A distributed control strategy for reactive power compensation in smart microgrids

Saverio Bolognani and Sandro Zampieri

Abstract

We consider the problem of optimal reactive power compensation for the minimization of power distribution losses in a smart microgrid. We first propose an approximate model for the power distribution network, which allows us to cast the problem into the class of convex quadratic, linearly constrained, optimization problems. Then, we design a randomized, gossip-like optimization algorithm based on that model. We show how a distributed approach is possible, where agents have a partial knowledge of the problem parameters and state, and can only perform local measurements. For the proposed algorithm, we provide conditions for convergence together with an analytic characterization of the convergence speed. The analysis shows that the best performance can be achieved when we command cooperation among agents that are neighbors in the electric topology. Numerical simulations are included to validate the proposed model and to confirm the analytic results about the performance of the proposed algorithm.
CONTENTS

I Introduction 3

II Mathematical preliminaries and notation 4

III Problem formulation 5
   III-A A model of a microgrid 5
   III-B Power losses minimization problem 8

IV A randomized distributed algorithm 9
   IV-A Optimization problem decomposition 10
   IV-B Hessian estimation from local topology information 11
   IV-C Gradient estimation via local voltage measurement 12
   IV-D Description of the algorithm 13

V Convergence of the algorithm 13
   V-A A necessary condition for convergence 13
   V-B Convergence and rate of convergence 16

VI Optimal communication hypergraph for a radial distribution network 19

VII Simulations 21

VIII Conclusions 24

Appendix: Proof of Proposition 2 25

References 27
I. INTRODUCTION

Most of the distributed optimization methods have been derived for the problem of dispatching part of a large scale optimization algorithm to different processing units [1]. When the same methods are applied to networked control systems (NCS) [2], however, different issues arise: the way in which decision variables are assigned to different agents is not part of the designer degrees of freedom; agents have a local and limited knowledge of the problem parameters and of the system state; moreover, the information exchange between agents can occur not only via a given communication channel, but also via local actuation and subsequent measurement performed on an underlying physical system.

The extent of these issues depends on the particular application. In this work we present a specific scenario, belonging to the motivating framework of smart electrical power distribution networks [3], [4], in which these features play a central role.

In the last decade, the introduction of distributed microgeneration of electric energy (driven by technological advances, economical reasons, and environmental issues), together with an increased demand and the need for higher quality of service, has been driving the integration of a large amount of information and communication technologies (ICT) into the power distribution network.

Among the many different aspects of this transition, we focus on the control of the microgenerators inside a smart microgrid [5], [6]. A microgrid is a portion of the low-voltage power distribution network that is managed autonomously from the rest of the network, in order to achieve better quality of the service, improve efficiency, and pursue specific economic interests. Together with the loads connected to the microgrid (both residential and industrial customers), we also have microgeneration devices (solar panels, combined heat-and-power plants, micro wind turbines, etc.). These devices are connected to the microgrid via electronic interfaces (inverters), whose main task is to enable the injection of the produced power into the microgrid.

However, these devices can also perform different other tasks, denoted as ancillary services [7], [8], [9]: reactive power compensation, harmonic compensation, voltage support.

In this work we consider the problem of optimal reactive power compensation. Loads belonging to the microgrid may require a sinusoidal current which is not in phase with voltage. A convenient description for this, consists in saying that they demand reactive power together with active power, associated with out-of-phase and in-phase components of the current, respectively. Reactive power is not a “real” physical power, meaning that there is no energy conversion involved nor fuel costs to produce it. Like active power flows, reactive power flows contribute to power losses on the lines, cause voltage drop, and may lead to grid instability. It is therefore preferable to minimize reactive power flows by producing it as close as possible to the users that need it.

We explore the possibility of using the electronic interface of the microgeneration units to optimize the flows of reactive power in the microgrid. Indeed, the inverters of these units are generally oversized, because most of the distributed energy sources are intermittent in time, and the electronic interface is designed according to the peak power production. When they are not working at the rated power, these inverters can be commanded to inject a desired amount of reactive power at no cost [10].

This idea has been recently investigated in the literature on power systems [11], [12], [13]. However, these works consider a centralized scenario in which the parameters of the entire power grid are known, the controller can access the entire system state, and the few agents deployed in the distribution network are dispatched by a
central processing unit. The contribution of this work consists in casting the same problem in the framework of networked control systems, distributed optimization, and large scale complex systems. This approach allows the design of algorithms and solutions which can guarantee scalability, robustness to node insertion and removal, and compliance with the actual communication capabilities of the agents. Up to now, the few attempts of applying these tools to the power distribution networks have focused on grids comprising a large number of mechanical synchronous generators, instead of power inverters. See for example the stability analysis for these systems in [14] and the decentralized control synthesis in [15]. Instead, preliminary attempts of designing distributed reactive power compensation strategies for power inverters have been performed only very recently. However, the available results in this sense are mainly supported by simulations and heuristic approaches [16], and in some cases do not implement any communication or synergistic behavior of the agents [17].

The contribution of this paper is twofold. On one side, it consists in proposing a rigorous analytic derivation of an approximate model of the power flows which is commonly adopted in the power system literature and which allows to cast the optimal reactive power flow (ORPF) problem into a quadratic optimization (Section III). The second contribution is to propose and to analyze a distributed optimization algorithm whose operation is strongly based on the developed model (Section IV). In Section V we show under which conditions the previous algorithm converges to the global optimal solution. In Section VI we show how this algorithm can be optimized by a proper choice of the communication strategy. In Section VII, we finally validate the proposed model and the proposed algorithm via simulations.

II. MATHEMATICAL PRELIMINARIES AND NOTATION

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \sigma, \tau)$ be a directed graph, where $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of edges, and $\sigma, \tau : \mathcal{E} \rightarrow \mathcal{V}$ are two functions such that edge $e \in \mathcal{E}$ goes from the source node $\sigma(e)$ to the terminal node $\tau(e)$. Two edges $e$ and $e'$ are consecutive if the intersection $\{\sigma(e), \tau(e)\} \cap \{\sigma(e'), \tau(e')\}$ is not empty. A path is a sequence of consecutive edges (regardless of their direction).

In the rest of the paper we will often introduce complex-valued functions defined on the nodes and on the edges. These functions will also be intended as vectors in $\mathbb{C}^n$ and $\mathbb{C}^r$ (where $n = |\mathcal{V}|$, $r = |\mathcal{E}|$). We denote by $\bar{\cdot}$ the operation of element-wise complex conjugation, and by $^T$ the operation of matrix transposition.

Let moreover $A \in \{0, \pm 1\}^{r \times n}$ be the incidence matrix of the graph $\mathcal{G}$, defined via its elements

$$[A]_{ev} = \begin{cases} -1 & \text{if } v = \sigma(e) \\ 1 & \text{if } v = \tau(e) \\ 0 & \text{otherwise.} \end{cases}$$

If $W$ is a subset of indices, we define by $1_W$ the column vector whose elements are

$$[1_W]_v = \begin{cases} 1 & \text{if } v \in W \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, if $w$ is an index, we denote by $1_w$ the column vector whose value is 1 in position $w$, and 0 elsewhere, and we denote by $1$ the column vector of all ones. Notice that, if the graph $\mathcal{G}$ is connected (i.e. for every pair of nodes $v, w \in \mathcal{V}$ there is a path connecting them), then $1$ is the only vector in $\ker A$. 

October 24, 2011 DRAFT
An undirected graph $G$ is a graph in which for every edge $e \in \mathcal{E}$, there exists an edge $e' \in \mathcal{E}$ such that $\sigma(e') = \tau(e)$ and $\tau(e') = \sigma(e)$. In case the graph has no multiple edges and no self loops, we can also describe the edges of an undirected graph as subsets $e \subseteq \mathcal{V}$ with cardinality $|e| = 2$. By extension, we define a hypergraph $\mathcal{H}$ as a pair $(\mathcal{V}, \mathcal{E})$ in which edges are subsets of $\mathcal{V}$ of arbitrary cardinality.

III. PROBLEM FORMULATION

A. A model of a microgrid

We define a smart microgrid as a portion of the power distribution network that is connected to the power transmission network in one point, and hosts a number of loads and micro power generators, as described for example in [5], [6].

Figure 1. Schematic representation of the adopted microgrid model. In the lower panel a circuit representation is given, where black diamonds are microgenerators, white diamonds are loads, and the left-most element of the circuit represents the PCC. The middle panel illustrates the adopted graph representation for the same microgrid. Circled nodes represent microgenerators and the edge orientation is consistent with the sign of the power line currents. The upper panel shows how the compensators can be divided into overlapping clusters in order to implement the control algorithm proposed in Section IV. Each microgenerator is capable of measuring and actuating the system locally, while each cluster is also provided with some computational capability.
For the purpose of this paper, we model a microgrid as a directed graph $G$, in which edges represent the power lines, and nodes represent both loads and generators that are connected to the microgrid. These include the residential and industrial consumers, microgenerators, and also the point of connection of the microgrid to the transmission grid (called point of common coupling, or PCC).

We limit our study to the steady state behavior of the system, when all voltages and currents are sinusoidal signals at the same frequency $\omega_0$. Each signal can therefore be represented via a complex number whose absolute value corresponds to the signal root-mean-square value, and whose phase corresponds to the phase of the signal at $t = 0$. Therefore the complex number $y = |y|e^{j\angle y}$ represents the signal $y(t) = |y|\sqrt{2}\sin(\omega_0 t + \angle y)$.

In this notation, the steady state of a microgrid is described by the following system variables:

- $u \in \mathbb{C}^n$, where $u_v$ is the grid voltage at node $v$;
- $i \in \mathbb{C}^n$, where $i_v$ is the current injected by node $v$ into the grid;
- $\xi \in \mathbb{C}^r$, where $\xi_e$ is the current flowing on the edge $e$.

The following constraints are satisfied by $u, i$ and $\xi$

$$A^T \xi + i = 0, \quad (1)$$
$$Au + Z \xi = 0, \quad (2)$$

where $A$ is the incidence matrix of $G$, and $Z = \text{diag}(z_e, e \in \mathcal{E})$ is the diagonal matrix of line impedances, $z_e$ being the impedance of the microgrid power line corresponding to the edge $e$. Equation (1) corresponds to Kirchhoff’s current law (KCL) at the nodes, while (2) describes the voltage drop on the edges of the graph.

Each node $v$ of the microgrid is then characterized by a law relating its injected current $i_v$ with its voltage $u_v$. We model the node corresponding to the PCC (which we assume to be node 0) as a constant voltage generator, i.e.

$$u_0 = U_0. \quad (3)$$

We assume instead that the voltage $u_v$ and the current $i_v$ of every node $v$, but the PCC, satisfy the following law

$$u_v i_v = s_v \left| \frac{u_v}{U_0} \right|^\eta_v, \quad \forall v \in \mathcal{V}\backslash\{0\}, \quad (4)$$

where $s_v$ is the nominal complex power and $\eta_v$ is a characteristic parameter of the node $v$. The model (4) is called exponential model [18] and is widely adopted in the literature on power flow analysis [19]. Notice that $s_v$ is the complex power that the node would inject into the grid if the voltage at its point of connection was $U_0$. Notice also that, by properly choosing the parameters $\eta_v$, equation (4) describes the behavior of constant power, constant current, and constant impedance devices (with $\eta_v = 0, 1, 2$ respectively). In this sense, this model is also a generalization of ZIP models [18], which also are very common in the power system literature.

It is correct to say that equation (4) models the steady state of the vast majority of residential and industrial loads. Microgenerators also fit in this model, with $\eta_v = 0$, as they generally are commanded via a complex power reference and they are capable of injecting it independently from the voltage at their point of connection [5], [6].

The task of solving the system of nonlinear equations given by (1), (2), (3), and (4) to obtain the grid voltages and currents, given the network parameters, the injected nominal powers $\{s_v, v \in \mathcal{V}\backslash\{0\}\}$ at every node, and
the nominal voltage $U_0$ at the PCC, has been extensively covered in the literature under the denomination of power flow analysis (see for example [20, Chapter 3]).

In the following, we derive an approximate model for the microgrid state, which will be used later for the setup of the optimization problem and for the derivation of the proposed distributed algorithm. To do so, a couple of technical lemmas are needed.

**Lemma 1.** Let $L$ be the complex valued Laplacian $L := A^T A^{-1}$. There exists a unique symmetric matrix $X \in \mathbb{C}^{n \times n}$ such that

$$\begin{cases} X L = I - 11^T_0 \\ X 1_0 = 0, \end{cases} \quad (5)$$

where, we recall, $[1_0]_v = 1$ for $v = 0$, and $0$ otherwise and $I$ is the identity matrix.

**Proof:** Let us first prove the existence of $X$. Note that $\ker L = \text{span} 1 = \ker(I - 11^T_0)$, and therefore there exists $X' \in \mathbb{C}^{n \times n}$ such that $X' L = (I - 11^T_0)$. Let $X = X'(I - 10^T)$. Then

$$XL = X'(I - 10^T)L = X'L = I - 11^T_0,$$
$$X 1_0 = X'(I - 10^T) 1_0 = 0.$$

Existence is therefore guaranteed. To prove uniqueness, notice that

$$\begin{bmatrix} X & 1 \\ 1^T & 0 \end{bmatrix} \begin{bmatrix} L & 1_0 \\ 1^T_0 & 0 \end{bmatrix} = \begin{bmatrix} X L + 11^T_0 & X 1_0 \\ 1^T L & 1^T 1_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix}. $$

Therefore

$$\begin{bmatrix} X & 1 \\ 1^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} L & 1_0 \\ 1^T_0 & 0 \end{bmatrix},$$

and uniqueness of $X$ follows from the uniqueness of the inverse. Moreover, as $L = L^T$, we have

$$\begin{bmatrix} X & 1 \\ 1^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} L & 1_0 \\ 1^T_0 & 0 \end{bmatrix}^{-T} = \begin{bmatrix} L^T & 1_0 \\ 1^T_0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} X & 1 \\ 1^T & 0 \end{bmatrix}$$

and therefore $X = X^T$. 

The matrix $X$ depends only on the topology of the microgrid power lines and on their impedance (compare it with the definition of Green matrix in [21]). Indeed, it can be shown that, for every pair of nodes $(u, v)$,

$$(1_u - 1_v)^T X (1_u - 1_v) = Z_{uv}^{\text{eff}}, \quad (6)$$

where $Z_{uv}^{\text{eff}}$ is a complex-valued matrix whose elements $Z_{uv}^{\text{eff}}$ represent the effective impedance of the power lines between node $u$ and $v$.

If we fix the nominal voltage $U_0$ at the PCC, all the currents $i$ and the voltages $u$ are therefore determined by the equations

$$\begin{cases} u = Xi + U_0 1 \\ 1^T i = 0 \\ u_v i_v = s_v \left| \frac{u_v}{U_0} \right|^{0^v}, \quad \forall v \in V \setminus \{0\}, \end{cases} \quad (7)$$

October 24, 2011 DRAFT
where the first equation results from (1), (2), and (3) together with Lemma 1.

We can see the currents $i$ and the voltages $u$ as functions $i(U_0)$, $u(U_0)$ of $U_0$. The following proposition provides the Taylor approximation of $i(U_0)$ for large $U_0$.

**Proposition 2.** Let $s_0 := -\sum_{v\in V\setminus\{0\}} s_v$. Then for all $v \in V$ we have that

$$i_v(U_0) = \left(s_v + \delta_v(U_0)\right) \frac{1}{U_0}$$

where $\delta_v(U_0)$ is infinitesimal when $U_0$ tends to infinity.

The proof of this proposition is based on elementary multivariable analysis but requires a quite involved notation. For this reason it is given in the appendix. The quality of the approximation proposed in the previous proposition relies on having large nominal voltage $U_0$ and relatively small currents injected by the inverters (or supplied to the loads). This assumption corresponds to correct design and operation of power distribution networks, where indeed the nominal voltage is chosen sufficiently large (subject to other functional constraints) in order to deliver electric power to the loads with relatively small power losses on the power lines.

Numerical simulation in Section VII will indeed show that the approximation is extremely good when power distribution networks are operated in their regular regime.

**Remark.** Notice that this approximation, in similar fashions, has been used in the literature before for the problem of estimating power flows on the power lines (see among the others [22], [17] and references therein). It is also possible to show the analogy between this approximation and the quite common assumption of power flow conservation across the power network. The analytical rigorous justification proposed here, however, allows us to estimate the approximation error, and, more importantly, will provide the tools to understand what information on the system state can be gathered by properly sensing the microgrid.

**B. Power losses minimization problem**

Similarly to what has been done for example in [23], we choose active power losses on the power lines as a metric for optimality of reactive power flows. The total active power losses on the edges are given by

$$J' = \sum_{e \in \mathcal{E}} |\xi(e)|^2 \text{Re} z(e) = \bar{u}^T \text{Re}(L)u,$$

where we used (2) to express $\xi$ as a function of $u$. Via (7) and by exploiting the properties (5) of $X$, we have

$$J' = \bar{i}^T X \text{Re}(L)X \bar{i} = \bar{i}^T \text{Re}(X)i.$$

Let $s$ be the vector of all nominal complex powers $s_v$, including $s_0 := -\sum_{v\in V\setminus\{0\}} s_v$. Then, by plugging in the approximate system state (8), we have

$$J' = \frac{1}{|U_0|^2} s^T \text{Re}(X)s + \frac{1}{|U_0|^2} \tilde{J}(U_0, s)$$

$$= \frac{1}{|U_0|^2} p^T \text{Re}(X)p + \frac{1}{|U_0|^2} q^T \text{Re}(X)q + \frac{1}{|U_0|^2} \tilde{J}(U_0, s)$$

where $s$ is decomposed into the injected active power $p = \text{Re}(s)$ and the injected reactive power $q = \text{Im}(s)$, and where $\tilde{J}(U_0, s) := \bar{s}^T \text{Re}(X)s + s^T \text{Re}(X)\delta + \bar{\delta}^T \text{Re}(X)\delta$ is infinitesimal as $U_0$ tends to infinity and so it can be neglected.
Therefore, via the approximate model (8), we have been able to

- approximate power losses as a quadratic function of the injected power;
- decouple the problem of optimal power flows into the problem of optimal active and reactive power injection.

In the scenario that we are considering, we are allowed to command only a subset \( C \subset V \) of the nodes of the microgrid (namely the microgenerators, also named compensators in this framework). We denote by \( m := |C| \) its cardinality. Moreover, we assume that for these agents we are only allowed to command the amount of reactive power injected into the grid, as the decision on the amount of active power follows imperative economic criteria (for example, in the case of renewable energy sources, any available active power is typically injected into the grid to replace generation from traditional plants, which are more expensive and exhibit a worse environmental impact – see [10]).

The problem of optimal reactive power injection at the compensators can therefore be expressed as a quadratic, linearly constrained problem, in the form

\[
\min_q J(q), \quad \text{where} \quad J(q) = \frac{1}{2} q^T \text{Re}(X)q,
\]

subject to

\[
1^T q = 0
\]

\[
q_v = \text{Im}(s_v) \quad \forall v \in V \setminus C,
\]

\( \{\text{Im}(s_v), v \in V \setminus C\} \) being the nominal amount of reactive power injected by the nodes that cannot be commanded.

The solution of the optimization problem (9) would not pose any challenge if the nodes knew the problem parameters \( X \) and \( \{\text{Im}(s_v), v \in V \setminus C\} \). These quantities depend both on the grid parameters and the power demand of all the microgrid devices, therefore it is unpractical that every agent is capable of retrieving all this information, and, in the case of a large scale system, it would be also preferable to avoid the presence of a centralized agent collecting all the necessary data.

IV. A RANDOMIZED DISTRIBUTED ALGORITHM

In this section, we present an algorithm that can be distributed across the agents of the microgrid, meaning that it can be implemented by the compensators without the supervision of any central controller, by performing local measurements, data processing, and actuation. The proposed approach consists in decomposing the optimization problem into smaller, tractable subproblems that are assigned to small groups of agents. We will show that the agents can solve their corresponding subproblems via local measurement, local knowledge of the grid topology, and minimal communication. We will finally show that the iterative, randomized, solution of these subproblems yields the solution of the original optimization problem.

It is worth remarking that the decomposition methods proposed in most of the literature on distributed optimization, for example in [24], cannot be applied to this problem because the cost function in (9) is not separable into a sum of individual terms for each agent. Approaching the decomposition of this optimization problem via its dual formulation (as proposed in many works including the recent [25]) is also unlikely to succeed, as the feasibility of the system state must be guaranteed at any time during the optimization process.
A. Optimization problem decomposition

Let the compensators be divided into \( \ell \) possibly overlapping sets \( C_1, \ldots, C_\ell \), with \( \bigcup_{i=1}^\ell C_i = C \). This family of subsets can be interpreted as the edges of a hypergraph \( \mathcal{H} \) defined over the set of compensator nodes \( C \). Nodes belonging to the same set (or cluster) are able to communicate each other, and they are therefore capable of coordinating their actions and sharing their measurements. Each cluster is also provided with some computational capabilities, in the form of a cluster supervisor, possibly one of the compensators.

The proposed optimization algorithm consists of the following, repeated steps:

1) a set \( C_i(t) \) is chosen at a certain discrete time \( t = 0, 1, 2, \ldots \) where \( i(t) \in \{1, \ldots, \ell \} \);
2) the agents in \( C_i(t) \), by coordinating their actions and communicating, determine the new feasible state that minimizes \( J(q) \), solving the optimization subproblem in which all the nodes that are not in \( C_i(t) \) keep their states constant;
3) the agents in \( C_i(t) \) actuate the system by updating their state (the injected reactive power).

In order to describe more precisely this procedure, we rewrite the optimization (9) distinguishing the controllable and the uncontrollable components of the vector \( q \). Assume with no loss of generality that the first \( m \) components of \( q \) are controllable (i.e. they describe the reactive power injected into the grid by the compensators) and that the remaining \( n - m \) are not controllable. Let us then partition \( q \) as

\[
q = \begin{bmatrix} q_C \\ q_{\mathcal{V}\setminus C} \end{bmatrix}
\]

where \( q_C \in \mathbb{R}^m \) and \( q_{\mathcal{V}\setminus C} \in \mathbb{R}^{n-m} \). According to this partition of \( q \), let us also partition the matrix \( \text{Re}(X) \) as

\[
\text{Re}(X) = \begin{bmatrix} M & N \\ N^T & Q \end{bmatrix}.
\]

Let us also introduce some convenient notation. Consider the subspaces

\[
S_i := \left\{ q_C \in \mathbb{R}^m : \sum_{j \in C_i} [q_C]_j = 0, [q_C]_j = 0 \forall j \not\in C_i \right\},
\]

and the \( m \times m \) matrices

\[
\Omega := \frac{1}{2m} \sum_{h,k \in C} (1_h - 1_k)(1_h - 1_k)^T = I - \frac{1}{m} 11^T,
\]

\[
\Omega_i := \frac{1}{2|C_i|} \sum_{h,k \in C_i} (1_h - 1_k)(1_h - 1_k)^T
\]

\[
= \text{diag}(1_{C_i}) - \frac{1}{|C_i|} 1_{C_i}1_{C_i}^T.
\]

Notice that \( S_i = \text{span} \Omega_i \). We list some useful properties of these matrices.

i) \( \Omega_i^\sharp = \Omega_i \) and \( \Omega_i^\dagger = \Omega_i \) (where \( \dagger \) means pseudoinverse).

ii) The matrices \( \Omega_i, \Omega_i M \Omega_i \) and \( (\Omega_i M \Omega_i)^\dagger \) have entries different from zero only in position \( h,k \) with \( h,k \in C_i \).

iii) \( \ker(\Omega_i M \Omega_i)^\dagger = \ker \Omega_i \) and \( \text{span}(\Omega_i M \Omega_i)^\dagger = \text{span} \Omega_i \), and thus \( (\Omega_i M \Omega_i)^\dagger = (\Omega_i M \Omega_i)^\sharp \Omega_i \).
If the system is in the state \( q \) and the nodes in the cluster \( C_i \) are activated, they perform the following optimization

\[
\arg \min_{q^C} J(q^C, q_{V\setminus C}).
\]

Using the standard formulas for quadratic optimization it can be shown that

\[
q^{opt, i}_C := \arg \min_{q^C} J(q^C, q_{V\setminus C})
\]

\[
= q_C - (\Omega_i M \Omega_i)^T (M q_C + N q_{V\setminus C})
\]

\[
= q_C - (\Omega_i M \Omega_i)^T \nabla J,
\]

where

\[
\nabla J = M q_C + N q_{V\setminus C} = [\text{Re}(X)q_C] \in \mathbb{R}^m
\]

is the gradient of \( J(q_C, q_{V\setminus C}) \) with respect to the decision variables \( q_C \).

We prove now that the computation of \( q^{opt, i}_C \) can be performed by the nodes in a distributed way via a suitable approximation. Notice first that, if \( h \not\in C_i \), then

\[
[q^{opt, i}_C]_h = q_h
\]

If instead \( h \in C_i \), we can write

\[
[q^{opt, i}_C]_h = q_h - \sum_{k \in C_i} \left[ (\Omega_i M \Omega_i)^T \right]_{hk} [\nabla J]_k.
\]

From the previous properties we get that \( \left[ (\Omega_i M \Omega_i)^T \right]_{hk} = 0 \) for all \( k \not\in C_i \), and so

\[
[q^{opt, i}_C]_h = q_h - \sum_{k \in C_i} \left[ (\Omega_i M \Omega_i)^T \right]_{hk} [\nabla J]_k.
\]

In the following, we show how each node \( h \in C_i \) can compute (14) on the basis of the information available to the cluster \( C_i \).

### B. Hessian estimation from local topology information

Agents belonging to the cluster \( C_i \) can infer some information about the Hessian \( M \). More precisely they can determine the matrix \( \Omega_i M \Omega_i \) appearing in (14). Define now \( R_{\text{eff}} \) as a \( m \times m \) matrix with entry in \( h, k \) equal to \( \text{Re} (Z_{hk}^{\text{eff}}) \), where \( Z_{hk}^{\text{eff}} \) are the mutual effective impedances between pairs compensators \( h, k \). From (6) and (10) we have that

\[
R_{hk}^{\text{eff}} = \text{Re} (Z_{hk}^{\text{eff}}) = (1_h - 1_k)^T M (1_h - 1_k)
\]

and so we can write

\[
R_{\text{eff}} = \sum_{h, k \in C} (1_h - 1_k)^T M (1_h - 1_k) 1_k^T
\]

\[
= \text{diag}(M) 11^T + 11^T \text{diag}(M) - 2M
\]

where \( \text{diag}(M) \) is the \( m \times m \) diagonal matrix having the same diagonal elements of \( M \). Consequently, since \( \Omega_{C_i} 1 = 0 \), we have that \( \Omega_i R_{\text{eff}} \Omega_i = -2 \Omega_i M \Omega_i \) and then

\[
\Omega_i M \Omega_i = -\frac{1}{2} \Omega_i R_{\text{eff}} \Omega_i.
\]

We can therefore argue that the elements of the matrix \( \Omega_i M \Omega_i \) can be obtained by the nodes from the mutual effective impedances \( Z_{hk}^{\text{eff}} \) for \( h, k \in C_i \). These impedances are assumed to be known by the nodes belonging
to the cluster. This is a reasonable hypothesis, since the mutual effective impedances can be obtained from either a priori knowledge of the local microgrid topology, or via online estimation procedures as in [26], or via ranging technologies over power line communications as suggested in [27].

Then the nodes need to compute the pseudo-inverse of $\Omega_i$. Notice that all these operations have to be executed offline once, as these coefficients depend only on the grid topology and impedances.

C. Gradient estimation via local voltage measurement

Assume that nodes in $C_i$ can measure the grid voltage at their point of connection (via phasor measurement units that return both amplitude and phase of the measured voltage – see [28], [29]). Let moreover assume the following about power line impedances.

Assumption 3. All power lines in the microgrid have the same inductance/resistance ratio, i.e.

$$Z = e^{j\theta}Z$$

where $Z$ is a diagonal real-valued matrix, whose elements are $Z_{ee} = |z_e|$. Consequently, $L = e^{-j\theta}A^T Z^{-1} A$, and $X := e^{-j\theta} X$ is a real-valued matrix.

This assumption seems quite reasonable in most practical cases (see for example the IEEE standard testbeds [30]). The possible effects of this approximation will be commented in Section VII.

Let then each agent $k \in C_i$ compute

$$\nu_k^{(i)} := \frac{1}{|C_i|} \sum_{v \in C_i} |u_v| |u_k| \sin(\angle u_k - \angle u_v - \theta)$$

(16)

Observe that

$$\nu_k^{(i)} = \text{Im} \left[ \left( e^{-j\theta} \frac{1}{|C_i|} \sum_{v \in C_i} \bar{u}_v \right) u_k \right].$$

By using (7), we have

$$\nu_k^{(i)} = \text{Im} \left[ e^{-j\theta} \left( e^{-j\theta} \frac{1}{|C_i|} \bar{1}_C^T X i + \bar{U}_0 \right) \left( e^{j\theta} \bar{1}_k^T X i + U_0 \right) \right]$$

$$= \text{Im} \left[ e^{-j\theta} \frac{1}{|C_i|} \bar{1}_C^T X i \bar{1}_k^T X i + e^{-j2\theta} \frac{1}{|C_i|} \bar{1}_C^T X i \bar{U}_0 \right.$$  

$$+ \bar{1}_k^T X \bar{s} + \bar{1}_k^T X \delta + e^{-j\theta} |U_0|^2 \right],$$

where we used the fact that, by (8), $\bar{U}_0 i = \bar{s} + \delta$. The third term of the sum is, according to (13),

$$\text{Im} \left[ \bar{1}_k^T X \bar{s} \right] = -\bar{1}_k^T X q = -\frac{1}{\cos \theta} |\nabla J|_k$$

The first and the fourth terms of the sum form the quantity

$$\bar{\nu}_k^{(i)} := \text{Im} \left[ e^{-j\theta} \frac{1}{|C_i|} \bar{1}_C^T X i \bar{1}_k^T X i + \bar{1}_k^T X \delta \right]$$

which is infinitesimal for large $U_0$ and so it can be ignored. The second and the last terms form the quantity

$$\alpha^{(i)} := \text{Im} \left[ e^{-j2\theta} \frac{1}{|C_i|} \bar{1}_C^T X i \bar{U}_0 + e^{-j\theta} |U_0|^2 \right]$$

which does not depend on the node $h$ but only on the selected cluster $C_i$. Therefore we can write

$$[\nabla J]_k = \cos \theta \left( -\nu_k^{(i)} + \alpha^{(i)} + \bar{\nu}_k^{(i)} \right).$$

(17)
D. Description of the algorithm

From the above consideration, starting from equation (14), we propose the following iterative algorithm. If we assume that at time \( t \) the system is in the state \( q(t) \in \mathbb{R}^n \) and that the cluster \( i \) is activated. Then the new state \( q(t+1) \) will be such that \( q_h(t+1) = q_h(t) \) for all \( h \not\in C_i \), while the node \( h \in C_i \) will inject the new reactive power

\[
q_h(t+1) = q_h(t) - 2\cos\theta \sum_{k \in C_i} \left[ (\Omega_i R_{eff}^i \Omega_i)^\sharp \right]_{hk} \nu_k^{(i)}(t),
\]

where we used that fact that, as \( \ker(\Omega_i R_{eff}^i \Omega_i)^\sharp = \ker \Omega_i \), the term \( \alpha^{(i)} \) in (17) disappears, while the terms \( \tilde{\nu}_k^{(i)} \) have been neglected since they are infinitesimal.

We have therefore shown as the proposed algorithm can be implemented by the agents of the microgrid in a distributed way. In a preliminary, offline phase, each cluster computes \( (\Omega_i R_{eff}^i \Omega_i)^\sharp \). Then, at every iteration of the algorithm:

- a cluster \( C_i \) is randomly chosen;
- every agent \( h \) not belonging to the cluster \( C_i \) holds its injected reactive power constant;
- every agent \( h \) belonging to the cluster \( C_i \) senses the grid voltage at its point of connection, computes \( \nu_h^{(i)} \) as in (16), and then updates its injected reactive power according to (18).

Remark. It is important to notice that the distributed implementation of the algorithm is only possible because of the specific physical system that we are considering. By sensing the system locally, nodes can infer some global information on the system state (namely, the gradient of the cost function) that otherwise would depend on the states of all other agents. For this reason, it is necessary that, after each iteration of the algorithm, the agents actuate the system implementing the optimization step, so that subsequent measurement will reflect the state change. This approach resembles some other applications of distributed optimization, e.g. the radio transmission power control in wireless networks [31] and the traffic congestion protocol for wired data networks [32], [33]. Also in these cases, the iterative tuning of the decision variables (radio power, transmission rate) depends on congestion feedback signals that are function of the entire state of the system. However, these signals can be detected locally by each agent by measuring error rates, signal-to-noise ratios, or specific feedback signals in the protocol.

V. Convergence of the algorithm

In this section we will give the conditions ensuring that the proposed iterative algorithm converges to the optimal solution, and we will analyze its convergence rate.

A. A necessary condition for convergence

Consider the discrete time system

\[
q(t+1) = q(t) - (\Omega_i M \Omega_i)^\sharp (M q(t) + N q_{\setminus C}).
\]

This system describes the ideal iterative optimization algorithm based on the steps described in (12). It therefore approximates the iterative update law (18), and we refer to it for studying the convergence of the proposed
algorithm. We introduce the auxiliary variable $x = q - q_{\text{opt}}^C \in \mathbb{R}^m$, where $q_{\text{opt}}$ is the solution of the optimization problem (9), and $q_{\text{opt}}^C$ are the elements of $q_{\text{opt}}$ corresponding to the nodes in $C$.

In this notation, it is possible to explicitly express the previous discrete time (19) system as

$$x(t + 1) = F_i(t)x(t), \quad x(0) \in \ker 1^T,$$

where

$$F_i = I - (\Omega_i M \Omega_i)^\# M.$$

The matrices $F_i$ have some nice properties which are listed below:

i) the matrices $F_i$ are projection operators, i.e. $F_i^2 = F_i$, and they are orthogonal projections with respect to the inner product $\langle \cdot, \cdot \rangle_M$, defined as $\langle x, y \rangle_M := x^T M y$. In other words, $\langle F_i x, F_i x - x \rangle_M = x^T M (F_i x - x) = 0$.

ii) The matrices $F_i$ are self-adjoint matrices with respect to the inner product $\langle \cdot, \cdot \rangle_M$, i.e. $F_i^T M = M F_i$, thus having real eigenvalues.

Notice that, as $\text{span}(\Omega_i M \Omega_i)^\# = \text{span} \Omega_i$, $\ker 1^T$ is a positively invariant set for (20), namely $F_i(\ker 1^T) \subseteq \ker 1^T$ for all $i$.

It is clear that $q(t)$ converges to the optimal solution $q_{\text{opt}}$ if and only if $x(t)$ converges to zero for any initial condition $x(0) \in \ker 1^T$. A necessary condition for this to happen is that there are no nonzero equilibria in the discrete time system (20), namely that the only $x \in \ker 1^T$ such that $F_i x = x$ for all $i$ must be $x = 0$. The following proposition provides a convenient characterization of this property. We need the following lemma first.

**Lemma 4.** $x = 0$ is the only point in $\ker 1^T$ such that $F_i x = x$ for all $i$ if and only if

$$\text{span}[\Omega_1 \ldots \Omega_\ell] = \ker 1^T.$$  

**Proof:** Let us prove the reverse implication first. Assume that $\text{span}[\Omega_1 \ldots \Omega_\ell] = \ker 1^T$ and that $x \in \ker 1^T$ is such that $F_i x = x$ for all $i$. Then we can find $y_i \in \mathbb{R}^n$ such that

$$x = \sum_i \Omega_i y_i.$$

Moreover, as $F_i x = x$ for all $i$, then $(\Omega_i M \Omega_i)^\# M x = 0$ and so, by the properties mentioned above, we can argue that $M x \in \ker \Omega_i$. We therefore have

$$x^T M x = \sum_i y_i^T \Omega_i M x = 0,$$

and so, since $M$ is positive definite, it yields that $x = 0$.

We prove now the other direction. The inclusion $\text{span}[\Omega_1 \ldots \Omega_\ell] \subseteq \ker 1^T$ is always true and follows from the fact that $\text{span} \Omega_i \subseteq \ker 1^T$. We need to prove only the other inclusion. Suppose that $x = 0$ is the only point in $\ker 1^T$ such that $F_i x = x$ for all $i$. This means that

$$\ker 1^T \cap \ker(I - F_1) \cap \cdots \cap \ker(I - F_\ell) = \{0\},$$

which implies that

$$\mathbb{R}^m = \text{span} 1 + \text{span}(I - F_1^T) + \cdots + \text{span}(I - F_\ell^T).$$
Take now any \( x \in \ker 1^T \). Then there exist a \( \alpha \in \mathbb{R} \) and \( y_i \in \mathbb{R}^m \) such that
\[
Mx = \alpha 1 + \sum_{i}^\ell (I - F_i)^Ty_i = \alpha 1 + \sum_{i}^\ell M(\Omega_i M\Omega_i)^y y_i
\]
Then \( \alpha 1 = Mw \) where
\[
w = x - \sum_{i}^\ell (\Omega_i M\Omega_i)^y y_i = x - \sum_{i}^\ell \Omega_i z_i
\]
where we have used the fact that \( \text{span}(\Omega_i X \Omega_i)^y = \text{span} \Omega_i \). From the fact that \( \text{span} \Omega_i \subseteq \ker 1^T \), we can argue that \( w \in \ker 1^T \) and so it follows that \( 0 = w^T \alpha 1 = w^T Mw \). Finally, since \( M \) is positive definite, we can conclude that \( w = 0 \) and so that \( x \in \text{span}[\Omega_1 \ldots \Omega_\ell] \).

The condition expressed in Lemma 4 can also be expressed as a connectivity test. Let \( \mathcal{H} \) be the hypergraph having the elements in \( \mathcal{C} \) as nodes and the sets \( \mathcal{C}_i, i = 1, 2, \ldots, \ell \) as hyperedges. This hypergraph describes the communication network needed for implementing the algorithm.

**Proposition 5.** \( x = 0 \) is the only point in \( \ker 1^T \) such that \( F_i x = x \) for all \( i \) if and only if the hypergraph \( \mathcal{H} \) is connected.

**Proof:** Consider the matrix \( W \in \mathbb{R}^{m \times m} \) with entries \( W_{hk} \) equal to the number of the sets \( \mathcal{C}_i \) which contain both \( h \) and \( k \). Let \( G_W \) be the weighted graph associated with \( W \). It is quite easy to see that the hypergraph having \( \mathcal{C}_i \) as edges is connected if and only if \( G_W \) is a connected graph.

Let us define \( \delta_{\mathcal{C}_i} : \mathcal{C} \to \{0, 1\} \) as the characteristic function of the set \( \mathcal{C}_i \), namely a function of the nodes that is 1 when the node belongs to \( \mathcal{C}_i \) and is zero otherwise. Then the Laplacian \( L_W \) of \( G_W \) can be expressed as follows
\[
L_W = \sum_{h,k \in \mathcal{C}} (1_h - 1_k)(1_h - 1_k)^T W_{hk}
\]
\[
= \sum_{h,k \in \mathcal{C}} (1_h - 1_k)(1_h - 1_k)^T \sum_{i=1}^\ell \delta_{\mathcal{C}_i}(h) \delta_{\mathcal{C}_i}(k)
\]
\[
= \sum_{i=1}^\ell \sum_{h,k \in \mathcal{C}_i} (1_h - 1_k)(1_h - 1_k)^T = \sum_{i=1}^\ell 2|\mathcal{C}_i|\Omega_i
\]
\[
= [\Omega_1 \ldots \Omega_\ell] \text{diag}(2|\mathcal{C}_1|I, \ldots, 2|\mathcal{C}_\ell|I)
\]
Notice that the graph connectivity is equivalent to the fact that \( \ker L_W = \text{span} 1 \) and so, by the previous equality, it is equivalent to
\[
\ker \begin{bmatrix} \Omega_1 \\ \vdots \\ \Omega_\ell \end{bmatrix} = \text{span} 1
\]
which is equivalent to (21). The proposition statement then follows from Lemma 4.
B. Convergence and rate of convergence

In order to prove that the connectivity of the hypergraph $H$ is not only a necessary but also a sufficient condition for the convergence of the algorithm, we introduce the following assumption on the sequence $i(t)$.

**Assumption 6.** The sequence $i(t)$ is a sequence of independently, identically distributed symbols in \( \{1, \ldots, \ell\} \), with non-zero probabilities \( \rho_i > 0, i = 1, \ldots, \ell \).

Let $v(t) := \mathbb{E}\left[x^T(t)Mx(t)\right] = J(q(t)) - J(q^{\text{opt}})$ and consider the following performance metric

\[
R = \sup_{x(0) \in \ker 1} \limsup_{t \to \infty} \{v(t)^{1/t}\}
\]

which describes the exponential rate of convergence to zero of $v(t)$. It is clear that $R < 1$ implies the exponential mean square convergence of $q_C(t)$ to the optimal solution $q^{\text{opt}}$.

Using (20), we have

\[
v(t) = \mathbb{E}\left[x(t)^T Mx(t)\right] = \mathbb{E}\left[x(t)^T \Omega M \Omega x(t)\right] = \mathbb{E}\left[x(t-1)^T F_{\sigma(t-1)}^T \Omega M \Omega F_{\sigma(t-1)} x(t-1)\right] = x(0)^T \mathbb{E}\left[F_{\sigma(0)}^T \cdots F_{\sigma(t-1)}^T \Omega M \Omega F_{\sigma(t-1)} \cdots F_{\sigma(0)}\right] x(0).
\]

Let us then define

\[
\Delta(t) = \mathbb{E}\left[F_{\sigma(0)}^T \cdots F_{\sigma(t-1)}^T \Omega M \Omega F_{\sigma(t-1)} \cdots F_{\sigma(0)}\right].
\]

Via Assumption 6, we can argue that $\Delta(t)$ satisfies the following linear recursive equation

\[
\begin{cases}
\Delta(t + 1) = \mathcal{L}(\Delta(t)) \\
\Delta(0) = \Omega M \Omega,
\end{cases}
\]

where $\mathcal{L}(\Delta) := \mathbb{E}\left[F_i^T \Delta F_i\right]$. Moreover we have that

\[
v(t) = x(0)^T \Delta(t) x(0).
\]

Equations (22) and (23) can be seen as a discrete time linear system with state $\Delta(t)$ and output $v(t)$.

**Remark.** The discrete time linear system (22) can be equivalently be described by the equation

\[
\text{vec} \left(\Delta(t + 1)\right) = \mathbf{F} \text{vec} \left(\Delta(t)\right),
\]

where $\text{vec}(\cdot)$ stands for the operation of vectorization and where

\[
\mathbf{F} := \mathbb{E}\left[F_i^T \otimes F_i^T\right] \in \mathbb{R}^{m^2 \times m^2}.
\]

Notice that $\mathbf{F}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{M^{-1} \otimes M^{-1}}$ and so $\mathbf{F}$, and consequently $\mathcal{L}$, has real eigenvalues.

Studying the rate $R$ (which can be proved to be the slowest reachable and observable mode of the system (22)-(23) [34]) is in general not simple. It has been done analytically and numerically for some special graph structures in [34]. It is convenient to analyze its behavior indirectly, through another parameter which is easier to compute and to study. Let us define

\[
\beta = \max\{\|\lambda\| \mid \lambda \in \lambda(F^{\text{ave}}), \lambda \neq 1\},
\]
where $F_{\text{ave}} := \mathbb{E} [F_i]$.

We have the following result.

**Theorem 7.** Assume that Assumption 6 holds true and that the hypergraph $H$ is connected. Then

$$R \leq \beta < 1.$$ 

Before proving Theorem 7, observe that the following result descends directly from it.

**Corollary 8.** Assume that Assumption 6 holds true and that the hypergraph $H$ is connected. Then the state of the iterative algorithm described in Section IV-A converges in mean square to the global optimal solution.

**Remark.** Observe that, by standard arguments based on Borel Cantelli lemma, the exponential convergence in mean square implies also almost sure convergence to the optimal solution.

In order to prove Theorem 7 we need a few technical lemmas.

**Lemma 9.** Let $P, Q \in \mathbb{R}^{m \times m}$ and $P \geq Q$. Then $L^t(P) \geq L^t(Q)$ for all $t \in \mathbb{Z}_{\geq 0}$.

**Proof:** From the definition of $L$, we have

$$x^T [L(P) - L(Q)] x = x^T [\mathbb{E} [F_i^T PF_i] - \mathbb{E} [F_i^T Q F_i]] x$$

$$= \mathbb{E} [x^T F_i^T (P - Q) F_i x] \geq 0.$$ 

By iterating these steps $t$ times we then obtain $L^t(P) \geq L^t(Q)$.

**Lemma 10.** For all $\Delta$ we have that $\Omega L^t(\Omega \Delta \Omega) \Omega = \Omega L^t(\Delta) \Omega$.

**Proof:** Proof is by induction. The statement is true for $t = 0$, as $\Omega^2 = \Omega$. Suppose it is true up to $t$. We then have

$$\Omega L^{t+1}(\Delta) \Omega = \Omega L(L^t(\Delta)) \Omega = \Omega L(\Omega L^t(\Delta) \Omega) \Omega$$

$$= \Omega L(\Omega L^t(\Omega \Delta \Omega) \Omega) \Omega = \Omega L^{t+1}(\Omega \Delta \Omega) \Omega.$$ 

**Lemma 11.** All the eigenvalues of $F_{\text{ave}}$ are real and have absolute value not larger than 1. If $\text{span} [\Omega_1, \ldots, \Omega_\ell] = \ker 1^T$ and if Assumption 6 holds, then the only eigenvalue of $F_{\text{ave}}$ on the unitary circle is $\lambda = 1$, with multiplicity 1 and with associated left eigenvector $1$ and right eigenvector $M^{-1}1$.

**Proof:** Recall that the matrices $F_i$ are projection operators, i.e. $F_i^2 = F_i$ and so they have eigenvalues 0 or 1. Recall moreover that $F_i$ are self-adjoint matrices with respect to the inner product $\langle \cdot, \cdot \rangle_M$. This implies that $\|F_i\|_M \leq 1$ where $\|\cdot\|_M$ is the induced matrix norm with respect to the vector norm $\|x\|_M := \langle x, x \rangle_M^{1/2}$. This implies that

$$\|F_{\text{ave}}\|_M = \|\mathbb{E} [F_i]\|_M \leq \mathbb{E} [\|F_i\|_M] \leq 1.$$
And so, since also $F_{\text{ave}}$ is self-adjoint, its eigenvalues are real and are smaller than or equal to 1 in absolute value.

Assume now that $F_{\text{ave}}x = \lambda x$, with $|\lambda| = 1$. Then we have
\[
\|x\|_M = \|F_{\text{ave}}x\|_M = \|E[F_i]x\|_M \leq \|x\|_M.
\]
If Assumption 6 holds (and thus the probabilities $\rho_i$’s are all strictly greater than 0), then the last inequality implies that $\|F_i x\|_M = \|x\|_M$ for all $i$’s. As $F_i$ are projection matrices, it means that $F_ix = x$ and so $Mx \in \ker \Omega_i^T, \forall i$. Using the fact that span $[\Omega_1 \ldots \Omega_d] = \ker 1^T$, we necessarily have $x = M^{-1}1$. By inspection we can verify that the left eigenvector corresponding to the same eigenvalue is 1.

We can now give the proof of Theorem 7.

Proof of Theorem 7: Let us first prove that $\Omega \mathcal{L}(\Omega M \Omega) \Omega \leq \beta \Omega M \Omega$. Indeed, we have
\[
x^T \Omega \mathcal{L}(\Omega M \Omega) \Omega x = \mathbb{E} \left[ x^T \Omega F_i^T \Omega M \Omega F_i \Omega x \right] = \mathbb{E} \left[ x^T \Omega M^{1/2} \Omega F_i^T M F_i \Omega x \right] = x^T \Omega M^{1/2} \mathbb{E} \left[ M^{1/2} F_i M^{1/2} \right] M^{1/2} \Omega x,
\]
where we use the fact that $\Omega F_i \Omega = F_i \Omega$ and $F_i^T M F_i = MF_i$. Notice moreover that $\mathbb{E} \left[ M^{1/2} F_i M^{1/2} \right] = M^{1/2} F_{\text{ave}} M^{1/2}$ is symmetric and, by Lemma 11, it has only one eigenvalue on the unit circle (precisely in 1), with eigenvector $M^{-1/2}1$. As $M^{1/2} \Omega x \perp M^{-1/2}1$ for all $x$, we have
\[
x^T \Omega \mathcal{L}(\Omega M \Omega) \Omega x \leq \beta x^T \Omega M \Omega x,
\]
with $\beta = \max \{|\lambda| \mid \lambda \in \lambda(F_{\text{ave}}), \lambda \neq 1\}$.

From this result, using Lemmas 9 and 10, we can conclude
\[
\Omega \mathcal{L}^t(\Omega M \Omega) \Omega = \Omega \mathcal{L}^{t-1}(\mathcal{L}(\Omega M \Omega)) \Omega \\
= \Omega \mathcal{L}^{t-1}(\Omega \mathcal{L}(\Omega M \Omega) \Omega) \\
\leq \Omega \mathcal{L}^{t-1}(\beta M M \Omega) \Omega \\
= \beta \Omega \mathcal{L}^{t-1}(\Omega M \Omega) \Omega \leq \cdots \leq \beta^t M M \Omega,
\]
and therefore $R \leq \beta$.

The tightness of $\beta$ as a bound for $R$ has been studied in [34], by evaluating both $\beta$ and $R$ analytically and numerically for some special graph topologies. In the following we consider $\beta$ as a reliable metric for the evaluation of the algorithm performances.

The following result shows what is the best performance (according to the bound $\beta$ on the convergence rate $R$) that the proposed algorithm can achieve.

**Theorem 12.** Consider the algorithm (19), and assume that $H$ describing the clusters $C_i$ is an arbitrary connected hypergraph defined over the nodes $C_i$. Let Assumption 6 hold. Then
\[
\beta \geq 1 - \frac{\sum_{i=1}^d \rho_i |C_i| - 1}{m - 1}.
\]
In case all the sets $C_i$ have the same cardinality $c$, namely $|C_i| = c$ for all $i$, then
\[
\beta \geq 1 - \frac{c - 1}{m - 1}.
\]
Proof: Let

\[ E_i := (\Omega_i M \Omega_i)^2 M \]  

(26)

\[ E_{\text{ave}} := \mathbb{E} [E_i] = \sum_{i=1}^{\ell} \rho_i E_i \]  

(27)

It is clear that \( \beta = 1 - \beta' \), where \( \beta' = \min\{|\lambda| \mid \lambda \in \lambda(E_{\text{ave}}), \lambda \neq 0\} \). We have that

\[ \sum_{\lambda_j \in \lambda(E_{\text{ave}})} \lambda_j = \text{Tr}(E_{\text{ave}}) = \text{Tr} \left( \sum_{i=1}^{\ell} \rho_i E_i \right) = \sum_{i=1}^{\ell} \rho_i \text{Tr}(E_i). \]

Notice now that \( E_i \) have eigenvalues zero or one and so

\[ \text{Tr}(E_i) = \text{rank}(E_i) = \dim(\text{span}(E_i)) = \dim(\text{span}(\Omega_i)) = |C_i| - 1. \]

This implies that

\[ \sum_{\lambda_j \in \lambda(E_{\text{ave}})} \lambda_j = \left( \sum_{i=1}^{\ell} \rho_i |C_i| \right) - 1, \]

and so

\[ \beta' = \min\{|\lambda| \mid \lambda \in \lambda(E_{\text{ave}}), \lambda \neq 0\} \]

\[ \leq \frac{\sum_{\lambda_j \in \lambda(E_{\text{ave}})} \lambda_j}{m - 1} = \frac{\left( \sum_{i=1}^{\ell} \rho_i |C_i| \right) - 1}{m - 1}. \]

\[ \blacksquare \]

VI. OPTIMAL COMMUNICATION HYPERGRAPH FOR A RADIAL DISTRIBUTION NETWORK

In this section we present a special case in which the optimal convergence rate of Theorem 12 is indeed achieved via a specific choice of the clusters \( C_i \). To obtain this result we need to introduce the following assumption, which is commonly verified in many practical cases (see for example the standard IEEE testbeds [30]).

Assumption 13. The distribution network is radial, i.e. the corresponding graph \( G \) is a tree.

We start by providing some useful properties of the matrices \( E_i \).

Lemma 14. The following properties hold true:

i) \( \text{span}(E_i) = \text{span}(\Omega_i) \);

ii) \( E_i^2 = E_i \);

iii) \( E_i x = x \) for all \( x \in \text{span}(\Omega_i) \).

Proof:

i) First observe that, since \( M \) is invertible, then \( \text{span}(E_i) = \text{span}(\Omega_i) M \Omega_i)^2 M = \text{span}(\Omega_i M \Omega_i)^2 \). Finally, from the properties of the matrices \( (\Omega_i M \Omega_i)^2 \), we obtain that \( \text{span}(E_i) = \text{span}(\Omega_i) \).
ii) First observe that $E_i^2 = (I - F_i)^2 = I - 2F_i + F_i^2$. Now, since $F_i^2 = F_i$ we obtain that $I - 2F_i + F_i^2 = I - F_i = E_i$.

iii) Let $x \in \text{span } \Omega_i = \text{span } E_i$. Then $x = E_i z$ for some vector $z$ and so $E_i x = E_i^2 z = E_i z = x$.

Let us now consider for any pair of nodes $h, k$ the shortest path $P_{hk} \subseteq \mathcal{E}$ (where, we recall, a path is a sequence of consecutive edges) connecting $h, k$. Define moreover

$$P_{C_i} := \bigcup_{h, k \in C_i} P_{hk}.$$ 

**Lemma 15.** Let Assumption 13 hold true. Then

i) $(1_h - 1_k)^T M (1_r - 1_s) = \sum_{e \in P_{hk} \cap P_{rs}} \text{Re}(z_e)$.

ii) $E_i (1_h - 1_k) = 0$ if $P_{C_i} \cap P_{hk} = \emptyset$.

**Proof:**

i) Observe that $(1_h - 1_k)^T M (1_r - 1_s) = (1_h - 1_k)^T X (1_r - 1_s)$, where with some abuse of notation we have denoted with the same symbol $1_h$ the vector in $\mathbb{R}^m$ and the corresponding vector in $\mathbb{R}^n$. Notice that, if we have a current $i = 1_h - 1_s$ in the network, then according to (7) we get a voltage vector $u = X i + U_0 = X (1_r - 1_s) + U_0$ and so $(1_h - 1_k)^T X (1_r - 1_s) = u_h - u_k$. Observing the following figure

we can argue that $u_h - u_k = u_h' - u_k' = \sum_{e \in P_{hk} \cap P_{rs}} z_e$ and so $(1_h - 1_k)^T M (1_r - 1_s) = \sum_{e \in P_{hk} \cap P_{rs}} \text{Re}(z_e)$.

ii) If $P_{C_i} \cap P_{hk} = \emptyset$, then

$$\Omega_i M (1_h - 1_k) = \frac{1}{2 |C_i|} \sum_{r, s \in C_i} (1_r - 1_s)^T (1_r - 1_s) = 0$$

since $P_{hk} \cap P_{rs} = \emptyset$ for all $r, s \in C_i$. Thus $E_i (1_h - 1_k) = (\Omega_i M \Omega_i)^2 \Omega_i M (1_h - 1_k) = 0$.

We give now the key definition to characterize, together with connectivity, the optimality of the communication hypergraph $\mathcal{H}$.

**Definition 16.** The hyperedges $\{C_i, i = 1, 2, \ldots, \ell\}$ of the hypergraph $\mathcal{H}$ are edge-disjoint$^1$ if for any $i, j$ such that $i \neq j$, we have that $P_{C_i} \cap P_{C_j} = \emptyset$.
In order to prove the optimality of having edge-disjoint clusters, we need the following lemma.

**Lemma 17.** Assume that Assumptions 13 holds true and that the clusters \( C_i \) are edge-disjoint. Then \( E_i E_j = 0 \) for all \( i \neq j \).

**Proof:** Let \( x \) be any vector in \( \mathbb{R}^m \). Notice that, since \( \text{span} \ E_j = \text{span} \Omega_j \), then there exists \( z \in \mathbb{R}^m \) such that
\[
E_i E_j x = E_i \Omega_j z = \frac{1}{2|C_i|} \sum_{r,s \in C_j} E_i (1_r - 1_s)(1_r - 1_s)^T z = 0
\]
where the last equality follows from the fact that \( \mathcal{P}_{C_i} \cap \mathcal{P}_{rs} = \emptyset \) for all \( r, s \in C_j \).

From the previous lemma we can argue that every vector in \( \text{span} \Omega_i \) is eigenvector of \( E_{\text{ave}} = \sum_{i=1}^\ell \rho_i E_i \) associated with the eigenvalue \( \rho_i \). On the other hand, recall that the hypergraph connectivity implies that \( \ker 1^T = \sum_{i=1}^\ell \text{span} \Omega_i = \sum_{i=1}^\ell \text{span} E_i \) and hence we have that, besides the zero eigenvalue, the eigenvalues of \( E_{\text{ave}} \) are exactly \( \rho_1, \ldots, \rho_\ell \) and so
\[
\beta = 1 - \min_{i=1,\ldots,\ell} \{ \rho_i \}.
\]
The bound \( \beta \) is minimized by taking \( \rho_i = 1/\ell \) for all \( i \), obtaining
\[
\beta = 1 - \frac{1}{\ell}.
\]
Notice finally that \( E_i E_j = 0 \) for all \( i \neq j \) implies that \( \sum_{i=1}^\ell \dim \text{span} E_i \) is a direct sum and then
\[
m - 1 = \sum_{i=1}^\ell \dim \text{span} E_i = \sum_{i=1}^\ell (|C_i| - 1) = \sum_{i=1}^\ell |C_i| - \ell.
\]
If all the sets \( C_i \) have the same cardinality \( c \), then \( \ell c = m + \ell - 1 \), and so
\[
\beta = 1 - \frac{1}{\ell} = 1 - \frac{c - 1}{m - 1}
\]
which, according to Theorem 12, shows the optimality of the hypergraph.

**VII. Simulations**

In this section we present numerical simulations to validate both the model presented in Section III, and the randomized algorithm proposed in Section IV.

To do so, we considered a 4.8 kV testbed inspired from the standard testbed IEEE 37 [30], which is an actual portion of power distribution network located in California. We however assumed that load are balanced, and therefore all currents and voltages can be described in a single-phase phasorial notation.

As shown in Figure 2, some of the nodes are microgenerators and are capable of injecting a commanded amount of reactive power. Notice that the PCC (node 0) belongs to the set of compensators. This means that the microgrid is allowed to change the amount of reactive power gathered from the transmission grid, if this reduces the power distribution losses. The nodes which are not compensators, are a blend of constant-power, constant-current, and constant-impedance loads, with a total power demand of almost 2 MW of active power and 1 MVAR of reactive power (see [30] for the testbed data).

\[\text{This definition of edge-disjoint for this specific scenario resembles a similar concept for routing problems in data networks, as for example [35].}\]
Figure 2. Schematic representation of the IEEE37 testbed. Circled nodes represent microgenerators taking part to the distributed reactive power compensation. The thick curved lines represent clusters of size $c = 2$ (i.e. edges of $\mathcal{H}$) connecting pair of compensators that can communicate and coordinate their behavior.

The length of the power lines range from a minimum of 25 meters to a maximum of almost 600 meters. The impedance of the power lines differs from edge to edge (for example, resistance ranges from $0.182 \ \Omega/km$ to $1.305 \ \Omega/km$). However, the inductance/resistance ratio exhibits a smaller variation, ranging from $\angle z(e) = 0.47$ to $\angle z(e) = 0.59$. This justifies Assumption 3, in which we claimed that $\angle z(e)$ can be considered constant across the network. The effects of this approximation will be commented later.

Without any distributed reactive power compensation, distribution power losses amount to $61.6 \ kW$, $3.11\%$ of the delivered active power.

Given this setup, we first estimate the quality of the linear approximated model proposed in Section III. As shown in Figure 3, the approximation error results to be negligible.

We then simulated the behavior of the algorithm proposed in Section IV. We considered the two following clustering choices:

- **edge-disjoint gossip**: motivated by the result stated in Section VI, we enabled pairwise communication between compensator in a way that guarantees that the clusters (pairs) are edge-disjoint; the resulting hypergraph $\mathcal{H}$ is represented as a dashed line in Figure 2;

- **star topology**: clusters are in the form $C_v = \{0, v\}$ for all $v \in C \backslash \{0\}$; the reason of this choice is that, as $0$ is the PCC, the constraint $1^T q = 0$ is inherently satisfied: whatever variation in the injected reactive power is applied by $v$, it is automatically compensated by a variation in the demand of reactive power from the transmission grid via the PCC.

In Figure 4 we plotted the result of a single execution of the algorithm in the *edge-disjoint gossip* case. One can see that the algorithm converges quite fast, reducing losses to a minimum that is extremely close to the best achievable solution. The results achieved by the proposed algorithm on this testbed are summarized in the following table.
Figure 3. Comparison between the network state (node voltages) computed via the exact model induced by (1), (2), (4), and (3), and the approximate model proposed in Proposition 2.

Figure 4. Power distribution losses resulting by the execution of the proposed algorithm. A edge-disjoint hypergraph, yielding optimal convergence speed, has been adopted. The dashed line represent the minimum losses that can be achieved by properly choosing the reactive power injected by the compensators.

|                      | Losses before optimization | 61589 W | Fraction of delivered power | 3.11 % |
|----------------------|----------------------------|---------|----------------------------|--------|
|                      | Losses after optimization  | 50338 W | Fraction of delivered power | 2.55 % |
|                      | losses reduction           | 18.27 % |

The minimum losses $J^{opt}$ has been obtained via numerical optimization performed on the exact microgrid model, and represent the minimum losses that can be achieved by properly choosing the amount of reactive power injected by the compensators (and retrieved from the PCC). The difference between this minimum and the
minimum achieved by the algorithm proposed in this paper is partly due to the approximation that we introduced when we modeled the microgrid (assuming large nominal voltage $U_0$), and partly due to the assumption that $\theta$ is constant across the network. Further simulative investigation on this testbed showed that the effect of this last assumption is largely predominant, compared to the effects of the approximation in the microgrid state equations. Still, the combined effect of these two terms is minimal, in practice.

In Figure 5 we compared the behavior of the two considered clustering strategies, plotted together with the curve representing the best achievable performance as given in Theorem 12. One can notice different things. First, both strategies converge to the same minimum, which is slightly larger than the minimum losses that could be achieved by solving the original, nonconvex optimization problem. As expected, the clustering choice do not affect the steady state performance of the algorithm. Second, the performance of the edge-disjoint gossip algorithm results indeed to be better than the star topology in the long-time regime, as the analysis in Section VI suggests, and corresponds to the fastest achievable performance, as predicted by Theorem 12. Third, one can see that, while the performance metric that we adopted is meaningful for describing the long-time regime, it does not describe the initial stage (or short-time behavior), when the full dynamics of the system (19) contribute to the algorithm behavior. This fact suggests that a different metric should possibly be adopted, in order to better characterize the two regimes. The choice of the most appropriate metric requires a preliminary characterization of how fast the reactive power demands change in the microgrid, compared with the rate of execution of the algorithm iterations. If the time-varying demands are much slower than the algorithm, then we expect that the analysis of the slowest eigenvalue remains the most meaningful choice. On the contrary, if the rate of change of demands is comparable with the rate of execution of the algorithm, then we expect that a different index will have to be considered to characterize the algorithm performance.

VIII. CONCLUSIONS

The proposed model for the problem of optimal reactive power compensation in smart microgrids exhibits two main features. First, it can be casted into the framework of quadratic optimization, for which robust solvers are available and the performance analysis becomes tractable; second, it shows how the physics of the system...
can be exploited to design a distributed algorithm for the problem.

On the basis of the proposed model, we have been able to design a distributed, leaderless and randomized algorithm for the solution of the optimization problem, requiring only local communication and local knowledge of the network topology.

We then proposed a metric for the performance of the algorithm, for which we are able to provide a bound on the best achievable performances. We are also able to tell which clustering choice is capable of giving the optimal performances. It is interesting that the optimal strategy requires short-range communications: this is somehow surprising, considered that consensus algorithm (which share many features with the proposed algorithm) benefit from long-range communication that shorten the graph diameter. At the same time, this result is also motivating from the technological point of view, because achieving cooperation of inverters that are far apart in the microgrid presents a number of issues (e.g. time synchronization, limited range of power line communication technologies).

Motivated by the final remarks of Section VII, we plan to add two elements to the picture in the next future in order to understand if the proposed metric is really meaningful for the real problem. These elements are the network dynamics (both the dynamic behavior of the power demands, and the electrical response of the system when actuated), and the presence of operational constraints for the compensators, whose capability of injecting reactive power is limited and possibly time varying.

APPENDIX

PROOF OF PROPOSITION 2

We first introduce the new variable \( \epsilon := 1/U_0 \). This way we can see the currents \( i \) and the voltages \( u \) as functions \( i(\epsilon), u(\epsilon) \) of \( \epsilon \), determined by the equations

\[
\begin{align*}
  u &= X i + \epsilon^{-1} 1 \\
  1^T i &= 0 \\
  u_v &= s_v |\epsilon u_v|^\eta_v, \quad \forall v \in V \setminus \{0\}
\end{align*}
\]

We want to give the Taylor approximation of \( i(\epsilon) \) around \( \epsilon = 0 \).

Let us define

\[
\begin{align*}
  \delta_v(\epsilon) &:= i_v(\epsilon) \epsilon^{-1} - s_v \\
  \lambda_v(\epsilon) &:= u_v(\epsilon) \epsilon - 1
\end{align*}
\]

and substitute them into (28). We get a system of equations in \( \delta, \lambda, \epsilon \) which can be written as

\[
\begin{align*}
  G_v(\delta, \lambda, \epsilon) &= 0, \quad \forall v \in V \\
  F_v(\delta, \lambda, \epsilon) &= 0, \quad \forall v \in V
\end{align*}
\]

where

\[
\begin{align*}
  G_0(\delta, \lambda, \epsilon) &:= \lambda_0 \\
  G_v(\delta, \lambda, \epsilon) &:= \lambda_v - \sum_{w \in V} X_{vw}(\delta_w + \delta_v)|\epsilon|^2, \quad \forall v \neq 0 \\
  F_0(\delta, \lambda, \epsilon) &:= \sum_{w \in V} \delta_w \\
  F_v(\delta, \lambda, \epsilon) &:= (1 + \lambda_v)(s_v + \delta_v) - s_v |1 + \lambda_v|^\eta_v, \quad \forall v \neq 0
\end{align*}
\]
We have $2n$ complex equations in $2n + 1$ complex variables.

We interpret now the complex numbers as vectors in $\mathbb{R}^2$ and $G_v, F_v$ as functions from $\mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R}^2$ to $\mathbb{R}^2$.

Observe that $G(\delta, \lambda, \epsilon)_{|\delta=0,\lambda=0,\epsilon=0} = 0$ and $F_v(\delta, \lambda, \epsilon)_{|\delta=0,\lambda=0,\epsilon=0} = 0$ for all $v \in V$. By the implicit function theorem, in order to prove that $\delta(\epsilon), \lambda(\epsilon)$ are infinitesimal in $\epsilon$, it is enough to prove that the matrix

\[
\begin{pmatrix}
\left(\frac{\partial G_v}{\partial \delta}\right)_{v, w \in V} & \left(\frac{\partial G_v}{\partial \lambda}\right)_{v, w \in V} \\
\left(\frac{\partial F_v}{\partial \delta}\right)_{v, w \in V} & \left(\frac{\partial F_v}{\partial \lambda}\right)_{v, w \in V}
\end{pmatrix}
\]

evaluated in $\delta = 0, \lambda = 0, \epsilon = 0$ is invertible.

In order to determine this matrix we need to introduce the following notation. If $a \in \mathbb{C}$, and $a = a^R + ja^I$ with $a^R, a^I \in \mathbb{R}$, then we define

\[
\langle a \rangle := \begin{pmatrix} a^R & -a^I \\ a^I & a^R \end{pmatrix}
\]

With this notation observe that, if $a, x \in \mathbb{C}$ and we consider those complex numbers as vectors in $\mathbb{R}^2$ and if we consider the function $f(x) := ax$ as a function from $\mathbb{R}^2$ to $\mathbb{R}^2$, then we have that $\partial f/\partial x = \langle a \rangle$. Notice moreover that the function $g(x) := \bar{x}$ has $\partial g/\partial x = N := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

From these observations we can argue that

\[
\frac{\partial G_0}{\partial \delta} = 0, \quad \frac{\partial G_0}{\partial \lambda} = \begin{cases} I & \text{if } w = 0 \\ 0 & \text{if } w \neq 0 \end{cases}
\]

\[
\frac{\partial F_0}{\partial \delta} = I, \quad \frac{\partial F_0}{\partial \lambda} = 0
\]

Instead for all $v \neq 0$ we have that

\[
\frac{\partial G_v}{\partial \lambda} = \begin{cases} I & \text{if } w = v \\ 0 & \text{if } w \neq v \end{cases}
\]

while

\[
\frac{\partial G_v}{\partial \delta} = \langle -X_{vw} \epsilon^2 \rangle \quad \forall w \in V
\]

which evaluated in $\epsilon = 0$ yields 0. Observe finally that, if $w \neq v$, then $\partial F_v/\partial \lambda_v = \partial F_v/\partial \delta_v = 0$ and that

\[
\frac{\partial F_v}{\partial \delta_v} = \langle 1 + \lambda_v \rangle N
\]

which evaluated in $\lambda_v = 0$ yields the matrix $N$. On the other hand,

\[
\frac{\partial F_v}{\partial \lambda_v} = (s_v + \delta_v) - s_v \frac{\partial [1 + \lambda_v]^{\eta_v}}{\partial \lambda_v}
\]

where $s_v$ is a 2-dimensional column vector and $\frac{\partial [1 + \lambda_v]^{\eta_v}}{\partial \lambda_v}$ is a 2-dimensional row vector. With some easy computations it can be seen that

\[
\frac{\partial [1 + \lambda_v]^{\eta_v}}{\partial \lambda_v} = \eta_v \left[ (1 + \lambda_v^R)^2 + (\lambda_v^I)^2 \right]^{\eta_v-1} \left[ 1 + \lambda_v^R \lambda_v^I \right]
\]
which evaluated in $\lambda_v = 0$ yields $[\eta_v \ 0]$. Hence in $\lambda_v = 0$ we have

$$\frac{\partial F_v}{\partial \lambda_v} = Q_v := \langle s_v \rangle - s_v[\eta_v \ 0]$$

We can conclude that for $\delta = 0$, $\lambda = 0$, $\epsilon = 0$ we have

$$\begin{bmatrix}
\left( \frac{\partial G_v}{\partial \delta_w} \right)_{v,w \in V} & \left( \frac{\partial G_v}{\partial \lambda_w} \right)_{v,w \in V} \\
\left( \frac{\partial F_v}{\partial \delta_w} \right)_{v,w \in V} & \left( \frac{\partial F_v}{\partial \lambda_w} \right)_{v,w \in V}
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & \cdots & 0 & I & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & I & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & I
\end{bmatrix}
\begin{bmatrix}
I & I & \cdots & I & 0 & 0 & \cdots & 0 \\
0 & N & \cdots & 0 & 0 & Q_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & N & 0 & 0 & \cdots & Q_{n-1}
\end{bmatrix}$$

which is invertible.

REFERENCES

[1] D. P. Bertsekas and J. N. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods. Athena Scientific, 1997.

[2] Special issue on technology of networked control systems, Proceedings of the IEEE, vol. 95, no. 1, Jan. 2007.

[3] E. Santacana, G. Rackliffe, L. Tang, and X. Feng, “Getting smart, with a clearer vision of the intelligent grid, control emerges from chaos,” IEEE Power Energy Magazine, vol. 8, no. 2, pp. 41–48, Mar. 2010.

[4] A. Ipakchi and F. Albuyeh, “Grid of the future. are we ready to transition to a smart grid?” IEEE Power Energy Magazine, vol. 7, no. 2, pp. 52 –62, Mar.-Apr. 2009.

[5] J. A. Lopes, C. L. Moreira, and A. G. Madureira, “Defining control strategies for microgrids islanded operation,” IEEE Transactions Power Systems, vol. 21, no. 2, pp. 916–924, May 2006.

[6] T. C. Green and M. Prodanovic, “Control of inverter-based micro-grids,” Electric Power Systems Research, vol. 77, no. 9, pp. 1204–1213, Jul. 2007.

[7] F. Katifra and M. R. Iravani, “Power management strategies for a microgrid with multiple distributed generation units,” IEEE Transactions on Power Systems, vol. 21, no. 4, pp. 1821–1831, Nov. 2006.

[8] M. Prodanovic, K. De Brabandere, J. Van den Keybus, T. Green, and J. Driesen, “Harmonic and reactive power compensation as ancillary services in inverter-based distributed generation,” IET Generation, Transmission & Distribution, vol. 1, no. 3, pp. 342–348, 2007.

[9] E. Tedeschi, P. Tenti, and P. Mattaveli, “Synergistic control and cooperative operation of distributed harmonic and reactive compensators,” in IEEE PESC 2008, 2008.

[10] A. Rabiee, H. A. Shayanfar, and N. Amjady, “Reactive power pricing: Problems and a proposal for a competitive market,” IEEE Power Energy Magazine, vol. 7, no. 1, pp. 18–32, Jan. 2009.

[11] H. Yoshida, K. Kawata, Y. Fukuyama, S. Takayama, and Y. Nakanishi, “A particle swarm optimization for reactive power and voltage control considering voltage security assessment,” IEEE Transactions on Power Systems, vol. 15, no. 4, pp. 1232–1239, Nov. 2000.

[12] B. Zhao, C. X. Guo, and Y. J. Cao, “A multiagent-based particle swarm optimization approach for optimal reactive power dispatch,” IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 1070–1078, May 2005.

[13] J. Lavaei, A. Rantzer, and S. H. Low, “Power flow optimization using positive quadratic programming,” in Proceedings of the 18th IFAC World Congress, 2011.

[14] F. Dorfler and F. Bullo, “Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators,” in Proceedings of the 2010 American Control Conference (ACC), 2010, pp. 930–937.

[15] Y. Guo, D. J. Hill, and Y. Wang, “Nonlinear decentralized control of large-scale power system,” Automatica, vol. 36, no. 9, pp. 1275–1289, Sep. 2000.

[16] P. Tenti, A. Costabber, P. Mattaveli, and D. Trombetti, “Distribution loss minimization by token ring control of power electronic interfaces in residential micro-grids,” IEEE Transactions on Industrial Electronics, to appear.
[17] K. Turitsyn, P. Šulc, S. Backhaus, and M. Chertkov, “Options for control of reactive power by distributed photovoltaic generators,” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.

[18] IEEE Task Force on Load representation for dynamic performance, “Load representation for dynamic performance analysis,” *IEEE Transactions on Power Systems*, vol. 8, no. 2, pp. 472–482, May 1993.

[19] M. H. Haque, “Load flow solution of distribution systems with voltage dependent load models,” *Electric Power Systems Research*, vol. 36, pp. 151–156, 1996.

[20] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, *Electric energy systems: Analysis and operation*. CRC Press, 2009.

[21] A. Ghosh, S. Boyd, and A. Saberi, “Minimizing effective resistance of a graph,” *SIAM Review*, vol. 50, no. 1, pp. 37–66, Feb. 2008.

[22] M. E. Baran and F. F. Wu, “Optimal sizing of capacitors placed on a radial distribution system,” *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735–743, Jan. 1989.

[23] A. Cagnano, E. De Tuglie, M. Liserre, and R. Mastromauro, “On-line optimal reactive power control strategy of PV-inverters,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4549–4558, Oct. 2011.

[24] A. Nedic and A. Ozdaglar, “Distributed subgradient methods for multi-agent optimization,” *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[25] A. G. Phadke and J. S. Thorp, *Synchronized phasor measurements and their applications*. Springer, 2008.

[26] A. Kelly, A. K. Mallik, and D. K. H. Tan, “Rate control for communication networks: shadow prices, proportional fairness and stability,” *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, Mar. 1998.

[27] S. Bolognani and S. Zampieri, “A gossip-like distributed optimization algorithm for reactive power flow control,” in *Proceedings of the IFAC World Congress 2011*, Milano, Italy, Aug. 2011.

[28] A. Srivastava, “Improved approximations for edge-disjoint paths, unsplittable flow, and related routing problems,” in *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, 1997.