Hierarchical Distributed Model Predictive Control of Interconnected Microgrids

C. A. Hans, P. Braun, J. Raisch, L. Grüne, and C. Reincke-Collon

Abstract—In this work, we propose a hierarchical distributed model predictive control strategy to operate interconnected microgrids (MGs) with the goal of increasing the overall infeed of renewable energy sources. In particular, we investigate how renewable infeed of MGs can be increased by using a transmission network allowing the exchange of energy. To obtain an model predictive control scheme which is scalable with respect to the number of MGs and preserves their independent structure, we make use of the alternating direction method of multipliers leading to local controllers communicating through a central entity. This entity is in charge of the power lines and ensures that the constraints on the transmission capacities are met. The results are illustrated in a numerical case study.

I. INTRODUCTION

Motivated by environmental, political, economic and technological aspects, electric power systems worldwide are undergoing large changes due to an increasing number of renewable energy sources [1]. Since many of these are small-scale distributed units connected to the grid via AC inverters, the power generation structure is moving from large, centralized plants to smaller, distributed units. Thus, strategies to operate future power systems in a safe and efficient way are needed [2].

In this context, the microgrid (MG) concept has recently attracted increasing attention. An MG gathers a combination of generation units, loads and energy storage elements into a system that can be controlled by a local operator [3]. To increase security of supply, MGs can be operated in a completely isolated manner from the main transmission grid if failures occur [4]. Motivated by existing control strategies in conventional power systems, a hierarchical approach has also been advocated for MGs (see, e.g., [5]). One typically distinguishes primary and secondary control (as in conventional power systems), while the top control level, which is mostly referred to as operation control or energy management, comprises tertiary control and scheduling in conventional power systems. The operation control level aims at maximizing infeed from renewable energy sources (RES) as, once installed, they provide power at a relatively low price and very small

C. A. Hans is with the Control Systems Group, Technische Universität Berlin, Germany, hans@control.tu-berlin.de.

P. Braun is with the Mathematical Institute, Universität Bayreuth, Germany and the School of Electrical Engineering and Computing, University of Newcastle, Australia, philipp.braun@newcastle.edu.au.

J. Raisch is with the Control Systems Group, Technische Universität Berlin, Germany and the Systems and Control Theory Group, Max-Planck-Institut für Dynamik komplexer technischer Systeme, Magdeburg, Germany, raisch@control.tu-berlin.de.

L. Grüne is with the Mathematical Institute, Universität Bayreuth, Germany, lars.gruene@uni-bayreuth.de.

C. Reincke-Collon is with Younicos AG, Berlin, Germany, reincke-collon@younicos.com.

negative ecological effects. Another objective of this layer is to minimize conventional thermal generation. A way to determine setpoints for operation control is by solving an optimization problem with these objectives while accounting for limits on power and energy [6]. A widely chosen strategy to achieve this is model predictive control (MPC), which allows to include forecasts and to plan the operation for several hours ahead.

Recent results for grid-connected operation control include mixed-integer MPC where the predictions of load and RES were assumed to be certain over the prediction horizon (see, e.g., [7]). Furthermore, in [8], [9] scenario-based MPC approaches for the grid-connected operation of MGs were proposed. Others concentrate on MGs in islanded mode, e.g., [10], [11] where a mixed integer program (MIP) is solved, assuming that RES and load predictions are certain. Furthermore, in [12], a mixed-integer MPC assuming bounded uncertain load and RES and in [13] a scenario-based approach were proposed. However, all these strategies focus on the operation of single MGs and do not consider possible benefits of connecting them.

For the operation of interconnected MGs several approaches have been proposed. For instance, in [14] a distributed optimization, based on the alternating direction method of multipliers (ADMM) was introduced. It was assumed that the power flow on every transmission line can be fully controlled. For the coordination of so called energy hubs a distributed algorithm was illustrated in [15]. Both approaches assume that the local controllers of the MGs or energy hubs communicate with each other to find an optimal operation regime. However, this does not comply with the widely present division of responsibilities for the local MGs and the transmission network.

To fill this gap, hierarchical approaches that tolerate the different responsibilities of the utilities used in the operation of interconnected MGs were presented. In [16], a conceptual scheme where each MG is operated by its own controller was proposed. On a higher hierarchical layer, a distribution management system coordinates the operation of different MGs that act on price signals. For optimal power flow and for charging electric vehicles similar schemes were introduced, respectively, in [17] and [18]. Furthermore, in [19], [20], [21] centralized and hierarchical distributed optimization algorithms for MPC of residential energy systems including storage devices with focus on scalability and flexibility with respect to the network size were proposed and investigated.

Even though some of the approaches allow for a division of responsibilities for the different utilities, they do not fully cover all aspects of the operation of complex interconnected MGs. For example, the approach presented by [17] concentrates on optimal power flow and does not include storage devices or a possible limitation of RES infeed. The focus of

[18], [20], [21] is less on complex interconnected MGs with high share of RES. They aim more at solutions for charging electric vehicles [18] or residential energy systems [20], [21], where RES cannot be limited and it is not necessary to include the power flow over transmission lines in the operation regime.

We aim to fill this gap by proposing a hierarchical distributed scheme that allows to optimize the operation of multiple interconnected MGs. Based on ideas from [20], [21], we use the ADMM (see, e.g., [22], [23]) to decompose and solve an optimization problem in a distributed way. On the MG layer, local optimization problems, similar to the ones considered in [7], [12], are solved. The resulting power provided to, or from, each MG is then reported to a central entity on a higher layer, which processes the data before it communicates results to the MGs again. In this context the ADMM allows to conceal the local cost functions from the central entity, as the individual MGs only communicate the power that they want to provide to the grid. Thus, MGs maintain autonomy in their own operation but can benefit from the transmission network. Since every MG considered in the overall optimization problem retains its local controller, flexibility and scalability in this approach is much better than in central approaches. Furthermore, the proposed scheme integrates well into the configuration of many present power systems, where the responsibilities for each MG and the transmission network lie within different entities. In a case study, four interconnected MGs are operated using the proposed scheme. The gain from interconnecting the MGs is assessed by comparing the simulation results to a case where all MGs are operated in islanded mode.

The focus of this paper is on distributed optimization, highlighting the flexibility to adapt to changing network structure and additionally taking privacy aspects in the information exchange between the MGs into account. Therefore, the model of the uncertain load and RES infeed is kept relatively simple. Throughout the paper we use a certainty equivalence approach that assumes perfect forecasts in the optimization and a naïve forecast of expected values in the case study. However, the model used is supposed to serve as an example and can be replaced by others, e.g., based on [8], [9], [12], [13]. More precisely, presuming that the interface with the central entity remains the same and the local optimizations are convex, different uncertainty models, e.g., scenario trees or sample average approximation and different MG structures can be used.

The remainder of this paper is structured as follows. In Section II we discuss the generic problem of operating a network of interconnected MGs. Then, in Section III a model for the local MGs and the electrical network connecting the MGs is proposed. Section IV introduces the costs for the operation of the network of MGs. In Section V the hierarchical distributed optimization algorithm is introduced and embedded into an MPC scheme. Finally, the use of the algorithm is illustrated in Section VI.

II. GENERAL DESCRIPTION OF A NETWORK OF INTERCONNECTED MGS AND NOTATIONS

Before we give a detailed description of the interconnected MGs and the distributed MPC algorithm, the general system and some notation are introduced in this section.
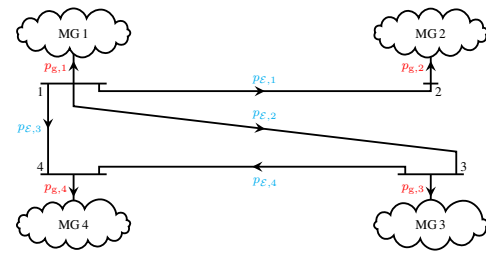

Fig. 1. Exemplary network of MGs connected by transmission lines.

Throughout this paper, we consider a network of interconnected MGs, which are able to share power over transmission lines. An example of such a network with four MGs is visualized in Figure 1. We assume that the MGs can operate independently from each other and thus, that the dynamics of a single MG are described by a discrete-time dynamical system

$$x_i(k+1) = f_i(x_i(k), u_i(k)), \tag{1a}$$
$$p_i(k) = g_i(u_i(k), w_i(k)). \tag{1b}$$

Here, $f_i : \mathbb{R}^{S_i} \times \mathbb{R}^{M_i} \to \mathbb{R}^{S_i}$ defines the dynamics of the system and $x_i(k) \in \mathbb{X}_i \subset \mathbb{R}^{S_i}$ describes the internal states of charge of the storage units. Furthermore, $u_i(k) \in \mathbb{U}_i \subset \mathbb{R}^{M_i}$ denotes the input, $p_i(k) \in \mathbb{R}^{Q_i}$ the power and $w_i(k) \in \mathbb{R}^{R_i}$ the exogenous signals at time $k \in \mathbb{N}$ of MG $i \in \mathcal{I}$, where $\mathcal{I} = \{1, \dots, I\} \subset \mathbb{N}$ is the set of all MGs.

The power contribution of the individual parts of a single MG, denoted by $p_i(k)$, is described by $g_i : \mathbb{R}^{M_i} \times \mathbb{R}^{R_i} \to \mathbb{R}^{Q_i}$. The vector $p_i(k)$ includes the scalar $p_{g,i}(k)$ which describes the power injected into or taken from the transmission network. To ensure global power balance, the constraint

$$\sum_{i=1}^{I} p_{g,i}(k) = 0 \tag{2}$$

needs to be satisfied for all $k \in \mathbb{N}$. Moreover, depending on the transmission lines, $p_g(k) = (p_{g,1}(k), \dots, p_{g,I}(k))^\top$ is bounded, i.e., we assume that $p_g(k) \in \mathbb{P}_\mathcal{E} \subset \mathbb{R}^I$ holds and the set $\mathbb{P}_\mathcal{E}$ is defined according to the capacities of the transmission lines. Equation (2) also motivates the use of a global controller in the case that the MGs want to exchange power.

In addition, we use the following notation. We define the set of real nonnegative numbers $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} | x \geq 0\}$, and the sets $\mathbb{R}_{\leq 0}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{<0}$ and $\mathbb{N}_{\geq 0}$ in a similar manner. The cardinality of a set $\mathcal{A}$ is $|\mathcal{A}|$. For a sequence $x^l \in \mathbb{R}^n$, $l \in \mathbb{N}$, the notation $(x^l)_{l \in \mathbb{N}} \subset \mathbb{R}^n$ is used. Furthermore, $\mathbf{0}_{nm}$ is the $n \times m$ matrix of all zeros and $\mathbf{1}_{nm}$ the matrix of all ones with the same size. Likewise, the identity matrix of size $n \times n$ is $\mathbf{I}_n$. Let $\mathbf{0}_n$ be the n-dimensional column vector of all zeros and $\mathbf{1}_n$ the vector of all ones with the same size. The diagonal matrix with entries a_i, $i = 1, \dots, n$, is denoted by $\mathrm{diag}(a_1, \dots, a_n)$.

III. SYSTEM MODEL

In this section we model the behavior of a single MG and the power flow over transmission lines. Before starting with the MG model, some assumptions are posed.

Throughout this work, it is assumed that the units of all MGs can run in an autonomous way for several minutes. Hence, providing power setpoints to the units with a sampling time
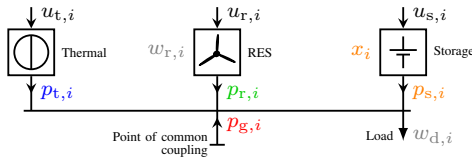
Fig. 2. Simple grid connected microgrid.

on the same time scale is sufficient. In addition, the startup and shutdown time of the thermal generators is assumed to be negligible compared to the sampling time of the MPC. For simplicity, no minimum on- or off-times of these units are assumed. This can be easily abandoned later by adding additional MIP constraints or using switching costs (see, e.g., [13]). Package loss and failure of communication lines will not be taken into account. Compared to the sampling time the communication delays are small, and thus will be neglected.

The voltage amplitudes of the nodes in the grid are assumed to be constant and the voltage phase angle differences small. Also, the lines are assumed to be purely inductive. Hence, the linearized DC power flow model can be used to describe the power flowing over the transmission lines. The error caused by the linearized model is assumed to be negligible compared to the uncertainty introduced by load and renewable generation. Furthermore, as reactive power flow is not considered, limits of the lines are only considered in an approximate way. The underlying control layers are assumed to ensure that the storage devices cover all the variations inside the local MGs. If a MG incorporates no storage units, the thermal units are assumed to cover all variations. Thus, the power provided to, or from, an MG is assumed to follow the value determined by the operational control.

A. Model of a single microgrid

In the model considered, we assume that an MG consists a number of units which can be classified in four categories: thermal generators, storage devices, RES and loads. An MG with one of each of these units is visualized in Figure 2.

Storage devices can be described by the state of charge $x_i(k) \in \mathbb{R}_{\geq 0}$ at time $k \in \mathbb{N}$ and an input variable $u_{\mathrm{s},i}(k) \in \mathbb{R}$ capturing the power supplied to or taken from the storage over a certain amount of time $\mathrm{T_s} \in \mathbb{R}_{>0}$, i.e.,

$$x_i(k+1) = x_i(k) - \mathrm{T_s} u_{\mathrm{s},i}(k). \tag{3}$$

Clearly, the power supplied to or taken from the storage, i.e., $u_{\mathrm{s},i}(k) \in \mathbb{R}$, is the power taken from or supplied to the MG, denoted by $p_{\mathrm{s},i}(k) \in \mathbb{R}$, i.e.,

$$p_{\mathrm{s},i}(k) = u_{\mathrm{s},i}(k). \tag{4a}$$

The state of charge as well as the power is bounded by

$$\underline{x}_i \leq x_i(k) \leq \overline{x}_i, \qquad \underline{p}_{\mathrm{s},i} \leq u_{\mathrm{s},i}(k) \leq \overline{p}_{\mathrm{s},i}, \tag{4b}$$

with $\underline{x}_i \in \mathbb{R}_{\geq 0}$, $\overline{x}_i \in \mathbb{R}_{>0}$, $\underline{p}_{\mathrm{s},i} \in \mathbb{R}_{<0}$ and $\overline{p}_{\mathrm{s},i} \in \mathbb{R}_{>0}$.

If a thermal generator is switched on, it can produce the power $p_{\mathrm{t},i}(k) \in \mathbb{R}_{\geq 0}$, which is bounded by lower and upper limits $0 < \underline{p}_{\mathrm{t},i} < \overline{p}_{\mathrm{t},i}$. If the thermal generator is switched off,

it does not contribute to the power flow of the MG. Thus, the power of a thermal generator can be described by

$$p_{\mathrm{t},i}(k) = u_{\mathrm{t},i}(k) \tag{4c}$$

and the bounds

$$\underline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}(k) \leq u_{\mathrm{t},i}(k) \leq \overline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}(k), \quad \delta_{\mathrm{t},i}(k) \in \{0,1\}. \tag{4d}$$

Here, $\delta_{\mathrm{t},i}(k)$ denotes a binary input variable. If $\delta_{\mathrm{t},i}(k) = 0$, then $u_{\mathrm{t},i}(k)$ and consequently also $p_{\mathrm{t},i}(k)$ is forced to zero. If $\delta_{\mathrm{t},i}(k) = 1$, then the power generated by the thermal generator can be set by $u_{\mathrm{t},i}(k)$ in the interval $[\underline{p}_{\mathrm{t},i}, \overline{p}_{\mathrm{t},i}]$.

The power $p_{\mathrm{r},i}(k) \in \mathbb{R}_{\geq 0}$ generated by RES, like wind turbines or solar generators, is limited due to the size of the plants and thus bounded by certain values $\underline{p}_{\mathrm{r},i} \in \mathbb{R}_{\geq 0}$, $\overline{p}_{\mathrm{r},i} \in \mathbb{R}_{\geq 0}$. Within these bounds the value of $p_{\mathrm{r},i}(k)$ can be controlled if weather conditions allow. The latter are captured in the signal $w_{\mathrm{r},i}(k)$, which represents the expected value of infeed under actual weather conditions at time k. By denoting the control variable, i.e., the desired infeed, by $u_{\mathrm{r},i}(k)$, the actual power infeed of the RES can be described by

$$p_{\mathrm{r},i}(k) = u_{\mathrm{r},i}(k), \text{ and} \tag{4e}$$

$$\underline{p}_{\mathrm{r},i} \leq u_{\mathrm{r},i}(k) \leq \min\{\overline{p}_{\mathrm{r},i}, w_{\mathrm{r},i}(k)\}. \tag{4f}$$

A load $w_{\mathrm{d},i}(k) \in \mathbb{R}_{\geq 0}$ only impacts the local power balance

$$0 = p_{\mathrm{r},i}(k) + p_{\mathrm{t},i}(k) + p_{\mathrm{s},i}(k) + p_{\mathrm{g},i}(k) + w_{\mathrm{d},i}(k) \tag{4g}$$

that has to be satisfied at every time step k. Here $p_{\mathrm{g},i}(k)$ denotes the power exchange with the transmission network. This power is limited by

$$\underline{p}_{\mathrm{g},i} \leq p_{\mathrm{g},i}(k) \leq \overline{p}_{\mathrm{g},i}. \tag{4h}$$

Remark III.1. Note that in the local power balance (4g) we assumed that the MG contains exactly one of each of the units introduced in this section. This is also the case for the MG visualized in Figure 2. However, this setting can be easily extended to MG with multiple or zero units of a certain type by including/excluding the corresponding power in (4g). For the remainder of this paper, we assume that every MG consists of exactly one thermal generator, one storage device, one RES and one load, for simplicity.

The equations derived in this section can be rewritten as

$$f_i(x_i, u_i) = x_i - \mathrm{T_s} u_{\mathrm{s},i},$$
$$g_i(u_i, w_i) = (u_{\mathrm{r},i}, u_{\mathrm{t},i}, u_{\mathrm{s},i}, -u_{\mathrm{r},i} - u_{\mathrm{t},i} - u_{\mathrm{s},i} - w_{\mathrm{d},i})^\top,$$
$$p_i = (p_{\mathrm{t},i}, p_{\mathrm{s},i}, p_{\mathrm{r},i}, p_{\mathrm{g},i})^\top,$$
$$u_i = (\delta_{\mathrm{t},i}, u_{\mathrm{t},i}, u_{\mathrm{s},i}, u_{\mathrm{r},i})^\top, \text{ and } w_i = (w_{\mathrm{r},i}, w_{\mathrm{d},i})^\top.$$

Additionally, with the definitions $\mathbb{X}_i = [\underline{x}_i, \overline{x}_i]$ and

$$\mathbb{U}_i(k) = \big((0,0) \cup (1, [\underline{p}_{\mathrm{t},i}, \overline{p}_{\mathrm{t},i}])\big)$$
$$\times [\underline{p}_{\mathrm{s},i}, \overline{p}_{\mathrm{s},i}] \times [\underline{p}_{\mathrm{r},i}, \min\{\overline{p}_{\mathrm{r},i}, w_{\mathrm{r},i}(k)\}]$$

the constraints $x_i \in \mathbb{X}_i$ and $u_i \in \mathbb{U}_i(k)$ are defined. Note that the constraints $\mathbb{U}_i(k)$ are time dependent due to $w_{\mathrm{r},i}(k)$.

B. Transmission network

The individual MGs are connected by transmission lines. We model the grid connecting the MGs as a weighted, undirected, connected graph $\mathcal{G} = (\mathcal{I}, \mathcal{E}, y)$. Every MG is represented by one node in the set $\mathcal{I} = \{1, \dots, I\}$. Additionally, $\mathcal{E} = \{E_1, \dots, E_{|\mathcal{E}|}\}$ is the set of vertices of $\mathcal{G}$, i.e., the transmission lines and $y : \mathcal{E} \to \mathbb{R}_{>0}$ is a weighting function assigning a susceptance to every line. We first choose an arbitrary direction for every edge $E_j \in \mathcal{E}$. The edge-node incidence matrix $\mathbf{F}$ of the resulting directed graph is defined element wise as

$$\mathbf{F}_{i,j} = \begin{cases} -1 & \text{if node } i \text{ is the sink of edge } E_j, \\ 1 & \text{if node } i \text{ is the source of edge } E_j, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Power flow over the lines in AC networks can be described in a nonlinear way (see, e.g., [24]). Assuming small differences in voltage amplitude and phase angles, as well as lossless lines, we can use the linearized DC power flow equations instead [25]. With these, the power flowing over the lines, $p_{\mathcal{E}}(k) \in \mathbb{R}^{|\mathcal{E}|}$, can be derived by

$$p_{\mathcal{E}}(k) = \mathrm{diag}(y_1, \dots, y_{|\mathcal{E}|})\mathbf{F}^\top \Theta(k), \tag{5}$$

where $y_j = y(E_j)$ is the susceptance of line E_j and $\Theta(k) = (\theta_1(k), \dots, \theta_I(k))^\top$ are the phase angles at the nodes. The power taken from or supplied to the transmission grid at the individual nodes is given by $p_{\mathrm{g}}(k) = \mathbf{F}p_{\mathcal{E}}(k)$. Therefore,

$$p_{\mathrm{g}}(k) = \mathbf{Y}\Theta(k), \tag{6}$$

where $\mathbf{Y} = \mathbf{F}\,\mathrm{diag}(y_1, \dots, y_{|\mathcal{E}|})\mathbf{F}^\top$ is the weighted Laplacian of $\mathcal{G}$. To consider the power flow over the lines, we need to derive $p_{\mathcal{E}}(k)$ from $p_{\mathrm{g}}(k)$. $\mathbf{Y}$ is symmetric and has rank $I-1$ as $\mathcal{G}$ is connected. To calculate $p_{\mathcal{E}}(k)$ as a linear function of $p_{\mathrm{g}}(k)$ we thus use the transformation

$$\begin{pmatrix} \tilde{\Theta}(k) \\ \theta_I(k) \end{pmatrix} = \begin{pmatrix} \theta_1(k) - \theta_I(k) \\ \vdots \\ \theta_I(k) \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{I}_{I-1} & -\mathbf{1}_{I-1} \\ \mathbf{0}_{I-1}^\top & 1 \end{pmatrix}}_{\mathbf{T}} \Theta(k). \tag{7}$$

Denoting $(p_{\mathrm{g},1}(k), \dots, p_{\mathrm{g},I-1}(k))^\top$ by $\tilde{p}_{\mathrm{g}}(k)$ and combining (6) and (7) yields

$$\begin{pmatrix} \tilde{p}_{\mathrm{g}}(k) \\ p_{\mathrm{g},I}(k) \end{pmatrix} = \mathbf{Y}\mathbf{T}^{-1}\begin{pmatrix} \tilde{\Theta}(k) \\ \theta_I(k) \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{Y}} & \mathbf{0}_{I-1} \\ b^\top & 0 \end{pmatrix}\begin{pmatrix} \tilde{\Theta}(k) \\ \theta_I(k) \end{pmatrix}, \tag{8}$$

where $\tilde{\mathbf{Y}} = \tilde{\mathbf{T}}\mathbf{Y}\tilde{\mathbf{T}}^\top$ with $\tilde{\mathbf{T}} = (\mathbf{I}_{I-1}, \mathbf{0}_{(I-1)})$ and $b^\top = (\mathbf{0}_{(I-1)}^\top, 1)\mathbf{Y}\tilde{\mathbf{T}}^\top$. Note that $\tilde{\mathbf{Y}}$ is invertible (see [26]). Inserting (7) and $\tilde{p}_{\mathrm{g}}(k) = \tilde{\mathbf{Y}}\tilde{\Theta}(k)$ into (5) yields

$$p_{\mathcal{E}}(k) = \mathrm{diag}(y_1, \dots, y_{|\mathcal{E}|})\mathbf{F}^\top \mathbf{T}^{-1}\begin{pmatrix} \tilde{\mathbf{Y}}^{-1}\tilde{p}_{\mathrm{g}}(k) \\ \theta_I(k) \end{pmatrix}. \tag{9a}$$

Additionally, the last row of (8),

$$p_{\mathrm{g},I}(k) = b^\top \tilde{\mathbf{Y}}^{-1}\tilde{p}_{\mathrm{g}}(k), \tag{9b}$$

needs to be satisfied.

Observe that $b^\top \tilde{\mathbf{Y}}^{-1} = -\mathbf{1}_{I-1}^\top$. This can be easily seen by using $b^\top = -\mathbf{1}_{I-1}^\top\tilde{\mathbf{Y}}$ and inserting $\tilde{\mathbf{Y}} = \tilde{\mathbf{T}}\mathbf{Y}\tilde{\mathbf{T}}^\top$ which results in

$$(\mathbf{0}_{(I-1)}^\top, 1)\mathbf{Y}\tilde{\mathbf{T}}^\top = -\mathbf{1}_{I-1}^\top\tilde{\mathbf{T}}\mathbf{Y}\tilde{\mathbf{T}}^\top.$$

With $\tilde{\mathbf{T}} = (\mathbf{I}_{I-1}, \mathbf{0}_{(I-1)})$ this becomes

$$\left((\mathbf{0}_{(I-1)}^\top, 1) + (\mathbf{1}_{I-1}^\top, 0)^\top\right)\mathbf{Y}\tilde{\mathbf{T}}^\top = \mathbf{0}_{I-1}^\top.$$

As $\mathcal{G}$ is connected, $\mathbf{Y}$ has a simple zero eigenvalue and corresponding right eigenvector $\mathbf{1}_I$. Hence, $\mathbf{1}_I\mathbf{Y} = \mathbf{0}_I$ and (9b) is equivalent to (2).

As power flow only depends on the phase angle differences, we can freely choose $\theta_I(k) = 0, \ \forall \ k \in \mathbb{N}$. Then, (9) can be expressed with $\tilde{b} = (-b^\top\tilde{\mathbf{Y}}^{-1}, 1)^\top = \mathbf{1}_I$ and $\tilde{\mathbf{F}} = \mathrm{diag}(y_1, \dots, y_{|\mathcal{E}|})\mathbf{F}^\top\mathbf{T}^{-1}\tilde{\mathbf{T}}^\top\tilde{\mathbf{Y}}^{-1}\tilde{\mathbf{T}} \in \mathbb{R}^{|\mathcal{E}|\times I}$ as

$$p_{\mathcal{E}}(k) = \tilde{\mathbf{F}}p_{\mathrm{g}}(k), \quad \text{and} \quad 0 = \tilde{b}^\top p_{\mathrm{g}}(k) = \mathbf{1}_I^\top p_{\mathrm{g}}(k). \tag{10}$$

The maximum power that can be transmitted over the lines is limited by the bounds $\overline{p}_{\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|}$, $\underline{p}_{\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|}$, i.e.,

$$\underline{p}_{\mathcal{E}} \leq p_{\mathcal{E}}(k) \leq \overline{p}_{\mathcal{E}}.$$

With $\underline{p}_{\mathrm{g}} = (\underline{p}_{\mathrm{g},1}, \dots, \underline{p}_{\mathrm{g},I})^\top$, $\overline{p}_{\mathrm{g}} = (\overline{p}_{\mathrm{g},1}, \dots, \overline{p}_{\mathrm{g},I})^\top$ and

$$\mathbb{P}_{\mathcal{E}} = \{p_{\mathrm{g}} \in [\underline{p}_{\mathrm{g}}, \overline{p}_{\mathrm{g}}] \,|\, \underline{p}_{\mathcal{E}} \leq \tilde{\mathbf{F}}p_{\mathrm{g}} \leq \overline{p}_{\mathcal{E}} \ \wedge \ 0 = \mathbf{1}_I^\top p_{\mathrm{g}}\} \tag{11}$$

we obtain the constraints for the power exchange, $p_{\mathrm{g}}(k) \in \mathbb{P}_{\mathcal{E}}$.

IV. OPERATIONAL COSTS

The costs to operate an MG at time k, i.e., to allocate the correct amount of energy, depends on the devices used to provide the power. We assume that the operation costs can be expressed in the variables u_i and $p_{\mathrm{g},i}$, which for all $i \in \mathcal{I}$ will be abbreviated by the variable $z_i(k) = (u_i(k)^\top, p_{\mathrm{g},i}(k))^\top$. Then the stage cost $\ell_i : \mathbb{R}^{M_i} \times \mathbb{R} \to \mathbb{R}$ of MG i is composed of

$$\ell_i(z_i) = \ell_{\mathrm{t},i}(z_i) + \ell_{\mathrm{r},i}(z_i) + \ell_{\mathrm{s},i}(z_i) + \ell_{\mathrm{g},i}(z_i). \tag{12a}$$

Here, motivated by [27], the cost of the thermal units $\ell_{\mathrm{t},i}(z_i(k)) \in \mathbb{R}$ is approximated by

$$\ell_{\mathrm{t},i}(z_i(k)) = c_{\mathrm{t},i}\delta_{\mathrm{t},i}(k) + c'_{\mathrm{t},i}u_{\mathrm{t},i}(k) + c''_{\mathrm{t},i}(u_{\mathrm{t},i}(k))^2, \tag{12b}$$

with the weights $c_{\mathrm{t},i} \in \mathbb{R}_{>0}$, $c'_{\mathrm{t},i} \in \mathbb{R}_{>0}$ and $c''_{\mathrm{t},i} \in \mathbb{R}_{>0}$.

Once installed, renewable units can provide energy at a very low price compared to other units. Even in times of low load it is advantages to store infeed from RES for later use in times of high load or low infeed. Thus, it is desirable to emphasize RES infeed and chose $\ell_{\mathrm{r},i}(z_i(k)) \in \mathbb{R}$ with $c_{\mathrm{r},i} \in \mathbb{R}_{>0}$ as

$$\ell_{\mathrm{r},i}(z_i(k)) = c_{\mathrm{r},i}(\overline{p}_{\mathrm{r},i} - u_{\mathrm{r},i}(k))^2. \tag{12c}$$

For simplicity, costs depending on the state of charge of storage device are not considered. However, depending on the efficiency of a storage unit, there are losses based on the usage that usually grow quadratically in power. With $c_{\mathrm{s},i} \in \mathbb{R}_{>0}$, costs due to conversion losses of storage units are included as

$$\ell_{\mathrm{s},i}(z_i(k)) = c_{\mathrm{s},i}(u_{\mathrm{s},i}(k))^2. \tag{12d}$$

Finally, the exchange of power with the transmission grid is included in the cost function by

$$\ell_{\mathrm{g},i}(z_i(k)) = c'_{\mathrm{g},i} p_{\mathrm{g},i}(k) + c''_{\mathrm{g},i}|p_{\mathrm{g},i}(k)|. \tag{12e}$$

Here, the first part of the equation, with $c'_{\mathrm{g},i} \in \mathbb{R}_{>0}$ assumes a certain price for energy. The second part, with $c''_{\mathrm{g},i} \in \mathbb{R}_{>0}$ represents a fixed cost per absolute value of traded power.

In addition to the costs assumed for every MG, costs for transmitting power are also taken into account. Assuming that line losses increase quadratically with the transmitted power and using $\tilde{\mathbf{F}}$ from Section III-B, they can be modeled as

$$\ell_{\mathcal{E}}(p_{\mathrm{g}}(k)) = (\tilde{\mathbf{F}}p_{\mathrm{g}}(k))^\top \mathbf{C}_{\mathcal{E}}(\tilde{\mathbf{F}}p_{\mathrm{g}}(k)) \tag{13}$$

with the positive definite real matrix $\mathbf{C}_{\mathcal{E}}$.

With $z(k) = (z_1(k)^\top, \dots, z_I(k)^\top)^\top$, the overall costs are

$$\ell(z(k)) = \ell_{\mathcal{E}}(p_{\mathrm{g}}(k)) + \sum_{i=1}^{I} \ell_i(z_i(j)). \tag{14}$$

V. MPC FOR A NETWORK OF INTERCONNECTED MGS

In Section III, we introduced the model of a grid consisting of independent MGs connected through transmission lines. The objective in the grid operation is the minimization of the costs (14). The MPC scheme used to achieve this is discussed in the first part of this section. In the second part we present a distributed optimization algorithm based on the ADMM to solve the optimization problem involved in the MPC scheme at every time step on the level of the individual MGs and a central entity, which corresponds to the transmission grid operator.

A. Optimal control problem and model predictive control

The time dependent exogenous signals $w_i(k)$, $i \in \mathcal{I}$ can only be predicted some time into the future with sufficient accuracy. Thus, we use an MPC scheme to compute the input $u_i(k)$ and thereby the power infeed $p_i(k)$ of all MGs at time instant $k \in \mathbb{N}$ in a receding fashion using a fixed prediction horizon $K \in \mathbb{N}_{>0}$. To emphasize immediate costs with lower uncertainty over far future costs with higher uncertainty we use a discount factor $\gamma \in (0, 1]$. To simplify notation, we use

$$\mathbf{u}_i(k) = (u_i(k)^\top, \dots, u_i(k+K-1)^\top)^\top, \text{ and}$$
$$\mathbf{z}_i(k) = (\mathbf{u}_i(k)^\top, p_{\mathrm{g},i}(k)^\top, \dots, p_{\mathrm{g},i}(k+K-1))^\top.$$

With the state measurement, $x_i^k \in \mathbb{R}_{\geq 0}^{S_i}$, and forecasts of the uncertain values, $w_i(j) \in \mathbb{R}^{R_i}$, at time $k \in \mathbb{N}$ for $j = k, \dots, k+K-1$, the predicted minimal costs to operate the network over the next K time steps can be obtained from the solution of the following optimization problem.

Problem 1. Mixed integer quadratic problem

$$\min_{\mathbf{u}_1(k), \dots, \mathbf{u}_I(k)} \sum_{j=k}^{k+K-1} \gamma^{(j-k)}\left(\ell_{\mathcal{E}}(p_{\mathrm{g}}(k)) + \sum_{i=1}^{I} \ell_i(z_i(j))\right)$$

s.t. $\quad x_i(k) = x_i^k, \; x_i(j+1) \in \mathbb{X}_i, \; u_i(j) \in \mathbb{U}_i(j), \tag{15}$
$$x_i(j+1) = f_i(x_i(j), u_i(j)), \tag{16}$$
$$p_i(j) = g_i(u_i(j), w_i(j)), \tag{17}$$
$$p_{\mathrm{g}}(j) \in \mathbb{P}_{\mathcal{E}} \tag{18}$$
$$\forall j = k, \dots, k+K-1, \; \forall i \in \mathcal{I}$$

Assuming that $\mathbb{X}_i$, $\mathbb{U}_i$ and $\mathbb{P}_{\mathcal{E}}$ are defined such that Problem 1 is feasible, in an MPC scheme, the problem is solved at every time step k to obtain an optimal solution $\mathbf{u}_i^\star(k)$ for every $i \in \mathcal{I}$. From this, $u_i^\star(k)$ is applied as a feedback in MG i. This leads to the next state measurement at time $k+1$ and a new optimization problem, which is solved to repeat the process.

Problem 1 is a mixed integer quadratic program (MIQP) due to the binary decision variables $\delta_{\mathrm{t},i}(k) \in \{0, 1\}$. Even though the MIQP is non-convex, efficient algorithms, e.g., branch-and-bound, exist to find an optimal solution of the problem. However, for the distributed ADMM investigated in the next section, a convex problem is needed to ensure convergence. Thus, we consider a second optimization problem, a relaxation of Problem 1. To this end, we relax the constraints (4d) by allowing $\delta_{\mathrm{t},i}(k)$ in the continuous interval $[0, 1]$, i.e.,

$$\underline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}(k) \leq u_{\mathrm{t},i}(k) \leq \overline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}(k), \quad \delta_{\mathrm{t},i}(k) \in [0, 1] \tag{19}$$

i.e., the constraint set $\mathbb{U}_i(k)$ becomes

$$\mathbb{U}_i^c(k) = \{(\delta_{\mathrm{t},i}, u_{\mathrm{t},i})|\delta_{\mathrm{t},i} \in [0, 1], u_{\mathrm{t},i} \in [\underline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}, \overline{p}_{\mathrm{t},i}\delta_{\mathrm{t},i}]\}$$
$$\times [\underline{p}_{s,i}, \overline{p}_{s,i}] \times [\underline{p}_{r,i}, \min\{\overline{p}_{r,i}, w_{r,i}(k)\}].$$

Using the constraint set $\mathbb{U}_i^c(k)$ instead of $\mathbb{U}_i(k)$ leads to a quadratic convex optimization problem.

Problem 2. Quadratic convex optimization problem

$$\min_{\mathbf{u}_1(k), \dots, \mathbf{u}_I(k)} \sum_{j=k}^{k+K-1} \gamma^{(j-k)}\left(\ell_{\mathcal{E}}(p_{\mathrm{g}}(j)) + \sum_{i=1}^{I} \ell_i(z_i(j))\right)$$

s.t. $\quad x_i(k) = x_i^k, \; x_i(j+1) \in \mathbb{X}_i, \; u_i(j) \in \mathbb{U}_i^c(j), \tag{20}$
$$\text{and } (16)\text{–}(18)$$
$$\forall j = k, \dots, k+K-1, \; \forall i \in \mathcal{I}.$$

Remark V.1. The only difference between Problem 2 and Problem 1 is in the constraint sets $\mathbb{U}_i^c(j)$ and $\mathbb{U}_i(j)$. $\mathbb{U}_i^c(j)$ is convex and the resulting Problem 2 is a convex problem. Since it is a relaxation of Problem 1, one cannot expect that their solutions coincide. However, to ensure that (4d) holds, Problem 2, is only used to determine p_{g}. After p_{g} is determined, local MIQPs can be solved to obtain a feasible solution of Problem 1 (see Algorithm 1). An exemplary quantification of the error introduced by the relaxation can be found in Section VI.

B. Hierarchical distributed optimization

To maintain autonomy of the individual MGs we use the ADMM to solve Problem 2 in a distributed way by exchanging information on p_{g} between the MGs and an additional central entity responsible for the transmission lines. For details on the ADMM see, e.g, [22], [23], and the references therein. Here, we only review the main ideas necessary for our application.

To solve Problem 2 using the ADMM, we introduce copies of $p_{\mathrm{g}}(j)$ denoted by $\hat{p}_{\mathrm{g}}(j)$ for $j = k, \dots, k+K-1$,

$$\hat{\mathbf{p}}_{\mathrm{g}}(k) = (\hat{p}_{\mathrm{g}}(k), \dots, \hat{p}_{\mathrm{g}}(k+K-1)^\top)^\top$$

and add the additional redundant constraints

$$p_{\mathrm{g}}(j) - \hat{p}_{\mathrm{g}}(j) = 0 \tag{21}$$

to Problem 2. With this definition and with a fixed parameter $\rho > 0$, the augmented Lagrangian of Problem 2 is given by

$$\mathcal{L}\big(\mathbf{z}_i(k), \hat{\mathbf{p}}_g(k), \Lambda(k)\big) =$$
$$\sum_{j=k}^{k+K-1} \Big(\gamma^{(j-k)}\big(\ell_{\mathcal{E}}(\hat{p}_g(j)) + \sum_{i=1}^{I} \ell_i(z_i(j))\big)$$
$$+ \sum_{i=1}^{I} \big(\lambda_i(j)(p_{g,i}(j) - \hat{p}_{g,i}(j)) + \tfrac{\rho}{2}\|p_{g,i}(j) - \hat{p}_{g,i}(j)\|_2^2\big)\Big).$$

Here, $\lambda(j) = (\lambda_1(j), \ldots, \lambda_I(j))^\top$ for $j = k, \ldots, k+K-1$ and $\Lambda(k) = (\lambda(k)^\top, \ldots, \lambda(k+K-1)^\top)^\top$ denote the Lagrange multipliers corresponding to the equality constraints (21).

The decoupled structure of the constraints of Problem 2 allows for an iterative solution of the optimization problem by sequentially performing the following updates for an arbitrary initialization $(\mathbf{z}_i^0(k), \hat{\mathbf{p}}_g^0(k), \Lambda^0(k))$. The individual MGs solve the decoupled optimization problems

$$\mathbf{z}_i^{l+1}(k) \in \operatorname*{argmin}_{\mathbf{z}_i(k)} \sum_{j=k}^{k+K-1} \big(\gamma^{(j-k)}\ell_i(z_i(j)) \qquad (22)$$
$$+ \lambda_i^l(j)p_{g,i}(j) + \tfrac{\rho}{2}\|p_{g,i}(j) - \hat{p}_{g,i}^l(j)\|_2^2\big)$$
$$\text{s.t.} \quad (16), (17) \text{ and } (20) \quad \forall j = k, \ldots, k+K-1.$$

in parallel for fixed $\hat{\mathbf{p}}_g^l(k)$ and $\Lambda^l(k)$. The updates $p_{g,i}^{l+1}(j)$, $j = k, \ldots, k+K-1$, extracted from $\mathbf{z}_i^{l+1}(k)$, are sent to the central entity, which then solves the optimization problem

$$\hat{\mathbf{p}}_g^{l+1}(k) \in \operatorname*{argmin}_{\hat{\mathbf{p}}_g(k)} \sum_{j=k}^{k+K-1} \big(\gamma^{(j-k)}\ell_{\mathcal{E}}(\hat{p}_g(j)) \qquad (23)$$
$$- \hat{p}_g(j)^\top \lambda(j)^l + \tfrac{\rho}{2}\|p_g^{l+1}(j) - \hat{p}_g(j)\|_2^2\big)$$
$$\text{s.t.} \quad \hat{p}_g(j) \in \mathbb{P}_{\mathcal{E}} \quad \forall j = k, \ldots, k+K-1.$$

and afterwards updates the Lagrange multipliers

$$\lambda_i^{l+1}(j) = \lambda_i^l + \rho(p_{g,i}^{l+1}(j) - \hat{p}_{g,i}^{l+1}(j)) \qquad (24)$$
$$\forall j = k, \ldots, k+K-1, \forall i \in \mathcal{I}.$$

For $j = k, \ldots, k+K-1$, $\hat{p}_{g,i}^{l+1}(j)$ and $\lambda_i^{l+1}(j)$ are communicated to the MGs and the steps are repeated for $l = l + 1$. For the scheme, the following convergence properties hold.

Theorem V.2. Let $(\mathbf{z}^l, \hat{\mathbf{p}}_g^l, \Lambda^l)_{l \in \mathbb{N}}$ be a sequence generated by iteratively computing the updates (22)–(24) for arbitrary initial values $(\mathbf{z}^0, \hat{\mathbf{p}}_g^0, \Lambda^0)$. Then the following holds.

(i) The sequence $(\mathbf{z}^l, \hat{\mathbf{p}}_g^l, \Lambda^l)_{l \in \mathbb{N}}$ is bounded, and every limit point of $(\mathbf{z}^l)_{l \in \mathbb{N}}$ is an optimal solution of Problem 2.

(ii) The sequence

$$\left(\sum_{j=k}^{k+K-1} \gamma^{(j-k)}\big(\ell_{\mathcal{E}}(\hat{p}_g^l(j)) + \sum_{i=1}^{I} \ell_i(z_i^l(j))\big)\right)_{l \in \mathbb{N}}$$

converges to the optimal value of Problem 2.

Proof. See [22, Proposition 4.2.] or [23, Chapter 3.2] for a general proof for convex optimization problems. □

Remark V.3 (Flexibility and privacy)**.** One main property of the ADMM scheme is that the essential optimization step is performed autonomously by each MG. By decomposing the global optimization problem in the proposed way, we obtain a control scheme that is highly flexible with respect to the local MGs since only the variable $p_{g,i}$ is shared with the network. The same flexibility holds for the topology and the

constraints of the transmission network which does not directly affect the individual MGs. Similarly, the cost function of MG i only needs to be known to MG i itself and the topology of the transmission network only needs to be known to the central entity implying the possibility of maintaining privacy of data within the scheme. The central entity can even change its update steps (23) and (24), without giving notice to the individual MGs, e.g., if the topology of the transmission grid changes or if an MG is connected to or disconnected from the grid. In summary, as only the power infeed of the MGs and the Lagrange multipliers are communicated, high flexibility and privacy of the different entities can be achieved.

C. Hierarchical distributed model predictive control algorithm

---
Algorithm 1 Distributed model predictive control algorithm
---
(1) **Initialization** at time $k \in \mathbb{N}$:
   - **MG** For $i, \ldots, I$, measure $x_i(k) = x_i^k$
   - **Central Entity**: Define grid constraints $\mathbb{P}_{\mathcal{E}}$
(2) **Main Loop (ADMM):** For $l = 1, \ldots, l_{\max} - 1 \in \mathbb{N}$
   (i) **MGs:** For $i = 1, \ldots, I$, solve (22) in parallel and send $p_{g,i}^{l+1}(k), \ldots, p_{g,i}^{l+1}(k+K-1)$ to the central entity.
   (ii) **Central Entity:**
      - Solve (23)
      - Update the Lagrange multipliers (24)
      - Communicate $\hat{p}_{g,i}(k), \ldots, \hat{p}_{g,i}(k+K-1)$ and $\lambda_i(k), \ldots, \lambda_i(k+K-1)$ to the MGs.
(3) **Mixed integer update:**
   For $l = l_{\max}$, for all MGs $i = 1, \ldots, I$
      - Solve (22) where $\mathbb{U}_i^c(j)$ is replaced by $\mathbb{U}_i(j)$ and the constraints $p_{g,i}(j) - \hat{p}_{g,i}^{l_{\max}}(j) = 0$ are added for $j = k, \ldots, k+K-1$.
(4) **Application** of $u_i^{l_{\max}+1}(k)$, increment k and go to (1).

At every time instance $k \in \mathbb{N}$ the relaxed optimization Problem 2 is solved in a distributed way using the ADMM updates (22)–(24) in step (2) of Algorithm 1 by performing a fixed number of iterations, $l_{\max} \in \mathbb{N}$, or until a stopping criterion, e.g., based on the difference of $(\mathbf{z}_i^l(k), \hat{\mathbf{p}}_g^l(k), \Lambda^l(k))$ and $(\mathbf{z}_i^{l-1}(k), \hat{\mathbf{p}}_g^{l-1}(k), \Lambda^{l-1}(k))$ is satisfied. Then, to obtain a feasible solution of the original Problem 1, a single MIQP is solved in every MG in step (3). Here, the constraints $p_{g,i}(j) - \hat{p}_{g,i}^{l_{\max}}(j) = 0$ are added to Problem 1 which implies that $p_g^{l_{\max}+1}(j) = \hat{p}_g^{l_{\max}}(j) \in \mathbb{P}_{\mathcal{E}}$ for all $j = k, \ldots, k+K-1$. The resulting solution is again denoted by $\mathbf{u}_i^\star(k)$. From this $u_i^\star(k)$ is extracted and applied in the standard MPC fashion.

Remark V.4 (Computational complexity)**.** Note that the complexity of the local optimization problems (22) in steps (2) and (3) in Algorithm 1 only depends on the prediction horizon and the number of devices in the MG, but not on the size (i.e., the number of MGs) of the overall network. The number of iterations in step (2) that are required to obtain a sufficiently accurate result, given forecasts of load and RES $w_i(k)$ with limited certainty, is usually in the order of 10 to 100 and only weakly depends on the number of MGs [23]. Furthermore,

the number of iterations can be reduced significantly in the context of MPC by using so-called warm-start techniques, see, e.g., [28, Section 10.5] or [20, Section V.A].

The numerical complexity of MIQPs can grow exponentially with the number of unknowns. Therefore, the complexity of Algorithm 1 highly depends on the MIQPs solved in step (3). Fortunately, in our case these are of the same complexity as the MIQPs solved in islanded mode. Since step (3) can be performed by each MG in parallel, the interconnection of MGs does not increase the numerical complexity of the single MIQPs. Furthermore, due to the potentially exponential growth in complexity, step (3) is computationally advantageous compared to a central solution of Problem 1 which can become hard to solve for a high number of MGs.

In general, it is possible that the MIQP (step (3) in Algorithm 1) is infeasible. This could occur if $u_{\mathrm{t},i}^{l_{\max}}(j) \in (0, \underline{p}_{\mathrm{t},i})$ holds for at least one $i \in \mathcal{I}$ and one $j \in \{k, \ldots, K + k - 1\}$.

Nevertheless, it is still possible to define feasible inputs $u_{\mathrm{t},i}(k)$ and $u_{\mathrm{s},i}(k)$ to go to the next time step $k + 1$ if

$$\underline{p}_{\mathrm{s},i} \leq -2\underline{p}_{\mathrm{t},i}, \quad \overline{p}_{\mathrm{s},i} \geq 2\underline{p}_{\mathrm{t},i}, \tag{25a}$$

$$(\overline{x}_i - \underline{x}_i) \geq 4\mathrm{T}_s\underline{p}_{\mathrm{t},i} \quad \text{and} \quad x_i^k \in \mathbb{X}_i \tag{25b}$$

holds for all $i \in \mathcal{I}$ and infeasibility is encountered in the MIP update in step (3) of Algorithm 1.

In this case, if $u_{\mathrm{t},i}^{l_{\max}}(k) \in \{0\} \cup [\underline{p}_{\mathrm{t},i}, \overline{p}_{\mathrm{t},i}]$ for MG $i \in \mathcal{I}$, i.e., the infeasibility only occurs in $k + 1, \ldots, k + K - 1$, then $u_{\mathrm{t},i}^{l_{\max}}(k)$ can be applied in step (4) of Algorithm 1 as a feasible control input. Thus we can concentrate on the case $u_{\mathrm{t},i}^{l_{\max}}(k) \in (0, \underline{p}_{\mathrm{t},i})$.

If one of the conditions

$$u_{\mathrm{s},i}^{l_{\max}}(k) \leq -\underline{p}_{\mathrm{t},i} \quad \text{or} \quad \begin{cases} -\underline{p}_{\mathrm{t},i} \leq u_{\mathrm{s},i}^{l_{\max}}(k) \leq \underline{p}_{\mathrm{t},i} \\ x_i^k > \frac{1}{2}(\underline{x}_i + \overline{x}_i) \end{cases} \tag{26}$$

holds, then we can define the input

$$u_{\mathrm{s},i}(k) = u_{\mathrm{s},i}^{l_{\max}}(k) + u_{\mathrm{t},i}^{l_{\max}}(k), \tag{27a}$$

$$u_{\mathrm{t},i}(k) = 0 \tag{27b}$$

to ensure that (4b) holds, i.e., that the state of charge $x_i(k+1)$ lies in $\mathbb{X}_i$ and that the local power balance (4g) is is satisfied.

If one of the conditions

$$u_{\mathrm{s},i}^{l_{\max}}(k) \geq \underline{p}_{\mathrm{t},i} \quad \text{or} \quad \begin{cases} -\underline{p}_{\mathrm{t},i} \leq u_{\mathrm{s},i}^{l_{\max}}(k) \leq \underline{p}_{\mathrm{t},i} \\ x_i^k \leq \frac{1}{2}(\underline{x}_i + \overline{x}_i) \end{cases} \tag{28}$$

is satisfied, we can define

$$u_{\mathrm{s},i}(k) = u_{\mathrm{s},i}^{l_{\max}}(k) + u_{\mathrm{t},i}^{l_{\max}}(k) - \underline{p}_{\mathrm{t},i}, \tag{29a}$$

$$u_{\mathrm{t},i}(k) = \underline{p}_{\mathrm{t},i} \tag{29b}$$

to ensure feasibility at time step $k + 1$. By using this heuristic update rule in the case of an infeasible mixed integer update in step (3), it is ensured that a feasible input $u_i(k)$ can always be derived by every MG locally.

In the following, we will exemplarily show that the input (27) is feasible if (26) is satisfied. The other cases, i.e., feasibility of (29) if (28) holds, follows the same arguments and is therefore omitted.

Assume that $u_{\mathrm{s},i}^{l_{\max}}(k) \leq -\underline{p}_{\mathrm{t},i}$ holds. Since $u_{\mathrm{s},i}^{l_{\max}}(k)$ and $u_{\mathrm{t},i}^{l_{\max}}(k)$ result from the solution of the convex Problem 2, they satisfy the local power balance (4g). The inputs $u_{\mathrm{s},i}(k)$ and $u_{\mathrm{t},i}(k)$ defined in (27) then also satisfy (4g) and $u_{\mathrm{t},i}(k)$ is feasible. Thus,

$$u_{\mathrm{s},i}^{l_{\max}}(k) \leq u_{\mathrm{s},i}^{l_{\max}}(k) + u_{\mathrm{t},i}^{l_{\max}}(k) = u_{\mathrm{s},i}(k) \leq -\underline{p}_{\mathrm{t},i} + \underline{p}_{\mathrm{t},i} = 0,$$

together with the feasibility of $u_{\mathrm{s},i}^{l_{\max}}(k)$ shows that $u_{\mathrm{s},i}(k)$ is feasible. Using this result in the storage dynamics leads to

$$x_i^{l_{\max}}(k + 1) = x_i^k - \mathrm{T}_s u_{\mathrm{s},i}^{l_{\max}}(k),$$
$$\geq x_i^k - \mathrm{T}_s u_{\mathrm{s},i}(k) \geq x_i^k.$$

Hence, $x_i(k + 1) = x_i^k - \mathrm{T}_s u_{\mathrm{s},i}(k)$ satisfies the constraints due to the feasibility of $x_i^{l_{\max}}(k + 1)$ from Problem 2 and the feasibility of the initial state x_i^k.

If $-\underline{p}_{\mathrm{t},i} \leq u_{\mathrm{s},i}^{l_{\max}}(k) \leq \underline{p}_{\mathrm{t},i}$ and $x_i^k > \frac{1}{2}(\underline{x}_i + \overline{x}_i)$ hold, then $u_{\mathrm{s},i}(k)$ can be obtained using (27a) and satisfies

$$-\underline{p}_{\mathrm{t},i} \leq u_{\mathrm{s},i}(k) < u_{\mathrm{s},i}^{l_{\max}}(k) + \underline{p}_{\mathrm{t},i} \leq 2\underline{p}_{\mathrm{t},i}$$

which again implies that $u_{\mathrm{s},i}(k)$ is feasible if (25) holds.

For the storage dynamics, using (25b), we can show that

$$x_i^k - \mathrm{T}_s u_{\mathrm{s},i}^{l_{\max}}(k) \geq x_i^k - \mathrm{T}_s u_{\mathrm{s},i}(k)$$
$$\geq x_i^k - 2\mathrm{T}_s\underline{p}_{\mathrm{t},i}$$
$$> \frac{1}{2}(\overline{x}_i + \underline{x}_i) - \frac{1}{2}(\overline{x}_i - \underline{x}_i) = \underline{x}_i$$

which implies that $x_i(k+1) = x_i^k - \mathrm{T}_s u_{\mathrm{s},i}(k)$ is feasible given feasible $x_i^{l_{\max}}(k + 1) = x_i^k - \mathrm{T}_s u_{\mathrm{s},i}^{l_{\max}}(k)$ from Problem 2.

The two remaining cases (28) and the corresponding input (29) lead to feasibility in a similar way. Broadly speaking, (25) ensures that the limits of the storage units are large enough to always compensate for either an increase or a decrease in thermal power. Even though the bounds provided only hold for our specific case, the proposed mechanism can also be used for more complex MG models. However, in the numerical case study in Section VI, infeasibility in the mixed integer update did not occur.

VI. CASE STUDY

To illustrate the performance of Algorithm 1 a numerical case study was performed. Before presenting the results, the grid and the simulation parameters are introduced.

In the case study, the topology in Figure 3 was used. It consists of four MGs, each containing a thermal generator, a storage unit and an RES. Every MG has a load connected to its busbar and a connection to the transmission network. The limits of the units and the point of common coupling are equal for all MGs as shown in Table I. Every MG i is equipped such that in case of islanded operation the power demand, $w_{\mathrm{d},i}(k)$, can always be provided by the units inside the MG. To ensure this, every MG has a thermal generator as backup unit in case of transmission line failures or low RES infeed. A sampling time of $\mathrm{T}_s = 0.5\,\mathrm{h}$ and a prediction horizon of $K = 12$ was used to predict a possible full charge or discharge of the storage units that have a capacity of $6\,\mathrm{pu\,h}$ and a nominal power of $1\,\mathrm{pu}$. For the prediction of load and RES infeed, a
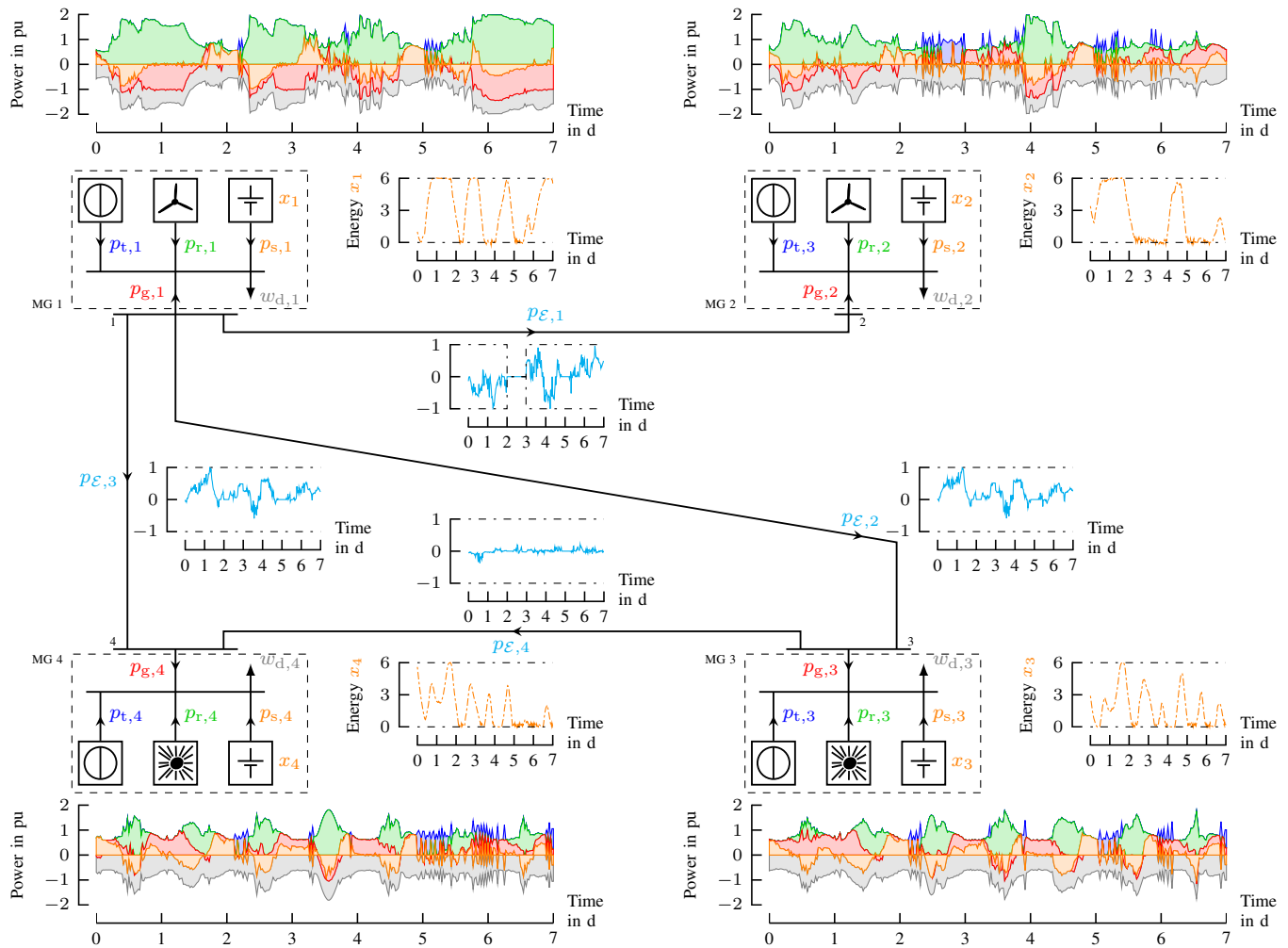
Fig. 3. Simulation topology and results.

naïve forecast (see, e.g., [29]) was used that assumes that all values in the future are equal to the ones that are currently measured. The time series for wind and irradiation were derived from [30], the load time series was emulated based on real load data. As indicated in Section III, every storage unit covers the uncertainties introduced by the imperfect forecast. In practice, the possible constraint violations caused by this, could be reduced by tightening the bounds of desired power or energy of the storage units. The weights for (12), (13) were chosen according to Table II. For the thermal generators, the costs were calculated by normalizing the values from [27] to 1 pu. The scheme was implemented in MATLAB® 2015a and YALMIP R20160930 [31] with Gurobi 7.0.2 as solver.

In Figure 3 the results of the operation using Algorithm 1 over a simulation interval of seven days is visualized. To show the flexibility of the MPC algorithm with respect to the network structure, a failure of the transmission line connecting MG 1 and MG 2 on the third day of the simulation is assumed. As pointed out in Section V, only the central entity has to change its optimization problem (23), or more precisely the set $\mathbb{P}_{\mathcal{E}}$. It can be seen that the power balance (4g) holds for all MGs and all time steps. Furthermore, the plots show that the

stored energy x_i, $i = 1, \ldots, 4$ stays inside the recommended bounds most of the time. The observed violations of the recommended bounds are due to imprecise forecasts. However, global power balance is always maintained. The plots further show that the photovoltaic units provide power during the day that is stored and provided during the night. Additionally, it can be observed that the power flow over the transmission lines $p_{\mathcal{E},i}$, $i = 1, \ldots, 4$ stays inside the bounds. Power exchange between MG 3 and 4 is close to zero almost all the time. This is probably due to similar patterns of photovoltaic infeed in MG 3 and 4. In contrast, the other lines are used extensively.

The results visualized in Figure 3 and those of simulation using different settings are summarized in Table III. In the first column the accumulated values of Figure 3 using Algorithm 1 are given. Note that in the closed-loop solution, the effect of imperfect forecasts causes an increase in costs of about 5 %.

In the second column of Table III, the central MIQP of Problem 1 is solved and directly applied to all MGs. Comparing column one and two, it can be seen that solving Problem 1 leads to a small improvement in RES and thermal infeed, as well as operation costs. However, this comes at the expense of a much higher computational complexity of

TABLE I
OPERATION LIMITS OF UNITS AND TRANSMISSION LINES ($i = 1, \ldots, 4$).

Parameter	Value
$(\underline{p}_{\mathrm{t},i}, \underline{p}_{\mathrm{r},i}, \underline{p}_{\mathrm{s},i}, \underline{p}_{\mathrm{g},i})^\top$	$(0.2, 0, -1, -1)^\top$ pu
$(\overline{p}_{\mathrm{t},i}, \overline{p}_{\mathrm{r},i}, \overline{p}_{\mathrm{s},i}, \overline{p}_{\mathrm{g},i})^\top$	$(1, 2, 1, 1)^\top$ pu
$(\underline{x}_i, \overline{x}_i)^\top$	$(0, 6)^\top$ pu h
x^0	$(1, 3.4, 2.9, 5.6)^\top$ pu h
$\underline{p}_{\mathcal{E}}$	$(-1, -1, -1, -1)^\top$ pu
$\overline{p}_{\mathcal{E}}$	$(1, 1, 1, 1)^\top$ pu
y	$(20, 20, 20, 20)^\top$ pu

TABLE III
ACCUMULATED RESULTS OF DIFFERENT SIMULATIONS.

		MG 2 one day islanded		
		local MIP	global MIP	Islanded
RES energy in pu h		404.3	404.7	314.3
Thermal energy in pu h		49.4	48.7	140.1
Costs (perfect forecast)		3136.0	3135.3	3515.2
Costs	MG 1	379.4	378.9	668.7
	MG 2	833.4	832.3	931.1
	MG 3	1021.4	1020.4	1028.9
	MG 4	1046.4	1045.9	1065.7
	Transmission	24.9	25.6	0.0
MG 1–4 & transmission		3305.5	3303.1	3694.4

the global MIQP. Furthermore, Algorithm 1 allows for more flexibility and autonomy of the MGs as well as scalability with respect to the essential optimization step.

In column three of Table III, the results using Algorithm 1 for the operation of four islanded MGs are provided. Here, no power is exchanged over the transmission lines, i.e., $\mathbb{P}_{\mathcal{E}} = \{\mathbf{0}_I\}$, and thus $p_g(k) = \mathbf{0}_I$ for all $k \in \mathbb{N}$. As to be expected, the costs of the connected operation in column one are significantly lower than in the islanded operation. This reduction results from an increase of RES infeed of 28.6 % and a decrease of thermal infeed of 64.7 % compared to the islanded case. In addition, the cost of each individual MG decreases when sharing power over the transmission network. Hence, trading energy between the MGs pays-off and every MG benefits from a collaboration. Nevertheless, there seems to be a difference in the decrease of costs for MGs equipped with wind turbines and MGs equipped with photovoltaic power plants (see Figure 3 and Table III). In the simulation, trading energy between the MGs especially pays-off, if their RES infeed differs. If all MGs were only equipped with photovoltaic units, an interconnection of the MGs might not result in such a high benefit as seen for $p_{\mathcal{E},4}$. Still, an interconnection can help decrease thermal generation and increase the share of RES.

VII. CONCLUSIONS

In this paper we proposed a hierarchical distributed control strategy based on the ADMM to operate interconnected MGs. The strategy allows to perform the essential optimization step autonomously by each individual MG. In particular, the structure and the local cost function of each individual MG does not need to be revealed to the other MGs or the central authority. It was shown that the algorithm used has the potential to increase the overall infeed from RES while maintaining flexibility in the operation of local MGs.

Next steps include testing the control scheme for a larger number of MGs and more complex MG setups to assess

how the computing time scales. Additionally, more accurate forecasts and a more comprehensive model that includes, e.g., reactive power flow and a more detailed model of the storage units, shall be employed. Furthermore, investigations for feasibility of the mixed integer update are planned. Also, the use of more complex forecasts of uncertain load and RES, e.g., robust intervals or scenario trees, shall be investigated.

REFERENCES

[1] REN21 Secretariat, "Renewables 2016 global status report," REN21, Paris, France, Tech. Rep., 2016.
[2] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," *IEEE Power Energy Mag.*, vol. 5, no. 4, pp. 78–94, 2007.
[3] R. H. Lasseter, "MicroGrids," in *IEEE PES WM*, 2002, pp. 305–308.
[4] J. A. Peças Lopes, C. L. Moreira, and A. G. Madureira, "Defining control strategies for microgrids islanded operation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 916–924, 2006.
[5] A. Bidram and A. Davoudi, "Hierarchical structure of microgrids control system," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1963–1976, 2012.
[6] F. Katiraei, R. Iravani, N. Hatziargyriou, and A. Dimeas, "Microgrids management," *IEEE Power Energy Mag.*, vol. 6, no. 3, pp. 54–65, 2008.
[7] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Trans. Control Syst. Technol.*, no. 99, 2014.
[8] ——, "Stochastic model predictive control for economic/environmental operation management of microgrids: An experimental case study," *J. Process Control*, vol. 43, pp. 24–37, 2016.
[9] S. Mohammadi, S. Soleymani, and B. Mozafari, "Scenario-based stochastic operation management of microgrid including wind, photovoltaic, micro-turbine, fuel cell and energy storage devices," *Int. J. Electr. Power Energy Syst.*, vol. 54, pp. 525–535, 2014.
[10] E. Mayhorn, K. Kalsi, M. Elizondo, W. Zhang, S. Lu, N. Samaan, and K. Butler-Purry, "Optimal control of distributed energy resources using model predictive control," in *IEEE PES GM*, 2012, pp. 1–8.
[11] D. Olivares, C. Cañizares, and M. Kazerani, "A centralized energy management system for isolated microgrids," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1864–1875, 2014.
[12] C. A. Hans, V. Nenchev, J. Raisch, and C. Reincke-Collon, "Minimax model predictive operation control of microgrids," in *IFAC WC*, 2014, pp. 10 287–10 292.
[13] C. A. Hans, P. Sopasakis, A. Bemporad, J. Raisch, and C. Reincke-Collon, "Scenario-based model predictive operation control of islanded microgrids," in *IEEE CDC*, 2015.
[14] W. Zheng, W. Wu, B. Zhang, H. Sun, Q. Guo, and C. Lin, "Dynamic economic dispatch for microgrids: A fully distributed approach," in *IEEE PES T&D*, 2016, pp. 1–3.
[15] M. Arnold, R. Negenborn, G. Andersson, and B. De Schutter, "Distributed predictive control for energy hub coordination in coupled electricity and gas networks," in *Intelligent Infrastructures*. Springer, 2010, vol. 42, pp. 235–273.
[16] A. G. Tsikalakis and N. D. Hatziargyriou, "Centralized control for optimizing microgrids operation," in *IEEE PES GM*, 2011, pp. 1–8.
[17] A. Jokić, "Price–based optimal control of electrical power systems," Ph.D. dissertation, Technische Universiteit Eindhoven, 2007.

TABLE II
WEIGHTS OF COST FUNCTIONS ($i = 1, \ldots, 4$).

Weight	Value	Weight	Value
$c_{\mathrm{t},i}$	0.1178	$c''_{\mathrm{s},i}$	0.05 $1/\mathrm{pu}^2$
$c'_{\mathrm{t},i}$	0.751 $1/\mathrm{pu}$	$c'_{\mathrm{g},i}$	0.5 $1/\mathrm{pu}^2$
$c''_{\mathrm{t},i}$	0.0048 $1/\mathrm{pu}^2$	$c''_{\mathrm{g},i}$	0.1 $1/\mathrm{pu}^2$
$c''_{\mathrm{r},i}$	1 $1/\mathrm{pu}^2$	$\mathbf{C}_{\mathcal{E}}$	$0.1 \cdot \mathrm{diag}(1, 2, 3, 6)$ $1/\mathrm{pu}^2$

[18] J. Rivera, P. Wolfrum, S. Hirche, C. Goebel, and H.-A. Jacobsen, "Alternating direction method of multipliers for decentralized electric vehicle charging control," in *IEEE CDC*, 2013, pp. 6960–6965.

[19] K. Worthmann, C. M. Kellett, P. Braun, L. Grüne, and S. R. Weller, "Distributed and decentralized control of residential energy systems incorporating battery storage," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1914–1923, 2015.

[20] P. Braun, L. Grüne, C. M. Kellett, S. R. Weller, and K. Worthmann, "A distributed optimization algorithm for the predictive control of smart grids," *IEEE Trans. Autom. Control*, pp. 3898–3911, 2016.

[21] P. Braun, T. Faulwasser, L. Grüne, C. M. Kellett, S. R. Weller, and K. Worthmann, "Hierarchical distributed ADMM for predictive control with applications in power networks," *IFAC J. Syst. Control*, vol. 3, pp. 10 – 22, 2018.

[22] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*.   Prentice-Hall, 1989.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[24] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*.   McGraw-Hill, 1994, vol. 7.

[25] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *IEEE PES GM*, 2005, pp. 454–459.

[26] J. Schiffer, R. Ortega, A. Astolfi, J. Raisch, and T. Sezi, "Conditions for stability of droop-controlled inverter-based microgrids," *Automatica*, vol. 50, no. 10, pp. 2457–2469, 2014.

[27] M. Živić Đurović, A. Milačić, and M. Kršulja, "A simplified model of quadratic cost function for thermal generators," *Ann. DAAAM 2012 Proc. 23 Int. DAAAM Symp.*, vol. 23, no. 1, pp. 25–28, 2012, costs for thermal generators.

[28] L. Grüne and J. Pannek, *Nonlinear model predictive control*.   Springer, 2011.

[29] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*.   OTexts, 2014.

[30] Atmospheric Radiation Measurement Climate Research Facility. Surface Meteorology System. June 2009–Oct. 2009, 39° 5' 28" N, 28° 1' 45" W: Eastern North Atlantic Facility, windspeed. ARM Data Archive: Oak Ridge, TN, USA. Data set accessed 2011-07-14 at www.arm.gov.

[31] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *IEEE CACSD*, 2004, pp. 284–289.