Journal of Process Control 21 (2011) 715–728

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

# Sensitivity-based coordination in distributed model predictive control

Holger Scheu*, Wolfgang Marquardt

*AVT – Process Systems Engineering, RWTH Aachen University, D-52056 Aachen, Germany*

ARTICLE INFO

*Article history:*
Received 1 October 2010
Received in revised form
20 December 2010
Accepted 24 January 2011
Available online 2 March 2011

*Keywords:*
Distributed model predictive control
Distributed optimization
Optimal control
Large-scale systems

ABSTRACT

A new distributed model-predictive control method is introduced, which is based on a novel distributed optimization algorithm, relying on a sensitivity-based coordination mechanism. Coordination and therefore overall optimality is achieved by means of a linear approximation of the objective functions of neighboring controllers within the objective function of each local controller. As for most of the distributed optimization methods, an iterative solution of the distributed optimal control problems is required. An analysis of the method with respect to its convergence properties is provided. For illustration, the sensitivity-driven distributed model-predictive control (S-DMPC) method is applied to a simulated alkylation process. An almost optimal control sequence can be achieved after only one iteration in this case.



## 1. Introduction

Chemical processes are usually described by dynamic mathematical models featuring several input and output variables. Typically, industrial plants are operated by decentralized control technology based on single-input single-output PID loops enhanced by supervisory controllers [1]. The decentralized controllers are designed without an explicit consideration of the interactions between the different loops in the plant. Hence, optimization-based methods such as linear or nonlinear model predictive control (MPC) or dynamic real-time optimization (DRTO) have become important technologies, as they maintain operational constraints. We refer to the papers of Kano and Ogawa [1] and Qin and Badgwell [2] for a summary on the industrial state of the art on MPC-technology. Model-predictive controllers consider all interactions due to their centralized implementation. Typically, favorable constraint pushing and good performance can be achieved. However, the design and maintenance of large-scale centralized predictive controllers is involved. Decentralization of the established constraint-handling controllers with negligible performance deterioration is therefore desirable.

Today, an increasing activity in research on model-based decentralized and distributed control methods can be observed. This trend is driven by opportunities to achieve better computational performance [3] and to remove possible communication bottlenecks. Furthermore, reliability and maintainability could be improved compared to a monolithic centralized solution [4]. Last but not least, completely new applications requiring decentralized control have to be addressed. These new applications are positioned mainly in the field of autonomous vehicles [5–7], such as aircrafts or satellites. Typically these systems are modeled by some linear mechanical systems, where each subsystem is modeled by the same dynamics [8]. An important characteristic of this class of systems, commonly referred to as multi-agent systems (MAS), is the autonomy of the different subsystems (or agents). They are not coupled and hence do not interact. However, these agents share some common goals such as consensus [8], pattern formation (e.g. flocking) or the avoidance of collisions [5,9]. In order to achieve these common goals, control methods are applied, that lead to a rational interaction between the agents in MAS. These methods use various modern control tools, including linear matrix inequalities [7] or control barrier functions [9]. While the systems considered can usually be described by low order models, one of the main challenges is the time-dependent communication topology. This time-dependence is caused by varying communication links due to changing distances, appearing obstacles or exogeneous disturbances. While in many papers a constant topology is assumed (e.g. [10,11]), first results exist on time-dependent topologies (e.g. [12,13]).

Process control problems usually feature several differences compared to MAS. In particular, they are characterized by high order dynamics with strong nonlinearities. In a faultless system, the (communication) topology is fixed. Typically, there is no complete interaction structure linking a subsystem with all others; rather, a subsystem interacts only with a few neighbors.

* Corresponding author. Tel.: +49 2418097007.
*E-mail addresses:* holger.scheu@avt.rwth-aachen.de (H. Scheu), wolfgang.marquardt@avt.rwth-aachen.de (W. Marquardt).

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.01.013

The subsystems of a process plant, i.e. the process units, exchange material, energy, or signals between each other. They are either controlled by a single centralized (multi-variable) controller or by a set of decentralized control systems. Typically, the design of the controllers of such a decentralized control structure does not account for the interactions between the subsystems explicitly. However, there is a desire to guarantee nominal stability, feasibility, optimality, reliability and maintainability for the control systems implemented in the plant. Distributed model predictive control (DMPC) methods are expected to contribute towards this aim because they can combine the optimality properties of centralized predictive controllers and the modularity and flexibility of decentralized control systems by means of communication, cooperation, or coordination [14].

The basis of distributed control, and in particular, distributed optimal control, has already been established in the early 1970s. Mesarovic et al. [15] presented a seminal monograph on distributed control systems. They formulated some very general principles on how a coordinated control structure can be implemented including the "interaction balance principle" and the "interaction prediction principle" [16]. At the same time, two main decomposition principles have been proposed as a starting point for the development of DMPC, namely primal decomposition [17,18] and dual decomposition [19]. Primal decomposition is based on resource allocation, where shared resources, which are described by primal variables, are allocated by a coordinator. On the other hand, dual decomposition is driven by price coordination. Each decomposed subproblem decides on the use of its resources. However, the use of these resources is coordinated by the prices, which are the dual variables to be adjusted by the coordinator. If a DMPC scheme includes a coordinator, it is often referred to as a hierarchical control system [20].

Various survey papers exist on distributed control. In 1975, Singh et al. [21] presented an early review on practical distributed control methods for interconnected systems, where all architectures presented feature both, hierarchical and distributed elements. A number of variations of dual decomposition [19] including the three-level method of Tamura [22] are reviewed by these authors. They also present indirect methods suggested by Smith and Sage [23] and by Takahara [24], the "pseudo-model coordination method" [25], as well as some suboptimal methods. In 1977, Mahmoud [26] presented a very comprehensive overview. While the main subject of that article covers multilevel optimization techniques, it also deals with early progress in multilevel systems identification as well as the application to water resource systems. In 1978, Sandell et al. [27] presented another survey on the topics of model simplifications, stability analysis of interconnected systems and decentralized control methods. In a recent survey, Rawlings and Stewart [14] summarize the present status of research in the field of coordinated optimization-based control, as well as opportunities and challenges for future research. The different control architectures to be considered in hierarchical and distributed control have been reviewed recently by Scattolini [20].

The main research focus in hierarchical and distributed model predictive control is geared towards linear systems. Wakasa et al. [28], for example, apply dual decomposition to SISO linear time-invariant systems, the subsystems of which are only coupled by their outputs. The output constraints correspond to the formation of a flock of multiple vehicles. The dual problem is solved using a subgradient optimization algorithm. Necoara and Suykens [29] propose a proximal center-based dual decomposition method, which is based on a smoothing of the Lagrange function to result in a continuous convex objective function. The method is successfully applied to a system of coupled linear oscillators [30]. Venkat et al. [31] propose a DMPC method for linear discrete-time systems based on primal decomposition, which they call "feasible cooperation-based control". The objective function of the complete system is included in all subsystem controllers to ensure coordination. In each iteration, the cooperative method generates a feasible solution of the optimal control problem. Based on this idea, Stewart et al. [32] have developed a distributed controller, which guarantees stability even in the suboptimal case. Not only the full objective function but also the full system dynamics have to be considered by each of the distributed controllers in this scheme, however. Another DMPC formulation based on resource allocation, i.e. a primal decomposition approach, is proposed by Marcos et al. [33]. Their method can be applied to a system described by linear step-response models. The method is successfully applied to a fluid catalytic cracking process, where fast convergence can be observed.

So far, there exist only few results related to nonlinear DMPC. Talukdar et al. [34] propose a cooperative distributed model predictive control method for nonlinear systems. They assume that the subsystems are only able to communicate and cooperate with neighboring subsystems. The method is applied to the IEEE 118 bus test case in order to study how to prevent cascading failures of that network. Liu et al. [35] propose a Lyapunov-based DMPC method. An additional equation ensuring a decrease of the Lyapunov function is added to the model to ensure closed-loop stability. In this method, only the decision variables are distributed, while the controllers of all the subsystems consider the full state vector and the objective function of the complete system. While in [35] the coordination of two Lyapunov-based model predictive controllers is described, the method is generalized in [36] for an arbitrary number of controllers in a sequential or iterative controller architecture. The authors successfully apply their method to a simulated alkylation process, which will also be picked up as an illustrative example in this paper. Dunbar [37] proposes a distributed MPC method and applies the method to a system of coupled nonlinear oscillators. Magni and Scattolini [38] introduce another stabilizing decentralized MPC scheme for nonlinear discrete-time systems. Since the method does not involve any coordination or cooperation scheme, the resulting control system is suboptimal.

In this paper, we will introduce a new sensitivity-driven distributed model predictive control (S-DMPC) scheme, which is based on a novel distributed dynamic optimization algorithm [39] motivated by so-called "goal-interaction operators" [15]. We assume the objective function of the complete system to be separable. The objective functions of the subsystems are modified using information on the complete system to achieve optimality of the distributed control scheme. The modification of the objective function of a given subsystem incorporates a linearization of the objective functions of the neighboring subsystems. Hence, coordination of the subsystem controllers is based on first order sensitivities. Each of the distributed controllers considers only a part of the full objective function and a reduced set of constraints and decision variables. In this paper, we focus on linear systems, though the idea also carries over to the nonlinear case. We prove, that the distributed optimization method converges under given assumptions to the solution of the complete system. The resulting DMPC method is applied to a simulated chemical process to illustrate its capabilities. In this case study, fast convergence can be observed: Almost optimal control performance can be achieved after only one iteration.

The remainder of this paper is organized as follows: In Section 2, we state the optimal control problem to be solved. Section 3 presents the distributed optimization method and a convergence analysis. Section 4 presents the formulation of closed-loop S-DMPC using the distributed optimization method of Section 3. In Section 5, the linear S-DMPC method is applied to a simulated nonlinear alkylation process. Finally, we conclude the paper with a summary and an outlook in Section 6.

## 2. Optimal control problem formulation

We consider a system $\Sigma$ consisting of $N$ linear time-invariant subsystems $\Sigma_i$:

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j=1}^{N} A_{ij}x_j(t) + B_{ij}u_j(t), \quad x_i(0) = x_{i,0}, \\ \forall i &\in \{1, \ldots, N\}. \end{aligned} \tag{1}$$

$t$ denotes time. $x(t) = \langle x_1(t), \ldots, x_N(t)\rangle$ [1], with $x_i(t) \in \mathbb{R}^{n_{x_i}}$, is the state vector with initial condition $x_0 = \langle x_{i,0}, \ldots, x_{N,0}\rangle$ and $u(t) = \langle u_1(t), \ldots, u_N(t)\rangle$ is the aggregated input vector, where $u_i(t) \in \mathbb{R}^{n_{u_i}}$ is the local input vector of subsystem $\Sigma_i$. $A = [A_{ij}]_{i,j\in\{1,\ldots,N\}}$, with $A_{ij} \in \mathbb{R}^{n_{x_i}\times n_{x_j}}$, refers to the system matrix, and $B = [B_{ij}]_{i,j\in\{1,\ldots,N\}}$, with $B_{ij} \in \mathbb{R}^{n_{x_i}\times n_{u_j}}$, denotes the input matrix. The mixed input and state constraints related to subsystem $\Sigma_i$ are described by

$$0 \le D_i\langle x, u\rangle + e_i, \quad \forall i \in \{1, \ldots, N\}, \tag{2}$$

with constant matrices $D_i$ and constant vectors $e_i$. Note, that we do not assume separable constraints. Hence, all constraints may include variables of the full system $\Sigma$. Separable constraints, i.e. those constraints which only include variables $x_i$ and $u_i$, are included in the $i$-th set of constraints. However, for non-separable constraints, there exist several possibilities for their allocation to control problem $i$. For example, these constraints could be pragmatically assigned redundantly to those local control problems, the states of which enter the constraint. This choice is part of control design and may affect the convergence properties of the methods proposed in Sections 3 and 4.

DMPC aims at the solution of the optimal control problem for the overall system $\Sigma$, i.e.

$$\min_u \Phi = \sum_{i=1}^{N} \Phi_i(x_i, u_i) \tag{3a}$$

$$\text{s.t. } \Phi_i = \frac{1}{2}\int_{t_0}^{t_f} x_i^T Q_i x_i + u_i^T R_i u_i dt, \tag{3b}$$

$$\dot{x} = Ax + Bu, \quad x(t_0) = x_0, \tag{3c}$$

$$0 \le D\langle x, u\rangle + e, \tag{3d}$$

on a finite (receding) horizon $[t_0, t_f]$, with a separable (or additive), quadratic objective function (3a)[14,21], symmetric positive definite weighting matrices $Q_i$ and $R_i$, $i \in \{1, \ldots, N\}$, with appropriate dimensions, and with[2]

$$D = \langle D_1, \ldots, D_N\rangle, \tag{3e}$$

$$e = \langle e_1, \ldots, e_N\rangle. \tag{3f}$$

## 3. Distributed optimization strategy

The basis for the sensitivity-driven distributed model-predictive control (S-DMPC) proposed in Section 4 is a new distributed optimization method, which relies on sensitivity-based coordination[3][39]. The open-loop optimal control problem (3) is transcribed into a quadratic programming problem (QP) in two steps [40,41]:

1. The input vector functions $u_i(t)$ are approximated by an expansion in a complete function space spanned by the base functions $\phi_l(t)$. B-splines [42,43] provide a flexible class of bases, which includes the piecewise constant representation used exemplarily in this work. In particular, the input functions are approximiated by finite sums $u_{i,j}(t) = \sum_l p_{i,j,l}\ \phi_l(t)$, where $i$ refers to the subsystem and $j$ to the $j$-th component of the corresponding input $u_i(t)$. The parameters $p_{i,j,l}$ are organized in parameter vectors $p_i = \langle p_{i,1}, \ldots, p_{i,n_{u_i}}\rangle$ with $p_{i,j}$ comprising the parameters associated with the $j$-th scalar input $u_{i,j}(t)$ of system $i$ and in the parameter vector $p = \langle p_1, \ldots, p_N\rangle$ concatenating the control vector parameterization of the overall system.
2. The state equations are solved analytically and the objective function and the constraints are evaluated.

This transcription is straightforward to result in the quadratic program (QP)

$$\min_p \sum_{i=1}^{N} \Phi_i(p) \tag{4a}$$

$$\text{s.t. } \Phi_i(p) = \frac{1}{2}p^T\mathbf{A}^i p + p^T\mathbf{B}^i + \mathbf{C}^i, \tag{4b}$$

$$c_i(p) = \mathbf{D}^{i\,T} p + \mathbf{E}^i \ge 0, \tag{4c}$$

with

$$\mathbf{A}^i = \begin{bmatrix} \mathbf{A}^i_{11} & \cdots & \mathbf{A}^i_{1N} \\ \vdots & & \vdots \\ \mathbf{A}^i_{N1} & \cdots & \mathbf{A}^i_{NN} \end{bmatrix} \in \mathbb{R}^{n_p\times n_p}, \tag{4d}$$

$$\mathbf{B}^i = \langle \mathbf{B}^i_1, \ldots, \mathbf{B}^i_N\rangle \in \mathbb{R}^{n_p}, \tag{4e}$$

$$\mathbf{C}^i \in \mathbb{R}, \tag{4f}$$

$$\mathbf{D}^i = \langle \mathbf{D}^i_1, \ldots, \mathbf{D}^i_N\rangle \in \mathbb{R}^{\kappa_i\times n_p}, \tag{4g}$$

$$\mathbf{E}^i \in \mathbb{R}^{\kappa_i}, \tag{4h}$$

$$\forall i \in \{1, \ldots, N\}.$$

$\kappa_i$ refers to the number of constraints associated with subsystem $\Sigma_i$, with $\kappa_i \le n_{p_i}$. $n_{p_i}$ is the number of control parameters related to input vector $u_i$ and $n_p = \sum_{i=1}^{N} n_{p_i}$ denotes the number of control parameters of the overall system. The matrices $\mathbf{A}^i$, $\mathbf{B}^i$, $\mathbf{C}^i$, $\mathbf{D}^i$, and $\mathbf{E}^i$ in QP (4) are computed from the matrices and vectors $A$, $B$, $D$, and $e$ appearing in the optimal control problem (3). The derivation of QP (4) is presented in Appendix A for reference.

The distributed solution of QP (4) requires its decomposition and the coordination of the resulting subproblems, which are described for a convex nonlinear program (a generalization of QP (4), if this QP is convex) in the remainder of this section. We propose a sensitivity-based coordination mechanism, where the overall objective $\Phi$ is approximated in each of the local optimal control problems by a linearization of the contributions $\Phi_j$, $j \ne i$, of the neighboring subsystems to the overall objective $\Phi$ while the local quadratic objective function $\Phi_i$ of the subsystem $\Sigma_i$ considered is retained. A convergence analysis is provided subsequently for the convex quadratic program.

### 3.1. Sensitivity-based coordination

The coordination algorithm for a parametric nonlinear optimization problem generalizing QP (4) is stated as follows:

[1] $\langle a, \ldots, b\rangle$ is used as a shorthand for $[a^T, \ldots, b^T]^T$.
[2] We again introduce the shorthand $\langle A, \ldots, B\rangle = [A^T, \ldots, B^T]^T$ for convenience.
[3] This optimization method has been referred to as "gradient-based distributed dynamic optimization" (GBDDO) method in [39].

**Algorithm 1.**

1. Choose feasible parameters $p^{[0]}$, and an initial guess of the Lagrange parameters $\lambda^{[0]}$ and set $k := 0$.
2. The control parameters $p_i^{[k]}$ and Lagrange parameters $\lambda_i^{[k]}$, $\forall \in \{1, \ldots, N\}$, are communicated to all local controllers.
3. Solve the local optimization problems

$$\min_{p_i} \Phi_i^*(p) \tag{5a}$$

$$\text{s.t. } c_i(p) \geq 0, \tag{5b}$$

with the strictly convex objective functions

$$\Phi_i^* = \Phi_i(p) + \left[ \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\partial \Phi_j}{\partial p_i}\bigg|_{p^{[k]}}^{T} - \lambda_j^{[k]\,T} \frac{\partial c_j}{\partial p_i}\bigg|_{p^{[k]}} \right] (p_i - p_i^{[k]}) \tag{5c}$$

to obtain the local minimizers $p_i^{[k+1]}$ and the Lagrange multipliers $\lambda_i^{[k+1]}$, $\forall i \in \{1, \ldots, N\}$. Note, that the solution of the local optimization problems for $p_i^{[k+1]}$ assumes that all other control parameters are fixed at the previous iterates $p_j^{[k]}$, $\forall j \neq i$.
4. Set $k := k + 1$ and go back to 2.
5. Stop, if $p^{[k]}$ satisfies some convergence criterion.

Here, $k$ refers to the iteration index, and $\partial \Phi_j / \partial p_i$ and $\partial c_j / \partial p_i$ are the first-order sensitivities of the objective function and the inequality constraints, respectively, corresponding to subsystem $\Sigma_j, j \neq i$, with respect to the control parameters $p_i$. $c_j$ are the constraint functions related to system $\Sigma_j$, and $\lambda_j^{[k]}$ are the Lagrange multipliers at iteration $k$ related to the $j$-th NLP.

There exist several possibilities for the choice of the stopping criterion in step 5. On the one hand, in order to limit the deviation of the parameters $p^{[k]}$ from the optimal parameters $p^*$, a relative change of the parameters can be calculated as a simple measure for convergence, i.e. $\epsilon_{\text{rel}} = ||p^{[k]} - p^{[k-1]}||_2 / ||p^{[k]}||_2$. Alternatively, the merit function [44] can be observed to define a stopping criterion for the iterative algorithm. On the other hand, since the computing time is always limited in a real-time implementation, a fixed number of iterations $k = 1, \ldots, I$ may be used as the stopping criterion, where an admissible $I$ is estimated from the computational demand of one iteration. Obviously, a suboptimal solution results in this case. In any case, due to the iterative nature of the algorithm and the presence of non-separable constraints (2) or (4c), a feasibility check of all inequality constraints has to be part of any convergence test, because feasibility of the problem cannot be guaranteed at present.

Equality of the minimizers obtained from the distributed Algorithm 1 and those of the centralized problem

$$\min_{p} \sum_{i=1}^{N} \Phi_i(p), \tag{6a}$$

$$\text{s.t. } c_i(p) \geq 0, \quad \forall i \in \{1, \ldots, N\} \tag{6b}$$

is stated in the following theorem.

**Theorem 2.** Assume the following conditions hold:

- The fixed-point iteration defined by Algorithm 1 is convergent and its fixed-point satisfies the linear independence constraint qualification (LICQ) [44].
- The minimizer $p^*$ solves the centralized problem (6) and satisfies the LICQ.
- The objective functions $\Phi_i$, $\forall i \in \{1, \ldots, N\}$, are strictly convex.
- The inequality constraint functions $c_i$, $\forall i \in \{1, \ldots, N\}$, are concave.

Then, the minimizer $p^{[k]}$, $k \to \infty$, of the local problems (5) and the minimizer $p^*$ of the centralized problem (6) are the same, i.e. $\lim_{k \to \infty} p^{[k]} = p^*$.

**Proof.** We define the Lagrange functions

$$\mathcal{L} = \sum_{i=1}^{N} (\Phi_i(p) - \lambda_i^T c_i(p))$$

for the centralized problem and

$$\mathcal{L}_i = \Phi_i + \left[ \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\partial \Phi_j}{\partial p_i}\bigg|_{p^{[k]}}^{T} - \lambda_j^{[k]\,T} \frac{\partial c_j}{\partial p_i}\bigg|_{p^{[k]}} \right] (p_i - p_i^{[k]}) - \lambda_i^T c_i(p)$$

for the local problems. At convergence, the necessary conditions of optimality (NCO) of the local problems are given by

$$\begin{aligned} 0 &= \frac{\partial \Phi_i(p)}{\partial p_i} + \sum_{\substack{j=1 \\ j \neq i}}^{N} \left( \frac{\partial \Phi_j(p)}{\partial p_i} - \frac{\partial c_j(p)}{\partial p_i}^{T} \lambda_j \right) - \frac{\partial c_i(p)}{\partial p_i}^{T} \lambda_i \\ &= \sum_{j=1}^{N} \left( \frac{\partial \Phi_j(p)}{\partial p_i} - \frac{\partial c_j(p)}{\partial p_i}^{T} \lambda_j \right), \\ c_i(p) &\geq 0, \\ \lambda_i &\geq 0, \\ \lambda_i^T c_i(p) &= 0, \\ \forall i &\in \{1, \ldots, N\}, \end{aligned}$$

which are the same as the NCO of the centralized problem. Since the NCO of the centralized and the local problems are the same and both problems have been assumed to be strictly convex, the minimizer computed from the iteration in Algorithm 1 is the same as the global minimizer of the centralized problem (6), provided the iterative solution of the former converges. □

Since the convex NLP (5) and (6) generalize QP (3) and (4), if the latter are convex, Theorem 2 is also valid for the distributed solution of convex QP. For each of the local optimization problems (5), the overall objective function is considered, though all nonlocal contributions (i.e. those which result from other subsystems) are simplified by linear approximation. Compared to the dual optimization method [45] no additional optimization layer has to be implemented for the sensitivity-based coordination suggested.

In order to improve the convergence properties or even enforce convergence of the iteration, the objective function (5c) can be extended to become

$$\Phi_i^+ = \Phi_i^* + \frac{1}{2} (p_i - p_i^{[k]})^T \Omega^i (p_i - p_i^{[k]}), \tag{7}$$

with a symmetric positive definite matrix $\Omega_i \in \mathbb{R}^{n_{p_i} \times n_{p_i}}$. This extension reproduces a secant method originally suggested by Wegstein [46,47] in the context of the derivative-free solution of nonlinear equations. Furthermore, the extension (7) is a generalization of the proximal minimization algorithm [48–50] and does not affect the singular points of the optimization problem. At convergence, $p_i$ approaches $p_i^{[k]}$, such that the differences in Eq. (7) vanish and

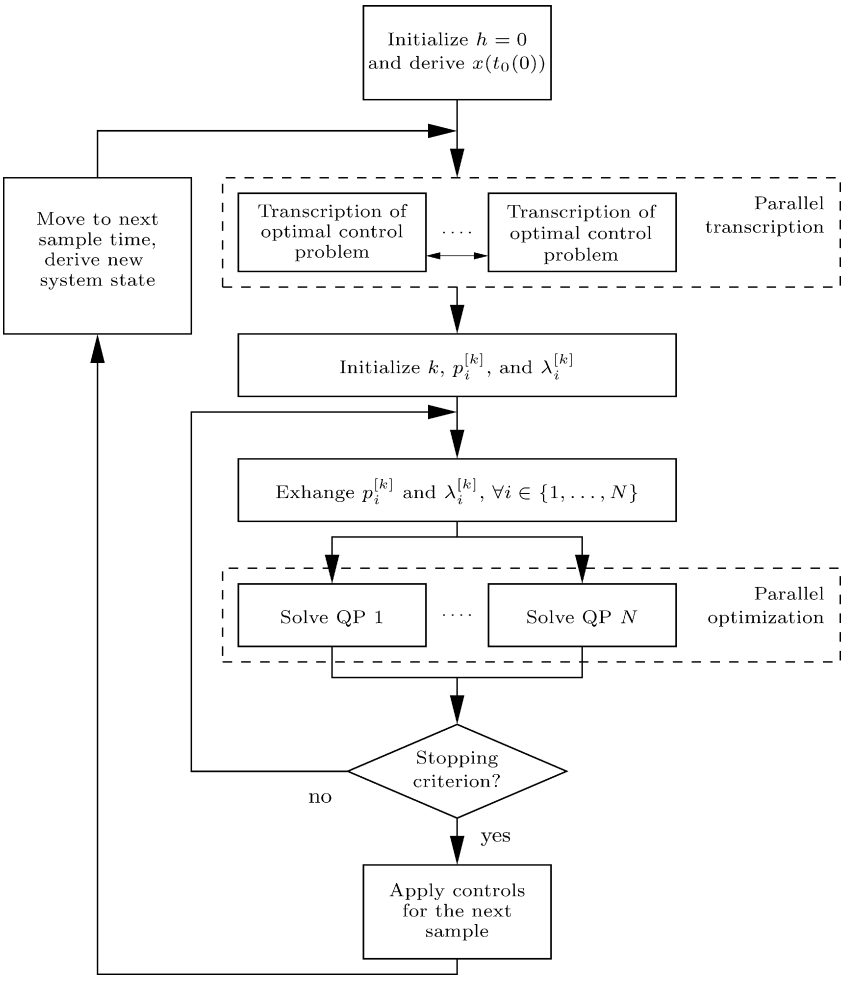


**Fig. 1.** Sensitivity-driven DMPC.

$\Phi_i^+ = \Phi_i^*$. Appropriate methods for the calculation of the matrices $\Omega^i$ are available [46,47]. Such computations only require an evaluation of the optimization problem (5) but not of its gradients.

### 3.2. Convergence analysis

In this section, convergence of the coordination Algorithm 1 introduced in Section 3.1 is analyzed for QP (4). The minimizer $p_i^{[k+1]}$ solves the local QP problems

$$\min_{p_i} \Phi_i^+ \tag{8a}$$

$$\text{s.t. } \Phi_i^+ = \frac{1}{2}\tilde{p}_i^{[k]T}\mathbf{A}^i\tilde{p}_i^{[k]} + \tilde{p}_i^{[k]T}\mathbf{B}^i + \mathbf{C}^i + \left[\sum_{\substack{j=1 \\ j\neq i}}^{N}\left(\mathbf{A}_{i1}^j \ldots \mathbf{A}_{iN}^j p^{[k]} + \mathbf{B}_i^j - \mathbf{D}_i^j\lambda_j^{[k]}\right)\right]^T (p_i - p_i^{[k]}),$$

$$+\frac{1}{2}(p_i - p_i^{[k]})^T\Omega^i(p_i - p_i^{[k]}), \tag{8b}$$

$$c_i(\tilde{p}_i^{[k]}) = \mathbf{D}^{iT}\tilde{p}_i^{[k]} + \mathbf{E}^i \geq 0, \tag{8c}$$

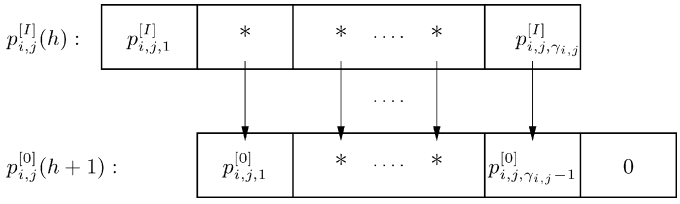


**Fig. 2.** Reinitialization of parameters $p$ for piecewise-constant input parameterization.

with $\tilde{p}_i^{[k]} = \langle p_1^{[k]}, \ldots, p_{i-1}^{[k]}, p_i, p_{i+1}^{[k]}, \ldots, p_N^{[k]} \rangle \in \mathbb{R}^{n_p}$. Each of the subproblems only involves the decision variables $p_i$ (and not the full vector $p$), as the parameters $p_j = p_j^{[k]}$, $j \neq i$, are assumed to be fixed. Furthermore, only the local inequality constraints, cf. Eq. (8c), are included in each subproblem. Hence, the complexity of the decomposed subproblems is much lower than the complexity of the original problem.

For the convergence analysis of the algorithm, the following assumptions are introduced:

**Assumption 3.** The matrices $\mathbf{A}^i$ and its submatrices $\mathbf{A}_{ii}^i$ are positive definite for all $i$.

**Assumption 4.** The matrices $\mathbf{D}_i^i$ have full rank, i.e. $\text{rank}(\mathbf{D}_i^i) = \kappa_i$.

Assumption 3 implies strict convexity of the centralized as well as the decentralized QP problems. Assumption 4 implies a proper

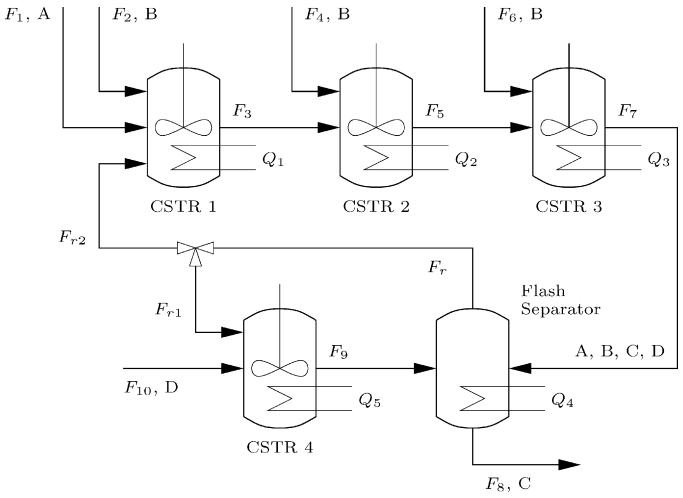


**Fig. 3.** Process flow diagram for alkylation of benzene [36].

assignment of the constraints to the local QP problems and assures that the constraints satisfy the LICQ. Consequently, it also guarantees a sequence of feasible points during the solution of the local QP problems if initialized with a feasible point. Assumption 4 could be relaxed to only refer to the set of active constraints. However, since the active set is unknown a priori, the chosen formulation is preferred.

Furthermore, the inequality constraints (8c) are separated into active and inactive constraints. $\mathbf{D}_{\mathcal{A}}^{i\,T}\tilde{p}_i^{[k]} + \mathbf{E}_{\mathcal{A}}^{i} = 0$ with $\lambda_{j|\mathcal{A}} > 0$ are the active constraints, while $\mathbf{D}_{\mathcal{I}}^{i\,T}\tilde{p}_i^{[k]} + \mathbf{E}_{\mathcal{I}}^{i} > 0$ with $\lambda_{j|\mathcal{I}} = 0$ are the inactive constraints.

With these preliminaries, the following theorem can be formulated:

**Theorem 5.** The sensitivity-driven coordination Algorithm 1 applied to QP (4) (see Eq. (8) for appropriate notation) converges under Assumptions 3 and 4, if

$$L = \left\| \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \mathbf{A}_{\text{diag}}^{\Omega} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}_{\text{diag}}^{T} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{N} \mathbf{A}^{j} & -\mathbf{D} \\ -\mathbf{D}^{T} & 0 \end{bmatrix} \right\|_{\mathcal{A}} < 1, \quad (9)$$

holds for all active sets. The matrices in Eqs. (4) and (9) are related as follows:

$$\mathbf{A}_{\text{diag}}^{\Omega} = \mathbf{A}_{\text{diag}} + \Omega, \quad (10a)$$

$$\mathbf{A}_{\text{diag}} = \text{diag}(\mathbf{A}_{11}^{1}, \ldots, \mathbf{A}_{NN}^{N}), \quad (10b)$$

$$\Omega = \text{diag}(\Omega^{1}, \ldots, \Omega^{N}), \quad (10c)$$

$$\mathbf{D}_{\text{diag}} = \text{diag}(\mathbf{D}_{1}^{1}, \ldots, \mathbf{D}_{N}^{N}), \quad (10d)$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \vdots \\ \mathbf{D}_N \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1^1 & \ldots & \mathbf{D}_1^N \\ \vdots & & \vdots \\ \mathbf{D}_N^1 & \ldots & \mathbf{D}_N^N \end{bmatrix} = \begin{bmatrix} \mathbf{D}^1 & \ldots & \mathbf{D}^N \end{bmatrix}. \quad (10e)$$

A check of the (conservative) convergence conditions (9) requires the evaluation of the inequality constraints for any possible active set $\mathcal{A}$. Though this analysis can be done during control system design, it may be computationally involved

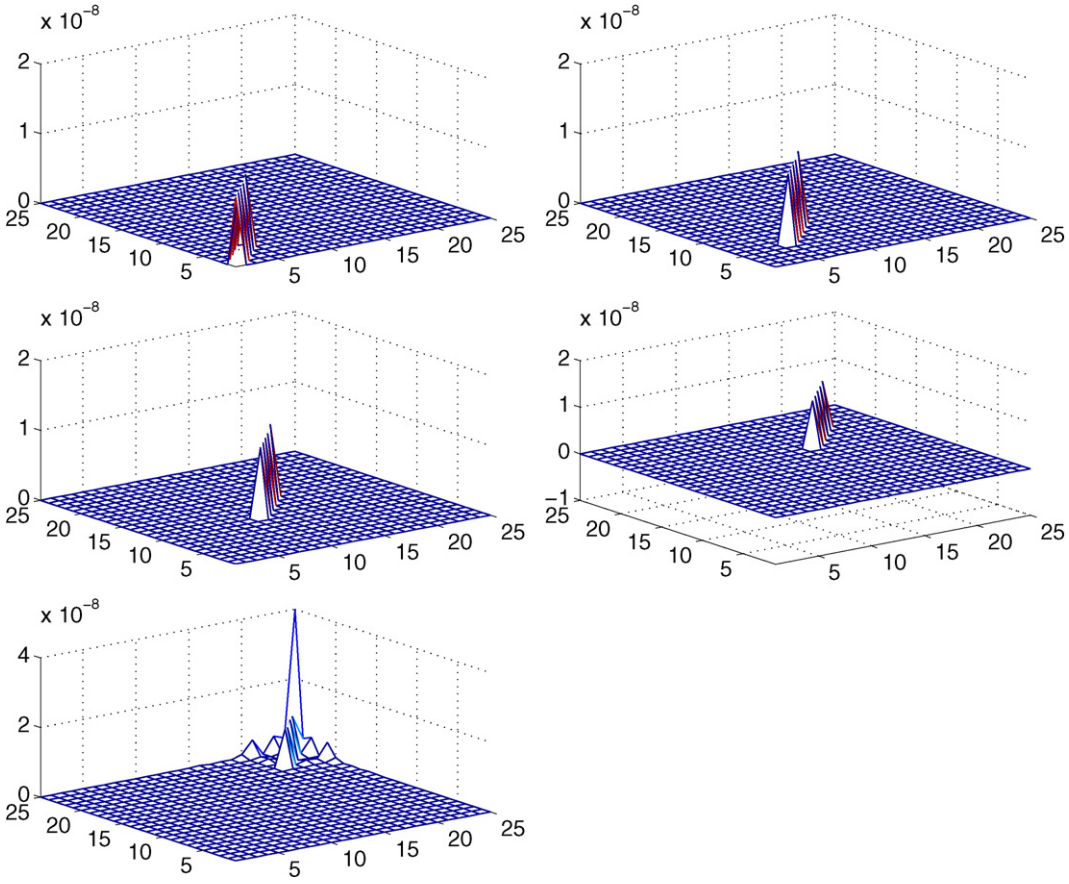


**Fig. 4.** Hessians $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$, $\mathbf{A}^4$ and $\mathbf{A}^5$ from top left to bottom right. The graphs display the magnitude of all $25 \times 25$ entries.

because of the combinatorial growth of the cardinality of $\mathcal{A}$ with the number of inequality constraints. Note that the convergence properties depend on the decomposition of the coupled constraints (see Eqs. (2) and (4c)), as the matrix $\mathbf{D}$ and in particular matrix $\mathbf{D}_{\text{diag}}$ depend on the decomposition of the constraints. This is an additional degree of freedom during control system design. The proof of Theorem 2 is provided in Appendix B.

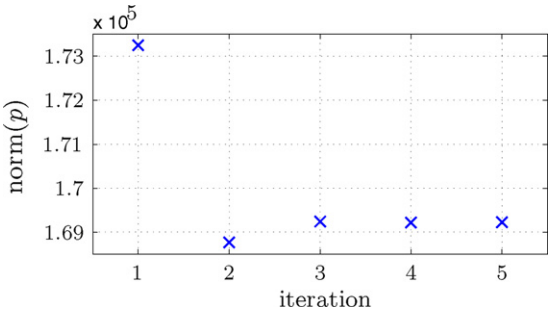


**Fig. 5.** Convergence of the distributed method.

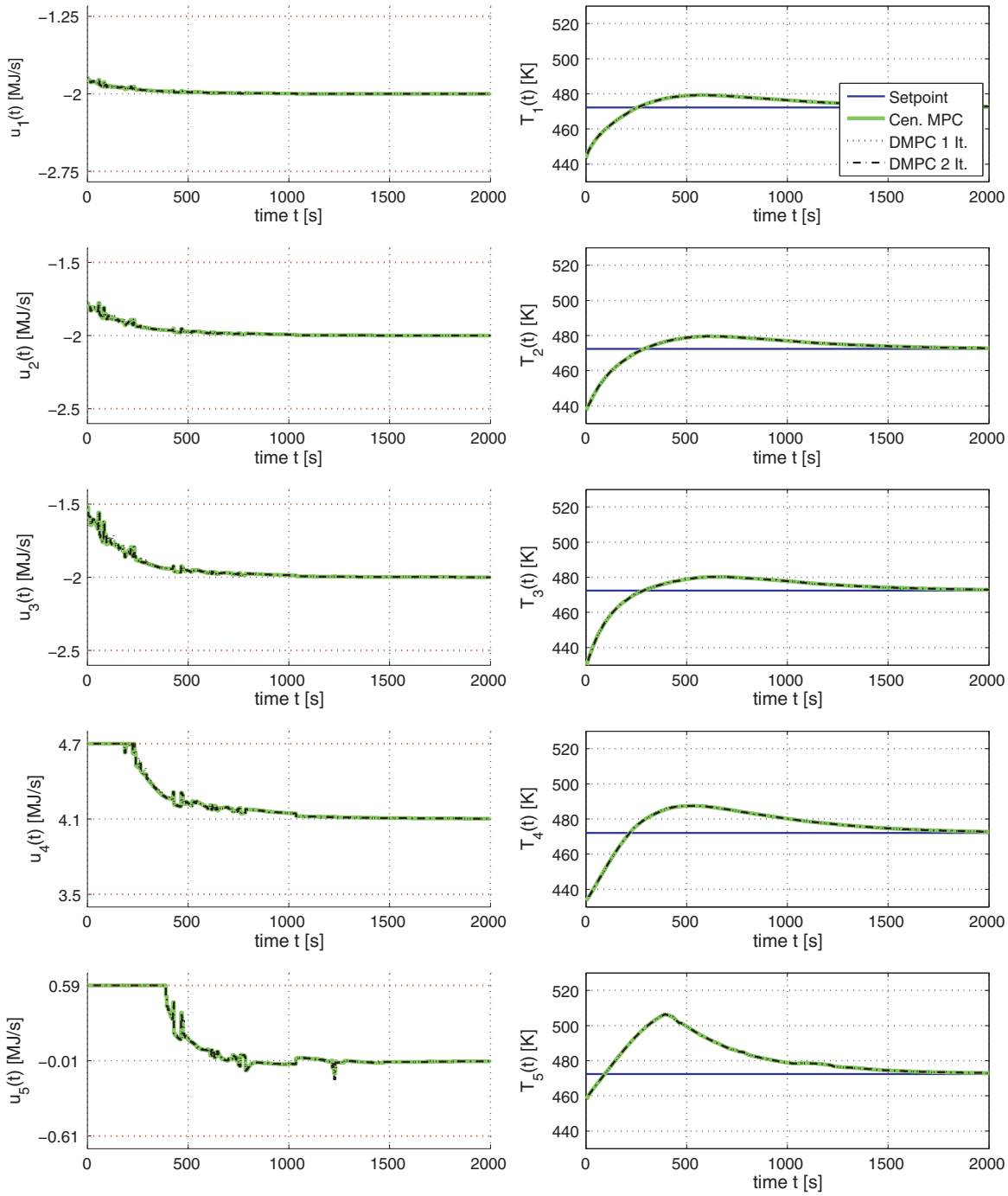


**Fig. 6.** Temperature set-point changes from steady state: manipulated variables (heat input) on the left; controlled variables (temperatures) on the right.

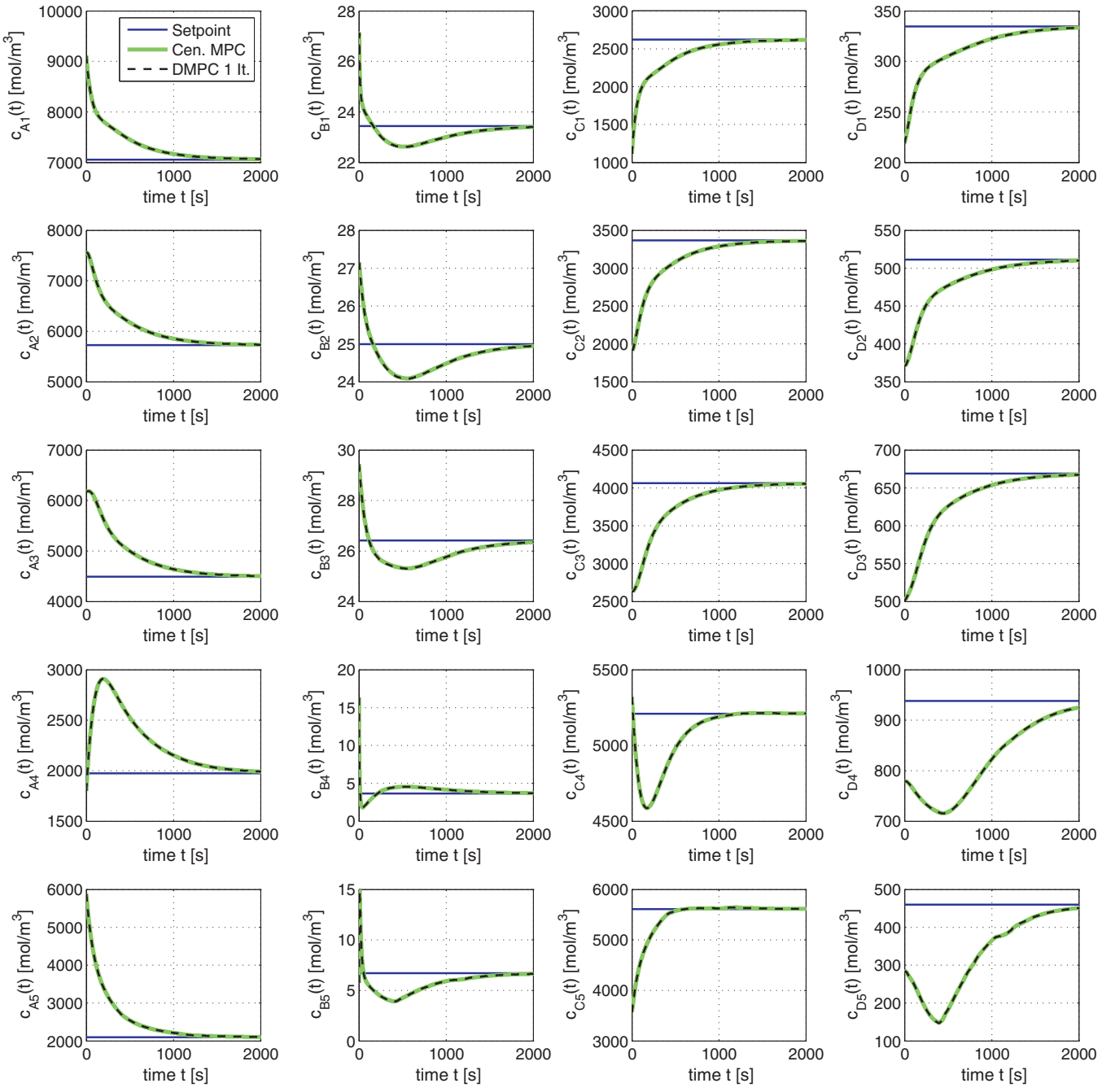


**Fig. 7.** Concentration trajectories in all 5 units (from top to bottom) after temperature setpoint change in closed-loop.

## 4. Sensitivity-driven distributed model-predictive control

So far, the solution of an open-loop optimal control problem has been considered. Here, the results are extended to the closed-loop S-DMPC formulation. An overview of the S-DMPC method is given in Fig. 1. S-DMPC relies on the distributed optimization method discussed in the preceding section on a moving horizon $[t_0(h), t_f(h)]$, where $h$ is the horizon index. Then, the S-DMPC method is given by the following

**Algorithm 6.**

1. Set $h := 0$ and fix the initial system state $x(t_0(0))$.
2. Transcribe the optimal control problem to compute $\mathbf{A}^i$, $\mathbf{B}^i(h)$, $\mathbf{C}^i(h)$, $\mathbf{D}^i$, and $\mathbf{E}^i(h)$; $\mathbf{A}^i$ and $\mathbf{D}^i$ do not depend on the initial state $x_0(h) = x_0(t_0(h))$ and need to be computed only once.
3. Select inital parameters $p^{[0]}(h)$ and an estimate of the initial Lagrange multipliers $\lambda^{[0]}(h)$ and set $k := 0$.
4. Send the control parameters $p_i^{[k]}(h)$ and the Lagrange multipliers $\lambda_i^{[k]}(h)$, $\forall i \in \{1, \ldots, N\}$, to the distributed controllers.
5. Solve the following QP on horizon $[t_0(h), t_f(h)]$ to obtain the minimizer $p_i^{[k+1]}$ and the Lagrange multiplier $\lambda_i^{[k+1]}$:

$$\min_{p_i} \quad \Phi_i^+ \text{ s.t.} \tag{11a}$$

$$\Phi_i^+ = \frac{1}{2}\tilde{p}_i^{[k]\,T}\mathbf{A}^i\tilde{p}_i^{[k]} + \tilde{p}_i^{[k]\,T}\mathbf{B}^i(h) + \mathbf{C}^i(h)$$

$$+\left[\sum_{\substack{j=1\\ j\neq i}}^{N}\left(\begin{bmatrix}\mathbf{A}_{i1}^j & \dots & \mathbf{A}_{iN}^j\end{bmatrix}p^{[k]} + \mathbf{B}_i^j(h) - \mathbf{D}_i^j(h)\lambda_j^{[k]}\right)\right]^T(p_i - p_i^{[k]}),$$

$$+\frac{1}{2}(p_i - p_i^{[k]})^T\Omega^i(p_i - p_i^{[k]}), \tag{11b}$$

$$c_i(\tilde{p}_i^{[k]}) = \mathbf{D}^{i\,T}\tilde{p}_i^{[k]} + \mathbf{E}^i(h) \geq 0, \tag{11c}$$

$$\forall i \quad \in \{1, \dots, N\}.$$

6. Increase $k := k+1$ and go back to 4.
7. Stop iteration, if $p^{[k]}$ satisfies some convergence criterion.
8. Apply the calculated optimal control inputs $u_{i,j}(t) = \sum_l p_{i,j,l}\phi_l(t)$ to the plant.
9. Set $h := h+1$, determine the new initial state $x(t_0(h))$ either from measurements or from state estimation, and go back to 2.

For the stopping criterion in step 7 the comments made on Algorithm 1 in Section 3.1 also apply.

On a new horizon, an almost optimal initialization of the method (see step 3), i.e. of the parameters $p^{[0]}(h+1)$ is possible by means of the old parameter set $p^{[l]}(h)$, as the QP change only slightly from one moving horizon to the next.

The parameters of the piecewise-constant input functions can be reinitialized as depicted in Fig. 2. Let $p_{i,j,l}$, $l = 1, \dots, \gamma_{i,j}$, represent the parameterization of the $j$-th input function $u_{i,j}(t)$ of subsystem $\Sigma_i$, where $\gamma_{i,j}$ denotes the number of parameters for input function $u_{i,j}(t)$. Then, all parameters but the first one, which has already been applied to the plant, are reused. The last parameters $p_{i,j,\gamma_{i,j}}$, $\forall i, j$, describing $u_{i,j}(t)$ are initialized to zero. The Lagrange multipliers are initialized with $\lambda_i^{[0]} = 0$, $\forall i \in \{1, \dots, N\}$, assuming all inequality constraints to be inactive. In order to ensure a feasible initialization, well-known methods such as the phase-I/phase-II approach [44] can be adapted for S-DMPC.

The system state $x(t)$ has to be measured or estimated from available measurements $y(x(t))$ in step 9 of Algorithm 6. We assume state feedback for simplicity in this work. If full state measurements are not available, the control system has to be extended by a state estimator such as the Luenberger Observer, the Kalman–Bucy-Filter, or a moving horizon estimator.

So far, closed-loop stability of S-DMPC cannot be guaranteed in the present formulation. Different strategies could be followed to extend the control problem formulation (3) for nominal closed-loop stability. Exemplarily, two alternatives are briefly mentioned:

1. Adding a proper terminal cost [51,52] to the objective function (3a), i.e.

$$\Phi = \sum_{i=1}^{N}\Phi_i(x_i, u_i) + x(t_f)^T P x(t_f). \tag{12}$$

2. Adding a terminal constraint [51,52] to the optimal control problem, e.g.

$$x(t_f) = 0. \tag{13}$$

The terminal state $x(t_f)$ can be written as a function of $p$ and hence does not complicate the decomposition problem. Both strategies have strengths and weaknesses. While in the first approach a proper matrix $P$ has to be derived during control system design using full system information, feasibility problems may arise in the second approach. A detailed consideration of formulations resulting in guaranteed stability for the distributed control problem is the subject of future work.

## 5. Simulation case study: process for the alkylation of benzene

We consider a simulated chemical process for the alkylation of benzene [36] as depicted in Fig. 3 to illustrate the performance of the S-DMPC method. The plant consists of five units, i.e. four continuous stirred-tank reactors (CSTR) and one flash separator. The purpose of the plant is to produce ethylbenzene (C) by the reaction of the raw materials benzene (A) and ethene (B). Benzene and ethene are fed into the cascaded CSTR 1, 2, and 3, where ethylbenzene is produced. In addition, the by-product diethylbenzene (D) is produced. The stream $F_7$ is fed into the flash separator, where unreacted benzene is separated from the product. The vapor stream is recycled, one part goes directly to CSTR 1, the other to CSTR 4. There, additional diethylbenzene (D) is fed and a transalkylation process leads to the reaction of benzene and diethylbenzene into ethylbenzene. The effluent of CSTR 4 is fed to the flash separator.

The mathematical model consists of material balances for each component and an energy balance for each unit of the plant. The plant behavior is described by a total of 25 differential variables. In addition, the model includes nonlinear reaction kinetics as well as a nonlinear description of the phase equilibrium in the flash separator, leading to a total of approximately 100 equations. We assume all volumetric feed flows to be constant, while the heat flows are assumed to be the manipulated variables $u_i(t)$. The model used in this work is different from the one reported in [36]. The derivation of the model and all data are available in the supplemental material (cf. Appendix C).

The goal is to stabilize the plant at its steady-state operating point $x_s$ described by the steady-state temperatures $T_{1s} = 472.32$ K, $T_{2s} = 472.35$ K, $T_{3s} = 472.39$ K, $T_{4s} = 472.00$ K, and $T_{5s} = 472.49$ K and the steady-state inputs $u_{1s} = -2.0$ MJ/s, $u_{2s} = -2.0$ MJ/s, $u_{3s} = -2.0$ MJ/s, $u_{4s} = 4.1$ MJ/s, $u_{5s} = -0.01$ MJ/s in the four CSTR (indices $i = 1, 2, 3, 5$) and the flash separator (index $i = 4$). The nonlinear model is linearized (by a finite-difference approach for convenience) at this operating point to result in the linear time-invariant model

$$\Delta\dot{x} = A\Delta x + B\Delta u, \quad \Delta x(0) = x_0 - x_s, \tag{14}$$

where $\Delta x = x - x_s$ and $\Delta u = u - u_s$ indicate the deviations of state and input variables from the set point $(x_s, u_s)$. $x_0$ indicates the initial condition of the plant. The linearized model is used as the internal model of the controller, while the nonlinear model is used to simulate the plant. For each unit $i$ of the plant, we consider the objective function

$$\Phi_i = \frac{1}{2}\int_{t_0}^{t_f}\Delta x_i^T Q_i \Delta x_i + \Delta u_i^T R_i \Delta u_i \, \mathrm{d}t \tag{15a}$$

with

$$Q_i = \mathrm{diag}[1, 1, 1, 1, 1], \tag{15b}$$

$$R_i = 1 \times 10^{-8}, \tag{15c}$$

$$\forall i \in \{1, \dots, N\}.$$

Thus, the objective function considers the deviation of the state and the inputs from the steady state values ($\Delta x_i$ and $\Delta u_i$). The inputs

$\Delta u_i(t)$ are discretized by piecewise constant base functions

$$\phi_l(t) = \begin{cases} 1, & \text{for } t_0 + \Delta t(k-1) \le t < t_0 + \Delta t\, k \\ 0, & \text{else} \end{cases} \tag{16}$$

with a horizon length of $\gamma_{i,j} = \gamma = 5$ samples, $\forall i = \{1, ..., 5\}$, $j = 1$, i.e. the scalar input of subsystem $\Sigma_i$ is represented by $\Delta u_i(t) = \sum_{l=1}^{\gamma} p_{i,l}\, \phi_l(t)$, $\forall i = 1, \ldots, 5$. The sampling time is $\Delta t = 5$ s. Finally, for the manipulated variables $u_i = u_{is} + \Delta u_i$, the following input constraints have been considered:

$$|\Delta u_1| < 0.75 \quad [\text{MJ/s}], \tag{17a}$$

$$|\Delta u_2| < 0.5 \quad [\text{MJ/s}], \tag{17b}$$

$$|\Delta u_3| < 0.5 \quad [\text{MJ/s}], \tag{17c}$$

$$|\Delta u_4| < 0.6 \quad [\text{MJ/s}], \tag{17d}$$

$$|\Delta u_5| < 0.6 \quad [\text{MJ/s}]. \tag{17e}$$

### 5.1. Convergence analysis

First the S-DMPC implementation is analyzed for convergence using the framework of Section 3.2. To this end, the weighting matrices $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$, $\mathbf{A}^4$ and $\mathbf{A}^5$ are calculated and graphically illustrated in Fig. 4. Obviously, $\mathbf{A}^1_{11}$ is the dominating part of $\mathbf{A}^1$. A similar observation can be made for the other weighting matrices. Consequently, the gain $L = 0.12 < 1$ is derived for the unconstrained case, which implies convergence of the proposed coordination mechanism.

In a second step, the numerical convergence of the method is experimentally analyzed for a set-point change, as depicted in Fig. 5. The diagram shows the evolution of the 2-norm of $p$ for an increasing number of iterations on a fixed horizon. In this example, convergence is achieved already after 3 iterations.

### 5.2. Closed-loop S-DMPC

The test cycle consists of one setpoint change, as suggested by Liu et al. [36]. The goal is to stabilize the process at the set-point given above, while the initial conditions are chosen at an operating point defined by the temperatures $T_{1,0} = 443.0$ K, $T_{2,0} = 437.1$ K, $T_{3,0} = 428.4$ K, $T_{4,0} = 433.1$ K, and $T_{5,0} = 457.6$ K. The performance of S-DMPC is compared to the performance of a centralized MPC as a reference. Fig. 6 shows the control inputs and the tracking behavior of the temperatures for the two MPC methods. The tracking behavior of the centralized controller and the S-DMPC are the same, although the S-DMPC is only implemented for $I = 1$ and $I = 2$ iterations respectively in order to limit the computing time. As a result, the control performance, described by the objective function $\Phi$, of the DMPC and the centralized MPC are the same. Hence, optimal control can be achieved by S-DMPC at a very low computational effort. In particular, for this example, and for $I = 1$, at each sample time one has to solve 5 QP involving 5 decision variables each in case of S-DMPC – possibly in parallel. In contrast, the centralized MPC solves one QP involving 25 decision variables. From one horizon to the next, the QP only vary due to the change of the initial conditions and due to external disturbances. Hence, the proper initialization of the parameters $p^{[0]}(h+1)$ using the old parameters $p^{[I]}(h)$ ensures fast convergence. For completeness, we provide the trajectories of the concentrations in Fig. 7.

Obviously, in case of this example the computational complexity is low and would not call for a distributed implementation of the predictive controller. This is different for very large-scale linear or even nonlinear problems where an analytical solution of the state equations (cf. Appendix A) is not possible anymore.

## 6. Conclusions

A novel sensitivity-driven distributed model predictive control method for linear time-invariant systems has been presented in this work. In order to achieve overall optimality, the local objective functions are extended by linear approximations of the contributions of the neighboring systems. A sufficient condition for the convergence of the method has been derived. The Wegstein method is suggested to extend the class of systems for which the S-DMPC method convergences or to improve the convergence rate. Continuous-time systems have been considered in this contribution. However, the method can easily be adapted to discrete-time systems.

A simulated case study, the control of a nonlinear chemical process has been used to illustrate the performance of the S-DMPC method. Convergence can be obtained in this case without the use of the Wegstein method; hence, no additional tuning parameters are required. Very fast convergence has been observed in this case.

Despite the very promising perspectives of the novel method, there are still a number of open questions to be addressed in the future. The analysis in this article provides a sufficient condition for convergence of the method. However, adaptive formulations are required to ensure convergence using appropriate tuning methods for the matrices $\Omega_i$. A basis for these formulations is given by Wegstein's method. Furthermore, such tuning methods have to cope with the combinatorial complexity in case inequality constraints occur. Alternatively, the convergence condition derived in this work can be used to directly derive a proper decomposition of the total process.

Future work will also concentrate on formulations which guarantee closed-loop stability of S-DMPC and which explicitly handle feasibility problems. The method can be extended to cover nonlinear systems, if an efficient implementation and a proper analysis of a nonlinear version of the method presented here can be provided. Finally, we will address large-scale applications of S-DMPC in process control.

## Acknowledgements

This research has been supported by the European 7th framework STREP project "Hierarchical and distributed model predictive control (HD-MPC)", contract number INFSO-ICT-223854. Fruitful discussions with Moritz Diehl, K.U. Leuven, as well as the comments of the reviewers are gratefully acknowledged.

## Appendix A. Control vector parameterization of the linear quadratic control problem

Without loss of generality, we assume $t_0 = 0$ as time-invariant systems are considered. The time horizon is given by $\Delta = t_f - t_0$. Further, the number of parameters (or basis functions) is assumed to be $\gamma_{i,j} = \gamma$, $\forall j \in \{1, \ldots, n_{u_i}\}$, and $\forall i \in \{1, \ldots, N\}$, for the sake of notational simplicity. Then, the objective functions $\Phi_i$ of the optimal control problem (3) are rewritten as

$$\Phi_i = \frac{1}{2} \int_0^{\Delta} x^T(\tau) \mathcal{Q}_i x(\tau) + u^T(\tau) \mathcal{R}_i u(\tau) \mathrm{d}\tau \tag{A.1a}$$

with

$$\mathcal{Q}_i = \mathrm{diag}([0, \ldots, 0, Q_i, 0, \ldots, 0]), \tag{A.1b}$$

$$\mathcal{R}_i = \mathrm{diag}([0, \ldots, 0, R_i, 0, \ldots, 0]), \tag{A.1c}$$

$\forall i \quad \in \{1, \ldots, N\}$.

In the following, the input vectors $u_i(t)$, $\forall i \in \{1, \ldots, N\}$, are assumed to be parameterized by means of piecewise constant functions or (or B-splines of order 1). Similar results can be derived for other parameterizations such as higher order B-Splines [42] and in particular piecewise linear functions [40].

Hence, the scalar input $u_{i,j}(t)$ to subsystem $\Sigma_i$ is a function of the parameter vector $p_{i,j}$. In particular, for zero-order hold discretizations, we get

$$u_{i,j}(t) = p_{i,j,\eta}, \quad \text{for } t_\eta \le t < t_{\eta+1}, \tag{A.2}$$

with $t_{\eta+1} = t_\eta + \Delta_\eta$, $\Delta_\eta \neq const.$, and $\Delta = \sum_{\eta=1}^{\gamma} \Delta_\eta$. This discretization results in the discrete-time system [40]:

$$x_{\cdot|\eta+1} = A_\eta^0 x_{\cdot|\eta} + B_\eta^0 p_{\cdot|\eta}, \tag{A.3a}$$

$$A_\eta^0 = e^{A\Delta_\eta}, \tag{A.3b}$$

$$B_\eta^0 = I_0(\Delta_\eta)\, B, \tag{A.3c}$$

$$I_0(t) = \int_0^t e^{A\tau}\, \mathrm{d}\tau. \tag{A.3d}$$

$x_{\cdot|\eta} = \langle x_{1|\eta}, \ldots, x_{N|\eta}\rangle$ represents the full discrete-time state at sample $\eta$. $p_{\cdot|\eta} = \langle p_{1\cdot|\eta}, \ldots, p_{N\cdot|\eta}\rangle$ with $p_{i\cdot|\eta} = \langle p_{i,1,\eta}, \ldots, p_{i,u_i,\eta}\rangle$, $\forall i \in \{1, \ldots, N\}$, indicates an aggregated parameter vector. Applying this discretization, the continuous-time cost function in Eq. (A.1) can be expressed as a finite sum of the control parameters $p_{\cdot|\eta}$ and the discrete-time state variables $x_{\cdot|\eta}$ [40,53,54]:

$$\begin{aligned}\Phi_i &= \frac{1}{2}\int_0^{\Delta} x^T(\tau)\mathcal{Q}_i x(\tau) + u^T(\tau)\mathcal{R}_i u(\tau)\mathrm{d}\tau \\ &= \frac{1}{2}\sum_{\eta=0}^{\gamma-1} x_{\cdot|\eta}^T \mathcal{Q}_{i,\eta}^0 x_{\cdot|\eta} + p_{\cdot|\eta}^T \mathcal{R}_{i,\eta}^0 p_{\cdot|\eta} + 2x_{\cdot|\eta}^T \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta}, \qquad \text{with}\end{aligned} \tag{A.4a}$$

$$\mathcal{Q}_{i,\eta}^0 = \int_0^{\Delta_\eta} (e^{At})^T \mathcal{Q}_i e^{At}\, \mathrm{d}t, \tag{A.4b}$$

$$\mathcal{R}_{i,\eta}^0 = \int_0^{\Delta_\eta} (\mathcal{R}_i + (I_0 B)^T \mathcal{Q}_i I_0 B)\, \mathrm{d}t, \tag{A.4c}$$

$$\mathcal{S}_{i,\eta}^0 = \int_0^{\Delta_\eta} (e^{At})^T \mathcal{Q}_i I_0 B\, \mathrm{d}t. \tag{A.4d}$$

For the elimination of the state variables $x_{\cdot|\eta}$ in the objective function (A.4a), they have to be expressed as functions of $x_0$ and $p_{\cdot|\eta}$. The solution of the time-varying discrete-time system (A.3a) is

$$x_{\cdot|\eta} = \underbrace{\left(\prod_{i=0}^{\eta-1} A_i^0\right)}_{=:\Xi_0^\eta} x_0 + \sum_{i=0}^{\eta-1} \underbrace{\left(\prod_{j=i+1}^{\eta-1} A_j^0\right)}_{=:\Xi_{i+1}^\eta} B_i^0 p_{\cdot|i} \tag{A.5}$$

With this result, $x_{\cdot|\eta}^T \mathcal{Q}_{i,\eta}^0 x_{\cdot|\eta}$ and $2x_{\cdot|\eta}^T \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta}$ in Eq. (A.4a) can be expressed as functions of $p_{\cdot|\eta}$:

$$\begin{aligned} x_{\cdot|\eta}^T \mathcal{Q}_{i,\eta}^0 x_{\cdot|\eta} &= x_0^T (\Xi_0^\eta)^T \mathcal{Q}_{i,\eta}^0 \Xi_0^\eta x_0 + 2x_0^T (\Xi_0^\eta)^T \mathcal{Q}_{i,\eta}^0 \sum_{q=0}^{\eta-1} \left(\Xi_{q+1}^\eta B_q^0 p_{\cdot|q}\right) \\ &\quad + \sum_{q=0}^{\eta-1}\sum_{j=0}^{\eta-1} \left(p_{\cdot|j}^T (B_j^0)^T (\Xi_{j+1}^\eta)^T \mathcal{Q}_{i,\eta}^0 \Xi_{q+1}^\eta B_q^0 p_{\cdot|q}\right), \\ 2x_{\cdot|\eta}^T \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta} &= 2\Xi_0^\eta x_0 \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta} + 2\sum_{q=0}^{\eta-1} p_{\cdot|q}^T (B_q^0)^T (\Xi_{q+1}^\eta)^T \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta}. \end{aligned}$$

The quadratic objective function can then be rewritten as a function of $x_0$ and the control parameters $p$:

$$\begin{aligned}\Phi_i(p) &= \frac{1}{2}\sum_{\eta=0}^{\gamma-1} \Big[ x_0^T (\Xi_0^\eta)^T \mathcal{Q}_{i,\eta}^0 \Xi_0^\eta x_0 \\ &\quad + 2\Xi_0^\eta x_0 \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta} + 2x_0^T (\Xi_0^\eta)^T \mathcal{Q}_{i,\eta}^0 \sum_{q=0}^{\eta-1} \left(\Xi_{q+1}^\eta B_q^0 p_{\cdot|q}\right) \\ &\quad + p_{\cdot|\eta}^T \mathcal{R}_{i,\eta}^0 p_{\cdot|\eta} + 2\sum_{q=0}^{\eta-1} p_{\cdot|q}^T (B_q^0)^T (\Xi_{q+1}^\eta)^T \mathcal{S}_{i,\eta}^0 p_{\cdot|\eta} \\ &\quad + \sum_{q=0}^{\eta-1}\sum_{j=0}^{\eta-1} \left(p_{\cdot|j}^T (B_j^0)^T (\Xi_{j+1}^\eta)^T \mathcal{Q}_{i,\eta}^0 \Xi_{q+1}^\eta B_q^0 p_{\cdot|q}\right) \Big]\end{aligned} \tag{A.6}$$

$$\Phi_i(p) = \frac{1}{2} p^T \mathbf{A}^i p + p^T \mathbf{B}^i + \mathbf{C}^i. \tag{A.7}$$

This derivation proves Eq. (4). Note that the local subproblems depend on $p$ and therefore on the parameterizations of all the control vectors despite the sparsity of the matrices $Q_i$ and $R_i$ in Eqs. (A.1b) and (A.1c).

Likewise, the linear constraints (2), can be transcribed to result in the linear inequality constraints

$$\begin{aligned} &D_i \langle x_{\cdot|\eta}, p_{\cdot|\eta}\rangle + e_i \\ &= D_i \begin{bmatrix} \left(\Xi_0^\eta x_0 + \sum_{i=0}^{\eta-1} \Xi_{i+1}^\eta B_i^0 p_{\cdot|i}\right)^T & p_{\cdot|\eta}^T \end{bmatrix}^T + e_i \\ &= \mathbf{D}^i p + \mathbf{E}^i \ge 0, \quad \forall i \in \{1, \ldots, N\}. \end{aligned} \tag{A.8}$$

## Appendix B. Proof of Theorem 5

The Lagrange functions corresponding to the QP subproblems (8) are

$$\mathcal{L}_{\mathrm{dist},i} = \Phi_i^* - \lambda_i^T c_i(p), \quad \forall i \in \{1, \ldots, N\}, \tag{B.1}$$

Hence, at iteration $k+1$, the minimizers of these problem are given by their necessary conditions of optimality (NCO):

$$\begin{aligned} 0 = \frac{\partial \mathcal{L}_{\mathrm{dist},i}}{\partial p_i} &= \frac{\partial \Phi_i^*}{\partial p_i} - \frac{\partial c_i}{\partial p_i}\lambda_i^{[k+1]} \\ &= \frac{\partial \Phi_i}{\partial p_i} + \Omega^i (p_i^{[k+1]} - p_i^{[k]}) + \sum_{\substack{j=1 \\ j\neq i}}^{N} \left( \frac{\partial \Phi_j}{\partial p_i}\bigg|_{p^{[k]}} - \frac{\partial c_j}{\partial p_i}\bigg|_{p^{[k]}} \lambda_j^{[k]} \right) \\ &\quad - \frac{\partial c_i}{\partial p_i}\lambda_i^{[k+1]} \end{aligned}$$

$$= \left( \begin{bmatrix} \mathbf{A}^i_{i1} & \dots & \mathbf{A}^i_{iN} \end{bmatrix} \tilde{p}_i^{[k+1]} + \mathbf{B}^i_i \right) + \sum_{\substack{j=1 \\ j \neq i}}^{N} \left( \begin{bmatrix} \mathbf{A}^j_{i1} & \dots & \mathbf{A}^j_{iN} \end{bmatrix} p^{[k]} + \mathbf{B}^j_i \right)$$

$$+\Omega^i (p_i^{[k+1]} - p_i^{[k]}) - \sum_{\substack{j=1 \\ j \neq i}}^{N} \left( \mathbf{D}^j_i \lambda_j^{[k]} \right) - \mathbf{D}^i_i \lambda_i^{[k+1]}, \tag{B.2a}$$

and

$$0 = \left[ \mathbf{D}^{i\,T} \tilde{p}_i^{[k+1]} + \mathbf{E}^i \right]_{\mathcal{A}}, \tag{B.2b}$$

$$0 = \lambda_i|_{\mathcal{I}} \tag{B.2c}$$

$\forall i \in \{1, \dots, N\}$, where $\tilde{p}_i^{[k+1]} = \langle p_1^{[k]} \cdots p_{i-1}^{[k]} \; p_i^{[k+1]} \, p_{i+1}^{[k]} \cdots p_N^{[k]} \rangle$. This notation means, that within subproblem $i$ only the parameters $p_i$ can be varied to a new parameter set $p_i^{[k+1]}$. All the other parameters are fixed and remain at the values of the previous iteration $k$. In the following Eqs. (B.3)–(B.13), we only consider the active constraints and omit the reference to the index set $\mathcal{A}$ for simplicity of notation. Eq. (B.2c) do not have to be reformulated and will be used later on. We rewrite the expressions

$$\begin{bmatrix} \mathbf{A}^i_{i1} & \dots & \mathbf{A}^i_{iN} \end{bmatrix} \tilde{p}_i^{[k+1]} = \mathbf{A}^i_{i,i} p_i^{[k+1]} + \begin{bmatrix} \mathbf{A}^i_{i1} & \dots & \mathbf{A}^i_{iN} \end{bmatrix} p^{[k]} - \mathbf{A}^i_{i,i} p_i^{[k]}, \quad \forall i, \tag{B.3}$$

$$\mathbf{D}^{i\,T} \tilde{p}_i^{[k+1]} = \mathbf{D}^{i\,T}_i p_i^{[k+1]} + \mathbf{D}^{i\,T} p^{[k]} - \mathbf{D}^{i\,T}_i p_i^{[k]}, \quad \forall i, \tag{B.4}$$

$$- \sum_{\substack{j=1 \\ j \neq i}}^{N} \left( \mathbf{D}^j_i \lambda_j^{[k]} \right) = - \sum_{j=1}^{N} \left( \mathbf{D}^j_i \lambda_j^{[k]} \right) + \mathbf{D}^i_i \lambda_i^{[k]}$$

$$= -\mathbf{D}_i \lambda^{[k]} + \mathbf{D}^i_i \lambda_i^{[k]}, \quad \forall i, \tag{B.5}$$

with

$$\mathbf{D}_i = \begin{bmatrix} \mathbf{D}^1_i & \dots & \mathbf{D}^N_i \end{bmatrix}, \quad \text{and} \quad \lambda^{[k]} = \langle \lambda_1^{[k]}, \dots, \lambda_N^{[k]} \rangle, \tag{B.6}$$

and simplify the NCO by resorting as follows:

$$\begin{cases} 0 &= \left( \mathbf{A}^i_{i,i} p_i^{[k+1]} + \left( \mathbf{A}^i_{i1} \; \dots \; \mathbf{A}^i_{iN} \right) p^{[k]} - \mathbf{A}^i_{i,i} p_i^{[k]} + \mathbf{B}^i_i \right) \\ & \quad + \sum_{\substack{j=1 \\ j \neq i}}^{N} \left( \begin{bmatrix} \mathbf{A}^j_{i1} & \dots & \mathbf{A}^j_{iN} \end{bmatrix} p^{[k]} + \mathbf{B}^j_i \right) + \Omega^i (p_i^{[k+1]} - p_i^{[k]}) \\ & \quad - \mathbf{D}_i \lambda^{[k]} + \mathbf{D}^i_i \lambda_i^{[k]} - \mathbf{D}^i_i \lambda_i^{[k+1]} \\ 0 &= \mathbf{D}^{i\,T}_i p_i^{[k+1]} + \mathbf{D}^{i\,T} p^{[k]} - \mathbf{D}^{i\,T}_i p_i^{[k]} + \mathbf{E}^i \end{cases}, \quad \forall i \tag{B.7}$$

$\Leftrightarrow$

$$\begin{cases} 0 &= \left( \mathbf{A}^i_{i,i} + \Omega^i \right) (p_i^{[k+1]} - p_i^{[k]}) \\ & \quad + \sum_{j=1}^{N} \left( \begin{bmatrix} \mathbf{A}^j_{i1} & \dots & \mathbf{A}^j_{iN} \end{bmatrix} p^{[k]} + \mathbf{B}^j_i \right) \\ & \quad - \mathbf{D}_i \lambda^{[k]} + \mathbf{D}^i_i \lambda_i^{[k]} - \mathbf{D}^i_i \lambda_i^{[k+1]} \\ 0 &= \mathbf{D}^{i\,T}_i p_i^{[k+1]} + \mathbf{D}^{i\,T} p^{[k]} - \mathbf{D}^{i\,T}_i p_i^{[k]} + \mathbf{E}^i \end{cases}, \quad \forall i \tag{B.8}$$

$\Leftrightarrow$

$$\begin{cases} 0 &= \begin{bmatrix} \mathbf{A}^i_{i,i} + \Omega^i & -\mathbf{D}^i_i \\ \mathbf{D}^{i\,T}_i & 0 \end{bmatrix} \begin{bmatrix} p_i^{[k+1]} \\ \lambda_i^{[k+1]} \end{bmatrix} \\ & \quad + \begin{bmatrix} \sum_{j=1}^{N} \begin{bmatrix} \mathbf{A}^j_{i1} & \dots & \mathbf{A}^j_{iN} \end{bmatrix} & -\mathbf{D}_i \\ \mathbf{D}^{i\,T} & 0 \end{bmatrix} \begin{bmatrix} p^{[k]} \\ \lambda^{[k]} \end{bmatrix} \\ & \quad + \begin{bmatrix} -\mathbf{A}^i_{i,i} - \Omega^i & \mathbf{D}^i_i \\ -\mathbf{D}^{i\,T}_i & 0 \end{bmatrix} \begin{bmatrix} p_i^{[k]} \\ \lambda_i^{[k]} \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^{N} \mathbf{B}^j_i \\ \mathbf{E}^i \end{bmatrix} \end{cases}, \quad \forall i. \tag{B.9}$$

We compose the full mapping of all subsystems, based on Eq. (B.9):

$$\begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^T_{\text{diag}} & 0 \end{bmatrix} \begin{bmatrix} p^{[k+1]} \\ \lambda^{[k+1]} \end{bmatrix} = - \begin{bmatrix} \sum_{j=1}^{N} \mathbf{A}^j & -\mathbf{D} \\ -\mathbf{D}^T & 0 \end{bmatrix} \begin{bmatrix} p^{[k]} \\ \lambda^{[k]} \end{bmatrix} + \begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^T_{\text{diag}} & 0 \end{bmatrix} \begin{bmatrix} p^{[k]} \\ \lambda^{[k]} \end{bmatrix} - \begin{bmatrix} \sum_{j=1}^{N} \mathbf{B}^j \\ -\mathbf{E} \end{bmatrix} \tag{B.10a}$$

with

$$\mathbf{A}^{\Omega}_{\text{diag}} = \mathbf{A}_{\text{diag}} + \Omega, \tag{B.10b}$$

$$\mathbf{A}_{\text{diag}} = \text{diag}(\mathbf{A}^1_{11}, \dots, \mathbf{A}^N_{NN}), \tag{B.10c}$$

$$\Omega = \text{diag}(\Omega^1, \dots, \Omega^N), \tag{B.10d}$$

$$\mathbf{D}_{\text{diag}} = \text{diag}(\mathbf{D}^1_1, \dots, \mathbf{D}^N_N), \tag{B.10e}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \vdots \\ \mathbf{D}_N \end{bmatrix} = \begin{bmatrix} \mathbf{D}^1_1 & \dots & \mathbf{D}^N_1 \\ \vdots & & \vdots \\ \mathbf{D}^1_N & \dots & \mathbf{D}^N_N \end{bmatrix} = \begin{bmatrix} \mathbf{D}^1 & \dots & \mathbf{D}^N \end{bmatrix}, \tag{B.10f}$$

$$\mathbf{E} = \langle \mathbf{E}^1, \dots, \mathbf{E}^N \rangle. \tag{B.10g}$$

In order to derive an explicit formulation for the updated control and Lagrange parameters $p^{[k+1]}$ and $\lambda^{[k+1]}$, we use the following

**Lemma 7.** *The matrix*

$$\begin{bmatrix} A^{\Omega}_{\text{diag}} & -D_{\text{diag}} \\ -D^T_{\text{diag}} & 0 \end{bmatrix} \tag{B.11}$$

*is regular under Assumptions 3 and 4. Thus it is invertible.*

**Proof.** Using the Schur complement, we can derive the determinant of matrix (B.11) as

$$\begin{vmatrix} A^{\Omega}_{\text{diag}} & -D_{\text{diag}} \\ -D^T_{\text{diag}} & 0 \end{vmatrix} = |A^{\Omega}_{\text{diag}}| \cdot |D^T_{\text{diag}} A^{\Omega -1}_{\text{diag}} D_{\text{diag}}|. \tag{B.12}$$

As $A^i_{ii}$ is positive definite for all $i$, the same holds for $A^i_{ii} + \Omega^i$, $A^{\Omega}_{\text{diag}}$ and its inverse matrix $A^{\Omega\,-1}_{\text{diag}}$. As a consequence $|A^{\Omega}_{\text{diag}}| \neq 0$. Additionally, as $D_{\text{diag}}$ has full rank and $A^{\Omega\,-1}_{\text{diag}}$ is positive definite, $|D^T_{\text{diag}} A^{\Omega\,-1}_{\text{diag}} D_{\text{diag}}| \neq 0$. Hence, the determinant of matrix (B.11) is non-zero, and the matrix is regular and invertible. □

Based on the results of Lemma B.1, we can solve equation (B.10a) for $\langle p^{[k+1]}, \lambda^{[k+1]} \rangle$:

$$\begin{bmatrix} p^{[k+1]} \\ \lambda^{[k+1]} \end{bmatrix} = -\begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^{T}_{\text{diag}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{N} \mathbf{A}^{j} & -\mathbf{D} \\ -\mathbf{D}^{T} & 0 \end{bmatrix} \begin{bmatrix} p^{[k]} \\ \lambda^{[k]} \end{bmatrix} + \begin{bmatrix} p^{[k]} \\ \lambda^{[k]} \end{bmatrix} - \begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^{T}_{\text{diag}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{N} \mathbf{B}^{j} \\ -\mathbf{E} \end{bmatrix} \tag{B.13}$$

This is a mapping for all control parameters and the Lagrange parameters of the active constraints during the course of the iteration. In order to derive the full mapping, we add the equations for the inactive constraints and get

$$\begin{bmatrix} p^{[k+1]} \\ \lambda_{\mathcal{A}}^{[k+1]} \\ \lambda_{\mathcal{I}}^{[k+1]} \end{bmatrix} = \underbrace{\begin{bmatrix} G_{\mathcal{A}}^{[k+1]} & 0 \\ 0 & 0 \end{bmatrix}}_{=G^{[k+1]}} \begin{bmatrix} p^{[k]} \\ \lambda_{\mathcal{A}}^{[k]} \\ \lambda_{\mathcal{I}}^{[k]} \end{bmatrix} + \begin{bmatrix} U^{[k+1]} \\ 0 \end{bmatrix} \tag{B.14}$$

with

$$G_{\mathcal{A}}^{[k+1]} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^{T}_{\text{diag}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{N} \mathbf{A}^{j} & -\mathbf{D} \\ -\mathbf{D}^{T} & 0 \end{bmatrix} \Bigg|_{\mathcal{A}^{[k+1]}} \tag{B.15}$$

$$U^{[k+1]} = -\begin{bmatrix} \mathbf{A}^{\Omega}_{\text{diag}} & -\mathbf{D}_{\text{diag}} \\ -\mathbf{D}^{T}_{\text{diag}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{N} \mathbf{B}^{j} \\ -\mathbf{E} \end{bmatrix} \Bigg|_{\mathcal{A}^{[k+1]}} \tag{B.16}$$

In Eq. (B.14), the mapping is described by its gain $G^{[k+1]}$. According to the small gain theorem [55], the mapping is convergent if $\left\| G^{[k+1]} \right\| < 1$, which simplifies to $\left\| G_{\mathcal{A}}^{[k+1]} \right\| < 1$. Hence, we have shown the result of Theorem 5.

## Appendix C. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jprocont.2011.01.013.

## References


[1] M. Kano, M. Ogawa, The state of the art in chemical process control in Japan: good practice and questionnaire survey, Journal of Process Control 20 (9) (2010) 969–982.
[2] S.J. Qin, T.A. Badgwell, A survey of industrial model predictive control technology, Control Engineering Practice 11 (7) (2003) 733–764.
[3] W. Al-Gherwi, H. Budman, A. Elkamel, An online algorithm for robust distributed model predictive control, in: Proceedings of ADCHEM 2009 (International Symposium on Advanced Control of Chemical Processes), vol. 7, Istanbul, Turkey, 2009.
[4] G. Pannocchia, S.J. Wright, B.T. Stewart, J.B. Rawlings, Efficient cooperative distributed MPC using partial enumeration, in: Proceedings of ADCHEM 2009 (International Symposium on Advanced Control of Chemical Processes), Istanbul, Turkey, 2009.
[5] F. Borrelli, T. Keviczky, K. Fregene, G.J. Balas, Decentralized Receding Horizon Control of Cooperative Vehicle Formations, in: Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005, Seville, Spain, 2005, pp. 3955–3960.
[6] G. Lionis, K.J. Kyriakopoulos, Approximate control of formations of multiagent systems, in: Proceedings of the 44th Conference on Decision and Control, and the European Control Conference 2005, Seville, Spain, 2005, pp. 4958–4963.
[7] P. Massioni, M. Verhaegen, Distributed Control of Vehicle Formations: a Decomposition Approach, in: Proceedings of the 47th IEEE Conference on Decision and Control 2008, Cancun, Mexico, 2008, pp. 2906–2912.
[8] P. Wieland, J.-S. Kim, H. Scheu, F. Allgöwer, On consensus in multi-agent systems with linear high-order agents, in: Proceedings of the 17th IFAC World Congress, vol. 17, 2008, pp. 1541–1546.
[9] P. Wieland, C. Ebenbauer, F. Allgöwer, Ensuring task-independent safety for multi-agent systems by feedback, in: Proceedings of the American Control Conference, 2007, pp. 3880–3885.
[10] J.A. Fax, R.M. Murray, Information flow and cooperative control of vehicle formations, IEEE Transactions on Automatic Control 49 (2004) 1465–1476.
[11] J.H. Seo, H. Shim, J. Back, Consensus of high-order systems using dynamic output feedback compensator: low gain approach, Automatica 45 (2009) 2659–2664.
[12] L. Moreau, Stability of muliagent systems with time-dependent communication links, IEEE Transactions on Automatic Control 50 (2005) 169–182.
[13] F. Xiao, L. Wang, Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays, IEEE Transactions on Automatic Control 53 (2008) 1804–1816.
[14] J.B. Rawlings, B.T. Stewart, Coordinating multiple optimization-based controllers: new opportunities and challenges, Journal of Process Control 18 (2008) 839–845.
[15] M.D. Mesarovic, D. Macko, Y. Takahara, Theory of Hierarchical, Academic Press, Multilevel, Systems, 1970.
[16] M.D. Mesarovic, D. Macko, Y. Takahara, Two coordination principles and their application in large scale systems control, Automatica 6 (1970) 261–270.
[17] A.M. Geoffrion, Primal resource-directive approach for optimizing nonlinear decomposable systems, Operations Research 18 (3) (1970) 375–403.
[18] G.J. Silverman, Primal decomposition of mathematical programs by resource allocation: I-basic theory and a direction-finding procedure, Operations Research 20 (1) (1972) 58–74.
[19] L.S. Lasdon, Optimization Theory for Large Systems, Macmillan Series for Operations Research, 1970.
[20] R. Scattolini, Architectures for distributed and hierarchical model predictive control—a review, Journal of Process Control 19 (2009) 723–731.
[21] M.G. Singh, S.A.W. Drew, J.F. Coales, Comparisons of practical hierarchical control methods for interconnected dynamical systems, Automatica 11 (1975) 331–350.
[22] H. Tamura, Decomposition techniques in large scale systems, Systems and Control 17 (6) (1973).
[23] N.J. Smith, A.P. Sage, An introduction to hierarchical systems theory, Computers & Electrical Engineering 1 (1973) 55–71.
[24] Y. Takahara, A multilevel structure for a class of dynamic optimisation problems, Master's thesis, Engineering Division, Case Western Reserve University, Cleveland, OH, 1965.
[25] M.G. Singh, A new algorithm for the on-line multi-level control of large interconnected systems with fast dynamics, International Journal of Control 21 (1975) 587–597.
[26] M.S. Mahmoud, Multilevel systems control and applications: A survey, IEEE Transactions on Systems, Man and Cybernetics SMC-7 (3) (1977) 125–143.
[27] N.R. Sandell, P. Varaiya, M. Athans, M.G. Safonov, Survey of decentralized control methods for large scale systems, IEEE Transactions on Automatic Control 23 (2) (1978) 108–128.
[28] Y. Wakasa, M. Arakawa, K. Tanaka, T. Akashi, Decentralized model predictive control via dual decomposition, in: Proceedings of the 47th IEEE Conference on Decision and Control 2008, Cancun, Mexico, 2008, pp. 381–386.
[29] I. Necoara, J.A. Suykens, A proximal center-based decomposition method for multi-agent convex optimization, in: Proceedings of the 47th IEEE Conference on Decision and Control 2008, Cancun, Mexico, 2008, pp. 3077–3082.
[30] I. Necoara, D. Doan, J.A. Suykens, Application of the proximal center decomposition method to distributed model predictive control, in: Proceedings of the 47th IEEE Conference on Decision and Control 2008, Cancun, Mexico, 2008, pp. 2900–2905.
[31] A.N. Venkat, J.B. Rawlings, S.J. Wright, Plant-wide optimal control with decentralized MPC, in: Proceedings of DYCOPS, Boston, MA, USA, 2004.
[32] B.T. Stewart, A.N. Venkat, J.B. Rawlings, S.J. Wright, G. Pannocchia, Cooperative distributed model predictive control, Systems & Control Letters 59 (8) (2010) 460–469.
[33] N.I. Marcos, J.F. Forbes, M. Guay, Coordination of distributed model predictive controllers for constrained dynamic processes, in: Proceedings of ADCHEM (International Symposium on Advanced Control of Chemical Processes), Istanbul, Turkey, 2009.
[34] S. Talukdar, D. Jia, P. Hines, B.H. Krogh, Distributed model predictive control for the mitigation of cascading failure, in: Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005, Sevilla, Spain, 2005, pp. 4440–4445.
[35] J. Liu, D. Muñoz de la Peña, P.D. Christofides, Distributed model predictive control of nonlinear process systems, AIChE Journal 55 (5) (2009) 1171–1184.
[36] J. Liu, X. Chen, D. Muñoz de la Peña, P.D. Christofides, Sequential and iterative architectures for distributed model predictive control of nonlinear process systems, AIChE Journal 56 (8) (2010) 2137–2149.
[37] W.B. Dunbar, Distributed receding horizon control of dynamically coupled nonlinear systems, IEEE Transactions on Automatic Control 52 (7) (2007) 1249–1263.
[38] L. Magni, R. Scattolini, Stabilizing decentralized model predictive control of nonlinear systems, Automatica 42 (2006) 1231–1236.
[39] H. Scheu, J. Busch, W. Marquardt, Nonlinear distributed dynamic optimization based on first order sensitivities, in: Proceedings of the American Control Conference 2010, Baltimore, MA, USA, 2010, pp. 1574–1579.

[40] G. Pannocchia, J. Rawlings, D. Mayne, W. Marquardt, On computing solutions to the continuous time constrained linear quadratic regulator, Automatic Control, IEEE Transactions on 55 (9) (2010) 2192–2198.
[41] R. Sargent, G. Sullivan, The development of an efficient optimal control package, in: J. Stoer (Ed.), Proceedings of the 8th IFIP Conference on Optimization Technique, Springer, 1978, pp. 158–168.
[42] C. de Boor, A Practical Guide to Splines, Springer, New York, 1978.
[43] M. Schlegel, K. Stockmann, T. Binder, W. Marquardt, Dynamic optimization using adaptive control vector parameterization, Computers & Chemical Engineering 29 (8) (2005) 1731–1751.
[44] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 2000.
[45] L.S. Lasdon, Duality and decomposition in mathematical programming, IEEE Transactions on Systems Science and Cybernetics 4 (2) (1968) 86–100.
[46] J.H. Wegstein, Accelerating convergence of iterative processes, Communications of the ACM 1 (6) (1958) 9–13.
[47] A.W. Westerberg, H.P. Hutchison, R.L. Motard, P. Winter, Process Flowsheeting, Cambridge University Press, 1979.
[48] G.J. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Mathematics Journal 29 (1962) 341–346.
[49] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM Journal on Control and Optimization 14 (1976) 877–898.
[50] Y. Censor, Proximal minimization algorithm with D-functions, Journal of Optimization Theory and Applications 73 (3) (1992) 451–464.
[51] D.Q. Mayne, J.B. Rawlings, C.V. Rao, P.O.M. Scokaert, Constrained model predictive control: stability and optimality, Automatica 36 (2000) 789–814.
[52] G.D. Nicolao, L. Magni, R. Scattolini, Stability and robustness of nonlinear receding horizon control, in: F. Allgöwer, A. Zheng (Eds.), Nonlinear Model Predictive Control, Birkhäuser, 2000, pp. 3–22.
[53] H. Kwakernaak, R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, 1972.
[54] J. Yuz, G. Goodwin, A. Feuer, J. De Doná, Control of constrained linear systems using fast sampling rates, Systems & Control Letters 54 (2005) 981–990.
[55] H.K. Khalil, Nonlinear Systems, Prentice Hall, 2000.