Hierarchical Distributed Model Predictive Control of Interconnected Microgrids


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 a 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 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
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 that will be 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 an extensive 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 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.

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

\[ x_i(k+1) = f_i(x_i(k), u_i(k)), \]

\[ p_i(k) = g_i(u_i(k), w_i(k)). \]

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 \( \mathcal{I} = \{1, \ldots , I\} \subset \mathbb{N} \) 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

\[ \sum_{i=1}^I p_{g,i}(k) = 0 \]

needs to be satisfied for all \( k \in \mathbb{N} \). Moreover, depending on the transmission lines, \( p_{g,i}(k) = (p_{g,1}(k), \ldots , p_{g,I}(k))^T \) is bounded, i.e., we assume that \( p_{g,i}(k) \in \mathbb{P}_E \subset \mathbb{R}^I \) holds and the set \( \mathbb{P}_E \) 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 \( \mathbb{R}_{< 0}, \mathbb{R}_{> 0}, \mathbb{R}_{< 0} \subset \mathbb{R}_{> 0} \) in a similar manner. The cardinality of a set \( A \) is \( |A| \). 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 \( \mathbf{1}_n \). Let \( a_i \), \( i = 1, \ldots , n \), be the \( n \)-dimensional column vector of all zeros and \( 1_n \) the vector of all ones with the same size. The diagonal matrix with entries \( a_i \), \( i = 1, \ldots , n \), is denoted by \( \text{diag}(a_1, \ldots , a_n) \).

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 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.

The power flow over transmission lines in the operation regime. The error caused by communication delays are small, and thus will be neglected.

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.
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 of 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 Fig. 2.

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

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

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

\[
p_{s,i}(k) = u_{s,i}(k). \quad \text{(4a)}
\]

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

\[
\underline{x}_{s,i} \leq x_i(k) \leq \overline{x}_{s,i}, \quad \underline{p}_{s,i} \leq u_{s,i}(k) \leq \overline{p}_{s,i}, \quad \text{with} \quad \underline{x}_{s,i} \in \mathbb{R}_{>0}, \quad \overline{x}_{s,i} \in \mathbb{R}_{>0}, \quad \underline{p}_{s,i} \in \mathbb{R}_{<0} \quad \text{and} \quad \overline{p}_{s,i} \in \mathbb{R}_{>0}. \tag{4b}
\]

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

\[
p_{t,i}(k) = u_{t,i}(k). \tag{4c}
\]

and the bounds

\[
\underline{p}_{t,i} \leq u_{t,i}(k) \leq \overline{p}_{t,i}, \quad \delta_{t,i}(k) \subseteq \{0, 1\}. \tag{4d}
\]

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

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

\[
p_{t,i}(k) = u_{t,i}(k), \quad \text{and} \quad \underline{p}_{t,i} \leq u_{t,i}(k) \leq \min \{ \overline{p}_{t,i}, w_{t,i}(k) \}. \tag{4f}
\]

A load \( w_{d,i}(k) \in \mathbb{R}_{\geq 0} \) only impacts the local power balance \( 0 = p_{c,i}(k) + p_{r,i}(k) + p_{s,i}(k) + p_{g,i}(k) + w_{d,i}(k) \quad \text{for} \quad k \in \mathbb{N} \)

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

\[
p_{g,i}(k) \leq \overline{p}_{g,i} \leq \min \{ p_{g,i}(k) \}. \quad \text{(4h)}
\]

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 Fig. 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 (4h).

For the remainder of this paper, we assume that every MG consists of exactly one thermal generator, one storage device, one RES unit and one load, for simplicity.

The equations derived in this section can be rewritten as

\[
f_i(x_i, u_i) = x_i - T_s u_{s,i},
\]

\[
g_i(u_i, w_i) = (u_{t,i}, u_{s,i}, u_{n,i} - u_{t,i} - u_{s,i} - w_{d,i}),
\]

\[
p_i = (p_{t,i}, p_{s,i}, p_{g,i}),
\]

\[
u_i = (\delta_{t,i}, u_{t,i}, u_{s,i}, u_{r,i}),
\]

\[
\lambda_i = (w_{c,i}, w_{d,i}).
\]

Additionally, with the definitions \( X_i = [\underline{x}_{s,i}, \overline{x}_{s,i}] \) and

\[
U_i(k) = \{(0, 0) \cup (1, [\underline{p}_{t,i}, \overline{p}_{t,i}]) \times (0, \underline{p}_{t,i}) \times \min \{ \overline{p}_{t,i}, w_{t,i}(k) \})\}
\]

the constraints \( x_i \in X_i \) and \( u_i \in U_i(k) \) are defined. Note that the constraints \( U_i(k) \) are time dependent due to \( w_{t,i}(k) \).

B. Transmission network

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

\[
F_{i,j} = \begin{cases} 
-1 & \text{if node } i \text{ is the sink of edge } E_j, \\
1 & \text{if node } i \text{ is the source of edge } E_j, \\
0 & \text{otherwise}.
\end{cases}
\]

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, \( p_{E}(k) \in \mathbb{R}_{|E|} \), can be derived by

\[
p_{E}(k) = \text{diag}(y_1, \ldots, y_{|E|}) F^T \Theta(k), \tag{5}
\]
where \( y_j = g(E_j) \) is the susceptance of line \( E_j \) and \( \Theta(k) = (\theta_1(k), \ldots, \theta_T(k))^\top \) are the phase angles at the nodes. The power taken from or supplied to the transmission grid at the individual nodes is given by \( p_g(k) = Y F p_E(k) \). Therefore,

\[
p_g(k) = Y \Theta(k),
\]

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

\[
\begin{pmatrix}
\tilde{\Theta}(k) \\
\tilde{\theta}_1(k)
\end{pmatrix} =
\begin{pmatrix}
\begin{pmatrix}
\theta_1(k) - \theta_1(k) \\
\vdots \\
\theta_T(k)
\end{pmatrix} \\
\begin{pmatrix}
1 & -1 & \cdots & -1 \\
0 & 1 & \cdots & 0
\end{pmatrix}
\end{pmatrix} \Theta(k).
\]

Denoting \((p_{g,1}(k), \ldots, p_{g,I-1}(k))\) by \( \tilde{p}_g(k) \) and combining (6) and (7) yields

\[
\begin{pmatrix}
p_{g,1}(k) \\
p_{g,2}(k) \\
\vdots \\
p_{g,I-1}(k)
\end{pmatrix} = Y \tilde{\Theta}(k) - \tilde{\theta}_1(k)
\]

where \( Y = T Y T^\top \) with \( T = (I_{I-1}, 0_{(I-1)}) \) and \( b^\top = (0_{(I-1)}, 1) Y T^\top \). Note that \( Y \) is invertible (see [24]). Inserting (7) and \( \tilde{p}_g(k) = Y \tilde{\Theta}(k) \) into (5) yields

\[
p_E(k) = \text{diag}(y_1, \ldots, y_{\|G\|}) F^\top T^{-1} \left( T^{-\top} \tilde{p}_g(k) \right),
\]

Additionally, the last row of (8),

\[
p_{g,I}(k) = b^\top \tilde{Y}^{-1} \tilde{p}_g(k),
\]

needs to be satisfied.

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

\[
(0_{(I-1)}, 1)^\top Y T^\top = -I_{I-1} Y T^\top.
\]

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

\[
(0_{(I-1)}, 1)^\top + I_{I-1,0}^\top Y T^\top = 0_{I-1}^\top.
\]

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

As power flow only depends on the angle differences, we can freely choose \( \theta_1(k) = 0, \forall k \in \mathbb{N} \). Then, (9) can be expressed with \( \tilde{b} = -b^\top \tilde{Y}^{-1} \) as

\[
p_E(k) = Y \tilde{\Theta}(k),
\]

and \( 0 = \tilde{b}^\top \tilde{p}_g(k) = I_{I-1} \tilde{p}_g(k) \). (10)

The maximum power that can be transmitted over the lines is limited by the bounds \( \underline{p}_E \in \mathbb{R}^{\|E\|} \), \( \bar{p}_E \in \mathbb{R}^{\|E\|} \), i.e.,

\[
\underline{p}_E \leq p_E(k) \leq \bar{p}_E.
\]

With \( \underline{p}_E = (\underline{p}_{g,1}, \ldots, \underline{p}_{g,I-1})^\top \), \( \bar{p}_E = (\bar{p}_{g,1}, \ldots, \bar{p}_{g,I-1})^\top \) and

\[
\underline{p}_E = \{p_E \in [\underline{p}_E, \bar{p}_E] \mid \underline{p}_E \leq F \bar{p}_E \leq \bar{p}_E \land 0 = \tilde{b}^\top \tilde{p}_g(k)\}
\]

we obtain the constraints for the power exchange, \( p_g(k) \in \mathbb{R}^{\|E\|} \).

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 \( p_{g,i} \) which for all \( i \in I \) will be abbreviated by the variable \( z_i(k) = (u_i(k)^\top, p_{g,i}(k))^\top \). Then the stage cost \( \ell_i \colon \mathbb{R}^{M_i} \times \mathbb{R} \rightarrow \mathbb{R} \) of MG \( i \) is composed of

\[
\ell_i(z_i) = \ell_i(z_i) + \ell_{t,i}(z_i) + \ell_{e,i}(z_i) + \ell_{g,i}(z_i).
\]

Here, motivated by [21], the cost of the thermal units \( \ell_{t,i}(z_i(k)) \in \mathbb{R} \) is approximated by

\[
\ell_{t,i}(z_i(k)) = c_{t,i} \delta_{t,i}(k) + c_{t,i}^\prime u_i(k) + c_{t,i}^\prime u_i(k)^2,
\]

with the weights \( c_{t,i} \in \mathbb{R}_{>0}, c_{t,i}^\prime \in \mathbb{R}_{>0} \) and \( c_{t,i}^\prime \in \mathbb{R}_{>0} \).

Once installed, RES units can provide energy at a very low price compared to other units. Even in times of low load it is advantageous 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 \( \ell_{e,i}(z_i(k)) \in \mathbb{R} \) with \( c_{e,i} \in \mathbb{R}_{>0} \) as

\[
\ell_{e,i}(z_i(k)) = c_{e,i} u_i(k)^2.
\]

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 \( c_{s,i} \in \mathbb{R}_{>0} \), costs due to conversion losses of storage units are included as

\[
\ell_{s,i}(z_i(k)) = c_{s,i} u_i(k)^2.
\]

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

\[
\ell_{E}(p_{g,i}(k)) = (F_{g,i} p_{g,i}(k))^\top C_E (F_{g,i} p_{g,i}(k))
\]

with the positive definite real matrix \( C_E \).

With \( z(k) = (z_1(k)^\top, \ldots, z_{\|G\|}(k)^\top) \), the overall costs are

\[
\ell(z(k)) = \ell_E(p_{g}(k)) + \sum_{i=1}^{\|I\|} \ell_i(z_i(j)).
\]

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 \( w_i(k), i \in I \) 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 future costs with higher uncertainty we use a discount factor $\gamma \in (0, 1]$. To simplify notation, we use

$$u_i(k) = (u_i(k)^\top, \ldots, u_i(k+K-1)^\top)^\top, \text{ and}$$

$$z_i(k) = (u_i(k)^\top, p_{g,i}(k)^\top, \ldots, p_{g,i}(k+K-1)^\top)^\top.$$

With the state measurement, $x_i^k \in \mathbb{R}^{2n}$, and forecasts of the uncertain values, $w_i(j) \in \mathbb{R}^{K_j}$, at time $k \in \mathbb{N}$ for $j = k, \ldots, k+K-1$, 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 program

$$\min_{u_i(k), \ldots, u_i(k)} \sum_{j=k}^{k+K-1} \gamma^{(j-k)} \left( \ell_E(p_{g,i}(k)) + \sum_{i=1}^{I} \ell_i(z_i(j)) \right)$$

s.t. 

$$x_i(k) = x_i^k, \quad x_i(j+1) \in \mathcal{X}_i, \quad u_i(j) \in \mathcal{U}_i(j),$$

$$p_{g,i}(j) = g_i(u_i(j), w_i(j)),$$

$$\forall j = k, \ldots, k+K-1, \forall i \in I.$$

Assuming that $\mathcal{X}_i, \mathcal{U}_i$ and $\mathcal{P}_\varepsilon$ 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 $u_i^*(k)$ for every $i \in I$. From this, $u_i^*(k)$ 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 2.** Quadratic convex optimization problem

$$\min_{u_i(k), \ldots, u_i(k)} \sum_{j=k}^{k+K-1} \gamma^{(j-k)} \left( \ell_E(p_{g,i}(j)) + \sum_{i=1}^{I} \ell_i(z_i(j)) \right)$$

s.t. 

$$x_i(k) = x_i^k, \quad x_i(j+1) = f_i(x_i(j), u_i(j)),$$

$$p_{g,i}(j) = g_i(u_i(j), w_i(j)),$$

$$\forall j = k, \ldots, k+K-1, \forall i \in I.$$

**Remark V.1.** The only difference between Problem 2 and Problem 1 is in the constraint sets $\mathcal{U}_i(j)$ and $\mathcal{U}_i(j)$. $\mathcal{U}_i(j)$ 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 $p_{g,i}$. After $p_{g,i}$ 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 $p_{g,i}$ 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 $\hat{p}_{g,i}(j)$ denoted by $\tilde{p}_{g}(j)$ for $j = k, \ldots, k+K-1$,

$$\hat{p}_{g,i}(k) = (\hat{p}_{g,i}(k), \ldots, \hat{p}_{g,i}(k+K-1))^\top$$

and add the additional redundant constraints

$$p_{g}(j) - \tilde{p}_{g}(j) = 0 \quad (21)$$

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

$$\mathcal{L}(z_i(k), \hat{p}_{g,i}(k), \Lambda(k)) = \sum_{j=k}^{k+K-1} \left( \gamma^{(j-k)} \left( \ell_E(\hat{p}_{g,i}(j)) + \sum_{i=1}^{I} \ell_i(z_i(j)) \right) + \sum_{i=1}^{I} \left( \lambda_i(j) (\hat{p}_{g,i}(j) - p_{g,i}(j)) + \rho \|p_{g,i}(j) - \hat{p}_{g,i}(j)\|_2^2 \right) \right).$$

Here, $\lambda_i(j) = (\lambda_1(j), \ldots, \lambda_{I}(j))^\top$ for $j = k, \ldots, K$. The vectors $\lambda(k)$ 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 $(z_i^0(k), \hat{p}_{g,i}^0(k), \Lambda^0(k))$. The individual MGs solve the decoupled optimization problems

$$z_i^{l+1}(k) \in \text{argmin} \sum_{j=k}^{k+K-1} \gamma^{(j-k)} \ell_i(z_i(j))$$

$$+ \lambda_i(j) p_{g,i}(j) + \frac{\rho}{2} \|p_{g,i}(j) - \hat{p}_{g,i}(j)\|_2^2 \quad (22)$$

s.t. 

$$\hat{p}_{g,i}(j)^\top = \hat{p}_{g,i}(j)$$

and $\hat{p}_{g,i}(j) \in \mathcal{P}_\varepsilon$, $j = k, \ldots, k+K-1.$

in parallel for fixed $\hat{p}_{g,i}^0(k)$ and $\lambda_i^0$. The updates $\hat{p}_{g,i}^{l+1}(j)$, $j = k, \ldots, k+K-1$, extracted from $z_i^{l+1}(k)$, are sent to the central entity, which then solves the optimization problem

$$\hat{p}_{g,i}^{l+1}(k) \in \text{argmin} \sum_{j=k}^{k+K-1} \gamma^{(j-k)} \ell_E(\hat{p}_{g,i}(j))$$

$$- \hat{p}_{g,i}(j)^\top \lambda_i(j)^l + \frac{\rho}{2} \|\hat{p}_{g,i}(j) - \hat{p}_{g,i}(j)^l\|_2^2 \quad (23)$$

s.t. 

$$\hat{p}_{g,i}(j) \in \mathcal{P}_\varepsilon \quad \forall j = k, \ldots, k+K-1.$$

and afterwards updates the Lagrange multipliers

$$\lambda_i^{l+1}(j) = \lambda_i^l + \rho (\hat{p}_{g,i}^{l+1}(j) - \hat{p}_{g,i}(j)^l) \quad (24)$$

$$\forall j = k, \ldots, k+K-1,$$

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 $\left(z', \mathbf{p}_{g}^i, \mathbf{\Lambda}^i\right)_{i \in \mathbb{N}}$ be a sequence generated by iteratively computing the updates (22)–(24) for arbitrary initial values $(z^0, \mathbf{p}_{g}^0, \mathbf{\Lambda}^0)$. Then the following holds.

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

(ii) The sequence

$$\left(\sum_{j=k}^{K-1} \gamma_{j-k}(f_E(p_{g}^j)) + \sum_{i=1}^{N} f_i(z^j_i)\right)_{i \in \mathbb{N}}$$

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. One property of the ADMM scheme is that the essential optimization step can be performed autonomously by each MG. As each MG only shares the variable $p_{g,i}$ with the network, the structure and cost function of MG $i$ only has to be known to MG $i$ itself. Furthermore, the size and complexity of the local optimization problems only depends on the prediction horizon and on the number of devices in the MG, but not on the size of the overall network. Another important aspect of the ADMM scheme is that the optimization problem of the central entity can change, e.g., if the topology of the transmission grid changes, without giving notice to the MGs.

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 $\mathbb{P}_E$

2. **Main Loop (ADMM):** For $l = 1, \ldots, l_{\text{max}} - 1 \in \mathbb{N}$
   (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 $p_{g,i}(k), \ldots, p_{g,i}(k + K - 1)$ and $\lambda_1(k), \ldots, \lambda_1(k + K - 1)$ to the MGs.

3. **Mixed integer update:**
   - For $l = l_{\text{max}}$, for all MGs $i = 1, \ldots, I$
     - Solve (22) where $U_j^i(k)\mathbb{P}_E$ is replaced by $U_j^i(k)$ and the constraints $p_{g,i}(j) - \hat{p}_{g,i}^{l_{\text{max}}}(j) = 0$ are added for $j = k, \ldots, k + K - 1$.

4. **Application of** $u_{i_{\text{max}}+1}^i(k)$, 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, $l_{\text{max}} \in \mathbb{N}$, or until a stopping criterion, e.g., based on the difference of $(z^l(k), p_{g,i}^l(k), \mathbf{\Lambda}^l(k))$ and $(z^{l-1}(k), p_{g,i}^{l-1}(k), \mathbf{\Lambda}^{l-1}(k))$ 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 $p_{g,i}(j) - \hat{p}_{g,i}^{l_{\text{max}}}(j) = 0$ are added to Problem 1 which implies that $\hat{p}_{g,i}^{l_{\text{max}}+1}(j) = \hat{p}_{g,i}^{l_{\text{max}}}(j) \in \mathbb{P}_E$ for all $j = k, \ldots, k + K - 1$. The resulting solution is again denoted by $u_*^t(k)$. From this $u_*^t(k)$ is extracted and applied in the standard MPC fashion.

Remark V.4. In general, it is possible that the mixed integer update (step (3) in Algorithm 1) is infeasible. However, such an infeasibility was not encountered in the numerical simulations in Section VI. Infeasibility could occur if $u_{i_{\text{max}}+1}^i(j) \in \{0, p_{g,i}^l\}$ holds for at least one $i \in I$ and one $j \in \{k, \ldots, K - 1\}$. Nevertheless, it is possible to set $u_{i}(j) = 0$ or $u_{i}(j) = p_{g,i}^l$ by changing $u_{i}(j)$, or by reducing the renewable infeed. However, a rigorous investigation on sufficient conditions for feasibility of the mixed integer update still needs to be performed.

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 is equipped such that in case of islanded operation the power demand, $w_{1i}(k)$, 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 $T_s = 0.5$ 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 pu h and a nominal power of 1 pu. For the prediction of load and RES infeed, a naïve forecast (see, e.g., [28]) 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 [29], 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 stage costs (12), (13) were chosen according to Table I. 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 with Gurobi 6.5.1 as solver.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(0.2, 0, -1, -1)^\top$ pu</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(1, 2, 1, 1)^\top$ pu</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(0.6)^\top$ pu h</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(1, 3, 4, 2.9, 5.6)^\top$ pu h</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(-1, -1, -1, -1)^\top$ pu</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(1, 1, 1, 1)^\top$ pu</td>
</tr>
<tr>
<td>$\mathbb{P}_E$</td>
<td>$(20, 20, 20, 20)^\top$ pu</td>
</tr>
</tbody>
</table>
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 $\mathcal{P}_c$. 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 passive transmission lines $p_{E,i}$, $i = 1, \ldots, 4$ 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 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.,

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

<table>
<thead>
<tr>
<th>Weight</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{t,i}$</td>
<td>0.1178</td>
</tr>
<tr>
<td>$c_{p,i}$</td>
<td>0.751 $/\text{pu}$</td>
</tr>
<tr>
<td>$c_{w,i}$</td>
<td>0.0048 $/\text{pu}^2$</td>
</tr>
<tr>
<td>$c_{x,i}$</td>
<td>1 $/\text{pu}^2$</td>
</tr>
<tr>
<td>$c_{E,i}$</td>
<td>0.05 $/\text{pu}^2$</td>
</tr>
<tr>
<td>$c_{g,i}$</td>
<td>0.5 $/\text{pu}^2$</td>
</tr>
<tr>
<td>$c_{s,i}$</td>
<td>0.1 $/\text{pu}^2$</td>
</tr>
<tr>
<td>$c_{w,i}$</td>
<td>0.1 $\cdot \text{diag}(1,2,3,6)$ $/\text{pu}^2$</td>
</tr>
</tbody>
</table>

Fig. 3. Simulation topology and results.
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 29% and a decrease of thermal infeed of 66% compared to the islanded case. 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.

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.