A Hierarchical MPC Scheme for Ensembles of Hammerstein Systems

Felix Petzke * **Marcello Farina** ** **Stefan Streif** *

* *Technische Universität Chemnitz, Automatic Control & System Dynamics Lab, Chemnitz, Germany (e-mail:* {felix.petzke, stefan.streif}@etit.tu-chemnitz.de*).*
** *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy (e-mail:* marcello.farina@polimi.it*).*

Abstract: Hierarchical control approaches have been one of the elective methods for the optimal control of large-scale systems in the last decades. In (Petzke et al., 2018) we presented a multirate hierarchical MPC scheme for linear systems, with remarkable flexibility and scalability properties. In this paper we extend the former approach to ensembles of Hammerstein systems and we complement the method by proposing a suitable high-level optimizer. The theoretical properties are discussed in the light of the theoretical properties of the former method. Lastly, an example case study is presented to show the effectiveness of the proposed method.

Keywords: Hierarchical control, model predictive control, Hammerstein systems.

1. INTRODUCTION

Recently, thanks to the widespread of distributed small-scale production units in many different realms, there has been a diffusion of *virtual plants*, i.e., complex plants where several similar systems operate in parallel to jointly produce a common product. This trend is presently impacting on different applications areas e.g. electrical generation systems, micro-grids, HVAC systems, steam generation, and water distribution. In all these applications, dedicated advanced control systems are required to coordinate an ever increasing number of subsystems that act towards a main goal, but also commonly operate in a limited range and with limited shared resources. Many other features must also be verified, e.g., adaptivity to changing environments, scalability, and flexibility.

Hierarchical approaches are an often necessary compromise between centralized methods, guaranteeing nominal system-wide optimality but impractical for large-scale systems, and decentralized/distributed approaches, that often trade optimality for scalability, flexibility and robustness. Furthermore, they can guarantee optimal performances in system supervision and in coordination of subunits.

Different hierarchical control schemes have been proposed in the recent years, with particular focus on constrained systems, i.e., based on model predictive control (MPC) and reference governor methods (Scattolini, 2009; Barcelli et al., 2010; Farina et al., 2018; Petzke et al., 2018; Garone et al., 2017; Kalabić et al., 2012).

All these approaches, however, have been designed for linear systems, which may limit the applicability range of the methods proposed therein. While the generalization of those methods to general-type nonlinear systems may be impractical, in view of the inherent computational complexity that this may produce, the use of block-oriented models, composed of linear dynamic parts and nonlinear static functions can be advantageous. In fact, besides being more prone to control design, they often provide good approximation accuracy, they are easy to develop and identify, and they may allow to easily incorporate a priori process knowledge.

Hammerstein models, for example, share all of these advantages and have been shown to be particularly suitable in many application areas, e.g., micro-turbines (Jurado, 2006), wind turbines (van der Veen et al., 2013), and synchronous generators (Sadabadi et al., 2007). MPC control algorithms have been also devised, during the last decades, to specifically address Hammerstein-type systems, e.g., (Fruzzetti et al., 1997; Patwardhan et al., 1998; Bloemen et al., 2001; Jurado, 2006; Chan and Bao, 2007; Harnischmacher and Marquardt, 2007; Lawrynczuk, 2015).

In this paper we extend the work in (Petzke et al., 2018) in two directions: First, we consider Hammerstein-type systems and, secondly, we complement the formerly presented scheme with a high-level resource sharing optimizer, inspired by (Farina et al., 2018).

In Sec. 2 the problem is formulated, while Sec. 3 illustrates the proposed three-layered hierarchical optimization control architecture. In Sec. 4 the theoretical properties are discussed, while a simulation example is provided in Sec. 5. Finally, some conclusions are drawn in Sec. 6.

Notation: Throughout the paper, the identity matrix of dimension $N \times N$ is denoted as I^N, and the zero matrix of dimension $N \times M$ as $O^{N\times M}$. Double-lined letters denote sets (e.g. $\mathbb{U}$), with $\oplus$ being the Minkowski sum of two sets and $\bigoplus_{i=1}^{N} \mathbb{W}_i = \mathbb{W}_1 \oplus \cdots \oplus \mathbb{W}_N$. We use two different timescales, where κ indicates steps on a fast timescale with step size τ, and k indicates steps on a slow timescale with step size T, respectively, with $\tau \ll T$ and $\frac{T}{\tau} = N_\tau \in \mathbb{N}$.

2. PROBLEM SETUP

The general focus of this paper is the optimal control of an ensemble of Hammerstein-type systems, i.e. a set of subsystems that have *similar* dynamics and act towards a joint goal (Morton, 2007). In this context we interpret *similarity* of the systems as variations in the model parameters and possibly dimensions.

We assume an ensemble of N Hammerstein-type subsystems modeled in discrete-time input-output form as

$$y_i(\kappa) = \sum_{\tau=1}^{P_i} A_{\tau,i} y_i(\kappa-\tau) + \sum_{\tau=1}^{M_i} B_{\tau,i} z_i(\kappa-\tau) \quad (1)$$

with $y_i \in \mathbb{R}^p$, $A_{\tau,i} \in \mathbb{R}^{p\times p}$, $B_{\tau,i} \in \mathbb{R}^{p\times m}$,

$$z_i(\kappa) = g_i(u_i(\kappa)), \quad (2)$$

where $g_i : \mathbb{R}^m \to \mathbb{R}^m$ is the static nonlinear input map, and, consistently, $z_i, u_i \in \mathbb{R}^m$. Note that while all outputs y_i of different subsystems are assumed to have the same dimension p, they might depend on different delay orders P_i and M_i.

We can derive the state-space model corresponding with (1) by defining $x_i(\kappa) = \left[y_i(\kappa), \ldots, y_i(\kappa-P_i+1), z_i(\kappa-1), \ldots, z_i(\kappa-M_i+1)\right]^\top \in \mathbb{R}^{n_i}$, $n_i = pP_i + m(M_i-1)$. We obtain

$$\mathcal{S}_i : \begin{cases} x_i(\kappa+1) = \mathcal{A}_i x_i(\kappa) + \mathcal{B}_i z_i(\kappa), \\ z_i(\kappa) = g_i(u_i(\kappa)), \\ y_i(\kappa) = \mathcal{C}_i x_i(\kappa), \end{cases} \quad (3)$$

where

$$\mathcal{A}_i = \left[\Theta_i^\top, Y_i^\top, Z_i^\top\right]^\top \in \mathbb{R}^{n_i\times n_i}, \quad (4)$$

$$\Theta_i = \left[A_{1,i}, \ldots, A_{P_i,i}, B_{2,i}, \ldots, B_{M_i,i}\right] \in \mathbb{R}^{p\times n_i}, \quad (5)$$

$$Y_i = \left[I^{p(P_i-1)}, O^{p(P_i-1)\times p+m(M_i-1)}\right] \in \mathbb{R}^{p(P_i-1)\times n_i}, \quad (6)$$

$$Z_i = \left[O^{m(M_i-1)\times pP_i}, I_{-m}^{m(M_i-1)}\right] \in \mathbb{R}^{m(M_i-1)\times n_i}, \quad (7)$$

where I_d^N denotes an $N\times N$-matrix with ones on the d-th subdiagonal, and

$$\mathcal{B}_i = \left[B_{1,i}^\top, O^{m\times p(P_i-1)}, I^m, O^{m\times m(M_i-2)}\right]^\top \in \mathbb{R}^{n_i\times m}, \quad (8)$$

$$\mathcal{C}_i = \left[I^p, O^{p\times p(P_i-1)+m(M_i-1)}\right] \in \mathbb{R}^{p\times n_i}. \quad (9)$$

We assume all systems to be asymptotically stable and controllable. Furthermore, we make the following assumptions on the nonlinear input mappings g_i:

Assumption 1. For all $u, v \in \mathbb{U}_i$ the functions $g_i(\cdot)$ are

(i) element-wise bounded, i.e. $\|g_i(u)\| \leq \bar{g}_i$, $\bar{g}_i > 0$,
(ii) Lipschitz continuous, i.e. $\|g_i(u) - g_i(v)\| \leq \gamma \|u - v\|$, with $\gamma > 0$,
(iii) element-wise strictly monotone, i.e. $u < v \Leftrightarrow g_i(u) < g_i(v)$, with $g_i(0) = 0$.

The subsystems are subject to the input constraints

$$u_i \in \mathbb{U}_i \subset \mathbb{R}^m, \quad (10)$$

$$\sum_{i=1}^{N} u_i \in \bar{\mathbb{U}}, \quad (11)$$

as well as an aggregated output constraint

$$\bar{y} = \sum_i y_i \in \bar{\mathbb{Y}}, \quad (12)$$

where we assume the sets $\mathbb{U}_i$, $\bar{\mathbb{U}}$, and $\bar{\mathbb{Y}}$ to be convex and compact. The overall goal is to drive the ensemble output $\bar{y}$ towards a total requested, constant output reference y_{ref}. In this work we employ a three-level hierarchy where

- at the *high level* a static optimization is employed to calculate optimal resource sharing (ORS) factors α_i for all subsystems given a constant reference signal y_{ref} (cf. Sec. 3.1),
- at the *medium level* an aggregated dynamic model of the ensemble is used to set up a model predictive controller (MPC) that computes an average control signal $\bar{z}$, which is broadcasted and applied to the low-level subsystems (cf. Sec. 3.2),
- at the *low level* a local shrinking horizon controller is used for each subsystem to optimize the individual performance and to ensure that the deviations between outputs of the subsystems and the medium level are within user-defined bounds (cf. Sec. 3.3).

Every level therefore gives a distinct contribution to the overall control signal u_i of each subsystem (cf. Fig. 1): the high level ORS factors α_i distribute the medium level average control signal $\bar{z}$ amongst all subsystems controlling the long-term behavior, while the signal v_i from the low level controls the short-term behavior. Consequently, the input signals u_i, $i = 1, \ldots, N$ are

$$u_i = \bar{u}_i + v_i, \; \bar{u}_i = g_i^{-1}(\alpha_i \bar{z}). \quad (13)$$

Note that the invertibility of the map g_i follows from Assumptions 1.(ii) and 1.(iii). A sketch of the resulting overall proposed control structure for the ensemble is shown in Fig. 2.

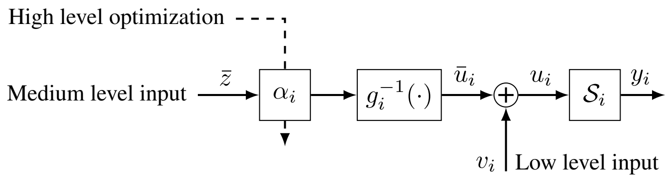

Fig. 1. Block diagram of a single subsystem of the ensemble with input contributions of the different layers.

3. HIERARCHICAL CONTROL STRUCTURE

The control structure employed in this paper aims to achieve three main goals, each one associated with one layer of the hierarchy: optimal resource sharing, reference tracking, and error minimization, where each layer operates on a different timescale.
The high level optimizes resource sharing between all systems of the ensemble in steady state. Since we assume the output reference to be constant, this can be done offline. The medium level tracking MPC runs on a slow timescale with steps of size T indicated by k. It aims to drive the ensemble output $\bar{y}$ towards the constant output reference y_{ref} in an optimal way. At the low level, the individual subsystems' performances are optimized and the error between the output predictions of the medium level and the actual output of the ensemble is minimized. This is achieved via a shrinking horizon MPC that operates on a fast timescale with steps of length τ indicated by κ. The according time scaling factor is defined as $N_\tau = \frac{T}{\tau}$, i.e. one (slow) step on the medium level corresponds to N_τ (fast) steps on the low level. This section provides details on how the optimization problems at each layer are constructed.

3.1 High Level: Optimal Resource Sharing

As stated in equation (13), the input signal $\bar{z}$ of the medium level is distributed between the low-level systems via the resource-sharing factors α_i, such that

$$\alpha_i \in [0,1]\ \forall i,\ \sum_{i=1}^{N} \alpha_i = 1. \tag{14}$$

To obtain their optimal values we adapt the scheme introduced in (Farina et al., 2018) for the presented setup. In steady-state, the high-level input $\bar{z} = z^{\mathrm{s}}$ has to fulfill

$$\hat{\mathcal{C}}(I^{n_{\mathrm{ref}}} - \hat{\mathcal{A}})^{-1} \sum_{i=1}^{N} \hat{\mathcal{B}}_i \alpha_i z^{\mathrm{s}} = r^s, \tag{15}$$

where $\hat{\mathcal{C}}(I^{n_{\mathrm{ref}}} - \hat{\mathcal{A}})^{-1}\hat{\mathcal{B}}_i \alpha_i z^{\mathrm{s}}$ is the contribution of subsystem i to the ensemble output and $r^s \in \bar{\mathbb{Y}}$ is the closest feasible reference to the given reference y_{ref} the ensemble can reach under constraint (10). When in steady state, $v_i = 0$ for all subsystems: in this case each subsystem applies input $\bar{u}_i = g_i^{-1}(\alpha_i \bar{z})$, and therefore the actual amount of resources consumed by the ensemble is given by $\sum_{i=1}^{N} \bar{u}_i$. In order to minimize this overall consumed resource and, at the same time, find an r^s as close as possible to y_{ref}, we define the nonlinear optimization problem n

$$\begin{aligned} \min_{r^s, z^{\mathrm{s}}, \boldsymbol{\alpha}}\ & \|y_{\mathrm{ref}} - r^s\|_T^2 + \sum_{i=1}^{N} q_i \|g_i^{-1}(\alpha_i z^s)\| \\ \text{s.t. } & (14),\ (15), \\ & r^s \in \bar{\mathbb{Y}} \\ & g_i^{-1}(\alpha_i z^s) \in \mathbb{U}_i \qquad (\to \text{cf. } (10)), \\ & \sum_{i=1}^{N} g_i^{-1}(\alpha_i z^s) \in \bar{\mathbb{U}} \qquad (\to \text{cf. } (11)), \end{aligned} \tag{16}$$

with $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_N\}$ and where the term q_i denotes a suitable cost associated with the usage of subsystem i, which yields the resource-optimal sharing factors $\alpha_i^\star$ for a given constant output reference y_{ref}.

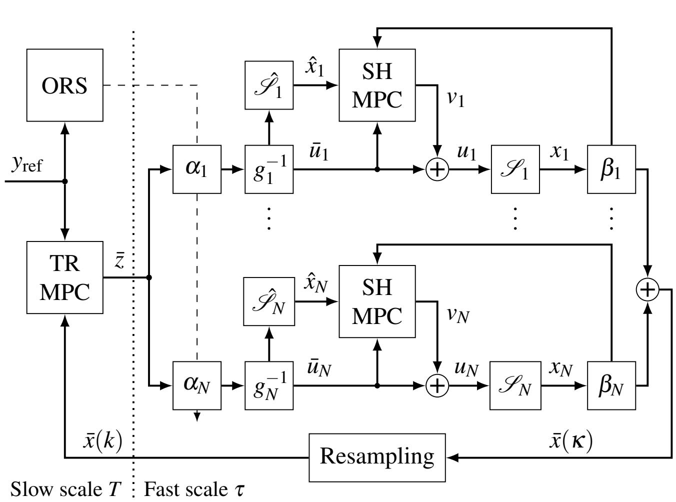

Fig. 2. Proposed hierarchical control structure. The high level optimizes weights α_i to achieve optimal resource sharing. The medium-level tracking (TR) MPC and the low-level shrinking-horizon (SH) MPCs run on different timescales.

3.2 Medium Level: Slow Tracking MPC

This section is concerned with the derivation of the medium-level MPC and the prediction model used therein. To this end, we need to derive an aggregated model of the entire ensemble. To do so, we first need to define the so-called *reference models* for the individual systems.

Individual Reference Models For each subsystem $\mathcal{S}_i$ we devise an associated reference model $\hat{\mathcal{S}}_i$. A key aspect in the derivation of these reference models is that they all have the same state dimension n_{ref}, being $n_{\mathrm{ref}} \le n_i$ for all i. The state of the i-th reference model is defined as $\hat{x}_i(\kappa) = \beta_i x_i(\kappa)$, where β_i, for each $i = 1, \ldots, N$, is a suitable map of rank n_{ref}. All reference models have the system matrix $\hat{\mathcal{A}}$ and output matrix $\hat{\mathcal{C}}$. Indeed, the i-th reference model is defined as

$$\hat{\mathcal{S}}_i : \begin{cases} \hat{x}_i(\kappa+1) = \hat{\mathcal{A}}\hat{x}_i(\kappa) + \hat{\mathcal{B}}_i \hat{z}_i(\kappa) + \hat{w}_i(\kappa), \\ \hat{z}_i(\kappa) = g_i(\bar{u}_i(\kappa)) \\ y_i(\kappa) = \hat{\mathcal{C}}\hat{x}_i(\kappa). \end{cases} \tag{17}$$

The term $\hat{w}_i(\kappa)$ (referred to as the *reference deviation*) has been introduced to account for the mismatch between systems $\mathcal{S}_i$ and their respective references $\hat{\mathcal{S}}_i$. As specified in (Petzke et al., 2018), two steady-state consistency conditions must be fulfilled by $\hat{\mathcal{S}}_i$ with respect to $\mathcal{S}_i$, and more specifically we must verify that $\beta_i(I^{n_i} - \mathcal{A}_i)^{-1}\mathcal{B}_i = (I^{n_{\mathrm{ref}}} - \hat{\mathcal{A}})^{-1}\hat{\mathcal{B}}_i$ and $\hat{\mathcal{C}}\beta_i = \mathcal{C}_i$. Note that these conditions do not account for the gains of the nonlinear input couplings $g_i(u_i)$ but only for the gain mismatch in the (linear) system dynamics. It is straightforward to verify them provided that

- β_i is a selection matrix that allows to define $\hat{x}_i$ as a collection of (possibly selected) past lags of y_i and $\hat{z}_i$, i.e., $\hat{x}_i(\kappa) = \left[y_i(\kappa), \ldots, y_i(\kappa - \hat{P} + 1), \hat{z}_i(\kappa - 1), \ldots, \hat{z}_i(\kappa - \hat{M} + 1)\right]^\top$ where $\hat{P} \le P_i$ and $\hat{M} \le M_i$,
- matrices $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}_i$, and $\hat{\mathcal{C}}$ have the same general structure as $\mathcal{A}_i$, $\mathcal{B}_i$, and $\mathcal{C}_i$, i.e.,

$$\begin{aligned} \hat{\mathcal{A}} &= \left[\hat{\Theta}^\top, \hat{Y}^\top, \hat{Z}^\top\right]^\top \in \mathbb{R}^{n_{\mathrm{ref}} \times n_{\mathrm{ref}}}, \\ \hat{\Theta} &= \left[\hat{A}_1, \ldots, \hat{A}_{\hat{P},i}, \hat{B}_2, \ldots, \hat{B}_{\hat{M}}\right] \in \mathbb{R}^{p \times n_{\mathrm{ref}}}, \\ \hat{Y} &= \left[I^{p(\hat{P}-1)}, O^{p(\hat{P}-1) \times p + m(\hat{M}-1)}\right] \in \mathbb{R}^{p(\hat{P}-1) \times n_{\mathrm{ref}}}, \\ \hat{Z} &= \left[O^{m(\hat{M}-1) \times p\hat{P}}, I_{-m}^{m(\hat{M}-1)}\right] \in \mathbb{R}^{m(\hat{M}-1) \times n_{\mathrm{ref}}}, \\ \hat{\mathcal{B}}_i &= \left[\hat{B}_{1,i}^\top, O^{m \times p(\hat{P}-1)}, I^m, O^{m \times m(\hat{M}-2)}\right]^\top \in \mathbb{R}^{n_{\mathrm{ref}} \times m}, \\ \hat{\mathcal{C}} &= \left[I^p, O^{p \times p(\hat{P}-1) + m(\hat{M}-1)}\right] \in \mathbb{R}^{p \times n_{\mathrm{ref}}}, \end{aligned} \tag{18}$$

- the first block $\hat{B}_{1,i}$ of matrix $\hat{\mathcal{B}}_i$ is adjusted in order to fulfill

$$\hat{\mathcal{B}}_i = (I^{n_{\mathrm{ref}}} - \hat{\mathcal{A}})\beta_i(I^{n_i} - \mathcal{A}_i)^{-1}\mathcal{B}_i. \tag{19}$$

Note that the matrix $\hat{\mathcal{A}}$ defines the "reference dynamics" and may be an arbitrarily chosen and possibly reduced-order system matrix $\mathcal{A}_i$ from the ensemble, while the matrices $\hat{\mathcal{B}}_i$ account for the static gain mismatch introduced by the mismatch in system dynamics. Furthermore, due to this static gain consistency, the previously introduced reference deviations $\hat{w}_i(k)$ are indeed nonzero only during the transient phase.

Aggregated Ensemble Model The reference model derived in this section is used as the prediction model for the output of the entire ensemble in the medium-level MPC. It is defined on a slower timescale than the individual references $\hat{\mathcal{S}}_i$ and - loosely speaking - is obtained by defining its state and its output as $\bar{x} = \sum_{i=1}^{N} \hat{x}_i$ and $\bar{y} = \sum_{i=1}^{N} y_i$, respectively. As a result, the aggregated ensemble model is

$$\bar{\mathcal{S}} : \begin{cases} \bar{x}(k+1) = \bar{\mathcal{A}}\bar{x}(k) + \bar{\mathcal{B}}\bar{z}(k) + \bar{w}(k), \\ \bar{y}(k) = \mathcal{C}\bar{x}(k) \end{cases} \tag{20}$$

with $\bar{\mathcal{A}} = \hat{\mathcal{A}}^{N_\tau}$, $\bar{w}(k) = \sum_{i=1}^{N} \hat{w}_i(N_\tau k)$, and

$$\bar{\mathcal{B}} = \sum_{i=1}^{N} \alpha_i \sum_{j=0}^{N_\tau - 1} \hat{\mathcal{A}}^j \hat{\mathcal{B}}_i. \tag{21}$$

The set where $\bar{w}(k)$ lies is denoted $\bar{\mathbb{W}}$ and will be defined in details later in the paper. Several thing are to be noted. Firstly, the definitions of $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$ correspond to a resampling of the fast individual reference systems to the slow timescale. Secondly, the ensemble reference does not include any information about the nonlinear input couplings $g_i(u_i)$ but only considers their resulting effects z_i on the linear parts of the subsystems.

Tracking MPC Problem The main objective of the medium level controller is to drive the ensemble output $\bar{y}$ towards a constant reference y_{ref} under the unknown reference deviation $\bar{w}(k)$ in (20) and while enforcing

$$\bar{y}(k) \in \bar{\mathbb{Y}}, \tag{22a}$$
$$\bar{z}(k) \in \bar{\mathbb{Z}}, \tag{22b}$$
$$g_i^{-1}(\alpha_i \bar{z}(k)) \in \mathbb{U}_i, \forall i \tag{22c}$$
$$\Delta\bar{z}(k) = \bar{z}(k) - \bar{z}(k-1) \in \Delta\bar{\mathbb{Z}}. \tag{22d}$$

The set $\Delta\bar{\mathbb{Z}}$ in (22d) is chosen compact and convex, containing the origin in its interior. In Sec. 4 we show that enforcing (22d) is necessary to constrain the reference deviations $\hat{w}_i(\kappa)$ for all $\kappa \geq 0$, and in turn that $\bar{w}(k) \in \bar{\mathbb{W}}$ for all $k \geq 0$. In order to guarantee all of the desired control properties listed above we employ a robust MPC scheme. The one presented in (Betti et al., 2013) is here used for a threefold reason.

- It guarantees offset-free tracking properties by reformulating the system in *velocity form*, i.e.,

$$\bar{\mathcal{S}}_\Delta : \begin{cases} \bar{\chi}(k+1) = \bar{\mathcal{A}}\bar{\chi}(k) + \bar{\mathcal{B}}\bar{\zeta}(k) + \bar{\omega}(k), \\ \bar{\epsilon}(k+1) \ = \bar{\mathcal{C}}\bar{\mathcal{A}}\bar{\chi}(k) + \bar{\epsilon}(k) + \bar{\mathcal{C}}\bar{\mathcal{B}}\bar{\zeta}(k) + \bar{\mathcal{C}}\bar{\omega} \end{cases} \tag{23}$$

  where $\bar{\chi}(k) = \bar{x}(k) - \bar{x}(k-1)$, $\bar{\epsilon}(k) = \bar{y}(k) - \bar{r}$, $\bar{\zeta}(k) = \bar{z}(k) - \bar{z}(k-1)$, $\bar{\omega}(k) = \bar{w}(k) - \bar{w}(k-1)$, and $\bar{r}$ is the current output reference value.
- It allows to address the presence of the noise term using the tube-based approach previously presented in (Mayne et al., 2005). This, in a nutshell, requires to define the nominal system corresponding with (23), i.e.,

$$\hat{\mathcal{S}}_\Delta : \begin{cases} \hat{\chi}(k+1) = \bar{\mathcal{A}}\hat{\chi}(k) + \bar{\mathcal{B}}\hat{\zeta}(k) \\ \hat{\epsilon}(k+1) \ = \bar{\mathcal{C}}\bar{\mathcal{A}}\hat{\chi}(k) + \hat{\epsilon}(k) + \bar{\mathcal{C}}\bar{\mathcal{B}}\hat{\zeta}(k) \end{cases} \tag{24}$$

  where $\hat{\zeta}(k)$ and $\bar{\zeta}(k)$ are related together. This approach makes it possible to recast the robust control problem as a non-robust one (expressed in terms of the variables of the nominal system (24)) by just suitably tightening constraints (22) and adding a further one relating $(\hat{\chi}, \hat{\epsilon})$ to $(\bar{\chi}, \bar{\epsilon})$.
- Similarly to (Limon et al., 2008) it accounts for the set-point $\bar{r}$, used in the optimization problem, as a further free variable, which allows to preserve recursive feasibility properties also in case of set-point variations and enhances the initial feasibility region.

The resulting cost function to be minimized at any slow sampling time k is

$$\begin{aligned} \bar{V}_H = &\sum_{j=k}^{k+N_H-1} \left(\|\hat{\xi}(j)\|_{\bar{Q}}^2 + \|\hat{\zeta}(j)\|_{\bar{R}}^2 \right) \\ &+ \|\hat{\xi}(N_H)\|_{\bar{P}}^2, + \|y_{\text{ref}} - \bar{r}(k)\|_{\bar{T}}^2 \end{aligned} \tag{25}$$

where $\hat{\xi} = [\hat{\chi}^\top, \hat{\zeta}^\top]^\top$, and where the weights $\bar{Q}$, $\bar{P}$, $\bar{R}$, and $\bar{T}$ are symmetric and positive definite.

3.3 Low level: Fast Shrinking Horizon MPC

The low-level systems employ local MPCs in a shrinking horizon fashion between each two consecutive slow time steps k of the medium level, which use the original system models $\mathcal{S}_i$ in (3) for their predictions. The control objective is to reduce the reference deviations $\hat{w}_i(N_\tau k)$ at the end of each shrinking horizon MPC instance. To formulate the low-level problems we first define, for each value of the medium-level control signal $\bar{z}(k)$, the corresponding steady state of the i-th subsystem, i.e.,

$$x_i^{\mathrm{s}}(k) = (I^{n_i} - \mathcal{A}_i)^{-1} \mathcal{B}_i \alpha_i \bar{z}(k). \tag{26}$$

The cost function to be minimized at a given fast sampling time $\kappa \in \{kN_\tau, \ldots, (k+1)N_\tau - 1\}$ is defined as

$$\begin{aligned} V_{L,i} = &\sum_{j=\kappa}^{(k+1)N_\tau - 1} \left(\|\beta_i x_i(j) - \hat{x}_i^o(j)\|_{Q_i}^2 + \|v_i(j)\|_{R_i}^2 \right) \\ &+ \|\beta_i x_i((k+1)N_\tau) - \hat{x}_i^o((k+1)N_\tau)\|_{P_i}^2, \end{aligned} \tag{27}$$

where the weights Q_i, P_i, and R_i are symmetric and positive definite and where $\hat{x}_i^o$ is computed as in (17) but setting $\hat{w}_i = 0$ and $\hat{x}_i^o(kN_\tau) = \beta_i x_i(kN_\tau)$. The i-th resulting low-level MPC problem is then given by

$$\begin{aligned} \min_{\boldsymbol{v}_i} \ & V_{L,i} \\ \text{s.t.} \ & \text{system } \mathcal{S}_i, \\ & g_i^{-1}(\alpha_i \bar{z}(k)) + v_i(j) \in \mathbb{U}_i, j = \kappa, \ldots, (k+1)N_\tau - 1 \\ & x_i((k+1)N_\tau) - x_i^{\mathrm{s}}(k) \in \mathbb{X}_{F,i} \\ & \hat{w}_i((k+1)N_\tau) \in \mathbb{W}_i \end{aligned} \tag{28}$$

where $\boldsymbol{v}_i = \{v_i(\kappa), \ldots, v_i(\kappa + N_\tau - 1)\}$ and the sets $\mathbb{X}_{F,i}$ are robust positively invariant (RPI), satisfying

$$\mathcal{A}_i \mathbb{X}_{F,i} \oplus \mathcal{A}_i \Delta\tilde{\mathbb{Z}} \subseteq \mathbb{X}_{F,i}, \tag{29}$$

with

$$\Delta\tilde{\mathbb{Z}}_i = -\left(I^{n_i} - \mathcal{A}_i\right)^{-1} \mathcal{B}_i \Delta\bar{\mathbb{Z}}_i. \tag{30}$$

If $\Delta\bar{\mathbb{Z}}$ is polytopic, an approximation of the minimal RPI set can be, e.g., calculated using the method proposed in (Rakovic et al., 2005). Finally, the sets $\mathbb{W}_i$, which bound the reference deviations $\hat{w}_i(k)$ of the low level subsystems $\mathcal{S}_i$ from their respective reference systems $\hat{\mathcal{S}}_i$, are defined as

$$\mathbb{W}_i = \left(\beta_i \mathcal{A}_i - \hat{\mathcal{A}}\beta_i\right) \left(\mathbb{X}_{F,i} \oplus \Delta\tilde{\mathbb{Z}}_i\right). \tag{31}$$

4. MAIN PROPERTIES

In this section we briefly discuss the main properties of the presented approach. For space reasons, we rely on the results presented in previous papers. A more detailed analysis will be the subject of future work.

Since the medium-level problem is formulated as a standard robust control one, recursive feasibility and the convergence properties of the ensemble controller follow directly from Theorem 1 in (Betti et al., 2013),

In turn, this is possible thanks to the fact that the disturbance term $\bar{w}$ can be guaranteed to lie in a time-invariant set. The proof of the latter is provided in (Petzke et al., 2018) where we showed that, for a similar hierarchical control structure, the deviation $\hat{w}_i(kN_\tau)$ between the systems $\mathcal{S}_i$ and their respective references $\hat{\mathcal{S}}_i$ is guaranteed to lie within the set $\mathbb{W}_i$, defined in (31). The latter property is implied by the constraints enforced by the low-level controller but, importantly, can be guaranteed even if no control action on the low level (i.e. $v_i \equiv 0$) is exerted.

Indeed, the recursive feasibility of the low level shrinking horizon controller follows directly from Lemma 1 in (Petzke et al., 2018), since by construction $v_i(\kappa) \equiv 0$ for all κ is always a feasible solution to problem (28).

While this result was formerly derived for linear systems, the setup at hand still underlies the same reasoning since under the assumption above the *effective* input of each subsystem is equal to $g_i(u_i) = g_i(\bar{u}_i) = g_i(g_i^{-1}(\alpha_i \bar{z})) = \alpha_i \bar{z}$, which is linear.

5. CASE STUDY

To illustrate the presented control scheme we adapt the MIMO distillation column example in Bloemen et al. (2001) with $y_1 = w_1$, $y_2 = w_2$, $D = O^{2\times 2}$, and the nonlinear input coupling functions

$$g_i(u_i;\mu_i) = \begin{bmatrix} \frac{\exp(\mu_{1,i}u_{1,i})-1}{\exp(\mu_{1,i}u_{1,i})+1}, & \frac{\exp(\mu_{2,i}u_{2,i})-1}{\exp(\mu_{2,i}u_{2,i})+1} \end{bmatrix} \quad (32)$$

with $u_i = [u_{1,i},\ u_{2,i}]^\top$ and slope parameters $\mu_{1,i}$ and $\mu_{2,i}$. Fig. 3 shows $g_i(u_i;\mu_i)$ for different values of $\mu_{j,i}$.

The first subsystem (which also served as the reference system) corresponds to the one presented in Bloemen et al. (2001) with $\mu_{1,i} = \mu_{2,i} = 1$. It is given by

$$y_1(\kappa) = \sum_{\tau=1}^{4} A_{\tau,1} y_1(\kappa-\tau) + \sum_{\tau=1}^{4} B_{\tau,1} z_1(\kappa-\tau) \quad (33)$$

with

$$A_{1,1} = \begin{bmatrix} 3.45 & 0 \\ 0 & 3.35 \end{bmatrix}, \qquad A_{2,1} = \begin{bmatrix} -4.45 & 0 \\ 0 & -4.18 \end{bmatrix},$$

$$A_{3,1} = \begin{bmatrix} 2.54 & 0 \\ 0 & 2.31 \end{bmatrix}, \qquad A_{4,1} = \begin{bmatrix} -0.54 & 0 \\ 0 & -0.47 \end{bmatrix},$$

$$B_{1,1} = \begin{bmatrix} -0.73 & 0.61 \\ 0.64 & -0.80 \end{bmatrix}, \qquad B_{2,1} = \begin{bmatrix} 1.08 & -0.93 \\ -0.87 & 1.17 \end{bmatrix},$$

$$B_{3,1} = \begin{bmatrix} -0.25 & 0.24 \\ 0.12 & -0.18 \end{bmatrix}, \qquad B_{4,1} = \begin{bmatrix} -0.11 & 0.09 \\ 0.13 & -0.20 \end{bmatrix},$$

which we then transformed into a state-space representation according to Eq. (3). We artificially created another 4 similar systems by randomly disturbing the nonzero

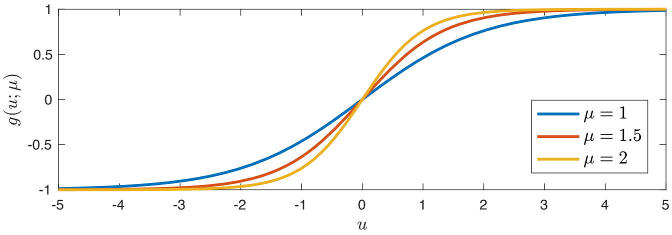

Fig. 3. Nonlinear input couplings $g(u;\mu)$ for different values of the slope-parameter μ (indices have been dropped for better readability).

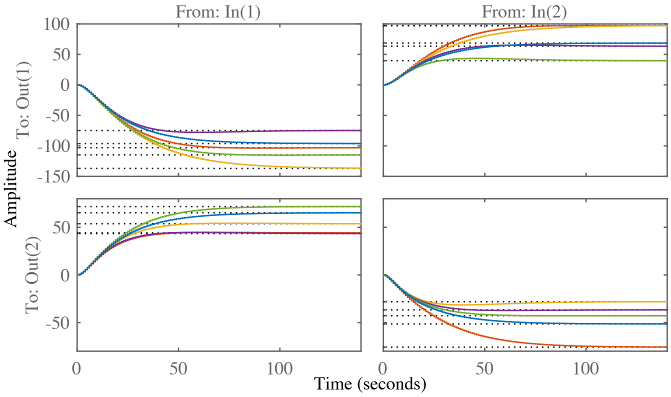

Fig. 4. Step responses of the subsystems in the ensemble.

elements of $A_{\tau,1}$ and $B_{\tau,1}$ as well as the parameters $\mu_{1,1}$ and $\mu_{2,1}$. The resulting step responses are shown in Fig. 4. In this case, all models – including the aggregated ensemble reference model for the medium-level MPC – have the same state dimension of $n = 14$. For the simulation we used the following parameters: $\tau = 1\,\mathrm{s}$, $T = 20\,\mathrm{s}$, $u_{1,i}, u_{2,i} \in [-5,\ 5]$, $\Delta\bar{\mathbb{Z}} = \{[-0.4, 0.4] \times [-0.4, 0.4]\}$ $\bar{Q} = Q_i = I^{14\times 14}$, $\bar{R} = R_i = I^{2\times 2}$. The output reference for the high-level ORS problem was set to $y_{\mathrm{ref}} = [100, -80]$, which returned $\boldsymbol{\alpha}^\star = [0.2091, 0.2901, 0.0737, 0.1381, 0.2889]$ as optimal resource sharing parameters. During the system simulation shown in Fig. 5 we changed the output reference after 150 s to $y_{\mathrm{ref}} = [100, -60]$ and again after 300 s to $y_{\mathrm{ref}} = [110, -60]$. This shows that, while the resource sharing might not be optimal anymore, the presented approach can still handle small changes around the operating point very well. Furthermore, note that the outputs of the subsystems in Fig. 5 are not just "scaled down" versions of the aggregated system model used on the medium level, but can show very different individual behavior (e.g. between 150 s and 250 s).

6. CONCLUSIONS

In this work we presented a three-layer hierarchical MPC scheme for Hammerstein systems. By employing the inverse of the nonlinear input map we were able to set up a linear optimization problem for the medium level controller, which allowed us to rely on previous results concerning recursive feasibility and convergence. However, this also meant that the thusly acquired results are generally suboptimal with respect to the consumed resource. In future work we will focus on analyzing and alleviating this optimality gap, either by using a nonlinear MPC formulation on the medium level or by employing the high

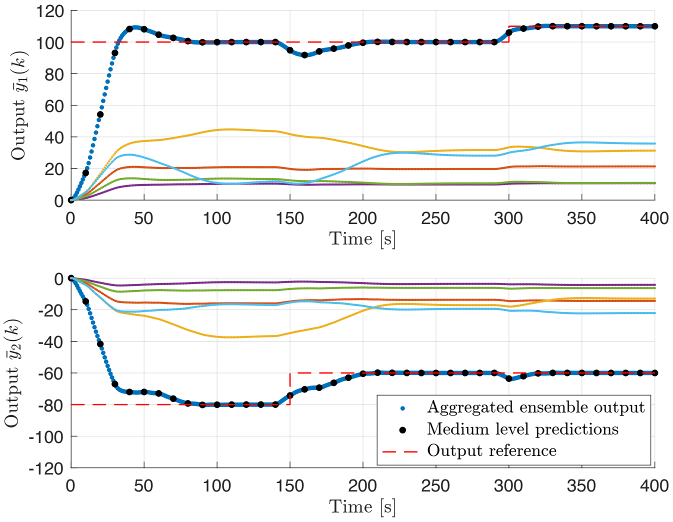

Fig. 5. Output of the controlled ensemble. Medium level predictions of the ensemble output are very close to the actual aggregated output. Outputs of individual subsystems are shown by thin colored lines. The output reference is the red dashed line.

level ORS problem in a more dynamic fashion.

The analysis of varying shares along the lines of the work in (Farina et al., 2018) is of particular interest since it would ensure overall optimality in case of significant changes in the output demand, and would allow for plug-and-play operations but, on the other hand, can compromise the feasibility of the scheme.

Furthermore, we will take a closer look at time-variant output references as well as suitable output constraint tightening approaches, and analyze their effect on the overall feasibility. Lastly, while the presented case study illustrated the general idea of the control approach, we plan on applying it to a more practical and larger scenario and also compare our results to a centralized approach.

REFERENCES

Barcelli, D., Bemporad, A., and Ripaccioli, G. (2010). Hierarchical multi-rate control design for constrained linear systems. In *49th IEEE Conference on Decision and Control*, 5216–5221.

Betti, G., Farina, M., and Scattolini, R. (2013). A robust MPC algorithm for offset-free tracking of constant reference signals. *IEEE Transactions on Automatic Control*, 58(9), 2394–2400.

Bloemen, H., van den Boom, T., and Verbruggen, H. (2001). Model-based predictive control for Hammerstein-Wiener systems. *International Journal of Control*, 74(5), 482–495.

Chan, K.H. and Bao, J. (2007). Model predictive control of Hammerstein systems with multivariable nonlinearities. *Industrial & Engineering Chemistry Research*, 168–180.

Farina, M., Zhang, X., and Scattolini, R. (2018). A hierarchical MPC scheme for coordination of independent systems with shared resources and plug-and-play capabilities. *IEEE Transactions on Control Systems Technology*, 1–12.

Fruzzetti, K., Palazoglu, A., and McDonald, K. (1997). Nolinear model predictive control using Hammerstein models. *Journal of Process Control*, 7(1), 31–41.

Garone, E., Di Cairano, S., and Kolmanovsky, I. (2017). Reference and command governors for systems with constraints: A survey on theory and applications. *Automatica*, 75, 306–328.

Harnischmacher, G. and Marquardt, W. (2007). Nonlinear model predictive control of multivariable processes using block-structured models. *Control Engineering Practice*, 15(10), 1238–1256.

Jurado, F. (2006). Hammerstein-model-based predictive control of micro-turbines. *International Journal of Energy Research*, 30(7), 511–521.

Kalabić, U., Kolmanovsky, I., Buckland, J., and Gilbert, E. (2012). Reduced order reference governor. In *51st IEEE Conference on Decision and Control (CDC)*, 3245–3251.

Lawrynczuk, M. (2015). Nonlinear predictive control for Hammerstein-Wiener systems. *ISA Transactions*, 55, 49–62.

Limon, D., Alvarado, I., Alamo, T., and Camacho, E. (2008). MPC for tracking piecewise constant references for constrained linear systems. *Automatica*, 44(9), 2382–2387.

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

Morton, A.B. (2007). *Model Aggregation of Wind Farms and Other Ensemble Systems.*

Patwardhan, R.S., Lakshminarayanan, S., and Shah, S.L. (1998). Constrained nonlinear MPC using Hammerstein and Wiener models: PLS framework. *AIChE Journal*, 44(7), 1611–1622.

Petzke, F., Farina, M., and Streif, S. (2018). A multirate hierarchical MPC scheme for ensemble systems. In *Proc. Conference on Decision and Control (CDC).*

Rakovic, S.V., Kerrigan, E.C., Kouramas, K.I., and Mayne, D.Q. (2005). Invariant approximations of the minimal robust positively invariant set. *IEEE Transactions on Automatic Control*, 50(3), 406–410.

Sadabadi, M.S., Karrari, M., and Malik, O.P. (2007). Nonlinear identification of synchronous generator using Hammerstein model with piecewise linear static maps. In *2007 IEEE Lausanne Power Tech*, 1067–1071.

Scattolini, R. (2009). Architectures for distributed and hierarchical model predictive control - a review. *Journal of Process Control*, 19(5), 723–731.

van der Veen, G., van Wingerden, J., Fleming, P., Scholbrock, A., and Verhaegen, M. (2013). Global data-driven modeling of wind turbines in the presence of turbulence. *Control Engineering Practice*, 21(4), 441–454.