - \gamma)\mathcal{U}$. Next, we define $E(\tilde{x}_e, \beta_3) \subset E(x_e^*, \beta_1)$ such that $x_e^* \in \text{int}(E(\tilde{x}_e, \beta_3))$. This set is invariant since it is defined using the level sets of a Lyapunov function. Moreover, for all $x \in E(\tilde{x}_e, \beta_3)$, $(x, K(x - \tilde{x}_e) + \tilde{u}_e) = (x, K(x - x_e^*) + u_e^*) + (0, K(\tilde{x}_e - x_e^*) + \tilde{u}_e - u_e^*) \in (\gamma\mathcal{X} \times \gamma\mathcal{U}) \oplus ((1 - \gamma)\mathcal{X} \times (1 - \gamma)\mathcal{U}) = \mathcal{X} \times \mathcal{U}$.

We now define $\beta_4 > 0$ such that $S > \beta_4\tilde{P}$ and the set $\hat{\mathcal{X}} = \{x : \|x - x_e^*\|_{\tilde{P}}^2 \leq \beta_5\|x_e^* - \tilde{x}_r\|_{\tilde{P}}^2 \text{ \& } (x_e^* - \tilde{x}_e)^\top \tilde{P} |x - x_e^*| \leq \beta_5\|x_e^* - \tilde{x}_r\|_{\tilde{P}}^2\}$ where $0 < \beta_5 < \beta_4(1 - \beta_2)/3$. Hence, $\|x - \tilde{x}_e\|_{\tilde{P}}^2 = \|x - x_e^*\|_{\tilde{P}}^2 + 2(x_e^* - \tilde{x}_e)^\top \tilde{P}(x - x_e^*) + (1 - \lambda)\|x_e^* - \tilde{x}_r\|_{\tilde{P}}^2 \leq \|x_e^* - \tilde{x}_r\|_{3\beta_5\tilde{P} + (1 - \lambda)\tilde{P}}^2$ for all $x \in E(x_e^*, \beta_6) \subset \hat{\mathcal{X}}$. Note that $E(x_e^*, \beta_6)$ can be also chosen to lie inside $E(\tilde{x}_e, \beta_3) \subset E(x_e^*, \beta_1)$ and hence is invariant. Note also that $\|x - \tilde{x}_e\|_{\tilde{P}}^2 + \|\tilde{x}_e - \tilde{x}_r\|_S^2 \leq \|x_e^* - \tilde{x}_r\|_S^2$ for all $x \in E(x_e^*, \beta_6)$ if $(1 - \beta_2)S - (1 - \lambda + 3\beta_5)\tilde{P} \geq 0$ or equivalently, $\beta_4(1 - \beta_2) - (1 - \lambda + 3\beta_5) \geq 0$. This can be achieved if $\lambda \geq 1 - (\beta_4(1 - \beta_2) - 3\beta_5)$. Since $\beta_5 < \beta_4(1 - \beta_2)/3$, this condition on λ does not contradict the aforementioned conditions on λ .

We conclude the proof using contradiction. Assume that $\lim_{t \rightarrow \infty} (x^*(t) - x_e^*(t)) = 0$ and $\lim_{t \rightarrow \infty} (x_e^*(t) - \tilde{x}_r) \neq 0$. Hence, it is always possible to pick an arbitrarily large t such that $x^* \in E(x_e^*, \beta_6) \subset E(\tilde{x}_e, \beta_3)$. In addition, there exists β_7 such that $\|x_e^* - \tilde{x}_r\| \geq \beta_7$ for infinitely many t . Since $E(x_e^*, \beta_6)$ and $E(\tilde{x}_e, \beta_3)$ are both invariant and satisfy the state and input constraints, let J^* and $\bar{J}$ be

the cost corresponding to the optimal solution and the feasible solution aiming to converge to $\tilde{x}_e$ and note that $J^* < \bar{J}$. Note also that $\bar{J} \leq \|x^* - \tilde{x}_e\|_{\tilde{P}}^2 + \|\tilde{x}_e - \tilde{x}_r\|_S^2$ due to the Lyapunov inequality that $\tilde{P}$ satisfies. Since $x^* \in E(x_e^*, \beta_6)$, then $\|x^* - \tilde{x}_e\|_{\tilde{P}}^2 + \|\tilde{x}_e - \tilde{x}_r\|_S^2 \leq \|x_e^* - \tilde{x}_r\|_S^2$. But, $\|x_e^* - \tilde{x}_r\|_S^2 \leq \|x_e^* - x_r\|_S^2 \leq J^*$ which contradicts the optimality of J^* . $\square$

Finally, we discuss the input-to-state stability of the global closed-loop system under the MPC scheme (15). Recall from (5) that the global system is given by $z(t + 1) = (A_d + E\tilde{A}_gC)z(t) + Bv(t) + d(t)$. Recall also that the optimal input is given by $v^*(t) = u^*(t) + K^*(t)(z^*(t) - x^*(t))$ where $K^*(t) = \text{diag}(K_{o1}T_1^*(t), \dots, K_{oM}T_M^*(t))$ and $u^*(t) = u^*(0|t) = [u_1^{*\top}(0|t), \dots, u_M^{*\top}(0|t)]^\top$. Hence, the closed-loop system is given by $z^*(t + 1) = A_dz^*(t) + Bu^*(t) + BK^*(t)(z^*(t) - x^*(t)) + E\tilde{A}_gCz^*(t) + d(t)$. By defining $e^*(t) = z^*(t) - x^*(t)$ and $\eta(t) = d(t) - x^*(t + 1) + Ax^*(t) + Bu^*(t)$, the closed-loop system becomes,

$$e^*(t + 1) = (A + BK^*(t))e^*(t) + \eta(t). \quad (16)$$

While $e^*(t)$ is the state of this system, $x^*(t)$ and $u^*(t)$, together with $d(t)$, are exogenous signals to this system. While $d(t)$ is bounded by assumption, $x^*(t)$ and $u^*(t)$ are bounded since they are the outputs of the proposed MPC problem (15) which consider compact constraint sets. Hence, we prove input-to-state stability with respect to these bounded exogenous signals. Note that $x^*(t + 1)$ is not necessarily equal to $Ax^*(t) + Bu^*(t)$ where $A = \text{diag}(A_1, \dots, A_M) + E\tilde{A}_gC$. This is because $x^*(\cdot)$ and $u^*(\cdot)$ are the nominal states and inputs which satisfy $x^*(t + 1) = \text{diag}(A_1, \dots, A_M)x^*(t) + Bu^*(t)$. Moreover, the considered system is switched since $K^*(t)$ is adapted at each time instant. Hence, the closed-loop system under the proposed MPC scheme can be considered as an uncertain switched system with bounded disturbances.

Theorem 11 *The MPC scheme (15) is recursively feasible and the global closed-loop system (16) under this scheme is input-to-state stable with respect to $\eta(t)$.*

PROOF. Recursive feasibility is proven using standard MPC arguments (Kouvaritakis & Cannon (2016)) by making use of the tail sequence. Regarding stability, we first consider the disturbance-free system $e^*(t + 1) = (A + BK^*(t))e^*(t)$. The matrix $A + BK^*(t)$ is always Schur for any $K^*(t)$ since the constraints derived in Propositions 5 and 6 ensure that $P - (A + BK^*(t))^\top P(A + BK^*(t)) \geq \Gamma(t) - C^\top \tilde{A}_gC - C^\top LD(t)LC \geq \epsilon I_n$ at the t -th time instant for all t . Since this condition is always satisfied using the same matrix P , then $V(e^*) = \|e^*\|_P^2$ is a Lyapunov function for the disturbance-free system independent of $K^*(t)$. Now, we consider the closed-loop

system (16). Using the Lyapunov function $V(e^*)$ and defining $\eta(t) = d(t) - x^*(t+1) + Ax^*(t) + Bu^*(t)$, we can deduce that $V((A + BK^*(t))e^*(t) + \eta(t)) - V(e^*) \leq -(\epsilon - \delta)\|e^*(t)\|^2 - \delta\|e^*(t)\|^2 + 2\|\eta(t)\|_P\|(A + BK^*(t))e^*(t)\|_P + \|\eta(t)\|_P^2$ where $0 < \delta < \epsilon$. To prove input-to-state stability, it suffices to ensure that $\delta\|e^*(t)\|_2^2 \geq 2\|\eta(t)\|_P\|(A + BK^*(t))e^*(t)\|_P + \|\eta(t)\|_P^2$ for some $\zeta > 0$ such that $\|e^*(t)\| \geq \zeta$. We denote the maximum eigenvalue of P by $\lambda_{\max}(P)$ and the maximum norm of $\eta(t)$ by $\eta_{\max}$. Since $\|(A + BK^*(t))e^*(t)\|_P < \|e^*(t)\|_P$ and $\|\cdot\|_P \leq \sqrt{\lambda_{\max}(P)}\|\cdot\|$, it suffices instead to ensure that $\delta\|e^*(t)\|^2 - 2\lambda_{\max}(P)\eta_{\max}\|e^*(t)\| - \lambda_{\max}(P)\eta_{\max}^2 \geq 0$. By computing the roots of the convex quadratic function on the left-hand side of the inequality, it is easy to verify that this inequality holds for $\|e^*(t)\| \geq \zeta = \left(\lambda_{\max}(P) + \sqrt{\lambda_{\max}^2(P) + \delta\lambda_{\max}(P)}\right)\eta_{\max}/\delta$. Hence, input-to-state stability is guaranteed. $\square$

5 Numerical Results

In this section, we investigate the efficacy of the proposed scheme by comparing it to existing schemes in the literature. In particular, we compare the proposed scheme (ADP+LRN) to robust MPC (ROB) (Mayne et al. (2005)), adaptive MPC (ADP) (Parsi et al. (2021)) and learning-based MPC (LRN) (Aswani et al. (2013)). Note that ADP is an extension of the adaptive MPC scheme in Lorenzen et al. (2019) to the class of interconnected systems. We solve all optimization problems using MATLAB with YALMIP (Lofberg (2004)), Mosek and Gurobi on a computer equipped with a 1.9-GHz Intel core i7-8550U processor.

In the simulation example, we use a network of five double integrators connected in series as in Parsi et al. (2021). We consider each double integrator as a sub-

system with the dynamics in (1) where $A_i = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$,

$B_i = [0.5 \ 1]^\top$, $C_i = [0 \ 1]$ and $E_i = [0.05 \ 0.1]^\top$ for all $i \in \{1, 2, 3, 4, 5\}$. All states and inputs are constrained between -2 and 2 . The cost function matrices are $Q_i = I_2$, $R_i = 0.1$ and $S_i = 10I_2$ for all subsystems and the prediction horizon is 5. The noise $d(t) \in \mathbb{R}^2$ is extracted uniformly with the first entry lying between -0.025 and 0.025 and the second between -0.05 and 0.05 . The unknown true value of the parameters a_{12} and a_{54} is 1, whereas that of a_{21} , a_{23} , a_{32} , a_{34} , a_{43} and a_{45} is 0.5. The parameters a_{12} and a_{54} are assumed to lie initially between 0 and 4, whereas all other parameters are assumed to lie initially between 0 and 2. The unknown true values of the parameters are not used within the proposed framework and, by design, communication is not used in the adaptation phase to locally solve the OCPs. Each subsystem is required to follow a randomly generated

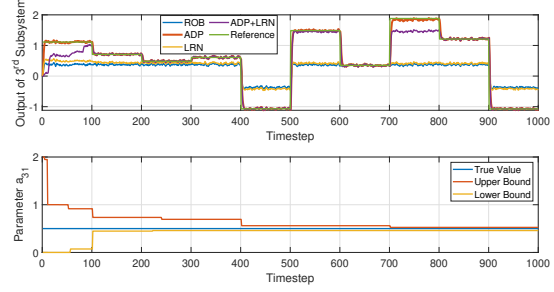


Fig. 1. Top: Output trajectory of the third subsystem. Bottom: Bound estimates of the parameter a_{31} .

piecewise constant reference changing every 100 seconds and constrained between -2 and 2 . The simulation time is 1000 seconds and the initial condition is $[0 \ 0]^\top$ for all subsystems.

Fig.1(Top) shows the closed-loop performance of the third subsystem for each scheme. As shown in this figure, ROB and LRN generally fail to track the reference trajectory since LRN updates the model only while keeping the initial tube and tightened constraint sets unchanged and ROB updates nothing. On the contrary, ADP follows the reference successfully since it updates the tube and the tightened constraint sets online. Similarly, ADP+LRN generally follows the reference (due to the online shrinkage of the uncertain parameter sets as shown in Fig.1(Bottom) for the parameter a_{31}) except for extreme references close to the actual state constraint bounds due to the conservative constraint tightening.

The resulting behavior of each scheme has a direct effect on the closed-loop cost shown in Fig.2(Top). ROB has the highest cost, followed by LRN whose cost is slightly lower due to updating the model. The costs of ADP+LRN and ADP are considerably lower with ADP having the lowest cost. This is because ADP can quickly track the reference, even in extreme cases, since this scheme can directly handle parametric uncertainties unlike ADP+LRN which lumps all uncertainties in an additive term. ADP+LRN, however, results in a decentralized OCP which requires significantly reduced computations and communication among neighbours. On the other hand, ADP yields a distributed OCP that is solved using distributed optimization techniques (e.g. ADMM) leading to extensive communication and computations.

Fig.2(Bottom) shows the mean and standard deviation of the computation time required by each scheme (including different versions of ADP which differ in the maximum ADMM iterations allowed) and the corresponding closed-loop cost. We use the ADP scheme whose computation time is around 1 second in the other figures. It is found, however, that if we run ADMM for around 0.1 second only, we still get a relatively satisfactory performance. Running ADMM for less time (e.g. same order of magnitude of the time required by

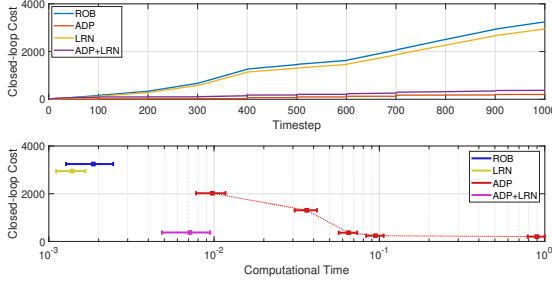


Fig. 2. Top: closed-loop cost, Bottom: computation time mean (square) and standard deviation (bar).

ADP+LRN) result in severe closed-loop performance deterioration. On the other hand, all other schemes have lower computation time due to their decentralized nature which do not require using distributed optimization techniques such as ADMM. Since the OCP of ADP+LRN is a semidefinite program, ADP+LRN requires slightly higher computation compared to LRN and ROB whose OCPs are quadratic programs.

6 Conclusion

We propose an adaptive learning-based MPC scheme for uncertain interconnected systems. The developed scheme is found to yield a better trade-off between computation complexity and closed-loop performance compared to existing schemes. Future work include applying the developed scheme to real-world applications.

Acknowledgements

The authors are grateful to Prof. Roy Smith, Prof. Andrea Iannelli, Dr. Goran Banjac and Dr. Anil Parsi for the fruitful discussions on the topic.

References

Aboudonia, A., Martinelli, A. & Lygeros, J. (2021), ‘Passivity-based decentralized control for discrete-time large-scale systems’, *IEEE Control Systems Letters* **5**(6), 2072–2077.

Aliyu, M. (2017), *Nonlinear H-infinity control, Hamiltonian systems and Hamilton-Jacobi equations*, CRC Press.

Aswani, A., Gonzalez, H., Sastry, S. S. & Tomlin, C. (2013), ‘Provably safe and robust learning-based model predictive control’, *Automatica* **49**(5), 1216–1226.

Betti, G., Farina, M. & Scattolini, R. (2013), Distributed mpc: A noncooperative approach based on robustness concepts, in ‘Distributed model predictive control made easy’, Springer, pp. 421–435.

Boyd, S., El Ghaoui, L., Feron, E. & Balakrishnan, V. (1994), *Linear matrix inequalities in system and control theory*, SIAM.

Conte, C., Jones, C. N., Morari, M. & Zeilinger, M. N. (2016), ‘Distributed synthesis and stability of cooperative

distributed model predictive control for linear systems’, *Automatica* **69**, 117–125.

Hewing, L., Wabersich, K. P., Menner, M. & Zeilinger, M. N. (2020), ‘Learning-based model predictive control: Toward safe learning in control’, *Annual Review of Control, Robotics, and Autonomous Systems* **3**, 269–296.

Kouvaritakis, B. & Cannon, M. (2016), ‘Model predictive control’, *Switzerland: Springer International Publishing* **38**.

Limón, D., Alvarado, I., Alamo, T. & Camacho, E. F. (2008), ‘MPC for tracking piecewise constant references for constrained linear systems’, *Automatica* **44**(9), 2382–2387.

Lofberg, J. (2004), YALMIP: A toolbox for modeling and optimization in MATLAB, in ‘2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)’, Taipei, Taiwan, pp. 284–289.

Lorenzen, M., Cannon, M. & Allgöwer, F. (2019), ‘Robust MPC with recursive model update’, *Automatica* **103**, 461–471.

Mayne, D. Q., Seron, M. M. & Raković, S. (2005), ‘Robust model predictive control of constrained linear systems with bounded disturbances’, *Automatica* **41**(2), 219–224.

Mesbah, A. (2016), ‘Stochastic model predictive control: An overview and perspectives for future research’, *IEEE Control Systems Magazine* **36**(6), 30–44.

Müller, M. A. & Allgöwer, F. (2017), ‘Economic and distributed model predictive control: Recent developments in optimization-based control’, *SICE Journal of Control, Measurement, and System Integration* **10**(2), 39–52.

Parsi, A., Aboudonia, A., Iannelli, A., Lygeros, J. & Smith, R. S. (2021), A distributed framework for linear adaptive MPC, in ‘2021 60th IEEE Conference on Decision and Control (CDC)’, Austin, TX, USA, pp. 460–465.

Parsi, A., Iannelli, A. & Smith, R. S. (2022), ‘Scalable tube model predictive control of uncertain linear systems using ellipsoidal sets’, *International Journal of Robust and Nonlinear Control*.

Raković, S. V., Kouvaritakis, B., Findeisen, R. & Cannon, M. (2012), ‘Homothetic tube model predictive control’, *Automatica* **48**(8), 1631–1638.

Ren, Y. M., Alhajeri, M. S., Luo, J., Chen, S., Abdullah, F., Wu, Z. & Christofides, P. D. (2022), ‘A tutorial review of neural network modeling approaches for model predictive control’, *Computers & Chemical Engineering* p. 107956.

Riverso, S., Farina, M. & Ferrari-Trecate, G. (2013), ‘Plug-and-play decentralized model predictive control for linear systems’, *IEEE Transactions on Automatic Control* **58**(10), 2608–2614.

Saltık, M. B., Özkan, L., Ludlage, J. H., Weiland, S. & Van den Hof, P. M. (2018), ‘An outlook on robust model predictive control algorithms: Reflections on performance and computational aspects’, *Journal of Process Control* **61**, 77–102.