

Distributed control for optimal 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. We also show how agents have a partial knowledge of the problem parameters and state via some local measurements. Then, we design a randomized, gossip-like optimization algorithm, providing conditions for convergence together with an analytical characterization of the convergence speed. The analysis shows that the best performance is achieved when we command cooperation among agents that are neighbors in the smart microgrid topology.

CONTENTS

I	Introduction	1
II	Mathematical preliminaries and notation	2
III	Microgrid model	3
III-A	Power flow approximated model	4
IV	Optimal Reactive Power Flow problem	5
V	A randomized distributed algorithm	6
V-A	Convergence results	7
V-B	Bound on the rate of convergence	8
V-C	Optimal strategy: nearest-neighbor gossip	10
VI	Simulations	11
VII	Conclusions	12
	References	12

I. INTRODUCTION

Most of the distributed optimization methods have been intended as a way of dispatching part of a large scale optimization algorithm to different processing units [1]. When the same methods are applied to *networked control systems*, 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; moreover, the information exchange between agents happens not only via a given communication channel, but also via local actuation and measurement performed on a common underlying physical system.

The extent of these issues depends on the particular application, and in this work we present the specific scenario of smart electrical power distribution networks.

In the last decade, the introduction of distributed microgeneration (driven by global environmental issues), together with the increased demand for electric power and for higher quality of service, has

The research leading to these results has received funding from the EU 7th Framework Programme [FP7/2007-2013] under grant agreements no. 257462 HYCON2 and no. 223866 FeedNetBack.

The authors are with the University of Padova, Italy, email: {saverio.bolognani, zampieri}@dei.unipd.it.

Accepted at IEEE CDC 2011, Orlando (FL), USA, December 2011.

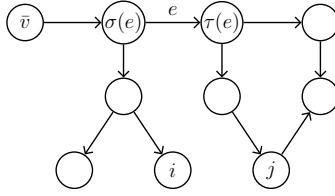


Fig. 1. Graph describing the adopted microgrid model.

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 distributed control of smart microgrids. A microgrid is a portion of the low-voltage power distribution network that is managed autonomously from the rest of the network, to achieve better quality of the service, to improve efficiency, and to pursue specific economic interests. Together with the loads connected to the microgrid (both residential and industrial customers), we also have microgeneration devices (solar panels, 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* [2], [3]: 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 [4].

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 [4].

In Section III and IV, we propose a model for the problem of optimal reactive power flow (ORPF) in a microgrid, showing how it can be casted into the framework of distributed convex optimization. In Section V we propose a distributed randomized algorithm for this problem, and we analyze its performance by providing relevant bounds and a result on the best achievable behavior. In Section VI, we finally validate the proposed model and we simulate the behavior of the proposed optimization method.

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 $\tau(e)$ (see Figure 1).

In the rest of the paper we will often introduce complex-valued functions defined on the nodes and on the edges. By indexing nodes and edges via the integers $1, \dots, N$, and $1, \dots, N_{\mathcal{E}}$ respectively (where $N = |\mathcal{V}|, N_{\mathcal{E}} = |\mathcal{E}|$), these functions will also be intended as vectors in \mathbb{C}^N and $\mathbb{C}^{N_{\mathcal{E}}}$.

Let moreover $A \in \{0, \pm 1\}^{N_{\mathcal{E}} \times N}$ 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 the graph \mathcal{G} is connected, then $\mathbf{1}$ (the vector of all 1’s) is the only vector in $\ker A$.

An undirected graph \mathcal{G} is a graph in which $e \in \mathcal{E} \Rightarrow e' \in \mathcal{E}$, where $\sigma(e') = \tau(e)$ and $\tau(e') = \sigma(e)$. Neglecting multiple edges and self loops, we can also intend 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. MICROGRID MODEL

For the purpose of this paper, we model a microgrid as a directed graph \mathcal{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 ω_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 represents the signal $\mathbf{y}(t) = |y|\sqrt{2}\sin(\omega_0 t + \angle y)$.

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

- $u : \mathcal{V} \rightarrow \mathbb{C}$ is the voltage at the nodes;
- $i : \mathcal{V} \rightarrow \mathbb{C}$ is the current injected by the nodes;
- $\xi : \mathcal{E} \rightarrow \mathbb{C}$ is the current flowing on the edges.

Moreover, let us define the following power flows in the microgrid:

- $s : \mathcal{V} \rightarrow \mathbb{C}$, where $s(v) := u(v)i(v)^*$ is the (complex) power injected by node v into the grid;
- $f_\sigma : \mathcal{E} \rightarrow \mathbb{C}$, where $f_\sigma(e) := u(\sigma(e))\xi(e)^*$ is the power flow entering the edge e ;
- $f_\tau : \mathcal{E} \rightarrow \mathbb{C}$, where $f_\tau(e) := u(\tau(e))\xi(e)^*$ is the power flow exiting the edge e .

The following constraints are satisfied by u, i and ξ .

$$A^T \xi + i = 0, \quad \mathbf{1}^T i = 0 \quad (1)$$

$$Au + Z\xi = 0, \quad (2)$$

where $Z = \text{diag}(z(e), e \in \mathcal{E})$, $z : \mathcal{E} \rightarrow \mathbb{C}$ being the impedance of the microgrid power lines. 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 characterized by a law relating its injected current $i(v)$ with its voltage $u(v)$. We assume that all the nodes, but the PCC, inject (or are supplied with, if negative) a constant power into the microgrid:

$$u(v)i(v)^* = \bar{s}_v, \quad \forall v \in \mathcal{V} \setminus \{\bar{v}\}. \quad (3)$$

This is a good approximation for all the devices that are connected to the grid via an electronic interface, or inverter. The vast majority of microgeneration devices fits in this category, together with most industrial and residential loads.

Instead, we model the node corresponding to the PCC (indexed as \bar{v}) as a constant voltage generator, i.e.

$$u(\bar{v}) = u_0. \quad (4)$$

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 power at every node, and the nominal voltage at the PCC, has been extensively covered in the literature under the denomination of *power flow analysis* (see for example [5, Chapter 3]).

A. Power flow approximated model

The power flows f_σ , f_τ , and s must satisfy the following equations, which descend directly from Kirchhoff's laws (1) and (2):

$$s(v) + \sum_{\tau(e)=v} f_\tau(e) - \sum_{\sigma(e)=v} f_\sigma(e) = 0, \quad \forall v \in \mathcal{V} \quad (5)$$

$$f_\sigma(e) = u(\sigma(e)) \frac{[u(\sigma(e)) - u(\tau(e))]^*}{z(e)^*}, \quad \forall e \in \mathcal{E} \quad (6)$$

Equation (5) guarantees power conservation at the nodes, while (6) expresses the power entering an edge as a function of the voltage drop on the same edge.

If the power losses $\ell(e) := f_\sigma(e) - f_\tau(e)$ on any edge $e \in \mathcal{E}$ are much smaller than the power delivered by the same edge, and voltage drops on the edges are much smaller than the nominal voltage u_0 , then we can replace (5) and (6) with the system of linear equations

$$A^T f_\sigma + s = 0 \quad (7)$$

$$f_\sigma + Y^* u_0 (Au)^* = 0, \quad (8)$$

where $Y = Z^{-1}$, and $s \in \mathbb{C}^N$ is the vector of all power injections $s(v)$, $v \in \mathcal{V} \setminus \{\bar{v}\}$, completed with $s(\bar{v}) = -\sum_{v \in \mathcal{V} \setminus \{\bar{v}\}} s(v)$.

In other words, the network constraints (7) and (8), together with the node equations (3) and (4), constitute an approximated model for the microgrid, linear in the variables s , f , and u . The following result expresses the approximation error analytically, showing how this error goes to zero as the nominal voltage of the grid grows in absolute value. This is consistent with the common design practice for power distribution network, where relatively high voltages are used to deliver power to the loads while having small losses.

We first introduce an assumption of the impedance of the edges of the microgrid, which simplifies the notation in the rest of this modeling section, and becomes a necessary assumption in the derivation of the proposed algorithm.

Assumption 1: The inductance-resistance ratio is fixed for all the edges, i.e.

$$z(e) = e^{j\theta} d(e), \quad d(e) \in \mathbb{R}^+, \quad \forall e \in \mathcal{E}.$$

We therefore have $Z = e^{j\theta} D$ and $Y = e^{-j\theta} D^{-1}$, where $D = \text{diag}(d(e), e \in \mathcal{E})$. We also introduce the weighted Laplacian matrix $L = A^T D^{-1} A$, and the *Green matrix* X [6], defined as the unique real, symmetric matrix satisfying both conditions¹

$$\begin{cases} XL = \Omega \\ X\mathbf{1} = 0, \end{cases}$$

where $\Omega = I - \mathbf{1}\mathbf{1}^T/N$. It is easy to show that $X \geq 0$ and that it has only one eigenvalue in zero. It also holds

$$Z_{\text{eff}}(u, v) = e^{j\theta} (e_u - e_v)^T X (e_u - e_v), \quad (9)$$

where $Z_{\text{eff}} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$ is the effective impedance between nodes, and where e_v is a vector of all zeros except for $e_v(v) = 1$.

The Green matrix allows us, via some simple algebraic manipulation of (7) and (8), to approximate node voltages u as a linear function of the injected power s :

$$\hat{u} = \frac{e^{j\theta}}{u_0^*} (Xs^* - e_{\bar{v}}^T Xs^* \mathbf{1}) + u_0 \mathbf{1}. \quad (10)$$

The following holds.

¹ X can be constructed as a pseudoinverse $L^\#$ of L , and in particular as $X = (L + \mathbf{1}\mathbf{1}^T/N)^{-1} - \mathbf{1}\mathbf{1}^T/N$.

Proposition 2: There exists $\hat{i} \in \mathbb{C}^N$ such that the solution (10) of the linear system (7), (8), satisfies the equations (1), (2), and (4) of the exact model. Moreover, \hat{u} and \hat{i} satisfy (3) up to an error that goes to zero as $|u_0|$ goes to infinity:

$$\lim_{|u_0| \rightarrow \infty} \left| \hat{u}(v) \hat{i}(v)^* - s(v) \right| = 0 \quad \forall v \in \mathcal{V} \setminus \{\bar{v}\}.$$

Proof: From (7) and (8) we have

$$\hat{i} = e^{-j\theta} L \hat{u} = \frac{1}{u_0^*} L X \hat{s}^* = \frac{1}{u_0^*} \hat{s}^*.$$

Therefore we have

$$\begin{aligned} \hat{u}(v) \hat{i}(v)^* &= \left(\frac{e^{j\theta}}{u_0^*} e_v^T X \hat{s}^* - \frac{e^{j\theta}}{u_0^*} e_{\bar{v}}^T X \hat{s}^* + u_0 \right) \frac{1}{u_0} \hat{s} \\ &= \frac{e^{j\theta}}{|u_0|^2} (e_v - e_{\bar{v}})^T X \hat{s}^* \hat{s}(v) + \hat{s}(v), \end{aligned}$$

from which the result follows. The fact that \hat{u} and \hat{i} satisfy (1), (2), and (4), can be verified by simple substitution. \blacksquare

A numerical validation of the approximated model proposed in this section is presented in Section VI.

IV. OPTIMAL REACTIVE POWER FLOW PROBLEM

The total active power losses on the edges are given by

$$\sum_{e \in \mathcal{E}} \text{Re } \ell(e) = \sum_{e \in \mathcal{E}} |\xi(e)|^2 \text{Re } z(e) = \cos \theta (u^*)^T L u,$$

where we used (2) to express ξ as a function of u . By using the approximate solution \hat{u} given in (10), we have

$$\sum_{e \in \mathcal{E}} \text{Re } \ell(e) = \frac{\cos \theta}{|u_0|^2} (s^*)^T X s, \quad (11)$$

where we exploited the properties of X and the fact that $\mathbf{1}^T s = 0$. Note that (11) is convex in s , as $X \geq 0$ and its zero-eigenvector $\mathbf{1}$ is orthogonal to the constraint $\mathbf{1}^T s = 0$.

Let us now consider the decomposition of the decision variable s into the injected active power $p = \text{Re}(s)$ and the injected reactive power $q = \text{Im}(s)$. Expression (11) for the power losses results to be separable in these two decision variables, as

$$\frac{\cos \theta}{|u_0|^2} (s^*)^T X s = \frac{\cos \theta}{|u_0|^2} (p^T X p + q^T X q),$$

while the constraint $\mathbf{1}^T s = 0$ decouples into $\mathbf{1}^T p = 0$ and $\mathbf{1}^T q = 0$.

We are allowed to command only a subset $\mathcal{C} \subset \mathcal{V}$ of the electronic interfaces (active agents, or compensators) connected to the microgrid. 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 generally injected into the grid to replace generation from traditional plants, which are more expensive and exhibit a worse environmental impact). Via a proper reordering of the nodes' indices, we can introduce a block-form for both q and X , obtaining

$$q = \begin{bmatrix} q_{\mathcal{C}} \\ q_{\bar{\mathcal{C}}} \end{bmatrix}, \quad X = \begin{bmatrix} X_{\mathcal{C}\mathcal{C}} & X_{\mathcal{C}\bar{\mathcal{C}}} \\ X_{\bar{\mathcal{C}}\mathcal{C}} & X_{\bar{\mathcal{C}}\bar{\mathcal{C}}} \end{bmatrix},$$

where the subscript \mathcal{C} identifies the variables corresponding to the compensators, while $\bar{\mathcal{C}}$ identifies the variables corresponding to the nodes in $\mathcal{V} \setminus \mathcal{C}$.

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

$$\min_{\mathbf{1}^T q_c = c} J(q_c), \quad \text{where} \quad J(q_c) = q_c^T M q_c + m^T q_c, \quad (12)$$

and where $M = \frac{\cos \theta}{|u_0|^2} X_{CC} > 0$, $m = \frac{2 \cos \theta}{|u_0|^2} X_{CC} q_{\bar{c}}$, and $c = -\mathbf{1}^T q_{\bar{c}}$.

The solution of the optimization problem (12) would not pose any challenge if the nodes knew the problem parameters M , m , and c . These quantities depend on both the grid parameters and the power demand of the whole microgrid, therefore it is impractical that every agent is capable of retrieving all this information, and we also exclude that a centralized agent is allowed to collect all the necessary data.

Instead, we want to explore what information on the optimization problem the agents can infer from local measurements. In particular, we focus on the gradient $\nabla J(q_c) = 2Mq_c + m$. The component of the gradient corresponding to a node v results to be

$$[\nabla J]_v = \frac{2 \cos \theta}{|u_0|^2} e_v^T X q.$$

Via equation (10), as $q = \text{Im}(\hat{s})$, we obtain

$$\begin{aligned} [\nabla J]_v &= \frac{2 \cos \theta}{|u_0|^2} e_v^T \text{Im}(X \hat{s}) \\ &= \frac{2 \cos \theta}{|u_0|^2} e_v^T \text{Im}(e^{j\theta} u_0 u^*) + \alpha \end{aligned}$$

where $\alpha = -\frac{2 \cos \theta}{|u_0|^2} \text{Im}(e^{j\theta} u_0 \mathbf{1}^T u^*) / N$ is common to all the nodes but unknown to any of them.

Therefore the nodes of a microgrid can obtain (element-wise) an estimate

$$\widehat{\nabla J} = \frac{2 \cos \theta}{|\hat{u}_0|^2} \text{Im}[e^{j\theta} \hat{u}_0 u^*]. \quad (13)$$

of the gradient of the cost function up to a common constant, provided that they also agree on estimate \hat{u}_0 of the nominal voltage.

V. A RANDOMIZED DISTRIBUTED ALGORITHM

Let the compensators be divided into ℓ possibly overlapping sets $\mathcal{C}_1, \dots, \mathcal{C}_\ell$, with $\bigcup_{i=1}^{\ell} \mathcal{C}_i = \mathcal{C} = \{1, \dots, N_C\}$. This family of subsets can be interpreted as the edges of a hypergraph \mathcal{H} defined over the set of nodes \mathcal{C} .

Nodes belonging to the same set are able to communicate each other, and they are therefore capable of coordinating their action and sharing measurements. We assume that, by using this information, nodes belonging to the same set are capable of driving their state in a new feasible state that minimizes $J(q_c)$, solving the optimization *subproblem* in which all the nodes that are not in \mathcal{C}_i keep their state constant:

$$\arg \min_{\Delta q_c \in \mathcal{S}_i} J(q_c + \Delta q_c), \quad (14)$$

where $\mathcal{S}_i := \left\{ q \in \mathbb{R}^{N_C} : \sum_{j \in \mathcal{C}_i} q_j = 0, q_j = 0 \forall j \notin \mathcal{C}_i \right\}$.

The optimization subproblem faced by the nodes in \mathcal{C}_i can then be rewritten as

$$\begin{aligned} \min_{q_{\mathcal{C}_i}} \quad & q_{\mathcal{C}_i}^T M_{\mathcal{C}_i \mathcal{C}_i} q_{\mathcal{C}_i} + (2q_{\bar{c}}^T M_{\bar{c} \mathcal{C}_i} + m_{\mathcal{C}_i}^T) q_{\mathcal{C}_i} \\ \text{subject to} \quad & \mathbf{1}^T q_{\mathcal{C}_i} = c - \mathbf{1}^T q_{\bar{c}_i}. \end{aligned} \quad (15)$$

It is easy to see that agents in \mathcal{C}_i can reach the optimal solution by adding to $q_{\mathcal{C}_i}$ the increment $\Delta q_{\mathcal{C}_i}$ given by

$$\Delta q_{\mathcal{C}_i} = -\frac{M_{\mathcal{C}_i \mathcal{C}_i}^{-1}}{2} [\nabla J]_{\mathcal{C}_i} + \frac{\mathbf{1}^T M_{\mathcal{C}_i \mathcal{C}_i}^{-1} [\nabla J]_{\mathcal{C}_i}}{\mathbf{1}^T M_{\mathcal{C}_i \mathcal{C}_i}^{-1} \mathbf{1}} \frac{M_{\mathcal{C}_i \mathcal{C}_i}^{-1}}{2} \mathbf{1},$$

where $M_{\mathcal{C}_i\mathcal{C}_i}^{-1}$ is the inverse of submatrix $M_{\mathcal{C}_i\mathcal{C}_i}$.

It has been shown in Section IV that an estimate of the gradient $\widehat{\nabla J}$ can be obtained by sensing the network voltages. Nodes in \mathcal{C}_i can therefore solve their corresponding optimization subproblem by performing the update

$$\Delta q_{\mathcal{C}_i} = -\frac{M_{\mathcal{C}_i\mathcal{C}_i}^{-1}[\widehat{\nabla J}]_{\mathcal{C}_i}}{2} + \frac{\mathbf{1}^T M_{\mathcal{C}_i\mathcal{C}_i}^{-1}[\widehat{\nabla J}]_{\mathcal{C}_i}}{\mathbf{1}^T M_{\mathcal{C}_i\mathcal{C}_i}^{-1}\mathbf{1}} \frac{M_{\mathcal{C}_i\mathcal{C}_i}^{-1}}{2}\mathbf{1},$$

as the unknown term $\alpha\mathbf{1}$ in the gradient estimate $\widehat{\nabla J}$ gets canceled in the expression.

The update law requires only information that can be gathered inside the subset \mathcal{C}_i , namely the matrix $M_{\mathcal{C}_i\mathcal{C}_i}^{-1} = \left(\frac{\cos\theta}{|u_0|^2} X_{\mathcal{C}_i\mathcal{C}_i}\right)^{-1}$ and the voltage measurements $u_{\mathcal{C}_i}$. This is possible because the physics of the system allow us to estimate the gradient, that otherwise would depend on the whole system state, from some local measurements. Moreover, according to (9), the elements of $M_{\mathcal{C}_i\mathcal{C}_i}$ depend only on the effective impedance between nodes in \mathcal{C}_i , and therefore we can assume one of the following:

- nodes in \mathcal{C}_i have a knowledge of the local topology and therefore of the effective mutual impedance;
- nodes estimate the values $M_{\mathcal{C}_i\mathcal{C}_i}$ by computing the gradient for different values of $\Delta q_{\mathcal{C}_i}$; the class of algorithms that implement this kind of estimation are called quasi-Newton methods, and their application to the linearly constrained case has been explored in [7];
- nodes estimate their mutual effective impedance by performing some identification on the network during an initialization stage, as suggested for example in [8].

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

- 1) a set \mathcal{C}_i is chosen according to a sequence of symbols $\eta(t) \in \{1, \dots, \ell\}$;
- 2) agents in \mathcal{C}_i sense the network and obtain an estimate of the gradient;
- 3) they determine a feasible update step that minimizes the given cost function;
- 4) they actuate the system by updating their state (the injected reactive power).

The iterated algorithm will then results in the following discrete time system for q

$$q_{\mathcal{C}}(t+1) = T_{\eta(t)}[q_{\mathcal{C}}(t)] := \arg \min_{\Delta q_{\mathcal{C}} \in \mathcal{S}_{\eta(t)}} J(q_{\mathcal{C}}(t) + \Delta q_{\mathcal{C}}), \quad (16)$$

with initial conditions $q_{\mathcal{C}}(0)$ such that $\mathbf{1}^T q_{\mathcal{C}}(0) = c$.

The following notation will be useful in the rest of the paper. Define the $N_{\mathcal{C}} \times N_{\mathcal{C}}$ matrices

$$\Omega_i = I_{\mathcal{C}_i} - \frac{1}{|\mathcal{C}_i|} \mathbf{1}_{\mathcal{C}_i} \mathbf{1}_{\mathcal{C}_i}^T$$

where $|\mathcal{C}_i|$ is the cardinality of the set \mathcal{C}_i , $I_{\mathcal{C}_i}$ is the diagonal matrix having diagonal entries 1 in positions belonging to \mathcal{C}_i and zero elsewhere and $\mathbf{1}_{\mathcal{C}_i}$ is the column vector having entries 1 in positions belonging to \mathcal{C}_i and zero elsewhere. Notice that

$$\Omega_i = \frac{1}{2|\mathcal{C}_i|} \sum_{h,k \in \mathcal{C}_i} (e_h - e_k)(e_h - e_k)^T$$

and that $\mathcal{S}_i = \text{Im } \Omega_i$. Being a projector matrix, we also have $\Omega_i^2 = \Omega_i$ and $\Omega_i^\# = \Omega_i$, where $\#$ means pseudo-inverse.

A. Convergence results

To study the convergence of the proposed algorithm and its speed, we introduce the auxiliary variable $x = q_{\mathcal{C}} - q_{\mathcal{C}}^{\text{opt}}$, where $q_{\mathcal{C}}^{\text{opt}}$ is the solution of the optimization problem (12). By substitution, it can be shown that the optimization problem (12) is equivalent to

$$\min_{\mathbf{1}^T x = 0} V(x) = x^T M x, \quad (17)$$

and that the subproblems described before are equivalent to the subproblems

$$\min_{\Delta x \in \text{Im } \Omega_i} V(x + \Delta x) \quad (18)$$

In this notation, it is possible to explicitly express the solution of the individual subproblems as a linear function of the starting point $x(t)$:

$$x(t+1) = F_i x(t), \quad F_i = I - (\Omega_i M \Omega_i)^\# M. \quad (19)$$

It is easy to verify from the properties of the pseudoinverse that $\ker(\Omega_i M \Omega_i)^\# = \ker \Omega_i$ and $\text{Im}(\Omega_i M \Omega_i)^\# = \text{Im} \Omega_i$, and thus $(\Omega_i M \Omega_i)^\# = (\Omega_i M \Omega_i)^\# \Omega_i$. 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 $\langle x, y \rangle_M := x^T M y$. Moreover, they are self-adjoint matrices with respect to the inner product $\langle \cdot, \cdot \rangle_M$, thus having real eigenvalues.

The discrete time system (16) in the x coordinates corresponds then to the linear time varying system

$$x(t+1) = F_{\eta(t)} x(t). \quad (20)$$

The following result, whose proof is in [9], characterizes the uniqueness of the equilibrium for all maps F_i .

Lemma 3: Consider the family of linear transformations $\{F_i\}$ as described in (19). $\bar{x} = 0$ is the only point in $\ker \mathbf{1}^T$ that is invariant for all F_i 's if and only if

$$\text{Im}[\Omega_1 \dots \Omega_\ell] = \ker \mathbf{1}^T.$$

The condition expressed in Lemma 3 corresponds to a necessary condition for the convergence of the algorithm, and can be also expressed as a connectivity requirement on the hypergraph \mathcal{H} .

Proposition 4: A necessary condition for the convergence of algorithm (16) to the solution q_C^{opt} of the optimization problem (12) is that the hypergraph \mathcal{H} is connected.

Proof: Consider the undirected graph $\mathcal{G}_{\mathcal{H}}$, defined as a weighted graph having nodes the compensators \mathcal{C} and weights on the edge h, k equal to the number of the sets \mathcal{C}_i which contain both h and k . It is quite easy to see that the hypergraph \mathcal{H} with edges \mathcal{C}_i is connected if and only if $\mathcal{G}_{\mathcal{H}}$ is a connected graph.

Let us define $\delta_{\mathcal{C}_i} : \mathcal{C} \rightarrow \{0, 1\}$ 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. Consider then the Laplacian matrix $L_{\mathcal{H}}$ of $\mathcal{G}_{\mathcal{H}}$:

$$\begin{aligned} L_{\mathcal{H}} &= \sum_{h,k \in \mathcal{C}} (e_h - e_k)(e_h - e_k)^T \sum_{i=1}^{\ell} \delta_{\mathcal{C}_i}(h) \delta_{\mathcal{C}_i}(k) \\ &= \sum_{i=1}^{\ell} \sum_{h,k \in \mathcal{C}_i} (e_h - e_k)(e_h - e_k)^T = \sum_i 2|\mathcal{C}_i| \Omega_i \\ &= [\Omega_1 \dots \Omega_\ell] \text{diag}\{2|\mathcal{C}_1|I, \dots, 2|\mathcal{C}_\ell|I\} [\Omega_1 \dots \Omega_\ell]^T. \end{aligned}$$

The condition $\text{Im}[\Omega_1 \dots \Omega_\ell] = \ker \mathbf{1}^T$, which is a necessary condition for the algorithm convergence according to Lemma 3, is equivalent to the fact that $L_{\mathcal{H}} + \mathbf{1}\mathbf{1}^T$ is positive definite, which is a characterization of connectivity of $\mathcal{G}_{\mathcal{H}}$. ■

B. Bound on the rate of convergence

For the study of the rate of convergence of the proposed algorithm, we introduce the following assumption of the random sequence $\eta(t)$.

Assumption 5: The sequence $\eta(t)$ is a sequence of independently, uniformly distributed symbols in $\{1, \dots, \ell\}$.

We consider the following performance metric:

$$R = \sup_{x(0) \in \ker \mathbf{1}^T} \limsup v(t)^{1/t}$$

where $v(t) = \mathbb{E} [V(x(t))]$. R describes the exponential rate of convergence to zero of $v(t)$ and so also the exponential rate of convergence of $q_{\mathcal{L}}(t)$ to the optimal solution $q_{\mathcal{L}}^{\text{opt}}$. Using (16), we have

$$\begin{aligned} v(t) &= \mathbb{E} [x(t)^T M x(t)] \\ &= \mathbb{E} [x(t)^T \Omega M \Omega x(t)] \\ &= \mathbb{E} [x(t-1)^T F_{\eta(t-1)}^T \Omega M \Omega F_{\eta(t-1)} x(t-1)] \\ &= x(0)^T \mathbb{E} [F_{\eta(0)}^T \cdots F_{\eta(t-1)}^T \Omega M \Omega F_{\eta(t-1)} \cdots F_{\eta(0)}] x(0). \end{aligned}$$

Let us then define

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

Via Assumption 5, we can derive the following linear system:

$$\begin{aligned} \Delta(t+1) &= \mathcal{L}(\Delta(t)) := \mathbb{E} [F^T \Delta F], \quad \Delta(0) = \Omega M \Omega \\ \Xi(t) &= \Omega \Delta(t) \Omega, \end{aligned} \tag{21}$$

and express the expected cost function as

$$\mathbb{E} [V(x(t))] = v(t) = x(0)^T \Xi(t) x(0).$$

Let denote by \mathbf{F} the $N^2 \times N^2$ matrix associated with the linear transformation \mathcal{L} , given by $\mathbf{F} = \mathbb{E} [F^T \otimes F^T]$. \mathbf{F} is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{M^{-1} \otimes M^{-1}}$, and therefore it has real eigenvalues. We can define the function

$$\lambda_{\mathcal{L}}(i) : \{1, \dots, N_{\mathcal{L}}^2\} \rightarrow \mathbb{R}$$

that returns the i -th eigenvalue of \mathbf{F} . We assume that the function is monotonically non increasing, i.e. $\lambda_{\mathcal{L}}(i) \geq \lambda_{\mathcal{L}}(i+1)$ for all i . Let moreover $\Delta_{\mathcal{L}}(i)$ be an eigenvector associated with the eigenvalue $\lambda_{\mathcal{L}}(i)$. By decomposing $\Omega M \Omega$ into $\sum_i \alpha_i \Delta_{\mathcal{L}}(i)$ we can then express the convergence rate R as

$$R = \max \{ |\lambda_{\mathcal{L}}(i)| \mid \alpha_i \neq 0, \Omega \Delta_{\mathcal{L}}(i) \Omega \neq 0 \}. \tag{22}$$

Computing R as defined in (22) is in general not simple. In [9] it has been shown how to compute it numerically and, for some special graph topologies, also analytically.

In the following, we will derive an upper bound for R that can be computed from $\bar{F} = \mathbb{E} [F]$. We first state a few technical lemmas, whose proofs can be found in [9].

Lemma 6: Let $P, Q \in \mathbb{R}^{N \times N}$ and $P \geq Q$. Then $\mathcal{L}^k(P) \geq \mathcal{L}^k(Q)$ for all $k \in \mathbb{Z}_{\geq 0}$.

Lemma 7: The following holds for all Δ :

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

Lemma 8: Let $\bar{F} = \mathbb{E} [F]$. If $\text{Im} [\Omega_1 \cdots \Omega_{\ell}] = \ker \mathbf{1}^T$, then all the eigenvalues of \bar{F} have absolute value not larger than 1, and its only eigenvalue on the unitary circle is $\lambda = 1$, with associated left eigenvector $\mathbf{1}$ and right eigenvector $M^{-1} \mathbf{1}$.

We can then state the following result.

Theorem 9: Consider the linear system (21) and its rate of convergence R . Define

$$\beta = \max \{ |\lambda| \mid \lambda \in \lambda(\bar{F}), \lambda \neq 1 \}$$

where $\bar{F} = \mathbb{E} [F]$. Then $R \leq \beta$.

Proof: Let us first prove that $\Omega \mathcal{L}(\Omega M \Omega) \Omega \leq \beta \Omega M \Omega$. Indeed, we have

$$\begin{aligned} x^T \Omega \mathcal{L}(\Omega M \Omega) \Omega x &= \mathbb{E} [x^T \Omega F^T \Omega M \Omega F \Omega x] \\ &= \mathbb{E} [x^T \Omega F^T M F \Omega x] \\ &= x^T \Omega M^{1/2} \mathbb{E} [M^{1/2} F M^{-1/2}] M^{1/2} \Omega x, \end{aligned}$$

where we use the fact that $\Omega F \Omega = F \Omega$ and that $F_i^T M F_i = M F_i$.

$\mathbb{E} [M^{1/2} F M^{-1/2}] = M^{1/2} \bar{F} M^{-1/2}$ is symmetric and, by Lemma 8, it has only one eigenvalue on the unit circle (precisely in 1), with eigenvector $M^{-1/2} \mathbf{1}$. As $M^{1/2} \Omega x \perp M^{-1/2} \mathbf{1}$ for all x , we have

$$x^T \Omega \mathcal{L}(\Omega M \Omega) \Omega x \leq \beta \Omega M \Omega,$$

with $\beta = \max\{|\lambda| \mid \lambda \in \lambda(\bar{F}), \lambda \neq 1\}$.

From this result, using Lemmas 6 and 7, we can conclude

$$\begin{aligned} \Omega \mathcal{L}^t(\Omega M \Omega) \Omega &= \Omega \mathcal{L}^{t-1}(\mathcal{L}(\Omega M \Omega)) \Omega \\ &= \Omega \mathcal{L}^{t-1}(\Omega \mathcal{L}(\Omega M \Omega) \Omega) \Omega \\ &\leq \Omega \mathcal{L}^{t-1}(\beta \Omega M \Omega) \Omega \\ &= \beta \Omega \mathcal{L}^{t-1}(\Omega M \Omega) \Omega \\ &\leq \dots \leq \beta^t \Omega M \Omega, \end{aligned}$$

and therefore $R \leq \beta$. ■

The tightness of β as a bound for R has been studied in [9], and in the following we therefore consider β as a reliable metric for the evaluation of the algorithm performances.

C. Optimal strategy: nearest-neighbor gossip

Consider the case in which all clusters \mathcal{C}_i are pair of nodes, i.e. $|\mathcal{C}_i| = 2, \forall i$. The following result shows what is the best performance (according to the bound β on the convergence rate R) that the proposed algorithm can achieve.

Theorem 10: Consider the algorithm (16), and assume that the hypergraph \mathcal{H} describing the clusters \mathcal{C}_i is an arbitrary connected, undirected graph defined over the nodes \mathcal{C} . Consider arbitrary triggering probabilities p_i for the clusters. Then the bound β on the convergence rate of the algorithm satisfies

$$\beta \geq 1 - \frac{1}{N_C - 1}.$$

Proof: By Theorem 9 we have $\beta = \lambda_2(\bar{F}) = 1 - \lambda_{N_C-1}(\bar{E})$, where $\bar{E} = \mathbb{E}[(\Omega_i M \Omega_i)^\# M]$. Let h, k be the compensators in \mathcal{C}_i , and let us define $e_{hk} = e_h - e_k$, so that $\Omega_i = e_{hk} e_{hk}^T / 2$. We have that

$$\begin{aligned} \sum_j \lambda_j(\bar{E}) &= \text{Tr} \left(\sum_i p_i (\Omega_i M \Omega_i)^\# M \right) \\ &= \text{Tr} \left(\sum_i p_i \left(\frac{e_{hk} e_{hk}^T}{2} M \frac{e_{hk} e_{hk}^T}{2} \right)^\# M \right) \\ &= \text{Tr} \left(\sum_i p_i \frac{2}{e_{hk}^T M e_{hk}} \Omega_i^\# M \right) \\ &= \sum_i p_i = 1, \end{aligned}$$

where in the last equation we used the facts that $\Omega_i^\# = \Omega_i$ and that $\text{Tr}(ABC) = \text{Tr}(CAB)$.

As $\lambda_{N_C}(\bar{E}) = 0$, then $\lambda_{N_C-1}(\bar{E}) \leq 1/(N_C - 1)$, and therefore $\beta \geq 1 - 1/(N_C - 1)$. ■

We now present a special case in which the optimal convergence rate of Theorem 10 is indeed achieved via a specific choice of the clusters \mathcal{C}_i .

Consider the special case in which the \mathcal{G} is a tree, and assume that $\mathcal{C} = \mathcal{V}$. Notice that the convergence rate of the algorithm depends only on $X_{\mathcal{C}\mathcal{C}}$, and therefore on the electrical distance between the compensators. The loss of generality of choosing $\mathcal{C} = \mathcal{V}$ is then minor: one can always add passive nodes between compensators, without affecting the convergence rate analysis. The choice $\mathcal{C} = \mathcal{V}$, however, simplifies the definition of neighbors among compensators.

Proposition 11: Consider the nearest-neighbor clustering choice, corresponding to the set of clusters $\{\mathcal{C}_e, e \in \mathcal{E}\}$, where $\mathcal{C}_e = \{\sigma(e), \tau(e)\}$. Assume that each set is triggered with the same probability. Then $R \leq \beta$, with

$$\beta = 1 - \frac{1}{N_C - 1}.$$

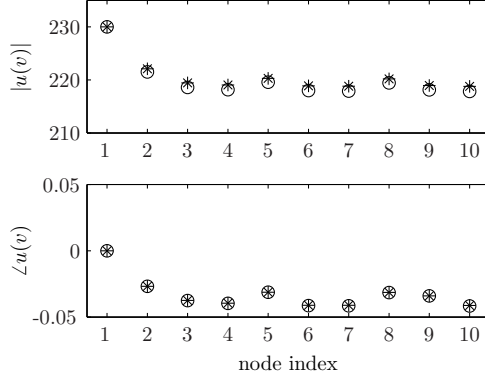


Fig. 2. Comparison between the network state (node voltages) computed via the exact model induced by (1) and (2) (circles), and the approximate model induced by (7) and (8) (stars), in the proposed testbed.

Proof: As before, we have $\beta = 1 - \lambda_{N-1}(\bar{E})$, where $\bar{E} = \mathbb{E} [(\Omega_e M \Omega_e)^\# M] = \mathbb{E} [(\Omega_e X \Omega_e)^\# X]$. Using the fact that

$$\Omega_e = \frac{1}{2}(e_{\sigma(e)} - e_{\tau(e)})(e_{\sigma(e)} - e_{\tau(e)})^T,$$

and the fact that

$$(e_{\sigma(e)} - e_{\tau(e)})^T X (e_{\sigma(e)} - e_{\tau(e)}) = e^{-j\theta} z(e) = d(e),$$

we have that $(\Omega_e X \Omega_e)^\# = \frac{2}{d(e)} \Omega_e$. Therefore, as the probability of triggering a set \mathcal{C}_e is $1/N_{\mathcal{E}} = 1/(N_{\mathcal{C}} - 1)$,

$$\begin{aligned} \mathbb{E} [(\Omega_e X \Omega_e)^\# X] &= \\ &= \frac{1}{N_{\mathcal{C}} - 1} \sum_{e \in \mathcal{E}} \frac{1}{d(e)} (e_{\sigma(e)} - e_{\tau(e)})(e_{\sigma(e)} - e_{\tau(e)})^T X \\ &= \frac{1}{N_{\mathcal{C}} - 1} A^T D^{-1} A X = \frac{1}{N_{\mathcal{C}} - 1} \Omega, \end{aligned}$$

and $\beta = 1 - \lambda_{N-1}(\Omega/(N_{\mathcal{C}} - 1)) = 1 - \frac{1}{N_{\mathcal{C}} - 1}$. ■

VI. SIMULATIONS

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

We consider the microgrid sketched in Figure 1. We assume that the nominal voltage at the PCC \bar{v} is 230 V, that the nominal operating frequency is $f_0 = 50$ Hz, and that the lines' characteristic resistance and inductance are 0.16 m Ω /m and 1 μ H/m, respectively. The length of the lines are uniformly distributed between 50 m and 200 m, while the injected powers has been set as $s(v) = |s(v)|e^{j\phi(v)}$, with $s(v)$ uniformly distributed between -10 kW and 0 kW, and $\cos \phi$ uniformly distributed between 0.7 and 1.

Given these parameter, we first estimate the quality of the linear approximated model proposed in Section III. As shows in Figure 2, the approximation error results to be negligible, even in the case in which voltage drops get close to the maximum that is generally allowed in power distribution networks (with smaller voltage drops, the approximation becomes even more accurate).

On the same testbed we then have validated the quality of the estimate (13) for the gradient of the cost function that we want to minimize. We considered again the microgrid in Figure 1, assuming that only 3 nodes can be commanded to inject the desired amount of reactive power: $\mathcal{C} = \{\bar{v}, i, j\}$. Because of the constraint $\mathbf{1}^T q = 0$, we are left with only 2 degrees of freedom, so we choose $q(i)$ and $q(j)$ as decision variables and we let $q(\bar{v})$ to satisfy the constraint.

In Figure 3 we reported, in thick line, the contour plot of the power distribution losses in the microgrid, computed according to the exact nonlinear model. As a thin line, we overlaid the contour plot of the (numerically computed) function whose gradient corresponds to the gradient estimate $\widehat{J}(q_{\mathcal{C}})$.

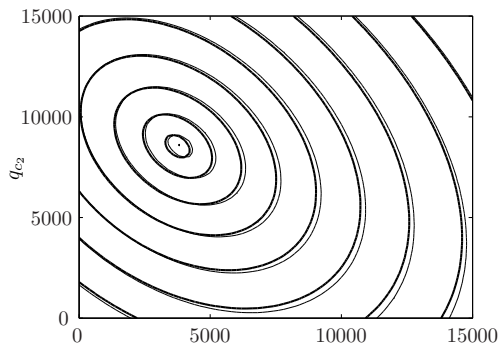


Fig. 3. Contour plot of the exact distribution losses (thick line) and of the cost function whose gradient is given by the voltage measures, according to (13) (thin line).

The optimum obtained by zeroing the estimated gradient practically corresponds to the solution of the exact optimization problem.

We then simulated the behavior of the algorithm proposed in Section V, and the performance of different clustering choices. We considered a larger network of 30 nodes, 11 of which are compensators (see Figure 4). We chose line impedances and loads similarly to before, and we considered the two following clustering choices:

- *nearest neighbor gossip*: based on the result stated in Proposition 11, we enabled pairwise communication between nodes whose distance in the electric grid is lower than a given threshold; notice however that the hypotheses of Proposition 11 are not precisely verified, as the graph is not a tree;
- *star topology*: clusters are in the form $\mathcal{C}_i = \{\bar{v}, v\}$ for all $v \in \mathcal{C}$. The reason of this choice is that, as \bar{v} is the PCC, the constraint $\mathbf{1}^T q = 0$ is inherently satisfied: whatever variation in the injected reactive power is applied by v , it is automatically compensated by a variation in the demand from the transmission grid via the PCC.

The result of the simulation are in Figure 4, together with the best achievable performance as given in Theorem 10. One can notice different things. First, the performances of the nearest-neighbor gossip algorithm result indeed to be better than the star topology, as Proposition 11 suggests. Second, one can see that, while the performance metric that we adopted is meaningful for its asymptotic behavior, it does not describe the initial stage, when the full dynamics of the system contribute to the algorithm behavior; we conjecture, however, that in the case in which the system is working close to the optimum, possibly tracking the optimal point (which changes according to variations in the demands), the analysis of the slowest dynamics of the system should be the most meaningful choice.

VII. 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.

We 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 algorithms (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 (consider for example the possibility of communicating over the power lines).

REFERENCES

- [1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.
- [2] 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. 432–438, 2007.

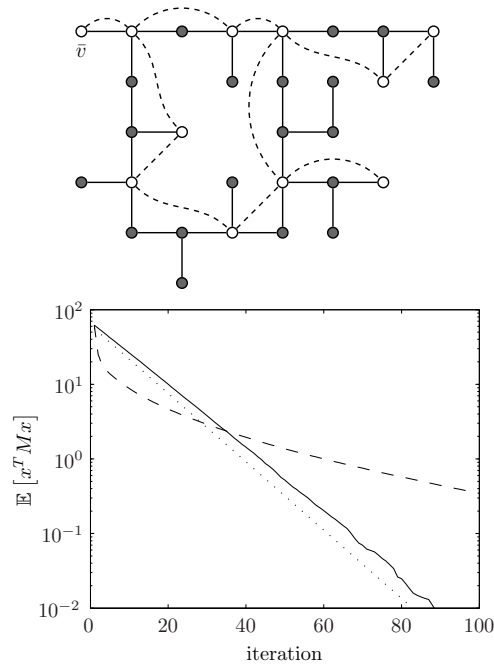


Fig. 4. Simulation of the behavior of the algorithm, when applied to the network above (where compensators are in white, loads in gray). The hypergraph \mathcal{H} is represented by a dashed line. The algorithm behavior (averaged over 1000 realizations) has been plotted for two different clustering choices: *nearest-neighbor gossip* (solid line) and *star topology* (dashed). The dotted line represent the best possible performance.

- [3] E. Tedeschi, P. Tenti, and P. Mattavelli, "Synergistic control and cooperative operation of distributed harmonic and reactive compensators," in *Proceedings of the IEEE Power Electronics Specialists Conference (PESC 2008)*, 2008.
- [4] A. Rabiee, H. A. Shayanfar, and N. Amjadi, "Reactive power pricing: Problems and a proposal for a competitive market," *IEEE Power Energy Magazine*, vol. 7, no. 1, pp. 18–32, Jan. 2009.
- [5] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, *Electric energy systems. Analysis and operation*. CRC Press, 2009.
- [6] A. Ghosh, S. Boyd, and A. Saberi, "Minimizing effective resistance of a graph," *SIAM Review*, vol. 50, no. 1, pp. 37–66, Feb. 2008.
- [7] S. Bolognani and S. Zampieri, "Distributed quasi-Newton method and its application to the optimal reactive power flow problem," in *Proc. NECSYS'10*, Annecy, France, Sep. 2010.
- [8] M. Ciobotaru, R. Teodorescu, P. Rodriguez, A. Timbus, and F. Blaabjerg, "Online grid impedance estimation for single-phase grid-connected systems using PQ variations," in *Proc. IEEE PESC 2007*, Jun. 2007.
- [9] S. Bolognani and S. Zampieri, "A gossip-like distributed optimization algorithm for reactive power flow control," in *Proc. IFAC World Congress 2011*, Milano, Italy, Aug. 2011.