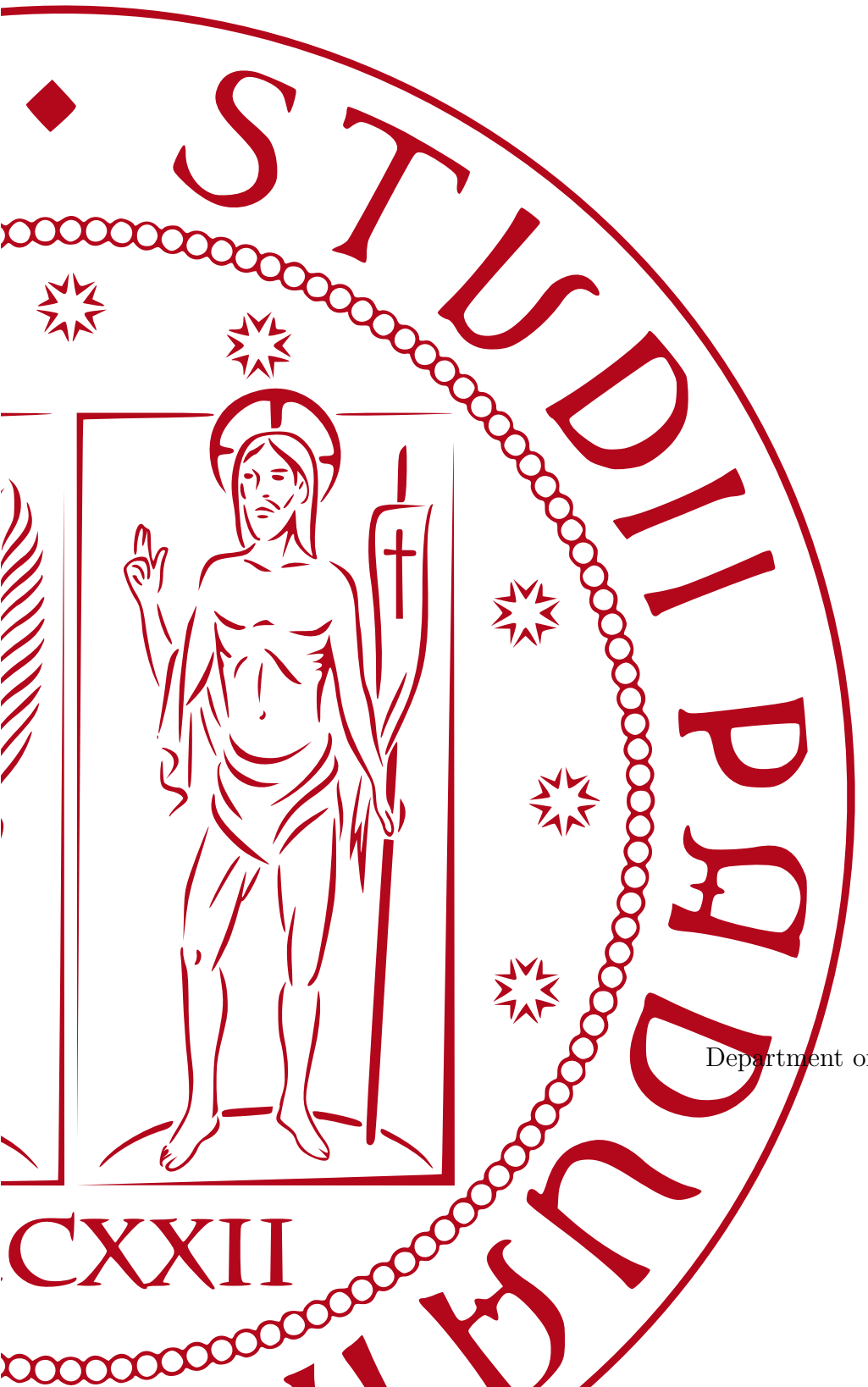
Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Modeling, Control and Identification of a Smart Grid



Ph.D. candidate
Guido Cavraro

Advisor
prof. Sandro Zampieri

Co-advisor
prof. Ruggero Carli

Director
prof. Matteo Bertocco

Coordinator
prof. Carlo Ferrari

Ph.D. School in
Information Engineering

Department of Information Engineering
University of Padova
2015

Get them to face danger, but do not reveal the advantages.
Throw them into danger and they will survive;
put them on deadly ground and they will live.
Only if the troops are in situations of danger will they turn defeat into victory.

Sun Tzu Ping Fa

Abstract

We are in front of an epochal change in the power distribution and generation scenario. The increasing request of energy, the energy dependency of several countries from few foreign nations endowed with oilfield or gas field, and, on the other hand, the climate change and environmental issues are the main explanation of the recent development and spread of renewable distributed energy generation technologies. Examples of them are photovoltaic panels, wind turbines or geothermal, biomass, or hydroelectric. They are called small-size generators, or micro-generator, since the amount of power they can produce is significantly lower than the one produced by the huge, classical power plants. These distributed energy resources (DERs) are located close to where electricity is used, in the distribution network. Furthermore, they are connected to the electrical grid via electronic interfaces, the inverters, that could allow us to control the power injected into the grid.

This thesis is focused on the study of some crucial aspects of this new energetic scenario:

1. **Modeling**: we recall the classical models and a recent linearized one of the power systems, that will be very useful for the design and the analysis of our algorithms.

2. **Optimal Reactive Power Flow (OPRF) problem**: in this part we recall classical and recent algorithms that deal with the reactive power regulation. In particular, we focus on the ones that solve the OPRF problem, i.e. the problem of the amount of reactive power to be injected by each micro-generators, in order to achieve "optimal" performance. We choose, as an optimality achievement, the minimization of the line losses. Finally we derive and propose our OPRF algorithm, providing formal proves of its convergence to the optimal solution.

3. **Optimal Power Flow (OPF) problem**: the OPF problem's aim is to find an operating point of the power system that optimize a cost function (tipically the generation cost) satisfying the power demand and some operative constraints. After

recalling the most popular algorithms that solve the OPF problem, we propose two of them. In this framework there are mainly two possible scenarios.

The first is related to the "utility point of view", where the total cost accounts for the production cost of the energy (that comes from big generation plants such as nuclear or hydro-electrical plants) and for the remuneration to be paid to the owners of DERs. In this framework, the utility imposes a behavior procedure to be followed by the producers to compute the amount of energy they have to inject into the grid to minimize the total cost. The first algorithm deal with this scenario.

The second one is related to the "producer point of view". Since the owners of the DERs are paid proportionally to the energy that they inject, they would like to maximize the power they inject, while keeping satisfied some operative constraints. The result is a game among the agents. A first treatment on this scenario is given by the second algorithm.

4. **Switches monitoring for topology identification**: in this part, after a literature review, we propose a algorithm for the identification of switches actions. They modify the topology of the electrical grid, whose knowledge is fundamental for monitoring, control and estimation. This algorithm works analyzing how the phasorial voltage profile vary and recognize a kind of signature left by the switches status change.

Sommario

Stiamo vivendo un cambiamento epocale dello scenario di produzione e distribuzione dell'energia. L'incremento della richiesta di energia, il fatto che molte Paesi dipendano energeticamente da poche nazioni ricche di giacimenti di gas o petroliferi e, inoltre, il cambiamento climatico e l'inquinamento costituiscono la ragione principale del recente sviluppo e diffusione di tecnologie per la generazione di energia da fonte rinnovabili. Alcuni esempi ne sono i pannelli fotovoltaici oppure generatori eolici, geotermici, idroelettrico o dalle biomasse. Essi sono generatori di piccole dimensione, o micro-generatori, visto che le loro dimensioni e la quantità di energia che producono sono decisamente inferiori a quelle dei grandi, classici impianti di generazioni. Queste fonti distribuite di energia (DERs) si trovano vicino agli utilizzatori, nella rete di distribuzione. Inoltre, essi sono collegati alla rete attraverso interfacce elettroniche, gli inverter, che ci potrebbero permettere di controllare la quantità di potenza che essi iniettano.

Questa tesi si concentra sullo studio di alcuni aspetti cruciali di di questo nuovo scenario energtico, e è composta da quattro parti principali, ciascuna delle quali tratta un aspetto diverso. Esse sono:

1. **Modellistica**: qui si richiamano i modelli classici e un recente modello linearizzato, che sarà utile per la progettazione e l'analisi degli algoritmi proposti, dei sistemi di potenza.

2. **Ottimizzazione dei flussi di potenza reattiva**: in questa parte si richiamano i classici e i più recenti algoritmi di gestione della potenza reattiva. In particolare ci si concentra su quelli che ne ottimizzano i flussi, cioè che si focalizzano sul problema di decidere quanta potenza reattiva ciascun micro-generatore deve iniettare se si vogliono ottenere delle prestazioni "ottime". Come indice di ottimalità è stata scelta la miniizzazione delle perdite sulle linee. Infine viene progettato e analizzato il nostro algoritmo di ottimizzazione, fornendo dimostrazione formale della sua convergenza.

3. **Ottimizzazione dei flussi di potenza**: lo scopo di questo problema è quello di

trovare una configurazione che ottimizza una funzione costo (di solito il costo di generazione) e che soddisfa alcuni vincoli operativi. Dopo aver richiamato i più famosi algoritmi che risolvono questo problema, ne vengono proposti due. Questo perchè vi sono principalmente due scenari.

Il primo è connesso al punto di "vista dell'utility", dove il costo tiene conto sia dell'effettivo costo di generazione dell'energia (che arriva dai grandi impianti di generazione, quali centrali nucleari o idroelettriche) e della remunerazione che deve essere data ai proprietari delle DERs per l'energia che producono. In questo caso , l'utility impone una procedura per calcolare la potenza da iniettare per minimizzare il costo totale. Il primo algoritmo rientra in questo scenario.

Il secondo è connesso al punto di "vista del proprietario di DERs". Poichè questi viene pagato proporzionalmente alla quantità di energia che inietta, vorrebbe massimizzare la potenza che inietta, soddisfando comunque alcuni vincoli operativi. Ne viene fuori un conflitto fra i diversi proprietari. Una trattazione di questo scenario è data dal secondo algoritmo.

4. **Controllo degli interruttori per l'identificazione della topologia**: in questa parte, dopo una revisione della letteratura, viene proposto un algoritmo per l'identificazione delle azioni degli interruttori nella rete. Queste modificano la topologia della rete elettrica, la cui conoscenza è fondamentale per il controllo, la supervisione e la stima. Questo algoritmo analizza le variazioni dei profili delle tensioni fasoriali e cerca di riconoscere in esse una sorta di firma della particolare azione degli interruttori.

Contents

Part I

Modeling

1 Introduction

1.1 Classical transmission and distribution grid

The actual transmission and distribution infrastructure (Harris Williams (Summer 2014)) has been the backbone of the electric power system, as it delivers the electricity from power plants to end customers. In this framework, energy is produced in the power plants and then its voltage level is increased by transformers to reduce energy loss during transportation along transmission lines. The former elements constitute the *transmission network*. Electricity is carried by transmission lines between regions to substations, where the voltage level is stepped down so that it can be distributed to end users, including residential, commercial, and industrial customers. Finally, distribution transformers decrease the voltage in order to safely distribute electricity to end users. Distribution lines then deliver electricity through overhead or underground power lines, and metering systems measure and record the locations and amounts of power transmitted. The latter

elements constitute the *distribution network*. Considering Figure 1.1, we have

1. **power plants** where the electricity is generated;

2. **high voltage transformers** that increase voltage from thousands to hundreds of thousands of volts to send power over long distances;

3. **transmission lines** at high voltage that carry energy from the power plant to substations. These lines are tipically inductive;

4. **transmission substations** that connect two or more transmission lines and are endowed with high-voltage switches that allow lines to be connected or isolated for maintanence;

5. **distribution substations**, that are basically transformers that reduce voltage to a lower level so power can be delivered on distribution lines to the surrounding community;

6. **distribution systems** that include that deliver electricity through overhead or underground wires to the users;

7. **service connections**, that provide are lines connecting meter at end-user location.

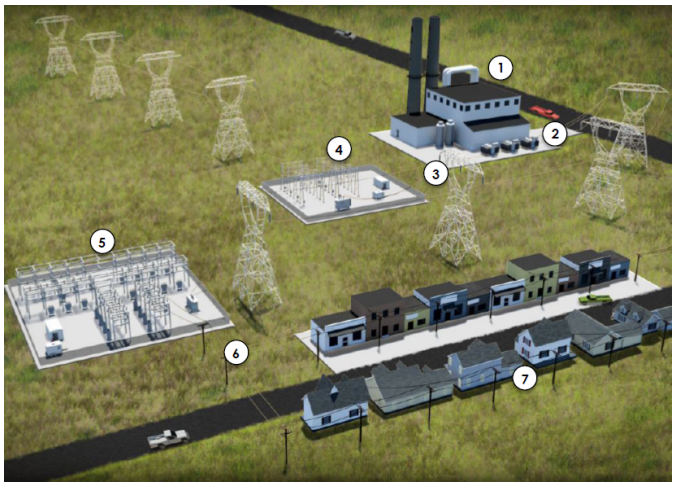


Figure 1.1: Transmission and distribution grid scheme

1.2 New energetic scenario and smart grid

The advent of *distributed energy resources* (DERs) is deeply changing the power distribution scenario (Jiayi, Chuanwen, and Rong (2008)). The need of reducing emissions in the electricity generation field, recent technological developments in the microgeneration domain, and electricity business restructuring are the main factors responsible for the growing interest in the use of microgeneration. In fact, the connection of small-size generators with power ratings less than a few tens of kilowatts to low voltage distribution networks could lead to the increase of the reliability to final consumers and could bring additional benefits for global system operation and planning, namely, regarding investment reduction for future grid reinforcement and expansion. Their use can lead to a number of benefits for the electrical distribution system, such as reduction of line losses, voltage profile improvements, and decrease of emissions of pollutants Chiradeja and Ramakumar (2004). However, application of individual distributed generators can cause as many problems as it may solve, e.g. leading to big voltage fluctuation or overheating of lines and transformers, basically due to the intermittance of the renewable resources, as wind or sun.

The changes that are happening are particularly significant for the electricity distribution grid, that is going to be transformed into a *smart grid* (Rahimi and Ipakchi (2010)). A feature of the smart grid is the increased and bidirectional interaction between transmission and distribution operations. The expected profusion of demand response, renewable resources, and distributed generation at distribution/retail level has direct implications on the operation of the transmission system and the wholesale energy markets. Enabling technologies, such as enhancements in communication and information technologies, make it possible to turn the potentially adverse autonomous operational impacts of these new resources into useful controllable products for wholesale market and transmission system operators.

In the traditional utility environment, power flow is almost unidirectional from centralized supply sources (big power plants) to the users, and information flow is from lower voltages to higher operational centers. In contrast, in the emerging utility environment, both power and information flows are bidirectional. These concepts are shown in Figure 1.2 and in Figure 1.3, both taken from Rahimi and Ipakchi (2010)

In this environment, the natural evolution of the distribution network is the *micro grid* (Lasseter (2002)). It is a portion of low-voltage network which can be considered autonomous from the rest of the distribution system, in the sense that it can sustain with the power produced within it the users demand. A microgrid is a localized grouping of electricity generation, energy storage, and loads that normally operates connected to a

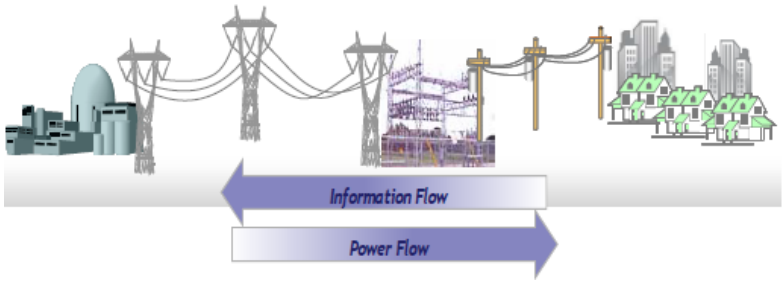


Figure 1.2: Transmission and distribution grid scheme

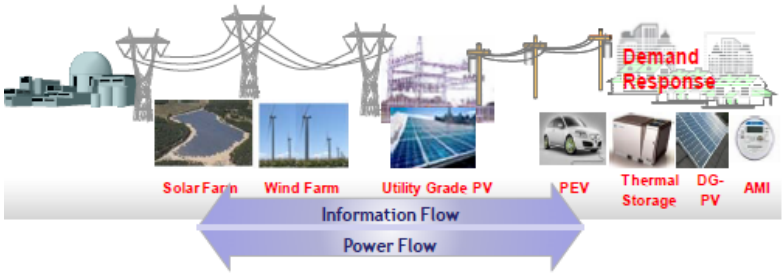


Figure 1.3: Transmission and distribution grid scheme

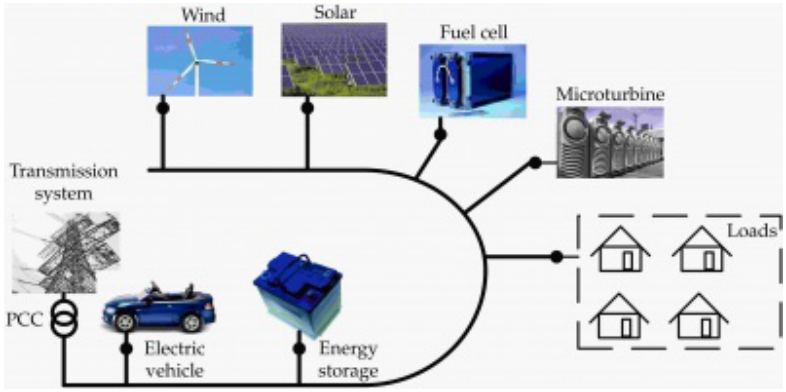


Figure 1.4: Microgrid scheme

traditional distribution or transmission grid in a single point, called *point of common coupling* (PCC). The PCC can be disconnected, for example in case in which the main grid has a fault, and the microgrid can then function autonomously. Generation and loads in a microgrid are usually interconnected at low voltage. Microgrid generation resources can include fuel cells, wind, solar, or other energy sources. The multiple dispersed generation sources and ability to isolate the microgrid from a larger network would make it possible to achieve better quality of the service, higher efficiency and lower costs since they exploits renewable sources. In Figure 1.4 we have a scheme of a microgrid, where we can see different DERs, some storage devices that store the power produced in excess and electric vehicles.

1.3 Thesis topics and outline

The thesis focuses on the control, optimization and monitoring of the distribution networks, possibly for microgrids. Novel algorithms that solve these problems are proposed. They all exploit a linearized grid model proposed in Bolognani and Zampieri (2013). Their main feature is that they operate in a scenario in which only a few buses (the smart Agents) can sense the grid, while the others remain unmonitored. However their goal can be achieved because from *local* (phasorial voltage) measurement, the agents can infer the *global* knowledge they need. In some sense, the physical layer works as a *analog computer* for the smart agents. Their action in fact spread through the electric lines and thus can be perceived indirectly by other agents, even if there is no direct communication.

Chapter 2 provides the model we used for the distribution network, recall the original linearized model of Bolognani and Zampieri (2013) and give other in-dept analysis developed to provide useful tools for the algorithms and a better comprehension of the grid characteristics. Chapter 4 shows a novel algorithm for the optimization of the reactive power flows in the distribution network. It exploits a *feedback* approach, in which there is an alternation of sensing, communication and actuation. Chapter 5 shows a novel algorithm for the optimization of the active power flows in the distribution network. It follows the same philosophy of the algorithm in Chapter 4. Chapter 6 instead provides a different algorithm for the active power managment. It position itself in an environment in which the DERs are in competition because they would like to maximize the power they inject, and thus the remuneration paid from the utility for the service. It turns out to be a game in which the algorithm determines a *fair* solution. Finally Chapter 7 provide a novel algorithm for the switches status monitoring, applied to the topology detection with only few sensors.

1.4 Mathematical preliminaries and notation

Given a matrix $W \in \mathbb{C}^{m\times n}$, we denote its element-wise complex conjugate by $\overline{W}$, its transpose by W^T and its conjugate transpose by W^*. We denote the matrix of the absolute values, of the real and of the imaginary parts of W by $|W|$, $\mathrm{Re}(W)$ and by $\mathrm{Im}(W)$. We denote the entry of W that belongs to the j-th row and to the k-th column by W_{jk}, and the trace of W as $\mathrm{Tr}(W)$.

Given a vector v, v_j will denote its j-th entry, while v_{-j} the subvector of v, in which the j-th entry has been eliminated. Sometimes, the subscript of a vector will not represent the index of a entry, but it will be a label, a part of "the name" of the vector. However it will be easy to distinguish the two cases. Given two vectors v and w, we denote by $\langle v, w\rangle$ their inner product v^*w, and by $v \odot w$ their Hadamard product, defined by $(v \odot w)_j = v_j\, w_j$. We define the column vector of all ones by $\mathbf{1}$, while the canonical base vector e_j is the vector whose elements are all zero exept for its j-th entry, which is set to 1. We will write $v \leq w$ if v is a vector whose entries are component-wise greater than the entries of w.

We define the *directed graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \sigma, \tau)$, where $\mathcal{V}$ is the set of nodes with cardinality $|\mathcal{V}|$, $\mathcal{E}$ is the set of edges with cardinality $|\mathcal{E}|$, and $\sigma, \tau : \mathcal{E} \to \mathcal{V}$ are two functions such that edge $e \in \mathcal{E}$ goes from the source node $\sigma(e)$ to the terminal node $\tau(e)$.

Given two nodes $h, k \in \mathcal{V}$, we define the path $\mathcal{P}_{hk}$ as the sequence of adjacent nodes, without repetitions, that connect node h to node k.

Let $A \in \{0, \pm 1\}^{|\mathcal{E}|\times|\mathcal{V}|}$ 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 (i.e. for every pair of nodes there is a path connecting them), then $\mathbf{1}$ is the only vector in the null space $\ker A$. We will denote the incidence matrix also as

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_{|\mathcal{E}|} \end{bmatrix}$$

where a_j is the j-th row of A.

2

Model of a smart grid

In the first part of this chapter we recall some basics of electrical circuits regarding the alternating current that will be used in the following. Most of the concepts here introduced can be found, with a deeper treatise, in Von Meier (2006), Guarnieri and Malesani (1998) and Andersson (2004). In the second part instead we provide the smart grid model used.

2.1 Basics of AC ciruits

Phasorial representation

Consider the generic sinusoidal waveform,

$$A(t) = A_M \sin(\omega t + \varphi) \tag{2.1}$$

where A_M is its maximum aplitude, with the same physical dimension of the waveform $a(t)$, ω is the pulsation measured in [rad/s], and φ is the initial phase, expressed in [rad], belonging to the set of isofrequential sinusoids with frequency $f = \frac{\omega}{2\pi}$ measured in [Hz]. The waveform (2.1) is determined by A_M and φ. These two parameters define one and only one complex numbers whose magnitude is $\frac{A_M}{\sqrt{2}}$ and whose phase is φ

$$a = \frac{A_M}{\sqrt{2}} e^{i\varphi} \tag{2.2}$$

Vice versa, the complex number A identifies one and only one waveform with pulsation ω, given by (2.1), i.e. there is a biunivocal relation between any complex number and the elements of the isofrequenzial sinusoids. The complex number (2.2) is called *phasor*. Notice that the phasor magnitude is equal to the *root mean square* of $a(t)$, i.e.

$$\frac{A_M}{\sqrt{2}} = \sqrt{\frac{1}{T}\int_{t_0}^{t_0+T} a^2(t)dt} \tag{2.3}$$

where T is the period of $a(t)$.

When the grid is working at an ideal sinusoidal regime, both the currents and the voltages are sinusoidal signals, waving at 50 Hz, in Europe, or 60 Hz, in the U.S.A.. Thus, they can be expressed as

$$I(t) = I_M \sin(\omega t + \varphi_I) \tag{2.4}$$

$$U(t) = U_M \sin(\omega t + \varphi_U) \tag{2.5}$$

and they are associated with the phasors $i = \frac{I_M}{\sqrt{2}} e^{i\varphi_I}$ and $u = \frac{U_M}{\sqrt{2}} e^{i\varphi_U}$ respectively. Currents and voltages are measured in amperes (A) and in volts (V), respectively.

Impedance and admittance

We define *impedance* of a electrical element the ratio between the currents passing through it, expressed by the phasor i, and the voltage drop between its ends, expressed by the phasor u

$$\mathrm{z} = \frac{i}{u} \tag{2.6}$$

measured in ohm (Ω). In general the impedance is a complex numbers, whose real part $\mathrm{r} = \mathrm{Re(z)}$ is called *resistance* and whose imaginary part $\mathrm{x} = \mathrm{Im(z)}$ is called *reactance*, both of them measured in ohm [Ω], i.e.

$$\mathrm{z} = \mathrm{r} + \mathrm{ix}.$$

Certain materials and electronic devices exhibit a nonlinear relationship between current and voltage, that is, z is not constant but varies for different values of u and i. A device instead for which the relation between voltage and current is linear, i.e. z is always the same, is called *constant impedance device*, e.g. wires of the overhead lines. The inverse of the impedance is called *admittance*

$$\mathrm{y} = \frac{i}{u}$$

measured in siemens (S). In general the impedance is a complex numbers, whose real part g = Re(y) is called *conductance* and whose imaginary part b = Im(y) is called *susceptance*, both of them measured in [S], i.e.

$$\mathrm{y} = \mathrm{g} + \mathrm{ib}$$

AC Power

Power is a measure of energy per unit of time, and thus it gives the rate of energy consumption or production. We define the *instantaneous power* as

$$\begin{aligned} P(t) &= V(t)I(t) && (2.7)\\ &= I_M \sin(\omega t + \varphi_I) V_M \sin(\omega t + \varphi_U) && (2.8)\\ &= \frac{I_M V_M}{2}\cos(\phi) - \frac{I_M V_M}{2}\cos(2\omega t + \varphi_U + \varphi_I) && (2.9) \end{aligned}$$

where $\phi = \varphi_U - \varphi_I$. However instantaneous power is not tipically very interesting, because we are interested about power transmitted or consumed in a time scale much greater than 1/60 of a second, and thus we need an expression for power as averaged over entire cycles of alternating current and voltage. The average value in a period of the instantaneous power is called *active power*

$$\begin{aligned} p &= \frac{1}{T}\int_{t_0}^{t_0+T} P(t)dt \\ &= \frac{I_M V_M}{2}\cos(\phi) && (2.10) \end{aligned}$$

and it corresponds to the power actually transmitted, consumed or generated by the loads or the generators. The active power is measured in watts (W).

There are other aspects of the transmitted power that we wish to specify. We denote as *apparent power*

$$s = u\bar{i} \qquad (2.11)$$

measured in volt-amperes (VA), and whose magnitude is given by $|s| = \frac{I_M U_M}{2}$. Apparent power is important in the context of equipment capacity: the crucial quantity w.r.t. devices thermal capacity limits is the current. Since usually the operating voltage of a given piece of equipment is quite constant, apparent power is a fair way to indicate the current. For this reason, utility equipment ratings are tipically given in VA.

Finally, we define the *reactive power*

$$q = \frac{I_M V_M}{2} \sin(\phi) \tag{2.12}$$

The reactive power is the component of power that oscillates back and forth through the lines or the devices, being exchanged between electric and magnetic fields and not getting dissipated.

It can be shown that the active and the reactive power are the real and the imaginary part of the complex power, respectively, i.e. $p = \mathrm{Re}(s), q = \mathrm{Im}(s)$, as we can see in Figure 2.1. By exploiting (2.6) and (2.11), the power absorbed by a passive element of

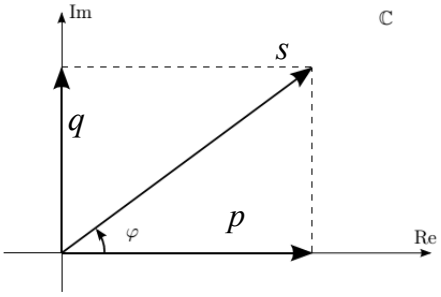


Figure 2.1: Representation in the complex plane of the apparent, active and reactive power.

impedance z is given by

$$s = \mathrm{z}|i|^2 \tag{2.13}$$

whose real part $\mathrm{r}|i|^2$ is the active power absorbed, while the imaginary part $\mathrm{x}|i|^2$ is the reactive power absorbed.

2.2 Smart power distribution grid as cyber-physical system

Computing and communication capabilities will soon be embedded in all types of objects and structures in the physical environment. Such systems that bridge the cyber-world of computing and communication with the physical world are referred as *cyber-physical*

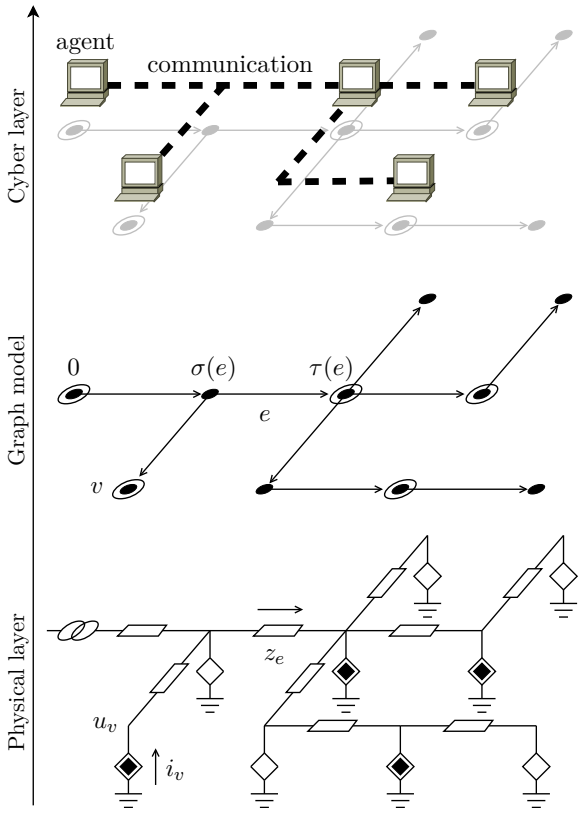


Figure 2.2: Schematic representation of the power distribution grid model. In the lower panel the *physical layer* is represented via a circuit representation, where black diamonds are microgenerators, white diamonds are loads, and the left-most element of the circuit represents the PCC. The middle panel illustrates the adopted *graph representation* for the same grid. Circled nodes represent both microgenerators and the PCC. The upper panel represents the *cyber layer*, where agents (i.e. microgenerator nodes and the PCC) are also connected via some communication infrastructure.

systems (Rajkumar, Lee, Sha, and Stankovic (2010)). They are physical and engineered systems whose operations are controlled, monitored, coordinated snd integrsted by a computing and communicating core. In this work, we envision a *smart distribution network* as one of this cyber-physical systems, where

- the **cyber layer** consists of intelligent agents, deployed in the grid, provided with actuation, sensing, communication, and computational capabilities.

- the **physical layer** consists of the power distribution infrastructure, including power lines, loads, microgenerators, and the distribution substations;

2.3 Electrical grid modeling

We model the physical layer (see Figure 2.2, lower panel), i.e. the electrical infrastructure of a smart power distribution network, as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \sigma, \tau)$, in which edges represent the power lines, and nodes represent buses (see Figure 2.2, middle panel). Thus we associate with the electric grid the sets

- $\mathcal{V}$ is the set of nodes (the buses), with cardinality n;
- $\mathcal{E}$ is the set of edges (the electrical lines connecting them), with cardinality $|\mathcal{E}|$;

Buses correspond to loads, microgenerators, and also the point of connection to the transmission grid. It can be a distribution substation or a point of common coupling (if we deal with microgrid), both of them called as PCC in the following, for the sake of simplicity.

The system is described by the following quantities:

- $u \in \mathbb{C}^n$, where u_v is the grid voltage at node v;
- $i \in \mathbb{C}^n$, where i_v is the current injected at node v;
- $\xi \in \mathbb{C}^{|\mathcal{E}|}$, where ξ_e is the current flowing on edge e.
- $s = p + \mathrm{i}q \in \mathbb{C}^n$, where s_v, p_v and q_v are the complex, the active and the reactive power injected at node v. If $p_v > 0$ ($q_v > 0$) we say that node v is injecting active (reactive) power, if $p_v < 0$ ($q_v < 0$) we say that node v is absorbing active (reactive) power;

In order to underline the difference among smart agents and passive loads, we introduce the following block decomposition of the vector of voltages u

$$u = \begin{bmatrix} u_1 \\ u_G \\ u_L \end{bmatrix}, \tag{2.14}$$

where u_1 is the voltage at the PCC, $u_G \in \mathbb{C}^{m-1}$ are the voltages at the microgenerators, and $u_L \in \mathbb{C}^{n-m}$ are the voltages at the loads. Similarly, we also define $s_G = p_G + jq_G$ and $s_L = p_L + jq_L$. The PCC and the microgenerators are collected in the smart agents set $\mathcal{C}$.

Let $A \in \{0, \pm 1\}^{|\mathcal{E}| \times n}$ be the incidence matrix of the graph $\mathcal{G}$ associated with the electric grid

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_{|\mathcal{E}|} \end{bmatrix}$$

We limit our study to the steady state behavior of the system, when all voltages and currents are sinusoidal signals waving at the same frequency ω_0. Thus, they can be expressed via a complex number whose magnitude corresponds to the signal root-mean-square value, and whose phase corresponds to the phase of the signal with respect to an arbitrary global reference.

For every edge e of the graph, we define by z_e and y_e the impedance of the corresponding power line. We model the grid power lines as series impedances, *neglecting* their shunt admittances. The standard Ohm law and Kirchoff currents law can be written, respectively, as

$$Au + Z\xi = 0, \tag{2.15}$$

$$A^T\xi + i = 0, \tag{2.16}$$

with $Z{=}\mathrm{diag}(\mathrm{z}_e, e{\in}\mathcal{E})$. From (2.16) and (2.15) we can also obtain the relation

$$i = \boldsymbol{Y}u \tag{2.17}$$

where $\boldsymbol{Y}$ is the bus admittance matrix of the grid, $\boldsymbol{Y} = A^T Z^{-1} A$.

Let h be a bus of the network. We define $\mathcal{E}_h$ as the subset of $\mathcal{E}$ containing every edge e such that $\sigma(e) = h$ or $\tau(e) = h$. The bus admittance matrix can be also defined as

$$\boldsymbol{Y}_{jk} = \begin{cases} \sum_{e \in \mathcal{E}_h} \mathrm{y}_e, & \text{if } j = k \\ -\mathrm{y}_e, & \text{if } e \text{ is connecting } j \text{ to } k \end{cases} \tag{2.18}$$

From (2.18) it can be seen that $\boldsymbol{Y}$ is symmetric, satisfies

$$\boldsymbol{Y}\mathbf{1} = 0, \tag{2.19}$$

i.e. $\mathbf{1}$ belongs to the Kernel of $\boldsymbol{Y}$, and it can be shown that if $\mathcal{G}$ is a connected graph, then the kernel of $\boldsymbol{Y}$ has dimension 1. Futhermore $\boldsymbol{Y}$'s sparsity pattern reflects the relation of proximity among the grid nodes.

Equation (2.17) allow us to write currents in terms of voltages. The converse is also possible, after defining the matrix $\boldsymbol{X}$.

Lemma 2.3.1 (Lemma 1 in Bolognani and Zampieri (2013))**.** *Let $\boldsymbol{Y}$ be the weighted Laplacian of $\mathcal{G}$. There exists a unique symmetric, positive semidefinite matrix $\boldsymbol{X} \in \mathbb{C}^{n\times n}$ such that*

$$\begin{cases} \boldsymbol{X}\boldsymbol{Y} = I - \mathbf{1}e_1^T \\ \boldsymbol{X}e_1 = 0. \end{cases} \tag{2.20}$$

Exploiting the previous lemma we can write

$$u = \boldsymbol{X}i + \mathbf{1}u_1 \tag{2.21}$$

where we recall that u_1 is the PCC voltage.

The matrix $\boldsymbol{X}$ depends only on the topology of the grid power lines and on their impedance, and it has some notable properties, including the fact that

$$|X_{hh}| \geq |X_{hk}| \quad h,k \in \mathcal{V}, \tag{2.22}$$

and the fact that

$$(e_h - e_k)^T X (e_h - e_k) = Z_{hk}^{\text{eff}}, \quad h,k \in \mathcal{V}, \tag{2.23}$$

where Z_{hk}^{eff} represents the *effective impedance* of the power lines between node h and k. If the grid is radial (i.e. $\mathcal{G}$ is a tree) then Z_{hk}^{eff} is simply the impedance of the only path from node h to node k, and $\boldsymbol{X}$ satisfy the following property.

Lemma 2.3.2. *Let X be the Green matrix associated with a weighted tree $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \sigma, \tau)$, and $e, e' \in \mathcal{E}$. If $e \neq e'$ then*

$$(e_{\sigma(e)} - e_{\tau(e)})^T X (e_{\sigma(e')} - e_{\tau(e')}) = 0 \tag{2.24}$$

Proof. If we consider equation (2.21), it is clear that (2.24) represent the voltage drop between the nodes $\sigma(e)$ and $\tau(e)$ when $\sigma(e')$ is injecting a current of 1 A, $\tau(e')$ is injecting a current of -1 A, while all the other nodes are injecting 0 A (see Fig. 2.3). Since the grid is radial, it is trivial to see that only the current flow ξ'_e is non-zero, and thus only the voltage drop among $\sigma(e')$ and $\tau(e')$ is different from zero. ∎

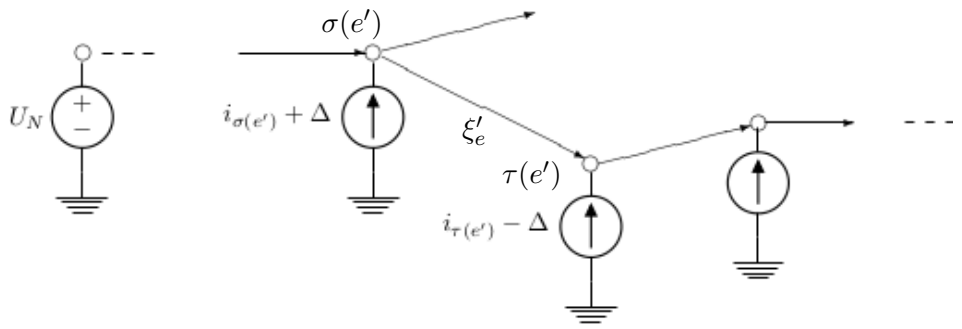


Figure 2.3: Figure explaining Lemma 2.3.2

By adopting a similar the block partition as (2.14), we have

$$\boldsymbol{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \boldsymbol{M} & \boldsymbol{N} \\ 0 & \boldsymbol{N}^T & \boldsymbol{Q} \end{bmatrix}, \tag{2.25}$$

with $\boldsymbol{M} \in \mathbb{C}^{(m-1)\times(m-1)}$, $\boldsymbol{N} \in \mathbb{C}^{(m-1)\times(m-n)}$, and $\boldsymbol{Q} \in \mathbb{C}^{(n-m)\times(n-m)}$. $\boldsymbol{M}$ will be very important for the algorithms proposed and it is related to the Kron reduction of the electric grid.

2.4 Cyber layer

We assume that every microgenerator, and also the PCC, correspond to an *agent* in the cyber layer (upper panel of Figure 2.2), and belong to the set $\mathcal{C} \subseteq \mathcal{G}$ (with $|\mathcal{C}| = m$). We assume furthermore that the agents are provided with some

1. *sensing capabilities*, in particular with a *voltage phasor measurement unit* (PMU) which is a device that measures voltage signals, processes them and returns their amplitude and phase (Phadke and Thorp (2008), Barchi, Macii, and Petri (2013)). allows them to take phasorial measurements of their voltages;

2. *computational capabilities* that will be exploited to implement the proposed algorithms;

3. *communication capabilities*, necessary for the information exchange needed for the algorithm execution. Possibly, agents can communicate via the same power lines (via *power line communication*).

4. *local and partial knowledge of the electrical grid*, e.g. the impedance of the electric path between neighbor agents.

2.5 Power flow problem

The power flow problem is concerned with describing the operating state of an entire power system. It is generally assumed that the impedances throughout the circuit are known, since these are more or less permanent properties of the hardware.

An alternate current (AC) circuit requires exactly two pieces of information per node in order to be completely determined. Thus, if we have a grid of n buses, we need to known n complex variables or $2n$ real variables to know exactly the grid state. Classically, a node can be of three types.

1. **PQ node**: it is a node whose power absorbed or injected $s = p + \mathrm{i}q$ is independent on its voltage, i.e. it satisfies

$$\mathrm{Re}(u\bar{i}) - p = 0 \tag{2.26a}$$
$$\mathrm{Im}(u\bar{i}) - q = 0 \tag{2.26b}$$

2. **PV node**: it is a node whose active power p (absorbed or injected) and whose voltage magnitude $|v|$ is fixed, i.e.

$$\mathrm{Re}(u\bar{i}) - p = 0 \tag{2.27a}$$
$$|u| - |v| = 0 \tag{2.27b}$$

3. **slack bus**: it is a node who imposes a fixed voltage v, i.e. it satisfies

$$\mathrm{Re}(u) - \mathrm{Re}(v) = 0 \tag{2.28a}$$
$$\mathrm{Im}(u) - \mathrm{Im}(v) = 0 \tag{2.28b}$$

   A grid, to be solved, must always have at least one slack bus. If we are considering a distribution network, the slack bus is the PCC.

With (2.26), (2.27) and (2.28), one can build a system of $2n$ equations. As we can see, apart of the voltages, in the former equation also currents appear, augmenting

the number of variables. We can reduce the system to have only voltages as unknown exploiting (2.17). In this way we finally obtain a system of $2n$ equation in $2n$ unknowns. Unfortunately, due to the non linearity of the equations, there is no closed form of the solution. Furthermore, the solution could be unique, multiple or even it might exist. However, it is usually straightforward, in case of multiple solutions, to identify the "true" solution among the mathematical possibilities based on physical plausibility and common sense, which is the one that leads to voltage magnitude closer to the nominal one.

2.6 Approximated solution of the power flow problem

In the following, we will model the PCC as a slack bus, while all the other nodes as PQ buses. This model describes the *steady state* of most loads, and also the behavior of microgenerators, which are typically connected to the grid via power inverters that can commande the amount of power to be injected. Rather then refer to the real equations (2.26) and (2.28), we will refer in the following to the complex equations

$$u\bar{i} = s, \tag{2.29}$$

for PQ nodes, while for the slack bus, modeled as an ideal sinusoidal voltage generator at the microgrid nominal voltage U_N with arbitrary fixed angle φ_0,

$$u_1 = U_N e^{\mathrm{i}\varphi_0}. \tag{2.30}$$

We can therefore describe the grid state with system

$$u = \boldsymbol{X} i + e_0 U_N e^{\mathrm{i}\varphi_0} \tag{2.31a}$$

$$u_1 = U_N e^{\mathrm{i}\varphi_0} \tag{2.31b}$$

$$u_v i_v = s_v \qquad v \neq 1 \tag{2.31c}$$

This is a static system, there are no differential equations, because we will consider only the steady states of the electrical grid,because the scale time of the algorithms proposed is slower enough to allow us to neglect the transitorial behaviors (in voltage frequency and magnitude) between two consecutive states.

In this section we review an approximate explicit solution of the nonlinear system (2.31) which has been proposed in Bolognani and Zampieri (2013). It provides a linearization of the relation among voltages and powers, that basically comes from a first order Taylor expansion.

Proposition 2.6.1. *Consider the physical model described by the system of nonlinear equations* (2.31)*. Node voltages then satisfy*

$$\begin{bmatrix} u_1 \\ u_G \\ u_L \end{bmatrix} = e^{j\varphi_0} \left(U_N \mathbf{1} + \frac{1}{U_N} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \boldsymbol{M} & \boldsymbol{N} \\ 0 & \boldsymbol{N}^T & \boldsymbol{Q} \end{bmatrix} \begin{bmatrix} 0 \\ \bar{s}_G \\ \bar{s}_L \end{bmatrix} \right) + o\left(\frac{1}{U_N}\right) \tag{2.32}$$

where the little-o notation means that $\lim_{U_N \to \infty} \frac{o(f(U_N))}{f(U_N)} = 0$.

The quality of this approximation relies on having large nominal voltage U_N and relatively small currents injected by the inverters (or supplied to the loads). This assumption is verified in practice, and 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. The model proposed in Proposition 2.32 extends the DC power flow model (Gómez-Expósito, Conejo, and Cañizares, 2009, Chapter 3) to the case in which lines are not purely inductive. This way, the model is able to describe the voltage drop on the lines, and therefore also the corresponding power losses, in a form that is conveniently linear in the complex power injections and demands.

2.7 Kron reduction of the grid w.r.t. microgenerators buses

The matrix $\boldsymbol{M}$ in (2.25) will be heavily exploited in the algorithms proposed. Thus in this section we provide an intuition of its physical nature. It is related to the *Kron reduction* (Dorfler and Bullo (2013)) of the electrical grid w.r.t. the microgenerators buses. Essentially, Kron reduction of a connected graph is again a graph whose Laplacian matrix is obtained by the Schur complement of the original Laplacian matrix with respect to a specified subset of nodes. The Kron reduction of networks is ubiquitous in circuit theory and related applications in order to obtain lower dimensional electrically equivalent circuits. It appears for instance in the behavior, synthesis, and analysis of resistive circuits.

We introduce the matrix $\boldsymbol{G}$.

Lemma 2.7.1. *There exists a unique symmetric matrix* $\boldsymbol{G} \in \mathbb{C}^{m\times m}$ *such that*

$$\begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{M} \end{bmatrix} \boldsymbol{G} = I - \mathbf{1}e_1^T \\ \boldsymbol{G}\mathbf{1} = 0. \end{cases}$$

Proof. The following matrix $\boldsymbol{G}$ satisfies the conditions.

$$\boldsymbol{G} = \begin{bmatrix} \mathbf{1}^T\boldsymbol{M}^{-1}\mathbf{1} & -\mathbf{1}^T\boldsymbol{M}^{-1} \\ -\boldsymbol{M}^{-1}\mathbf{1} & \boldsymbol{M}^{-1} \end{bmatrix}. \tag{2.33}$$

The proof of uniqueness, that we omit here, follows exactly the same steps as in the proof of Lemma 2.3.1. ∎

$\boldsymbol{G}$ has the following notable property, that allow us to claim it is the Laplacian of the Kron reduced network with respect to the smart agents nodes.

Lemma 2.7.2. $\boldsymbol{G}$ *is the Schur-complement of* $\boldsymbol{Y}$ *with respect to the indices in* $\mathcal{C}$*.*

Proof. Let partition the bus admittance matrix as

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{cc} & \boldsymbol{Y}_{cl} \\ \boldsymbol{Y}_{cl}^T & \boldsymbol{Y}_{ll} \end{bmatrix} \tag{2.34}$$

where $\boldsymbol{Y}_{cc} \in \mathbb{C}^{m\times m}$ includes both the PCC and the generators components, $\boldsymbol{Y}_{cl} \in \mathbb{C}^{m\times n-m}$, $\boldsymbol{Y}_{ll} \in \mathbb{C}^{n-m\times n-m}$. This partitioning is similar to the one introduced in (2.25), but we are not distinguishing the PCC from the other agents. Accordingly, let us partition the green matrix as

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_{cc} & \boldsymbol{X}_{cl} \\ \boldsymbol{X}_{cl}^T & \boldsymbol{X}_{ll} \end{bmatrix} \tag{2.35}$$

where trivially,

$$\boldsymbol{X}_{cc} = \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{M} \end{bmatrix}$$

Let $\boldsymbol{\Phi} = \boldsymbol{Y}_{cc} - \boldsymbol{Y}_{cl}\boldsymbol{Y}_{ll}^{-1}\boldsymbol{Y}_{cl}^T$ be the Schur complement of $\boldsymbol{Y}$ with respect to the indices in $\mathcal{C}$. If we show that $\boldsymbol{\Phi}\boldsymbol{X}_{cc} = I - e_1\mathbf{1}^T$ and $\boldsymbol{\Phi}\mathbf{1} = 0$ then it will follow, from Lemma 2.7.1,

that $\boldsymbol{G} = \boldsymbol{\Phi}$. The latter follows trivially from the equation $\boldsymbol{Y}\mathbf{1} = 0$. In fact

$$\boldsymbol{Y}\mathbf{1} = \begin{bmatrix} \boldsymbol{Y}_{cc} & \boldsymbol{Y}_{cl} \\ \boldsymbol{Y}_{cl}^T & \boldsymbol{Y}_{ll} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_{cc}\mathbf{1} + \boldsymbol{Y}_{cl}\mathbf{1} \\ \boldsymbol{Y}_{cl}^T\mathbf{1} + \boldsymbol{Y}_{ll}\mathbf{1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.36}$$

from which $\boldsymbol{Y}_{cl}^T\mathbf{1} = -\boldsymbol{Y}_{ll}\mathbf{1}$. Then $\boldsymbol{\Phi}\mathbf{1} = (\boldsymbol{Y}_{cc} - \boldsymbol{Y}_{cl}\boldsymbol{Y}_{ll}^{-1}\boldsymbol{Y}_{cl}^T)\mathbf{1} = 0$. Furthermore, from (2.20) we have

$$\begin{aligned} \boldsymbol{Y}\boldsymbol{X} &= \begin{bmatrix} \boldsymbol{Y}_{cc} & \boldsymbol{Y}_{cl} \\ \boldsymbol{Y}_{cl}^T & \boldsymbol{Y}_{ll} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_{cc} & \boldsymbol{X}_{cl} \\ \boldsymbol{X}_{cl}^T & \boldsymbol{X}_{ll} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{Y}_{cc}\boldsymbol{X}_{cc} + \boldsymbol{Y}_{cl}^T\boldsymbol{X}_{cl} & \boldsymbol{Y}_{cc}\boldsymbol{X}_{cl} + \boldsymbol{Y}_{cl}^T\boldsymbol{X}_{ll} \\ \boldsymbol{Y}_{cl}^T\boldsymbol{X}_{cc} + \boldsymbol{Y}_{ll}\boldsymbol{X}_{cl}^T & \boldsymbol{Y}_{cl}^T\boldsymbol{X}_{cl} + \boldsymbol{Y}_{ll}\boldsymbol{X}_{ll} \end{bmatrix} = \begin{bmatrix} I - e_1\mathbf{1}^T & -e_1\mathbf{1}^T \\ 0 & I \end{bmatrix} \end{aligned} \tag{2.37}$$

that implies $\boldsymbol{Y}_{cl}^T\boldsymbol{X}_{cc} = -\boldsymbol{Y}_{ll}\boldsymbol{X}_{cl}^T$. Thus

$$\boldsymbol{\Phi}\boldsymbol{X}_{cc} = \boldsymbol{Y}_{cc}\boldsymbol{X}_{cc} - \boldsymbol{Y}_{cl}\boldsymbol{Y}_{ll}^{-1}\boldsymbol{Y}_{cl}^T\boldsymbol{X}_{cc} = I - e_1\mathbf{1}$$

■

Lemma II.1, Theorem III.4 and Theorem III.8 in Dorfler and Bullo (2013), show some properties of $\boldsymbol{G}$ that can be exploited for its construction. We summarize them in the following

Lemma 2.7.3. *The matrix $\boldsymbol{G}$ has the sparsity pattern induced by the Definition 4.1.1 of neighbor agents in the cyber layer, i.e.*

$$\boldsymbol{G}_{hk} \neq 0 \quad \Leftrightarrow \quad k \in \mathcal{N}(h).$$

Furthermore, if h and h are neighbors

$$\boldsymbol{G}_{hk} = \frac{1}{z_{hk}^{eff}} \tag{2.38}$$

where z_{hk}^{eff} is the effective impedance between h and k

Summarizing, $\boldsymbol{G}$ is the bus admittance matrix of the reduced grid with only smart agents, while $\boldsymbol{M}$ is its Green matrix.

2.8 Homogeneus line impedance case

In order to design the control and the identification strategies presented in this thesis, in the following we will make the following assumption:

Assumption 2.8.1. All the power lines in the grid have the same inductance over resistance (X/R) ratio, but possibly different impedance magnitude, i.e.

$$\mathrm{z}_e = e^{\mathrm{i}\theta}|\mathrm{z}_e|$$

for any e in $\mathcal{E}$ and for a fixed θ.

This assumption is satisfied when the X/R ratio of the power lines of the grid is relatively homogeneous, which is reasonable in many practical cases (see for example the IEEE standard testbeds Kersting (2001)). It describes a special, ideal case, that is useful since it simplifies the electric model and make the control strategies easier to be developed. It is a very common assumption that the algorithms designer usually adopts. One example of such simplification is to consider purely inductive or purely resistive lines.

Under this assumption, we can write

$$\boldsymbol{Y} = e^{-\mathrm{i}\theta}Y, \tag{2.39}$$

$$\boldsymbol{G} = e^{-\mathrm{i}\theta}G, \tag{2.40}$$

while the Green matrix as

$$\begin{aligned}\boldsymbol{X} &= e^{\mathrm{i}\theta}X \\ &= e^{\mathrm{i}\theta}\begin{bmatrix} 0 & 0 & 0 \\ 0 & M & N \\ 0 & N^T & Q \end{bmatrix}.\end{aligned} \tag{2.41}$$

where $Y = |\boldsymbol{Y}|$, $X = |\boldsymbol{X}|$, $M = |\boldsymbol{M}|$, $N = |\boldsymbol{N}|$ and $Q = |\boldsymbol{Q}|$. As a consequence, equations (2.17) and (2.21) can be rewritten as

$$i = e^{-\mathrm{i}\theta}Yu \tag{2.42}$$

$$u = e^{\mathrm{i}\theta}Xi + \mathbf{1}u_1. \tag{2.43}$$

Assumption 2.8.1 implies that there exist $\alpha \in \mathbb{R}$ such that $\mathrm{Im}(\boldsymbol{Y}_{jk}) = \alpha\,\mathrm{Re}(\boldsymbol{Y}_{jk}), \forall\, \boldsymbol{Y}_{jk}$. It can be shown that $\mathrm{Re}(\boldsymbol{Y})$ is a real, symmetric, positive semidefinite matrix. Therefore,

we can write

$$\begin{aligned}\boldsymbol{Y} &= \operatorname{Re}(\boldsymbol{Y}) + \mathrm{i}\alpha \operatorname{Re}(\boldsymbol{Y}) \\ &= (1 + \mathrm{i}\alpha) \operatorname{Re}(\boldsymbol{Y}) \\ &= (1 + \mathrm{i}\alpha) U \Sigma_R U^* \end{aligned} \tag{2.44}$$

where $\Sigma_R \in \mathbb{R}^{n-1 \times n-1}$ is a diagonal matrix, whose diagonal entries are the $n-1$ non-zero eigenvalues of $\operatorname{Re}(\boldsymbol{Y})$, while $U \in \mathbb{R}^{n \times n-1}$ is a matrix that includes all the associated $n-1$ normalized eigenvector, such that $U^*U = I$. From (2.19), it can be showed that U spans the image of $I - \mathbf{1}\mathbf{1}^T/n$, i.e. the space orthogonal to $\mathbf{1}$. As a consequence, $UU^* = I - \mathbf{1}\mathbf{1}^T/n$.

When Assumption 2.8.1 holds, the Green matrix can be written as the following lemma states

Lemma 2.8.2. *Let $\boldsymbol{Y} = (1 + i\alpha) U \Sigma_R U^*$ be the bus admittance matrix. Then the Green matrix can be written as*

$$\boldsymbol{X} = \frac{1}{1 + i\alpha} (I - \mathbf{1} e_1^T)(U \Sigma_R^{-1} U^*)(I - e_1 \mathbf{1}^T) \tag{2.45}$$

Proof. The right hand side of (2.45) trivially describes a symmetric matrix. Furthermore, it satisfies the equation system (2.20). In fact:

$$\begin{aligned}&\left(\frac{1}{1 + \mathrm{i}\alpha} (I - \mathbf{1} e_1^T)(U \Sigma_R^{-1} U^*)(I - e_1 \mathbf{1}^T)\right) \boldsymbol{Y} = \\ &\qquad = \left(\frac{1}{1 + \mathrm{i}\alpha} (I - \mathbf{1} e_1^T)(U \Sigma_R^{-1} U^*)(I - e_1 \mathbf{1}^T)\right)\left((1 + \mathrm{i}\alpha) U \Sigma_R U^*\right) \\ &\qquad = (I - \mathbf{1} e_1^T)(U \Sigma_R^{-1} U^*) U \Sigma_R U^* \\ &\qquad = (I - \mathbf{1} e_1^T) U U^* \\ &\qquad = (I - \mathbf{1} e_1^T)(I - \mathbf{1}\mathbf{1}^T/n) = (I - \mathbf{1} e_1^T)\end{aligned}$$

and

$$\frac{1}{1 + \mathrm{i}\alpha} (I - \mathbf{1} e_1^T)(U \Sigma_R^{-1} U^*)(I - e_1 \mathbf{1}^T) e_1 = 0.$$

Being the Green matrix unique, the lemma is proved. ∎

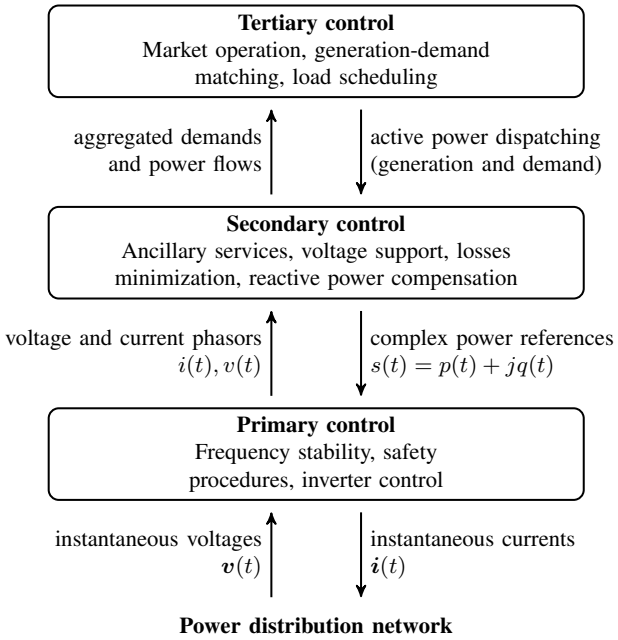


Figure 2.4: A possible layered architecture for the simultaneous execution of different algorithms in a smart microgrid.

2.9 Electrical grid control layers

We've seen that one of the major features that is going to characterize future *smart grids* is the appearance of a large number of microgenerators connected to the LV power distribution network. This scenario poses a number of nontrivial challenges, together with exciting opportunities. The management of future smart grids requires that many control and optimization algorithms are executed at the same time: generation-demand matching protocols, energy market mechanisms, algorithms for optimal energy use and quality of service, and many others. The complexity of such scenario and the different time scales of these control tasks, suggest that a *layered* architecture should be adopted. In a layered architecture different algorithms coexist at different levels and at different time scales. Lower level algorithms, in charge of controlling the specific physical devices, obtain references and commands from higher levels of the architecture. Higher level algorithms command many instances of the lower level ones, based on a simplified model of their behavior and on the aggregated data provided by the underlying layers.

One possible application of this structure to the control of a smart microgrid has been depicted in Figure 2.4 (see also De Brabandere, Vanthournout, Driesen, Deconinck, and Belmans (2007) and references therein).

Tertiary control algorithms (residing in the top layer) dispatch active power generation

on the basis of economic reasons, measured and predicted aggregate demand, and availability of the microgenerators. The reference signals they dispatch come from some optimization process

Secondary control algorithms, on the other hand, take care of sharing the commanded power references among microgenerators, while satisfying a number of operational constraints (voltage limits, stability, congestion avoidance) and optimizing the network operation (power losses minimization, improved power quality). These algorithms require some knowledge of the grid parameters and on the system state.

Primary control algorithms are executed on a local level (at the single inverter), with very little or no communication between the devices, and on a faster time scale; on the basis of the reference signals that they receive, they practically command the power converters that equip each inverter, ensuring frequency stability, avoiding detrimental interactions, and ensuring safety of operation. The aim of primary control is to quickly stabilize the state of the grid, mostly in terms of voltage frequency and amplitude. The most popular approach is the droop control of the inverters (Chandorkar, Divan, and Adapa (1993); Li and Kao (2009); Simpson-Porco, Dorfler, and Bullo (2012)), which does not require direct communications between the agents, but rather exploits voltage frequency and magnitude as information carrier.

The algorithms proposed in the following reside naturally in the secondary and in the tertiary layer, and then live on slower time scale, thus justifying the neglect of fast transitorial behaviors.

Part II

Control and optimization

3

Electrical grid optimization

3.1 Managing the electrical grid

A *transmission system operator* (TSO), called *independent system operator* (ISO) in the U.S., is an entity entrusted to transport energy using fixed infrastructure, from generation plants, over the electrical grid, to regional or local electricity distribution operators. Power flow analysis is an essential tool for transmission system operators to operate a power system, as it predicts how the system would response to certain operations, in the context of either day-to-day operations or longer-term planning.

Sometimes it is necessary to compare several hypothetical operating scenarios, i.e. different hypothetical dispatches of generation units that could meet a given loading condition, to guide operating and planning decisions. This is the motivation that leads to the *optimal power flow* (OPF) problem. The OPF problem constitutes a *static*, *constrained* optimization problem which computes, given settings of loads and system

parameters, optimal settings for electrical variables in a power network. Its objective is to identify the operating configuration or "solution" that best meets a particular set of evaluation criteria. These criteria may include the cost of generation, transmission line losses, and various requirements concerning the system's security, or resilience with respect to disturbances.

It is both economically and computationally a very hard problem, Cain, O'Neill, and Castillo (2012). On one side, an efficient market equilibrium requires multipart non-linear pricing. On the other side the power flow is intrinsically non-linear, and furthermore the optimization problem has nonconvexities, binary variables (e.g. associated with the decision of starting up or shutting down generators) which makes the problem difficult to solve. For investment planning purposes, the problem needs binary investment variables and a multiple years horizon.

Citing Huneault, Galiana, and St Bruno (1991), "the history of optimal power flow (OPF) research can be characterized as the application of increasingly powerful optimization tools to a problem which basically has been well-defined since the early 1960's (Carpentier (1962)), and has been one of the fundamental problems in power system operation ever since".

The OPF problem is becoming more and more important for distribution networks. In the near future, a massive number of small power generators are envisioned to be deployed in the low voltage and medium voltage power distribution grid. Distributed generation is difficult to predict, affecting the traditional control strategy which aim is to follow the power demand. Furthermore, power injection of several renewable energy sources could lead, if not properly regulated, to system instability, thus requiring that also the *distribution system operator* (DSO) tackles with the OPF problems in the low and medium voltage power distribution networks.

3.2 OPF formulation

The OPF problem can be generally stated as a constrained optimization problem, with both equality and inequality constraint

$$\min_{\mathfrak{u}} \quad f(\mathfrak{x}, \mathfrak{u}) \tag{3.1a}$$

$$s.t. \quad h(\mathfrak{x}, \mathfrak{u}) = 0 \tag{3.1b}$$

$$g(\mathfrak{x}, \mathfrak{u}) \leq 0 \tag{3.1c}$$

where $\mathfrak{x}$ is the vector collecting the state variables and $\mathfrak{u}$ the vector collecting the control variables. Problem (3.1) is in general non convex, and hard to solve.

State and control variables

The most common variables that describe the grid state are the complex voltages, or their real and imaginary part, of each node. From the voltage knowledge, exploiting (2.16) and (2.17), one can compute the currents flowing through the lines and injected by the nodes, and thus all the powers.

The optimization (or control) variables depend instead both on the type of devices that perform the control action. They can be

- the active powers and the voltage magnitudes (the classical generators controlled outputs);
- the active and the reactive powers (outputs of the devices interfaced with the grid by inverters);
- the (complex) voltages.

There is a particular case in which the control variables are just the generators reactive power output. In this case the OPF problem is named *optimal reactive power flow* (OPRF) problem, and it will be examined in depth in Chapter 4.

Objective function

There are several objective function that have been proposed. Rather it is dependent on the particular application of the OPF problem. Some of the most common are:

- the *generation cost*, i.e. the cost of the active power produced by the generators. Every generator j of the grid is associated with its own cost, most of the time modeled as a quadratic convex function of the power produced, e.g. $f_j = \alpha_2^j p^2 + \alpha_1^j p + \alpha_0^j, \quad \alpha_2^j > 0$.
- the *grid losses*, i.e. the power dissipated during the transport through the electric lines. Minimizing the losses is equivalent to minimize both the sum of the power generated and the generation cost when all the generators have the same production cost;
- the *number of control action*. Some of the devices that actuate the control actions are electro-mechanical devices, like tap-changers, circuit breakers or capacitor banks.

Thus, by minimizing the number of actions, utility companies aim to delay the devices degradation.

- the *voltage drop along electric lines*. Maintaining the voltage as close as possible to its nominal value is sometimes a very important task, especially in the OPRF problem.

Equality and inequality constraints

The equality constraints (3.1b) typically are the equations that model the system physics, i.e. the power flow equation. Their aim is to enforce the solution to be physically consistent. In the case in which we have just the PCC and PQ-nodes, the equality constraints are given by the set of equations (2.31).

The inequality constraints (3.1c) define instead the acceptable configurations, e.g. establishing limitations on

- the power injected or absorbed by the nodes. For what concerns the instantaneous generation capability of an inverter j attached to a DER generator, it is limited by its fixed apparent power capability $|s_j^M|$, and then it has to satisfy

$$|s_j| \leq |s_{j,max}| \tag{3.2}$$

  i.e., the phasor representing the power injected must lie inside a circle of ray $|s_j^M|$ centered in the origin. Usually, constraint (3.2) is replaced with the couple of box constraints

$$p_{j,min} \leq p_j \leq p_{j,max} \tag{3.3a}$$
$$q_{j,min} \leq q_j \leq q_{j,max} \tag{3.3b}$$

  Equation (3.3) can be used also to model the fixed power absorbed by a load, just setting $p_{j,min} = p_{j,max}$ and $q_{j,min} = q_{j,max}$. This type of constraint models also the maximum power that can be transferred through the PCC, via

$$|s_1| \leq |s_{1,max}|. \tag{3.4}$$

- the voltage magnitude of each node. For the correct operation of the devices connected to node j, the voltage magnitude must be near the nominal U_N, resulting in the constraints

$$U_{min} \leq |u_j| \leq U_{max}. \tag{3.5}$$

Usually we have $U_{min} = (1 - \beta)U_N, U_{max} = (1 + \beta)U_N$, with $\beta = 0.05$ in the U.S., or $\beta = 0.1$ in Europe.

- the magnitude of the current flowing in the lines, i.e. for each branch e

$$|\xi_e| \leq |\xi|_{e,max}. \tag{3.6}$$

Sometimes this constraint is replaced by a limitation on the maximum apparent power $|S_{e,max}|$ that can be injected into a branch. If we consider the edge e, the magnitude of the powers injected from node $\sigma(e)$ (called $s_{\sigma(e),e}$) and from node $\tau(e)$ ($s_{\tau(e),e}$) are

$$s_{\sigma(e),e} = u_{\sigma(e)}\xi_e^*$$
$$s_{\tau(e),e} = u_{\tau(e)}\xi_e^*$$

and the constraint (3.6) can be replaced by

$$|s_{\sigma(e),e}| \leq |S_{e,max}| \tag{3.7a}$$
$$|s_{\tau(e),e}| \leq |S_{e,max}| \tag{3.7b}$$

The inequality constraints can be further divided into two categories: *hard constraints* and *soft constraints*. The first are constraints that can't actually be violated, coming for example from the physics of the devices, like the generation capability of a DER. The latter are instead constraints that can physically be violated, like voltage magnitude constraints, but whose satisfaction assures a correct and safe grid operation and behavior.

The former observations lead to the following general OPF formulation:

$$\min_{s_G} \quad \sum_{j \in \mathcal{C}} f_j([p]_j) \tag{3.8a}$$

$$\text{subject to } p_{j,min} \leq p_j \leq p_{j,max} \quad \forall j \in \mathcal{V} \tag{3.8b}$$

$$q_{j,min} \leq q_j \leq q_{j,max} \quad \forall j \in \mathcal{V} \tag{3.8c}$$

$$U_{min} \leq |u_j| \leq U_{max} \quad \forall j \in \mathcal{V} \tag{3.8d}$$

$$|s_{\sigma(e),e}| \leq |S_{e,max}| \quad \forall e \in \mathcal{E} \tag{3.8e}$$

$$|s_{\tau(e),e}| \leq |S_{e,max}| \quad \forall e \in \mathcal{E} \tag{3.8f}$$

$$u = \boldsymbol{X} i + \mathbf{1} u_1 \tag{3.8g}$$

$$u_1 = U_N \tag{3.8h}$$

3.3 OPF problem solution for distribution networks

In the past, algorithms for the solution of the OPF problem have been applied to the transmission networks, the high voltage networks transporting the electrical power from the power plants to the distribution networks.

Several algorithms solving the OPF problem have been designed specifically for distribution networks, recently. Many of them exploit a powerful optimization technique, the ADMM (Erseghe (2014)). They typically require a large number of iterations and a high computational burden to converge, mainly due to the nonlinear relations among powers and voltages which make the OPF problem non-convex. To overcome these drawbacks, one of the most popular solution is to reformulate the OPF problem as a rank-constrained semidefinite program, which is convexified (Lavaei and Low (2012); Low (2014)) by dropping the rank constraint and it is finally solved in a distributed manner, via a primal or dual optimization (Lam, Zhang, and Tse (2012b)) or via the ADMM (Dall'Anese, Zhu, and Giannakis (2013)).

However all these approaches are based on the standing assumptions, unrealistic in many scenarios, that all the buses of the grid are monitored and all the grid parameters (topology, line impedances etc.) are perfectly known. Additionally, the OPF solution is applied only after a number of communication rounds which are needed by the optimization process to provide the solution. For these reasons, these approaches can be seen as *open loop* strategies, i.e. strategies that control the system by using only a-priori

knowledge of both the model and the power demand. Thus, the errors coming from an imperfect knowledge of the electrical model (e.g. impedance values) or from a wrong load forecast, affect irreparably the output of the algorithms, which imposes setpoints that are not actually optimal.

Chapter 4 and Chapter 5 offer instead a novel solution approach for the OPF approach. The proposed optimization algorithms work as *closed loop strategies*. They still require some a-priori knowledge, even if more limited than the one required by the former approaches. However they inherit the disturbance rejection property, being *feedback* algorithms, and thus the effect of the initial errors can be partially compensated, since the algorithms adjust their output by exploiting the field measurements.

4

A distributed feedback algorithm for the Optimal Reactive Power Flow (ORPF) problem

We introduced the OPRF problem in section 3.2. While the production of active power is costly and heavily dependent on economic reasons, since it comes from actual energy transformation, the reactive power can be produced "for free". Furthermore, reactive power control can be performed with passive electro-mechanical devices, like tap changers or capacitor banks (Baran and Wu (1989)).

In a distribution grid with high-penetration of DERs, like photovoltaic panels, customer demands and generation capabilities throughout the day are rapidly varying. As a result, we may have degradation of power quality, voltage sags and swells that cannot be compensated by slowly responding utility traditional equipment. Although not allowed under current standards for interconnection of distributed generation, fast-reacting, VAR-capable PV inverters may provide the necessary reactive power control to maintain voltage regulation under difficult transient conditions. As side benefit, the control of

reactive power injection at each PV inverter provides an opportunity and a new tool for distribution utilities to optimize the performance of distribution circuits, e.g., by minimizing power losses. In order to properly command the operation of these devices, the distribution network operator is required to solve an *optimal reactive power flow* (ORPF) problem.

The more direct control strategies are purely local: each device responds to its voltage measurements (see for example Turitsyn, Sulc, Backhaus, and Chertkov (2011), Turitsyn, Sulc, Backhaus, and Chertkov (2010) or Kundu and Hiskens (2013)). These strategies basically come from the *standard decoupling approximation*, in which the reactive power is assumed to affect the voltage magnitudes, while the active power the voltage phase. A trivial consequence of the standard decoupling approximation, exploited in these approaches, is that an increment of the reactive power injected by a compensator leads to a voltage rise, while an increment of the reactive power absorbed leads to a voltage decrease. Despite the fact that these approaches are very effective for what concern voltage rise mitigation, fundamental in a DERs high-penetrated scenario (see Carvalho, Correia, and Ferreira (2008); Keane, Ochoa, Vittal, Dent, and Harrison (2011)), the drawback of using only local information and of the absence of coordination is that the steady state condition not always satisfies the voltage magnitude constraints and never achieves optimal performance. Basically because neither optimization problem nor any *feasibility problem* are solved.

Powerful solvers of the ORPF problem have been designed for the ORPF problem, and advanced optimization techniques have been recently specialized for this task (Zhao, Guo, and Cao (2005); Villacci, Bontempi, and Vaccaro (2006); Lavaei, Rantzer, and Low (2011)). However, these approaches assume that an accurate model of the grid is available, that all the grid buses are monitored, that loads announce their demand profiles in advance, and that generators and actuators can be dispatched on a day-ahead, hour-ahead, and real-time basis. For this reason, these solvers are in general offline and centralized, and they collect all the necessary field data, compute the optimal configuration, and dispatch the reactive power production at the generators.

These tools cannot be applied directly to the ORPF problem faced in distribution network, because not all the buses of the grid are monitored, individual loads are unlikely to announce they demand profile in advance, the availability of small size generators is hard to predict (being often correlated with the availability of renewable energy sources). Moreover, the grid parameters, and sometimes even the topology of the grid, are only partially known, and generators are expected to connect and disconnect, requiring an automatic reconfiguration of the grid control infrastructure (the so called *plug and play*

approach).

Recently, some distributed approaches that do not require any central controller, but still require measurements at all the buses of the distribution grid, have been proposed. In order to derive a distributed algorithm for this problem, different convex relaxation methods (Bai, Wei, Fujisawa, and Wang (2008); Zhang and Tse (2013); Lavaei and Low (2012)) have been applied, and various distributed optimization algorithms have been specialized for the resulting convex ORPF problem (Lam, Zhang, Dominiguez-Garcia, and Tse (2012a); Zhang, Dominguez-Garcia, and Tse (2013); Farivar, Neal, Clarke, and Low (2012); Šulc, Backhaus, and Chertkov (2013)).

Our contribution is to provide an algorithm truly scalable in the number of generators. In the proposed strategy, each microgenerator is a smart agent that can measure its phasorial voltage, shares these data with the other agents, and adjusts the amount of reactive power injected into the grid, according to a feedback control law that descends from duality-based methods applied to the optimal reactive power flow problem. The main feature is that we use only a partial knowledge of the grid, since only microgenerators can sense the grid, while the loads are not monitored. However the algorithm reaches optimal performance because from the local phasorial measurements, the agents can infer the global knowledge they need. In some sense, the physical layer works as an analog computer for the smart agents. Their action in fact spreads through the electric lines and thus can be perceived indirectly by other agents, even if there is no direct communication. The algorithm, that solve a particular versions of the optimization problem 3.8, was first presented in Bolognani, Cavraro, and Zampieri (2013d) with only voltage magnitude constraints, in Bolognani, Carli, Cavraro, and Zampieri (2013c) with only power constraints, and finally in Bolognani, Carli, Cavraro, and Zampieri (2013b) and in Bolognani, Carli, Cavraro, and Zampieri (2015) with both the constraints types.

4.1 Preliminary assumptions

In order to develop our control strategy, we assume that Assumption 2.8.1 holds. This assumption is necessary for the algorithm design. However in Section 4.6 we will investigate what is the effect of a possible variability of the R/X ratio, showing the robustness of the proposed strategy against this kind of uncertainty.

We assume that each agent (a microgenerator or the PCC) is provided with some computational capability, with some sensing capability, in the form of a phasor measurement unit (PMU), and finally with a microgenerator that allows them to actuate the system, by injecting the set point for the amount of reactive power.

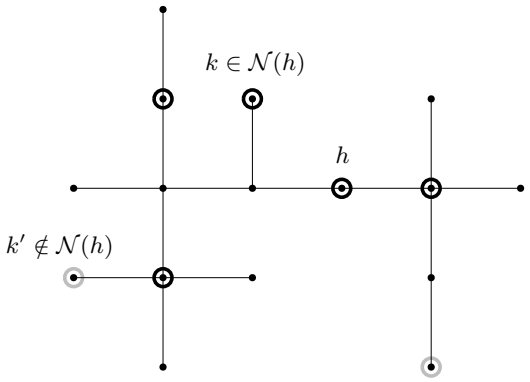


Figure 4.1: An example of neighbor agents in the cyber layer. Circled nodes (both gray and black) are agents (nodes in $\mathcal{C}$). Nodes circled in black belong to the set $\mathcal{N}(h) \subset \mathcal{C}$. Node circled in gray are agents which do not belong to the set of neighbors of h. For each agent $k \in \mathcal{N}(h)$, the path that connects h to k does not include any other agent besides h and k themselves.

We define the neighbors in the cyber layer in the following way:

Definition 4.1.1 (Neighbors in the cyber layer)**.** Let $h \in \mathcal{C}$ be an agent of the cyber layer. The set of agents that are neighbors of h, denoted as $\mathcal{N}(h)$, is the subset of $\mathcal{C}$ defined as

$$\mathcal{N}(h) = \{k \in \mathcal{C} \mid \exists\, \mathcal{P}_{hk}, \mathcal{P}_{hk} \cap \mathcal{C} = \{h, k\}\}\,.$$

Figure 4.1 gives an example of such set. We will assume that every agent $h \in \mathcal{C}$ knows its set of neighbors $\mathcal{N}(h)$. Notice that this architecture can be constructed by each agent in a distributed way, for example by exploiting the *power line communication* (PLC) channel (as suggested for example in Costabeber, Erseghe, Tenti, Tomasin, and Mattavelli (2011)). This allows also a plug-and-play reconfiguration of such architecture when new agents are connected to the grid. Finally, the agents can communicate, via some communication channel that could possibly be the same power lines (via power line communication) with each one of their neighbors.

4.2 Optimal reactive power flow problem

We consider the problem of controlling the reactive power injection of the microgenerators in order to minimize power distribution losses on the power lines and to guarantee that the voltage magnitude and the reactive power injection stay within pre-assigned intervals. The decision variables (or, equivalently, the inputs of the system) are therefore the reactive power setpoints $q_h, h \in \mathcal{C}\backslash\{1\}$, compactly written as q_G.

Power distribution losses can be expressed as a function of the voltages, in a matricial quadratic form, as

$$J_{\text{losses}} = \bar{u}^T Y u \cos\theta. \tag{4.1}$$

In fact we have

$$\begin{aligned} J_{\text{losses}} &= \sum_{e\in\mathcal{E}} |\xi_e|^2 \operatorname{Re}(\mathrm{z}_e) \\ &= \bar{\xi}^T \operatorname{Re}(Z)\xi \\ &= \bar{\xi}^T |Z| \xi \cos\theta \\ &= (-e^{\mathrm{i}\theta}\bar{u}^T A^T |Z|^{-1}) |Z| (-e^{-\mathrm{i}\theta} |Z|^{-1} A\bar{u}) \cos\theta \\ &= \bar{u}^T Y u \cos\theta. \end{aligned}$$

where, we recall, Z=diag$(\mathrm{z}_e, e \in \mathcal{E})$.

Given a lower bound $U_{\min}$ and an upper bound $U_{\max}$ for the voltage magnitudes, and a lower bound $q_{\min}$ and an upper bound $q_{\max}$ for the reactive power injected by each microgenerator, we can therefore formulate the following optimization problem,

$$\begin{aligned} \min_{q_G} \quad & \bar{u}^T L u && \text{(4.2a)} \\ \text{subject to} \quad & |u_h| \geq U_{\min} \quad \forall h \in \mathcal{C} && \text{(4.2b)} \\ & |u_h| \leq U_{\max} \quad \forall h \in \mathcal{C} && \text{(4.2c)} \\ & q_h \geq q_{\min} \quad \forall h \in \mathcal{C}\backslash\{1\} && \text{(4.2d)} \\ & q_h \leq q_{\max} \quad \forall h \in \mathcal{C}\backslash\{1\} && \text{(4.2e)} \end{aligned}$$

where voltages u are a function of the decision variables q_G, via the implicit relation defined by the system of nonlinear equations (2.31). Problem (4.2) can be seen as a particular case of problem (3.8), in which

1. the control variables are just the reactive powers q_G, and not the whole comlex powers s_G;
2. the voltage magnitude constraints are required to be satisfied only by the agents in $\mathcal{C}$;
3. the branch current magnitude limitations are not taken into account.

From a control design prospective, the system-wide problem is characterized by the input variables q_G, the measured output variables $\begin{bmatrix} u_0 \\ u_G \end{bmatrix}$, and the unmeasured disturbances p_L,

q_L, p_G.

In this chapter we design a control algorithm to tackle the ORPF problem in a distributed fashion, where each microgenerator h is allowed to communicate only with its neighbors $\mathcal{N}(h)$.

Remark 4.2.1. While the decision variables of the ORPF problem (i.e. the input variables q_G) do not include the reactive power supplied by the PCC (i.e. $q_0 = u_0\bar{i}_0$), this quantity will also change every time the reactive power set points of the generators are updated by the algorithm, because the inherent physical behavior of the slack bus (the PCC) guarantees that the system of equations (2.31) are satisfied at every time.

Remark 4.2.2. In the above formulation we have assumed that all the microgenerators have the same reactive power injection capability. This scenario can be easily generalized to the case of heterogeneous microgenerators, by replacing (4.2d) and (4.2e) with $q_{\text{h,min}} \leq q_h \leq q_{\text{h,max}}$ being the values $q_{\text{h,min}}$, $q_{\text{h,max}}$, $h \in \mathcal{C}$, in general different for different microgenerators.

4.3 A synchronous algorithm based on dual decomposition

In this section we apply the tool of dual decomposition to (4.2), in order to design a distributed feedback control strategy to solve the ORPF problem. Specifically, we use the approximate explicit solution of the nonlinear system of equations (2.31) introduced in Proposition 2.6.1, to derive update steps for a *dual ascent algorithm* (Bertsekas (1999)) that can be implemented distributively by the agents and that can be used as a feedback control update law.

It is convenient to do following the change of coordinates,

$$\begin{aligned} v_h &= |u_h|^2/U_N^2 & w_h &= 2q_h/U_N^2 \\ v_{\min} &= U_{\min}^2/U_N^2 & w_{\min} &= 2q_{\min}/U_N^2 \\ v_{\max} &= U_{\max}^2/U_N^2 & w_{\max} &= 2q_{\max}/U_N^2. \end{aligned}$$

Basically, we have squared and normalized the constraints on the voltage magnitude and we have normalized the constraints on the power injection. While these modifications does not have any effect on the optimization problem, they will allow us to simplify the derivation of the algorithm we are going to present. We can now rewrite the problem

(4.2) as

$$\min_{q_G} \quad \bar{u}^T L u \tag{4.3a}$$
$$\text{subject to} \quad v_h \geq v_{\min} \quad \forall h \in \mathcal{C} \tag{4.3b}$$
$$v_h \leq v_{\max} \quad \forall h \in \mathcal{C} \tag{4.3c}$$
$$w_h \geq w_{\min} \quad \forall h \in \mathcal{C} \backslash \{1\} \tag{4.3d}$$
$$w_h \leq w_{\max} \quad \forall h \in \mathcal{C} \backslash \{1\} \tag{4.3e}$$

The Lagrangian of (4.3) is

$$\mathcal{L}(q_G, \lambda_{\min}, \lambda_{\max}, \mu_{\min}, \mu_{\max}) = \bar{u}^T L u + \lambda_{\min}^T (v_{\min}\mathbf{1} - v_G) + \lambda_{\max}^T (v_G - v_{\max}\mathbf{1}) + \\ + \mu_{\min}^T (w_{\min}\mathbf{1} - w_G) + \mu_{\max}^T (w_G - w_{\max}\mathbf{1}) \tag{4.4}$$

where $\lambda_{\min}$, $\lambda_{\max}$, $\mu_{\min}$, $\mu_{\max}$ are the Lagrange multipliers (i.e. the dual variables of the problem) and u, v_G, w_G are functions of the decision variables q_G, even if the dependence has been omitted. To have a more compact notation let

$$\nu = \begin{bmatrix} \lambda_{\min}^T & \lambda_{\max}^T & \mu_{\min}^T & \mu_{\max}^T \end{bmatrix}^T.$$

A dual ascent algorithm consists in the iterative execution of the following alternated steps

1. dual gradient ascent step on the dual variables

$$\lambda_{\min}(t+1) = \left[\lambda_{\min}(t) + \gamma \frac{\partial \mathcal{L}(q_G(t), \nu(t))}{\partial \lambda_{\min}}\right]_+$$
$$\lambda_{\max}(t+1) = \left[\lambda_{\max}(t) + \gamma \frac{\partial \mathcal{L}(q_G(t), \nu(t))}{\partial \lambda_{\max}}\right]_+$$
$$\mu_{\min}(t+1) = \left[\mu_{\min}(t) + \gamma \frac{\partial \mathcal{L}(q_G(t), \nu(t))}{\partial \mu_{\min}}\right]_+$$
$$\mu_{\max}(t+1) = \left[\mu_{\max}(t) + \gamma \frac{\partial \mathcal{L}(q_G(t), \nu(t))}{\partial \mu_{\max}}\right]_+,$$

where the $[\cdot]_+$ operator corresponds to the projection on the positive orthant, and where γ is a suitable positive constant;

2. minimization of the Lagrangian with respect to the primal variables q_G

$$q_G(t+1) = \arg\min_{q_G} \mathcal{L}(q_G, \nu(t+1)). \tag{4.5}$$

Observe that the updates of the Lagrange multipliers can be performed naturally in a distributed way by the agents. Indeed let $\lambda_{\min,h}$, $\lambda_{\max,h}$, $\mu_{\min,h}$ and $\mu_{\max,h}$ be the components of the the Lagrange multipliers $\lambda_{\min}$, $\lambda_{\max}$, $\mu_{\min}$ and $\mu_{\max}$, respectively, related to the compensator h. Then it easily follows that the dual step can be implemented as

$$\begin{aligned}
\lambda_{\min,h}(t+1) &= [\lambda_{\min,h}(t) + \gamma(v_{\min} - v_h)]_+ \\
\lambda_{\max,h}(t+1) &= [\lambda_{\max,h}(t) + \gamma(v_h - v_{\max})]_+ \\
\mu_{\min,h}(t+1) &= [\mu_{\min,h}(t) + \gamma(w_{\min} - w_h)]_+ \\
\mu_{\max,h}(t+1) &= [\mu_{\max,h}(t) + \gamma(w_h - w_{\max})]_+ .
\end{aligned}$$

being the partial derivative of the Lagrangian w.r.t. the Lagrange multipliers equal to the voltage and power constraints violation. The following proposition provide instead an approximated form of the Lagrangian partial derivative w.r.t. the primal variables, that will be exploited to approximate the minimizer.

Proposition 4.3.1. *Consider the Lagrangian $\mathcal{L}(q_G, \lambda)$ defined in (4.4). The partial derivative with respect to the primal variables q_G is*

$$\begin{aligned}
\frac{\partial \mathcal{L}(q_G, \nu)}{\partial q_G} = \frac{2}{U_N^2}\Big(&M q_G + N q_L + \sin\theta M(\lambda_{max} - \lambda_{min}) + \\
&+ \mu_{max} - \mu_{min}\Big) + o\left(\frac{1}{U_N^2}\right).
\end{aligned}$$

Proof. From (4.4) we have that

$$\frac{\partial \mathcal{L}(q_G, \nu)}{\partial q_G} = \frac{\partial \bar{u}^T Y u}{\partial q_G} + \left(\frac{\partial v_G}{\partial q_G}\right)^T (\lambda_{\max} - \lambda_{\min}) + \left(\frac{\partial w_G}{\partial q_G}\right)^T (\mu_{\max} - \mu_{\min}) \tag{4.6}$$

In order to derive $\frac{\partial \bar{u}^T Y u}{\partial q_G}$, we introduce the orthogonal decomposition $u = (u' + \mathrm{i}u'')e^{\mathrm{i}(\psi+\theta)}$,

with $u', u'' \in \mathbb{R}^n$. We then have that, via Proposition 2.6.1,

$$\begin{aligned} u' &= \operatorname{Re}\left(ue^{-\mathrm{i}(\phi_0+\theta)}\right) \\ &= \cos\theta U_N \mathbf{1} + \frac{1}{U_N}\begin{bmatrix} 0 & 0 & 0 \\ 0 & M & N \\ 0 & N^T & Q \end{bmatrix}\begin{bmatrix} 0 \\ p_G \\ p_L \end{bmatrix} + o\left(\frac{1}{U_N}\right), \end{aligned}$$

and similarly

$$\begin{aligned} u'' &= \operatorname{Im}\left(ue^{-\mathrm{i}(\phi_0+\theta)}\right) \\ &= -\sin\theta U_N \mathbf{1} - \frac{1}{U_N}\begin{bmatrix} 0 & 0 & 0 \\ 0 & M & N \\ 0 & N^T & Q \end{bmatrix}\begin{bmatrix} 0 \\ q_G \\ q_L \end{bmatrix} + o\left(\frac{1}{U_N}\right). \end{aligned}$$

By using the fact that $\bar{u}^T Y u = u'^T Y u' + u''^T Y u''$, we have

$$\begin{aligned} \frac{\partial \bar{u}^T Y u}{\partial q_G} &= 2\left(\frac{\partial u''}{\partial q_G}\right)^T Y u'' + 2\left(\frac{\partial u'}{\partial q_G}\right)^T Y u' \\ &= -2\left[\frac{1}{U_N}\begin{bmatrix} 0 & M & N \end{bmatrix} + o\left(\frac{1}{U_N}\right)\right] Y u'' + o\left(\frac{1}{U_N^2}\right) \\ &= \frac{2}{U_N^2}\begin{bmatrix} 0 & M & N \end{bmatrix} Y \begin{bmatrix} 0 & 0 & 0 \\ 0 & M & N \\ 0 & N^T & Q \end{bmatrix}\begin{bmatrix} 0 \\ q_G \\ q_L \end{bmatrix} + o\left(\frac{1}{U_N^2}\right) \\ &= \frac{2}{U_N^2}\left(M q_G + N q_L\right) + o\left(\frac{1}{U_N^2}\right), \end{aligned} \tag{4.7}$$

where we used the fact that $Y\mathbf{1} = 0$ and that, by Lemma 2.3.1, $YX = I - e_1 \mathbf{1}^T$.

The same approximate solution (2.32), via some algebraic manipulations, allows us to express v_G as

$$v_G = \mathbf{1} + \frac{2}{U_N^2}\operatorname{Re}\left(e^{\mathrm{i}\theta} M \bar{s}_G + e^{\mathrm{i}\theta} N \bar{s}_L\right) + o\left(\frac{1}{U_N^2}\right). \tag{4.8}$$

We therefore have that

$$\frac{\partial v_G}{\partial q_G} = \frac{2}{U_N^2}\sin\theta M + o\left(\frac{1}{U_N^2}\right), \tag{4.9}$$

while, trivially

$$\frac{\partial w_G}{\partial q_G} = \frac{2}{U_N^2} I, \tag{4.10}$$

and finally, from (4.6), (4.7), (4.9) and (4.10),

$$\frac{\partial \mathcal{L}(q_G,\nu)}{\partial q_G} = \frac{2}{U_N^2}\Big(Mq_G + Nq_L + \sin\theta M(\lambda_{\max} - \lambda_{\min}) + \\ + \mu_{\max} - \mu_{\min}\Big) + o\left(\frac{1}{U_N^2}\right).$$

■

The crucial point is to derive an expression for the minimizer $q_G(t+1)$ in (4.5) that can be computed distributively by the compensators. We are able to do that by exploiting Proposition 2.6.1, that allow us to obtain a value of $q_G(t+1)$ equivalent to the one in (4.5) up to a term which vanishes to zero for large U_N. We assume here that the agents are coordinated, i.e., they can update their state variables q_h, $\lambda_{\min,h}$, $\lambda_{\max,h}$, $\mu_{\min,h}$ and $\mu_{\max,h}$, synchronously. The algorithm we propose is now presented.

Synchronous algorithm

Let all agents (except the PCC) store the auxiliary scalar variables $\lambda_{\min,h}$, $\lambda_{\max,h}$, $\mu_{\min,h}$ and $\mu_{\max,h}$. Let γ be a positive scalar parameter, and let θ be the impedance angle defined in Assumption 2.8.1. Let G_{hk} be the elements of the sparse matrix G defined in Lemma 2.7.1. At every synchronous iteration of the algorithm, each agent $h \in \mathcal{C}\backslash\{0\}$ executes the following operations in order:

- it measures its voltage u_h and it gathers the voltage measurements

$$\{u_k = |u_k|\exp(j\angle u_k), k \in \mathcal{N}(h)\}$$

  from its neighbors;

- it updates the auxiliary variables $\lambda_{\min,h}$, $\lambda_{\max,h}$, $\mu_{\min,h}$, $\mu_{\max,h}$, as

$$\lambda_{\min,h} \leftarrow \left[\lambda_{\min,h} + \gamma\left(\frac{U_{\min}^2}{U_N^2} - \frac{|u_h|^2}{U_N^2}\right)\right]_+$$
$$\lambda_{\max,h} \leftarrow \left[\lambda_{\max,h} + \gamma\left(\frac{|u_h|^2}{U_N^2} - \frac{U_{\max}^2}{U_N^2}\right)\right]_+$$
$$\mu_{\min,h} \leftarrow \left[\mu_{\min,h} + \gamma\left(\frac{q_{\min}}{U_N^2} - \frac{q_h}{U_N^2}\right)\right]_+$$
$$\mu_{\max,h} \leftarrow \left[\mu_{\max,h} + \gamma\left(\frac{q_h}{U_N^2} - \frac{q_{\max}}{U_N^2}\right)\right]_+;$$

- it gathers from its neighbors the updated values of the Lagrange multipliers $\mu_{\min,k}$, $\mu_{\max,k}$, $k \in \mathcal{N}(h)$;
- based on the new values of $\lambda_{\max,h}$, $\lambda_{\min,h}$ and of $\mu_{\min,k}$, $\mu_{\max,k}$, $k \in \mathcal{N}(h)$, it updates the injected reactive power q_h as

$$\begin{aligned} q_h \leftarrow q_h &- \sin\theta(\lambda_{\max,h} - \lambda_{\min,h}) \\ &+ \sum_{k\in\mathcal{N}(h)} G_{hk}|u_h||u_k|\sin(\angle u_k - \angle u_h - \theta) \\ &- \sum_{k\in\mathcal{N}(h)\setminus\{0\}} G_{hk}(\mu_{\max,k} - \mu_{\min,k}). \end{aligned} \tag{4.11}$$

Observe that the above algorithm can be implemented in a completely distributed fashion. Indeed each agent is required to exchange information only with its neighbors in the cyber layer.

The following Proposition shows how the update (4.11) approximates the primal step (4.5).

Proposition 4.3.2. *Consider the synchronous algorithm above described. Then*

$$\frac{\partial\mathcal{L}(q_G(t+1),\nu(t+1))}{\partial q_G} = o\left(\frac{1}{U_N^2}\right),$$

namely, the update (4.11), minimizes the Lagrangian with respect to the primal variables, up to a term that vanishes for large U_N.

Proof. It can be shown, by using Proposition 2.7.1 and via some algebraic manipulation, that the update (4.11) can be also rewritten as

$$\begin{aligned} q_G \leftarrow q_G(t) &- \sin\theta(\lambda_{\max}(t+1) - \lambda_{\min}(t+1)) + \\ &- M^{-1}(\mu_{\max}(t+1) - \mu_{\min}(t+1)) + \\ &+ \mathrm{Im}\left(e^{-j\theta}\,\mathrm{diag}(\bar{u}_G)\begin{bmatrix} M^{-1}\mathbf{1} & M^{-1}\end{bmatrix}\begin{bmatrix} u_0 \\ u_G\end{bmatrix}\right) \end{aligned}$$

which, by using the expression for u provided by Proposition 2.6.1, is equal to

$$\begin{aligned} q_G \leftarrow q_G(t) &- \sin\theta(\lambda_{\max}(t+1) - \lambda_{\min}(t+1)) + \\ &- M^{-1}(\mu_{\max}(t+1) - \mu_{\min}(t+1)) + \\ &- (q_G(t) + M^{-1}Nq_L) + o\left(\frac{1}{U_N}\right). \end{aligned} \tag{4.12}$$

Then, after the update, by plugging the former into the expression for the partial derivative of the Lagrangian with the respect to q_G, provided in Proposition 4.3.1, we obtain

$$\frac{\partial \mathcal{L}(q_G(t+1), \nu(t+1))}{\partial q_G} = \frac{2}{U_N^2}\Big(Mq_G(t+1) + Nq_L + \sin\theta M(\lambda_{\max}(t+1) - \lambda_{\min}(t+1)) + \\ + \mu_{\max}(t+1) - \mu_{\min}(t+1)\Big) + o\left(\frac{1}{U_N^2}\right) = o\left(\frac{1}{U_N^2}\right).$$

and therefore the update minimized the Lagrangian with respect to the primal variables, up to a term that vanishes for large U_N. ■

Via these steps, we therefore specialized the dual ascent steps to the ORPF problem that we are considering, and we obtained a distributed feedback control law for the system. We study the convergence of the closed loop system in Section 4.5.

Remark 4.3.3. It is important to notice that the proposed synchronous algorithm requires that the agents actuate the system at every iteration, by updating the set point for the amount of reactive power injected by the microgenerators. Only by doing so, the subsequent measurement of the voltages will be informative of the new state of the system. The resulting control strategy is thus a *feedback strategy* that necessarily requires the real-time interaction of the controller (the cyber layer) with the plant (the physical layer), as depicted in Figure 4.2. This tight interaction between the cyber layer and the physical layer is the fundamental feature of the proposed approach, and allows to drive the system towards the optimal configuration, which in principle depends on the reactive power demands of the loads, without collecting this information from them. In a sense, the algorithm is inferring this hidden information from the measurement performed on the system during its execution.

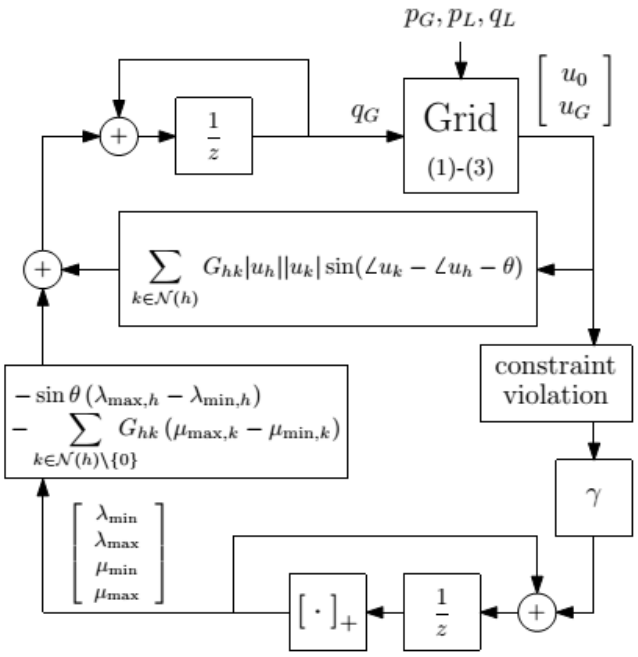


Figure 4.2: A block diagram representation of the synchronous control algorithm proposed in Section 4.3, where the tight interconnection of the cyber and the physical layer (i.e. the feedback strategy) is evident.

4.4 Asynchronous algorithm

In order to avoid the burden of system-wide coordination among the agents, we also propose an asynchronous version of the algorithm, in which the agents corresponding to the microgenerators update their state $(q_h, \lambda_{\max,h}, \lambda_{\min,h}, \mu_{\max,h}, \mu_{\min,h})$ independently one from the other, based on the information that they can gather from their neighbors.

We assume that each agent (except for the agent located at the PCC) is provided with an individual timer, by which it is triggered, and that no coordination is present between these timers: they tick randomly, with exponentially, identically distributed waiting times.

Asynchronous algorithm

Let all agents (except the PCC) store four auxiliary scalar variables $\lambda_{\max,h}$, $\lambda_{\min,h}$, $\mu_{\max,h}$, $\mu_{\min,h}$. Let γ be a positive scalar parameter, and let θ be the impedance angle defined in Assumption 2.8.1. Let G_{hk} be the elements of the matrix G defined in Lemma 2.7.1.

When agent $h \in \mathcal{C}\setminus\{0\}$ is triggered by its own timer, it performs the following actions in order:

- it measures its voltage u_h and it gathers from its neighbors the voltage measurements

$$\{u_k = |u_k| \exp(j\angle u_k), k \in \mathcal{N}(h)\}$$

and the values of the Lagrange multipliers

$$\{\mu_{\min,k}, \mu_{\max,k},\ k \in \mathcal{N}(h)\};$$

- it updates the auxiliary variables $\lambda_{\min,h}$, $\lambda_{\max,h}$, $\mu_{\min,h}$, $\mu_{\max,h}$, as

$$\begin{aligned}
\lambda_{\min,h} &\leftarrow \left[\lambda_{\min,h} + \gamma\left(\frac{U_{\min}^2}{U_N^2} - \frac{|u_h|^2}{U_N^2}\right)\right]_+ \\
\lambda_{\max,h} &\leftarrow \left[\lambda_{\max,h} + \gamma\left(\frac{|u_h|^2}{U_N^2} - \frac{U_{\max}^2}{U_N^2}\right)\right]_+ \\
\mu_{\min,h} &\leftarrow \left[\mu_{\min,h} + \gamma\left(\frac{q_{\min}}{U_N^2} - \frac{q_h}{U_N^2}\right)\right]_+ \\
\mu_{\max,h} &\leftarrow \left[\lambda_{\max,h} + \gamma\left(\frac{q_h}{U_N^2} - \frac{q_{\max}}{U_N^2}\right)\right]_+ ;
\end{aligned}$$

- based on the new value of $\lambda_{\min,h}$, $\lambda_{\max,h}$, it updates the injected reactive power q_h as

$$\begin{aligned}
q_h \leftarrow\ & q_h - \sin\theta(\lambda_{\max,h} - \lambda_{\min,h}) + \\
& + \sum_{k\in\mathcal{N}(h)} G_{hk}|u_h||u_k| \sin(\angle u_k - \angle u_h - \theta) + \\
& - \sum_{k\in\mathcal{N}(h)\setminus\{0\}} G_{hk}(\mu_{\max,k} - \mu_{\min,k}).
\end{aligned} \tag{4.13}$$

The update equations for the asynchronous algorithm are exactly the same of the synchronous case. Here, however, we update both the primal and the dual variable of the agents independently and asynchronously. Also the analysis of the convergence of this algorithm is postponed to the next section.

4.5 Convergence analysis

In this section, we investigate the convergence of both the synchronous algorithm proposed in Section 4.3 and of the asynchronous algorithm proposed in Section 4.4.

In order to do so, we rewrite the terms that appeared in the dual ascent update step, namely

$$\frac{\partial \mathcal{L}(q_G,\nu)}{\partial \lambda_{\min}}, \quad \frac{\partial \mathcal{L}(q_G,\nu)}{\partial \lambda_{\max}}, \quad \frac{\partial \mathcal{L}(q_G,\nu)}{\partial \mu_{\min}}, \quad \frac{\partial \mathcal{L}(q_G,\nu)}{\partial \mu_{\max}},$$

using the expression introduced in Proposition 2.6.1 for the voltages. We start from

$$\frac{\partial \mathcal{L}(q_G,\nu)}{\partial \lambda_{\min}} = v_{\min}\mathbf{1} - v_G.$$

By plugging in the approximate solution (2.32), via some algebraic manipulations, we can express v_G as

$$\begin{aligned} v_G &= \mathbf{1} + \frac{2}{U_N^2}\operatorname{Re}\left(e^{j\theta}M\bar{s}_G + e^{j\theta}N\bar{s}_L\right) + o\left(\frac{1}{U_N^2}\right) \\ &= \mathbf{1} + \frac{2}{U_N^2}(\sin\theta M q_G + \cos\theta M p_G) \\ &\quad + \frac{2}{U_N^2}(\cos\theta N p_L + \sin\theta N q_L) + o\left(\frac{1}{U_N^2}\right) \end{aligned} \tag{4.14}$$

and, in turn, we have that

$$\frac{\partial \mathcal{L}(q_G,\nu)}{\partial \lambda_{\min}} = \frac{2}{U_N^2}(b_{\min} - \sin\theta M q_G) + o\left(\frac{1}{U_N^2}\right) \tag{4.15}$$

where

$$b_{\min} = \frac{U_N^2}{2}(v_{\min} - 1)\mathbf{1} - (Mp_G\cos\theta + N(p_L\cos\theta + q_L\sin\theta)).$$

Similar calculations lead to

$$\frac{\partial \mathcal{L}(q_G,\nu)}{\partial \lambda_{\max}} = \frac{2}{U_N^2}(\sin\theta M q_G - b_{\max}) + o\left(\frac{1}{U_N^2}\right),$$

where

$$b_{\max} = \frac{U_N^2}{2}(v_{\max} - 1)\mathbf{1} - (Mp_G\cos\theta + N(p_L\cos\theta + q_L\sin\theta)).$$

Furthermore, observe that

$$\begin{aligned} \frac{\partial \mathcal{L}(q_G,\nu)}{\partial \mu_{\min}} &= w_{\min}\mathbf{1} - w_G = \frac{2}{U_N^2}(\mathbf{1}q_{\min} - q_G), \\ \frac{\partial \mathcal{L}(q_G,\nu)}{\partial \mu_{\max}} &= w_G - w_{\max}\mathbf{1} = \frac{2}{U_N^2}(q_G - \mathbf{1}q_{\max}). \end{aligned}$$

The proposed dual ascent step can therefore be rewritten in compact form as

$$\nu(t+1) = \left[\nu(t) + \gamma \frac{2}{U_N^2}\left(\Phi q_G(t) + b\right) + o\left(\frac{1}{U_N^2}\right)\right]_+, \tag{4.16}$$

where

$$\Phi = \begin{bmatrix} -\sin\theta M \\ \sin\theta M \\ -I \\ I \end{bmatrix}, \quad b = \begin{bmatrix} b_{\min} \\ -b_{\max} \\ \mathbf{1}q_{\min} \\ -\mathbf{1}q_{\max} \end{bmatrix}. \tag{4.17}$$

The update step for the primal variables can be rewritten based on Proposition 4.3.2 and Proposition 4.3.1, obtaining

$$q_G(t+1) = -M^{-1}Nq_L - M^{-1}\Phi\nu(t+1) + o\left(\frac{1}{U_N}\right). \tag{4.18}$$

In the analysis that follows, we study the approximated description of the closed loop system in which we neglect the infinitesimal terms. Notice that, by doing so, both the voltages u and the squared voltage magnitudes v_G become affine functions of the decision variables q_G. By plugging those expressions in the formulation of (4.3), one obtains the following strictly convex quadratic problem with linear inequality constraints.

$$\min_{q_G} \quad q_G^T \frac{M}{2} q_G + q_G^T N q_L \tag{4.19a}$$

$$\text{subject to} \quad \Phi q_G + b \leq 0, \tag{4.19b}$$

for which strong duality holds. The rest of the section is split into two subsections: in the first one we consider the synchronous version of the algorithm, in the second one the asynchronous version.

Remark 4.5.1. Basically we obtain a convexified problem, exploiting the approximation introduced in Proposition 2.6.1. For the resulting problem, holding strong duality, there always exists a solution, even if not exact but approximated. This convexification is quite different w.r.t. the popular semidefinite programming relaxation (e.g. see Lavaei and Low (2012)). The semidefinite programming relaxation leads to an optimization problem whose solution is exact, but sometimes it does not exist, as shown in Lesieutre, Molzahn, Borden, and DeMarco (2011)

Syncronous case

For the synchronous version of the algorithm, we consider the update equations

$$\nu(t+1) = \left[\nu(t) + \gamma \frac{2}{U_N^2}\left(\Phi q_G(t) + b\right)\right]_+, \tag{4.20}$$

for the dual variables, and

$$q_G(t+1) = -M^{-1}Nq_L - M^{-1}\Phi^T\nu(t+1). \tag{4.21}$$

for the primal variables. Observe that (4.20) and (4.21) differ from (4.16) and from (4.18) only by infinitesimal terms, and they correspond to the standard equation for the dual ascent steps for (4.19). Indeed, the equilibrium (q_G^*, ν^*) of (4.20)-(4.21) is characterized by

$$\Phi q_G^* + b \leq 0 \quad \text{and} \quad q_G^* + M^{-1}Nq_L + M^{-1}\Phi^T\nu^* = 0,$$

which correspond to the necessary conditions for the optimality according to Uzawa's saddle point theorem Uzawa (1958).

It will be useful in the following to define $\sigma_{\min}$ and $\sigma_{\max}$ as the minimum and the maximum eigenvalue of M, respectively. The following result characterizes the convergence of the algorithm described by (4.20) and (4.21).

Theorem 4.5.2. *Consider the optimization problem (4.19) and the dynamic system described by the update equations (4.20) and (4.21). Then the trajectory $t \to q(t)$ converges to the optimal primal solution q_G^* if*

$$\gamma \leq \frac{U_N^2}{\rho(\Phi M^{-1}\Phi^T)},$$

where

$$\rho(\Phi M^{-1}\Phi^T) = 2\max\{\sigma_{min}^{-1} + \sin^2\theta\sigma_{min}, \sigma_{max}^{-1} + \sin^2\theta\sigma_{max}\}.$$

Proof. It is straightforward to see that the dual of (4.19) is

$$\max_{\nu \geq 0} \quad g(\nu) \tag{4.22}$$

where

$$g(\nu) = -\nu^T \frac{\Phi M^{-1} \Phi^T}{2} \nu - \nu^T (\Phi M^{-1} N q_L - b) - q_L^T \frac{N^T M^{-1} N}{2} q_L. \tag{4.23}$$

Since (4.19) is quadratic optimization problem that we have assumed feasible and the constraint is expressed by a linear affine inequality, the Slater's condition (Boyd and Vandenberghe, 2004, p. 226) holds and then there is zero duality gap between (4.19) and (4.22). Observe that, by plugging (4.21) into (4.20) we obtain

$$\begin{aligned} &\nu(t+1) = \\ &= \left[(I - \gamma \frac{2}{U_N^2} \Phi M^{-1} \Phi^T) \nu(t) - \gamma \frac{2}{U_N^2} (\Phi M^{-1} N q_L - b) \right]_+ \\ &= \left[\nu(t) + \gamma \frac{2}{U_N^2} \frac{\partial g(\nu)}{\partial \lambda} \right]_+ \end{aligned} \tag{4.24}$$

that is the update of ν is a projected gradient ascent algorithm for the dual function $g(\lambda)$. Then any optimal solution ν^* of (4.22) is a fixed point for (4.24) and satisfies

$$\nu^* = \left[(I - \gamma \frac{2}{U_N^2} \Phi M^{-1} \Phi^T) \nu^* - \gamma \frac{2}{U_N^2} (\Phi M^{-1} N q_L - b) \right]_+ \tag{4.25}$$

while the primal optimal solution, from (4.21), has the form

$$q_G^* = -M^{-1} N q_L - M^{-1} \Phi^T \nu^*. \tag{4.26}$$

It is worth to notice that (4.23) has not necessarily a unique solution: given a particular solution ν_1^*, if there exists $v \in \ker(\Phi^T)$ such that $\nu_2^* = \nu_1^* + v = [\nu_2^*]_+$, then also ν_2^* is an optimal solution. Despite that, q_G^* is unique. In fact we have

$$\begin{aligned} q_G^* &= -M^{-1} N q_L - M^{-1} \Phi^T \nu_1^* \\ &= -M^{-1} N q_L - M^{-1} \Phi^T (\nu_1^* + v) \\ &= -M^{-1} N q_L - M^{-1} \Phi^T \nu_2^*. \end{aligned}$$

Notice that we have, $\forall \nu_1, \nu_2 \geq 0$, that

$$\begin{aligned}\|\nabla g(\nu_1) - \nabla g(\nu_2)\| &= \|(\Phi M^{-1}\Phi^T)(\nu_1 - \nu_2)\| \\ &\leq \|\Phi M^{-1}\Phi^T\|\|\nu_1 - \nu_2\|\end{aligned}$$

and then the gradient $\nabla g(\nu)$ is Lipschitz continuous with Lipschitz constant equals to $\|\Phi M^{-1}\Phi^T\|$. Then, from Prop. 2.3.2 in Bertsekas (1999), if

$$\gamma \leq \frac{U_N^2}{\rho(\Phi M^{-1}\Phi^T)} \tag{4.27}$$

the algorithm (4.24) converges to a maximizer of $g(\nu)$ and then we reach the optimal solution q_G^* of the primal optimization problem. We have

$$\Phi M^{-1}\Phi^T = \\ = \begin{bmatrix} \sin^2\theta M & -\sin^2\theta M & \sin\theta I & -\sin\theta I \\ -\sin^2\theta M & \sin^2\theta M & -\sin\theta I & \sin\theta I \\ \sin\theta I & -\sin\theta I & M^{-1} & -M^{-1} \\ -\sin\theta I & \sin\theta I & -M^{-1} & M^{-1} \end{bmatrix}.$$

Being $\Phi M^{-1}\Phi^T$ a positive semi-definite symmetric matrix, its norm is equal to its spectral radius. It can be shown that the spectrum of $\Phi M^{-1}\Phi^T$ is given by $\Lambda(\Phi M^{-1}\Phi^T) = \{0\} \cup \Lambda(2\Xi)$ where

$$\Xi = \begin{bmatrix} \sin^2\theta M & -\sin\theta I \\ -\sin\theta I & M^{-1} \end{bmatrix}.$$

The characteristic polynomial of Ξ is

$$P(z) = \prod_{j=1}^{n} z(z - \sigma_j^{-1} - \sin^2\theta\sigma_j) \tag{4.28}$$

where σ_j is the j-th eigenvalues of M and where, without loss of generality we assume that $0 < \sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_{m-1}$. Thus we can see that

$$\Lambda(\Phi M^{-1}\Phi^T) = \{0, 2(\sigma_1^{-1} + \sin^2\theta\sigma_1), \ldots, 2(\sigma_{m-1}^{-1} + \sin^2\theta\sigma_{m-1})\}$$

and the spectral radius is

$$\rho(\Phi M^{-1}\Phi^T) = 2\max\{\sigma_1^{-1} + \sin^2\theta\sigma_1, \sigma_{m-1}^{-1} + \sin^2\theta\sigma_{m-1}\}$$

■

We conclude this subsection by specializing the above result to the case where either only voltage constraints or only power constraints are considered. Observe that if we take into account only voltage constraints then the matrix Φ and the vector b become

$$\Phi = \begin{bmatrix} -\sin\theta M \\ \sin\theta M \end{bmatrix}, \quad b = \begin{bmatrix} b_{\min} \\ -b_{\max} \end{bmatrix} \tag{4.29}$$

and only the multipliers $\lambda_{\min}$ and $\lambda_{\max}$ are employed in the algorithm, while if we consider only power constraints then

$$\Phi = \begin{bmatrix} -I \\ I \end{bmatrix}, \quad b = \begin{bmatrix} \mathbf{1}q_{\min} \\ -\mathbf{1}q_{\max} \end{bmatrix} \tag{4.30}$$

and only the multipliers $\mu_{\min}$ and $\mu_{\max}$ are needed. The following results follow from Theorem 4.5.2.

Corollary 4.5.3. *Consider the optimization problem (4.19), where Φ and b are given as in (4.29), and the dynamic system described by the update equations (4.20) and (4.21). Then the trajectory $t \to q(t)$ converges to the optimal primal solution q_G^* if*

$$\gamma \leq \frac{U_N^2}{2\,\sin^2\theta\;\sigma_{max}},$$

Corollary 4.5.4. *Consider the optimization problem (4.19), where Φ and b are given as in (4.30), and the dynamic system described by the update equations (4.20) and (4.21). Then the trajectory $t \to q(t)$ converges to the optimal primal solution q_G^* if*

$$\gamma \leq \frac{\sigma_{min}\,U_N^2}{2}.$$

Proof of Corollories 4.5.3 and 4.5.4. Consider the case where only voltage constraints are considered. Then we have that

$$\Phi M^{-1}\Phi^T = \sin^2\theta \begin{bmatrix} M & -M \\ -M & M \end{bmatrix}$$

from which it follows that $\rho\left(\Phi M^{-1}\Phi^T\right) = 2\sin^2\theta\sigma_{\max}$.

Instead when only power constraints are taken into account we have that

$$\Phi M^{-1}\Phi^T = \begin{bmatrix} M^{-1} & -M^{-1} \\ -M^{-1} & M^{-1} \end{bmatrix}$$

from which we get that $\rho\left(\Phi M^{-1}\Phi^T\right) = 2\sigma_{\min}^{-1}$. ■

Asynchronous case

We introduce the following assumption.

Assumption 4.5.5. Let $\{T_i^{(h)}\}$, $i \in \mathbb{N}$, be the time instants in which the agent h is triggered by its own timer. We assume that the timer ticks with exponentially distributed waiting times, identically distributed for all the agents in $\mathcal{C}\backslash\{0\}$.

Let us define the random sequence $h(t) \in \mathcal{C}\backslash\{0\}$ which tells which agent has been triggered at iteration t of the algorithm. Because of Assumption 4.5.5, the random process $h(t)$ is an i.i.d. uniform process on the alphabet $\mathcal{C}\backslash\{0\}$. If we repeat the same analysis, neglecting the infinitesimal terms, we obtain the following update equations for the primal and dual variables, instead of (4.20) and (4.21). In these equations, only the component $h(t)$ of the vectors $\lambda_{\min}$, $\lambda_{\max}$, $\mu_{\min}$, $\mu_{\max}$ and q_G is updated at time t, namely,

$$\begin{aligned}
\lambda_{\min,h(t)}(t+1) &= \Big[\lambda_{\min,h(t)}(t) + \\
&\quad + \gamma\frac{2}{U_N^2}\mathbf{1}_{h(t)}^T(b_{\min} - \sin\theta M q_G(t))\Big]_+ \\
\lambda_{\max,h(t)}(t+1) &= \Big[\lambda_{\max,h(t)}(t) + \\
&\quad + \gamma\frac{2}{U_N^2}\mathbf{1}_{h(t)}^T(\sin\theta M q_G(t) - b_{\max})\Big]_+ \\
\mu_{\min,h(t)}(t+1) &= \Big[\mu_{\min,h(t)}(t) + \gamma\frac{2}{U_N^2}(q_{\min} - q_h(t))\Big]_+ \\
\mu_{\max,h(t)}(t+1) &= \Big[\mu_{\max,h(t)}(t) + \gamma\frac{2}{U_N^2}(q_h(t) - q_{\max})\Big]_+
\end{aligned} \tag{4.31}$$

while

$$\begin{aligned}
\lambda_{\min,k}(t+1) &= \lambda_{\min,k}(t) \\
\lambda_{\max,k}(t+1) &= \lambda_{\max,k}(t) \\
\mu_{\min,k}(t+1) &= \mu_{\min,k}(t) \\
\mu_{\max,k}(t+1) &= \mu_{\max,k}(t)
\end{aligned} \qquad \forall k \neq h(t), \tag{4.32}$$

and

$$\begin{aligned} q_{h(t)}(t+1) &= -\mathbf{1}_{h(t)}^T(M^{-1}Nq_L + M^{-1}\Phi^T\nu(t+1)) \\ q_k(t+1) &= q_k(t) \qquad \forall k \neq h(t). \end{aligned} \tag{4.33}$$

Notice that, also in the asynchronous case, Uzawa's necessary conditions for optimality are satisfied at the equilibrium of (4.31), (4.32) and (4.33). For the asynchronous version of the algorithm we can provide theoretical results only when Φ and b assume the form in (4.29) or in (4.30), namely, when we consider either only voltage constraints or only power constraints. However in the numerical section we show the effectiveness of the asynchronous algorithm when Φ and b assume the general form in (4.17).

The following convergence results hold.

Proposition 4.5.6. *Consider the optimization problem (4.19), where Φ and b are given as in (4.29), and the dynamic system described by the update equations (4.31) and (4.32) (for the multipliers λ_{min} and λ_{max}) and (4.33). Let Assumption 4.5.5 hold. Then the evolution $t \to q(t)$ converges almost surely to the optimal primal solution q_G^* if*

$$\gamma \leq \frac{U_N^2}{2\,\sin^2\theta\;\sigma_{max}}.$$

Proposition 4.5.7. *Consider the optimization problem (4.19), where Φ and b are given as in (4.30), and the dynamic system described by the update equations (4.31) and (4.32) (for the multipliers μ_{min} and μ_{max}) and (4.33). Let Assumption 4.5.5 hold. Then the evolution $t \to q(t)$ converges almost surely to the optimal primal solution q_G^* if*

$$\gamma \leq \frac{\sigma_{min}\,U_N^2}{2}.$$

Proof of Propositions 4.5.6 and 4.5.7. Consider the update equations (4.31), (4.32) and (4.33) for the dual variables ν and for the primal variables q_G. Let (q_G^*, ν^*) be a solution of the optimization problem, which satisfies (4.25) and the KKT conditions

$$q_G^* + M^{-1}Nq_L + M^{-1}\Phi^T\nu^* = 0 \tag{4.34a}$$

$$\Phi q^* + b \leq 0 \quad \forall h \in \mathcal{C} \tag{4.34b}$$

$$\Phi q^* + b < 0 \quad \Leftrightarrow \quad \nu_h^* = 0. \tag{4.34c}$$

We introduce the following two quantities

$$x(t) = q_G(t) - q_G^* \quad \text{and} \quad y(t) = \nu(t) - \nu^*.$$

Without loss of generality let us assume that node h is the node performing the update at the t-th iteration. The update for the variable x is given by

$$\begin{aligned} x_h(t+1) &= q_h(t+1) - q_h^* \\ &= -\mathbf{1}_h^T M^{-1} N q_L - \mathbf{1}_h^T M^{-1} \Phi \nu(t+1) - q_h^* \\ &= -\mathbf{1}_h^T M^{-1} N q_L - \mathbf{1}_h^T M^{-1} \Phi \nu(t+1) \\ &\qquad + \mathbf{1}_h^T M^{-1} N q_L + \mathbf{1}_h^T M^{-1} \Phi^T \nu^* \\ &= -\mathbf{1}_h^T M^{-1} \Phi^T (\nu(t+1) - \nu^*) \\ &= -\mathbf{1}_h^T M^{-1} \Phi^T y(t+1), \end{aligned} \tag{4.35}$$

where we used (4.33) and (4.34a). Now, let us consider first the case where only voltage constraints are taken into account. Via some algebraic manipulations we can write from (4.31) that

$$\begin{aligned} &\lambda_{\min,h}(t+1) - \lambda^*_{\min,h} = \\ &= \left[\lambda_{\min,h}(t) - \lambda^*_{\min,h} - \gamma \frac{2}{U_N^2} \sin\theta \mathbf{1}_h^T M x(t) + \alpha_{\min,h} \right]_+ \\ &\quad - [\alpha_{\min,h}]_+ \end{aligned}$$

where

$$\alpha_{\min,h} = \lambda^*_{\min,h} + \gamma \frac{2}{U_N^2} \mathbf{1}_h^T (b_{\min,h} - \sin\theta M q_G^*(t)).$$

and

$$\begin{aligned} &\lambda_{\max,h}(t+1) - \lambda^*_{\max,h} \\ &= \left[\lambda_{\max,h}(t) - \lambda^*_{\max,h} + \gamma \frac{2}{U_N^2} \sin\theta \mathbf{1}_h^T M x(t) + \alpha_{\max,h} \right]_+ \\ &\quad - [\alpha_{\max,h}]_+ \end{aligned}$$

where

$$\alpha_{\max,h} = \lambda^*_{\max,h} + \gamma \frac{2}{U_N^2} \mathbf{1}_h^T (\sin\theta M q_G^*(t) - b_{\max,h}).$$

Thanks to the fact that $|a_+ - b_+| \leq |a - b|$ we can write that

$$\begin{aligned} &|\lambda_{\min,h}(t+1) - \lambda^*_{\min,h}| \\ &\qquad \leq |\lambda_{\min,h}(t) - \lambda^*_{\min,h} - \gamma \frac{2}{U_N^2} \sin\theta \mathbf{1}_h^T M x(t)| \end{aligned}$$

and, in turn,

$$\begin{aligned}&\|\lambda_{\min}(t+1)-\lambda^*_{\min}\|^2\\&\quad\le\left\|\lambda_{\min}(t)-\lambda^*_{\min}+\gamma\frac{2}{U_N^2}\mathbf{1}_h\mathbf{1}_h^T(-\sin\theta M)x(t)\right\|^2\end{aligned}$$

Similarly we have

$$\begin{aligned}&\|\lambda_{\max}(t+1)-\lambda^*_{\max}\|^2\\&\quad\le\left\|\lambda_{\max}(t)-\lambda^*_{\max}+\gamma\frac{2}{U_N^2}\mathbf{1}_h\mathbf{1}_h^T\sin\theta Mx(t)\right\|^2\end{aligned}$$

Now observe that Assumption 4.5.5 implies there exists almost surely a positive integer T such that any node has performed an update within the window $[0,T]$. Moreover observe from (4.35) that $x_h(t+1)=-\mathbf{1}_h^T[-\sin\theta I\ \sin\theta I]^Ty(t+1)$. It follows that, for $t\ge T$, $x(t)=-M^{-1}\Phi^Ty(t)$. Hence we can write

$$\|y(t+1)\|^2\le\left\|\left(I-\gamma\frac{2}{U_N^2}D_h\Phi M^{-1}\Phi^T\right)y(t)\right\|^2 \tag{4.36}$$

where

$$D_h=\begin{bmatrix}\mathbf{1}_h\mathbf{1}_h^T & 0\\ 0 & \mathbf{1}_h\mathbf{1}_h^T\end{bmatrix}.$$

Let $P_h=I-\gamma\frac{2}{U_N^2}D_h\Phi M^{-1}\Phi^T$. Consider the evolution of the quantity $\mathbb{E}\left[\|y(t)\|^2\right]$. We have

$$\begin{aligned}\mathbb{E}\left[\|y(t+1)\|^2\right]&\le\mathbb{E}\left[y(t)^TP_h^TP_hy(t)\right]\\&=\text{trace}\mathbb{E}\left[y(t)^TP_h^TP_hy(t)\right]\\&=\text{trace}\left\{\mathbb{E}\left[P_h^TP_hy(t)y(t)^T\right]\right\}\end{aligned}$$

Let $\chi=\mathbb{E}\left[P_h^TP_h\right]$. Observe that

$$\chi=I-\frac{4\gamma}{(m-1)U_N^2}\Phi M^{-1}\Phi^T+\frac{4\gamma^2}{(m-1)U_N^4}(\Phi M^{-1}\Phi^T)^2$$

Let us adopt the decomposition $y=y_\perp+y_\parallel$ where $y_\perp\perp\ker\Phi^T$ and $y_\parallel\in\ker\Phi^T$. It

follows

$$\begin{aligned}\mathbb{E}\left[\|y(t+1)\|^2\right] &\leq \text{trace}\left\{\mathbb{E}\left[\chi y(t)y(t)^T\right]\right\} \\ &= \mathbb{E}\left[y^T(t)\chi y(t)\right] \\ &\leq \omega^2\mathbb{E}\left[\|y_\perp(t)\|^2\right] + \mathbb{E}\left[\|y_\parallel(t)\|^2\right]\end{aligned} \tag{4.37}$$

where

$$\omega = \max_{v\perp\ker\Phi^T,\|v\|=1} \left\|v^T\chi v\right\|.$$

It can be checked that $\omega < 1$ if

$$\gamma \leq \frac{U_N^2}{\rho(\Phi M^{-1}\Phi^T)}. \tag{4.38}$$

The condition $\omega < 1$ implies that $\mathbb{E}[\|y(t+1)\|] < \mathbb{E}[\|y(t)\|]$. Hence $\mathbb{E}[\|y(t)\|]$ is a decreasing sequence that lives in the compact set $[0, \mathbb{E}[\|y(0)\|]$, which, then, admits a limit, i.e.,

$$\lim_{t\to\infty} \mathbb{E}[\|y(t)\|] = c.$$

If $c = 0$, then y goes to 0 which implies that also x goes to zero and then q_G tends to the optimal primal solution q_G^*. Otherwise if $c \neq 0$, we have from (4.37) that

$$\lim_{t\to\infty} \mathbb{E}[\|y(t)\|] = \mathbb{E}[\|y^\infty\|] = \mathbb{E}[\|y_\parallel^\infty\|] = c.$$

This implies that the trajectory $\nu(t)$ tends to the set

$$\mathcal{S} = \{\nu^* + v, v \in \ker\Phi^T, \|\nu^* + v\| = c\}$$

This implies the boundedness of the sequence $\nu(t)$ and the convergence of $q_G(t)$ to q_G^*, being from (4.33)

$$\begin{aligned}\lim_{t\to\infty} q_G(t) &= \lim_{t\to\infty} -M^{-1}Nq_L - M^{-1}\Phi^T\nu(t) \\ &= -M^{-1}Nq_L - M^{-1}\Phi^T\nu^*.\end{aligned}$$

We now briefly repeat similar steps for the case where only power constraints are

taken into account. In this case we have from (4.31) that

$$\begin{aligned}
&\mu_{\min,h}(t+1) - \mu^*_{\min,h} = \\
&= \left[\mu_{\min,h}(t) - \mu^*_{\min,h} - \gamma\frac{2}{U_N^2}(q_h(t) - q_h^*) + \mu^*_{\min,h}\right]_+ \\
&\quad - \left[\mu^*_{\min,h}\right]_+ \\
&= \left[\mu_{\min,h}(t) - \mu^*_{\min,h} - \gamma\frac{2}{U_N^2}x_h(t) + \mu^*_{\min,h}\right]_+ \\
&\quad - \left[\mu^*_{\min,h}\right]_+
\end{aligned}$$

From the expression for x_h in (4.35), it follows that

$$\begin{aligned}
&\|\mu_{\min}(t+1) - \mu^*_{\min}\|^2 \\
&\leq \left\|\mu_{\min}(t) - \mu^*_{\min} - \gamma\frac{2}{U_N^2}\mathbf{1}_h\mathbf{1}_h^T M^{-1}\Phi^T y(t)\right\|^2
\end{aligned}$$

Reasoning similarly we obtain that

$$\begin{aligned}
&\|\mu_{\max}(t+1) - \mu^*_{\max}\|^2 \\
&\leq \left\|\mu_{\max}(t) - \mu^*_{\max} - \gamma\frac{2}{U_N^2}\mathbf{1}_h\mathbf{1}_h^T M^{-1}\Phi^T y(t)\right\|^2
\end{aligned}$$

Recalling that, in this case, $\Phi = [-I \;\; I]^T$, from the above inequalities we get that, as in (4.36), that

$$\|y(t+1)\|^2 \leq \left\|(I - \gamma\frac{2}{U_N^2}D_h\Phi M^{-1}\Phi^T)y(t)\right\|^2.$$

Again the convergence of $q_G(t)$ to the optimal solution q_G^* is guaranteed if condition (4.38) is satisfied. ■

4.6 Simulations

The algorithm has been tested on the testbed IEEE 37 Kersting (2001), which is an actual portion of 4.8kV power distribution network located in California. The load buses

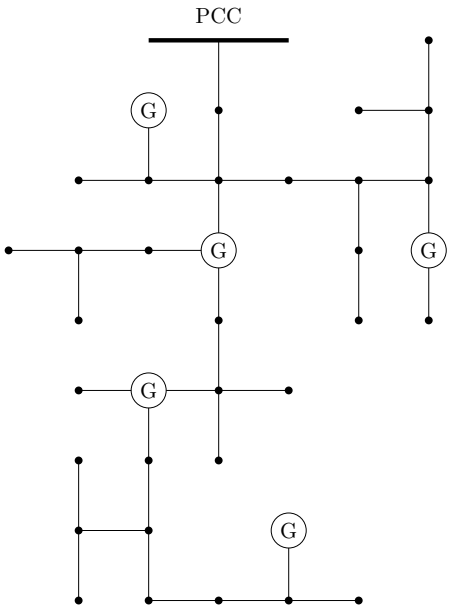


Figure 4.3: Schematic representation of the IEEE 37 test feeder Kersting (2001), where 5 microgenerators have been deployed.

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 Kersting (2001) for the testbed data). 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 Ω/km to 1.305 Ω/km). However, the inductance/resistance ratio exhibits a smaller variation, ranging from $X/R = 0.5$ to 0.67. This justifies Assumption 2.8.1, in which we claimed that $\angle z_e$ can be considered constant across the network. We considered the scenario in which 5 microgenerators have been deployed in this portion of the power distribution grid (see Figure 4.3).

The lower bound for voltage magnitudes has been set to 4700 V. Both the synchronous and the asynchronous algorithm presented in Section 4.3 and 4.4 have been simulated on a nonlinear exact solver of the grid Zimmerman, Murillo-Sánchez, and Thomas (2011). The approximate model presented in Proposition 2.6.1 has not been used in these simulations, being only a tool for the design of the algorithm and for the study of the algorithm's convergence.

A time-varying profile for the loads has been generated, in order to simulate the effect of slowly varying loads (e.g. the aggregate demand of a residential neighborhood), fast changing / intermittent demands (e.g. some industrial loads).

The results of the simulation have been plotted in Figure 4.4 for the asynchronous case,

while the synchronous case has not been reported, being very similar. In order to tune the parameter γ, based on the similarity between the conditions in Propositions 4.5.6 and 4.5.7, and those in Corollaries 4.5.3 and 4.5.4, we conjecture that the bound derived in Theorem 4.5.2 is also valid in the asynchronous case. We have therefore chosen γ to be one half of such bound. The power distribution losses, the lowest voltage magnitude measured by the microgenerators, and the reactive power injection of one of the microgenerators, are reported. The dashed line represents the case in which no reactive power compensation is performed. The thick black line represents the best possible strategy that solves the ORPF problem (4.2) (computed via a numerical centralized solver that have real time access to all the grid parameters and load data). The thin red line represents the behavior of the proposed algorithm.

It can be seen that the proposed algorithm achieves practically the same performance of the centralized solver, in terms of power distribution losses. Notice however that the proposed algorithm does not have access to the demands of the loads, which are unmonitored. The agents, located only at the microgenerators, can only access their voltage measurements and share them with their neighbors. Notice moreover that, as expected for duality based methods, the voltage constraints can be momentarily violated. Therefore, in the time varying case simulated in this example, the voltage sometimes falls slightly below the prescribed threshold, when the power demand of the loads present abrupt changes. It should be remarked, however, that the extent of this constraint violation depends on the rate at which the algorithm is executed, compared with the rate of variation of loads, and on the fact that an exact (and thus aggressive) primal update step has been implemented. The same behavior cannot be observed for the power constraints, as the reactive power set-point has been saturated in order to simulate the typical implementation of power inverters, which cannot accept set-point references that exceed their rated power. Notice that the reactive power reference is almost constant when voltage constraints are not active (as the primal step is exact, and therefore the algorithm reaches the optimal point immediately). When the constraints are active, the evolution depends instead on the update of the Lagrange multipliers.

Finally, in Figure 4.5, we investigated the robustness of the algorithm with respect to possible larger variations of the R/X ratio of the power lines. The same testbed has been modified in order to have R/X ratios ranging from 0.36 to 2.6. Despite the fact that Assumption 2.8.1 is needed for the technical results of the paper, simulations show how the effect on the closed loop behavior of the controlled systems is minimal. Intuitively, this is due to the feedback nature of the control strategy: as the violation of the constraints is integrated in the feedback loop (see Figure 4.2), this violation is

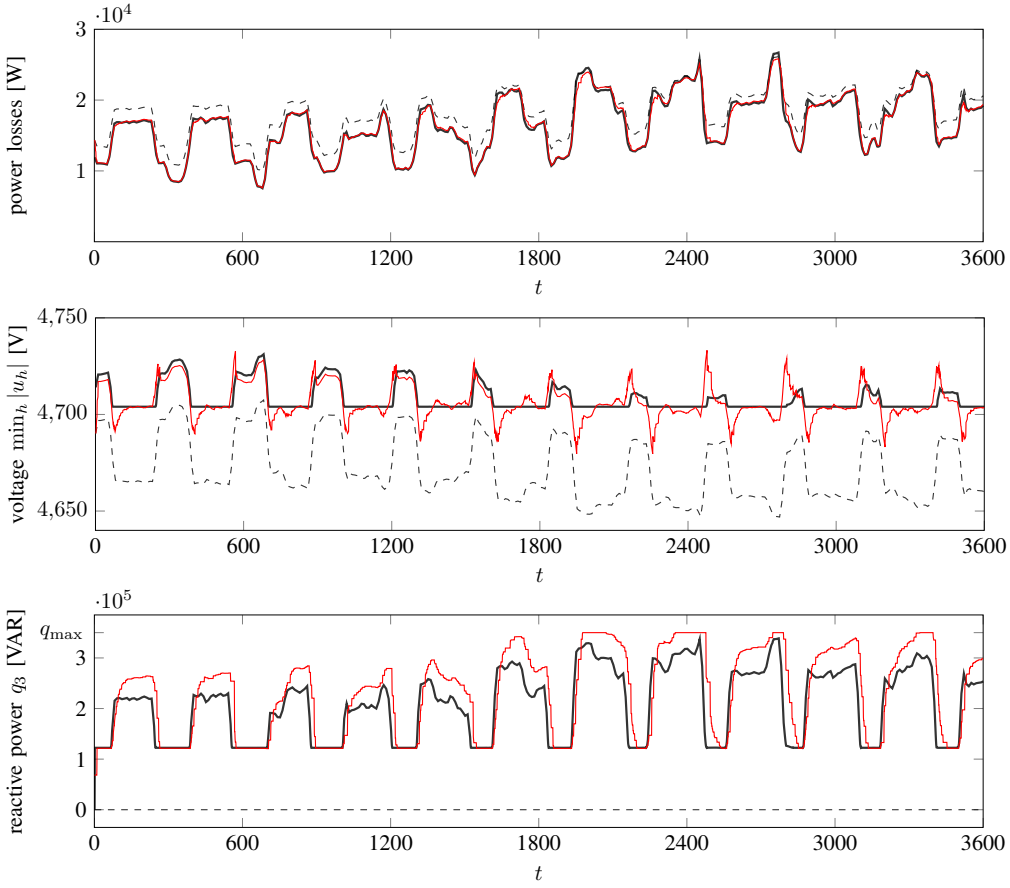


Figure 4.4: Power distribution losses, the lowest measured voltages, and the reactive power setpoint of generator 3, have been plotted for the following cases: when no reactive power compensation is performed (dashed line), when an ideal centralized numerical controller commands the microgenerators (thick black line), and for the proposed algorithm, where microgenerators are commanded via a feedback law from the voltage measurements (thin red line).

guaranteed to go to zero, as long as the closed loop system is stable.

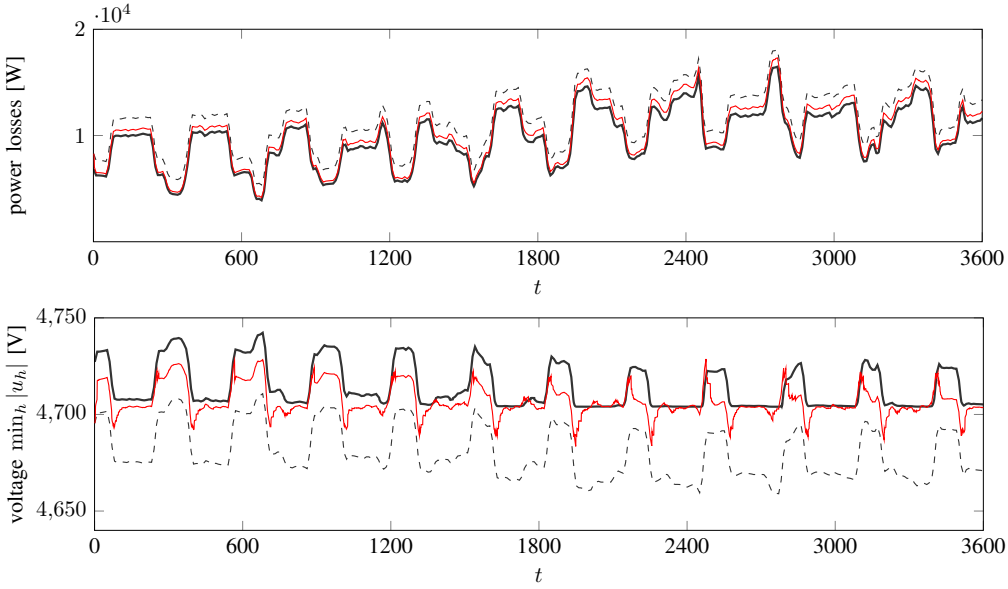


Figure 4.5: The same simulation of Figure 4.4 has been repeated for a larger variation of the inductance/resistance ratio of the lines, from $X/R = 0.36$ to 2.6.

4.7 Chapter conclusions

In this chapter we proposed a distributed control law for optimal reactive power flow in a smart power distribution grid, based on a *feedback strategy*. Such a strategy requires the *interleaving of actuation and sensing*, and therefore the control action (the reactive power injections $q_h, h \in \mathcal{C}\backslash\{0\}$) is a function of the real time measurements (the voltages $u_h, h \in \mathcal{C}$). According to this interpretation, the active power injections in the grid ($p_h, h \in \mathcal{V}$) and the reactive power injection of the loads ($q_h, h \in \mathcal{V}\backslash\mathcal{C}$) can be considered as *disturbances* for the control system. As explained in Remark 4.3.3, these quantities do not need to be known to the controller, and the agents are implicitly inferring them from the measurements. It is also well known that the presence of feedback in the control action makes the closed loop behavior of the system less sensitive to model uncertainties, as shown in the simulations. These features differentiate the proposed algorithm from most of the ORPF algorithms available in the power system literature, with the exception of some works, like Turitsyn et al. (2011), where however the feedback is only local, with no communication between the agents, and of Tenti, Costabeber, Mattavelli, and Trombetti (2012) and Bolognani and Zampieri (2013). Moreover, in the proposed feedback strategy, the controller does not need to solve any model of the grid in order to find the optimal solution. The computational effort required for the execution of the proposed algorithm is therefore minimal. These features are extremely interesting for the scenario of power

distribution networks, where real time measurement of the loads is usually not available, and the grid parameters are partially unknown.

While a feedback approach to the ORPF problem is a recent approach, similar methodologies have been used to solve other tasks in the operation of power grids. In particular, in order to achieve realtime power balance of demand and supply, synchronous generators are generally provided with a local feedback control that adjusts the input mechanical power according to frequency deviation measurements (see Chandorkar et al. (1993); De Brabandere et al. (2007)). By adding a communication channel (a cyber layer) that enables coordination among the agents, it is possible to drive the system to the configuration of minimum generation costs Barklund, Pogaku, Prodanovic, Hernandez-Aramburo, and Green (2008). Notice that, in this scenario, generators do not have access to the aggregate active power demand of the loads, but infer it from the purely local frequency measurements. In this sense, this example share some qualitative similarities with the original approach presented in this paper.

As suggested in Wang and Elia (2011), a control-theoretic approach to optimization problems (including ORPF) enables a number of analyses on the performance of the closed loop system that are generally overlooked. Examples are L_2-like metrics for the resulting losses in a time-varying scenario (e.g. the preliminary results in Bolognani, Cavraro, and Zampieri (2012)), robustness to measurement noise and parametric uncertainty, stability margin against communication delays. These analyses, still not investigated, are also of interest for the design of the cyber architecture, because they can provide specifications for the communication channels, communication protocols, and computational resources that need to be deployed in a smart distribution grid.

5

A distributed feedback algorithm for the Optimal Power Flow (OPF) problem

In section 3.1 we have seen that the ultimate goal of the optimal power flow (OPF) problem is to find an operating point of the power system that minimizes a cost function while satisfying the power demand and some operative constraints, such as bus voltage limits or generators generation limits.

The algorithm we propose in this chapter, extends the approach of Bolognani et al. (2015) to the OPF problem, and thus can be considered as a feedback control strategy. In the OPF problem we consider, the goal is to minimize the global generation cost by controlling the amount of active power injected in the grid by the generators. The active powers are subject to box constraints modeling the generation capability of each generator, while the objective function is given by the sum of the generation cost functions associated with the generators. One of the aspects that differs our approach to the other in the recent literature is to propose a more realistic generation cost that is a linear

piecewise function, while the generation cost is usually modeled as a quadratic or a differentiable function. This choice will be motivated in the following. In our setup we consider only two types of such functions: the ones associated with the micro-generators dispersed in the grid and the one associated with the utility.

We tackle the problem via a projected gradient-based approach. Applying at each iteration a projected gradient descent step, the algorithm is shown to be provably convergent to an approximated optimal solution of the OPF problem. Furthermore, we provide a characterization of the optimal solution that can be useful to design algorithms that solve the OPF problem.

In the following, given $u, v, w \in \mathbb{R}^\ell$, with $v_h \leq w_h, h = 1, \ldots, \ell$ we define the operator $\text{proj}(u, v, w)$ as the component wise projection of u in the set

$$\left\{ x \in \mathbb{R}^\ell : v_h \leq x_h \leq w_h, h = 1, \ldots, \ell \right\},$$

i.e.,

$$\Big(\text{proj } (u, v, w)\Big)_h = \begin{cases} u_h & \text{if } v_h \leq u_h \leq w_h \\ v_h & \text{if } u_h < v_h \\ w_h & \text{if } u_h > w_h \end{cases} \tag{5.1}$$

5.1 Preliminary assumptions

We assume that Assumption 2.8.1 holds, in order to develop our control strategy. This assumption is necessary for the algorithm design. However, simulations in Section 5.5 will be made over a standard testbed, that does not satisfy Assumption 2.8.1, showing the effectiveness of the algorithm also in the realistic scenario.

We assume that each agent (a microgenerator or the PCC) is provided with some computational capability, and with some sensing capability, in the form of a phasor measurement unit (PMU).

We assume that there is a communication link between the PCC and every agent, as showed by Figure 5.1. The PCC in fact, in order to perform the algorithm, will broadcast some information, and the agents are assumed to be able to receive this information.

In this chapter, agents are modeled as *prosumers* (Grijalva and Tariq (2011)), i.e. nodes having both generation capability and load, e.g. a house with a photovoltaic panel in the roof. As a consequence the prosumer j

- generates the power $p_j^g (> 0)$ (is a *producer*).
- requires the power $p_j^r (< 0)$ (is a *customer*);

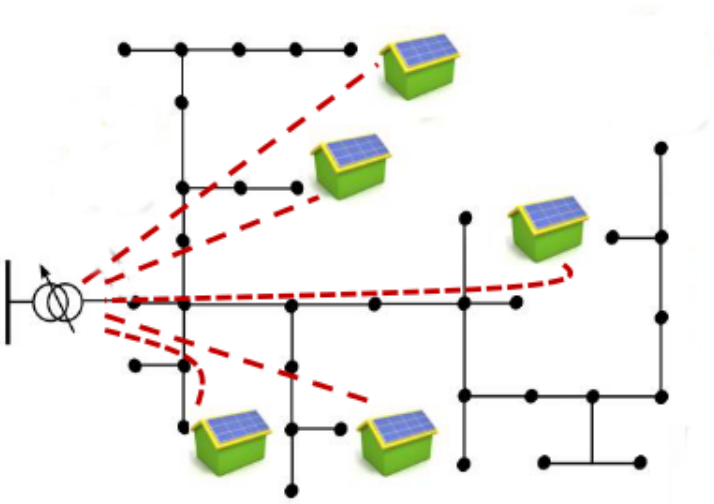


Figure 5.1: The red dotted lines represent the communication links among agents in the smart grid. The communication architecture has a *star topology*.

If the generation capability p_j^g is lower than the power required p_j^r, prosumer j will behave as an uncontrollable load that absorbs the difference $p_j = p_j^p - p_j^r < 0$. Otherwise it will behave like a generator, whose generation capability is limited to $p_{j,max} = p_j^g - p_j^r > 0$, and thus it can also actuate the system, by commanding a set point for the amount of active power to be injected.

5.2 Optimal power flow problem

The goal is to design a distributed control algorithm that leads to the cost minimization of the power supplied to the loads, that require s_L. Formally the problem we are interested into can be stated as

$$\min_{p_j, j \in \mathcal{C}} f = \sum_{j=2}^{m} f_j(p_j) + f_1(p_1) \tag{5.2a}$$

$$\text{subject to} \quad p_1 = -(\mathbf{1}^T p_G + \mathbf{1}^T p_L) + \ell(s_G, s_L, U_N) \tag{5.2b}$$

$$p_{j,min} \leq p_j \leq p_{j,max} \qquad j \in \mathcal{C} \tag{5.2c}$$

$$u_1 = U_0 \tag{5.2d}$$

where

- constraint in (5.2b) models the active power conservation in the grid, being $\ell(s_G, s_L, U_N)$ the active power losses in the grid;

- constraints in (5.2c) model the prosumers generation capabilities. In the case in which the prosumer j is behaving like a load, $p_{j,min} = p_{j,max} < 0$. If, otherwise, prosumer j is behaving like a generator, $p_{j,min} = 0, p_{j,max} > 0$.
- the objective function f is the sum of the cost of the power produced by the utility and injected into the microgrid through the PCC ($f_1(p_1)$), and of the microgenerators' payments for the power that they inject.

Problem is a particular case of Problem 3.8, in which

1. the control variables are just the active powers p_G, and not the whole complex powers s_G;
2. the branch current magnitude limitations and the voltage magnitude constraints are not taken into account.

Observe that (5.2b) provides a expression of p_1 as a function of s_G, s_L, U_N. In the following p_1 or $p_1(p_G)$ stand for $p_1(s_G, s_L, U_N)$), since q_G, s_L, U_N are fixed.

In several studies of the OPF problem and in recent papers (Devane and Lestas (2013); Lavaei and Low (2012); Mallada and Tang (2013)), the cost functions are typically chosen quadratic (e.g $f_j(p_j) = \alpha_j^2 p_j^2 + \alpha_j^1 p_j + \alpha_j^0$), or anyway convex, continuous and differentiable, because they derive from models of the energy production costs of classical power plants and generators. However these choices do not capture the features of scenarios with high penetration of distributed renewable energy resources, such as photovoltaic panels, that produce energy at zero cost, but whose owners receive a reward from the utility for the energy they inject into the grid. In this new scenario the cost function does not model anymore the physical cost of energy production, but rather the remuneration that the prosumers receive when they are able to produce. If the prosumer is behaving like a generator, the reward is proportional to the quantity of energy injected into the grid, with a proportionality constant that depends on the contractual agreement among the prosumer and the utility. Otherwise, if the prosumer is behaving like a load, it will receive no remuneration, rather it has to pay the power consumed, whose price is proportional to the power required, with a proportionality constant that depends on the contractual agreement among the prosumer and the utility. As a result, the utility must sustain a negative price, i.e. it earns money. The model of the remuneration (see Figure 5.2) comes straightforward from the previous reasoning, and it is

$$f_j(p_j) = \begin{cases} c_G^j p_j & p_j \geq 0 \\ c_L^j p_j & p_j < 0 \end{cases} \qquad \forall j \in \mathcal{C} \tag{5.3}$$

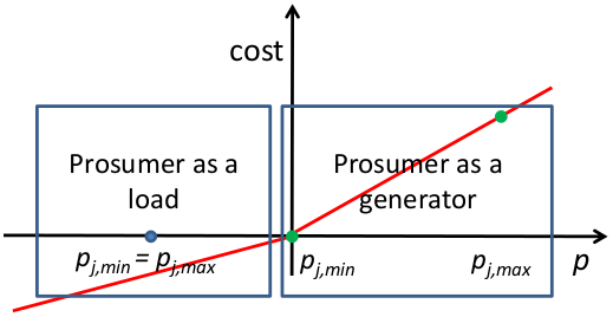


Figure 5.2: Power cost function associated with a prosumer.

We assume for simplicity that all the agents are paid in the same way, or otherwise that pay the energy in the same way, that is

$$c_G^j = c_G, c_L^j = c_L \quad \forall j \in \mathcal{C}$$

For what concern the PCC, if it is injecting the power $p_1 \geq 0$, the cost sustained by the utility is the cost of the generation of p_1 in the power plants, modeled as $c_1 p_1$. If instead the PCC is absorbing the power $p_1 \leq 0$, it means that the power utilized in the grid comes just from the microgenerators and that there is a reverse power flow. In this framework, for the sake of simplicity, we assume that the power in excess that flows from the grid to the PCC is dissipated. As a result, the cost function associated with p_1 must be zero when $p_1 \leq 0$, and we model $f_1(p_1)$ as

$$f_1(p_1) = \begin{cases} c_1 p_1 & p_1 \geq 0 \\ 0 & p_1 < 0 \end{cases} \tag{5.4}$$

The cost c_1 is in general different from c_G. The cost c_G is fixed and dependent on the particular contract between the prosumers and the utility. The cost c_1 instead may vary from one day to another or during the same day: energy can be cheaper when there is a surplus of production in the power plants, or it can be more expansive when the demand is higher.

However our treatment can be easily generalized to the case in which the reward is different from one prosumer to another and in which either the reward or the power generation cost at the PCC are modeled with different functions. In the following we will assume, without loss of generality, that the generation capability of the prosumers is greater than their load demand, and thus that the active power they inject is controllable. Therefore, for every prosumer j, $p_{j,min} = 0, p_{j,max} > 0$.

As it will be clear in the following, the utility, differently from the classical centralized schemes, does not impose to the agents the active power reference points (and even does not know them). But it just provides them the information needed to compute and to actuate (all by their own) the optimal power input.

Based on (2.32) and by exploiting the properties of X we can express the active power losses in the grid as

$$\ell(s_G, s_L, U_N) = \left(\begin{bmatrix} p_G^T & p_L^T \end{bmatrix} \begin{bmatrix} M & N \\ N^T & Q \end{bmatrix} \begin{bmatrix} p_G \\ p_L \end{bmatrix} + \begin{bmatrix} q_G^T & q_L^T \end{bmatrix} \begin{bmatrix} M & N \\ N^T & Q \end{bmatrix} \begin{bmatrix} q_G \\ q_L \end{bmatrix} \right) \frac{\cos(\theta)}{U_N^2} + o\left(\frac{1}{U_N^2}\right) \tag{5.5}$$

Defining

$$\tilde{\ell}(s_G, s_L, U_N) = \left(\begin{bmatrix} p_G^T & p_L^T \end{bmatrix} \begin{bmatrix} M & N \\ N^T & Q \end{bmatrix} \begin{bmatrix} p_G \\ p_L \end{bmatrix} + \begin{bmatrix} q_G^T & q_L^T \end{bmatrix} \begin{bmatrix} M & N \\ N^T & Q \end{bmatrix} \begin{bmatrix} q_G \\ q_L \end{bmatrix} \right) \frac{\cos(\theta)}{U_N^2} \tag{5.6}$$

we can rewrite (5.2b) as

$$\begin{aligned} p_1 + \mathbf{1}^T(p_G + p_L) &= \ell(s_G, s_L, U_N) \\ &\simeq \tilde{\ell}(s_G, s_L, U_N). \end{aligned}$$

Hence we can approximate (or convexify) problem (5.2) with the following one

$$\min_{p_j, j \in \mathcal{C}} \quad \hat{f} \tag{5.7a}$$

$$\text{subject to} \quad \hat{p}_1 = -(\mathbf{1}^T p_G + \mathbf{1}^T p_L) + \tilde{\ell}(s_G, s_L, U_N) \tag{5.7b}$$

$$0 \le p_j \le p_{j,max} \qquad j \in \mathcal{C} \tag{5.7c}$$

$$u_1 = U_0 \tag{5.7d}$$

where

$$\hat{f} = c_G \mathbf{1}^T p_G + f_1(-\mathbf{1}^T p_G - \mathbf{1}^T p_L + \tilde{\ell}(s_G, s_L, U_N))$$

and where $\hat{p}_1$ is an approximation of p_1 that differs from it just up to infinitesimal terms. It is convenient to express $\hat{f}$ in the following way:

$$\hat{f}(p_G) = \begin{cases} \hat{f}^+(p_G) & \text{if } p_G \in \mathcal{P}^+ \\ \hat{f}^-(p_G) & \text{if } p_G \in \mathcal{P}^- \end{cases} \tag{5.8}$$

where

$$\begin{aligned}
\hat{f}^+(p_G) &= (c_G - c_1)\mathbf{1}^T p_G + c_1 \tilde{\ell}(s_G, s_L, U_N) \\
\hat{f}^-(p_G) &= c_G \mathbf{1}^T p_G \\
\mathcal{P}^+ &= \{p_G : -\mathbf{1}^T p_G - \mathbf{1}^T p_L + \tilde{\ell}(s_G, s_L, U_N) > 0\} \\
\mathcal{P}^- &= \{p_G : -\mathbf{1}^T p_G - \mathbf{1}^T p_L + \tilde{\ell}(s_G, s_L, U_N) \leq 0\}
\end{aligned}$$

i.e., $\mathcal{P}^+$ is the set of all the p_G such that $\hat{p}_1(p_G) > 0$, while $\mathcal{P}^-$ is the set of all the p_G such that $\hat{p}_1(p_G) < 0$. We furthermore define

$$\mathcal{P}^0 = \{p_G : -\mathbf{1}^T p_G - \mathbf{1}^T p_L + \tilde{\ell}(s_G, s_L, U_N) = 0\}$$

i.e., the set of all the p_G such that $\hat{p}_1(p_G) = 0$.

We conclude this section with a characterization of the optimal solution of (5.2).

Proposition 5.2.1. *Consider problem (5.2). If p_G^* is an optimal solution, then*

$$p_1(p_G^*) \geq 0.$$

Proof. Let p_G^* be an optimal solution such that $p_1^* = p_1(p_G^*) \leq 0$ and let $\ell_A(s_G^*, s_L^*, U_N)$ be the active power losses of the system in this configuration. We will show that there exist another configuration $p_{\tilde{G}}$ such that $f(p_{\tilde{G}}) < f(p_G^*)$ and $\tilde{p}_1 = p_1(\tilde{p}_G) \geq p_1^*$. If $p_1^* < 0$, from (5.2b), it follows that there exist at least a compensator k such that $p_k > 0$. Let now k decrease its active power injection by $-\delta_k, \delta_k > 0$. Let δ_k such that

$$p_1(p_G^*) \leq p_1(p_G^* - 1_k \delta_k) \leq 0.$$

Now let us examine the change Δi_1 in the current injected by the PCC. We have

$$\begin{aligned}
\Delta i_1 &= \mathbf{1}_1^T Y \Delta u \\
&= \mathbf{1}_1^T Y X (1_1 - \mathbf{1}_k) \frac{\delta_k}{U_N^2} + o\left(\frac{1}{U_N}\right) \\
&= \mathbf{1}_1^T (I - \mathbf{1}_1 \mathbf{1}^T)(1_1 - \mathbf{1}_k) \frac{\delta_k}{U_N^2} + o\left(\frac{1}{U_N}\right) \\
&= \frac{\delta_k}{U_N^2} + o\left(\frac{1}{U_N}\right) \simeq \frac{\delta_k}{U_N^2}
\end{aligned}$$

It follows that the active power absorption of the PCC after the injection drop δ_p changes

by

$$\Delta p_1 = \mathrm{Re}\left(U_N \Delta \xi_e\right) = U_N \frac{\delta_k}{U_N} > 0$$

that is, $p_1(p_G^* - 1_k\delta_k, q_G^*) > p_1(p_G^*)$. Finally, we point out that the new configuration, with a greater p_1 has a cost that is lower than the starting one:

$$\begin{aligned}
&\hat{f}(p_G - \mathbf{1}_k\delta_k,) - \hat{f}(p_G) = \\
&- c_G\delta_k + f_1(p_G - \mathbf{1}_k\delta_k, q_G) - f_1(p_G, q_G) \\
&\simeq -c_G\delta_k \leq 0
\end{aligned}$$

■

The same proof can be applied also if we consider problem (5.7) instead of (5.2). The above proposition implies that the optimal solution p_G belongs to $\mathcal{P}^+ \cup \mathcal{P}^0$. Moreover, using a reasoning similar to the one exploited in the above proof, we can easily prove the following

Lemma 5.2.2. *Consider problems (5.2) and (5.7), if $p_1^* = 0$, then $c_1 > c_G$.*

5.3 A distributed gradient projected descent algorithm

The algorithm we propose is based on a gradient descent strategy that we derive considering the approximated problem (5.7). Furthermore, the algorithm is distributed, in the sense that it can be performed *without a coordinating central unit* by the agents, just exploiting local measurements and information broadcasted by the PCC.

Let us compute the gradient of the cost function $\hat{f}$. Observe that

$$\frac{\partial\left(\sum_j^m f_j(p_j)\right)}{\partial p_G} = c_G\mathbf{1},$$

while, concerning $\frac{\partial f_1(p_1)}{\partial p_j}$, by exploiting the chain rule we have that

$$\frac{\partial f_1(p_1)}{\partial p_G} = f_1'(p_1)\left(-\mathbf{1} + 2\frac{\cos(\theta)}{U_N^2}(Mp_G + Np_L)\right)$$

Plugging together the above expressions we get

$$\frac{\partial \hat{f}}{\partial p_G} = f_G'(p_G) + f_1'(p_1)\left(-\mathbf{1} + 2\frac{\cos(\theta)}{U_N^2}(Mp_G + Np_L)\right) \tag{5.9}$$

which, using (5.3) and (5.4), can be finally rewritten as

$$\frac{\partial \hat{f}}{\partial p_G}(p_G) = c_G \mathbf{1}, \tag{5.10}$$

if $p_G \in \mathcal{P}^-$, otherwise if $p_G \in \mathcal{P}^+$.

$$\frac{\partial \hat{f}}{\partial p_G}(p_G) = 2c_1 \frac{\cos(\theta)}{U_N^2}(Mp_G + Np_L) + (c_G - c_1)\mathbf{1} \tag{5.11}$$

Observe that , while the values U_N and θ can be assumed known a priori, the quantities $Mp_G + Np_L$ depend on all the active powers injected into the grid (also on the unmonitored active powers of the loads) and on the topology of the grid. However, by exploiting again (2.32), we have that

$$u_G = \frac{e^{i\theta}}{U_N}\begin{bmatrix} M & N \end{bmatrix}\begin{bmatrix} p_G - iq_G \\ p_L - iq_L \end{bmatrix} + \mathbf{1}U_N + o\left(\frac{1}{U_N}\right)$$

from which it follows that

$$\mathrm{Re}(e^{-i\theta}(u_G - \mathbf{1}U_N)) = \frac{Mp_G + Np_L}{U_N} + o\left(\frac{1}{U_N}\right) \tag{5.12}$$

and then

$$\frac{\partial \hat{f}}{\partial p_G}(p_G) \simeq f_G'(p_G) + f_1'(p_1)\left(-\mathbf{1} + 2\frac{\cos(\theta)}{U_N}\,\mathrm{Re}(e^{-i\theta}(u_G - \mathbf{1}U_N))\right) \tag{5.13}$$

where in this last expression we are neglecting the terms that vanish to zero for large U_N. It turns out that the gradient of $\hat{f}$ can be computed only by *local voltage measurements*. Indeed $\forall k \in \mathcal{C}$ we have that

$$\Big(\mathrm{Re}(e^{-i\theta}(u_G - \mathbf{1}U_N))\Big)_k = |u_k|\cos(\angle u_k - \theta) - |u_N|\cos(\angle u_N - \theta)$$

and then each compensator, in order to obtain its component of $\frac{\partial \hat{f}}{\partial p_G}$, needs only to know its own voltage, the PCC voltage and $f_1'(p_1)$. Next we formally describe the algorithm we propose in this paper. For simplicity, in the following, $\hat{g}_h$ denotes the component of the approximated gradient related to agent h.

Let γ be a positive scalar parameter. At every iteration of the algorithm, each agent $h \in \mathcal{C}\backslash\{0\}$ executes the following operations in order:

1. senses the system obtaining its voltage phasorial measurement u_h;
2. receives the PCC voltage phasorial measurement $u_1 = U_N$, the PCC active power injected p_1 and the cost coefficient c_1;
3. computes the approximated gradient direction

$$\hat{g}_h = f_h'(p_h) + f_1'(p_1)\Big(-1 + 2\frac{\cos\theta}{U_N}\Big(|u_k|\cos(\angle u_k - \theta) - |u_N|\cos(\angle u_N - \theta)\Big)\Big) \tag{5.14}$$

4. computes the active power to be injected in the grid performing the following gradient descent steps

$$p_h \leftarrow p_h - \gamma \hat{g}_h \tag{5.15}$$

5. projects p_h into the feasible region and actuates the projected values

$$p_h \leftarrow \text{proj}(p_h, 0, p_h^M) \tag{5.16}$$

Based on the above description, it is clear what is the feedback scheme that underlies the procedure we propose: during each iteration each agent senses the grid, communicates with the PCC, computes the power set-point and then actuates it.

5.4 Convergence analysis

In this section we consider the gradient projected descent of $\hat{f}$

$$p_G(t+1) = \text{proj}\Big(p_G(t) - \gamma\frac{\partial \hat{f}}{\partial p_G}, 0, p_G^M\Big) \tag{5.17}$$

in spite of its approximated version given in (5.14), (5.15) and (5.16), that we perform in practice.

We consider two different scenarios:

1. the one in which $p_G^* \in \mathcal{P}^+$;
2. the one in which $p_G^* \in \mathcal{P}^0$, that is the minimum argument of $\hat{f}$ is such that $p_1(p_G^*) = 0$;

We were able to provide a formal proof of convergence of (5.17) for the first scenario. Concerning the third scenario, we can prove the convergence of the continuous-time

version of (5.17) (which corresponds to have a γ which tends to zero) by resorting to the tools of sliding mode control. The crucial point is to theoretically quantify the difference among the continuous-time trajectory and the discrete-time trajectory. These considerations are summarized in the following propositions.

Proposition 5.4.1. *Consider the optimization problem (5.7) and the dynamic system described by the update equation (5.17). Let p_G^* be the optimal configuration and assume that $p_G^* \in \mathcal{P}^+$ and $\frac{\partial \hat{f}^+}{\partial p_G}(p_G^*) = 0$. Then if*

$$\gamma \leq \frac{U_N^2}{\cos\theta \rho(M) c_1}.$$

the trajectory $t \to p_G(t)$ converges to the optimal value p_G^*, where $\rho(M)$ is the spectral radius of M

Proposition 5.4.2. *Consider the optimization problem (5.7) and the continuous-time version of the dynamic system described by the update equation (5.17). Let p_G^* be the optimal configuration and suppose that $p_G^* \in \mathcal{P}^0$. Then the continuous-time trajectory $t \to p_G(t)$ converges to the optimal value p_G^*.*

Proof of Proposition 5.4.1. Let $\mathcal{B}$ be the feasible set, that is

$$\mathcal{B} = \{p_G \in \mathbb{R}^m : 0 \leq p_j \leq p_{j,max}\}.$$

In this scenario, $p_G \in \mathcal{P}^+ \cap \mathcal{B}$. Being p_G^* the solution of (5.7) and then a fixed point for (5.17), it satisfies

$$\begin{aligned} p_G^* = \Bigg[\left(I - \frac{2\gamma\cos(\theta)c_1}{U_N^2} M \right) p_G^* + \\ - \frac{2\gamma\cos(\theta)}{U_N^2} N p_L - \gamma(c_G - c_1)\mathbf{1} \Bigg]_{\mathcal{P}^+ \cap \mathcal{B}} \end{aligned} \tag{5.18}$$

where $[\cdot]_{\mathcal{P}^+}$ is the projection into $\mathcal{P}^+ \cap \mathcal{B}$. Notice that the gradient of $\hat{f}^+$ is a Lipschitz continuous function, with Lipschitz constant $K = \frac{2c_1\cos\theta}{U_N^2}\rho(M)$. In fact

$$\left\| \frac{\partial \hat{f}^+}{\partial p_G}(u) - \frac{\partial \hat{f}^+}{\partial p_G}(v) \right\| = \left\| \frac{2c_1\cos\theta}{U_N^2} M(u-v) \right\| \leq \frac{2c_1\cos\theta}{U_N^2} \rho(M) \left\| u - v \right\|$$

where $\rho(M)$ is the spectral radius of M. Let us define $d(t) = p_G(t) - p_G^*$, and let $d_{\|}$ an $d_{\perp}$ be the components of d parallel or orthogonal to $\mathbf{1}$, respectively. Let us perform a projected gradient descent of $\hat{f}$ with

$$\gamma \leq \frac{U_N^2}{2c_1 \cos\theta \rho(M)} \tag{5.19}$$

If $p_G(t) \in \mathcal{P}^+$ then

1. the distance among p_G and p_G^* *always* decreases, that is

$$\|d(t+1)\| \leq \|d(t)\| \tag{5.20}$$

   In fact

$$\begin{aligned} \|d(t+1)\| &= \|p_G(t+1) - p_G^*\| \\ &= \|p_G(t+1) - [p_G^*]_{\mathcal{P}^+ \cap \mathcal{B}}\| \leq \|d(t)\| \end{aligned}$$

   where we exploit (5.11), (5.18) and the fact that the projection is a non expansive map, that is

$$\|[x]_{\mathcal{P}^+ \cap \mathcal{B}} - [y]_{\mathcal{P}^+ \cap \mathcal{B}}\| \leq \|x - y\|;$$

2. if $p_G(t+1) \in \mathcal{P}^+$, then $f(p_G(t+1)) \leq f(p_G(t))$ (it comes from Proposition 3.3 in Bertsekas and Tsitsiklis (1997));

3. if $p_G(t+1) \in \mathcal{P}^-$, we cannot know if $f(p_G(t+1)) \leq f(p_G(t))$.

   Otherwise, if $p_G(t) \in \mathcal{P}^-$ then

4. if $p_G(t+1) \in \mathcal{P}^-$, then $\|d(t+1)\| \leq \|d(t)\|$ and $f(p_G(t+1)) \leq f(p_G(t))$. In fact, it is trivial to see that $\|d(t+1)_{\|}\| \leq \|d(t)_{\|}\|$ while $\|d(t+1)_{\perp}\| = \|d(t)_{\perp}\|$.

5. if $p_G(t+1) \in \mathcal{P}^+$, then we cannot know if $\|d(t+1)\| \leq \|d(t)\|$ and if $f(p_G(t+1)) \leq f(p_G(t))$. We point out that this is the only situation in which d could *increase*;

From the above considerations, if the trajectory lies always in $\mathcal{P}^+$, equation (5.19) guarantees the convergence of the algorithm. Furthermore, if $\exists\, T : \{p_G : \|p_G - p_G^*\| \leq d(T)\} \subset \mathcal{P}^+$, then from 1) and 2) it follows the convergence of the algorithm. It is clear that, in principle, the only possible case in which the algorithm does not converge, is the one in which there is a continuing sequence of crossing of $\mathcal{P}^0$, from $\mathcal{P}^-$ to $\mathcal{P}^+$, because, we point it out again, this is the only situation in which d can increase. That would be a motion with the following characteristics:

(i) when $p_G \in \mathcal{P}^+$, the trajectory of p_G approaches $\mathcal{P}^-$, always diminishing however the distance to p_G^* (due to 1)), until p_G crosses $\mathcal{P}^0$. The only condition that make it happens is that

$$\langle \frac{\partial \hat{f}}{\partial p_G}, \mathbf{1} \rangle \leq 0 \tag{5.21}$$

that is

$$\begin{aligned} \mathbf{1}^T \left(2c_1 \frac{\cos(\theta)}{U_N^2} (Mp_G + Np_L) + (c_G - c_1)\mathbf{1} \right) &= (c_G - c_1)\mathbf{1}^T\mathbf{1} + o\left(\frac{1}{U_N}\right) \\ &\simeq (c_G - c_1)m \leq 0. \end{aligned}$$

Thus, this type of trajectory is possible only when $c_1 > c_G$. If otherwise $c_G > c_1$, then the update of p_G does not point $\mathcal{P}^-$ and the trajectory will evolve only in $\mathcal{P}^+$ and (5.19) guarantees the convergence of the algorithm.

(ii) when $p_G \in \mathcal{P}^-$, then the trajectory of p_G approaches $\mathcal{P}^+$, always diminishing however the distance to p_G^* (due to 4)), until p_G crosses $\mathcal{P}^0$, producing an increasing in the distance among p_G and p_G^*;

(iii) points (i) and (ii) repeat continuously, and it is just the transition from $\mathcal{P}^-$ to $\mathcal{P}^+$ that makes the distance increase and the algorithm not to converge.

Now we will show that, if (5.19) holds the former motion cannot last forever. First of all, exploiting (5.6) and being all $p_G \in \mathcal{P}^0$ such that

$$\begin{aligned} 0 &= -\mathbf{1}^T p_G - \mathbf{1}^T p_L + \ell(s_G, s_L, U_N) \\ &= -\mathbf{1}^T p_G - \mathbf{1}^T p_L + o\left(\frac{1}{U_N}\right) \end{aligned}$$

we approximate $\mathcal{P}^0$ with the set of all p_G such that $\mathbf{1}^T p_G + \mathbf{1}^T p_L = 0$. Suppose now that a trajectory like the one described above takes place. Consider the evolution of $d_\perp(t)$. We have that $d_\perp(t) = (I - \mathbf{1}\mathbf{1}^T/m)d(t) = P_\perp d(t)$, where $P_\perp$ is the projection matrix onto the space orthogonal to $\mathbf{1}$. If we define the variety

$$\mathcal{V} = \{p_G : p_G = p_G^* + v, v \in (\ker \mathbf{1})^\perp\}$$

then $d_\perp$ represents the distance among p_G and $\mathcal{V}$.

1. if $p_G \in \mathcal{P}^+$, then (5.20) and (5.21) force $d_\perp$ to decrease. The condition is well depicted in Figure 1.

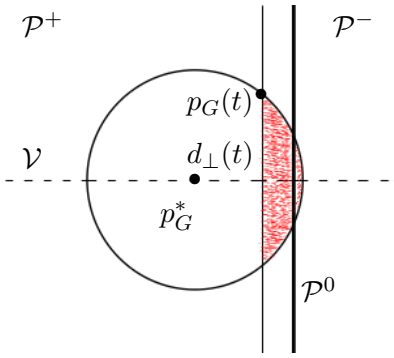


Figure 5.3: Here we are in the case in which $m = 2$. Due to (5.20), $p_G(t+1)$ will be inside the circle centered in p_G^* and passing through $p_G(t)$, while due to (5.21), $p_G(t+1)$ will be on the right side of the vertical line passing through $p_G(t)$. As a consequence, $p_G(t+1)$ will lie in the red dotted region, and then $\|d_\perp(t+1)\| \leq \|d_\perp(t)\|$. The dashed line represents the variety $\mathcal{V}$.

2. if $p_G \in \mathcal{P}^-$, then until p_G remain in $\mathcal{P}^-$ then $d_\perp$ maintains its value, because

$$P_\perp \frac{\partial \hat{f}}{\partial p_G} = c_G P_\perp \mathbf{1} = 0$$

   i.e. the motion takes place in a subspace orthogonal to $\mathbf{1}$.

From the previous consideration, it turns out that the trajectory tends to the variety $\mathcal{V}$ that intersects $\mathcal{P}^0$ in $[p_G^*]_{\mathcal{P}^0}$, where $[\cdot]_{\mathcal{P}^0}$ is the projection into $\mathcal{P}^0$. Notice that the set

$$\Phi = \{p_G : \|p_G - p_G^*\| \leq \|[p_G^*]_{\mathcal{P}^0} - p_G^*\|\}$$

that represents the ball centered in p_G^* and with radius $\|[p_G^*]_{\mathcal{P}^0} - p_G^*\|$ lies inside $\mathcal{P}^+$ and it is tangent to $\mathcal{P}^0$ in $[p_G^*]_{\mathcal{P}^0}$. The decreasing of $d_\perp$ and the crosses from $\mathcal{P}^-$ to $\mathcal{P}^+$ force the trajectory to enter Φ, and then (5.20) makes p_G to converge to p_G^*. ■

Proof of Proposition 5.4.2. The solution p_G^* of (5.7) belongs to $\mathcal{P}^0$, that is the discontinuity variety of $\frac{\partial \hat{f}}{\partial p_G}$, and the minimum of $\hat{f}^+$ belongs to $\mathcal{P}^-$. This implies that, both in $\mathcal{P}^+$ and $\mathcal{P}^-$, the opposite of the gradient of $\hat{f}$ points $\mathcal{P}^0$. Then, if $p_G(t) \in \mathcal{P}^+$ and we perform the projected gradient descent of $\hat{f}^+$ ($\hat{f}(p_G)$ is equal to $\hat{f}^+(p_G)$ in $\mathcal{P}^+$), in a finite time the trajectory cross the boundary $\mathcal{P}^0$. If, otherwise, $p_G(t) \in \mathcal{P}^-$ again the projected gradient descent of $\hat{f}^-$ ($\hat{f}(p_G)$ is equal to $\hat{f}^-(p_G)$ in $\mathcal{P}^-$), will make p_G to cross $\mathcal{P}^0$. It is clear as a result that it arises a motion in which there is a continuous cross of $\mathcal{P}^0$. Consider now the update equation

$$p_G(t+1) = p_G(t) - \gamma \frac{\partial \hat{f}}{\partial p_G}(p_G(t)) \tag{5.22}$$

that is simply (5.17) without the projection. It can be interpreted as the forward-Euler discrete-time version of the continuous time update

$$\epsilon \dot{p}_G(t) = -\gamma \frac{\partial \hat{f}}{\partial p_G}(p_G(t)) \tag{5.23}$$

where ϵ is the discretization time interval. If we apply the change of timescale

$$\tau = \frac{t}{\epsilon} \Rightarrow \frac{d\tau}{d\epsilon} = \frac{1}{\epsilon}$$

we obtain

$$\dot{p}_G(\tau) = -\gamma \frac{\partial \hat{f}}{\partial p_G}(p_G(\tau)) = \gamma \varphi(p_G(\tau)) \tag{5.24}$$

where

$$\varphi(p_G) = \begin{cases} \varphi^+ = c_1 \left(2\frac{\cos(\theta)}{U_N^2}(Mp_G + Np_L) \right) + \\ \qquad +(c_G - c_1)\mathbf{1} \text{ if } p_G \in \mathcal{P}^+ \\ \varphi^- = c_G \mathbf{1} \qquad \text{if } p_G \in \mathcal{P}^- \end{cases}$$

Equation (5.24) represents the continuous time gradient descent of $\hat{f}$. If we fix a initial p_G, from the previous considerations in a finite time we reach $\mathcal{P}^0$ and then, due to the nature of $\hat{f}$ gradient, we have a sliding mode on the variety. Let us approximate, as we did in the previous proof, $\mathcal{P}^0$ with the set of all p_G such that $\mathbf{1}^T p_G + \mathbf{1}^T p_L = 0$. Following Filippov's continuation method Utkin (1978), in order to find a *equivalent velocity*, we have to find, among the convex combination of φ^+ and φ^- "near" $\mathcal{P}^0$ the one that maintains the trajectory in the variety. That is, we have to find a φ^0 such that

$$\varphi^0 = \mu\varphi^+ + (1-\mu)\varphi^-, \quad \mathbf{1}^T\varphi^0 = 0$$

The former leads to the condition

$$\mu = \frac{c_G}{c_1} + o\left(\frac{1}{U_N}\right) \simeq \frac{c_G}{c_1}$$

from which it turns out that

$$\varphi^0 = 2c_G \frac{\cos(\theta)}{U_N^2}(Mp_G + Np_L)$$

and then, for all $p_G(\tau)$ belonging to $\mathcal{P}^0$, the equivalent velocity is described by

$$\dot{p}_G(\tau) = -\gamma 2 c_G \frac{\cos(\theta)}{U_N^2}(M p_G(\tau) + N p_L) \tag{5.25}$$

describes the sliding mode on $\mathcal{P}^0$. Notice that, from Lemma 5.2.2, we have that $c_1 > c_G$ and then $0 < \mu < 1$. Now consider the discretized version of (5.25) constrained to the feasible set $\mathcal{B}$

$$p_G(t+1) = \left[p_G(\tau) - \epsilon\gamma 2 c_G \frac{\cos(\theta)}{U_N^2}(M p_G(t) + N p_L) \right]_{\mathcal{P}^0 \cap \mathcal{B}} \tag{5.26}$$

with ϵ arbitrary small, and where $[\cdot]_{\mathcal{P}^0 \cap \mathcal{B}}$ is the projection into $\mathcal{P}^0 \cap \mathcal{B}$. Let's define $d(t) = [p_G(t)]_{\mathcal{P}^0} - p^*$, i.e. $d(t)$ represents the distance among the agents state $p_G(t)$ and the solution of (5.7). The equilibrium p_G^* satisfies

$$p_G^* = \left[p_G^* - \epsilon\gamma 2 c_G \frac{\cos(\theta)}{U_N^2}(M p_G^* + N p_L) \right]_{\mathcal{P}^0 \cap \mathcal{B}} \tag{5.27}$$

We have that

$$\|d(t+1)\| \leq \left\| \left(I - \epsilon\gamma 2 c_G \frac{\cos(\theta)}{U_N^2} M \right) \right\| \|d(t)\|$$

being the projection a non expansive map ($\|[x]_{\mathcal{P}^0} - [y]_{\mathcal{P}^0}\| \leq \|x - y\|$). Thus, being ϵ arbitrary small and M a symmetric positive definite matrix, we have $\|d(t+1)\| < \|d(t)\|$ and then $p_G(t)$ converges to p_G^* ■

5.5 Simulations

The algorithm has been tested on the testbed IEEE 37 Kersting (2001), which is an actual portion of 4.8kV power distribution network located in California. The load buses 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 Kersting (2001) for the testbed data). The length of the power lines range from a minimum of 25 meters to a maximum of almost 600 meters. We considered the scenario in which 5 microgenerators have been deployed in this portion of the power distribution grid (see Figure 5.4).

The maximum active power capabilities of each generator has been set to values that go from 85 to 490 kW. The algorithm presented in Section 5.3 have been simulated on a nonlinear exact solver of the grid Zimmerman et al. (2011).

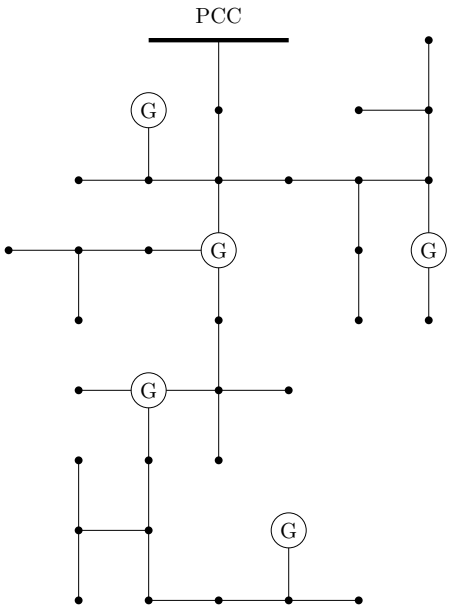


Figure 5.4: Schematic representation of the IEEE 37 test feeder Kersting (2001), where five microgenerators have been deployed.

Firstly, we simulate the first scenario, and we choose γ as one half of the bound indicated in the statement of Proposition 5.4.1, $\gamma = 4.8 \cdot 10^6$. The results of the simulation have been plotted in Figure 5.5, in which we can see a smooth convergence. The dashed black line represents the cost of the OPF solution, computed via a numerical centralized solver (Zimmerman et al. (2011)) that have access to all the grid parameters and load data, while the red line represents the behavior of the proposed algorithm. Finally, we simulate the case in which we are in the scenario treated by Proposition 5.4.2. We choose two different value of γ, much smaller than the one chosen in the other scenario, in order to reduce the difference among the discrete-time and the continuous-time algorithm implementation. The cost trajectories is depicted in Figure 5.6, where it is clear the *chattering* typically produced by the sliding mode control near the optimal solution. In Figure 5.6 and most of all in Figure 5.7, where we plot the PCC active power injected trajectory, one can sees that the "reaching phase", in which the trajectory of p_G reaches the variety $\mathcal{P}^0$. In this phase, the p_G's move towards $\mathcal{P}^0$ until they cross it. Then, the sliding mode on $\mathcal{P}^0$ with the characteristic chattering begins, and because of it p_1 starts to chatter near zero. The chattering amplitude of the trajectory with $\gamma = 8 \cdot 10^3$ is greater than the one of the trajectory with $\gamma = 8 \cdot 10^3$. Thus it could seems that the smaller is γ, the better are the performance. On the other hand, the trajectory with $\gamma = 8 \cdot 10^3$ approaches the optimal solution quicker than the other. In a realistic scenario the loads, and thus the optimal energy dispatch, are time varying. If the step size parameters

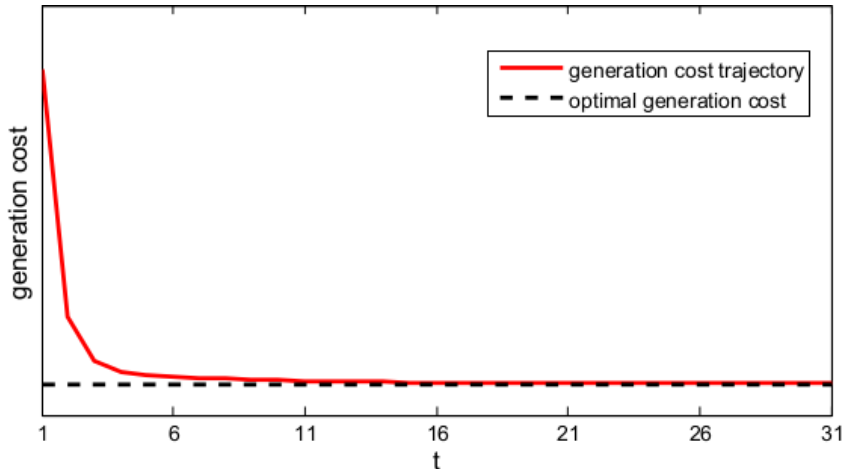


Figure 5.5: Trajectory of a algorithm run, in the first scenario, with $\gamma = 4.8 \cdot 10^6$

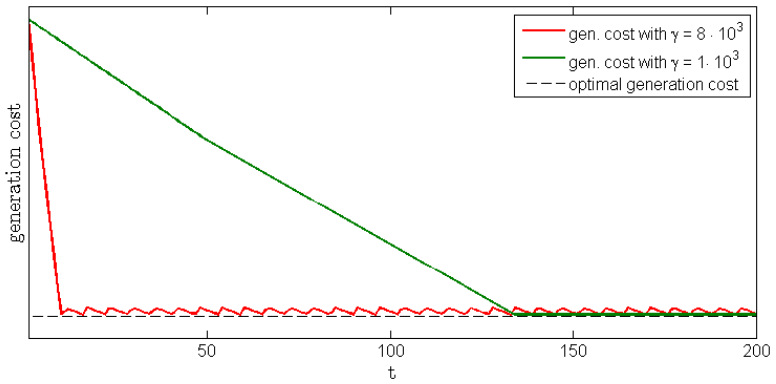


Figure 5.6: Trajectories of a algorithm run, in the third scenario considered

γ is too small, the algorithm could not be able to chase efficiently the optimum. For this reason, in the algorithm implementation on a real testbed, a tradeoff between the velocity in reaching a neighborhood of the optimum and the smallness of the chattering amplitude must be sought.

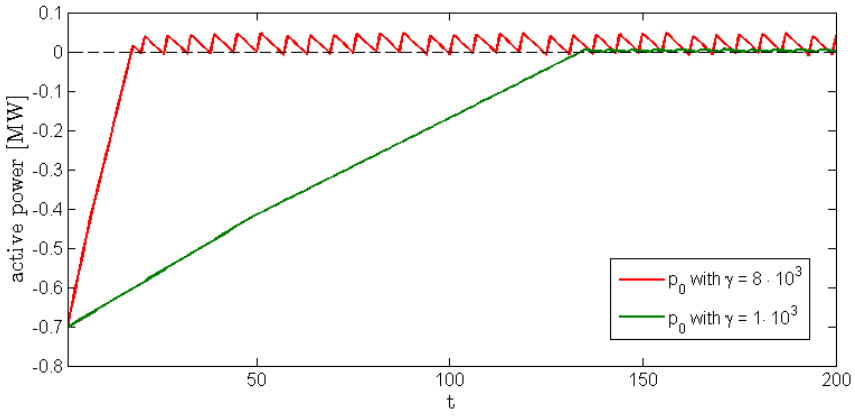


Figure 5.7: Trajectories of p_1 in the run depicted in Figure 5.6.

5.6 Chapter conclusions

In this paper we propose a feedback control strategy to solve the OPF problem in smart micro-grids with high penetration of DERs. We model the cost function, we state the OPF problem and we derive its convex approximation. Furthermore we characterize its optimal solution. Then we tackle the OPF problem by deriving a projected gradient ascent, and finally we provide some simulations in order to show and explain its behavior. We envision as future plans to introduce in this framework also the control of reactive power, to deal with other operative constraints (such as node voltage magnitude) and to study the interaction among various micro-grid connected, by their PCC, to a higher level network (e.g. high voltage grid).

6

A game theoretic approach for the OPF problem

The natural user of the algorithms presented in Chapter 4 and in Chapter 5, is a central supervisor entity, like the transmission system operator or the distribution system operator. The problem these algorithms solve, in fact, aims to minimize a function, e.g. power losses or energy production cost, that, in the end, models how much each configuration is favorable for the whole grid performance. In some sense these algorithms solve the problem as seen from the "utility point of view", as several other OPF algorithms in the literature, since the actual aim of utility companies is to optimize the grid performance. This is quite natural, since the recent control strategies for distribution networks descend from the strategies for transmission network, managed by the utilities. In order to obtain optimal performance, the solution proposed by these algorithms could present big differences among the amount of power injected by each generator. Moreover, often the optimal power injected is heavily dependent on the position of the generators.

The microgenerators owners receive a remuneration for the power they inject, that

is typically proportional to the amount injected. Thus, from the "producers point of view", the favorable, for the whole system, solutions, may not be advantageous, in terms of remuneration received. In fact, microgenerators owners would like to maximize the power they inject. The result is a conflict among the agents.

We envision Game Theory, the popular branch of mathematics that studies the interactions and conflicts between multiple players within a common system Owen (2001), as a possible approach for solving the power injection problem. In the following, the players are the microgenerators, that want to maximize the amount of active power that they inject into distribution network; the economic counterpart of this injection is that the injected power is sold to the main provider, and thus the microgenerators aim at increasing their revenues. Other works use game theory for similar goals involving control of the injected power. For example, in Weaver and Krein (2009) and Weaver (2011), the authors use a *minmax* approach, in which only one agent plays at each step, while the others maintain the injected power constant. In Saad, Han, and Poor (2011) the micro-generators try to construct the optimal *coalitions* that will exchange the power produced with the aim to minimize the power losses. In all the previous example, the compensators try to improve the global distribution system performance, and their goal functions are related to some performance index of the grid.

In the following, it will be presented a control algorithm in which the agents are only interested in maximizing a egoistic function, i.e. the amount of power that they inject and, as a consequence, the remuneration received. Of course their conflict must not interfere with the adequate grid functioning, and in particular, the algorithm will take into account the voltage magnitude constraints. This algorithm was presented for the first time in Cavraro and Badia (2013).

6.1 Problem statement

We assume that Assumption 2.8.1 holds, in order to develop our control strategy. However, simulations in Section 6.3 will be made over a standard testbed, which does not satisfy Assumption 2.8.1.

We assume that each agent is provided with some computational capability, with some communication capability, with a phasor measurement unit (PMU) and that each agent can actuate the system, by commanding the active power injected.

We assume that there is no limit on the generation capability of the microgenerators. The only constraints come from the voltage profile quality and the voltage values that are guaranteed to the customers. The customers are assumed to require a constant, fixed

amount of power, while, for the sake of simplicity, the agents are assumed to not inject reactive power. Thus, the voltages will be a function just of the active power injected by the agents. In this chapter, it is convenient sometimes to make this dependence explicit, by writing the voltage resulting from the injection of p_G as $u(p_G)$.

A compromise must be sought among the microgenerators on the amount of active power that they will inject in the distribution system, in order to satisfy their objectives to inject the maximum possible amounts of energy. Also, the operative constraints must be respected; the latter can be translated into voltage magnitude boundaries, i.e., it has to be guaranteed

$$(1-\beta)U_N\mathbf{1} \leq |u_G| \leq (1+\beta)U_N\mathbf{1} \tag{6.1}$$

where β is a confidence margin, e.g. β=0.05 in the U.S.. We recall that, by exploiting Proposition 2.32, the agents voltage can be approximated by

$$u_G = \frac{e^{\mathrm{i}\theta}}{U_N}\begin{bmatrix} M & N \end{bmatrix}\begin{bmatrix} \bar{s}_G \\ \bar{s}_L \end{bmatrix} + \mathbf{1}U_N \tag{6.2}$$

while its squared magnitude can be approximated by

$$|u_G|^2 = U_N^2\mathbf{1} + 2\,\mathrm{Re}\left(e^{\mathrm{i}\theta}M\bar{s}_G + e^{\mathrm{i}\theta}N\bar{s}_L\right) \tag{6.3}$$

where we assumed, without loss of generality, that the PCC voltage is real, i.e. $u_1 = U_N e^{\mathrm{i}\varphi_0}$ with $\varphi_0 = 0$. Now, assume the network state at some instant to be represented by the voltages $u(p_G)$ and the powers s, and that the compensators active power injection changes by $\Delta \in \mathbb{R}^m$. By (6.2), the new voltages are

$$u(p_G + \Delta) = u(p_G) + \frac{e^{\mathrm{i}\theta}}{U_N}M\Delta \tag{6.4}$$

The voltage squared magnitude becomes

$$|u_G(p_G + \Delta)|^2 = U_N^2\mathbf{1} + 2\,\mathrm{Re}\left(e^{\mathrm{i}\theta}M\bar{s}_G + e^{\mathrm{i}\theta}N\bar{s}_L\right) + 2\cos\theta M\Delta \tag{6.5}$$

Being M a matrix whose entries are all positive numbers, from equation (6.5) we can see that if a compensator increases its power injection, as a results, the voltage magnitudes rise up. Instead, if a compensator decreases its power injection, as a results, the voltage magnitudes decrease.

After this consideration, we relax the constraint and we consider only the voltage

magnitude upper bound constraint

$$|u_G| \leq (1+\beta)U_N\mathbf{1} \tag{6.6}$$

Indeed, as compensators have no limits on their generation capability and they desire to inject the maximum allowed active power, the minimum voltage bound would never be active. For the sake of simplicity, we do not consider other possible constraints such as loss limits on individual lines or thermal losses of the line.

Let $\mathcal{B}$ be the feasible set, that is

$$\mathcal{B} = \{p_G \in \mathbb{R}^m : |u_G| \leq (1+\beta)U_N\mathbf{1}\}.$$

The boundary between $\mathcal{B}$ and the unfeasible set $\mathbb{R}^m \backslash \mathcal{B}$ is called the *Pareto boundary*. In our setup, the Pareto boundary, or the *Pareto front*, is the set of all the points p_G belonging to $\mathcal{B}$ such that any of their coordinate cannot be increased further without decreasing at least one other coordinate. That is, points of the Pareto front are such that the agents voltage magnitudes satisfy the inequality (6.6), but there exists at least one agent k such that $u(p_G + \Delta)_k = (1+\beta)U_N$.

We define δ_j^M as the maximum change, in the active power injected by j such that, after the injection changing $\Delta_j^M = \delta_j^M e_j$, $p_G + \Delta_j^M$ lies into the Pareto Front. We can interpret δ_j^M as the minimum argument of the following optimization problem:

$$\max_{\delta} \quad p_j + \delta e_j \tag{6.7a}$$

$$\text{subject to } |u_G| \leq (1+\beta)U_N\mathbf{1} \tag{6.7b}$$

The value of δ_j^M can be negative in the case of over voltage, when we need to decrease the voltage magnitude. Since $\forall j, k \in \mathcal{C}$, after changing the injection by $\Delta_j = \delta_j \mathbf{1}_j$, the voltage magnitude at node k is

$$|u(p_G + \Delta_j)_k|^2 = |u(p_G)|_k^2 + 2\cos\theta M_{kj}\delta_j, \tag{6.8}$$

if $\delta_j > 0$, $|u|_k$ will increase, and conversely decrease only if δ_j is negative. As a consequence, given voltages u_G, the quantities δp_i^M will be either all positive or all negative.

Assume there are ℓ compensators that want to change their active power injection to share the active power generation.

Proposition 6.1.1. *Let $\Delta_1 = \delta_1\mathbf{1}_1, \ldots, \Delta_\ell = \delta_\ell\mathbf{1}_\ell$ be a sequence of possible changes on the injection of active power, each of them leading to a feasible point that satisfies (6.6).*

Let $\lambda \in \mathbb{R}^\ell$, with all the $\lambda(i)$'s greater than or equal to zero and such that $\mathbf{1}^T\lambda = 1$. Then the convex combination

$$\Delta_\lambda = \lambda_1\Delta_1 + \cdots + \lambda_\ell\Delta_\ell \tag{6.9}$$

leads, with respect to the model (6.4), to a feasible point.

Proof. We define $\psi^M = ((1+\beta)U_N)^2\mathbf{1} - |u(p_G)|^2$ and $\psi_\Delta = |u(p_G+\Delta)|^2 - |u(p_G)|^2$. The elements of ψ_Δ and ψ^M can be negative. If $u(p_G+\Delta)$ is a feasible point, i.e. it satisfies (6.6), it holds

$$\psi_\Delta \leq \psi^M \tag{6.10}$$

As a result, $\psi_{\Delta_j^M} \leq \psi^M$. Let u be the voltages of the compensator at some instant. If we change the amount of injected active power by (6.9), we obtain

$$|u(p_G+\Delta_\lambda)|^2 = |u(p_G)| + 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\Delta_\lambda,$$

from which

$$\psi_{\Delta_\lambda} = 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\Delta_\lambda. \tag{6.11}$$

Being the entries of M all real numbers greater than zero and $\Delta_\lambda \leq \Delta^M$, it follows trivially that

$$\begin{aligned}
\psi_{\Delta_\lambda} &= 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\Delta_\lambda \\
&= 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\sum_{j=1}^{\ell}\lambda_j\Delta_j \\
&\leq 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\sum_{j=1}^{\ell}\lambda_j\Delta^M \\
&= 2\,\mathrm{Re}(e^{\mathrm{i}\theta})M\Delta^M \leq \psi_{\Delta^M}
\end{aligned}$$

which means that $u(p_G+\Delta_\lambda)$ is feasible. ∎

Proposition 4 can be used to construct an m-dimensional region of the active powers injected by the compensators that is strictly contained in the region of feasible states of the distribution network $\mathcal{B}$. The key is to compute the δ_i's exploiting the voltages phasorial measurements taken by phasor measurement unit (PMU). This could be done by a central control unit that receives the voltage measurements by the compensators and exploits the knowledge of the grid topology, or in a distributed iterative way, even exploiting the approximation (6.5).

6.2 Active Power Injection Game

Given any state of the grid, the compensators either can or cannot increase the amount of active power injected, depending on whether (6.6) is satisfied or not. The ℓ compensators wanting to change their injected active power need to reach a (possibly fair) agreement. To this end, we propose to employ a repeated games framework. The game will choose the active power change between the convex combinations of the Δ_j^M's. This is, from the analysis of Section 6.1, within the Pareto front.

Moreover, from (6.3) the voltages magnitudes are almost linear in the active power. Thus, if each δ_j^M leads the same compensator to have a voltage magnitude of $(1+\beta)U_N$, the Pareto front of the feasible region induced the constraint (6.6) can be approximated by the convex combination of the Δ_j^M's.

Now, we analyze *three different games* that can be used to decide the amount of the injected active power change. We assume that the number of playing agents $\ell = 2^\alpha$ is a power of 2, and that the PCC does not take part in the game. This assumption is not restrictive since we could always add b dummy compensators with $\delta = 0$ such that $\ell + b = 2^\alpha$. At first, each compensator receives or computes the value $\delta_i^0 = \delta_i^M$. Then, the agents are divided into pairs playing the following game. Let $j, k \in \mathcal{C}$ be the compensators that form one of these couples. They play the game

$$\max_{0\leq\lambda\leq 1} \varphi(\lambda\delta_j^0, (1-\lambda)\delta_k^0) \tag{6.12}$$

i.e., maximize a function that models the fairness of their agreement moving on the Pareto boundary approximation, obtaining a first stage game where

$$\delta_j^1 = \lambda_j^1\delta_j^M = \lambda\delta_j^M, \quad \delta_k^1 = \lambda_k^1\delta_k^M = (1-\lambda)\delta_k^M \tag{6.13}$$

After this, each pair elects a representative, which will be paired with the representative of another pair. For instance, let the representatives of (j,k) and (f,g) be j and g respectively. Then, they play a second stage where

$$\max_{0\leq\lambda\leq 1} \varphi(\lambda\delta_1^j, (1-\lambda)\delta_1^g) \tag{6.14}$$

where φ is a proper function discussed later, obtaining

$$\delta_j^2 = \lambda_j^2\delta_j^1 = \lambda\delta_j^1 = \lambda_j^2\lambda_j^1\delta_j^M$$

$$\delta_g^2 = \lambda_g^2\delta_g^1 = (1-\lambda)\delta_g^1 = \lambda_g^2\lambda_g^1\delta_g^M$$

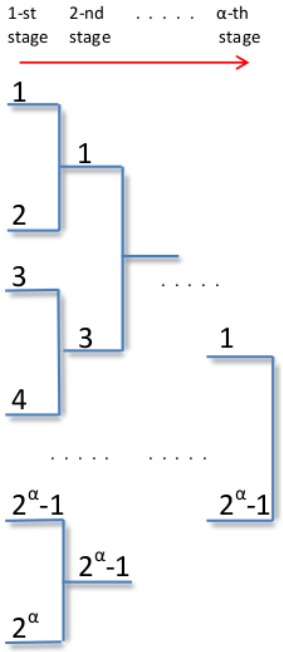


Figure 6.1: Schematic representation of the game process.

Now, j and g notify λ_j^2 and λ_g^2 to the node they were originally paired with (k and f, respectively), which compute

$$\delta_k^2 = \lambda_j^2 \delta_k^1 = \lambda_j^2 \lambda_k^1 \delta_k^M, \quad \delta_f^2 = \lambda_g^2 \delta_f^1 = \lambda_g^2 \lambda_f^1 \delta_f^M$$

This process is further iterated up to α times, and it is showed in Figure 6.1.

A suitable φ is the square of the geometric average, leading to a solution akin to Nash bargaining solution Nash (1950),

$$\varphi(x, y) = xy \ . \tag{6.15}$$

It is easy to see that the solution of the l-th stage (played for example by j and k) using (6.15) is

$$\lambda_j^l = \frac{1}{2}, \quad \lambda_k^l = \frac{1}{2} \tag{6.16}$$

and that the algorithm induces

$$\delta_1 = \frac{\delta_1^M}{2^\alpha} = \frac{\delta_1^M}{\ell}, \ldots, \delta_\ell = \frac{\delta_\ell^M}{2^\alpha} = \frac{\delta_\ell^M}{\ell} \tag{6.17}$$

and so it can be computed just knowing the number of playing agents in the distribution

network (that would not be constrained to a power of 2) and the values $\delta_1^M, \dots, \delta_\ell^M$.

Another function that can be used is

$$\varphi(x, y) = -\left|\frac{x}{\zeta_x} - \frac{y}{\zeta_y}\right| \tag{6.18}$$

which trivially forces a solution characterized by the fact that the ratio δ_j/ζ_j is equal for each compensators j. Notice that, if the ζ_j's are chosen proportionally to the nominal generation capabilities of the inverters, the solution is similar to the one obtained by the classical droop control. It is easy to see that the solution of the l-th stage (played for example by j and k) using (6.18) is

$$\lambda_i^k = \frac{\zeta_i \delta_j^{k-1}}{\zeta_j \delta_i^{k-1} + \zeta_i \delta_j^{k-1}}, \quad \lambda_j^k = \frac{\zeta_j \delta_i^{k-1}}{\zeta_j \delta_i^{k-1} + \zeta_i \delta_j^{k-1}} \tag{6.19}$$

Instead, the choice of all equal ζ_i's compels the compensators to make the same change of injected active power, ending in a solution

$$\delta_1 = \cdots = \delta_\ell \triangleq \delta_{eq} \tag{6.20}$$

where δ_{eq} is the common value of all δ_i's.

As will be seen in Section 6.3, values $\delta_1^M, \dots, \delta_\ell^M$ can be very different, and are heavily dependent by the compensators position in the grid. So, solution (6.17), in spite of its apparent fairness, can actually lead to very unbalanced outcomes, which means that some compensators are privileged because of their locations. On the other hand, the solution of (6.20) seems the most egalitarian, being all the δ_i's equal. Yet, if we compute the total injected active power $\sum_{j=1}^{\ell} \delta_j = \ell \delta_{eq}$, it can be much smaller than what computed as per (6.17), i.e., $\frac{1}{\ell} \sum_{j=1}^{\ell} \delta_j^M$. A possible trade-off is to move on the Pareto boundary between the solutions induced by (6.16) and (6.19), i.e. at the l-th stage played by j and k, to choose

$$\begin{aligned} \lambda_j^l &= \frac{\delta_k^{l-1}}{\delta_j^{l-1} + \delta_k^{l-1}} + \eta\Big(\frac{1}{2} - \frac{\delta_k^{l-1}}{\delta_j^{l-1} + \delta_k^{l-1}}\Big) \\ \lambda_k^l &= \frac{\delta_j^{l-1}}{\delta_j^{l-1} + \delta_k^{l-1}} + \eta\Big(\frac{1}{2} - \frac{\delta_j^{l-1}}{\delta_j^{l-1} + \delta_k^{l-1}}\Big) \end{aligned} \tag{6.21}$$

with $0 \leq \eta \leq 1$, η being a parameter that has to be properly designed to obtain the desired solution.

Table 6.1: Initial allocation

node	$\lvert u\rvert$ [V]	δ^M [kW]
1	4742	7478
2	4744	3190
3	4708	2121
4	4723	4909
5	4722	2603
6	4688	2881
7	4673	1614
8	4647	1624

6.3 Numerical Results

As a low voltage testbed is currently missing in the literature, we considered a 4.8 kV testbed inspired from the standard test feeder IEEE37 Kersting (2001). However, we assume that the loads are balanced, and therefore all the signals can be described in a single-phase phasorial notation. As shown in Fig. 6.2, some of the nodes are microgenerators connected to the microgrid via power inverters. Following the model proposed in Section 2.31, we consider every node (but the PCC) behaving as a constant-power device. At first the ℓ playing compensators obtain or compute $\delta_1^M, \ldots, \delta_\ell^M$. This

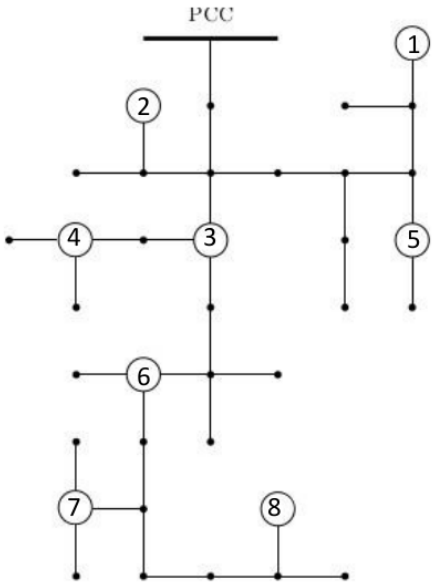


Figure 6.2: Graph of a microgrid based on the IEEE37 test feeder Kersting (2001)

computation is done considering constraint (6.6), with $\beta = 0.05$. The results are reported in Table 6.1. After that, to compare the three possible games, we compute the λ_j's playing the games determined by (6.16), (6.20), or (6.21) with $\eta = 0.7$, as well as the values

Table 6.2: Resulting Allocation by using (6.16)

node	λ	δ [kW]	$\lvert u' \rvert$ [V]
1	0.125	934.81	4873
2	0.125	398.75	4896
3	0.125	265.12	4870
4	0.125	613.68	4897
5	0.125	325.43	4914
6	0.125	360.18	4906
7	0.125	201.81	4896
8	0.125	203.00	4899

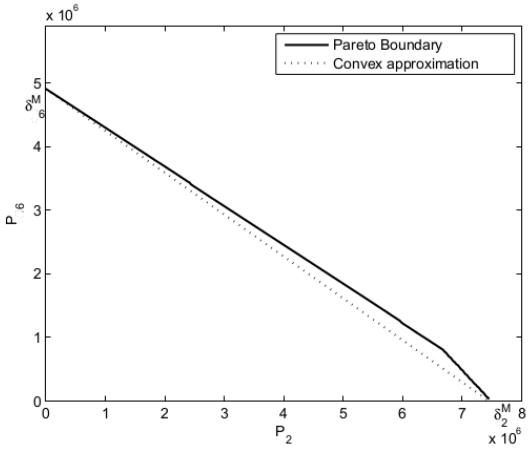


Figure 6.3: In this case, only compensators 2 and 6 change their injected power, starting from $P_2 = P_6 = 0$. The solid line represents the Pareto Front, the dotted line is its approximation by the convex combination of Δ_2 and Δ_5 (thus shown to be feasible).

injected, i.e., $\delta_j = \lambda_j \delta_j^M$. The results are reported in Tables 6.2, 6.3, 6.4, respectively. Table 6.2 shows very different values of the δ_j's. In this scenario the compensators closer to the PCC take advantage from their locations. Conversely, the δ_j in Table 6.3 are all equal, but there is a pronounced difference among the λ_j's, which implies a dissimilarity in the quantities that each compensators avoid to inject. Furthermore, the total amount of the change in the active power injection, which in some sense represents the "global fairness", is $\sum_{i=1}^{\ell} \delta_j = 2.59$ MW, much lower than the one computed with the δ_j's of Table 6.2 (3.30 MW). The δ_j's in Table 6.4 represent globally a mediation between the requirements of "individual fairness" and "global fairness", and sum to 2.88 MW. As a final consideration, in all the cases the $|u_j'|$'s are all feasible as Proposition 6.1.1 holds.

Finally. in Fig. 6.3 it is shown how the linear combination of the δ_j's approximate the Pareto front, in the case in which only compensators 2 and 6 play.

Table 6.3: Resulting Allocation by using (6.20)

node	λ	δ [kW]	$\lvert u' \rvert$ [V]
1	0.0433	353.87	4827
2	0.1015	323.76	4845
3	0.1526	323.76	4864
4	0.0659	323.76	4848
5	0.1244	323.76	4866
6	0.1124	323.76	4884
7	0.2005	323.76	4887
8	0.1994	323.76	4900

Table 6.4: Resulting Allocation by using (6.21) with $\eta = 0.7$

node	λ	δ [kW]	$\lvert u' \rvert$ [V]
1	0.0739	552.76	4857
2	0.1360	433.97	4881
3	0.1289	273.37	4854
4	0.0873	428.53	4878
5	0.1282	333.85	4896
6	0.1141	328.81	4893
7	0.1548	249.88	4917
8	0.1767	287.02	4899

6.4 Chapter conclusions

In this chapter we analyzed the problem of sharing the active power generation among the micro-generators in a smart grid. We aim at a fair, ethical sharing. First of all, we characterized a feasible region for the active power injected by the micro-generators in a smart grid. Then we studied a repeated game setup to reach an agreement among the agents about the amount of active power that they will inject, moving along a Pareto boundary approximation. We derived and analyzed three types of solutions. Finally we evaluated the resulting approach through simulation over the standard test feeder IEEE37.

Part III

Identification

7

Switches monitoring for topology identification

7.1 Topology detection problem

Topology detection in transmission networks

The knowledge of the exact grid topology is fundamental for the grid operator, in order to generate and dispatch power in the right location and in the right amount, to match load demands and especially to avoid operation that are insecure or might destabilize the system. Furthermore, as pointed out in Chen, Fuller, Diao, Zhou, Huang, and Tuffner (2011), topology changes also have significant impact on the dynamic features (especially small signal stability) of a power system.

Transmission grid operators use state estimation for supervisory control and scheduling. Through it, they find the best estimate of every node voltage magnitudes and phases, and thus the generators outputs, the loads demands, and branch power and current flows, based on telemetered measurements. Dealing with this class of problem, it is usually

assumed that the power system model and the electrical parameters are perfectly known, and that there are highly redundant measurements. This assumption was justified by the fact that the ratio of available measurements to the $2n$ (e.g. the nodes complex voltages, see Section 2.5) state variables is of about 2 in practice (Monticelli (2000)).

The problem of estimating the state of the grid is usually divided into two interrelated phases:

1. the first is topology processing and topology-error detection. In this phase, breaker status information is used to track the current topology of the grid, and errors in the calculated topology are detected and corrected. The *network topology processor* deals with it;

2. the second is state estimation proper in which analog quantities are estimated.

These two stages iterate, and the combined process is known as generalized state estimation. The robustness and reliability of state estimation is a critical issue and concern of power utilities. A wrong output given by the network topology processor will affect the estimation process, making the state estimator not reach a solution and diverge. The unavailability of state estimation solution may cause the occurrence of cascading failures or blackouts for considerable time periods, if disturbance occurs during the period of unavailability and thus can not be closely monitored. A famous case in which the estimator did not converge due to the existence of a topology error was a indirect contributing factor to the August Blackout in Northeastern U.S. in 2003, whose damage was reported to be in excess of 10 billion dollars.

Topology detection in distribution networks

Power transmission networks tend to be far better equipped with measurement devices than distribution systems, for economic reasons and necessity, than the distribution network. This is because they used to be the focus of all the control challenges, while the distribution networks were just passive, as sinks absorbing power, with power flowing from the distribution substation to the final customers, and with trivial voltage profile (the further from the PCC was the node, the lower the voltage magnitude) and currents flows (the further from the PCC was the branch, the lower the current flows).

With the advent of DERs in the distribution level, fast intermittent generators make the voltage and the currents fluctuate, in a way that can be dangerous for the grid. Furthermore, a survey from utilities experts von Meier, Brown, Arghandeh, Mehrmanesh, Cibulka, and Russ (2014b) shows on average 10 to 15 switching actions, being thus not rare, happens under an urban distribution substation. Moreover, from a practical

point of view, in typical distribution systems with low penetration of distributed energy sources and communication devices, topology estimation is probably more important than state estimation (Hoffman (2006)). Lueken, Carvalho, and Apt (2012) shows how the control of the topology of the grid can help to integrate renewables and improve the grid performance, thus underlining the importance of the topology. One of the main tasks of a distribution dispatcher is to quickly deal with forced outages on distribution networks. However, for many distribution networks there are so few telemetered measurements that often the only indication of a network outage is telephone calls from customers reporting loss of supply. Switches may not reliably communicate their status to the distribution operator, so the topology can only be determined by sending crews into the field, and the reported switch status may be faulty due to switch malfunction and unreported maintenance crew manipulation. As a consequence, it is not possible to find and fix one topology beyond any kind of uncertainty.

The tools for the detection of circuit breakers status in transmission networks do not apply to distribution networks. Distribution network today may have only few measurements, usually at the substation. Thus classical state estimation tools cannot be applied, since the number of measurements is lower than the number of states, resulting in a ill posed estimation problem. The cost of monitoring systems in distribution networks remains a barrier to equip all nodes with measurement devices. To some extent, a capable Distribution System State Estimation can compensate for the lack of direct sensor data to support observability. However, switches status errors will easily be misinterpreted as analog measurement errors (e.g. voltage or current readings). Thus topology detection is, as already remarked, an important enabling component for state estimation as well as a host of other operation and control functions based on knowledge of the system operating states in real-time.

Literature review

Most of literature on distribution network topology detection topology are based on state estimation results and measurement matching with different topologies. In Korres and Manousakis (2012) authors propose a state estimation algorithm that incorporates switching device status as additional state variables. A normalized residual test is used to identify the best estimate of the topology. Also in Kekatos and Giannakis (2012), the breaker status is incorporated in the state to be estimated, but the resulting optimization problem is solved using the ADMM. In Arya, Seetharam, Kalyanaraman, Dontas, Pavlovski, Hoy, and Kalagnanam (2011), the authors use power flow analysis for matching substation loading and aggregated household meter load data for network

connectivity modeling. They assumed metering load data are time synchronized with a measurement device on each transformer. The assumption is still far from an actual load metering system. Moreover, convergence of the proposed optimization is too sensitive to bad data. In Sharon, Annaswamy, Motto, and Chakraborty (2012), the authors solve the detection problem again exploiting state estimation, but they explicitly assume to have limited measurements, which is a realistic scenario. They provide a tool for choosing sensor placement for topology detection, too. Even though it is not an automatic tool for optimal placement of sensors, given a placement, it quickly reveals at what confidence one can detect the status of breakers. This then allows to compare easily several configurations of sensors placement, and to select the one with the highest level of confidence. Another technique for placement is showed in Wiel, Bent, Casleton, and Lawrence (2014). In Bolognani, Bof, Michelotti, Muraro, and Schenato (2013a), the authors reconstruct from voltage measurements the power distribution grid topology. They assume to have no a-priori information, but they can collect voltage phasorial measurements from every node, and they need the grid to be radial. The measurements are collected over a period of time, processed, and finally an estimate of the network topology is inferred. The proposed algorithm is derived from the methodologies derived for the identification of Markov random fields (graphical models), and is based on some conditional correlation properties that characterize voltage measurements in a radial grid. In Weng, Faloutsos, and Ilić (2014), the authors exploit an expected database with vast amounts of data from the field, to improve topology identification accuracy against "the ever-changing hard-to-predict uncertainties" in smart grids. Instead of using only a single data point (the current time), they use the historical data (i.e. topology and measurement) in the database for robustness. The idea is that two similar system measurement sets usually indicate two similar topology configurations.

Usually the measurement set consists of:

1. actual measurements of voltages and power and current flows at distribution substation transformers and often at feeder heads;
2. actual measurements of voltage and generation at distributed generator locations;
3. estimated values of loads at all unmeasured buses, from historical or forecast data.

Proposed novel approach

Most of the proposed methods in literature are post-processing methods, which depend on data collected before the detection. There is still a need for practical methods for on-line (Kezunovic (2006)) topology detection in a reliable, fast and robust approach. In

the following, a novel approach for topology detection is proposed based on time series analysis of phasor measurement unit (PMU) data.

The former works typically require the knowledge of the grid parameters and of the possible topologies, resulting from each breaker status, and the statistical knowledge of the unmonitored variables, i.e. the load demands. The developed solution *just* need the first of them, thus resulting free from the high uncertainty which usually describes each load behavior (Sevlian and Rajagopal (2014)).

This approach is inspired by high-precision phasor measurement units for distribution systems micro-synchrophasors or μ-PMU (von Meier, Culler, McEachern, and Arghandeh (2014a)). The main idea derives from the fact that time-series data from a dynamic system show specific patterns regarding system state transitions, a kind of signature left from each topology change. The algorithm is based on the comparison of the trend vector, built from system observations, with a library of signatures derived from the possible topologies transitions, and it works in *real time*. The topology detection results are impacted by load uncertainty and measurement device accuracy. A set of actual load measurements are used for a number of residential houses in the United States. The derived statistical load model is applied for topology detection scenarios to validate the proposed algorithm. The topology detection accuracy is also dependent on μ-PMU placement. A μ-PMU placement approach for topology detection application is proposed in the following. The analysis shows that topology detection converges robustly even in the realistic environment with limited sensors.

7.2 Electrical grid with switches model

In this section we model a electrical grid endowed with switches. The model can be seen as a more general version of the one proposed in Section 2.3. The electrical infrastructure is modeled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \sigma, \tau)$, in which edges represent the power lines, and nodes represent buses. We associate with the electric grid the sets

- $\mathcal{V}$ is the set of nodes (the buses), with cardinality n;
- $\mathcal{E}$ is the set of edges (the electrical lines connecting them), with cardinality $|\mathcal{E}|$;
- $\mathcal{S}$ is the set of the switches (or breakers) deployed in the electrical grid, with cardinality r
- $\mathcal{Y}$ is the set of the electrical grid nodes endowed with a sensor, with cardinality $|\mathcal{Y}|$

The system is described, beside the variables u, i, ξ and s, by the following quantities:

- $w \in \{0,1\}^r$, where $w_v = 0$ if the switch v is open, $w_v = 1$ if the switch v is closed;
- $y \in \mathbb{C}^{|\mathcal{Y}|}$, where y_v is the voltage measured by the sensor v at the frequency f.

We define as $\mathcal{E}_\mathcal{S} \subset \mathcal{E}$ the subset of the edges that are endowed with a switch. Each electric line e is characterized by the physical impedance z_e. The incidence matrix of $\mathcal{G}$ is $A \in \{0, \pm 1\}^{|\mathcal{E}| \times n}$ is

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_{|\mathcal{E}|} \end{bmatrix}^T \tag{7.1}$$

where the rows $a_1, \ldots, a_r$ are associated with the edges in $\mathcal{E}_\mathcal{S}$. We collect the physical impedances of the edges endowed with a switch in the vector $\mathrm{z}_s \in \mathbb{C}^r$, with the same order of w, while all the other in the vector $\mathrm{z}_n \in \mathbb{C}^{|\mathcal{E}|-r}$ However, a line endowed with a switch appear as having different impedance, in relation with the switch status, e.g. infinite impedance when the breaker is open. In order to model this fact, we build the function

$$\hat{\mathrm{z}}_s(w) := w \odot \mathrm{z}_s \tag{7.2}$$

resulting $(\hat{\mathrm{z}}_s(w))_j = (\mathrm{z}_s)_j$ if the switch j is closed, $(\hat{\mathrm{z}}_s(w))_j = \infty$ if the switch j is open. The status w is associated with a topology completely described by the bus admittance matrix

$$\boldsymbol{Y}_w := A^T (Z_w)^{-1} A \tag{7.3}$$

where

$$Z_w = \begin{bmatrix} z_1 & & 0 & & & \\ & \ddots & & & & \\ 0 & & z_r & & & \\ & & & z_{r+1} & & 0 \\ & & & & \ddots & \\ & & & 0 & & z_{|\mathcal{E}|} \end{bmatrix}$$

and whose Green Matrix is denoted with $\boldsymbol{X}_w$.

We assume that in the grid there are deployed some PMUs, that provide voltage measurements regularly and synchronously at frequency f, at the times h_k, with $h_k \in \mathbb{R}$, $k \in \mathbb{Z}$. On the other hand, during the operation of the electrical grid, there can be changes of the breakers status, mainly due to

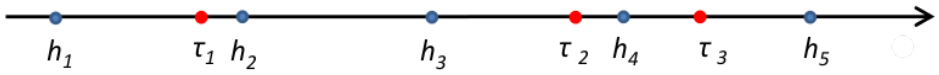


Figure 7.1: The blue points represent the time instants in which the PMUs provide a voltage measurements, while the red points represent the time instants in which there is a switching action

1. overload;
2. over-generation;
3. faults;
4. maintenance.

We denote by τ_k, with $\tau_k \in \mathbb{R}$, $k \in \mathbb{Z}$, the time instants when there is a change of the breaker status. While the time interval between two measurements is always the same, i.e. $h_{k+1} - h_k = 1/f$, the time interval between two switches actions, i.e. $\tau_{k+1} - \tau_k$, is not regular, as we can see in Figure 7.1.

7.3 Identification of Switching Actions

In this section we discuss how topology changes appear in the network equation set. The basic idea behind our proposed approach is that changes in switching status will create specific signatures in the grid voltages. In order to develop the theoretical base for the proposed algorithm, beside Assumption 2.8.1 we make the following assumptions.

Assumption 7.3.1. $\mathbf{1}^T |w(\tau_k^+) - w(\tau_k)| \leq 1$, $\forall \tau_k$, i.e. only one switch can change its status at each time (we recall that given a vector v, $|v|$ is the vector such that $|v|_j = |(v_j)|$).

Assumption 7.3.2. The electrical network is always connected, i.e. there are no admissible states in which any portion of the grid remains disconnected.

Assumption 7.3.3. The time instants in whiche there is a topology change are always different from the time instants in which the PMUs take the measurements, i.e. $\tau_m \neq h_n, \forall m, n \in \mathbb{Z}$.

Assumption 2.8.1 will be relaxed in Section 7.8, in order to test the algorithm in a more realistic scenario. Assumption 7.3.1 is instead reasonable for the proposed algorithm framework: it works in the scale of some seconds, while typically the switches are electro-mechanical devices and their actions are not simultaneous. Finally, Assumption 7.3.2 is always satisfied during the normal operation.

It is convenient to describe the value of the breaker status at the times when the PMUs measure the voltages, by setting

$$w(h_n) = w(\tau_m), \text{ if } \tau_m < h_n, \tau_{m+1} > h_n \tag{7.4}$$

In this way we build a sequence of breaker states $\{w(h_k)\}$. We sample furthermore voltages and powers these variables synchronously with the voltage measurements, thus obtaining ordered sequences, e.g. the active powers sequence $\{p(h_k)\}$. We can therefore express all the system variables as discrete time function:

$$\begin{aligned} w(t) &= w(h_t) \\ p(t) &= p(h_t) \\ q(t) &= q(h_t) \\ u(t) &= u(h_t) \end{aligned}$$

The following quantities will be useful for the algorithm development.

Definition 7.3.4 (Trend vector)**.** The *trend vector* $\delta(t_1, t_2) \in \mathbb{C}^n$ is the difference between phasorial voltages taken at the discrete time instants t_1 and t_2, i.e.

$$\delta(t_1, t_2) = u(t_1) - u(t_2)$$

.

Definition 7.3.5 (Signature matrix)**.** Given a breaker status w associated with the topology $\boldsymbol{Y}_w = (1 + \mathrm{i}\alpha)U\Sigma_R U^*$, and fixed an edge ℓ endowed with a breaker, associated with the row a_ℓ in the incidence matrix (7.1), we define the signature matrix $\Phi_{w_{-\ell}}$ as

$$\Phi_{w_{-\ell}} = U\Sigma_R^{-1}U^* - U(\Sigma_R + \mathrm{Re}(\mathrm{z}_\ell^{-1})U^* a_\ell a_\ell^T U)^{-1}U^* \tag{7.5}$$

We introduce the main idea that underlies our algorithm with the following example.

Example 7.3.6. Assume that at time t the switches status is described by $w(t)$, and that the switch ℓ is open, that is $w(t)_\ell = 0$, resulting in the topology $\boldsymbol{Y}_{w(t)}$. We recall that, holding Assumption 2.8.1, the bus admittance matrix can be written as

$$\boldsymbol{Y}_{w(t)} = (1 + \mathrm{i}\alpha)U\Sigma_R U^*,$$

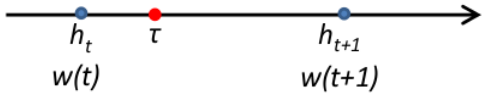


Figure 7.2: w(t) and w(t+1) describe the breakers status at the times h_t and h_{t+1}, i.e. the discrete times t and $t+1$

while the Green matrix can be written as

$$\boldsymbol{X}_{w(t)} = \frac{1}{1+\mathrm{i}\alpha}(I - \mathbf{1}e_1^T)(U\Sigma_R^{-1}U^*)(I - e_1\mathbf{1}^T),$$

as showed in Section 2.8. Applying Proposition 2.6.1 and neglecting the infinitesimal term, the voltages can be expressed as

$$u(t) = \boldsymbol{X}_{w(t)}\frac{\bar{s}}{U_N} + \mathbf{1}U_N \tag{7.6}$$

Then, the ℓ-th switch changes its status. At the discrete time $t+1$, the new status is described by $w(t+1) = w_2$, with $w(t+1)_\ell = 1$, associated with the topology $\boldsymbol{Y}_{w(t+1)}$. The time sequence of the events, i.e the measurements at the times h_t and h_{t+1} (the discrete times t and $t+1$) and the topology change at time τ, is depicted in Figure 7.2 We have basically added the edge in which switch ℓ is placed. Its admittance is $(\mathrm{z}_\ell)^{-1}$, and it is such that $\mathrm{Im}(\mathrm{z}_\ell^{-1}) = \alpha\,\mathrm{Re}(\mathrm{z}_\ell^{-1})$.

By exploiting (7.2) and (7.3), we can write

$$\boldsymbol{Y}_{w(t+1)} = \boldsymbol{Y}_{w(t)} + \mathrm{z}_\ell^{-1}a_\ell a_\ell^T$$

where a_ℓ is the ℓ-th row of the incidence matrix (7.1). Since a_ℓ is orthogonal to $\mathbf{1}$, it can be expressed as

$$\begin{aligned} a_\ell &= Ia_\ell \\ &= \left(UU^* + \frac{\mathbf{1}\mathbf{1}^T}{n}\right)a_\ell \\ &= UU^*a_\ell. \end{aligned}$$

Therefore, we can write

$$\begin{aligned} \boldsymbol{Y}_{w(t+1)} &= (1+\mathrm{i}\alpha)(U\Sigma_R U^* + \mathrm{z}_\ell^{-1}UU^*a_\ell a_\ell^T UU^*) \\ &= (1+\mathrm{i}\alpha)U(\Sigma_R + \mathrm{Re}(\mathrm{z}_\ell^{-1})U^*a_\ell a_\ell^T U)U^* \end{aligned} \tag{7.7}$$

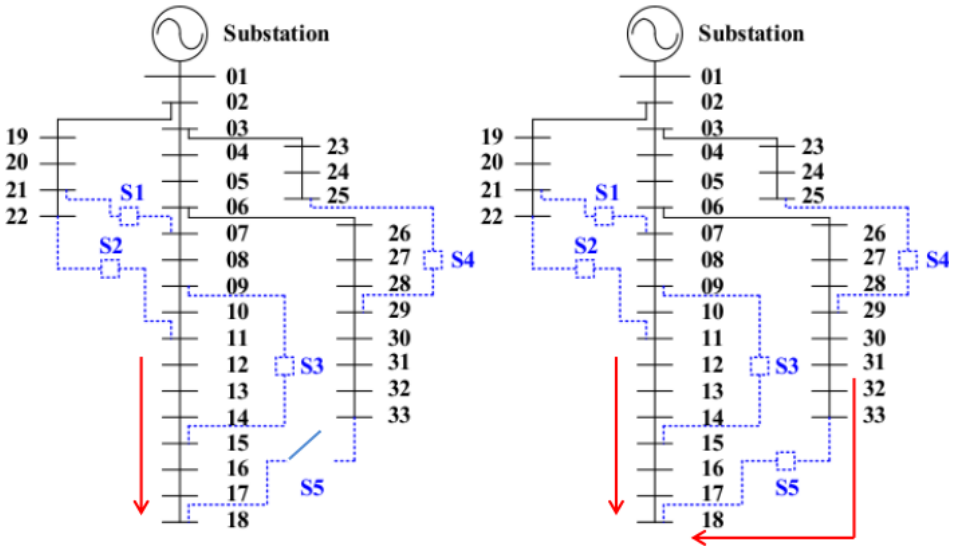


Figure 7.3: In the left-hand side of the figure, the red arrow represent the path of the current feeding node 18 with breaker 5 open. When, instead we close breaker 5, the current come from both the central and the right branch, as showed in the right-hand side of the figure.

and, retracing the proof of Lemma 2.8.2

$$\boldsymbol{X}_{w(t+1)} = \frac{1}{1+\mathrm{i}\alpha}(I - \mathbf{1}e_1^T)U(\Sigma_R + \mathrm{Re}(\mathrm{z}_\ell^{-1})U^* a_\ell a_\ell^T U)^{-1}U^*(I - e_1\mathbf{1}^T) \tag{7.8}$$

The voltages are

$$u(t+1) = \boldsymbol{X}_{w(t+1)}\frac{\bar{s}}{U_N} + \mathbf{1}U_N \tag{7.9}$$

From (7.6) and (7.9), the trend vector can be written as

$$\delta(t, t-1) = \frac{1}{1+\mathrm{i}\alpha}(I - \mathbf{1}e_1^T)\Phi_{w(t)_{-\ell}}(I - e_1\mathbf{1}^T)\frac{\bar{s}}{U_N} \tag{7.10}$$

where we exploited the signature matrix $\Phi_{w(t)_{-\ell}}$.

Example 7.3.7. Consider now the opposite situation: at time $t-1$ the ℓ-th switch is closed and at time t it is open, i.e. $w(t-1) = w_2, w(t) = w_1$. In this case that

$$\delta(t, t-1) = -\frac{1}{1+\mathrm{i}\alpha}(I - \mathbf{1}e_1^T)\Phi_{w(t)_{-\ell}}(I - e_1\mathbf{1}^T)\frac{\bar{s}}{U_N} \tag{7.11}$$

We can observe that when there is a switching action, the voltage profile varies in a way determined by the particular transition from a topology to another. That's because we are basically varying the currents paths, as showed in Figure 7.3. In particular, the trend vector $\delta(t_1, t_2)$ depends on $\Phi_{w(t)_{-\ell}}$. Furthermore, opening or closing a determined

switch makes the system vary in the opposite way, e.g. compare (7.10) with (7.11). This is basically due to the fact that $(w_1)_{-\ell} = (w_2)_{-\ell}$. The following lemma (Miller (1981)) will be useful to prove a characteristic of $\Phi_{w(t)_{-\ell}}$ that is fundamental for the development of our topology detection algorithm.

Lemma 7.3.8 (K. Miller Lemma)**.** *Let G and $G + E$ be non-singular matrices where E is a matrix of rank one. Let $g = Tr(EG)$. Then*

$$(G + E)^{-1} = G^{-1} - \frac{1}{1 + g}(G^{-1}EG^{-1}) \tag{7.12}$$

Proposition 7.3.9. *For every transition from the state described by $w(t)$ to the one described by $w(t + 1)$, due to a action of the switch ℓ, $\Phi_{w(t)_{-\ell}}$ is a rank one matrix.*

Proof. Following the same reasoning of Example 7.3.6, and exploiting K. Miller Lemma, with some simple computations we can write

$$\Phi_{w(t)_{-\ell}} = \frac{1}{1 + \mathrm{Tr}(\mathrm{Re}(\mathrm{z}_\ell^{-1})U^* a_\ell a_\ell^T U \Sigma_R)} U \Sigma_R^{-1} U^* a_\ell a_\ell^T U \Sigma_R^{-1} U^* \tag{7.13}$$

It's trivial to see that $\Phi_{w(t)_{-\ell}}$ is a rank one matrix with eigenvector

$$\hat{g}_{w(t)_{-\ell}} = U \Sigma_R^{-1} U^* a_\ell \tag{7.14}$$

associated with the non-zero eigenvalue

$$\lambda_{w(t)_{-\ell}} = \frac{1}{1 + \mathrm{Tr}(\mathrm{Re}(\mathrm{z}_\ell^{-1})U^* a_\ell a_\ell^T U \Sigma_R)} \| U \Sigma_R^{-1} U^* a_\ell \|^2 \tag{7.15}$$

Thus, we have

$$\Phi_{w(t)_{-\ell}} = \lambda_{w(t)_{-\ell}} \hat{g}_{w(t)_{-\ell}} \hat{g}^*_{w(t)_{-\ell}}$$

■

The trend vector $\delta(t + 1, t)$ represents how the opening or the closure of a switch spreads on the voltages profile. Thanks to Proposition 7.3.9 we can write it as

$$\begin{aligned} \delta(t + 1, t) &= \frac{\lambda_{w(t)_{-\ell}}}{1 + \mathrm{i}\alpha} (I - \mathbf{1} e_1^T) \hat{g}_{w(t)_{-\ell}} \left[\hat{g}^*_{w(t)_{-\ell}} (I - e_1 \mathbf{1}^T) \frac{\bar{s}}{U_N} \right] \\ &= \frac{\lambda_{w(t)_{-\ell}}}{1 + \mathrm{i}\alpha} \left[\hat{g}^*_{w(t)_{-\ell}} (I - \mathbf{1} e_1^T) \frac{\bar{s}}{U_N} \right] (I - \mathbf{1} e_1^T) \hat{g}_{w(t)_{-\ell}} \end{aligned} \tag{7.16}$$

from which we see that

$$\delta(t + 1, t) \propto (I - \mathbf{1} e_1^T) \hat{g}_{w(t)_{-\ell}}. \tag{7.17}$$

Therefore, every specific switching action pattern that appears on the voltage profile is proportional to the eigenvector $\hat{g}_{w(t)_{-\ell}}$. Thus, $g_{w(t)_{-\ell}}$ can be seen as the *particular signature* of the switch action. This fact is the cornerstone for the topology detection algorithm in this paper.

Remark 7.3.10. Equation (7.16) points out the fact that the trend vector after a switch action will always be proportional to $g_{w(t)_{-\ell}}$, irrespective of other variables such as voltages u and loads s that describe the network operating state at the time.

Remark 7.3.11. The opening and the closure of the switch ℓ, once the other switches status is fixed, share the same signature $g_{w(t)_{-\ell}}$, thus in principle they are indistinguishable. Without any other information, the trend vector can just identify which breaker has changed its status.

7.4 Topology Detection Algorithm

Assuming the distribution network physical infrastructure, i.e. conductor impedances and breakers locations, known, we can construct a *library* $\mathcal{L}$ in which we collect all the normalized products between $(I - \mathbf{1}e_1^T)$ and the eigenvectors (7.14) for all the admissible switching action

$$\mathcal{L} = \{g_{w_{-\ell}} : w \text{ satisfies Assumption } 7.3.2\} \tag{7.18}$$

where

$$g_{w(t)_{-\ell}} = \frac{(I - \mathbf{1}e_1^T)\hat{g}_{w(t)_{-\ell}}}{\|(I - \mathbf{1}e_1^T)\hat{g}_{w(t)_{-\ell}}\|} \tag{7.19}$$

It is natural then to compare at each time the trend vector $\delta(t+1,t)$ with the elements in the library to identify if and which switch changed its status at time $t+1$. As stated in Remark 7.3.11, if we want to identify which is the topology after the transition, we need additional information, i.e. the knowledge of the topology before the transition.

If we know the exact switches status, we could compare the trend vector with a restricted portion of the library $\mathcal{L}$, since there are only r possible transitions, each of one caused by the action of one of the r switches. That is, we can compare $\delta(t+1,t)$ with the *particular library*

$$\mathcal{L}_{w(t)} = \{g_{w(t+1)_{-\ell}} : w(t+1)_{-\ell} = w(t)_{-\ell}\} \tag{7.20}$$

that is peculiar to the state $w(t)$. The procedure is stated in Algorithm 1.

The comparison is made by projecting the normalized measurements-based trend vector $\frac{\delta(t+1,t)}{\|\delta(t+1,t)\|}$ onto the topology library $\mathcal{L}_{w(t)}$. The projection is performed with the

inner product, and it allows us to obtain for each vector in $\mathcal{L}_{w(t)}$

$$c_{w(t)_{-\ell}} = \left\| \left\langle \frac{\delta}{\|\delta\|}, g_{w(t)_{-\ell}} \right\rangle \right\|, \quad \forall \ell \in \mathcal{S}. \tag{7.21}$$

All the coefficient $c_{w(t)_{-\ell}}$ are collected in the set $\mathcal{C}$. If $c_{w(t)_{-\ell}} \simeq 1$, it means that $\delta(t+1,t)$ is spanned by $g_{w(t)_{-\ell}}$ and then that the switch ℓ changed its status. Because of the approximation (2.32), the projection will never be exactly one. Therefore, we introduce a heuristic threshold, called *min_proj*. If the projection is greater than the threshold, switch ℓ is assumed to have changed its status and the topology change time is detected. If there is no switches action, the trend vector will be zero as all the $c_{w(t)_{-\ell}}$, and the algorithm will not reveal any topology transition. Thus, the projection value is used by the algorithm to detect the change time too, differently of what proposed in Cavraro, Arghandeh, Barchi, and von Meier (2015), where the authors used the norm of a matrix built from measurements (the *trend matrix*). The new approach is more reliable in the realistic, noisy case. With a slight abuse of notation, we will say that the maximizer of $\mathcal{C}$, denoted by $\max \mathcal{C}$, is the switches status w such that $w_{-\ell} = w(t)_{-\ell}$, $w_\ell = 1$ if $w(t)_{-\ell} = 0$ or vice-versa $w_\ell = 0$ if $w(t)_{-\ell} = 1$ and $c_{w(t)_{-\ell}}$ its the maximum element in $\mathcal{C}$.

We tacitly assumed so far, just to show the main idea, all the buses endowed with a PMU. But this is not a realistic scenario for distribution network. Now we show how the former approach can be generalized in presence of limited information, in which we are allowed to take a few voltage measures. In this case, the measurement vector is

$$\begin{aligned} y &= I_{\mathcal{Y}} u \\ &= I_{\mathcal{Y}} \boldsymbol{X}_{w(t)} \frac{\bar{s}}{U_N} + \mathbf{1} U_N \end{aligned} \tag{7.22}$$

where $I_{\mathcal{Y}} \in [0,1]^{|\mathcal{Y}| \times n}$ is a matrix that select the entries of u where a PMU is placed. In that case, the algorithm works as previously, with the trend vector

$$\delta(t_1, t_2) = y(t_1) - y(t_2) \tag{7.23}$$

The library vectors and their dimension change too. In fact one can easily shows, using (7.22) and retracing (7.9) and (7.13) that (7.19) becomes

$$g_{w(t)_{-\ell}} = \frac{I_{\mathcal{Y}}(I - e_0 \mathbf{1}^T)\hat{g}_{w(t)_{-\ell}}}{\|I_{\mathcal{Y}}(I - e_0 \mathbf{1}^T)\hat{g}_{w(t)_{-\ell}}\|} \tag{7.24}$$

collected in the particular library $\mathcal{L}_{\mathcal{Y},w(t)}$.

Algorithm 1 Topology Changes Detection

Require: At each time $t+1$, $w(t)$, $min\_proj = 0.98$

1: each PMU records the voltage phasor measurement $y_j(t)$
2: the algorithm builds the trend vector $\delta(t, t-1)$
3: the algorithm projects $\delta(t, t-1)$ in the library $\mathcal{L}_{\mathcal{Y},w(t)}$ obtaining the set of values

$$\mathcal{C} = \left\{ c_{[w(t)]_{-\ell}} = \left\| \left\langle \frac{\delta}{\|\delta\|}, g_{[w(t)]_{-\ell}} \right\rangle \right\|, g_{[w(t)]_{-\ell}} \in \mathcal{L}_{\mathcal{P},w(t)} \right\};$$

4: **if** $\max \mathcal{C} \geq min\_proj$ **then**
5: $\quad w(t) \leftarrow \arg\max \mathcal{C}$
6: **else**
7: $\quad w(t+1) \leftarrow w(t)$
8: **end if**

7.5 PMUs placement

When we have a limited number of sensors to be deployed in the distribution grid, the first requirement to be satisfied is the system *observability*. It is related to the capability of the algorithm of detecting every topology transition. Since our algorithm is basically based on the comparison between trend vectors and the library vectors, the trivial condition for the observability of the network is that each vector of the library is not proportional to any of the others.

Definition 7.5.1 (Observability)**.** Given $\mathcal{Y}$, the set of nodes endowed with PMUs, let, with a slight abuse of notation, the juxtaposition of the library vectors be denoted by $\mathcal{L}_{\mathcal{Y}}$. Then if

$$(|\mathcal{L}_{\mathcal{Y}}^{*}\mathcal{L}_{\mathcal{Y}}|)_{uv} < 1, \forall u \neq v \tag{7.25}$$

the switch that changes its status can be identified and thus the grid is observable.

Notice that the element $(\mathcal{L}_{\mathcal{Y}}^{*}\mathcal{L}_{\mathcal{Y}})_{uv}$ is simply the projection of $g_{w(t)_{-u}}$ onto $g_{w(t)_{-v}}$. If they are note purely proportional, trivially the magnitude of their inner product is smaller than one.

Condition (7.25) can be too restrictive, if we know the switches status before the topology change. In that case, in fact, we can just check that each vector of the *particular* library is not proportional to any of the others.

Definition 7.5.2 (Weak observability)**.** Let the switches status be w, and let it be known. Let, with a slight abuse of notation, the juxtaposition of the particular library

vectors be denoted by $\mathcal{L}_{\mathcal{Y},w}$. Then if

$$(|\mathcal{L}^*_{\mathcal{Y},w}\mathcal{L}_{\mathcal{Y},w}|)_{uv} < 1, \forall u \neq v, \forall w \tag{7.26}$$

each switch action can be identified and thus the placements $\mathcal{Y}$ leads to system observability.

Equations (7.25) and (7.26) can be used to infer, given the electrical grid model, which is the minimum number of PMUs to be deployed in order to have the observability.

A second issue is to find an "optimal" placement. The notion of optimal is always relative to objective function to be maximized. However, here we propose a simply, even if onerous, strategy for placement. After finding a minimal number of sensors and a placement that guarantee the (weak) observability, we propose a greedy PMU placement procedure based on the sequential addition of the PMU able to provide the best performance improvement, verified by Monte Carlo simulations. For every possible new place for PMU, we run num_run Monte Carlo simulation, of length TSTOP. For each of them we choose , randomly, the initial w, the switch ℓ that changes its status and the time τ of the action. The place for the new PMU that performs the minor number of errors is then chosen.

Algorithm 2 Greedy PMU placement

Require: $\mathcal{Y}$ such that the network is (weak) observable, num_run, TSTOP
min $\leftarrow \infty$
minimizer $\leftarrow \emptyset$
for every possible place $j \notin \mathcal{Y}$ **do**
    err $\leftarrow 0$
    for $t = 1$ to num_run **do**
        choose $w \sim \mathcal{U}(\{w : w$ satisfies Assumption 7.3.2$)$
        choose $\ell \sim \mathcal{U}(\{0, \ldots, r\})$
        choose $\tau \sim \mathcal{U}(\{0, \ldots, \text{TSTOP}\})$
        run Monte Carlo simulations
        err = number of errors
        if err $\leq$ min **then**
            min $\leftarrow$ err
            minimizer $\leftarrow j$
        end if
    end for
end for
$\mathcal{Y} \leftarrow \mathcal{Y} \cup \{$minimizer$\}$

7.6 Switching action in the non-ideal scenario

So far, we considered the case in which the measurements devices were not affected by noise and the loads were static. In reality, there is some noise associated with PMUs, i.e. the output of our PMU placed at bus j is

$$y_j = u_j + e_j \tag{7.27}$$

where $e_j \in \mathbb{C}$ is the error that the measurements device commits. A common characterization of the error is the *total vector error*. For example if x is the variable to be measured, and x^N is the measured value, the total vector error TVE is

$$TVE = \frac{|x - x^N|}{|x|} \tag{7.28}$$

Furthermore, the loads are not static but they have a natural dynamic. We assume that the loads have constant *power factor*, and consequently

$$\frac{q(t)_j}{p(t)_j} = \gamma_j, \quad \forall t, j = 2, \dots, n. \tag{7.29}$$

We model the active power and the reactive power consumption at each load by

$$p(t+1)_{-1} = p(t)_{-1} + n_p(t) \tag{7.30}$$

$$q(t)_{-1} = \text{diag}(\gamma_2, \dots, \gamma_n)\, p(t)_{-1} \tag{7.31}$$

where $n_p(t)$ is a Gaussian random vector . This could seem a too rough and unsophisticated, but we will see in Section 7.7 that is actually enough accurate.

If we take into account (7.27), (7.30) and (7.31), the trend vector becomes

$$\begin{aligned}\delta(t_1, t_2) = I_{\mathcal{Y}}(\boldsymbol{X}_{w(t_1)} - \boldsymbol{X}_{w(t_2)})\frac{\bar{s}(t_2)}{U_N} + e_{t_1} - e_{t_2} + \\ + \frac{I_{\mathcal{Y}}\boldsymbol{X}_{w(t_1)}}{U_N}\sum_{t=t_2}^{t_1-1}(I - \mathrm{i}\ \text{diag}(\gamma_2, \dots, \gamma_n))n_p(t)\end{aligned} \tag{7.32}$$

being

$$p(t_1)_{-1} = p(t_2)_{-1} + \sum_{t=t_2}^{t_1-1} n_p(t)$$

Therefore, because of measurement noise and load dynamics, the trend vector is typically non-zero even if there has not been any switching action. Thus, the projection

(7.21) may be almost one, leading to false topology detection. When any switch is closed or opened, since we are adding or deleting a branch, we are changing the currents flows, reflecting on an abrupt, greater voltages variation. Therefore, a first strategy to avoid false positive is to not consider trend vector whose norm is lower than a defined threshold, called in the following *min_norm*.

Moreover the additive noise can make the maximum projection value of the trend vector onto the library vectors, $\max \mathcal{C}$, considerably lower than one, even if a topology change occurred. This fact induces us to use a value for *min_proj* lower than the one considered in Algorithm 1. Of course, the use of a lower threshold make the algorithm more vulnerable to false positive. The following example will give the idea for a possible, solution.

Example 7.6.1. Assume the grid without load variation and measurements noise, and that at time t_1 the ℓ-th switch change its status. Consider the trend vector

$$\delta(t, t-T) = y(t) - y(t-T).$$

For $t < t_1$ and $t \geq t_1 + T$ the projections of the trend vector onto the library are all equal to zero, because

$$\delta(t, t-T) = y(t) - y(t-T) = 0$$

Instead for $t_1 \leq t < t_1 + T$, the trend vector is

$$\delta(t, t-T) = (I - e_0 \mathbf{1}^T)\Phi_{w(t)_{-\ell}} \frac{\bar{s}}{U_N}$$

leading to a cluster of algorithm time instant of length T (or $\frac{T}{f}$ seconds), in which the maximum projection coefficient will be almost one.

A possible solution is thus to consider a trend vector built using not two consecutive measures, but considering measures separated by T algorithm time instants , i.e.

$$\delta(t, t-T) = y(t) - y(t-T).$$

and to assume that a topology change has happened at time t when we have a cluster whose length length_cluster is T of consecutive values of $\max \mathcal{C}$ greater than min_proj. This idea will be clarified with some simulations in section 7.8. The former observations lead to the Algorithm 3 for topology detection with measurements noise and load variation.

Algorithm 3 Topology transition detection with noise

Require: At each time $t+1$, we are given the variables $w(t)$, minimizer(t), length_cluster(t)

$w(t+1) \leftarrow w(t)$
each PMU at each node j record voltage phasor measurements $y_j(t+1)$
the algorithm builds the trend vector $\delta(t+1, t+1-T)$
if $\|\delta(t+1, t+1-T)\| <$ min_norm **then**
    $\delta(t+1, t+1-T) \leftarrow 0$
    minimizer$(t+1) = 0$
    length_cluster$(t+1) = 0$
else
    the algorithm projects $\delta(t+1, t+1-T)$ in the particular library $\mathcal{L}_{\mathcal{Y},w(t)}$ obtaining the set of values

$$\mathcal{C} = \left\{ c_{[w(t)]_{-\ell}} = \left\| \left\langle \frac{\delta}{\|\delta\|}, g_{[w(t)]_{-\ell}} \right\rangle \right\|, g_{[w(t)]_{-\ell}} \in \mathcal{L} \right\};$$

    if $\max \mathcal{C} >$min_proj **then**
        minimizer$(t+1) = \arg\min \mathcal{C}$
        if minimizer$(t+1) =$ minimizer(t) **then**
            length_cluster$(t+1) \leftarrow$ length_cluster$(t) + 1$
            if length_cluster$(t+1) = T$ **then**
                $w(t+1) \leftarrow$ minimizer$(t+1)$
            end if
        else
            length_cluster$(t+1) \leftarrow 1$
        end if
    end if
end if

7.7 Actual loads dynamic characterization

Load behavior at household level is usually a critical question in distribution networks. Due to lack of accurate and high resolution measurement at meters, there is no clear answer to that question. The most available data from loads come from meters with hourly (or 15 minutes) time intervals. To add more practicality to the proposed topology detection algorithm in this paper, a load measurement data set for five residential houses in the United States is used. Load demand (kW) are recorded every second for a week. Statistical analysis of these load data are presented in Table 7.1. For our topology detection algorithm, the load variation is more important, since it affects topology detection capability, as (7.32) shows. Table 7.2 reports instead the statistical analysis of the variation of the load demand between two consecutive seconds. The field "Relative SD" expresses the standard deviation of the load or of the load change as a percentage of the load peak value, written Table 7.1. The Figure 7.4 shows the histogram of load duration and load changes. The orange bars represents measurements data points over the percentage of pick load values. The blue bars show load change in one second intervals frequency over the percentage of pick load values. In case of load changes, more than 90 percent of measurement time points have loads with less than one percents of pick load. It means that during one second time intervals, there is not huge difference in load data.

In the United States, a number of houses are connected to one distribution transformer. Therefore, the aggregated five houses loads are considered as the reference for load variability in this paper. The lower the measurements frequency, the higher the load variability. Re-sampling the data, we obtain the aggregated characterization reported in Table 7.3 for different measurements frequencies, i.e. for one measure every second, one measure every five seconds and one measure every ten seconds.

Further analysis shows that the load differences, for all the frequencies, can be modeled as Gaussian random variables thus validating equations (7.30) and (7.31), that can be assumed as a realistic model of load variation. Figure 7.5 is the output, for the 1 second

Table 7.1: Load values

	Mean (kW)	SD (kW)	Max (kW)	Relative SD (%)
House 1	0.541	0.415	5.990	6.93
House 2	0.917	1.309	10.956	11.95
House 3	1.787	1.955	11.578	16.88
House 4	2.937	2.566	12.364	20.75
House 5	0.846	1.309	8.155	16.06
Aggregate	7.029	4.204	27.229	15.44

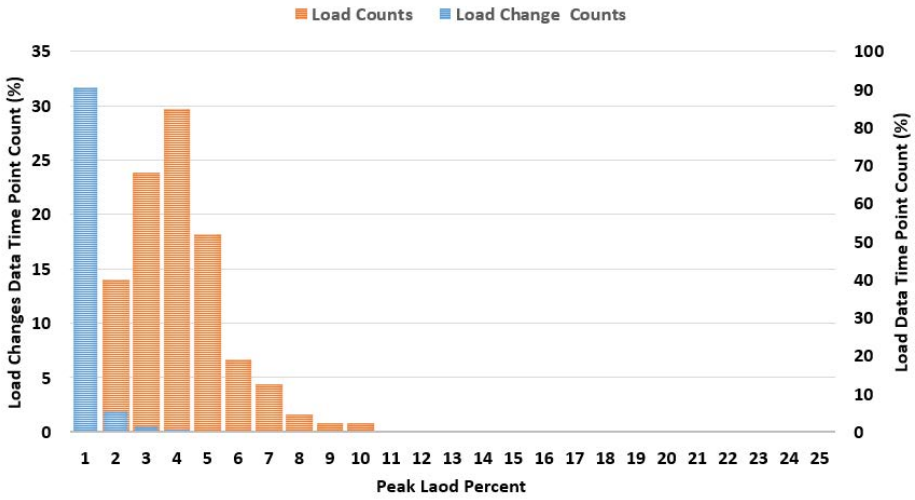


Figure 7.4: Load variability and load change variability based on normalized load values.

Table 7.2: Load variation

	Mean (kW)	SD (kW)	Relative SD (%)
House 1	0.000	0.045	0.11
House 2	0.000	0.070	0.64
House 3	0.000	0.113	0.98
House 4	0.000	0.110	0.89
House 5	0.000	0.046	0.56
Aggregate	0.000	0.184	0.68

Table 7.3: Load variation for different frequency

	Mean (kW)	SD (kW)	Relative SD (%)
$f = 1$ Hz	0.000	0.184	0.68
$f = 0.2$ Hz	0.000	0.425	1.56
$f = 0.1$ Hz	0.000	0.604	2.22

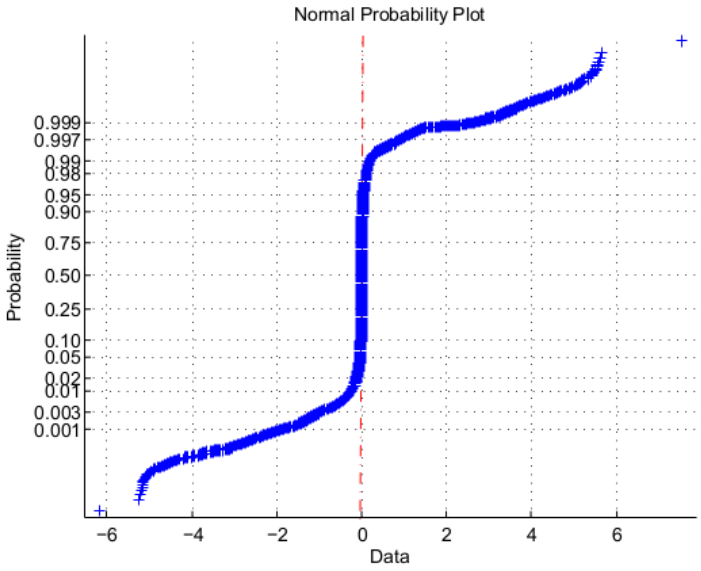


Figure 7.5: Output of `normplot` for the 1 second data

data, of the MATLAB function `normplot`, whose purpose is to graphically assess whether the data could come from a normal distribution. More similar to the normal distribution, more linear the plot will be. We can see how the plot is almost linear.

7.8 Results, Discussions and Conclusions

We tested our algorithm for topology detection on the IEEE 33-bus distribution test feeder Parasher (2014), which is illustrated in the Figure 7.6. In this testbed, there are five switches (namely S_1, S_2, S_3, S_4, S_5) that can be opened or closed, thus leading to the set of 32 possible topologies $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_{32}$. Because of the ratio between the number of buses and the number of switches, some very similar topologies can occur (for example the topology where only S_1 is closed and the one in which only S_2 is closed). In the IEEE33-bus test case, Assumption 2.8.1 about line impedances does not hold, making the test condition more realistic. Each bus of the network represents an aggregate of five houses, whose power demand is described by the statistical Gaussian model derived in section 7.7, dependent on the sampling frequency, and whose characterization is reported in Table 7.3. For what concern the measurement noise, we assume that the buses are endowed with high precision devices, the μPMU (see PQube3). PMU measurements are effected by Gaussian noise such that $TVE \leq 0.05\%$, based on the PMU manufacturer test information. It is also comply the IEEE standard C37.118.1-2011 for PMUs (IeeeStandard).

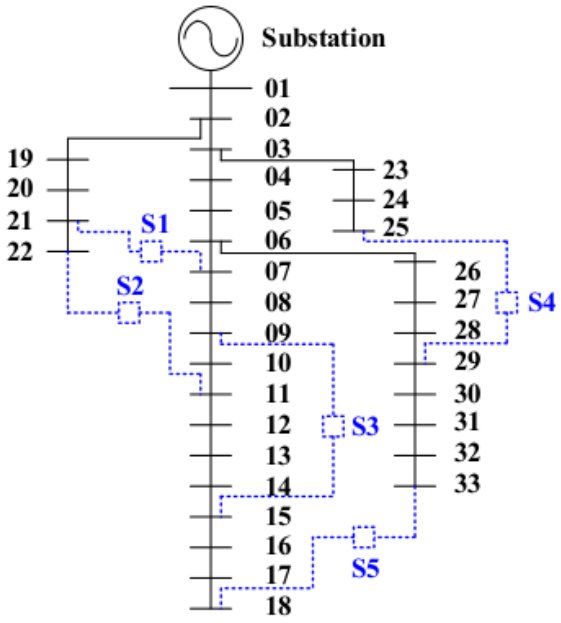


Figure 7.6: Schematic representation of the IEEE33 buses distribution test case with the five switches

Trend vectors and noise treatment strategy

Here we provide a simulations that show the problems caused by the noise and that validate the strategy we use to overcome them. The overall time window simulated is of 1000 seconds, the measures are done is every 10 seconds and the topology transition ($f = 1/10$ sec^{-1}), from $w_1 = (1, 1, 1, 0, 1)$ to $w_2 = (1, 1, 0, 0, 1)$, happens at $t = 480$ sec. In Figure 7.7 we see what happens to the trend vector norm, while in Figure 7.8 we plot $\max \mathcal{C}$, when we are in a noiseless scenario with loads not time varying. We can see that the switch action instant shows itself clearly.

In the case with measurements noise and time varying loads, the trajectory of the trend vector norm, compared with *min_norm* numerically setted to 0.05, is reported in Figure 7.9. In Figure 7.10 instead we plot $\max \mathcal{C}$, without putting to zero the trend vectors whose norm is lower than *min_norm* before the projection. We see that there are several cluster of time instants in which $\max \mathcal{C}$ is greater than *min_norm*, and this leads to a number of false positive. In Figure 7.11 we plot $\max \mathcal{C}$, after we put to zero the trend vectors whose norm is lower than *min_norm* before the projection. We still have more than one cluster, but only one has a length of $T\frac{1}{f}$, thus revealing the switch action instant, and the validity of our strategy.

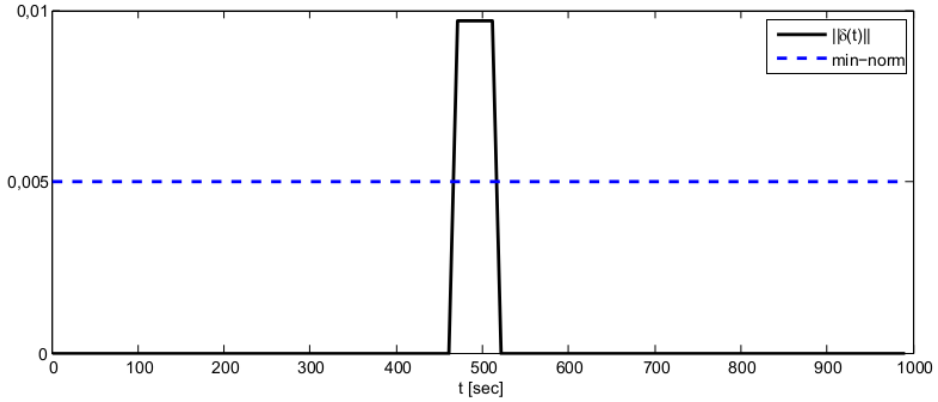


Figure 7.7: Trend vector norm with noiseless measurements and loads non time varying

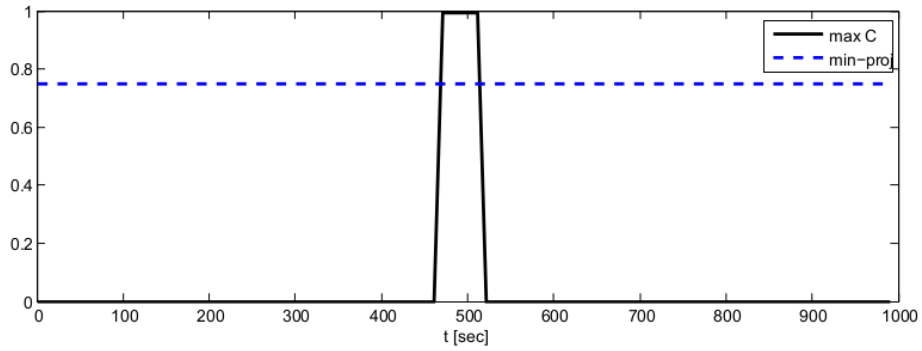


Figure 7.8: Trajectory of the $\max \mathcal{C}$ with noiseless measurements and loads non time varying

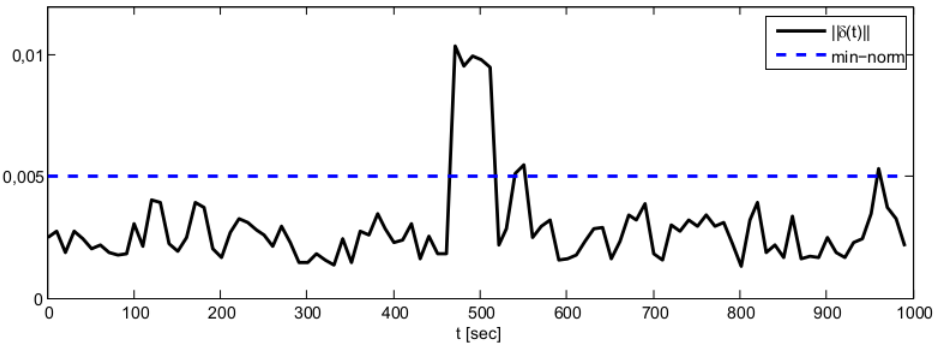


Figure 7.9: Trend vector norm with noisy measurements and time varying loads

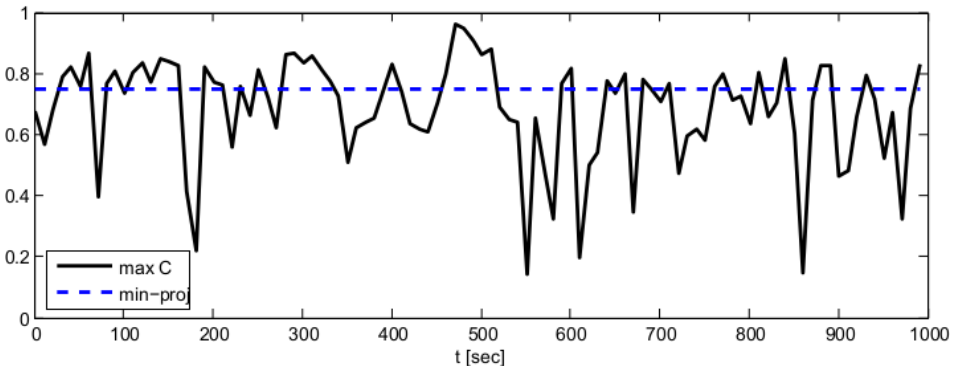


Figure 7.10: Trajectory of the $\max \mathcal{C}$ with noisy measurements and time varying loads

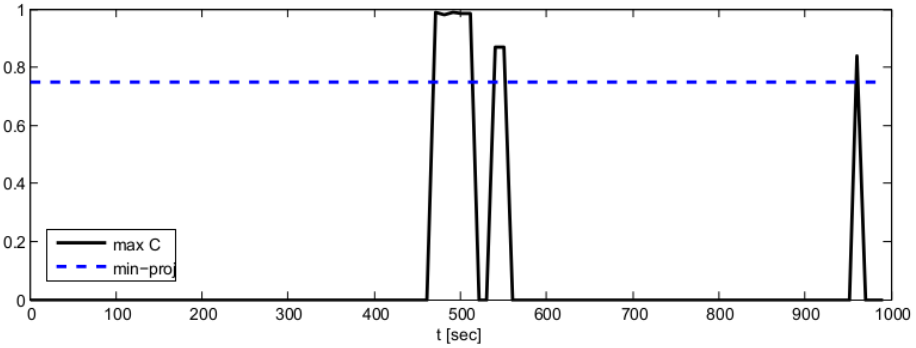


Figure 7.11: Trajectory of the $\max \mathcal{C}$ with noisy measurements and time varying loads after putting to zero the vectors whose norm is lower than *min_norm*

Simulation of the algorithm

Here we tested the entire switches monitoring algorithm, in different situations and with different placements of PMUs:

- $\mathcal{Y}_{33} = \{1, \ldots, 33\}$, where there is a PMU for every node;
- $\mathcal{Y}_{15} = \{3, 8, 9, 10, 12, 15, 16, 17, 18, 19, 21, 24, 25, 27, 30\}$, where there are 15 PMUs, almost one every two nodes;
- $\mathcal{Y}_7 = \{9, 12, 15, 18, 24, 27, 30\}$, where there are 7 PMUs, almost one every four nodes

$\mathcal{Y}_{15}$ and $\mathcal{Y}_7$ has been computed using Algorithm 2. The algorithm has been tested in each condition via 10000 Monte Carlo simulations..

Firstly we tested the algorithm in the ideal case without load variation nor measurement noise, resulting in no errors. Therefore, we can see that in the steady-state condition and in the absence of noise, the algorithm is extremely efficient and robust for the 33-bus test case. It also overcomes the linearization from Proposition 2.6.1 and the initial Assumption 2.8.1.

Secondly we added the measurement noise and different levels of load variation (or alternatively of measures frequency). The results are reported in Table 7.4, Table 7.5 and Table 7.6. The field "non detection" refers to the number of run in which the algorithm doesn't comprehend that there is a switching action, the field "wrong detection" refers to the number of times it detect a false action, while "decision errors" is the number of times the algorithm estimates a wrong switches status. Of course every time there is a wrong detection, we have a decision error, too. If we subtract the values of the second column to the values of the third, we find very small number, meaning that, once detected the exact action time, the algorithm works very well. The main challenge is therefore to detect the topology transition time. However our approach is very robust with the number of sensors: the percentage of errors with only 7 PMUs (where we have partial information on the grid state) is always worst than the one with 33 PMUs (where we have global information on the grid state) less than 2.5%, and probably with the best parameter tuning (*min_norm* and *min_proj*), it can still be improved.

The results show how the performance degrades with the frequency decrease, suggesting that the procedure is reliable only with high sampling frequencies. In fact, as soon as the topology change has occurred, the voltage profile variation follow the shape of the particular excited eigenvector in the library. Then, as time goes by, this shape become less recognizable due to the effect of noise and load variation. This is the price to be paid for not using additional statistical information.

Table 7.4: Simulation Results with 33 PMUs

SD [kV]	non detections	wrong detection	decision errors	total errors	perc. of errors (%)
0	0	50	50	100	1.00
0.68, ($f = 1$ Hz)	0	64	67	131	1.31
1.56, ($f = 0.2$ Hz)	17	131	152	300	3.00
2.22, ($f = 0.1$ Hz)	72	211	244	527	5.27

Table 7.5: Simulation Results with 15 PMUs

Relative SD (%)	non detections	wrong detection	decision errors	total errors	perc. of errors (%)
0	0	52	52	104	1.04
0.68, ($f = 1$ Hz)	0	73	76	149	1.49
1.56, ($f = 0.2$ Hz)	29	135	158	322	3.22
2.22, ($f = 0.1$ Hz)	74	213	252	539	5.39

Table 7.6: Simulation Results with 7 PMUs

Relative SD (%)	non detections	wrong detection	decision errors	total errors	perc. of errors (%)
0	0	56	56	112	1.12
0.68, ($f = 1$ Hz)	2	180	185	367	3.65
1.56, ($f = 0.2$ Hz)	31	199	209	441	4.41
2.22, ($f = 0.1$ Hz)	76	245	298	619	6.19

7.9 Chapter conclusions

We propose a novel strategy for the monitoring and the identification of switches action in a distribution grid. The novelty of this algorithm is the possibility of running it in real time and the fact that it can work satisfactorily also with a partial knowledge of the grid state (few phasorial voltages) without any other statistical characterization. Furthermore, it allows us, exploiting a voltage phasorial measurements, to understand both the time of the switches action and the new topology, by comparing the trend vector, built by data, with other vectors contained in a library, that represent the a-priori knowledge of the electrical network.

The algorithm has been tested in a truthful scenario, where both measurement noise and load variation had realistic characterization. In particular, we provide a simple but plausible model of the fast time scale load variation, validated using real field load data. The simulations show the algorithm behavior in different scenarios. In the ideal one, with static loads and perfect measurements devices, the algorithm gives no errors. When instead we have noisy PMUs and load variations, we still have satisfactory results, strengthened by the fact that the algorithm performance is very similar both in the case of one PMU per bus and of one PMU every four nodes.

8

Conclusions

In this dissertation we propose novel algorithms for the optimization, the control and the identification of smart grids.

In Chapter 4 we propose a control strategy that solves the OPRF problem, while in Chapter 5 we propose an algorithm that solves the OPF problem. Even if their aims are different, they are similar since they use local phasorial measurements and local communication between the agents to infer the global information (e.g. the losses function gradient) they need. Furthermore they can be implemented in a distributed way, avoiding the burden of computation to a central control unit. The algorithm proposed in Chapter 6, can be implemented without a central control unit, too. However, it needs a more complicated communication architecture between the agents, in order to play the game whose aim is to share the active power generation (see Figure 6.1). The algorithm stated in Chapter 7 needs instead a central unit, that collects all the data taken by the agents and infers the breakers status.

The former algorithms have in common the following features:

1. they work in real time;

2. they don't need any measurement nor any statistical model of the power demanded by the loads.

These characteristics make the algorithms proposed more suitable than others in the literature for distribution networks, where the behavior of a renewable source or of a load demand is less predictable and more intermittent than the behavior of a node in the transmission networks, which is the aggregate of several users. The use of the proposed algorithms could lead to remarkable improvement of a distribution network performance. Indeed the algorithms can react to (or detect the) new operating condition of the grid quickly exploiting the feedback that comes from the phasorial measurements. Thus, the algorithms implementation in real testbeds is an important future work, in order to prove their effectiveness.

An interesting future work is the study of the parallel execution of the algorithms, e.g. the simultaneous execution of the OPF algorithm and of the ORPF algorithm.

Beside the deployment of the DERs, a large use of storage devices is envisioned in the future distribution network. A very important challenge thus will be the integration and the management of these devices in the proposed control procedures. This integration is not trivial, since e.g. the resulting optimization problem would be no more static but dynamic, because we should take into account the charge and the discharge of the batteries.

Other important future developments are also the study of the algorithms behavior in presence of model errors or in presence of communication failure.

References

Andersson G. Modelling and analysis of electric power systems. *EEH-Power Systems Laboratory, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland*, 2004.

Arya V., Seetharam D., Kalyanaraman S., Dontas K., Pavlovski C., Hoy S., and Kalagnanam J. R. Phase identification in smart grids. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, pages 25–30, 2011.

Bai X., Wei H., Fujisawa K., and Wang Y. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30 (6-7):383–392, Jul.-Sep. 2008.

Baran M. E. and Wu F. F. Optimal sizing of capacitors placed on a radial distribution system. *IEEE Trans. Power Del.*, 4(1):735–743, January 1989.

Barchi G., Macii D., and Petri D. Synchrophasor estimators accuracy: A comparative analysis. *IEEE Transactions on Instrumentation and Measurement*, 62(5):963–973, 2013.

Barklund E., Pogaku N., Prodanovic M., Hernandez-Aramburo C., and Green T. C. Energy management in autonomous microgrid using stability-constrained droop control of inverters. *IEEE Transactions on Power Electronics*, 23(5):2346–2352, September 2008.

Bertsekas D. P. *Nonlinear programming.* Athena Scientific, Belmont (MA), 2nd edition, 1999.

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

Bolognani S., Cavraro G., and Zampieri S. Performance analysis of a distributed algorithm for dynamic reactive power compensation. In *Proc. 51st IEEE Conf. on Decision and Control (CDC)*, 2012.

Bolognani S. and Zampieri S. A distributed control strategy for reactive power compensation in smart microgrids. *Automatic Control, IEEE Transactions on*, 58(11): 2818–2833, Nov 2013. ISSN 0018-9286.

Bolognani S., Bof N., Michelotti D., Muraro R., and Schenato L. Identification of power distribution network topology via voltage correlation analysis. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 1659–1664. IEEE, 2013a.

Bolognani S., Carli R., Cavraro G., and Zampieri S. A distributed control strategy for optimal reactive power flow with power and voltage constraints. In *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pages 115–120. IEEE, 2013b.

Bolognani S., Carli R., Cavraro G., and Zampieri S. A distributed control strategy for optimal reactive power flow with power constraints. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 4644–4649. IEEE, 2013c.

Bolognani S., Carli R., Cavraro G., and Zampieri S. Distributed reactive power feedback control for voltage regulation and loss minimization. *Automatic Control, IEEE Transactions on*, 2015.

Bolognani S., Cavraro G., and Zampieri S. A distributed feedback control approach to the optimal reactive power flow problem. In *Control of Cyber-Physical Systems*, pages 259–277. Springer International Publishing, 2013d.

Boyd S. P. and Vandenberghe L. *Convex optimization*. Cambridge university press, 2004.

Cain M. B., O'Neill R. P., and Castillo A. History of optimal power flow and formulations. *FERC Staff Technical Paper*, 2012.

Carpentier J. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 3(8):431–447, 1962.

Carvalho P. M. S., Correia P. F., and Ferreira L. A. F. Distributed reactive power generation control for voltage rise mitigation in distribution networks. *IEEE Transactions on Power Systems*, 23(2):766–772, 2008.

Cavraro G., Arghandeh R., Barchi G., and von Meier A. Distribution networ topology detection with time-series measurements. In *Innovative Smart Grid Technologies (ISGT), 2015 IEEE PES*. IEEE, 2015.

Cavraro G. and Badia L. A game theory framework for active power injection management with voltage boundary in smart grids. In *European Control Conference ECC 2013*, 2013.

Chandorkar M. C., Divan D. M., and Adapa R. Control of parallel connected inverters in standalone ac supply systems. *IEEE Transactions on Industry Application*, 1993. Submitted.

Chen Y., Fuller J., Diao R., Zhou N., Huang Z., and Tuffner F. The influence of topology changes on inter-area oscillation modes and mode shapes. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–7. IEEE, 2011.

Chiradeja P. and Ramakumar R. An approach to quantify the technical benefits of distributed generation. *Energy Conversion, IEEE Transactions on*, 19(4):764–773, 2004.

Costabeber A., Erseghe T., Tenti P., Tomasin S., and Mattavelli P. Optimization of micro-grid operation by dynamic grid mapping and token ring control. In *Proc. 14th European Conf. on Power Electronics and Applications (EPE)*, Birmingham, UK, 2011.

Dall'Anese E., Zhu H., and Giannakis G. B. Distributed optimal power flow for smart microgrids. *Smart Grid, IEEE Transactions on*, 4(3):1464–1475, 2013.

De Brabandere K., Vanthournout K., Driesen J., Deconinck G., and Belmans R. Control of microgrids. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–7, June 2007.

Devane E. and Lestas I. Stability and convergence of distributed algorithms for the opf problem. In *52nd IEEE Conference on Decision and Control*, 2013.

Dorfler F. and Bullo F. Kron reduction of graphs with applications to electrical networks. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 60(1): 150–163, 2013.

Erseghe T. Distributed optimal power flow using admm. *Power Systems, IEEE Transactions on*, 29(5):2370–2380, Sept 2014. ISSN 0885-8950.

Farivar M., Neal R., Clarke C., and Low S. Optimal inverter VAR control in distribution systems with high PV penetration. In *Proc. of the IEEE Power and Energy Society General Meeting*, July 2012.

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

Grijalva S. and Tariq M. U. Prosumer-based smart grid architecture enables a flat, sustainable electricity industry. In *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*, pages 1–6. IEEE, 2011.

Guarnieri M. and Malesani G. *Elementi di elettrotecnica: reti elettriche*. Edizioni progetto, 1998.

Harris Williams C. Transmission & distribution infrastructure. In *A Harris Williams & Co. White Paper*, Summer 2014.

Hoffman R. Practical state estimation for electric distribution networks. In *Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES*, pages 510–517. IEEE, 2006.

Huneault M., Galiana F., and St Bruno Q. A survey of the optrial power flow literature. *IEEE Transactions on Power Systems*, 6(2), 1991.

IeeeStandard . Ieee standard for synchrophasor measurements for power system. *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*.

Jiayi H., Chuanwen J., and Rong X. A review on distributed energy resources and microgrid. *Renewable and Sustainable Energy Reviews*, 12(9):2472–2483, 2008.

Keane A., Ochoa L. F., Vittal E., Dent C. J., and Harrison G. P. Enhanced utilization of voltage control resources with distributed generation. *IEEE Transactions on Power Systems*, 26(1):252–260, 2011.

Kekatos V. and Giannakis G. Joint power system state estimation and breaker status identification. In *North American Power Symposium (NAPS), 2012*, pages 1–6, Sept 2012.

Kersting W. H. Radial distribution test feeders. In *IEEE Power Engineering Society Winter Meeting*, volume 2, pages 908–912, January 2001.

Kezunovic M. Monitoring of power system topology in real-time. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 10, pages 244b–244b. IEEE, 2006.

Korres G. N. and Manousakis N. M. A state estimation algorithm for monitoring topology changes in distribution systems. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE, 2012.

Kundu S. and Hiskens I. A. Distributed control of reactive power from photovoltaic inverters. In *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, pages 249–252. IEEE, 2013.

Lam A. Y. S., Zhang B., Dominiguez-Garcia A., and Tse D. Optimal distributed voltage regulation in power distribution networks. *arXiv*, 1204.5226 [math.OC], 2012a.

Lam A., Zhang B., and Tse D. N. Distributed algorithms for optimal power flow problem. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 430–437. IEEE, 2012b.

Lasseter R. H. Microgrids. In *Power Engineering Society Winter Meeting, 2002. IEEE*, volume 1, pages 305–308. IEEE, 2002.

Lavaei J. and Low S. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 2012.

Lavaei J., Rantzer A., and Low S. H. Power flow optimization using positive quadratic programming. In *Proc. 18th IFAC World Congr.*, 2011.

Lesieutre B. C., Molzahn D. K., Borden A. R., and DeMarco C. L. Examining the limits of the application of semidefinite programming to power flow problems. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1492–1499. IEEE, 2011.

Li Y. W. and Kao C.-N. An accurate power control strategy for power-electronics-interfaced distributed generation units operating in a low-voltage multibus microgrid. *IEEE Transactions on Power Electronics*, 2009. Submitted.

Low S. H. Convex relaxation of optimal power flow (part i, ii). *IEEE Trans. Control of Network Systems*, 1(Part I: 1; Part II: 2):Part–I, 2014.

Lueken C., Carvalho P. M., and Apt J. Distribution grid reconfiguration reduces power losses and helps integrate renewables. *Energy Policy*, 48(0):260 – 273, 2012. ISSN 0301-4215. URL http://www.sciencedirect.com/science/article/pii/S0301421512004351. Special Section: Frontiers of Sustainability.

Mallada E. and Tang A. Dynamics-aware optimal power flow. In *52nd IEEE Conference on Decision and Control*, 2013.

Miller K. S. On the inverse of the sum of matrices. *Mathematics Magazine*, pages 67–72, 1981.

Monticelli A. Electric power system state estimation. *Proceedings of the IEEE*, 88(2): 262–282, 2000.

Nash J. The bargaining problem. *Econometrica*, 28:152–155, 1950.

Owen G. *Game Theory*. New York: Academic, 3rd edition, 2001.

Parasher R. Load flow analysis of radial distribution network using linear data structure. *arXiv preprint arXiv:1403.4702*, 2014.

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

PQube3 . New low cost, high-precision power quality, energy and environment monitoring. In *Power Standard Lab. Inc. (PSL),http://www.powerstandards.com/*.

Rahimi F. and Ipakchi A. Overview of demand response under the smart grid and market paradigms. In *Innovative Smart Grid Technologies (ISGT), 2010*, pages 1–7. IEEE, 2010.

Rajkumar R., Lee I., Sha L., and Stankovic J. Cyber-physical systems: The next computing revolution. In *Design Automation Conference (DAC), 2010 47th ACM/IEEE*, pages 731–736, June 2010.

Saad W., Han Z., and Poor H. V. Coalitional game theory for cooperative micro-grid distribution networks. In *Proc. IEEE ICC Workshops*, 2011.

Sevlian R. and Rajagopal R. Short term electricity load forecasting on varying levels of aggregation. *arXiv preprint arXiv:1404.0058*, 2014.

Sharon Y., Annaswamy A. M., Motto A. L., and Chakraborty A. Topology identification in distribution network with limited measurements. In *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, pages 1–6. IEEE, 2012.

Simpson-Porco J. W., Dorfler F., and Bullo F. Droop-controlled inverters in microgrids are kuramoto oscillators. In *IFAC Workshop on Distributed Estimation and Control in Networked Systems, Santa Barbara, California, USA, pages 264-269, September 2012*, 2012.

Šulc P., Backhaus S., and Chertkov M. Optimal distributed control of reactive power via the alternating direction method of multipliers. *arXiv*, 1310.5748 [math.OC], 2013.

Tenti P., Costabeber A., Mattavelli P., and Trombetti D. Distribution loss minimization by token ring control of power electronic interfaces in residential microgrids. *IEEE Trans. Ind. Electron.*, 59(10):3817–3826, October 2012.

Turitsyn K., Sulc P., Backhaus S., and Chertkov M. Local control of reactive power by distributed photovoltaic generators. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 79–84. IEEE, 2010.

Turitsyn K., Sulc P., Backhaus S., and Chertkov M. Options for control of reactive power by distributed photovoltaic generators. *Proceedings of the IEEE*, 99(6): 1063–1073, 2011.

Utkin V. I. *Sliding modes and their application in variable structure systems.* Mir Publishers, 1978.

Uzawa H. *Studies in linear and nonlinear programming*, chapter The Kuhn-Tucker theorem in concave programming, pages 32–37. Stanford University Press, 1958.

Villacci D., Bontempi G., and Vaccaro A. An adaptive local learning-based methodology for voltage regulation in distribution networks with dispersed generation. *Power Systems, IEEE Transactions on*, 21(3):1131–1140, 2006.

von Meier A., Culler D., McEachern A., and Arghandeh R. Micro-synchrophasors for distribution systems. In *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*, Feb 2014a.

Von Meier A. *Electric power systems: a conceptual introduction.* John Wiley & Sons, 2006.

von Meier A., Brown M. L., Arghandeh R., Mehrmanesh L., Cibulka L., and Russ B. Electric distribution advanced monitoring plan (amp). In *Electric distribution Advanced Monitoring Plan*, 2014b.

Wang J. and Elia N. A control perspective for centralized and distributed convex optimization. In *Proceedings of the 50th IEEE Conference on Decision and Control (CDC)*, pages 3800–3805, Orlando, FL (USA), December 2011.

Weaver W. Dynamic energy resource control of power electronics in local area power networks. *IEEE Trans. on Power Electronics*, 26(3):852–859, March 2011.

Weaver W. and Krein P. Game-theoretic control of small-scale power systems. *IEEE Trans. on Power Delivery*, 24(3):1560–1567, July 2009.

Weng Y., Faloutsos c., and Ilić M. Data-driven topology estimation. In *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, 2014.

Wiel S. V., Bent R., Casleton E., and Lawrence E. Identification of topology changes in power grids using phasor measurements. *Applied Stochastic Models in Business and Industry*, 2014.

Zhang B. and Tse D. Geometry of injection region of power networks. *IEEE Transactions on Power Systems*, 28(2):788–797, May 2013.

Zhang B., Dominguez-Garcia A. D., and Tse D. A local control approach to voltage regulation in distribution networks. In *Proc. of the North American Power Symposium (NAPS)*, Manhattan, KS, September 2013.

Zhao B., Guo C. X., and Cao Y. J. A multiagent-based particle swarm optimization approach for optimal reactive power dispatch. *IEEE Trans. Power Syst.*, 20(2):1070–1078, May 2005.

Zimmerman R. D., Murillo-Sánchez C. E., and Thomas R. J. MATPOWER: steady-state operations, planning and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, February 2011.

Acknowledgments

First of all, I would like to thank my advisor, Prof. Sandro Zampieri, for his guidance during the PhD. Sandro is a great scientist. What I have particularly appreciated and tried to learn working with him was his solid methodological attitude. I am very grateful also to my co-advisor Prof. Ruggero Carli. Ruggero is very skilled and brilliant, and guided me with passion. We had fun working together.

I would like to express my sincere gratitude to Prof. Alexandra von Meier and Prof. Luca Schenato: they gave me the opportunity of visit the California Institute for Energy and Environment (CIEE) at U.C. Berkeley, and they helped me during my abroad period.

My PhD years wouldn't have been so pleasant without all the officemates, the collaborators and the friends I had, in Padova and in Berkeley. I shared with them a important, wonderful part of my life and I'll never forget them, our coffee breaks, our lunch together in the canteen or in Euclid street, and our discussion.

I would like to thank also my parents Paolo and Leopoldina, and my sister Alessandra. They always supported me and my family during these years, with love and unconditioned confidence. Special thanks also to Antonio and Marisa for their help.

Finally I want to thank my girlfriend Elisa and our son Isaac. They are my strength, my happy thought. They are always by my side, and they are the main reasons for my strength and my desire of improving.