This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A Hierarchical MPC Scheme for Coordination of Independent Systems With Shared Resources and Plug-and-Play Capabilities

Marcello Farina, Xinglong Zhang, and Riccardo Scattolini

***Abstract*— This paper describes a hierarchical control scheme for the coordination of independent, asymptotically stable, and square systems with joint input and output constraints. At the higher layer of the control structure, a centralized model predictive control (MPC) problem is solved at a slow rate based on a reduced order model of the overall system to fulfill the global output request under shared resources. At the lower layer, fast MPC controllers are designed for each subsystem to remove the effects of the model mismatch introduced at the higher layer, satisfy local constraints and optimize the individual performance. The proposed algorithm allows to dynamically modify the control configuration according to plug-and-play operations. Theoretical results of feasibility and convergence are proven and a simulation study is reported to witness the potentialities of the method.**

***Index Terms*— Coordination, hierarchical control, model predictive control (MPC), plug-and-play control.**

I. Introduction

In many industrial and civil contexts, there is the need to coordinate a number of independent subsystems with limited resources to guarantee a given behavior of the overall system. Examples can be found in buildings, where different thermal power generators must be controlled to provide the required cooling or heating and to minimize an economic cost (see [15]). In [11], the considered problem consists in coordinating different oxygen generators available in a distribution network to satisfy a given request under shared resources. In isolated microgrids, possibly including renewable energy sources, the available dispatchable diesel generators must be managed to satisfy the load request and eventually provide frequency and voltage regulation (see [2], [14]). Other examples, concerning irrigation systems and chemical processes, are discussed in [1]. In all of these problems, it can also be useful to allow for plug-and-play operations to improve the economic performance and guarantee better flexibility, adaptation to changing conditions, and fault tolerance (see [18], [20]).

A simple approach to the solution of these coordination problems consists in considering a unique model of the entire system and to design a centralized regulator. However, this could be computationally expensive when the overall system is large-scale, as recognized in [1], since the computational complexity would scale with the number of subsystems. For this reason, hierarchical control schemes have widely been used in many industrial control problems. The general idea is to design a centralized high-level regulator for a simplified model with the scope to consider the long-term behavior of the system, while local regulators are used at the lower levels for stability and disturbance rejection. Research in hierarchical control can be traced back to the early contributions of [3] and [13] and has received a renewed interest in the design of multilayer model predictive control (MPC) systems, see the significant contributions of [6] and [10] in the field of plantwide control, the cascade MPC design methods reported in [16] and [21], and the survey [19] on distributed and hierarchical control. Among the many recent contributions to the design of multilayer control systems, in [1], a three-level control scheme was designed: at the higher layer, a simplified centralized regulator was used, while the concept of aggregators, each one containing a number of subsystems, was introduced at the intermediate layer to simplify the size of the problem; finally, at the lower layer, autonomous regulators were used. By resorting to the definition of aggregators, a distributed algorithm was presented in [5] for the power balancing of smart grids with flexible consumers. Finally, the management of flexible thermal loads was considered in [4], where at high level, a centralized regulator was designed for an estimated low-order ARX model, while local linear controllers were used at the low level.

In this scenario, this paper presents a novel hierarchical control algorithm for the coordination of independent asymptotically stable (or prestabilized), square systems with possibly joint input and output constraints. At the higher layer of the control structure, a centralized MPC problem is solved at a slow rate and considering an average model of the overall system having the same size of the individual sybsystems, which computes the input to the independent subsystems so as to fulfill the global output request. At the lower layer, a decentralized scheme is proposed: indeed, for any subsystem,

Manuscript received February 27, 2018; revised July 23, 2018; accepted November 6, 2018. Manuscript received in final form November 27, 2018. The work of X. Zhang was supported by the China Scholarship Council under Grant 201403170238. Recommended by Associate Editor G. Papafotiou. *(Corresponding author: Marcello Farina.)*

The authors are with the Dipartimento di Elettronica, Informazione, e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: marcello.farina@polimi.it; xinglong.zhang@polimi.it; riccardo.scattolini@polimi.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2018.2884243

a local fast MPC controller is designed to remove the effects of the model mismatch introduced at the higher layer, to satisfy local constraints and to optimize the individual performance. The proposed design method allows to modify the system configuration, in terms of the contribution provided by any subsystem to the overall system performance, and to implement plug-and-play operations. The recursive feasibility of the MPC problems to be solved at the high and low levels is guaranteed also during plug-in and plug-out operations, and overall convergence of the system output to the setpoint is proven.

The main advantages with respect to a centralized control scheme are its scalability and flexibility. Concerning scalability, in this setup, the number of subsystems to be coordinated can be arbitrarily large, without affecting the computational complexity of the high-layer problem, to be solved at a centralized level; also, at the low level, each subproblem is solved locally. Concerning flexibility, the proposed scheme is able to account, automatically and in a flexible and reliable way, for structural system changes and variations in the control setup, e.g., plug and play requests.

The adopted control structure is similar to the one considered in [17], but the solution proposed here represents a significant improvement for the following reasons: 1) the problem at the high level is fully scalable with the number of subsystems, so allowing for plug-and-play operations; 2) the high-level model is easily determined from the impulse responses of the subsystems; 3) constraints on the shared resources (inputs) are included; and 4) the possibility to perform static high-level optimization is explicitly considered to optimize the subsystems' usage and provide flexibility to the control configuration.

This paper is organized as follows. In Section II, the problem is stated; also, the system model is introduced, together with a sketch of the overall controller structure. The design of the high and low-level MPC regulators is presented in Section III where the main theoretical results of recursive feasibility and convergence are given. The static optimization problem to be solved for the definition of the optimal usage of the subsystems is presented in Section IV together with the plug-and-play procedure to be followed to guarantee the properties of the system. A simulation study is discussed in Section V to illustrate the performance of the method, while some conclusions are drawn in Section VI. The proof of the main result is collected in the Appendix.

Notation: For a given set of variables $z_i \in \mathbb{R}^{q_i}$, $i = 1, 2, \ldots, M$, we define the vector whose vector components are z_i in the following compact form: $(z_1, z_2, \ldots, z_{\mathrm{M}}) = [\, z_1^T \; z_2^T \; \ldots \; z_{\mathrm{M}}^T \,]^T \in \mathbb{R}^q$, where $q = \sum_{i=1}^{M} q_i$. The symbols $\oplus/\ominus$ denote the Minkowski sum/difference (see [12]). We denote by $\|\cdot\|$ the Euclidean norm. The expression $\|x\|_\Gamma^2$ stands for $x^T \Gamma x$, where x is a column vector. Matrix Γ is the positive semidefinite, i.e., $\Gamma \geq 0$, if, for any x, $x^T \Gamma x \geq 0$. Given two symmetric positive semidefinite matrices A and B, it holds that $A \geq B$ if, for any x, $A - B \geq 0$. A ball with radius ρ and centered at $\bar{x}$ in the $\mathbb{R}^{dim}$ space is defined as $\mathscr{B}_\rho(\bar{x}) := \{x \in \mathbb{R}^{dim} : ||x - \bar{x}|| \leq \rho\}$. Finally, I_n denoted the identity matrix of dimension n, $\mathbb{1}$ is the unit vector and $\otimes$ denotes the Kronecker matrix product, i.e., $A \otimes B$ is the matrix whose ijth block entry is $a_{ij} B$, where a_{ij} is the ijth element of A.

II. Statement of the Problem and Models

In this section, we present the model of the overall system under study, and the main idea allowing to simplify the control problem both at high and at low hierarchical levels.

A. System Model and Control Objectives

We assume that the overall system $\mathscr{S}$ is composed by M discrete time, linear, independent subsystems described by

$$\mathscr{S}_i : \begin{cases} x_i(h+1) = A_i x_i(h) + B_i u_i(h) \\ y_i(h) = C_i x_i(h) \end{cases} \tag{1}$$

$i = 1, 2, \ldots, M$, where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^m$, and $y_i \in \mathbb{R}^p$ are the state, input, and output vectors, respectively, while h is the discrete-time index.

Subsystems $\mathscr{S}_i$, $i = 1, \ldots, M$ are similar, in the sense that the inputs $u_i(h)$ [respectively, the outputs $y_i(h)$] are homogeneous vectors. The following properties are assumed to hold.

Assumption 1:

1) A_i is Schur stable, $i = 1, \ldots, M$;
2) $m = p$;
3) $\det(C_i(I_{n_i} - A_i)^{-1} B_i) \neq 0$, $i = 1, \ldots, M$.

We will also use, for all $i = 1, \ldots, M$, the equivalent infinite impulse response forms

$$y_i(h) = \sum_{j=1}^{+\infty} G_j^i u_i(h-j) \tag{2}$$

where $G_j^i = C_i A_i^{(j-1)} B_i \in \mathbb{R}^{p \times m}$ is the impulse response matrix of the ith subsystem. The control scheme to be designed must allow to coordinate the M subsystems in such a way that the following objectives are attained.

1) *Collective Output Tracking and Constraint Satisfaction:* Solve a constrained control problem for the collective output

$$y(h) = \sum_{i=1}^{M} \chi_i y_i(h) \tag{3}$$

where χ_i can be either zero or one depending whether the ith subsystem is in use or not. Specifically, we aim to drive $y(h)$ to a desired reference value y_{ref} while verifying collective output constraint

$$y(h) \in \mathscr{Y} \tag{4}$$

where $\mathscr{Y}$ is a specified compact and convex output constraint set.

2) *Local Constraints Satisfaction:* Verify local input constraints, for each subsystem $\mathscr{S}_i$, $i = 1, \ldots, M$, of the type

$$u_i(h) \in \mathscr{U}_i. \tag{5}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

3) *Resource Sharing:* Assuming that a subset of agents $\sigma \subseteq \{1, \ldots, M\}$ shares the same (limited) input resource, we may require that

$$u_{\text{shared}}(h) = \sum_{i\in\sigma} u_i(h) \in \mathscr{U}_{\text{shared}}. \tag{6}$$

The sets $\mathscr{U}_i$ and $\mathscr{U}_{\text{shared}}$ are compact and convex, and they contain the origin, possibly on their boundary.

B. Input Signal Contributions

To address the overall design problem, we define each input signal $u_i(h)$ as the sum of two contributions

$$u_i(h) = \alpha_i \bar{u}(h) + v_i(h) \tag{7}$$

where: 1) the common input $\bar{u}(h)$ will be computed by a centralized high-level, low-dimensional, and slow-timescale controller designed to pursue collective output tracking and cnstraint satisfaction; 2) the terms $v_i(h)$ will be defined by M local low-level, fast controllers to enforce local input constraints and optimize local dynamic performances; and 3) the weights $\alpha_i \geq 0$ will be chosen offline to guarantee system-wide optimality. Their values are temporarily assumed to be fixed; their choice and possible adaptive tuning, also allowing for plug-and-play operations, will be discussed in Section IV. In (3), it is assumed that $\chi_i = 0$ if $\alpha_i = 0$, while $\chi_i = 1$ if $\alpha_i > 0$.

The feasibility properties of the two schemes are strictly correlated. In particular, letting $\bar{\mathscr{U}}$ and $\mathscr{V}_i$ be compact and convex sets and in view of (7), in order to satisfy constraints (5), (6) we can enforce

$$\bar{u}(h) \in \bar{\mathscr{U}} \tag{8a}$$

$$v_i(h) \in \mathscr{V}_i, \quad i = 1, \ldots, M \tag{8b}$$

and select $\bar{\mathscr{U}}$, $\mathscr{V}_i$, and the parameters α_i such that

$$\mathscr{V}_i \oplus \alpha_i \bar{\mathscr{U}} \subseteq \mathscr{U}_i, \quad \text{for all } i = 1, \ldots, M \tag{9a}$$

$$\left(\sum_{i\in\sigma} \alpha_i\right) \bar{\mathscr{U}} \oplus \left(\bigoplus_{i\in\sigma} \mathscr{V}_i\right) \subseteq \mathscr{U}_{\text{shared}}. \tag{9b}$$

C. Output Signal Contributions

To develop the models used for control, we rewrite (2) as follows:

$$y_i(h) = \sum_{j=1}^{T_L} G_j^i u_i(h-j) + \sum_{j=T_L+1}^{+\infty} G_j^i u_i(h-j) \tag{10}$$

where the positive integer T_L may be selected so that $G_{T_L+j}^i \simeq 0$, $j > 0$. From (7) and (10) and applying the superposition principle, for each subsystem $i = 1, \ldots, M$, we obtain that

$$y_i(h) = \sum_{j=1}^{T_L} \alpha_i G_j^i \bar{u}(h-j) + w_i(h) \tag{11}$$

where $w_i(h) = \alpha_i \sum_{j=T_L+1}^{+\infty} G_j^i \bar{u}(h-j) + y_i^{(v)}(h)$ and where $y_i^{(v)}(h)$ is the output of system (1) with input $v_i(h)$. As discussed in Section II-B, the input contribution $\bar{u}(h)$ is managed by the high-level controller that runs at a slower timescale.

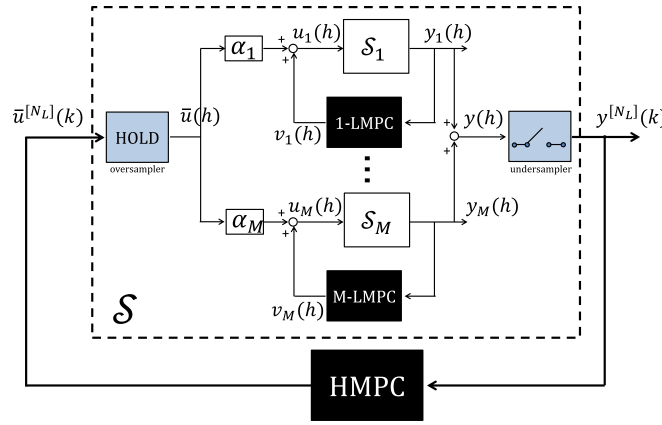

Fig. 1. Sketch of the overall control system.

Namely, it is assumed that there exist two positive integer values T_H and N_L such that $T_L = T_H N_L$. Then, $\bar{u}(h)$ is kept constant for N_L subsequent time steps and equal to $\bar{u}^{[N_L]}(k)$, i.e., $\bar{u}(h) = \bar{u}^{[N_L]}(k)$ for all $h \in \{kN_L, \ldots, (k+1)N_L - 1\}$ for all $k \geq 0$. Defining $y_i^{[N_L]}(k) = y_i(kN_L)$, from (11) we obtain, for all $k \geq 0$, that

$$y_i^{[N_L]}(k) = \sum_{l=1}^{T_H} \alpha_i G_l^{i,[N_L]} \bar{u}^{[N_L]}(k-l) + w_i(kN_L) \tag{12}$$

where $G_l^{i,[N_L]} = \sum_{j=(l-1)N_L+1}^{lN_L} G_j^i$. From (12), we realize that the signal $y_i^{[N_L]}(k)$ has two additive contributions, listed as follows:

1) A pure finite impulse response (FIR) system $\sum_{l=1}^{T_H} \alpha_i G_l^{i,[N_L]} \bar{u}^{[N_L]}(k-l)$: this component has $\bar{u}^{[N_L]}(k-l)$ as input, and therefore it will be managed by the high-layer (slow) control system (denoted HMPC in the sequel) at a centralized level.
2) The term $w_i(kN_L)$: this component will be managed by the fast, local, low-level control system (denoted i-LMPC in the sequel) acting on input $v_i(h)$, committed to guaranteeing that, for all $k \geq 0$

$$w_i(kN_L) = 0. \tag{13}$$

A sketch of the overall control scheme is depicted in Fig. 1, highlighting the role of each control layer in the definition of the control variables to be applied to each subsystem.

III. Regulator Design

In this section, we present the algorithms for the design of the MPC controllers at the two layers of the control structure and we prove their feasibility and convergence properties.

A. High-Level Regulator

In order to derive the model used for the design of the controller at the high level, define $y^{[N_L]}(k) = y(kN_L)$. In view of (3) and (12), and under (13), we can write

$$y^{[N_L]}(k) = \sum_{l=1}^{T_H} G_l^{[N_L]} \bar{u}^{[N_L]}(k-l) \tag{14}$$

where $G_l^{[N_L]} = \sum_{i=1}^{M} \alpha_i G_l^{i,[N_L]}$. Note that the reformulation of the individual systems as FIR ones has allowed us to easily obtain the slow-timescale collective system model (14) to be controlled in a scalable way (i.e., through the input component $\bar{u}^{[N_L]}(k)$) by the high-level regulator. The following assumption, guaranteeing that the high-level model has an invertible gain matrix, is required for the well-posedness of the high-level control problem.

Assumption 2: The matrix $\sum_{l=1}^{T_H} G_l^{[N_L]}$ is invertible.

Before presenting the optimization problem solved at each time instant, recall that y_{ref} is the reference that should be ideally tracked by the collective output, and note that it could not be reached within the considered prediction horizon due to the presence of the input constraints. For this reason, we define by r a feasible output reference, which must be regarded as a further optimization variable and defined by the control algorithm to be as near as possible to y_{ref}. Therefore, the optimization problem to be solved at the high layer at a generic time instant t aims to minimize the cost function

$$J_{HL} = \sum_{k=t}^{t+N_H-1} \left\{ \| y^{[N_L]}(k) - r(t) \|_{Q_y}^2 + \| \bar{u}^{[N_L]}(k) - \bar{u}^{[N_L]}(k-1) \|_R^2 \right\} + \gamma \, \| r(t) - y_{\text{ref}} \|^2. \quad (15)$$

In (15), $N_H > T_H$ is the prediction/control horizon, Q_y and R are positive definite symmetric matrices, while γ is a positive constant, defined such that

$$\gamma I_p \geq P \quad (16a)$$

where

$$P = \left(\sum_{l=1}^{T_H} G_l^{[N_L]} \right)^{-T} \times \left(\sum_{k=0}^{T_H-1} \left(\sum_{l=k+1}^{T_H} G_l^{[N_L]} \right)^T Q_y \left(\sum_{l=k+1}^{T_H} G_l^{[N_L]} \right) + R \right) \times \left(\sum_{l=1}^{T_H} G_l^{[N_L]} \right)^{-1}. \quad (16b)$$

The optimization problem is formulated including the terminal constraints

$$y^{[N_L]}(t + N_H) = r(t) \quad (17a)$$

and for all $k = t + N_H - T_H, \ldots, t + N_H - 1$

$$\bar{u}^{[N_L]}(k) = \left(\sum_{l=1}^{T_H} G_l^{[N_L]} \right)^{-1} r(t) \quad (17b)$$

that allow to guarantee that a solution exists leading to the actually computed reference $r(t)$ at $t + N_H$ at a steady-state condition.

Under the specifications mentioned above, the overall High-level MPC optimization problem (therefore denoted HMPC) is

$$\min_{r(t), \bar{u}^{[N_L]}(t:t+N_H-1)} J_{HL} \quad (18)$$

subject to the dynamical model (14). The constraints are

$$\bar{u}^{[N_L]}(k) \in \bar{\mathscr{U}} \quad (19)$$

$$y^{[N_L]}(k) \in \mathscr{Y} \quad (20)$$

for all $k = t, \ldots, t + N_H - 1$, while the terminal constraint (17) can be reformulated as

$$\begin{bmatrix} I_p \\ \left(\sum_{l=1}^{T_H} G_l^{[N_L]} \right)^{-1} \end{bmatrix} r(t) \in \mathscr{Y}_u(\varepsilon) \quad (21)$$

where $\mathscr{Y}_u(\epsilon)$ is a closed and convex set satisfying $\mathscr{Y}_u(\epsilon) \oplus \mathscr{B}_\epsilon(0) \subseteq \mathscr{Y} \times \bar{\mathscr{U}}$.

Remark 1: For enforcing a feedback term in the controller, it could be beneficial to introduce an equivalent autoregressive formulation. This requires a slight reformulation of model (14) into the following equivalent one:

$$y^{[N_L]}(k) = y^{[N_L]}(k-1) + \sum_{l=1}^{T_H} G_l^{[N_L]} \Delta \bar{u}^{[N_L]}(k-l) \quad (22)$$

where $\Delta \bar{u}^{[N_L]}(l) = \bar{u}^{[N_L]}(l) - \bar{u}^{[N_L]}(l-1)$. Note that the equivalence of (14) and (22) holds in case the weights $\alpha_1, \ldots, \alpha_M$ are time-invariant. In case they are time-varying (as it is considered in Section IV), this reformulation leads to slightly more involved computations. For this reason, it has not been explicitly used in the following.

B. Low-Level Regulators

The behavior of the low-level controller will be different in the two cases $\alpha_i > 0$ (i.e., $\chi_i = 1$ in (3)) and $\alpha_i = 0$ (i.e., $\chi_i = 0$), which are considered separately in this section.

Connected Subsystem, with $\alpha_i > 0$: The role of the low-level regulators is twofold. First, they are needed to remove the mismatch of the high-level simplified system from its model (14) by enforcing (13). Second, they optimize the performances of the subsystems in transient conditions.

The state-space model describing the evolution of the variable $w_i(h)$ is required for control purposes. In particular, $w_i(h)$ can be regarded as the output of the following model:

$$\begin{cases} x_i^{(w)}(h+1) = A_i x_i^{(w)}(h) + A_i^{T_L} B_i \alpha_i \bar{u}(h - T_L) + B_i v_i(h) \\ w_i(h) = C_i x_i^{(w)}(h). \end{cases} \quad (23)$$

For well-posedness, the following rather technical assumption is required to guarantee the convergence of the state of the system at fast sampling time (see Appendix and [17]).

Assumption 3: For all $i = 1, \ldots, M$, the transfer matrix $C_i(zI_{n_i} - A_i^{N_L})^{-1}$ has no zeros on the unitary circle. □

At the low level, we adopt a shrinking horizon approach and, for each $i = 1, \ldots, M$, at any time $h \in \{kN_L, (k+1)N_L - 1\}$ we solve the following low-level MPC problem (therefore denoted i-LMPC)

$$\min_{v_i(h:(k+1)N_L-1)} J_{LL}^i \quad (24a)$$

where

$$J_{LL}^i = \sum_{s=h}^{(k+1)N_L-1} \left\{ \|w_i(s)\|_{Q_i}^2 + \|v_i(s) + \alpha_i g_i^{[T_L]} \bar{u}^{[N_L]}(k-T_H)\|_{R_i}^2 \right\} \tag{24b}$$

subject to the model (23) and the constraints

$$v_i(s) \in \mathscr{V}_i, \quad s = h, \dots, (k+1)N_L - 1 \tag{24c}$$

$$w_i((k+1)N_L) = 0 \tag{24d}$$

$$x_i^{(w)}((k+1)N_L) \in \mathbb{X}_i^F. \tag{24e}$$

In (24b), weights Q_i and R_i are symmetric and positive definite arbitrary matrices. Also, we have defined

$$g_i^{[T_L]} = (C_i(I_{n_i} - A_i)^{-1}B_i)^{-1} C_i (I_{n_i} - A_i)^{-1} A_i^{T_L} B_i$$

in such a way that if, at steady state, $v_i(h) + \alpha_i g_i^{[T_L]} \bar{u}^{[N_L]}(k - T_H) = 0$, then $w_i(h) = 0$. Note that $g_i^{[T_L]}$ is well defined in view of Assumption 1.

The set $\mathbb{X}_i^F$ used in (24e) is a suitable terminal set for $x_i^{(w)}$ that guarantees the feasibility of the low-level problem at the subsequent long sampling time, i.e., for the problem at $h = (k+1)N_L$. Its definition is deferred to Section III-C.

Also, note that, since $w_i(h) = y_i(h) - \alpha_i \sum_{j=1}^{T_L} G_j^i \bar{u}(h - j)$ is given, the estimate of $x_i^{(w)}(h)$ can be computed at all sampling times. This is possible thanks to the detectability of the pair (A_i, C_i), guaranteed by Assumption 1. A more robust and possibly fast estimation of variable $x_i^{(w)}$ can be obtained provided that the pair (A_i, C_i) is observable.

Disconnected Subsystem, with $\alpha_i = 0$: In case $\alpha_i = 0$, the input of subsystem $\mathscr{S}_i$ is set to zero, i.e., $v_i(h) = 0$ for all h. In view of the fact that A_i is Schur stable (see Assumption 1), then $x_i^{(w)}(h) \to 0$ as $h \to 0$. Importantly, since (see in Section III-C) set $\mathbb{X}_i^F$ contains the origin in its interior, for all possible initial conditions there exists $\bar{h} \geq 0$ such that $x_i^{(w)}(h) \in \mathbb{X}_i^F$ for all $h \geq \bar{h}$.

C. Design Requirements and Main Results

We now introduce the main conditions required to make the two-layer control scheme consistent and well posed and we derive the main feasibility and convergence results of the proposed design method.

Up to this point, the values of the parameters $\alpha_i \geq 0$, $i = 1, \dots, M$, have been assumed to be fixed. However, in the final part of this paper, they will be considered as additional tuning knobs to be possibly retuned online and/or to be modified to allow for plug-and-play operations. For this reason, the conditions discussed in the following will be formulated to be consistent with all the values that α_i can take. The range of values that α_i can take when $\mathscr{S}_i$ is in operation (i.e., $\alpha_i > 0$) is defined by the lower and upper bounds $\underline{\alpha}_i$ and $\bar{\alpha}_i$, respectively, while, as discussed in Section III-B, $\alpha_i = 0$ when $\mathscr{S}_i$ is essentially disconnected. Overall, the set where α_i can take values is defined as follows:

$$\alpha_i \in \{0\} \cup [\underline{\alpha}_i, \bar{\alpha}_i]. \tag{25}$$

Remark 2: The minimum value $\underline{\alpha}_i$ means that, if subsystem $\mathscr{S}_i$ is in use, it is required to provide a minimum contribution to the overall control action. This represents the fact that, in many applications, for economic reasons, it is not worth using an actuator below a given threshold of its operating conditions. □

We should design $\underline{\alpha}_i$, $\bar{\alpha}_i$ such that

$$0 \leq \underline{\alpha}_i < \bar{\alpha}_i \leq 1, \quad \sum_{i=1}^{M} \bar{\alpha}_i \geq 1. \tag{26a}$$

The further requirements that $\bar{\mathscr{U}}, \mathscr{V}_i, \underline{\alpha}_i, \bar{\alpha}_i, \mathbb{X}_i^F, i = 1, \dots, M$ must fulfill are now listed.

1) In order to guarantee the fulfillment of both the local (5) and the shared (6) input constraints, conditions (9a) and (9b) must be verified for all the values of $\alpha_i \in [\underline{\alpha}_i, \bar{\alpha}_i]$.
2) The set $\mathbb{X}_i^F$ in (24e) must guarantee that there exists a scalar $\lambda_i \in [0, 1)$ such that
$$A_i^{N_L} \mathbb{X}_i^F \subseteq \lambda_i \mathbb{X}_i^F. \tag{26b}$$
This is always possible in view of Assumption 1.
3) When $\mathscr{S}_i$ is in operation mode (i.e., when $\alpha_i > 0$, $\chi_i = 1$), it must be imposed, at the low level, that $w_i(kN_L) = 0$ in a recursive fashion. This implies that the set of states that can be reached in N_L steps by using $v_i(h)$ in system (23) must allow to cancel the effect of the input $\bar{u}$ and of possible nonnull initial conditions of $x_i^{(w)}$. More specifically, we must guarantee that, for all $x_i^{(w)}(kN_L) \in \mathbb{X}_i^F$ and for all admissible inputs $\bar{u}(h - T_L) = \bar{u}^{[N_L]}(k - T_H) \in \bar{\mathscr{U}}$ constant for all $h = kN_L, \dots, (k+1)N_L - 1$, there exists a feasible sequence $v_i(kN_L), \dots, v_i((k+1)N_L - 1)$ such that $x_i^{(w)}((k+1)N_L) \in \mathbb{X}_i^F$ and that $w_i((k+1)N_L) = C_i x_i^{(w)}((k+1)N_L) = 0$. The following condition is therefore required for all $\alpha_i \in [\underline{\alpha}_i, \bar{\alpha}_i]$:
$$\begin{bmatrix} C_i \\ I_{n_i} \end{bmatrix} (\lambda_i \mathbb{X}_i^F \oplus \alpha_i \bar{R}_i \bar{\mathscr{U}}) \subseteq \begin{bmatrix} 0 \\ I_{n_i} \end{bmatrix} \mathbb{X}_i^F \oplus \begin{bmatrix} C_i \\ I_{n_i} \end{bmatrix} (-R_i^{(v)} (\mathscr{V}_i)^{N_L}) \tag{26c}$$
where
$$\bar{R}_i = A_i^{T_L} \sum_{j=0}^{N_L-1} A_i^j B_i, \quad R_i^{(v)} = \left[A_i^{N_L-1} B_i, \dots, B_i \right]$$
and $(\mathscr{V}_i)^{N_L} = \mathscr{V}_i \times \cdots \times \mathscr{V}_i$, i.e., the N_L-times Cartesian product of sets $\mathscr{V}_i$. Note that
$$R_i^{(v)} (\mathscr{V}_i)^{N_L} = \bigoplus_{j=0}^{N_L-1} A_i^j B_i \mathscr{V}_i.$$
4) Finally, to make such compensation possible also at steady-state conditions, the following must hold for all $\alpha_i \in [\underline{\alpha}_i, \bar{\alpha}_i]$:
$$-\alpha_i g_i^{[T_L]} \bar{\mathscr{U}} \subseteq \mathscr{V}_i \tag{26d}$$
$$\alpha_i g_i^x \bar{\mathscr{U}} \subseteq \mathbb{X}_i^F \tag{26e}$$
where $g_i^x = (I_{n_i} - A_i)^{-1}(A_i^{T_L} B_i - B_i g_i^{[T_L]})$.

Remark 3: Note that, if set $\bar{\mathscr{U}}$ contains the origin, to verify conditions (26) it is enough to check their validity when $\alpha_i = \bar{\alpha}_i$. Also, in this case, the value $\underline{\alpha}_i = 0$ is possible, while in the general case when $\bar{\mathscr{U}}$ does not contain the origin for some values of i, it may be critical to set $\underline{\alpha}_i = 0$. □

The following main result can now be stated (the proof is in the Appendix).

Theorem 1: Under Assumptions 1–3, if the feasibility of the high-layer problem (17)–(21) and of the low-layer ones (24), for all $i = 1, \dots, M$, is verified at time $t = 0$, then feasibility is guaranteed:

1) for (17)–(21) at all time instants $h = kN_L$, $k \geq 0$;
2) for (24), for all $i = 1, \dots, M$, at all time instants h, $h \geq 0$.

Also, (5) and (6) hold for all $h \geq 0$ and (4) holds for all $h = kN_L$, where $k \geq 0$. Finally, as $h \to \infty$, the output $y(h) \to y_{\text{feas.ref}}$, where

$$y_{\text{feas.ref}} = \underset{\begin{bmatrix} I_p \\ \left(\sum_{l=1}^{T_H} G_l^{[N_L]}\right)^{-1} \end{bmatrix} z \in \mathscr{Y}_u(\epsilon)}{\arg\min} \|z - y_{\text{ref}}\|^2. \quad (27)$$

Remark 4: Theorem 1 guarantees that $y(kN_L) \in \mathscr{Y}$ for all $k \geq 0$, i.e., the collective output constraint (4) is enforced in the slow time scale, which is acceptable in many application scenarios. However, it could be unacceptable that even small-amplitude violations may occur in the fast timescale, i.e., that $y(h) \notin \mathscr{Y}$ for $h \neq kN_L$. In these cases, a possible solution for guaranteeing (4) for all $h \geq 0$ consists of slightly modifying the proposed control scheme at the price of a more conservative problem statement. In fact, in view of (11), we can define the set $\mathscr{W} = \bigoplus_{i=1}^{M} \chi_i \mathscr{W}_i$, where $\mathscr{W}_i = \alpha_i (\bigoplus_{j=T_L+1}^{+\infty} G_j^i \bar{\mathscr{U}}) \oplus (\bigoplus_{j=1}^{+\infty} G_j^i \mathscr{V}_i)$. Constraint (4) can be enforced by adding the following constraint to the high-layer optimization problem (17)–(21):

$$\sum_{j=1}^{T_L} \left(\sum_{i=1}^{M} \alpha_i G_j^i \right) \bar{u}(h - j) \in \mathscr{Y} \ominus \mathscr{W} = \bar{\mathscr{Y}}$$

for all $h = tN_L, \dots (t + N_H)N_L - 1$. Recursive feasibility is simply guaranteed by redefining, in (21), set $\mathscr{Y}_u(\varepsilon)$ as the closed and convex set satisfying $\mathscr{Y}_u(\varepsilon) \oplus \mathscr{B}_\varepsilon(0) \subseteq \bar{\mathscr{Y}} \times \bar{\mathscr{U}}$.

IV. Static and Dynamic Optimization of the Weights α_i and Plug-and-Play Operations

In the control law (7), the term $\bar{u}(t)$ may be regarded as the total input request to the set of subsystems, and the parameters α_i represent the share of input assigned to each subsystem $\mathscr{S}_i$. Their values can be chosen according to global optimality criteria by solving a static higher layer optimization problem, periodically, or based on an event-driven rationale. The optimization problem proposed here has the role of minimizing the steady-state values of the control signals, in order to minimize the overall cost for controlling the plant, but other alternative cost functions can be proposed and used, and other constraints can be included.

At steady state, from (14), the input $\bar{u}$ must take the value $\bar{u}^{ss}$ such that

$$\left(\sum_{i=1}^{M} \alpha_i \sum_{l=1}^{T_H} G_l^{i,[N_L]} \right) \bar{u}^{ss} = y_{\text{ref}}. \quad (28)$$

Note that the matrix $\sum_{i=1}^{M} \alpha_i \sum_{l=1}^{T_H} G_l^{i,[N_L]}$ has full rank in view of Assumption 2. Also, to make $w_i = 0$, from (23), $v_i = -\alpha_i g_i^{[T_L]} \bar{u}^{ss}$, Therefore, at steady state, the input u_i must take the value

$$u_i^{ss} = \alpha_i \left(I_p - g_i^{[T_L]} \right) \bar{u}^{ss}. \quad (29)$$

The proposed minimization problem reads

$$\min_{\bar{u}^{ss}, \{u_i^{ss}, \alpha_i\}_{i=1,\dots,M}} \sum_{i=1}^{M} q_i^{\alpha} \left\| u_i^{ss} \right\|^2 \quad (30)$$

where q_i^{α} is a suitable cost associated with the use of subsystem $\mathscr{S}_i$, under constraints (25) and

$$\sum_{i=1}^{M} \alpha_i = 1. \quad (31)$$

Note that (31) enforces the uniqueness of the solution to (30).

A. Time Varying Weights α_i

We consider now the case where a change in the weights α_i is required during the system operation, and how this impacts on the control scheme at the two dynamic control layers. This will pave the way for the application of the method in plug-and-play scenarios, as described in Section IV-B. At this point, we will only make the assumption that the optimization problem (30) is run during the system operation, and that the changes are applied only at the beginning of long sampling times, i.e., at $h = kN_L$. In general, we denote by $\alpha_i(k)$ the values taken by α_i at time $h = kN_L$, for all $i = 1, \dots, M$, which are kept constant during the high-layer sampling time.

1) High-Level Control During Weight Changes: At time $h = kN_L$, the model (14) must be rewritten as

$$y^{[N_L]}(k) = \sum_{l=1}^{T_H} G_{l,k}^{[N_L]} \bar{u}^{[N_L]}(k - l) \quad (32)$$

where $G_{l,k}^{[N_L]} = \sum_{i=1}^{M} \alpha_i(k - l) G_l^{i,[N_L]}$. A variation of the weights α_i, $i = 1, \dots, M$ during the system operation may compromise the feasibility properties of the high-layer control scheme. For this reason, a three-step procedure is proposed.

a) Based on periodic or event-based call, a weight change is proposed by the optimizer running (30). The candidate new weights, denoted α_i^*, $i = 1, \dots, M$, are transmitted at time $\bar{t}$ to the high-level dynamic controller but are not directly applied.
b) The feasibility of the optimization problem (18) is checked at a time instant $t = \bar{t}$ where the output variable is computed based on (32), with $\alpha_i(k) = \alpha_i^-$ for all $k < t$, α_i^- is the value taken by the weight before the variation request, while $\alpha_i(k) = \alpha_i^*$ for all $k \geq t$. If the feasibility of the problem is verified, then we can set

$\alpha_i(t) = \alpha_i^*$ for all $i = 1, \ldots, M$, for all $t \geq \bar{t}$. Otherwise, go to step 3.

c) For any time instant $t \geq \bar{t}$, replace problem (18) with the following one:

$$\min_{r(t),\bar{u}^{[N_L]}(t:t+N_H-1),\alpha_1(t),\ldots,\alpha_M(t)} J_{HL}+\gamma_\alpha \sum_{i=1}^{M} \left\| \alpha_i(t)-\alpha_i^* \right\|^2 \tag{33}$$

subject to the dynamical system (32). In the optimization problem solved at time t (in a receding horizon fashion), it is assumed that $\alpha_i(k) = \alpha_i(t)$ is kept constant for all $k \geq t$. The constraints are (17) and (19)–(21), where the term $G_l^{[N_L]}$ is replaced by $G_{l,k}^{[N_L]}$, and (25), (31).

Note that the feasibility of the problem (33) is guaranteed by the fact that one can always keep $\alpha_i(t)$ constant at the last feasible value.

2) Low-Level Control During Weight Changes: The recursive feasibility of the low-level optimization problem (24) is less critical than the high-level one during weight changes. Indeed, as already remarked, it is assumed that the weight α_i remains constant over the long sampling time, i.e., during the low-level shrinking-horizon optimization problem for all $h \in [kN_L, \ldots, (k+1)N_L - 1]$. Therefore, the results obtained on recursive feasibility still hold thanks to the terminal constraint (24e) and to the definition of $\mathbb{X}_i^F$, which is given for all admissible values of $\alpha_i \in [\underline{\alpha}_i, \bar{\alpha}_i]$.

The only peculiar cases are the switch from $\alpha_i > 0$ to $\alpha_i = 0$ (i.e., the disconnection of subsystem $\mathscr{S}_i$) and viceversa (i.e., the connection of subsystem $\mathscr{S}_i$). Although the first switch does not entail any problem as far as the low-level optimization is concerned, the connection may imply minor feasibility issues for (24). Indeed, if the feasibility of (24) is not verified at time $h = kN_L$, however, this problem is just temporary. In fact, feasibility is guaranteed if $x_i^{(w)} \in \mathbb{X}_i^F$, which is proven to occur after a finite number of steps. These cases are more thoroughly discussed in Section III-B.

B. Plug-and-Play Operations

The cases in which one or more subsystems join or leave the network may be naturally included in the scenario in which the weights α_i are time varying. We can distinguish two main cases, the plug-in and the unplug cases.

1) Plug-In Requests: Assume that, at time instants $t \leq \bar{t}$, the system is controlled using the scheme proposed in this paper and is composed of M subsystems. At time $\bar{t}$, a plug-in request is received, i.e., we want to include subsystem $M+1$ in the network. Note that the case in which $\mathscr{S}_{M+1}$ is not plugged-in is equivalent to the case in which $\mathscr{S}_{M+1}$ is plugged-in, but with weight $\alpha_{M+1} = 0$. Thanks to this simple remark, we can define a plug-in procedure, consisting of the following steps.

a) *Structural Plug-In Design:* Define a tuple $(\underline{\alpha}_{M+1}, \bar{\alpha}_{M+1}, \mathscr{V}_{M+1}, \mathbb{X}_{M+1}^F)$ satisfying the conditions (26).

b) *Feasibility Plug-In Test:* Set $\alpha_{M+1} = 0$ and wait until $x_{M+1}^{(w)}(kN_L) \in \mathbb{X}_i^F$.

If these steps have been successfully carried out, then the subsystem $\mathscr{S}_{M+1}$ is formally plugged in. At this point, we may run the procedure sketched in Section IV-A to properly take advantage of the newly plugged device.

2) Unplug Requests: Assume that, at time instants $t \leq \bar{t}$, the system is controlled using the scheme proposed in this paper and is composed of M subsystems. At time $\bar{t}$, an unplug request is received, i.e., we want to exclude subsystem M from the network. Note that the situation where $\mathscr{S}_M$ gets unplugged can be achieved by setting $\alpha_M = 0$. The unplug operation is done by taking the following steps.

a) *Structural Unplug Condition:* Solve the optimization problem (30) with the further constraint $\alpha_M = 0$. The candidate new weights α_1^*, α_2^*, ..., $\alpha_M^* = 0$ are transmitted to the high-level dynamic controller but are not directly applied.

b) According to the procedure sketched in Section IV-A, we must further check the feasibility of the high-level optimization problem (18) (see step 2 in Section IV-A) with the candidate weights $\alpha_1^*, \alpha_2^*, \ldots, \alpha_M^*$. If the feasibility of the problem is verified, then we can set $\alpha_i(t) = \alpha_i^*$ for all $i = 1, \ldots, M$. Otherwise, we need to solve (33) at each time instant. Unfortunately, using this procedure, the condition required for unplugging subsystem $\mathscr{S}_M$ (i.e., that $\alpha_M(t) = 0$) may not be attained after a finite number of steps. To try to enforce the attainment of the unplug condition after a finite number of steps, we suggest to (periodically, in case) check if (33) admits a solution with the further constraint $\alpha_M(t) = 0$, which allows to unplug $\mathscr{S}_M$, while the weights α_i related to the other subsystems converge to their steady-state values.

If the steps mentioned above are successfully carried out, then we can remove the subsystem $\mathscr{S}_M$ from the overall plant. Otherwise, the system cannot support an unplug event and the request is denied.

V. Simulation Example

The hierarchical control algorithm described in the this paper has been used for coordination of a number of synchronous machines.

A. Description of the Models

Consider six diesel generators connected to a network with terminal voltage 240 V and frequency 60 Hz, which must track a total electrical power ($y_{\text{ref}} = 98.32$ kW). The values of their rated powers are $\{16.1, 25.0, 26.5, 30.7, 40.8, 47.6\}$ kW, respectively. The linear continuous models of the generators are obtained from [22], and are of orders, $n_1 = n_6 = 9$, $n_2 = n_3 = 8$, $n_4 = 4$, and $n_5 = 5$. The input u_i and output y_i, for all $i = 1, \ldots, 6$, are the fuel flow rate and the produced power, respectively, with $m = 1$ and $p = 1$. The six generators' linear continuous-time models have been sampled with $\Delta t = 1$ *s* to obtain their discrete-time counterpart of the fast time scale. Then, these discrete-time subsystems have been used as the models in (1) for the implementation of the hierarchical control structure. The control variables, as well

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

as the controlled variables, are limited by $0 \leq (u_1, \ldots, u_6) \leq (0.89, 1.39, 1.48, 1.71, 2.27, 2.64)$ g/s and $0 \leq (y_1, \ldots, y_6) \leq (16.1, 25.0, 26.5, 30.7, 40.8, 47.6)$ kW. The considered constraint on resource sharing is $0 \leq \sum_{i=3}^{5} u_i \leq 4.4$ g/s.

B. Design of the Sets

The following design choices have been taken. First, we set $T_{\rm L} = 60$, $T_{\rm H} = 6$, and $N_{\rm L} = 10$. Also, to verify (9) for all admissible values of α_i, the parameters $\underline{\alpha}_i$, $\bar{\alpha}_i$, and the sets $\bar{\mathscr{U}}$, $\mathscr{V}_i$, $i = 1, \cdots, 6$, have been selected as $(\underline{\alpha}_1, \ldots, \underline{\alpha}_6) = (0.1, 0.1, 0.13, 0.13,\ 0.15, 0.15)$, $(\bar{\alpha}_1,\ \cdots, \bar{\alpha}_6) = (0.17, 0.25, 0.23, 0.23,\ 0.25, 0.25)$, $\bar{\mathscr{U}} = [2,\ 4]$, and $\mathscr{V}_1 = [-0.2, 0.21]$, $\mathscr{V}_2 = [-0.2, 0.39]$, $\mathscr{V}_3 = [-0.26, 0.50]$, $\mathscr{V}_4 = [-0.26, 0.41]$, $\mathscr{V}_5 = [-0.3, 0.57]$, and $\mathscr{V}_6 = [-0.3, 1.64]$. Finally, the sets $\mathbb{X}_i^F$, $i = 1, \cdots, 6$, have been chosen as $\mathbb{X}_i^F = \{x_i^{(w)} | (x_i^{(w)})' P_i x_i^{(w)} \leq \mu_i\}$ where P_i is the solution to the Riccati equation related to the infinite horizon control problem with state weight $Q_{x,i} = I_{n_i}$ and input weight $R_i = 0.1 I_m$, while $\mu_i = 0.1$.

C. Simulation Results Without Plug-and-Play Operations

The hierarchical control structure has been applied to the system with only five generators, i.e., $M = 5$, assuming that the sixth generator is disconnected from the network. The optimization problem (30) has been solved with $(q_1^\alpha, \ldots, q_5^\alpha) = (0.78, 0.81, 0.82, 0.83, 0.84)$. The values of the parameters $\bar{u}^{ss}$, u_i^{ss}, and α_i, $i = 1, \ldots, 5$, are: $\bar{u}^{ss} = 3.14$, $(u_1^{ss}, \ldots, u_5^{ss}) = (0.447, 0.59, 0.583, 0.576, 0.565)$, $(\alpha_1, \ldots, \alpha_5) = (0.1496, 0.2107, 0.2191, 0.2028, 0.2178)$. The high-layer MPC has been designed with prediction horizon $N_{\rm H} = 10$, penalties $Q_{\rm y} = I_p$ and $R = 0.1 I_m$, while $\gamma = 24.99$. The low-layer shrinking horizon optimization algorithms have been solved with state and input penalties $Q_i = 10 I_p$, $R_i = 0.1 I_m$, $i = 1, \ldots, 5$. It is worth mentioning that a penalty on the deviation of the input of the form $\|v_i(h) - v_i(h-1)\|^2$ has also been added for each instant within the prediction interval during the transient phase to enforce smooth dynamic behavior and has been removed once the regulator was close to its steady state. For comparison, a centralized state-feedback MPC has been designed at any fast time h with cost function $J_c = \sum_{k=h}^{h+N-1} \| \sum_{i=1}^{5} C_i x_i(k) - y_{\rm ref}\|_{Q_c}^2 + \sum_{i=1}^{5} q_i^\alpha \|u_i(k)\|^2$ where the penalty $Q_c = 500$ and prediction horizon $N = N_{\rm H} \cdot N_{\rm L} = 100$. The (centralized) terminal set has been chosen as $\mathbb{X}_{\rm F} = \{x | \sum_{i=1}^{5} C_i x_i(t+N) = y_{\rm ref}\}$. All the simulation tests have been implemented using MATLAB, YALMIP, and MPT toolbox, see [7] and [9], in a PC with Intel Core i5-4200U 2.30 GHz and with Windows 10 operating system. The MATLAB QUADPROG solver has been used for the implementation of the centralized MPC and the HMPC problem of the proposed optimization algorithms, while SDPT3 solver has been used for the local i-LMPC regulators. The detailed online computational time required for each controller is reported in Table I, showing the computational advantages of the proposed hierarchical scheme. Note that the low-level controller calculations can be run in parallel thanks to the decentralized implementation.

TABLE I

ONLINE COMPUTATION TIME COMPARISON

Approach			Opt. activated at	Av. comp. time
Proposed	HMPC		$h = kN_{\rm L}$	0.360 s
	i-LMPC	$i = 1$	$\forall h$	0.453 s
		$i = 2$	$\forall h$	0.329 s
		$i = 3$	$\forall h$	0.333 s
		$i = 4$	$\forall h$	0.345 s
		$i = 5$	$\forall h$	0.356 s
Centralized MPC			$\forall h$	1.756 s

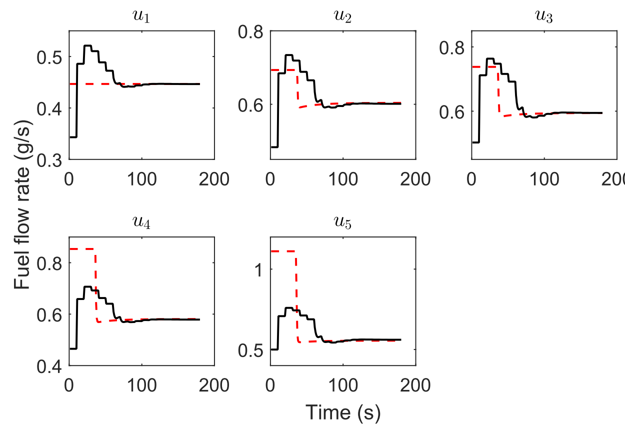

Fig. 2. Control variables of the controlled generators: overall control actions computed by the two-layer scheme (black solid lines) and control variables computed by the centralized scheme (red dashed lines).

The evolution of the control and output variables of the controlled subsystems are reported in Figs. 2 and 3, which show that, after an initial transient, inputs and outputs return to their nominal values, and both the separations of total electrical power and the control performance in terms of the proposed two-layer approach are close to those of centralized MPC. For a numerical comparison, we have computed the mean-square tracking error $J_y = (1/N_{sim}) \sum_{k=0}^{N_{sim}-1} \| \sum_{i=1}^{5} C_i x_i(k) - y_{\rm ref}\|^2$ and the input-related cost $J_u = (1/N_{sim}) \sum_{k=0}^{N_{sim}-1} (\sum_{i=1}^{5} q_i^\alpha \|u_i(k)\|^2)$ with N_{sim} simulation steps in case of our hierarchical scheme and in case of centralized MPC. For the proposed scheme, $J_y = 0.65 \cdot 10^3$ and $J_u = 2.8 \cdot 10^3$, while for centralized MPC, $J_y = 0.43 \cdot 10^3$ and $J_u = 3.1 \cdot 10^3$. This shows that, with respect to the tested centralized scheme, in transient conditions, the proposed scheme displays slightly worse performances in terms of tracking capabilities, but at the price of a more limited control effort. In steady-state, however, the performances are equivalent, both in term of steady-state tracking error and in terms of use of inputs.

D. Simulation Results With Plug-and-Play Operations

The two-layer control structure has also been applied with plug-and-play operations. Starting from the simulation conditions of scenario 1, at time $t = 180$ s, the sixth generator has been added to the network; then, at time $t = 360$ s, the third generator has been disconnected by setting $\alpha_3 = 0$. The evolution of the output and control variables of the controlled system are reported in Figs. 4 and 5. These figures show that,

TABLE II
ONLINE COMPUTATION TIME COMPARISON

Approach		Opt. activated at	Av. comp. time (s)					
			4 sys	8 sys	12 sys	16 sys	20 sys	24 sys
Proposed one	HMPC	$h = kN_L$	0.33	0.35	0.38	0.38	0.43	0.47
	i-LMPC	each fast time instant h	0.53	0.52	0.54	0.54	0.58	0.62
Centralized MPC		each fast time instant h	1.76	4.07	8.09	12.75	20.54	30.70

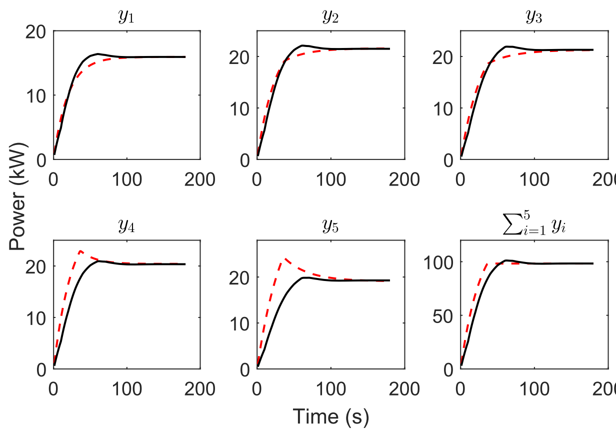

Fig. 3. Outputs of the controlled generators: outputs obtained with the two-layer scheme (black solid lines) and outputs obtained with the centralized scheme (red dashed lines).

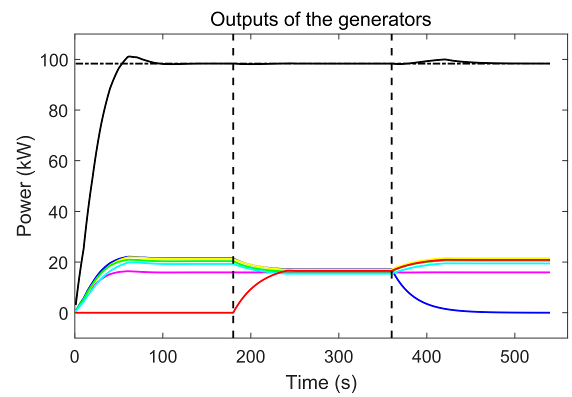

Fig. 5. Outputs of the controlled generators: outputs $(y_1, \ldots, y_6, \sum_{i=1}^{6} y_i)$ (magenta, blue, yellow, green, cyan, red, and black solid lines) and reference power (black dashed-dotted lines). Vertical dashed lines: plug-in/unplug instants.

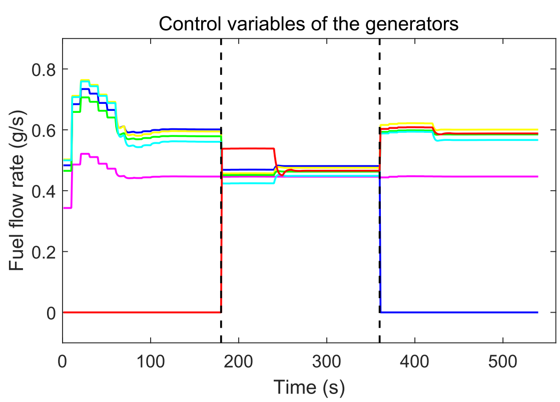

Fig. 4. Control variables of the controlled generators: overall control actions $(u_1, \ldots, u_6)$ (magenta, blue, yellow, green, cyan, and red solid lines). Vertical dashed lines: plug-in/unplug instants.

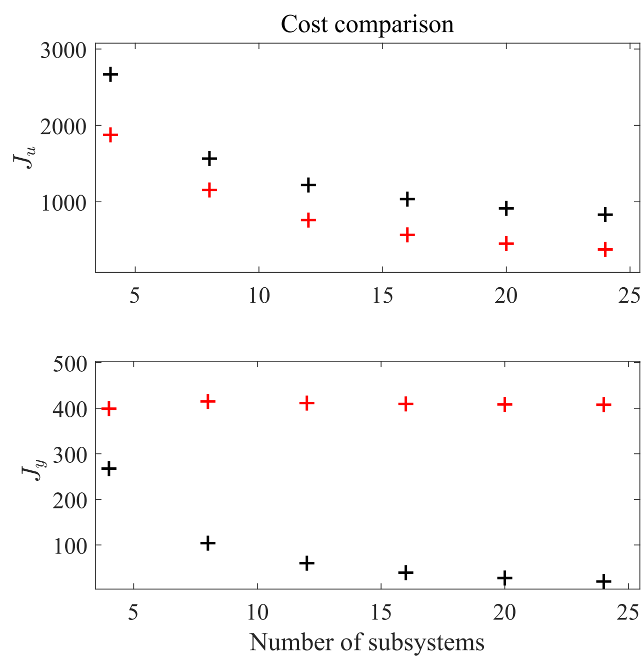

Fig. 6. Cost comparison. Red (black) + markers are the values of the input-related cost, i.e., J_u, computed with the proposed two-layer approach (centralized MPC) (top). Red (black) + markers are the values of the mean-square tracking error cost, i.e., J_y, computed with the proposed two-layer approach (centralized MPC) (bottom).

after an initial transient due to the plug-in procedure, inputs and outputs return to their current nominal values until the next plug-out operation occurs, when the two-layer control system properly reacts to bring the controlled variables to their new reference values.

E. Scalability of the Algorithm

The scalability of the proposed two-layer approach with respect to centralized MPC has been analyzed as follows. First, the simulation was run with a group of four subsystems (generators) of the same size (i.e., $n_i = 9$ for $i = 1, 2, 3, 4$), all with initial conditions equal to the ones in the previous simulations and with null past input values. Then, the simulation tests were repeated with 8, 12, 16, 20, 24 subsystems of the same size with the control objective to steer the collective output $\sum_{i=1}^{M} y_i$ to the reference value $y_{ref} = 100.2$. The computational times in different cases are listed in Table II. In particular, in Table II, we show the average (on all subsystems and on time) computational time required by the i-LMPC problems, the one required by the HMPC, and compare them to the average computational time required by a corresponding centralized MPC algorithm. This table shows that the time required to solve the control problems defined by our two-layer approach undergoes very small variations as a function of the subsystems number thanks to its distributed nature, while the time requires to solve centralized MPC scales with the number of subsystems.

The mean-square tracking error and input-related costs in different cases are reported in Fig. 6, it is apparent that the suboptimality of the solution, i.e., $|J_y - J_y^c|$, increases with the number of subsystems, but at price of a fully scalable computational load and a long-lasting more limited control effort.

VI. Conclusion

In this paper, a hierarchical control scheme has been proposed for the coordination of independent systems. This scheme is fully scalable: the size of the control problem to be solved at the high layer does not grow with the number of subsystems, and so does the complexity of the low-level control scheme, which is fully decentralized. Also, a procedure is given to determine the optimal usage of the subsystems and to account for structural changes, e.g., plug-and-play operations, so as to provide flexibility and practical fault tolerance to the scheme.

Among its main characteristics, we recall the possibility to consider shared input constraints as well as joint output constraints and the simplicity of the procedure required to obtain the simplified model at the high layer of the hierarchical structure. In addition, the algorithm allows to consider long-term objectives at the high layer, for instance, based on economic criteria, and short-term goals at the low layer, typically fast compensation of disturbances and model mismatch.

The recursive feasibility and convergence properties of the closed-loop system have been established and a simulation example has been reported to illustrate the algorithm's behavior. Future work will consider the use of multirate low-level algorithms and the application of the approach to industrial control problems.

Appendix
Proof of Theorem 1

Without lack of generality, the proof is conducted under the assumption that $\alpha_i > 0$ (i.e., $\chi_i = 1$) for all $i = 1, \dots, M$. In fact the case where, for some i, $\alpha_i = \chi_i = 0$ is trivial: the input-output pairs (u_i, y_i) are not involved in the HMPC problem (i.e., $u_i = 0$ and y_i does not concur to the output y) and, at low level, feasibility is not an issue, as described in Section III-B.

A. Recursive Feasibility of the i-LMPC Problems (24)

Assume that the problem (24) is feasible for subsystem $\mathscr{S}_i$ at time $h \in \{kN_L, \dots, (k+1)N_L - 1\}$, i.e., an optimal sequence $\upsilon_i(h|h), \dots, \upsilon_i((k+1)N_L - 1|h)$ is available, allowing to satisfy the constraints (24d), (24e) in face of input $\bar{u}(s - T_L) = \bar{u}^{[N_L]}(k - T_H)$, constant for all $s = kN_L, \dots, (k+1)N_L - 1$. At time h, the input value $\upsilon_i(h|h)$ is applied, and the remaining sequence $\upsilon_i(h+1|h), \dots, \upsilon_i((k+1)N_L - 1|h)$ is feasible at time h, since the problem is a shrinking-horizon one.

Note that, at time $h = (k+1)N_L$ (i.e., at the beginning of the subsequent high-level sampling time), the state $x_i^{(w)}((k+1)N_L)$ enjoys (24e). To guarantee the recursive feasibility, for any input $\bar{u}(l - T_L) = \bar{u}^{[N_L]}(k + 1 - T_H)$ (constant for all $l = (k+1)N_L, \dots, (k+2)N_L - 1$) there must exist an input sequence $\vec{\upsilon}_i(k+1) = \upsilon_i((k+1)N_L), \dots, \upsilon_i((k+2)N_L - 1)$ such that $w_i((k+2)N_L) = 0$ and $x_i^{(w)}((k+2)N_L) \in \mathbb{X}_i^F$. Therefore, we require the existence of $\vec{\upsilon}_i(k+1)$ such that, at the same time

$$
\begin{aligned}
&C_i(A_i^{N_L} x_i^{(w)}((k+1)N_L) + \alpha_i \bar{R}_i \bar{u}^{[N_L]}(k+1-T_H) \\
&\qquad\qquad + R_i^{(\upsilon)} \vec{\upsilon}_i(k+1)) = 0 \\
&A_i^{N_L} x_i^{(w)}((k+1)N_L) + \alpha_i \bar{R}_i \bar{u}^{[N_L]}(k+1-T_H) \\
&\qquad\qquad + R_i^{(\upsilon)} \vec{\upsilon}_i(k+1) \in \mathbb{X}_i^F.
\end{aligned}
$$

Note that if $x_i^{(w)}((k+1)N_L) \in \mathbb{X}_i^F$, in view of (26b), then $A^{N_L} x_i^{(w)}((k+1)N_L) \in \lambda_i \mathbb{X}_i^F$. Thanks to (26c), there exists a sequence $\vec{\upsilon}_i(k+1) \in (\mathscr{V}_i)^{N_L}$ such that both (24d) and (24e) can be verified for all inputs $\bar{u}^{[N_L]}(k+1-T_H) \in \bar{\mathscr{U}}$.

B. Recursive Feasibility and Convergence of the HMPC Problem (17)-(21)

Thanks to the recursive feasibility properties of the low-level problems (24), it is possible to guarantee that $w_i(kN_L) = 0$ for all $i = 1, \dots, M$ and for all $k \geq 0$. Thanks to this, it is possible to describe the evolution of variable $y^{[N_L]}(k)$ using the FIR model (14).

To show the recursive feasibility of the HMPC problem, we assume that a solution to (17)–(21) is available at time t, i.e., $\bar{u}^{[N_L]}(t|t), \dots, \bar{u}^{[N_L]}(t + N_H - 1|t), r(t)$, satisfying (17), (19)–(21). Here, $y^{[N_L]}(k)$, $k = t, \dots, t + N_H$, is obtained through (14) with inputs $\bar{u}^{[N_L]}(k|t)$. It is easy to see that, at time $t+1$, the sequence $\bar{u}^{[N_L]}(t+1|t+1) = \bar{u}^{[N_L]}(t+1|t), \dots, \bar{u}^{[N_L]}(t + N_H - 1|t+1) = \bar{u}^{[N_L]}(t + N_H - 1|t)$, $\bar{u}^{[N_L]}(t + N_H|t+1) = \bar{u}^{[N_L]}(t + N_H - 1|t)$, $r(t+1) = r(t)$ is admissible (but possibly not optimal).

Therefore, the optimal value taken by the cost function J_{HL} at time $t+1$ (denoted $J_{HL}^*(t+1)$) satisfies $J_{HL}^*(t+1) \leq \sum_{k=0}^{N_H-1}\{\|y^{[N_L]}(t+1+k) - r(t+1)\|_{Q_y}^2 + \|\bar{u}^{[N_L]}(t+1+k|t+1) - \bar{u}^{[N_L]}(t+k|t+1)\|_R^2\} + \gamma\|r(t+1) - y_{\mathrm{ref}}\|^2 = \sum_{k=1}^{N_H-1}\{\|y^{[N_L]}(t+k) - r(t)\|_{Q_y}^2 + \|\bar{u}^{[N_L]}(t+k|t) - \bar{u}^{[N_L]}(t+k-1|t)\|_R^2\} + \gamma\|r(t) - y_{\mathrm{ref}}\|^2$ since $\bar{u}^{[N_L]}(t+N_H|t+1) = \bar{u}^{[N_L]}(t+N_H-1|t+1) = \bar{u}^{[N_L]}(t+N_H-1|t)$ and since, in view of (17), $y^{[N_L]}(t+N_H) = r(t+1) = r(t)$. In view of this, it follows that $J_{HL}^*(t+1) \leq J_{HL}^*(t) - \{\|y^{[N_L]}(t) - r(t)\|_{Q_y}^2 + \|\bar{u}^{[N_L]}(t) - \bar{u}^{[N_L]}(t-1)\|_R^2\}$. Therefore, as $t \to +\infty$, a steady-state condition (where $\bar{u}^{[N_L]}(t) = \bar{u}^{[N_L]}(t-1)$) is asymptotically achieved, characterized by $y^{[N_L]}(t) = r(t)$.

At this point, similar to [8], we show that the only closed-loop stable equilibrium point compatible with (18) is the one corresponding to $y^{[N_L]} = r = y_{\mathrm{feas.ref}}$. To do so, we consider a feasible steady-state condition with $r(t) = \hat{r} \neq y_{\mathrm{feas.ref}}$, $y^{[N_L]}(k) = \hat{r}$, and $\bar{u}^{[N_L]}(k) = (\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1}\hat{r}$ for all

$k = t, \dots, t + N_H - 1$ (assuming that this is also valid for $k = t - T_H, \dots, t - 1$), whose corresponding cost is $\hat{J}_{HL} = \gamma \|\hat{r} - y_{\text{ref}}\|^2 = \gamma \|\hat{r} - y_{\text{feas.ref}}\|^2 + \gamma \|y_{\text{feas.ref}} - y_{\text{ref}}\|^2 + 2\gamma (y_{\text{feas.ref}} - y_{\text{ref}})^T (\hat{r} - y_{\text{feas.ref}})$.

At the same time, instead of keeping the system at steady state, at time t, we can consider the following alternative solution to the high-level optimization problem: $r(t) = \tilde{r} = \lambda\hat{r} + (1 - \lambda)y_{\text{feas.ref}}$ (with $\lambda \in [0, 1)$), and $\bar{u}^{[N_L]}(k) = (\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1}\tilde{r}$ for $k = t, \dots, t + N_H - 1$. Correspondingly, $y^{[N_L]}(t) = \hat{r}$, while for $k = 1, \dots, T_H$, $y^{[N_L]}(t + k) = \hat{r} + (\sum_{l=1}^{k} G_l^{[N_L]})(\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1}(\tilde{r} - \hat{r})$ and, for $k > T_H$, $y^{[N_L]}(t + k) = \tilde{r}$. Note that, for a value of λ sufficiently close to 1, this alternative solution is always feasible. The corresponding cost is $\tilde{J}_{HL} = \sum_{k=0}^{T_H-1}\{\|(\sum_{l=k+1}^{T_H} G_l^{[N_L]})(\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1}(\tilde{r} - \hat{r})\|_{Q_y}^2\} + \|(\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1}(\tilde{r} - \hat{r})\|_R^2 + \gamma \|\tilde{r} - y_{\text{ref}}\|^2 = \|\tilde{r} - \hat{r}\|_P^2 + \gamma \|\tilde{r} - y_{\text{feas.ref}}\|^2 + \gamma \|y_{\text{feas.ref}} - y_{\text{ref}}\|^2 + 2\gamma (y_{\text{feas.ref}} - y_{\text{ref}})^T (\tilde{r} - y_{\text{feas.ref}})$, where P is defined in (16b). In view of the fact that $P \leq \gamma I_p$, then $\hat{J}_{HL} - \tilde{J}_{HL} \geq \gamma \{\|\hat{r} - y_{\text{feas.ref}}\|^2 - \|\tilde{r} - \hat{r}\|^2 - \|\tilde{r} - y_{\text{feas.ref}}\|^2 + 2(y_{\text{feas.ref}} - y_{\text{ref}})^T (\hat{r} - \tilde{r})\}$. Note that $\tilde{r} - y_{\text{feas.ref}} = \lambda(\hat{r} - y_{\text{feas.ref}})$ and $\hat{r} - \tilde{r} = (1 - \lambda)(\hat{r} - y_{\text{feas.ref}})$. In view of the latter and of the optimality of $y_{\text{feas.ref}}$ with respect to the function $\|r - y_{\text{ref}}\|^2$ in the convex feasible region

$$\begin{aligned}2(y_{\text{feas.ref}} - y_{\text{ref}})^T (\hat{r} - \tilde{r}) \\ = (1 - \lambda)2(y_{\text{feas.ref}} - y_{\text{ref}})^T (\hat{r} - y_{\text{feas.ref}}) > 0.\end{aligned}$$

From this $\hat{J}_{HL} - \tilde{J}_{HL} \geq \gamma \|\hat{r} - y_{\text{feas.ref}}\|^2(1 - \lambda^2 - (1 - \lambda)^2) = 2\gamma \lambda(1 - \lambda)\|\hat{r} - y_{\text{feas.ref}}\|^2$. This means that any steady-state different from the one corresponding to $r(t) = y_{\text{feas.ref}}$ is not compatible with the optimality of the problem (18). Since a steady state is eventually attained, this implies that $r(t) \to y_{\text{feas.ref}}$ as $t \to +\infty$. This, in turn, implies that $y^{[N_L]}(t) \to y_{\text{feas.ref}}$ and $\bar{u}^{[N_L]}(t) \to \bar{u}_{\text{feas.ref}} = (\sum_{l=1}^{T_H} G_l^{[N_L]})^{-1} y_{\text{feas.ref}}$ as $t \to +\infty$.

Since constraints (19) and (20) are respected at all $k \geq 0$, then $y^{[N_L]}(k) = y(kN_L) \in \mathscr{Y}$ and $\bar{u}^{[N_L]}(k) \in \bar{\mathscr{U}}$ for all $k \geq 0$. Furthermore, since the constraint (24c) is verified at low level, then, from (9a), (9b), and (25)–(31), for all $i = 1, \dots, M$ and for all $h \geq 0$, $u_i(h) = \alpha_i \bar{u}^{[N_L]}(\lfloor(h/N_L)\rfloor) + \upsilon_i(h) \in \mathscr{U}_i$ and $\sum_{i\in\sigma} u_i(h) = \sum_{i\in\sigma} \alpha_i \bar{u}^{[N_L]}(\lfloor(h/N_L)\rfloor) + \sum_{i\in\sigma} \upsilon_i(h) \in \mathscr{U}_{\text{shared}}$, as required for satisfying (5) and (6), respectively, for all $h \geq 0$.

C. Convergence of the LMPC Problem (24)

The evolution of variable $x_i^{(\omega)}(h)$ on a N_L-steps sampling time is

$$\begin{aligned}x_i^{(\omega)}((k + 1)N_L) &= A_i^{N_L} x_i^{(\omega)}(kN_L) + R_i^{(\upsilon)}\overrightarrow{\upsilon}_i(k) \\ &\quad + \alpha_i \bar{R}_i \bar{u}^{[N_L]}(k - T_H) \\ \omega_i(kN_L) &= C_i x_i^{(\omega)}(kN_L). \qquad (34)\end{aligned}$$

Now, define $\epsilon_i^x(k) = x_i^{(\omega)}(kN_L) - \alpha_i g_i^x \bar{u}^{[N_L]}(k - T_H)$, where $g_i^x = (I_{n_i} - A_i)^{-1}(A_i^{T_L} B_i - B_i g_i^{[T_L]})$. Also, define $\epsilon_i^u(k) = \overrightarrow{\upsilon}_i(k) + \mathbb{1} \otimes (\alpha_i g_i^{[T_L]}\bar{u}^{[N_L]}(k - T_H))$. Under this change of coordinates, we can rewrite (34) as

$$\begin{aligned}\epsilon_i^x(k + 1) &= A_i^{N_L}\epsilon_i^x(k) + R_i^{(\upsilon)}\epsilon_i^u(k) + z_i(k) \\ \bar{\omega}_i(k) &= \omega_i(kN_L) = C_i\epsilon_i^x(k) \qquad (35)\end{aligned}$$

where $z_i(k) = \alpha_i g_i^x(\bar{u}^{[N_L]}(k - T_H) - \bar{u}^{[N_L]}(k + 1 - T_H))$ can be accounted for as a (vanishing) disturbance. In view of constraint (24d), $\bar{\omega}_i(k) = 0$ for all $k \geq 0$. We can write $\bar{\omega}_i(k) = \omega_i^{FREE}(k) + \omega_i^{FORCED}(k) = 0$, where $\omega_i^{FREE}(k)$ and $\omega_i^{FORCED}(k)$ are the free and forced, respectively, motions of variable $\bar{\omega}_i(k)$. Therefore $\omega_i^{FORCED}(k) = -\omega_i^{FREE}(k) = -C_i A_i^{k(N_L)}\epsilon_i^x(0) \to 0$ as $k \to +\infty$ in view of Assumption 1. Similar to Proposition 3 in [17], if $C_i(zI_{n_i} - A_i^{N_L})^{-1}$ has no zeros on the unitary circle (i.e., thanks to Assumption 3), then

$$\lim_{k\to+\infty} \left(R_i^{(\upsilon)}\epsilon_i^u(k) + z_i(k)\right) = 0.$$

Consequently, $\lim_{k\to+\infty} \epsilon_i^x(k) = 0$ (i.e., $\lim_{k\to+\infty} x_i^{(\omega)}(kN_L) = \alpha_i g_i^x \bar{u}_{\text{feas.ref}}$) follows from the asymptotic stability of (35) and from the fact that $\lim_{k\to+\infty} \bar{u}^{[N_L]}(k) = \bar{u}_{\text{feas.ref}}$. This also implies that $\lim_{h\to+\infty} \upsilon_i(h) + \alpha_i g_i^{[T_L]}\bar{u}_{\text{feas.ref}} = 0$, in view of the fact that, when $x_i^{(\omega)}(kN_L) = \alpha_i g_i^x \bar{u}_{\text{feas.ref}}$, in stationary conditions $\upsilon_i(h) = -\alpha_i g_i^{[T_L]}\bar{u}_{\text{feas.ref}}$ (which is feasible thanks to (26d) and (26e)) minimizes the cost function (24b). This, in turns, implies that $\lim_{h\to+\infty} \omega_i(h) = 0$ and that $\lim_{h\to+\infty} y(h) = \bar{y}_{\text{feas.ref}}$.

References

[1] J. Bendtsen, K. Trangbaek, and J. Stoustrup, "Hierarchical model predictive control for resource distribution," in *Proc. 49th IEEE Conf. Decis. Control (CDC)*, Dec. 2010, pp. 2468–2473.

[2] S. Conti, R. Nicolosi, S. A. Rizzo, and H. H. Zeineldin, "Optimal dispatching of distributed generators and storage systems for MV islanded microgrids," *IEEE Trans. Power Del.*, vol. 27, no. 3, pp. 1243–1251, Jul. 2012.

[3] W. Findeisen, F. N. Bailey, M. Bryds, K. Malinowski, P. Tatjewski, and A. Wozniak, *Control and Coordination in Hierarchical Systems*. Hoboken, NJ, USA: Wiley, 1980.

[4] R. Halvgaard, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Thermal storage power balancing with model predictive control," in *Proc. Eur. Control Conf.*, Jul. 2013, pp. 2567–2572.

[5] R. Halvgaard, L. Vandenberghe, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Distributed model predictive control for smart energy systems," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1675–1682, May 2016.

[6] V. Havlena and J. Lu, "A distributed automation framework for plant-wide control, optimisation, scheduling and planning," *Proc. IFAC*, vol. 38, no. 1, pp. 232–246, 2005.

[7] M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, "Multi-parametric toolbox 3.0," in *Proc. Eur. Control Conf. (ECC)*, Jul. 2013, pp. 502–510.

[8] D. Limon, I. Alvarado, T. Alamo, and E. F. Camacho, "MPC for tracking piecewise constant references for constrained linear systems," *Automatica*, vol. 44, no. 9, pp. 2382–2387, 2008.

[9] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Conf. Robot. Autom.*, Sep. 2004, pp. 284–289.

[10] J. Z. Lu, "Closing the gap between planning and control: A multiscale MPC cascade approach," *Annu. Rev. Control*, vol. 40, pp. 3–13, 2015.

[11] R. Martí, D. Sarabia, D. Navia, and C. de Prada, "A method to coordinate decentralized NMPC controllers in oxygen distribution networks," *Comput. Chem. Eng.*, vol. 59, pp. 122–137, Dec. 2013.

[12] D. Q. Mayne, M. M. Seron, and S. V. Raković, "Robust model predictive control of constrained linear systems with bounded disturbances," *Automatica*, vol. 41, no. 2, p. 219–224, 2005.

[13] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical, Multilevel, Systems*. New York, NY, USA: Academic, 1970.

[14] H. Morais, P. Kádár, P. Faria, Z. A. Vale, and H. M. Khodr, "Optimal scheduling of a renewable micro-grid in an isolated load area using mixed-integer linear programming," *Renew. Energy*, vol. 35, pp. 151–156, Jan. 2010.

[15] P. D. Moroşan, R. Bourdais, D. Dumur, and J. Buisson, “A distributed MPC strategy based on benders’ decomposition applied to multi-source multi-zone temperature regulation,” *J. Process Control*, vol. 21, pp. 729–737, Jun. 2011.
[16] B. Picasso, D. De Vito, R. Scattolini, and P. Colaneri, “An MPC approach to the design of two-layer hierarchical control systems,” *Automatica*, vol. 46, pp. 823–831, May 2010.
[17] B. Picasso, X. Zhang, and R. Scattolini, “Hierarchical model predictive control of independent systems with joint constraints,” *Automatica*, vol. 74, pp. 99–106, Dec. 2016.
[18] S. Riverso, “Distributed and plug-and-play control for constrained systems,” Ph.D. dissertation, Dept. Elect., Comput., Biomed. Eng., Univ. Pavia, Pavia, Italy, 2013.
[19] R. Scattolini, “Architectures for distributed and hierarchical model predictive control—A review,” *J. Process Control*, vol. 19, no. 5, pp. 723–731, 2009.
[20] J. Stoustrup, “Plug & play control: Control technology towards new challenges,” *Eur. J. Control*, vol. 15, nos. 3–4, pp. 311–330, 2009.
[21] C. Vermillion, A. Menezes, and I. Kolmanovsky, “Stable hierarchical model predictive control using an inner loop reference model and λ-contractive terminal constraint sets,” *Automatica*, vol. 50, pp. 92–99, Jan. 2014.
[22] T. Zabaiou, L.-A. Dessaint, and P. Brunelle, “Development of a new library of IEEE excitation systems and its validation with PSS/E,” in *Proc. IEEE Power Energy Soc. General Meet.*, Jul. 2012, pp. 1–8.

Marcello Farina received the master’s degree in electronic engineering and the Ph.D. degree in information engineering from the Politecnico di Milano, Milan, Italy, in 2003 and 2007, respectively.

He was Visiting Student with the Institute for Systems Theory and Automatic Control, Stuttgart, Germany, in 2005. He is currently an Associate Professor with the Dipartimento di Elettronica, Informazione, e Bioingegneria, Politecnico di Milano. His current research interests include distributed, hierarchical and decentralized state estimation and control, stochastic model predictive control, and applications, including mobile robots, sensor networks, and energy supply systems.

Xinglong Zhang was born in Bozhou, China, in 1990. He received the B.E. degree from Zhejiang University, Hangzhou, China, in 2011, the M.S. degree in mechanical engineering from the PLA University of Science and Technology, Nanjing, China, in 2014, and the Ph.D. degree in system and control from the Politecnico di Milano, Milan, Italy, in 2018.

He is currently an Assistant Professor with the National University of Defense Technology, Changsha, China. His current research interests include multivariable optimization, hierarchical and distributed model predictive control of large-scale systems.

Riccardo Scattolini was born in Milan, Italy, in 1956.

He is a Full Professor of automatic control with the Politecnico di Milano, Milan. From 1984 to 1985, he was a Visiting Researcher with Oxford University, Oxford, U.K. He was involved in industry on the simulation and control of chemical plants. He has authored more than 100 papers published in the main international journals of control, and of more than 150 papers presented at the main international conferences. His current research interests include modeling, identification, simulation, diagnosis, and control of industrial plants and electrical systems, with emphasis on the theory and applications of model predictive control and fault detection methods.

Mr. Scattolini was a recipient of the Quazza Premium and the Heaviside Premium of the Institution of Electrical Engineers, U.K. He has been an Associate Editor of the *IFAC Journal Automatica* and the *International Journal of Adaptive Control and Signal Processing*.