Frédéric Olivier
Advisor: Prof. Damien Ernst

Solutions for Integrating Photovoltaic Panels Into Low-voltage Distribution Networks

2018

Faculty of Applied Sciences
Liège Université

First printing, July 2018

If I have seen further,

it is by standing on the shoulders of giants.

ISAAC NEWTON

Abstract

As more PhotoVoltaic (PV) units are being installed, some low-voltage (LV) distribution networks have already attained their maximum hosting capacity, i.e. the maximum amount of distributed energy resources that they can accommodate during regular operations without suffering problems, such as overvoltages. This thesis presents different solutions to increase the hosting capacity and to prevent the disconnection of PV units, without resorting to network reinforcements.

The first action to increase the hosting capacity of an LV network with a high penetration of PV panels is to better balance production on its three phases. Indeed, the hosting capacity is reduced by the imbalance caused by the currents consumed by single-phase household appliances, or produced by single-phase distributed generation units. To do so, the Distribution System Operators (DSOs) have to know to which phase of their networks the houses and the PV units are connected. In other words, they need to know the phase identification of the meters. In this thesis, we propose two novel algorithms to identify the phases of the smart meters. Both algorithms solely rely on the correlation between the voltage measurements of the smart meters to solve the phase identification problem. The first one is dedicated to the identification of three-phase smart meters and uses graph theory as well as the notion of maximum weight spanning tree to associate the smart meters that have the most correlated voltages. The second algorithm overcomes the main drawback of the first one, its inability to identify both single-phase and three-phase smart meters. Our algorithm improves the quality of the clustering by taking into account the underlying structure of the LV distribution networks beneath the voltage measurements without a priori knowledge on the topology of the network. The performance of this algorithm is compared using real measurements to those of the constrained k-means clustering method, which has been previously used for phase identification.

The second action to increase the hosting capacity is to control the power flows inside the network. This is called Active Network Management. We propose a distributed scheme that adjusts the reactive and active power output of inverters to prevent or alleviate the overvoltage problems that might arise from the integration of photovoltaic panels in LV networks. The proposed scheme is model free and makes use of limited communication between the controllers, in the form of a distress signal, only during emergency conditions. It prioritizes the use of reactive power, while active power curtailment is performed only as a last resort. The behaviour of the scheme is studied using balanced three-phase dynamic simulations on a single low-voltage feeder and on a larger network composed of 14 low-voltage feeders. This control scheme is then extended to the case of unbalanced three-phase four-wire distribution networks with single- and/or three-phase inverters. It works by first partitioning the inverters into four groups, three for the single-phase inverters (one for each phase), and one for the three-phase ones. Each group then independently applies a distributed algorithm similar to the one previously presented. Its performance is compared to those of two reference schemes, an on-off algorithm that models the default behaviour of PV inverters when there is an overvoltage, and the other one based on an unbalanced OPF.

Finally, as a way to implement active network management at the level of LV distribution, we introduce the concept of electricity prosumer communities, which are groups of people producing, sharing and consuming electricity locally. We focus on building a rigorous mathematical framework in order to formalise sequential decision-making problems that may soon be encountered within electricity prosumer communities. After introducing our formalism, we propose a set of optimisation problems reflecting several types of theoretically optimal behaviours for energy exchanges between prosumers. We then discuss the advantages and disadvantages of centralised and decentralised schemes and provide illustrations of decision-making strategies, allowing a prosumer community to generate more distributed electricity (compared to commonly applied strategies) by mitigating overvoltages over a low-voltage feeder. We finally investigate how to design distributed control schemes that may contribute to reaching (at least partially) the objectives of the community, by resorting to machine learning techniques to extract, from centralised solution(s), decision-making patterns to be applied locally.

Résumé

Avec l'accroissement du nombre de panneaux photovoltaïques (PV), certains réseaux de distribution basse tension (BT) atteignent leur capacité d'accueil maximale, c'est-à-dire la quantité maximale de production décentralisée d'électricité qu'ils peuvent accueillir sans souffrir de problèmes opérationnels, comme des surtensions. Cette thèse présente différentes solutions pour augmenter la capacité d'accueil et prévenir la déconnexion des installations PV, sans recourir à un renforcement du réseau.

La première action pour augmenter la capacité d'accueil d'un réseau BT souffrant d'une forte pénétration des panneaux PV est de mieux équilibrer la production sur ses trois phases. En effet, la capacité d'accueil est réduite par le déséquilibre des réseaux de distribution BT qui est causé par les courants des appareils électroménagers monophasés ou des unités de production décentralisées monophasées. Pour ce faire, les gestionnaires de réseaux de distribution (GRD) doivent savoir à quelle phase de leurs réseaux les maisons et les installations PV sont connectées, c'est-à-dire qu'ils doivent connaître l'identification de phase des compteurs. Dans cette thèse, nous proposons deux nouveaux algorithmes pour identifier les phases des compteurs communicants. Les deux algorithmes reposent uniquement sur la corrélation entre les mesures de tension des compteurs intelligents pour résoudre le problème d'identification de phase. Le premier est dédié à l'identification des compteurs intelligents triphasés, et utilise la théorie des graphes ainsi que la notion d'arbre de poids maximum pour associer les compteurs qui ont les tensions les plus corrélées. Le second algorithme surmonte l'inconvénient principal du premier, son incapacité à identifier à la fois les compteurs monophasés et les compteurs triphasés. L'algorithme proposé améliore la qualité du *clustering* en prenant en compte la structure sous-jacente des réseaux de distribution BT sans avoir préalablement connaissance de la topologie du réseau. Les performances de cet algorithme sont comparées, à l'aide de mesures réelles, à celles de la méthode de *constraint k-means clustering*, qui a déjà été utilisée pour l'identification de phase.

La deuxième stratégie est de contrôler les flux d'énergie dans le réseau, connue sous le nom de gestion active des réseaux. Nous proposons un schéma de contrôle distribué qui ajuste la puissance réactive et active des onduleurs pour prévenir ou atténuer les problèmes de surtension qui pourraient survenir avec l'intégration des panneaux photovoltaïques dans les réseaux BT. Le schéma proposé ne nécessite pas de modèle détaillé du réseau et fait appel à une communication limitée entre les contrôleurs, sous la forme d'un signal de détresse envoyé uniquement en cas d'urgence. Il donne la priorité à l'utilisation de la puissance réactive, tandis que la réduction de la puissance active n'est effectuée qu'en dernier recours. Le comportement du schéma est illustré grâce à des simulations dynamiques triphasées équilibrées sur un seul départ basse tension et sur un réseau plus large composé de 14 départs basse tension. Ce schéma de commande est ensuite étendu aux réseaux de distribution triphasés non équilibrés à quatre conducteurs avec onduleurs monophasés et/ou triphasés. Il fonctionne en divisant d'abord les onduleurs en quatre groupes, trois pour les onduleurs monophasés (un pour chaque phase) et un pour les onduleurs triphasés. Chaque groupe applique ensuite indépendamment un algorithme distribué similaire à celui introduit plus haut. Ses performances sont comparées à celles de deux schémas de référence, un algorithme on-off qui modélise le comportement par défaut des onduleurs PV en cas de surtension, et l'autre, basé sur un écoulement de charge optimal non équilibré.

Enfin, pour mettre en place une gestion active au niveau des réseaux de distribution BT, nous introduisons les communautés de prosommateurs d'électricité, qui sont des groupes de personnes produisant, partageant et consommant de l'électricité localement. Nous nous concentrons sur la construction d'un cadre mathématique rigoureux afin de formaliser les problèmes de prise de décision séquentielle qui y seront rencontrés. Après avoir introduit notre formalisme, nous proposons un ensemble de problèmes d'optimisation reflétant plusieurs types de comportements théoriquement optimaux pour les échanges d'énergie entre prosommateurs. Nous discutons ensuite les avantages et les inconvénients des schémas centralisés et décentralisés et fournissons des illustrations de stratégies de prise de décision, permettant à une communauté de prosommateurs de générer plus d'électricité de manière décentralisée (par rapport aux stratégies couramment appliquées) en atténuant les problèmes de surtensions dans le réseau. Nous étudions enfin comment concevoir des schémas de contrôle distribués qui peuvent atteindre (du moins partiellement) les objectifs de la communauté, en recourant à des techniques d'apprentissage pour extraire, à partir de solutions centralisées, des modèles de prise de décision à appliquer localement.

Remerciements

Le doctorat est une aventure de longue haleine, un parcours semé d'embuches et de doutes, mais aussi de moments extrêmement positifs et valorisants. Il est le fruit d'une réflexion aussi bien personnelle que partagée, et je souhaiterais remercier les nombreuses personnes avec qui j'ai partagé cette aventure.

Merci à mon promoteur, Damien Ernst, qui m'a proposé cette thèse et a été présent quand j'en avais vraiment besoin, merci pour sa sagacité et ses remarques pertinentes sur mon travail.

Merci à Raphaël pour nos innombrables discussions, qu'elles soient techniques ou philosophiques. Elles m'ont permis d'ouvrir le champ des possibles.

Merci aux personnes avec qui j'ai eu la chance de cosigner des articles scientifiques : Petros Aristidou, Bertrand Cornélusse, Antoine Dubois, Damien Ernst, Alexis Fabre, Raphaël Fonteneau, Pierre Geurts, María-Emilia Hervás, Ramón López-Erauskin, Daniele Marulli, Sébastien Mathieu, Antonio Sutera, Thierry Van Cutsem, et Antoine Wehenkel. Chacun d'entre eux a apporté sa pierre à l'édifice.

Merci aux membres de mon Jury, Emmanuel De Jaeger, Damien Ernst, Raphaël Fonteneau, Jean-Claude Maun, Luc Warichet et Louis Wehenkel, de s'être intéressé à mon travail.

Merci à mes collègues qui ont mis une bonne ambiance dans l'équipe, et qui avaient toujours une idée intéressante quand j'étais bloqué sur un sujet.

Merci à Yannick avec qui j'ai partagé mon bureau pendant plusieurs années.

Merci aux différents Gestionnaires de Réseau de Distribution et de Transport avec qui j'ai discuté et collaboré. Je les remercie également pour toutes les données qu'ils m'ont transmises à propos de leurs réseaux basse-tension.

Merci au F.R.S.-FNRS de m'avoir octroyé une bourse FRIA.

Merci à ma maman et Étienne pour leur accueil lorsque j'avais besoin de m'isoler pour me concentrer sur mon travail.

Merci à ma famille et à mes amis pour leur soutien et leurs encouragements, en particulier mon papa qui m'a donné le goût d'aller plus loin, et ma maman qui m'a montré qu'il fallait être bien entouré pour le faire.

Vous me faites grandir !

Contents

Introduction

1 The rise of distributed energy resources and the power system

Everywhere around the world, governments take actions to favour the development of green energy. In many countries, especially European ones, this development was initially favoured by generous subsidies. However, the cost of renewable energy sources has now dropped, making it competitive with non-renewable generation in many parts of the world. In particular, the cost of PhotoVoltaic (PV) panels has been continuously decreasing during the last few decades, which, in certain sunny countries, makes it possible to offer a price for photovoltaic energy that is largely competitive with that of the grid.[1]

[1] EPIA (*Connecting the Sun: Solar photovoltaics on the road to large-scale grid integration*)

Renewable energy sources and distributed energy resources usually come hand in hand when thinking about small-scale PV units or small wind turbines. However, they should not be mistaken, as they are not one and the same.

Indeed, Distributed Energy Resources (DERs) are composed of three parts: distributed generation, storage and demand response or flexibility.

Distributed generation (DG) is composed of relatively small-scale electricity production units connected to the distribution network, such as PV panels, small wind turbines, natural gas fuel cells, combined heat and power (CHP), diesel and gasoline engines, gas turbines, small hydro impoundments or run of the river hydro plants, etc. Even if the latest developments of DG are strongly linked to renewable energy sources such as PV panels, DG should not be confused with renewable energy sources, as large off-shore wind farms, for example, can hardly be called distributed generation, but are considered renewable. In addition to renewability, DG can be polarized along another axis: schedulability, which will impact the way DG is integrated in the power system. For example, small hydro impoundments which are renewable and schedulable are integrated differently in the system to PV panels, whose power production depends on solar irradiance. Properly integrated, DG can allow the generation of power close to the consumption points, alleviate grid constraints such as congestions, reduce power losses, etc. If it is microgrid enabled, it can provide a reliable power supply in the event of a default of the power system. If it is renewable, it can help reach the target many countries have set in terms of percentage of production that should be renewable in order to reach the desired level of decarbonization.

Storage can be seen as a hybrid between production and consumption as it can both inject power into the network and withdraw it when needed. There are several types of storage, such as batteries, capacitors, thermal or

high-pressure air storage, flywheels, etc. which can be used to match peak load, compensate for the variations of non-schedulable production units (PV, wind...). They can reduce or offset investment in peak production and alleviate grid constraints.

Finally, demand response (DR) is the last part of DERs as it allows for shedding or shifting consumption. It can be used to match demand to supply, to reduce demand in the event of a plant shut-down, alleviate grid constraints, etc. DR can be implemented via direct load control, rolling blackouts, or time-of-use tariffs, etc.

DERs are not new concepts as they have been deployed for a long period of time, even successfully accommodated on a small scale and occasional basis. What has changed is the number of DERs that are installed in the networks, as DG is strongly connected to the rise of renewable energy sources. This is the result of the development of new technologies and the continuing decrease in their prices. Moreover, DG has specific characteristics that, if massively deployed, become no longer negligible compared to those of the bulk power system, e.g. the low inertia of inverter-based production compared to that of rotating machines. The massive installation of power electronics-interfaced DG changes the behaviour of the system on all voltage levels.

The improper integration of DERs in electricity networks can lead to numerous problems.

First of all, DG changes the loading of the feeders. If the penetration level is low, DG offsets the load, thus decreasing the section loading. However, if the penetration level is high, the production can exceed the load and a reverse power flow can occur. Distribution networks were not designed with two-way power flows in mind, so the coordination of protection devices might be erroneous. For example, overcurrent protections might not be suited for reverse power flows and mistakenly trip.[2] Regarding the losses, a low penetration level can reduce them as power is generated close to the consumption point. If the penetration level is high and the feeder is exporting power, the losses can increase. A reverse power flow can also lead to unacceptable voltage rises that significantly endanger the security of the electrical network. Some basic protection schemes are utilized to protect the devices from over/under-voltages, such as automatically switching off the PV units when the voltage at the terminal is too high.[3] Norms impose the disconnection of a PV when the ten-minute average voltage at its connection point is higher than 110% of the nominal voltage, or when the instantaneous voltage surpasses 115% of the nominal voltage (cf. EN 50160). In networks with high penetration of PV units, they can be repeatedly switched off during a sunny day, leading to a loss of earnings for the producer and a loss in renewable energy production, calling into question their profitability.

[2] Barker and Mello ("Determining the impact of distributed generation on power systems. I. Radial distribution systems")

[3] IEEE (*IEEE 1547 standard for interconnecting distributed resources with electric power systems IEEE Std 1547-2003.*)

The rapid variations of unschedulable DG production cause voltage fluctuations, that leads to the prohibitive operations of capacitor banks, On-Load Tap Changer (OLTC) and voltage regulators, considerably reducing their lifetime.

Inverter-based DG can inject harmonics into the network, i.e. currents with frequencies that are multiples of the fundamental frequency. They con-

tribute to transformer overheating, degradation of meter accuracy and communication system interference.

Furthermore, DG must be equipped with an algorithm able to identify any fault in the power system and disconnect the unit in order to prevent islanding, i.e. a DG unit producing power in a section of the network that is no longer connected to the rest of the power system and maintaining wires live, thus endangering the personnel in charge of repairs, etc.

Finally, single-phase connected DG can increase the imbalance between the phases of the feeder. Some imbalance is naturally present in distribution networks, especially the low-voltage ones, because household appliances and small-scale DG are generally single-phase, but strong imbalances can damage three-phase motors, and more easily create under- or overvoltage situations.

This section is not intended to be exhaustive, but it details a few of the main problems arising when integrating DG in distribution networks. Those problems are tightly linked to the concept of hosting capacity, i.e. the maximum amount of DG that can be connected to a network without having the above-mentioned issues becoming prohibitive.

In this thesis, we will focus on the integration of DG in low-voltage (LV) distribution networks, the part of the power system with the lowest voltage (400 V), connecting the majority of customers[4] to the rest of the power system. We will also limit ourselves to the study of PV panels since it is a type of DG commonly found in LV networks as they are primarily installed in private residences and small companies, either on their land, or on their buildings. However, the work done in this thesis could be extended to other small-scale inverter-based DG.

[4] Some large industrial customers are directly connected to higher voltages.

At the heart of this thesis is the desire to find solutions to increase the hosting capacity of LV networks and limit the occurrences of overvoltages that force PV units to disconnect from the network and create loss of earnings for the owners. This, in turn, can considerably slow down the energy transition if DERs are not properly integrated.

2 *How to increase the hosting capacity?*

Distribution networks are traditionally managed according to the *fit and forget* doctrine. This consists of ensuring that investments made in network infrastructure (cables, lines, transformers, etc.) make it possible to avoid infringing operational limits (i.e. avoiding congestion or voltage problems) in all circumstances, without requiring permanent monitoring and control of energy flows or voltages in certain network locations.

2.1 Network reinforcement

The classical approach to increase the hosting capacity of networks is to rely on hefty investments to upgrade and reinforce the networks, in order to continue to manage them historically. However, the significant increase in the number of DGs connected to the distribution networks calls into question the maintenance of this doctrine because it would be much too expensive. Indeed, the current electricity grid is the result of years of evolution and con-

tinuous investment. Updating the network would present a significant cost in terms of infrastructure, personnel and time. ORES, one of the largest Belgian DSOs which supplies energy to more than 1.3 million households and businesses, estimated the cost of updating its network to be several hundred million euros.[5]

[5] ORES ("Les coulisses du photovoltaïques")

For this reason, many companies and researchers are looking at better ways to use existing equipment by developing and designing flexible and inexpensive control schemes to limit those investments and increase the hosting capacity of the networks.[6] This second strategy can be pursued in two different ways:

[6] J. P. Lopes et al. ("Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities")

1. By better balancing the production over the three phases of the network.
2. By controlling the power injections in the network,

   by controlling the production of PV panels,

   by storing part of the electricity production in batteries,

   by promoting self-consumption at critical moments.

The following sections will further discuss these options.

2.2 Balancing the phases of the network

LV distribution networks are intrinsically unbalanced due to the currents that are consumed by single-phase household appliances, or produced by single-phase distributed generation units. This imbalance reduces the hosting capacity of the network, i.e. the maximum amount of distributed power generation that can be connected without violating the operational constraints. Thus, the first action to increase the hosting capacity of an LV network with high penetration of photovoltaic (PV) panels[7] is to better balance production on the three phases of the network. To do so, the Distribution System Operators (DSOs) have to know to which phase of their networks the houses and the PV units are connected, i.e. they need to know the phase identification of the meters. Without it, it is not possible to know from the measurements which phase of the network is the most loaded and how they can better balance their network by changing the phase(s) to which their customers are connected. This can be achieved by solving the phase identification problem, i.e. associating the phases of the network to those of the smart meter. It is equivalent to clustering the measurements into three groups, one for each phase of the network, keeping in mind that the physical nature of the problem imposes constraints on the clustering process.

[7] Walling et al. ("Summary of distributed resources impact on power delivery systems")

The phase information is also crucial for multiple reasons: in the absence of the latter, (i) phase measurements such as line-to-neutral voltage cannot be used to identify key features of the network. For example, it is impossible to display a voltage profile of the feeder for each phase. (ii) For research purposes, it is impossible to properly use the measurements to perform load flow on a three-phase model of the network and compare the measured and simulated voltages. To obtain simulations accurately reflecting real operations, the phase information is needed in order to know to which phase of the network the active and reactive power measurements correspond.

Two types of solution oppose each other to solve the phase identification problem: manual ones, which require a technician to proceed, manually, to the phase identification and automatic ones which either use a built-in function of the smart meters or perform an analysis of the measurements.

Manual methods. On the one hand, phase identifiers comprise equipment usually composed of two parts, with one of them connected at a reference point in the network and the other one used by a technician to identify the phases at the customers' premises. The main drawbacks are the equipment cost and the need to send a technician to the premises. Moreover, these methods may also be prone to human error.

Automatic methods. On the other hand, recent smart meters, which use Power Line Carrier (PLC) technology to transfer the measurements, have a built-in function to identify the phases. However, if they do not have this function or uses a technology other than PLC, e.g. General Packet Radio Service (GPRS), a remote solution would instead be to use the measurements collected by the smart meter to perform the phase identification.

2.3 Active network management as an alternative to the fit and forget doctrine

After attempting to balance the phases, as a second alternative to costly investments to upgrade and reinforce the distributions networks, the current trend is to design and develop flexible, inexpensive control schemes for dealing with voltage problems. In the case of PV units, recent inverters offer the possibility of setting an active power production set point lower than the maximum production. This value is achieved by deviating the inverters from their optimal operating mode. It is also possible to establish a consumption or reactive power production set point, which also allows voltage control.

These schemes, referred to under the term of Active Network Management, usually control the distribution networks' power generation, consumption or storage to prevent or mitigate overvoltage and congestion problems. Several schemes have been suggested in the literature, acting on the PV active and reactive power outputs, to control the voltage profile of the system.[8]

Some of these schemes can be qualified as centralised, because modulation actions are calculated by a common entity responsible for collecting information (e.g. network type data, consumption and production information, etc.), calculating modulation orders according to optimisation objectives and constraints, and sending these modulation orders to actuators. For these schemes, the calculation of modulation orders is often based on solving an Optimal Power Flow (OPF) formulation of the problem. Such schemes usually require a large communication infrastructure and a detailed network model. As an example of such schemes, one can mention the one described by Liew[9] which proposes a formulation of the OPF problem where the objective is to minimise the market value of energy curtailed by wind turbines. A multi-step OPF can be used to determine the optimal location and management of distributed generation to minimize losses in the network.[10] Gemine[11] develops a sequential decision-making problem under uncertainty

[8] Malekpour, Pahwa, and S. Das ("Inverter-based var control in low voltage distribution systems with rooftop solar PV")

[9] Liew and Strbac ("Maximising penetration of wind generation in existing distribution networks")

[10] L. F. Ochoa and Harrison ("Minimizing energy losses: optimal accommodation and smart operation of renewable distributed generation")

[11] Gemine et al. ("Active network management: planning under uncertainty for exploiting load modulation")

where the distribution system operator calculates on D-1 the amount of flexible load it wishes to reserve and activates these flexible loads and/or decreases production in real time. Finally, in the field of LV PV panels, Tant[12] studies the possibility of battery storage of electricity that would otherwise have been lost. A multi-objective centralized optimization method is proposed to visualize the trade-off between voltage regulation, peak consumption reduction and annual costs.

[12] Tant et al. ("Multiobjective battery storage to improve PV integration in residential distribution grids")

On the other hand, many of these schemes can also be qualified as distributed, since the entities are controlled in a distributed manner without a central control entity and with little or no communication between the entities participating in the control scheme. In this case, distributed controllers often use local measurements to modulate the production and consumption of individual entities. In the field of photovoltaic panel integration, Xin[13] proposes a self-organised cooperative strategy for modulating the power of several installations. It requires the exchange of information to operate groups of installations at a certain percentage of the available power. Turitsyn[14] describes a distributed scheme that controls reactive power generation by inverters. Its actions are based on local voltage, power consumption and line parameters, in order to minimize network losses or regulate voltage.

[13] H. Xin et al. ("Cooperative control strategy for multiple photovoltaic generators in distribution networks")

[14] Turitsyn et al. ("Options for control of reactive power by distributed photovoltaic generatorss")

3 Contributions

This section describes the different contributions of each chapter and the global organization of the thesis. It is organized as a collection of research papers divided into three main parts: the phase identification of the smart meters, the control of the active and reactive power outputs of the PV units with a distributed control scheme, and the formalization of the concept of Energy Prosumers Communities.

3.1 Phase identification

We have developed two novel algorithms to identify the phases of the smart meters. Both algorithms solely rely on the correlation between the voltage measurements of the smart meters to solve the phase identification problem.

The first chapter details the first algorithm which is dedicated to the identification of three-phase smart meters and uses graph theory as well as the notion of maximum weight spanning tree to associate the smart meters that have the most correlated voltages. The material of this chapter was presented at the 2017 International Conference on Electricity Distribution (CIRED) in Glasgow (Scotland) with the paper:

> Frédéric Olivier, Damien Ernst, Raphaël Fonteneau. "Automatic phase identification of smart meter measurement data". In: *Proc. of the 2017 International Conference on Electricity Distribution (CIRED)*, Glasgow, Scotland, 2017.

Chapter 2 presents the second algorithm which overcomes the main drawback of the first algorithm, its inability to identify both single-phase and three-phase smart meters. The performance of this algorithm is compared to those of the constrained k-means clustering method, which has been previously used for phase identification.[15] Our algorithm improves the quality

[15] W. Wang, Yu, Foggo, et al. ("Phase identification in electric power distribution systems by clustering of smart meter data")

of the clustering, compared to the constrained k-means clustering. Indeed, it takes into account the underlying structure of the low-voltage distribution networks beneath the voltage measurements without a priori knowledge on the topology of the network. Both methods are analysed with real measurements from a distribution network in Belgium in order to give insight on how to parametrise them, and generalise their performance in different settings, by varying the ratios of single-phase over three-phase meters in the network, by increasing the period over which the voltages are averaged, etc.

The contributions are four-fold: (i) We propose a novel algorithm using the advantages of both graph theory and correlation to identify the measurements that should be linked together and cluster them. (ii) The performance of the algorithm is assessed in comparison to those of a constrained k-means clustering performed on the voltage measurements. (iii) Unlike previous references, the algorithms are designed for the specificities of European LV distribution networks. (iv) The algorithms are tested on real measurements from a distribution network in Belgium, in a variety of settings.

This algorithm was detailed in the article:

> Frédéric Olivier, Antonio Sutera, Pierre Geurts, Raphaël Fonteneau, Damien Ernst. "Phase identification of smart meters by clustering voltage measurements". In: *Proc. of the XX Power Systems Computation Conference (PSCC)*, Dublin, Ireland, 2018.

3.2 Active network management

In Chapter 3, a *distributed* control scheme that changes the active and reactive power injected by PV units into LV DNs is proposed. The objective of the control algorithm is to mitigate overvoltage problems by directing PV units to consume reactive power and, if necessary, to curtail active power generation. The distributed controllers are implemented on the PV inverters with several modes of operation.

First, if there are no overvoltage problems in the LV feeder, the controllers act by preventively adjusting the PV units' reactive power to avert the occurrence of overvoltages while performing Maximum Power Point Tracking (MPPT) for active power. In this mode of operation, only local measurements are used and no communication among the controllers is needed. Second, if overvoltage problems occur, the controllers make use of limited communication for coordinating their reactive power consumption within each DN LV feeder. Third, if the overvoltage persists, even after all PV units have utilized their maximum reactive capabilities, the controllers switch to active power curtailment. Finally, the fourth and fifth modes of operation restore active and reactive power production to normal operation. The proposed scheme is model free, as knowledge of the topology, the parameters or even the PVs' relative position in the LV feeder is not required by the controllers. Moreover, it is robust and fault-tolerant, as the lack of a centralised controller allows for the algorithm to work even if a number of PV units is malfunctioning or not participating.

The main contribution of the proposed scheme is that it does not require a network model or remote measurements, thus making it easy to deploy. Moreover, the use of communication is limited to a distress signal that can

be easily implemented with the use of Power-Line Communication (PLC).[16] This technology has been exploited for several decades to provide low-cost remote switching capabilities to utilities; one such example is the day/night tariff signal used in some countries.[17]

[16] Galli, Scaglione, and Z. Wang ("Power line communications and the smart grid")

[17] Carcelle (*Power Line Communications in Practice*)

Additionally, it does not require a centralised entity to collect measurements, compute set-points and dispatch the PV units. Thus, it is more robust to component failure. Finally, when active power curtailment is required to secure the system, the proposed scheme is designed to proportionally share the burden among all the PV units in the problematic feeder.

Assuming balanced three-phase network operations, the behaviour of the proposed control scheme is studied first on a single LV feeder and then on a larger network composed of 14 low-voltage feeders. Following, it is compared to a centralized OPF-based scheme applied to the same systems. Their performance is assessed based on their ability to alleviate the overvoltage problems and the amount of active power curtailed to achieve this. Both control schemes manage to ensure the security of the system. However, the OPF-based scheme does this with a little less curtailed active power.

The material of this chapter was published in:

> Frédéric Olivier, Petros Aristidou, Damien Ernst, Thierry Van Cutsem. "Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units". In: *IEEE Transaction on Smart Grid* 7.2 (2016), pp. 926-936.

Imbalances can be generally neglected in High Voltage (HV) and Medium Voltage (MV) networks in Europe, so a simplification commonly made, when simulating power systems, is the balanced three-phase operations of the network, i.e the currents in the three phases are equal; the three phases are balanced. This hypothesis simplifies the modelling of the system and the analysis of the simulation results. However, imbalances are often present in LV DNs since household appliances and small-scale PV units are usually single phase. So, the hypothesis that the three phases are balanced in LV networks is questionable but provides a good starting point. Chapter 4 is thus dedicated to the unbalanced operations of LV networks, explicitly taking into account and comprehensively modelling the three phases and the neutral, as well as the coupling between them.

In this chapter, the algorithm of Chapter 3 is extended to be used in unbalanced three-phase four-wire distribution networks with single- and/or three-phase inverters. Its key principles remains as follows: (i) it should first make use of reactive power, (ii) it should use active power curtailment as a last resort, (iii) it should only need communication in the form of a distress signal sent throughout the feeder to pool available resources, (iv) it should no require a detailed model of the network.

Using multi-step steady-state simulations, the behaviour of the algorithm is illustrated and its performance compared to that of two other algorithms. The first one is the default on-off behaviour of PV inverters when there is an overvoltage. They momentarily disconnect from the network until they try to restore production back to normal. The second one is based on a three-phase optimal power flow that is solved at each time step. The former represents the easiest solution to implement, whereas the latter is a more evolved scheme,

assuming perfect knowledge of the network, and thus providing a minimum bound on the curtailed energy

Their performances are compared in different PV connection scenarios impact the balance of the network, in terms of curtailed energy, reactive power usage and ohmic losses in the network.

The extension of the algorithm is presented in the article:

> Frédéric Olivier, Raphaël Fonteneau, Damien Ernst. "Distributed control of photovoltaic units in unbalanced LV networks to prevent overvoltages". *To be published.*

3.3 Implementing active network management in the form of energy prosumer communities

Chapter 5 presents a way to implement active network management in the form of Electricity Prosumer Communities (EPCs), which are groups of people producing, sharing and consuming electricity locally. Their members can be considered as prosumers, in the sense that they are both energy consumers and producers. They share resources in order to achieve a community goal. This goal may be related to electricity bills, but it can also be related to the nature of consumed electricity (for instance, maximising the consumption of renewable energy). The objectives that energy prosumers communities should explicitly or implicitly optimise, are related to the consumption of locally produced electricity, i.e. energy produced from renewable resources that is also consumed close to its place of production. Resources can be seen as distributed electricity generation using PV panels, flexibility means, distribution storage capacities such as batteries, and generation curtailment for PV units to mitigate possible over-voltages. In the case of a community of houses under the same low-voltage feeder, consuming locally means that when PV panels produce electricity, the power injection in the low-voltage feeder to grid should be avoided or minimised. Energy prosumer communities are a way to perform active network management at the level of an LV feeder: at the same time maximizing the goal of the community and minimizing the occurrence of operational constraints violations, such as overvoltages.

One of the main triggers for the emergence of the concept of energy communities is distributed electricity generation: PV units, small wind turbines and Combined Heat and Power (CHP) that may be installed close to consumers. In addition to electricity production and storage technology improvements, one should also mention the emergence of information technologies facilitating interactions between prosumers[18] by offering a new way of sharing goods and services. One should also note the existence of projects related to the use of distributed ledgers for managing energy exchanges[19] between microgrids.[20]

As an example, a community could organise itself as follows:

1. At every time-step, algorithmic tools (either distributed and / or centralised) should be able to identify which behaviour the community should adopt (consuming, producing, storing, curtailing, etc);

2. Members of the community should be informed in order to process the recommendation and act accordingly (either actively, either automatically);

[18] J. Wang, Costa, and Cisse ("From distribution feeder to microgrid: an insight on opportunities and challenges")

[19] Beitollahi and Deconinck ("Peer-to-peer networks applied to power grid")

[20] See for instance the Brooklyn Microgrid project.

3. The community should be rewarded for acting in accordance with the recommendation.

In the case of EPCs, our goal is to propose a rigorous mathematical framework for studying their behaviour. We first propose a mathematical framework for modelling the interactions between several prosumers. We then formalise a few optimisation problems targeting several different objectives (e.g., maximising distributed production, taking losses into account, optimising costs and revenues, etc). We address two ways to target these objectives: centralised and distributed control schemes, and we provide examples for each. In the centralised approach, we propose the design of a community strategy dedicated to the maximisation of the local renewable energy production by formalising it as an optimal power flow problem (OPF). Then, in the context where we want to minimise community costs, we propose to design a distributed strategy that may still approach community optimality. To do so, we build upon Fortenbacher's work[21] which proposes a centralised optimal strategy named Forward-Backward Sweep OPF (FBS-OPF). We use time series provided by FBS-OPF solutions to build learning sets in the form $(input, output)$, where $input$ contains only local measurements related to one single prosumer, and where $output$ contains an optimal value that was output by the FBS-OPF for this input configuration. The learning sets are processed by machine learning techniques in order to build regressors able to compute suggestions for any input configuration.

[21] Fortenbacher, Zellner, and Andersson ("Optimal sizing and placement of distributed storage in low voltage networks")

The formalization of EPCs was presented at the 2017 IREP Symposium in Espinho (Portugal) with the article:

> Frédéric Olivier, Daniele Marulli, Damien Ernst, Raphaël Fonteneau. "Foreseeing New Control Challenges in Electricity Prosumer Communities". In: *Proc. of the 10th bulk Power Systems Dynamics and Control Symposium (IREP)*, Espinho, Portugal, 2017.

3.4 Contributions in the Appendix

In addition to the chapters of the main body, this thesis collects four articles in the Appendix.

Chapter A aims at numerically assessing the influence of the modelling of the earth and the connection between the neutral and the earth, in terms of voltages and currents. Indeed, in order to simulate and optimise the behaviour of LECs in unbalanced networks, the three phases and neutral must be modelled explicitly. The simulations are performed on an existing Belgian low-voltage feeder supplying 19 houses which are all equipped with a smart meter measuring the mean voltage, current, and active and reactive power every minute for each phase. The simulations show that the explicit modelling of the earth using Carson's equations has a moderate effect on the simulation results. In particular, it creates differences in the simulations that are around ten-times smaller than the errors between simulations and measurements. The material of this chapter is detailed in the paper:

Frédéric Olivier, Raphaël Fonteneau, Damien Ernst. "Modelling of three-phase four-wire low-voltage cables taking into account the neutral connection to the earth". In: *Proc. of CIRED Workshop on Microgrids and Local Energy Communities*, Ljubljana, Slovenia, 2018.

In Chapter B, the penetration of grid-connected photovoltaic systems is studied, experimentally tested and compared to simulation results. In particular, how the reverse current flow and imbalance situations affect the voltage in the low-voltage grid. A test platform has been developed for obtaining experimental results with grid-tied commercial inverters. Photovoltaic arrays are emulated and subjected to different irradiance profiles and the inverters are controlled to produce under different power conditions. An electrical model has been developed in order to reproduce the same operating conditions and working environment. The simulations are performed with the software PowerFactory and the results compared to the experimental ones. The results were presented in the article:

Ramón López-Erauskin, Johan Gyselinck, Frédéric Olivier, Damien Ernst, María-Emilia Hervás, Alexis Fabre. "Modelling and emulation of an unbalanced LV feeder with photovoltaic inverters". In: *Proc. of the 8th IEEE Benelux Young researchers symposium in Electrical Power Engineering*, Eindhoven, The Netherlands, 2016.

Focusing on the flexibility aspect of DERs, in Chapter C, we studied different possibilities to organise exchanges of flexibility within a distribution system at the medium voltage level. Such a possibility is called an *interaction model*. The DSIMA (Distribution System Interaction Model Analysis) testbed allows one to compare quantitatively candidate interaction models by simulating a distribution system with actors taking decisions to maximise their own profit or minimise their costs. The original testbed focused on establishing the procedures to exchange information between actors and used a network flow model considering only active power. This paper extends DSIMA with a linear approximation of the power flow equations, the line limits, and the voltage constraints. This linear flow model is compared to a network flow model through the simulation of three different interaction models governing the exchange of flexibility services within a Belgian power distribution system. Results show that changing the network model may significantly impact the quantitative results obtained from the simulations. The material of this chapter was presented at the 2018 ISGT Conference in Washington (USA) with the paper:

Sébastien Mathieu, Frédéric Olivier, Damien Ernst, Bertrand Cornélusse. "Effect of voltage constraints on the exchange of flexibility services in distribution networks". In: *Proc. of the 9th Conference on Innovative Smart Grid Technology North America (ISGT)*, Washington, USA, 2018.

Finally, since the emergence of electric vehicles (EVs), combined with the rise of renewable energy production capacities, will strongly impact the way electricity is produced, distributed and consumed, Chapter D focuses on the problem of dispatching a fleet of electric vehicles in order to match,

at best, the harvesting of local, renewable energy. The combination of (mobile) EVs and (static) electricity sources is modelled as a multi-agent system, where mobile agents (the EVs) should gather, at best, renewable energy under travel distance and electricity production fluctuation constraints. The chapter exposes how a mobile application may offer an efficient solution for addressing this problem. Such an app would play two main roles: Firstly, it would incite and help people to play a more active role in the energy sector by allowing PV unit owners to sell their electrical production directly to consumers. Secondly, it would help distribution system operators to more efficiently modulate the load due to the EV deployment by allowing them to influence EV charging behaviour in real time. It advocates for the introduction of a two-sided market-type model between EV drivers and electricity producers. The material of this chapter was presented at the 2017 ICAART Conference with the article:

Antoine Dubois, Antoine Wehenkel, Raphael Fonteneau, Frédéric Olivier, Damien Ernst. “An app-based algorithmic approach for harvesting local and renewable energy using electric vehicles”. In: *Proc. of the 9th International Conference on Agents and Artificial Intelligence (ICAART)*, Porto, Portugal, 2017.

First Chapter
Automatic phase identification of smart meter measurement data

This paper highlights the importance of the knowledge of the phase identification for the different measurement points inside a low-voltage distribution network. Besides considering existing solutions, we propose a novel method for identifying the phases of the measurement devices, based exclusively on voltage measurement correlation. It relies on graph theory and the notion of maximum spanning tree. It has been tested on a real Belgian LV network, first with simulated unbalanced voltage for which it managed to correctly identify the phases of all measurement points, second, on preliminary data from a real measurement campaign for which it shows encouraging results.

1 Introduction

PREMASOL is a project which aims a predicting, analysing and controlling photovoltaic production in low-voltage distribution networks. Within the framework of this project, a measurement campaign is ongoing in a sub-urban low-voltage (LV) distribution network in Belgium. Each house is equipped with a smart meter measuring the voltages, the currents, the active and reactive power of the three phases. This measurement campaign has two goals: (i) The first goal is to validate the modelling that has been developed in a LV distribution system analysis tool. The active and reactive powers will be used as input and the resulting voltages will be compared to the measured ones. (ii) Another goal is for the distribution system operator (DSO) to gain insight on the three phase power flows inside its networks. Since all measurements are three-phase, it is critical to be able to associate each of them to a physical phase of the network. However, this information may not be easily available depending on the technology used to transfer the data from the smart meters. Several solutions exist and will be discussed. However, none of them suited our requirements. As result, in this paper, we propose a methodology for automatically pairing the phases of any number of measurement devices with those of the network, solely using measurement data. The paper is organized as follows. We first discuss the importance of being able to identify the phases of the smart meters. We then formalize the phase identification problem, shortly introduce the existing solutions, and describe our methodology. We finally provide an illustration of our method using both pseudo-measurements generated for the purpose of this study and real smart

meter data. On the importance of phase information When a measurement device — smart meter or other — is placed in a low-voltage distribution network, it is critical to know to which phase of the network the three phase measurements correspond. This information is crucial for multiple reasons: In the absence of the latter, (i) phase measurements such as line to neutral voltage cannot be used to identify key features of the network. For example, it is impossible to display a voltage profile of the feeder for each phase. (ii) It is not possible to know from the measurements which phase of the network is the most loaded and how the DSO can better balance its network by changing the phases at customer place. This can help to increase the hosting capacity of a LV network with high penetration of photovoltaic (PV) panels[1] by mitigating the overvoltages that cause PV inverters to disconnect and induce a loss of earnings for the owner.[2] (iii) For research purposes, it is impossible to properly use the measurement to perform load flow on a model of the network and compare the resulting voltages.

[1] Walling et al. ("Summary of distributed resources impact on power delivery systems")

[2] Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

2 The phase identification problem

For the purpose of explaining the phase identification problem, we will denote N_A, N_B and N_C the phases of the network and M^i_A, M^i_B and M^i_C the phases of the measurement device M^i. We already have established that the identity pairing $N_A - M^i_A$, $N_B - M^i_B$ and $N_C - M^i_C$ cannot be guaranteed because it depends on the connection of the cables of the smart meter and of the cables between the network and the house, an information not in the possession of the DSO of the network. The goal of the phase identification problem is, for each measurement point, to uniquely associate one phase of the measurements to one phase of the network. Since the reference is relative, the problem consists in clustering all the phase of the measurement points in three groups and then arbitrarily deciding which one corresponds to N_A, N_B or N_C.

3 Existing solutions

Different approaches can be used to solve the phase identification problem.

3.1 Smart meters with PLC capabilities

To begin with, the measurements of the smart meter can be repatriated thanks to Power Line Communication (PLC). In such a case, the measurements are gathered at the secondary substation and PLC can be used to directly identify the measured phases for each smart meter. The smart meter sends a different signal on each phase, a signal that is then used by the receiver to find the matching between it and the smart meter.

3.2 Phase identifiers

The second solution is to use a specific device which is based on GPS timing signals to compare the phase of an unknown voltage to a phase reference at the same instant of time. For this purpose, a base station is connected

to a known phase, for example at the distribution transformer. A technician then proceeds to the phase identification of each house by reading the phase shift between the phase reference and the phase of the voltages at the smart meter. Examples of such phase identification devices are the "Phase ID 6000" by *Power Systems Integrity*, the "PVS 100i" by *Megger* or the "Phase Identifier" from *Orgo Corporation*. The drawbacks are their cost and the manpower required to perform the identification. Moreover, this operation requires access to smart meters located inside houses, which depends on customer availability.

4 Motivations for an alternative

For the present application, one would certainly advise in favour of the deployment of PLC-enabled smart meters to repatriate the data. However, there exists several cases where this is not a viable option:

1. The technical infrastructure is already up and running and does not use PLC (e.g. GPRS).
2. The measurement devices are not smart meters but mobile measurement devices storing data in a local memory that has to be manually harvested, such as PQ boxes.
3. The measurements to be analysed are from previous measurement campaigns and no record of the phases has been included.

This highlights the need for an alternative. Two directions oppose each other: one based on equipment and man power, the other based on the analysis of the measurements themselves, which is the topic of this paper. Of course, the two are not mutually exclusive and the solution chosen by DSOs will most likely be based on their financial appeal. In that regard, the proposed method should be of interest as it can be significantly cheaper. This situation was observed within the PREMASOL project. One of the partners specialises in photovoltaic monitoring and already possesses an infrastructure to repatriate data through GPRS so this option was favoured over PLC. Moreover, the DSO wanted a less expensive solution than to invest in phase identification equipment and dedicate manpower to this task. In this paper, we propose an alternative based on the measurements and the unbalanced nature of the power flow inside the networks, to cluster the measurements by phase using correlation. Although time-series clustering is a well-studied problem,[3] it has, to the best of our knowledge, never been applied to the phase identification problem.

[3] Warren Liao ("Clustering of time series data—a survey")

5 Assumptions

The methodology is designed for unbalanced three-phase low-voltage distribution networks where houses have a smart meter — and/or where mobile measurement devices are placed — measuring phase-to-neutral voltages for each phase. We also assume that the measurements are synchronized and have the same sampling period of at least one minute to be able to capture the voltages variations. "Averaging data over periods longer than a minute

is shown to under-estimate the proportions of both [electricity] export and import."[4]

[4] Wright and Firth ("The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations")

6 The identification algorithm

The key idea on which the identification algorithm relies is that LV distribution networks are intrinsically unbalanced. Indeed, house appliances are mainly single phase. This unbalanced load creates an unbalance in voltages which can be used to cluster the measurements, phase by phase. The algorithm relies on graph theory, more specifically on the notion of a maximum spanning tree (MST), to select the pairings that will maximize the correlation between the voltages of the nodes. Using graph theory and maximum spanning trees comes naturally since any electrical network can be viewed as graph where nodes are buses and branch electrical lines, and distribution networks are usually operated in a radial fashion, so they can be viewed as trees. The different steps of the algorithm are the selection of the data, the creation of a complete graph where the branches are weighted by a correlation coefficient, and the selection of the most relevant branches by a maximizing spanning tree algorithm. The algorithm ends with the clustering of the measurements into three groups. The next subsections discuss these steps in more detail.

6.1 Selecting a time window for the data

First, a time window is selected; for example, one day, and to each measurement point three voltage time series are associated, one for each phase-to-neutral voltage.

6.2 Creating the graph

Let us define a pairing between a measurement M^i and a measurement M^j as a set of two mappings: a mapping from the three phases of M^i to the three phases of M^j, and an associated mapping from the three phases of M^j to three phases of M^i. Given that there are three phases, there are at most $3! = 6$ different sets of pairings, detailed in Table 1.1.

Pairing 1	$M^i_A - M^j_A$	$M^i_B - M^j_B$	$M^i_C - M^j_C$
Pairing 2	$M^i_A - M^j_A$	$M^i_B - M^j_C$	$M^i_C - M^j_B$
Pairing 3	$M^i_A - M^j_B$	$M^i_B - M^j_A$	$M^i_C - M^j_C$
Pairing 4	$M^i_A - M^j_B$	$M^i_B - M^j_C$	$M^i_C - M^j_A$
Pairing 5	$M^i_A - M^j_C$	$M^i_B - M^j_A$	$M^i_C - M^j_B$
Pairing 6	$M^i_A - M^j_C$	$M^i_B - M^j_B$	$M^i_C - M^j_A$

Table 1.1: Different pairings between the measurements triplets of M^i and M^j.

At this point, it can be noted that pairings 1 to 3, and 6, are symmetrical in the sense that the two associated mappings are identical. This is not the case for pairings 4 and 5: pairing 4 becomes pairing 5 and *vice versa*.

The next step of the algorithm is the creation of a complete graph where each node is connected to all the others. It means that, for a graph with n nodes, there are $n(n-1)/2$ branches. For each branch and its two end nodes, there are 6 different possible pairings as defined above. For each pairing, the

sum of the correlation coefficients between the three pairs of voltage time series is computed. For example, the correlation coefficient for the first and second pairing are computed in Equations (1.1) and (1.2).

$$\rho_1^{i,j} = corr(M_A^i, M_A^j) + corr(M_B^i, M_B^j) + corr(M_C^i, M_C^j) \tag{1.1}$$

$$\rho_2^{i,j} = corr(M_A^i, M_A^j) + corr(M_B^i, M_C^j) + corr(M_C^i, M_B^j) \tag{1.2}$$

The same formula can be applied for the other pairings. Pearson's measure is used is assess the correlation between the time series:

$$corr(X, Y) = \frac{\sum_{t=1}^{T} (X_t - \bar{X})(Y_t - \bar{Y})}{\sqrt{\sum_{t=1}^{T} (X_t - \bar{X})^2} \sqrt{\sum_{t=1}^{T} (Y_t - \bar{Y})^2}} \tag{1.3}$$

where X and Y are to time series of length T with a mean value of $\bar{X}$ and $\bar{Y}$.

Each branch of the graph represents the pairing that results in the maximum correlation coefficient, the weight of the branch being this maximum coefficient.

$$w_{i,j} = \max(\rho_1^{i,j}, \rho_2^{i,j}, \rho_3^{i,j}, \rho_4^{i,j}, \rho_5^{i,j}, \rho_6^{i,j}) \tag{1.4}$$

When storing the information on the pairing that maximizes correlation between two nodes, it is important to have a data structure that can differentiate between the two directions the branch can be traversed, because, as explained above, pairing 4 in one direction is pairing 5 in the other.

6.3 Finding the maximum spanning tree

Next, the Prim algorithm[5] is used to find the maximum spanning tree of the complete graph. It selects the edges that will bring the maximum total correlation between the nodes. The result is a tree where each branch represents the pairing which must be used to link measurements from the parent node to the child node. By using a tree, we ensure that there is no cycle inside the network and that there is only one possible succession of pairings from one node to any other.

[5] Prim ("Shortest connection networks and some generalizations")

6.4 Selecting the reference and clustering the phases

Once the tree is computed, we know how the phases of two adjacent nodes are paired, but this information is relative to the phases of the parent. The final step is to select a reference node to start from, and to traverse the entire tree structure, from parent to children, applying the pairing of each branch to uniquely select which cluster each phase measurements belongs to. The final results of the algorithm are three sets, $\mathcal{C}_A$, $\mathcal{C}_B$ and $\mathcal{C}_C$ where, for example,

$$\mathcal{C}_A = \{M_A^1, M_B^2, M_B^3, ...\} \tag{1.5}$$

$$\mathcal{C}_B = \{M_B^1, M_A^2, M_C^3, ...\} \tag{1.6}$$

$$\mathcal{C}_C = \{M_C^1, M_C^2, M_A^3, ...\} \tag{1.7}$$

that can be arbitrarily associated to the phases of the network N_1, N_B and N_C.

6.5 Including information on the topology of the distribution network

If more information on the structure of the electrical network is available, it can be used to reduce the number of branches from the complete graph. If nodes are far apart and on different feeders, the branch linking them can be discarded, reducing the number of edges and thus simplifying the resolution of the maximum spanning tree problem. One extreme option could be to directly define the tree by linking the nodes that are closest to each other, thus eliminating the need for the maximum spanning tree step. However, it is obviously not recommended because this eliminates a powerful step of the algorithm which can restructure the network based on correlation. For example, if measurement data were corrupted or if one house was associated with the wrong feeder, the algorithm will be able to circumvent those errors. Finally, it would be a strong shortcut to assume that the closer the measurement points are geographically, the stronger the correlation between the voltages as voltage variations are mainly due to the line impedance. So, we advise the reader to suppress branches with parsimony and only those that are without doubt irrelevant.

7 Test network

The test network used for this study is an existing Belgian low-voltage distribution network, composed of three feeders made with underground cables of the type EVAVB-F2 3x95 + 1x50. It is located in a suburban area and has been modelled according to Ciric's work[6] based on the data provided by the DSO (topology, line length, cable type, etc.). Detailed unbalanced three-phase four-wire modelling of the network has been used.

[6] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

7.1 Model for the dwellings: load and photovoltaic units

The network is composed of 32 houses, all of which have a three-phase 400/230 V connection of various length with a cable of type EXAVB 4x10. Five of these houses are equipped with photovoltaic units. The energy consumption of the house is modelled using consumption profiles created with Widen's tool.[7] Several alterations have been made to the code created by Widén and Wäckelgård in order to allow the creation of unbalanced load profiles. First, the appliances have been classified as single phase or three phase. Each single-phase appliance has been allocated to one of three groups based on good practice, trying to balance the load in each group as optimally as possible. Each time the profile generator is run, appliances are clustered in the same groups, however, the clusters are randomly allocated to a specific phase. At this point, rather than calculate the sum of the consumption of all appliances, appliance consumptions are summed phase by phase, adding one third of the three-phase appliances. An example of a load profile can be seen in Figure 1.1.

[7] Widén and Wäckelgård ("A high-resolution stochastic model of domestic activity patterns and electricity demand")

The production of the photovoltaic panels is based on the production of a typical photovoltaic unit in Belgium, scaled with respect to the peak power of each unit. The consumption and production profile have been generated for an arbitrarily chosen day: Thursday, the 5th of May, 2016.

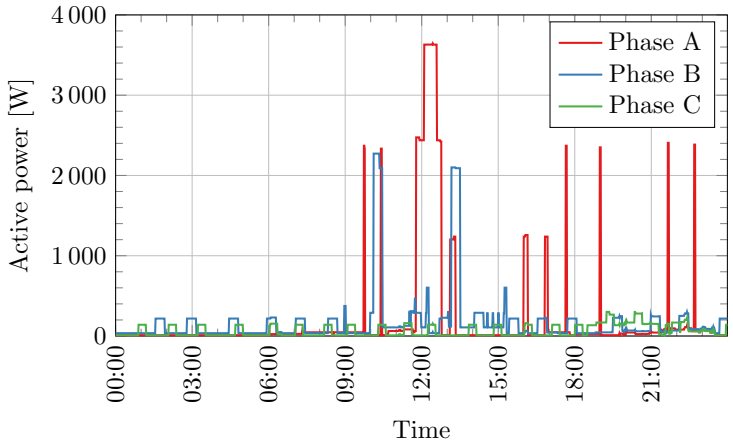


Figure 1.1: Example of an unbalanced load profile.

7.2 *Model for the medium voltage network*

The medium voltage network is modelled as a Thevenin equivalent. The phase-to-phase voltage of the equivalent is fixed at 420 V and the impedance at $0.0059 + j \cdot 0.0094\ \Omega$.

7.3 *Pseudo-measurement generation*

An unbalanced load flow algorithm[8] implemented in Python is used to compute the currents in the lines and the voltages at each node with a resolution of 1 minute to capture the variability of the loads. The three phase-to-neutral voltages at the bus where a house is connected are exported so that they can be used in the next step. Such voltages are displayed in Figure 1.2 for measurement point M^{25}.

[8] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

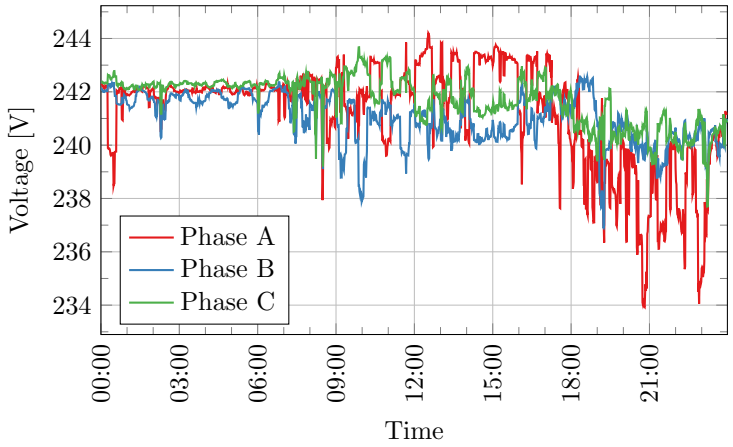


Figure 1.2: Voltages for measurement point M^{25}.

8 *Results*

The algorithm was implemented in Matlab. To compute the maximum spanning tree, the Prim algorithm (minimum spanning tree) is used with the opposite of the branches' weight.

8.1 With pseudo-measurements

Because they are generated using a LV network simulator, phase measurements are already sorted, so the first step is to randomly permute them for each measurement point. Then, the identification algorithm is run on the data. The most important result is that the algorithm successfully manages to identify all the phases and to cluster them the proper way, regardless of the initial permutation. As an illustration, Figure 1.3 shows the voltages that were clustered by the algorithm into one single group.

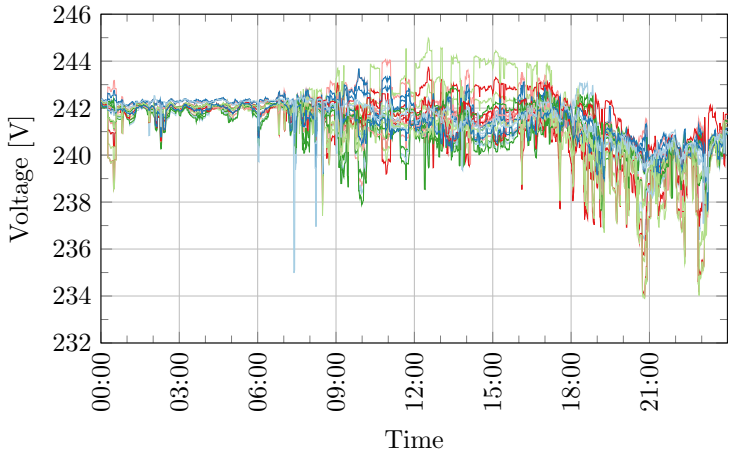


Figure 1.3: Example of a cluster of time series after the algorithm.

The maximum spanning tree output by the algorithm is displayed in Figure 1.4, superimposed on the structure of the electrical network. It can be seen that the structure of the MST is coherent with the structure of the network as no nodes are connected between different feeders except at the root of the network. The edges are weighted by the correlation coefficient of the correlation maximizing pairing. It can be seen that all values are close to 3, the maximum, indicating an excellent correlation between the nodes.

8.2 With real measurements

At the time of writing this paper, the roll-out of the smart meters for the purpose of the measurement campaign is still ongoing and all houses have not yet been equipped. Regardless, the algorithm has been applied to the measurements that have already been collected. The resulting tree is displayed in Figure 1.5.

It can be seen that the majority of edges have a strong correlation coefficient, except the one linking nodes 8 and 24. The reason is certainly that the two measurement points are too far apart, either due to the lack of a smart meter in between, or due to a large impedance between them. It is information that can be further investigated by the DSO in its phase identification process. In any case, we advise the installation of a smart meter at the low-voltage side of the distribution transformer to provide a measurement point that can be used to link different feeders. Finally, the method can be used to check if a measurement point's location was incorrect. It could be observed when a branch of the MST is not coherent with the topology of the network, and has a low correlation coefficient.

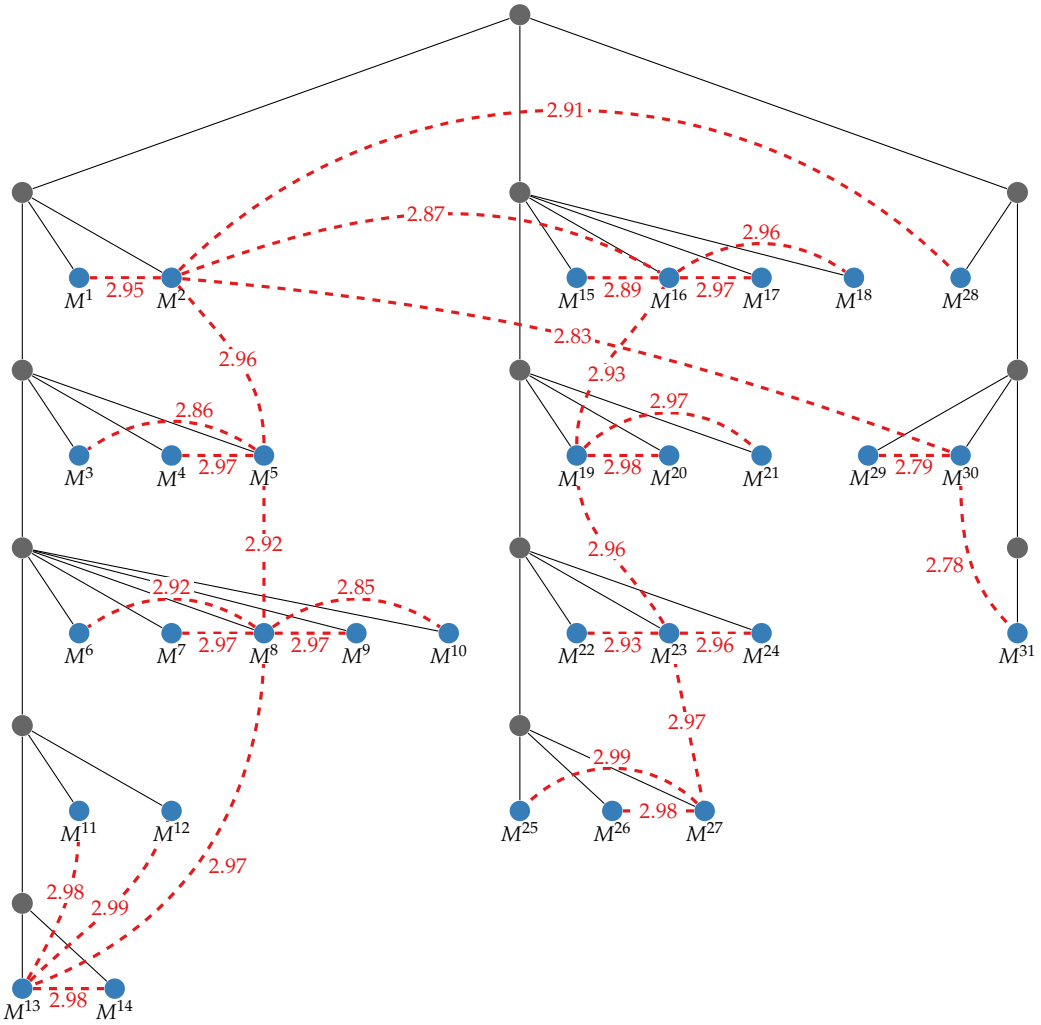


Figure 1.4: The topology of the electrical network is in grey. The upper node corresponds to the distribution transformer and the nodes that are numbered are measurement points (blue). The maximum spanning tree is in red (dashed) and the weight of the edges is the maximum correlation coefficient.

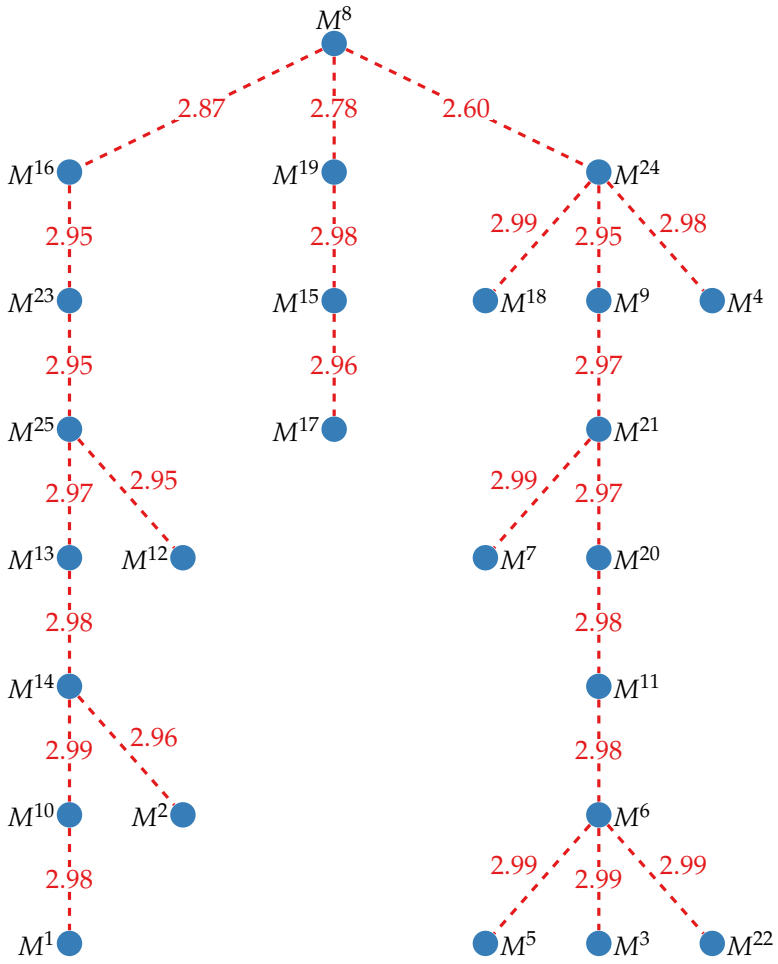


Figure 1.5: Maximal spanning tree for the real measurement data where the weight of the edges is the maximum correlation coefficient.

9 Conclusion

We have proposed a phase identification algorithm which performs exactly as planned when applied to measurements purposely generated. In addition to phase identification, the algorithm outputs a maximum spanning tree which provides insight into the structure of the electrical network and the measurements that are more correlated. Furthermore, the first results from real measurements are extremely encouraging as the correlation coefficients are close to their maximum. Several research questions arise from this study. It would be interesting to investigate the behaviour of the algorithm when the sampling period of the measurement is increased (from 1 to 10 minutes for example). This should lead to a decrease in the imbalance of the voltages and lead to a decrease of the overall correlation coefficients. The next steps are to further analyse the real measurement data in regards to their location in the network, and to provide an explanation and/or a solution to the links of the graph which have a poor correlation coefficient.

Chapter 2
Phase identification of smart meters by clustering voltage measurements

When a smart meter, be it single-phase or three-phase, is connected to a three-phase network, the phase(s) to which it is connected is (are) initially not known. This means that each of its measurements is not uniquely associated with a phase of the distribution network. This phase information is important because it can be used by Distribution System Operators to take actions in order to have a network that is more balanced.

In this work, the correlation between the voltage measurements of the smart meters is used to identify the phases. To do so, the constrained k-means clustering method is first introduced as a reference, as it has been previously used for phase identification. A novel, automatic and effective method is then proposed to overcome the main drawback of the constrained k-means clustering, and improve the quality of the clustering. Indeed, it takes into account the underlying structure of the low-voltage distribution networks beneath the voltage measurements without a priori knowledge on the topology of the network. Both methods are analysed with real measurements from a distribution network in Belgium. The proposed algorithm shows superior performance in different settings, e.g. when the ratio of single-phase over three-phase meters in the network is high, when the period over which the voltages are averaged is longer than one minute, etc.

1 Introduction

When a smart meter, be it single-phase or three-phase, is connected to a three-phase network, the phase(s) to which it is connected is (are) initially not known. This means that each of its measurements is not uniquely associated with a phase of the distribution network. This can be achieved by solving the phase identification problem, i.e. associating the phases of the network to those of the smart meter. It is equivalent to clustering the measurements in three groups, one for each phase of the network. This phase information is important because it can be used by Distribution System Operators (DSO) to take actions in order to have a network that is more balanced.

In this work, the correlation between the voltage measurements of the smart meters is used to solve the phase identification problem. To do so, the constrained k-means clustering method is first introduced. It will be used as a reference for the performance as it has been previously used for phase

identification.[1] A novel, automatic and effective method is then proposed to overcome the main drawback of the constrained k-means clustering and improve the quality of the clustering. Indeed, it takes into account the underlying structure of the low-voltage (LV) distribution networks beneath the voltage measurements without a priori knowledge on the topology of the network. Both methods are analysed with real measurements from a distribution network in Belgium, in order to give insight on how to parametrise them, and generalise their performance in different settings, by varying the ratios of single-phase over three-phase meters in the network, increasing the period over which the voltages are averaged, etc.

[1] W. Wang, Yu, Foggo, et al. ("Phase identification in electric power distribution systems by clustering of smart meter data")

The paper is organised as follows: the phase identification problem is motivated in the first section and formalised in the second. In section 3, the two clustering algorithms are presented. We then describe the LV distribution network used for the computations and the assessment of the algorithms. Sections 4 and 5 present the parametrisation of the algorithms and their performance in different settings. Finally, section 6 summarises the advantages and pitfalls of the proposed clustering methods to identify the phase of smart meters in a LV distribution network.

2 *On the importance of phase identification*

LV distribution networks are intrinsically unbalanced due to the currents that are consumed by single-phase household appliances, or produced by single-phase distributed generation units. This imbalance reduces the hosting capacity of the network, i.e. the maximum amount of distributed generation that can be connected without violating the operational constraints in the network.[2] Thus, the first action to increase the hosting capacity is to better balance production on the three phases of the network. The phase information allows the distribution system operators (DSOs) to do so by changing the phase(s) to which their customers are connected.

[2] Dubey, Santoso, and Maitra ("Understanding photovoltaic hosting capacity of distribution circuits"), Bletterie et al. ("Understanding the effects of unsymmetrical infeed on the voltage rise for the design of suitable voltage control algorithms with PV"), Cundeva, M. Bollen, and D. Schwanz ("Hosting capacity of the grid for wind generators set by voltage magnitude and distortion levels"), and Daphne Schwanz et al. ("Stochastic assessment of voltage unbalance due to single-phase-connected solar power")

Moreover, three-phase power flow simulations take as input the active and reactive powers that are consumed or produced at each connection point, phase by phase. To obtain simulations accurately reflecting real operations, the phase information is required to know to which phase of the network the active and reactive power measurements correspond. This also allows for the comparison between the voltages that are measured and voltages that are simulated.

3 *The phase identification problem*

In this paper, we consider that the meters can either be single phase, i.e. connected between a phase and the neutral, or three phase, i.e. connected to the three phases and the neutral. This implies that the network to which the meters are connected is a four-wire, three-phase and neutral star-shaped distribution network. Let

$$\mathcal{S} = \{S_j\}, j \in \mathcal{J} = 1...J$$

be the set of all meters, where J is the total number of smart meters. The meters measure several electrical quantities that can be aggregated, such

the total active power that is consumed, or phase related, such as phase-to-neutral voltages. The latter measurements are identified by the letters r_j, s_j, t_j, among which each corresponds to one phase of the meter. When a smart meter is single phase, only the index r_j is used.

As explained in the introduction, there is no coherence between r, s and t at the network level. They are only valid at the meter level. If the phases of the network are labelled a, b, c, the goal of the phase identification is thus to associate the phase r_j, s_j and t_j of each smart meter, S_j, to one phase of the distribution network.

It is equivalent to clustering the phases of the meters in three groups, which will be denoted $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$, and then associating each group to one phase of the network by taking a common reference. Moreover, the physical nature of the problem imposes constraints on the clustering process. Since meter phases are electrically connected to different phases of the network, two indices linked to the same smart meter cannot be placed in the same cluster, i.e. r_j, s_j, t_j must be placed in different clusters. Of course, single-phase meters do not impose such constraints and the phase index can be placed in a cluster without restrictions.

4 Existing solutions

Two types of solution oppose each other to solve the phase identification problem: *manual* ones, which require a technician to proceed, manually, to the phase identification and *automatic* ones which either use a built-in function of the smart meters or perform an analysis of the measurements.

4.1 Manual methods

On the one hand, phase identifiers are equipment usually composed of two parts, with one of them connected at a reference point in the network and the other one used by a technician to identify the phases at the customers' premises. A primary technology uses clocks that are synchronised with GPS, to compare the voltage angle between the reference voltage and the voltage measured by the technician. A second technology broadcasts different signals in the three phase conductors, and the technician proceeds to the phase identification by reading the signal in each of the smart meter's conductors. The main drawbacks are the equipment cost and the need to send a technician to the premises. Moreover, these methods may also be prone to human error.

4.2 Automatic methods

On the other hand, recent smart meters, which uses Power Line Carrier (PLC) technology to transfer the measurements, have a built-in function to identify the phases. However, if it does not have this function or uses a technology other than PLC, e.g. General Packet Radio Service (GPRS), a remote solution would instead be to use the measurements collected by the smart meter to perform the phase identification.

In the literature, it has been proposed to use energy measurements[3] or voltage measurements[4] to (re)-discover the network topology, including the

[3] Pappu et al. ("Identifying topology of low voltage (LV) distribution networks based on smart meter data")

[4] Mitra et al. ("Voltage correlations in smart meter data") and J. D. Watson, Welch, and N. R. Watson ("Use of smart-meter data to determine distribution system topology")

phase identification of the smart meters. Pappu[5] proposes a method using graph theory based on a principal component analysis (PCA), whereas several papers[6] use a correlation-based approach exploiting the similarities between the voltage measurements of the same phase (k-means clustering). Some papers[7] explicitly identify the phases with either energy or voltage measurements. Rajeswaran[8] applies graph theory and PCA to identify the phases of a simulated dataset, taking into account the effect of noise. Pezeshki[9] uses correlation to find how to locally associate the phases of each smart meter to the phases of a reference meter. Finally, instead of directly employing correlation, Wang[10] extracts features on which they perform a k-means clustering.

This paper extends our previous work.[11] Its contributions are four-fold: (i) It proposes a novel algorithm with a global network, using the advantages of both graph theory and correlation to identify the measurements that should be linked together and cluster them. (ii) The performance of the algorithm is assessed in comparison to those of a constrained k-means clustering performed on the voltage measurements. (iii) Unlike previous references, the algorithms are designed for the specificities of European LV distribution networks. (iv) The algorithms are tested on real measurements from a distribution network in Belgium, in a variety of settings.

[5] Pappu et al. ("Identifying topology of low voltage (LV) distribution networks based on smart meter data")

[6] Mitra et al. ("Voltage correlations in smart meter data") and J. D. Watson, Welch, and N. R. Watson ("Use of smart-meter data to determine distribution system topology")

[7] Rajeswaran, Bhatt, Pasumarthy, et al. ("A novel approach for phase identification in smart grids using graph theory and principal component analysis"), Pezeshki and Peter J Wolfs ("Consumer phase identification in a three phase unbalanced LV distribution network"), and W. Wang, Yu, Foggo, et al. ("Phase identification in electric power distribution systems by clustering of smart meter data")

[8] Rajeswaran, Bhatt, Pasumarthy, et al. ("A novel approach for phase identification in smart grids using graph theory and principal component analysis")

[9] Pezeshki and Peter J Wolfs ("Consumer phase identification in a three phase unbalanced LV distribution network")

[10] W. Wang, Yu, Foggo, et al. ("Phase identification in electric power distribution systems by clustering of smart meter data")

[11] Olivier, Ernst, and Fonteneau ("Automatic phase identification of smart meter measurement data")

5 *Description of the proposed methods*

After a description of the distribution network characteristics that provides an insight on how to measure similarity between two voltage measurements, the constrained k-means and the novel approach are presented. It should be mentioned that both methods are automatic and do not require additional equipment. Moreover, they are solely based on voltage measurements and do not require a priori knowledge of the network topology, such as the partition of smart meters between feeders. This point is important since this information is often not reliable due to the various reconfigurations of LV distribution networks.

5.1 *Definition of the distance between two voltage measurements*

Correlation-based approaches seem very well adapted to the phase identification problem and have been successfully used.[12] The proposed algorithms also use Pearson's correlation to assess the similarities between two time series M_1 and M_2, gathering the measurements of the phase-to-neutral voltage magnitudes:

$$PC(M_1, M_2) = \frac{\sum_{\tau=1}^{n} \left(M_{1,\tau} - \hat{M}_1\right)\left(M_{2,\tau} - \hat{M}_2\right)}{\sqrt{\sum_{\tau=1}^{n} \left(M_{1,\tau} - \hat{M}_1\right)^2} \sqrt{\sum_{\tau=1}^{n} \left(M_{2,\tau} - \hat{M}_2\right)^2}}$$

where $\hat{M}_1$ and $\hat{M}_2$ are the mean values of M_1 and M_2, and n the length of both time series. Indeed, the voltage at the end of an electrical line is equal to the sum of the voltage at its beginning and the voltage drop caused by the current flowing through the impedance. Thus, points that are separated by a line with a small impedance, or which are traversed by a small current, have voltages that are more correlated. Since phases are not electrically connected to each other, voltage measurement from different phases are less correlated.

[12] J. D. Watson, Welch, and N. R. Watson ("Use of smart-meter data to determine distribution system topology") and Pezeshki and Peter J Wolfs ("Consumer phase identification in a three phase unbalanced LV distribution network")

Other similarity metrics could be of interest, such as the cosine similarity, if the same approach were to be adapted to other applications with a different type of data.

Finally, let the set of voltage measurements be $\mathcal{M} = \{M_i\}$, with $i \in \mathcal{I} = \{r_1, s_1, t_1, r_2, ..., r_J, s_J, t_J\}$, where M_i is a voltage time series and i is its index in the total set of time series, not the time index. The distance between two voltage measurements is defined as

$$d(M_l, M_i) = 1 - PC(M_l, M_i), \quad \forall l, i \in \mathcal{I}$$

so that the distance is equal to 0 if the measurements are perfectly correlated.

5.2 *Reference algorithm: Constrained k-means Clustering*

k-means clustering[13] aims at partitioning a set of observations into k clusters. At each iteration, the observations are associated to the cluster with the *closest* centre, in the sense of a distance metric (e.g., Euclidian distance, correlation measure, etc.) that assesses the similarity between an observation and a centre. The centres are usually computed by averaging the observations already associated to the cluster. In the phase identification case, the number of clusters, k, is equal to the number of phases, i.e. $k = 3$. However, this algorithm does not take into account the constraints that the measurements from a same three-phase smart meter must be placed in different clusters. The k-means clustering algorithm is thus modified to correspond to the one proposed by Wagstaff,[14] which introduces background knowledge through constraints..

[13] Lloyd ("Least squares quantization in PCM")

[14] Wagstaff et al. ("Constrained k-means clustering with background knowledge")

Formally, the constrained k-means (CKM) clustering algorithm for phase identification (as used by Wagstaff and Wang,[15] detailed in Algorithm 1 and illustrated in Figure 2.6) works as follows: Three empty clusters are defined based on the number of phases. The initial centres are equal to the measurements of the root three-phase smart meter given as input. This ensures that initial centres are measurements from three distinct phases. At each iteration, all smart meters are examined: each voltage time series of the smart meter is associated with the cluster whose center is the closest, in such a way that only one measurement can belong to a cluster to satisfy the three-phase constraint. For the sake of implementation, the pair measurement-cluster with the smallest distance is first selected, then the second pair (excluding the measurement and cluster from the first pair), and then the last pair is trivially associated to the remaining cluster. For a single-phase element, the measurement is directly associated to the closest centre. Once all smart meters have been examined, centres are updated by averaging all the measurements that have been assigned to them. If those updates are still above the given convergence threshold, or if the maximal number of iterations is not reached, then the process keeps iterating.

[15] Wagstaff et al. ("Constrained k-means clustering with background knowledge"), W. Wang, Yu, Foggo, et al. ("Phase identification in electric power distribution systems by clustering of smart meter data"), and W. Wang, Yu, and Lu ("Advanced metering infrastructure data driven phase identification in smart grid")

A drawback of the k-means clustering approach is the fact that, between two subsequent iterations, measurements associated with a given cluster are averaged to compute the cluster centre for the next iteration. Averaging measurements may destroy some information by eliminating small variations in the measurements.

Algorithm 1: Constrained k-means Algorithm

Inputs:

1. Set of measurements $\mathcal{M}$
2. Set of smart meters $\mathcal{S}$
3. Three-phase root smart meter S_{j_0}
4. Distance metric $d(\cdot,\cdot)$

Output:

- Three clusters satisfying the smart meter constraints: $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$.

Algorithm:

1. Let $c_1 = M_{r_{j_0}}, c_2 = M_{s_{j_0}}, c_3 = M_{t_{j_0}}$ be the initial cluster centres of $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$.
2. Let $\mathcal{J}_1$ be the subset of the single-phase smart meter indices and $\mathcal{J}_3$ the subset of three-phase smart meter indices.
3. $\mathcal{C}_k = \emptyset$ with $k = 1, 2, 3$.
4. $\forall S_j$, with $j \in \mathcal{J}_1$,

$$k^* = \underset{k=\{1,2,3\}}{arg\,mind}(c_k, M_{r_j})$$
$$\mathcal{C}_{k^*} = \mathcal{C}_{k^*} \cup r_j$$

5. $\forall S_j$ with $j \in \mathcal{J}_3$,

a)
$$\{k_1^*, i_1^*\} = \underset{\substack{k=\{1,2,3\}\\ i=\{r_j,s_j,t_j\}}}{arg\,min}\, d(c_k, M_i)$$
$$\mathcal{C}_{k_1^*} = \mathcal{C}_{k_1^*} \cup i_1^*$$

b)
$$\{k_2^*, i_2^*\} = \underset{\substack{k=\{1,2,3\}\backslash k_1^*\\ i=\{r_j,s_j,t_j\}\backslash i_1^*}}{arg\,min}\; d(\mathcal{C}_k, M_i)$$
$$\mathcal{C}_{k_2^*} = \mathcal{C}_{k_2^*} \cup i_2^*$$

c)
$$k_3^* = \{1,2,3\}\backslash\{k_1^*, k_2^*\},\ i_3^* = \{r_j, s_j, t_j\}\backslash\{i_1^*, i_2^*\}$$
$$\mathcal{C}_{k_3^*} = \mathcal{C}_{k_3^*} \cup i_3^*$$

6. For each cluster $\mathcal{C}_k$, update its center by averaging all the measurements M_i, $\forall i \in \mathcal{C}_k$.
7. Iterate between (3) and (6) until the algorithm reaches convergence (i.e., changes in centres are smaller than a given threshold ϵ) or the maximal number of iterations.

5.3 Proposed algorithm: Constrained Multi-tree Clustering

This paragraph explains the motivation for a tree-structured algorithm to cluster the phase measurements. Distribution systems are usually operated radially, i.e. there are no electrical loops, with one point of connection between the MV network and the LV network (the distribution transformer). From a graph theory perspective, the network can be seen as a tree, where the distribution transformer is the root and the connection points are the leaves. Indeed, a tree is defined as an undirected graph in which any two nodes are connected by exactly one path, in other words without cycles. It makes sense to use a tree structure to cluster the measurements, as explained previously,[16] because the tree structure of the network creates an underlying structure between the voltage measurements. Indeed, the voltages at two points that are neighbours in the network tree are more correlated.

[16] Olivier, Ernst, and Fonteneau ("Automatic phase identification of smart meter measurement data")

The Constrained Multi-tree (CMT) algorithm was inspired by Prim's algorithm, which is used to calculate minimum spanning trees in weighed graphs. Prim's algorithm starts at a root node and makes the tree grow gradually by adding the node whose branch will add the lowest weight to the tree. Prim's algorithm cannot be applied in this form to cluster the measurements because its output is only one tree. The algorithm is modified to output three trees, hence three clusters.

Let

$$\Delta(\mathcal{C}_k, M_i) = \min_{l \in \mathcal{C}_k} d(M_l, M_i)$$

be the definition of the distance between a cluster $\mathcal{C}_k$ and a measurement M_i. Initially, the clusters contain only one measurement, those of the root smart meter given as input to the algorithm. For the same reason as for the previous algorithm, the root smart meter needs to be three-phase. Measurements will be sequentially added to the root measurements to form the trees in the cluster, as explained in Algorithm 2 and illustrated in Figure 2.7. At each iteration, a distance δ_i is associated to the measurement M_i from a smart meter, that is the closest to a cluster, while satisfying the constraints. A potential cluster κ_i and a potential predecessor π_i are also associated, corresponding to the cluster and its measurement, which minimises the distance. The constraints are taken into account by selecting the best pair between the remaining measurements for a smart meter and the remaining clusters. If the measurements of a smart meter are already in a cluster, the distance, potential cluster, and potential predecessor are computed without considering those measurements and corresponding clusters.

Once the distances have been computed, the measurement, which is not yet in a cluster and whose distance is the smallest, is added to its potential cluster, and edge is created with its potential predecessor. By doing so, the trees grow at each iteration by adding the measurement with the minimum cost, i.e. whose connection to the minimum spanning tree will make the sum of the weight of the edges increase by the least amount. The process is repeated until all measurements are clustered.

Algorithm 2: Constrained Multi-Tree Algorithm

Inputs:

1. Set of measurements $\mathcal{M}$
2. Set of smart meters $\mathcal{S}$
3. Three-phase root smart meter S_{j_0}
4. Distance metric $d(\cdot,\cdot)$

Output:

- Three clusters satisfying the smart meter constraints: $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$.

Algorithm:

1. Let $\mathcal{C}_1 = \{r_{j_0}\}, \mathcal{C}_2 = \{s_{j_0}\}, \mathcal{C}_3 = \{t_{j_0}\}$ be the initial clusters, where $r_{j_0}, s_{j_0}, t_{j_0}$ are the phase indices of the root smart meter S_{j_0}.
2. Let $\mathcal{J}_1$ be the subset of single-phase smart meters and $\mathcal{J}_3$ the subset of three-phase smart meters.
3. $\delta_i = +\infty, \forall i \in \mathcal{I}$
4. $\forall S_j$ with $j \in \mathcal{J}_1$, whose measurement index is not yet in a cluster,

   a) $$k^* = \underset{k=\{1,2,3\}}{arg\,min} \Delta(\mathcal{C}_k, M_{r_j})$$

   b) $$\delta_{r_j} = \Delta(\mathcal{C}_{k^*}, M_{r_j})$$
   $$\pi_{r_j} = \underset{l \in C_{k^*}}{arg\,min}\ d(M_l, M_{r_j})$$
   $$\kappa_{r_j} = k^*$$

5. $\forall S_j$, with $j \in \mathcal{J}_3$, whose three measurements are not yet in a cluster,

   - Let $\mathcal{I}^* = \{r_j, s_j, t_j\} \cap (\cup_{k=1}^{3} \mathcal{C}_k)$ be the set of measurement indices from the smart meter S_j that are already in a cluster, and the set of corresponding clusters $\mathcal{K}^*$.
   - Then,

   $$\{k^*, i^*\} = \underset{\substack{k=\{1,2,3\}\setminus K^* \\ i=\{r_j,s_j,t_j\}\setminus I^*}}{arg\,min}\ \Delta(\mathcal{C}_k, M_i)$$
   $$\delta_{i^*} = \Delta(\mathcal{C}_{k^*}, M_{i^*})$$
   $$\pi_{i^*} = \underset{l \in C_{k^*}}{arg\,min}\ d(M_l, M_{i^*})$$
   $$\kappa_{i^*} = k^*$$

6. The measurement index with the smallest distance δ is added to the corresponding cluster κ:

   $$i^* = \underset{i \in I \setminus \cup_{k=1}^{3} C_k}{arg\,min}\ \delta_i$$
   $$\mathcal{C}_{\kappa_{i^*}} = \mathcal{C}_{\kappa_{i^*}} \cup i^*$$

   Create an edge between i^* and π_{i^*}.

7. Iterate between (3) and (6) until all measurements are assigned to a cluster.

6 Test system

6.1 The low-voltage distribution network

The algorithm is tested on voltage measurements from a Belgian LV distribution network, which is composed of five feeders with a star configuration 400V/230V.

The network supplies 89 houses, among which 74 are equipped with a three-phase smart meter and two with a single-phase smart meter. Three houses are equipped with both a regular smart meter and a night-exclusive one, the latter not providing voltage measurements during the day from 7:00 to 22:00.

Each smart meter provides phase-to-neutral voltage measurements as a one-minute average that are transferred using GPRS. The beginning of each feeder is also equipped with a smart meter and each feeder is associated with voltage and current measurements at the transformer. So, the total number of voltage measurement points is 81. It is important to note that the phase identification of all smart meters is known, in order to evaluate the performance of the algorithms.

6.2 Voltage measurement test set

To deal with night-exclusive meters, the phase identification is done with night measurements. The test set is composed of night measurements from all houses and transformer from September 15th, 2017, 23:00 to September 18th, 2017, 4:00, without any measurements from 7:00 to 23:00. This corresponds to 239 time series of 1260 one-minute voltage measurement.

6.3 Performance measure and empirical assessment

In order to assess the validity of the phase identification, two cases need to be distinguished: (i) an empirical verification (the true phase identification) is available or possible, at a potentially high cost, and can be used to verify the results, (ii) no information is available.

With a (partial) solution, two accuracy metrics can be used benefiting from the true solution. The first one considers each pair (measurement-cluster) as an individual element. The second one considers network elements (house or transformer) as indivisible elements and therefore, an element is either fully correctly identified or not at all. This second metric is not sensitive to the number of measurements per element, while the first one allows to have score for partial identification. In the following, we only focus on the first metric in order to differentiate single-phase and three-phase smart meters. Moreover, if an empirical verification – by sending a technician for example – is possible, but expensive, it can be used to verify the network at some critical points.

Without a (partial) solution, no clues are available to evaluate the quality of the prediction. In this case, some confidence scores may be extracted from the phase identification process in order to identify what may or may not be correctly identified. For example, instead of using all the measurements, one can divide the measurements into several (overlapping) windows of a

given length. The consistency of the predictions is assessed by comparing the results of the phase identification for all time windows. To do so, the number of occurrences on which two measurements are in the same cluster are counted to obtain a confidence measure on the association of these two measurements.

7 *Results for the test set: discussions on the selection of the root*

Both algorithms need an initialisation with the measurements from a three-phase smart meter. The current section aims at assessing the impact of the root on the performance of the algorithms. The performance index is equal to the ratio between the number of voltage measurements which are correctly identified and the total number of measurements (i.e. 239 for the full data set). When the root is the meter at the distribution transformer, the performance index is 1 for both algorithms. When it is a house, the performance is between 0.67 and 1 for CKM and remains at 1 for the CMT. When half the number of three-phase meters are converted to single-phase meters by randomly selecting one phase, the performance index lies between 0.67 and 1 for both algorithms, while the performance index remains at 1 when the transformer is the root. This highlights the influence of the root on the performance of the algorithms, and hints that the transformer provides a strong root. Indeed, the voltages at the distribution transformer have a particular importance because they influence all the other voltages of the LV network. To further show this behaviour, the algorithms were tested in more demanding settings. The phase identification was run 100 times. For each run, half the meters were randomly removed and half of the remaining were converted to single-phase smart meters by randomly selecting a phase. Furthermore, the 1-minute-average voltages were converted to 15-minute-average voltages by taking the mean of 15 samples at the time. In one case, the root was the transformer and in the other case, the root was randomly selected among the remaining three-phase smart meters. The mean performance index is shown in Figure 2.1.

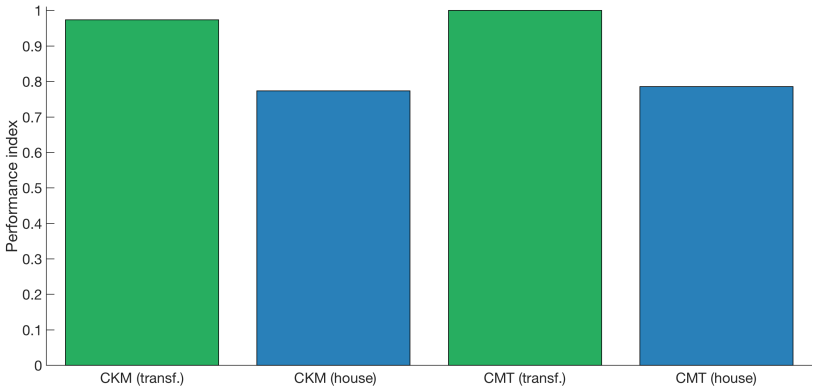


Figure 2.1: Mean performance indices comparing Constrained k-means (CKM) and Constrained Multi-trees (CMT) in demanding settings (i.e., fewer smart meters among which more are single-phase smart meters, and the sampling period is equal to fifteen minutes).

As can be seen, the mean performance indices are higher with CMT, and for both algorithms the performance is better when the root is the transformer. This is explained thanks to the radial structure of the distribution

network and the central role that is played by the distribution transformer. Indeed, the order, with which the measurements are added, and the final tree structures in each cluster are coherent with the network topology, especially with the separation of smart meters between feeders. Since, the CMT algorithm works by sequentially adding the measurement with the lowest cost and making the corresponding tree grow, it may happen that all single-phase measurements are added to the same tree, which grows unequally compared to the two other trees. Indeed, which tree is growing is only selected by the cost and no other method, which may not be optimal. Because of the high number of nodes for that tree, it may occur by chance that the similarity between one of these many nodes is high enough to capture a wrong association and thus gather two phases in the same cluster, leading to three incorrect trees. This could not occur with three-phases measurements because, by definition, one voltage measurement is associated to each tree/phase. This is probably the pitfall of the method and on its own justifies why a central element should be used as root to avoid such a situation as often as possible. Nevertheless, let us note that the multi-trees algorithm cannot reach the optimal solution if the best way to make trees grow is to connect measurements from different phases according to the given distance measure, reflecting the inherent limitation in the data that restrains the performances of the algorithm.

8 Performance in different settings

In this section, the performances of both algorithms are analysed when the original data set is modified in four different ways: (i) when one-minute average voltage time series are converted to time series averaging voltages over a longer period of time, (ii) when some random smart meters are removed from the data set, (iii) when three-phase smart meters are converted to single-phase ones by conserving a randomly selected phase, (iv) when the time window over which the phase identification is performed is less than the one from the original data set. Given the results of the previous section, the measurements from the distribution transformer are used as roots of the three clusters. Both algorithms are applied on exactly the same data, and thus results are comparable across methods. In all experiments, the performance index is the one defined in the previous section, i.e. the ratio between the measurements correctly identified and the total number of measurements.

Moreover, all experiments were repeated 100 times in order to obtain average results, since the process is partially randomised, except for the first case where only one experiment is carried out. Finally, all results are represented with box plots: on each box, the central mark is the median, while the edges of the box are the 25th and 75th percentiles. Outliers are identified and plotted individually, while the whiskers extend to the most extreme data points not considered as outliers.

8.1 Influence of the voltage-averaging period

If the smart meters' internal clocks are well synchronised, averaging the voltages over a short period of times is more reliable because it keeps more vari-

ations in the voltages.However, due to technical limitations, some equipment is not able to measure voltages at a high frequency, e.g. instead of one measure per minute, one measure is taken every five minutes, or even every 15 minutes, which is established as the standard in the power system community.

Even if it seems better to gather as much data as possible, it also leads to huge amounts of data that need to be transferred and then stored, potentially inducing a trade-off between the amount of data required to achieve good performances in phase identification, while being reasonable on the amount of data.

Figure 2.2 illustrates the performance index when the sampling period increases up to one measure per hour. When analysing the performances of the reference method, it appears that the ability to solve the phase identification decreases as the sampling period increases. As expected, the loss of information caused by the lower resolution makes it harder to correctly gather measurements. However, it also shows that (i) the performances of the constrained multi-trees are more stable, and (ii) this method is able to perfectly solve the phase identification problem, even with an extremely high sampling period.

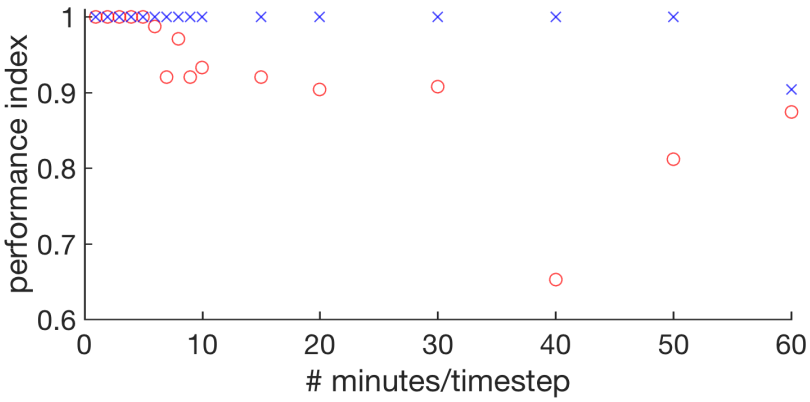


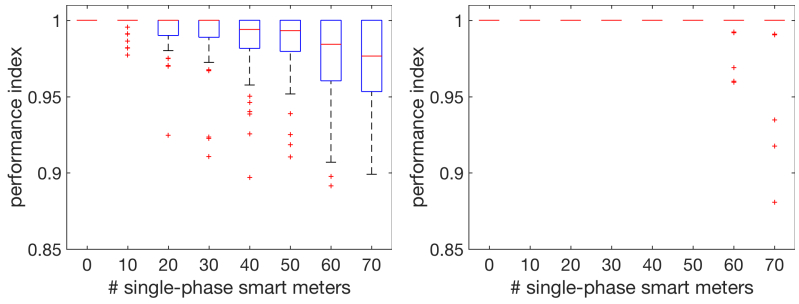
Figure 2.2: Performance index as a function of the decrease of the sampling frequency. With red circle markers, the reference (constrained k-means clustering) and with blue cross markers, the proposed method (constrained multi-trees) which outperforms the former.

8.2 Influence of ratio single-phase – three-phase smart meters

From a correlation point of view, the three-phase constraints may help to find the phase identifications for measurements. Indeed, if, in the three voltage measurements from a three-phase smart meter, two are strongly correlated to the other measurements, while the other one is not. The constraint will ensure that the final measurement is properly identified by correctly identifying the first two. However, if the same voltage measurements were to belong to three different single-phase smart meters, and thus not be constrained, the third measurement could be associated to a wrong cluster because it could be more closely correlated to measurements from the clusters to which the first and second measurements belong. Single-phase smart meters complicate the phase identification process.

Figure 2.3 displays the performance index when a certain number of randomly selected three-phase meters (the maximum of three-phase meters being 79) are converted to single-phase ones.

From Figure 2.3, the stability of the constrained multi-trees algorithm seems obvious when the number of single-phase meters increases, while the performance of the reference method decreases significantly. Indeed, measurements from single-phase smart meters are harder to correctly identify, but growing tree structures prove to be effective with a great number of single-phase smart meters.



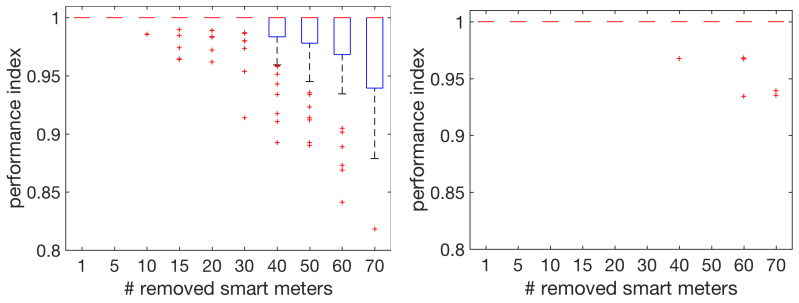
(a) Constrained k-means (b) Constrained multi-trees

Figure 2.3: Performance index according to the evolution of the number of single-phase smart meters.

8.3 *Influence of missing smart meters*

In this section, smart meters (without distinction between single phase and three phase) are randomly removed in order to virtually reduce the observability on the network, and evaluate the critical number of meters that is required to solve the identification problem. This is equivalent to increasing the number of houses that are not monitored.

Similarly, as in results from section 8.2, constrained multi-trees performance are more stable and significantly better than the reference when the data set is small (i.e. less than 40 smart meters). Given that the results for CMT only shows outliers and the median is equal to 1, the outliers probably correspond to cases where the remaining smart meters are not correlated to other measurements, e.g. where a single smart meter is the only one remaining at the end of a long feeder.



(a) Constrained k-means (b) Constrained multi-trees

Figure 2.4: Performance index according to the number of removed smart meters (i.e., a decrease in the total number of smart meters).

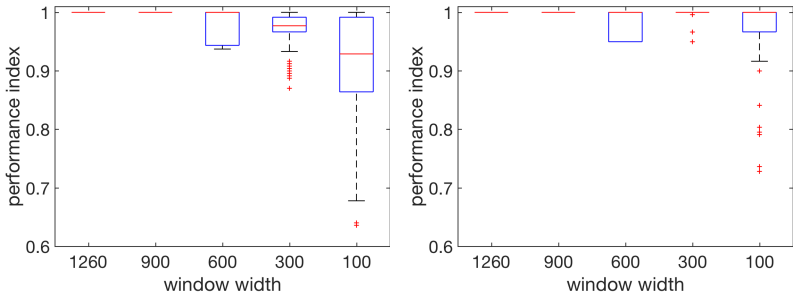
8.4 *Influence of sliding time windows*

Measurements of the original data set gather 1260 samples each, which may seem enough when measuring similarities thanks to correlation. However, one may ask if the algorithm is able to solve the identification problem with a smaller set of measurements.

To perform this experiment, only a random sub-window of a given length L is kept. In other words, L consecutive measurements were randomly ex-

tracted in order to produce a new dataset with a smaller recording time window.

As expected, smaller time windows complicate the task of phase identification. However, constrained multi-trees still manage to achieve better performance with a very small number of consecutive measurements, i.e. a very small measurement time window.



(a) Constrained k-means

(b) Constrained multi-trees

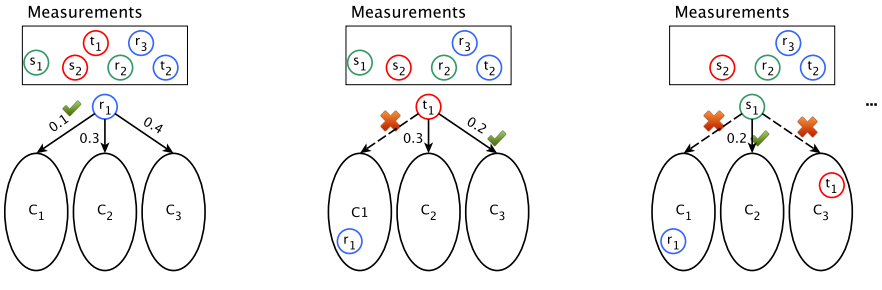
Figure 2.5: Performance index according to the decrease in the time window width.

9 *Conclusion*

This paper introduces a novel method to identify the phases of smart meters in LV distribution networks by clustering the voltage measurements using graph theory and the correlation between measurements. The algorithm, named constrained multi-tree clustering, successfully manages to identify the phases of smart meters based on real voltage measurements from an LV network in Belgium. It takes, as input, a root smart meter upon which the clustering process is done. Computations show that a good choice for the root is the meter at the distribution transformer. This is mainly due to the central position played by the distribution transformer in LV networks. The performance of the algorithm in various settings is compared to those of constrained k-means clustering. The constrained multi-tree method performs better regardless of the ratio between single-phase and three-phase smart meters, or the increasing number of houses that are not monitored, i.e. a very small set of measurements. Finally, another one of its advantages is its capacity to properly handle voltages that are averaged over a longer period of time without wrongly identifying smart meters.

In this paper, the algorithms have been tested on voltage measurements from a 3-phase 4-wire network with ungrounded neutral. Future works could include tests on measurements from other network configurations, such as (i) 3-phase 4-wire with grounded neutral, (ii) 3-phase 3-wire (3x230V) and no ground. The proposed algorithm need not be adapted since the identification is based on the fact that voltage measurements between the same phases are more correlated, than measurements between different phases. On the one hand, the performance are expected to increase in case (i) because the neutral voltage is kept low thanks to its repeated connection to the ground. On the other hand, the chance of misidentification should increase in configuration (ii) because the correlation between two phase-to-phase voltage measurements that share a common phase should be higher than the one between two phase-to-neutral voltage measurements.

We could also improve the method to handle which cluster is growing and at what pace, to avoid the growth of a cluster at the expense of the others. Finally, it could be interesting to use this novel method to infer network topology, especially since the tree-structured assumption seems very well adapted to distribution networks.



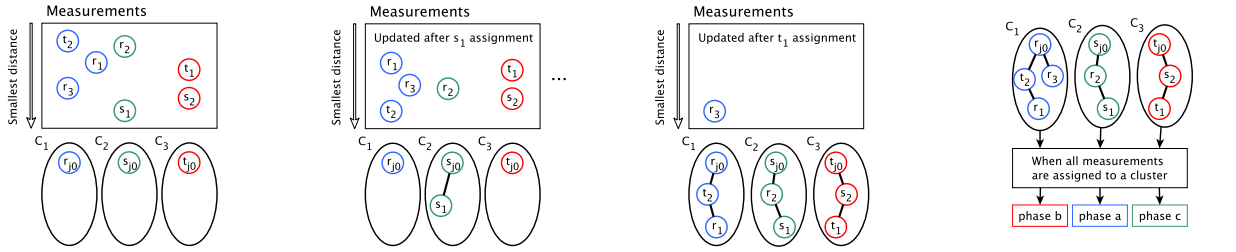
(a) To associate the first measurement from a three-phase smart meter, its distance to all three cluster centres is computed. The closest one (i.e. the one with the smallest distance) is chosen.

(b) The second measurement from a three-phase smart meter is associated based on its distance to the two remaining cluster centres, the other one being unavailable because of the first measurement.

(c) The constraints force the last measurement from a three-phase smart meter to be trivially associated to the remaining cluster.

(d) Since single-phase smart meters do not impose constraints, its measurement is associated to the closest cluster centre.

Figure 2.6: Schematic representation of the Constrained k-means algorithm applied to phase identification: initially, the cluster centres are defined by the selected root. At the beginning of every iteration, all three clusters are empty. After each iteration, the centres are computed as the average of the measurements associated to the cluster. Measurements are associated to clusters following steps (a) to (c) for three-phase smart meters and step (d) for single-phase smart meters.



(a) The measurements are sorted according to the distance δ_i between them and the closest cluster, while satisfying the constraints and the potential association of the other measurements from the same smart meter.

(b) The measurement with the smallest distance is associated to its potential cluster, expending the tree by being connected to its potential predecessor. The distances δ_i of the remaining measurements are recomputed.

(c) The process of measurement selection and association is repeated until all measurements are in a cluster.

(d) The three clusters are associated to the phases of the network either arbitrarily or by selecting a reference.

Figure 2.7: Schematic representation of the Constrained Multi-Tree algorithm.

Chapter 3
Active Management of Low-Voltage Networks for Mitigating Overvoltages due to Photovoltaic Units

In this paper, the overvoltage problems that might arise from the integration of photovoltaic panels into low-voltage distribution networks is addressed. A distributed scheme is proposed that adjusts the reactive and active power output of inverters to prevent or alleviate such problems. The proposed scheme is model-free and makes use of limited communication between the controllers, in the form of a distress signal, only during emergency conditions. It prioritizes the use of reactive power, while active power curtailment is performed only as a last resort. The behavior of the scheme is studied using dynamic simulations on a single low-voltage feeder and on a larger network composed of 14 low-voltage feeders. Its performance is compared to a centralized scheme based on the solution of an Optimal Power Flow problem, whose objective function is to minimize the active power curtailment. The proposed scheme successfully mitigates overvoltage situations due to high photovoltaic penetration and performs almost as well as the Optimal Power Flow based solution with significantly less information and communication requirements.

1 Introduction

During the last decades, the cost of PhotoVoltaic (PV) panels has been continuously decreasing. It is estimated that with each doubling of installed capacity, the cost of PV installations decreases by 20%.[1] This leads to the rapid growth of PV installations in Low-Voltage (LV) Distribution Networks (DNs).

However, the presence of power generation inside LV DNs changes the voltage profile of the feeders.[2] If the total installed PV power is larger than the feeder hosting capacity, i.e. the maximum amount of PV that can be accommodated, network security cannot be guaranteed.[3] Specifically, when the production of a feeder surpasses its consumption, a reverse power flow occurs which leads to overvoltages and might cause problems to the coordination of protective devices and disconnection of equipment for security reasons.[4] This important problem is studied in this paper.

The classical approach for addressing this issue is to rely on hefty investments to upgrade and reinforce the networks. However, many companies and researchers are looking at better ways to use existing equipment by develop-

[1] Maycock (*The future of energy summit 2013. bloomberg new energy finance*), Bazilian et al. ("Re-considering the economics of photovoltaic power"), and Schleicher-Tappeser ("How renewables will change electricity markets in the next five years")

[2] Masters ("Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines")

[3] Schwaegerl et al. ("Voltage control in distribution systems as a limitation of the hosting capacity for distributed energy resources")

[4] Barker and Mello ("Determining the impact of distributed generation on power systems. I. Radial distribution systems")

ing and designing flexible and inexpensive control schemes to limit those investments and increase the hosting capacity of the networks.[5]

These schemes, referred to under the term of Active Network Management, usually control the DNs' power generation, consumption or storage to prevent or mitigate overvoltage problems.

1.1 Literature review

Some of the schemes proposed in the literature can be classified as centralized: the control actions are computed by a common entity responsible for gathering information about the network, processing them according to some optimization objectives and constraints, and sending the set-points back to the actuators. In such schemes, the computation of the control actions often relies on an Optimal Power Flow (OPF) formulation of the problem and requires an extended communication infrastructure as well as a network model.

Liew[6] proposes an OPF formulation with the objective to minimize the market value of the curtailed energy of embedded wind generation is proposed. Other objectives include minimizing the voltage profile deviation from a reference,[7] the transformer tap changer switching,[8] or the network losses.[9] Yeh [10] studies a combination of those objectives. In Gemine's work[11], a sequential decision making problem under uncertainty is formulated where the Distribution System Operator has the choice to reserve the availability of flexible demand and, in the subsequent steps, to curtail generation and vary flexible loads. The aforementioned centralized schemes are defined and simulated in the framework of Medium Voltage (MV) networks.

Following, a second category of schemes are distributed: the units are controlled in a distributed way with no centralized control entity. The distributed controllers often use local information to adjust each unit individually.

Concerning the overvoltage problems considered in this work, it has been proposed to change the reactive power production[12] or power factor[13] of PV units, as a function of their terminal voltage. Alternatively, the use of active power curtailment[14] or storage of the excess energy in batteries[15] have been suggested.

Carvalho and Aghaterhrani[16] suggested to compensate the voltage variations caused by changes in the PV generation with the use of reactive power and appropriate voltage sensitivity models. An adaptive control of reac-

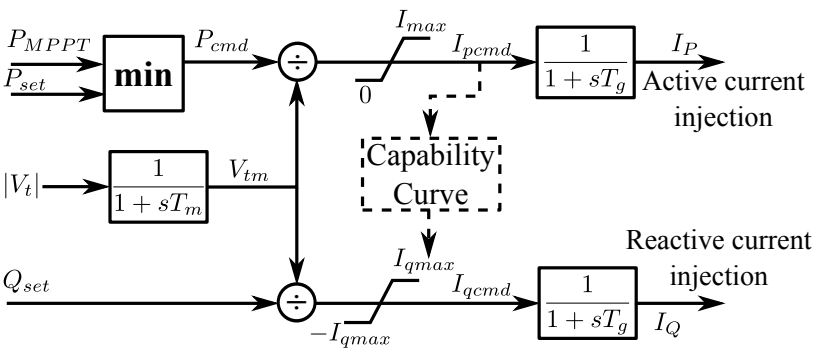


Figure 3.1: PV dynamic model

[5] J. P. Lopes et al. ("Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities")

[6] Liew and Strbac ("Maximising penetration of wind generation in existing distribution networks")

[7] Bonfiglio et al. ("Optimal control and operation of grid-connected photovoltaic production units for voltage support in medium-voltage networks")

[8] Agalgaonkar, Pal, and Jabr ("Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators")

[9] Farivar et al. ("Optimal inverter VAR control in distribution systems with high PV penetration") and L. F. Ochoa and Harrison ("Minimizing energy losses: optimal accommodation and smart operation of renewable distributed generation")

[10] Yeh, Gayme, and Low ("Adaptive VAR control for distribution circuits with photovoltaic generators")

[11] Gemine et al. ("Active network management: planning under uncertainty for exploiting load modulation")

[12] M.H.J. Bollen and Sannino ("Voltage control with inverter-based distributed generation") and Turitsyn et al. ("Options for control of reactive power by distributed photovoltaic generatorss")

[13] Demirok et al. ("Local reactive power control methods for overvoltage prevention of distributed solar inverters in low-voltage grids") and Garcia-Gracia et al. ("Integrated control technique for compliance of solar photovoltaic installation grid codes")

[14] Tonkoski, L. A. C. Lopes, and El-Fouly ("Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention")

[15] M. J. E. Alam, Muttaqi, and Sutanto ("Mitigation of rooftop solar PV impacts and evening peak support by managing available capacity of distributed energy storage systems")

[16] Carvalho, Correia, and L. Ferreira ("Distributed reactive power generation control for voltage rise mitigation in distribution networks") and Aghatehrani and Golnas ("Reactive power control of photovoltaic systems based on the voltage sensitivity analysis")

tive power production is proposed by Kakimoto,[17] providing a compromise between the operation of the PV unit at unitary power factor and voltage control. Xin[18] explicitly uses communication to make a group of PV units converge to a percentage of the available power and regulate a critical bus. In Farag's work,[19] the voltage of the MV network is regulated thanks to a two-way communication between the transformer tap changer, the capacitor banks and the distributed generation units.

Varma[20] presents the benefits of using an inverter with STATCOM-like capabilities to regulate voltage variations caused by other sources or loads during night. Finally, Vovos[21] performs a comparison between an OPF-based centralized scheme and Kiprakis's distributed scheme.[22]

A third category of control schemes, consisting of a combination of centralized and distributed schemes, are referred to as decentralized. More specifically, they are composed of local controllers and a centralized entity which computes the control law to be sent to them, so some communication is always needed. Once the control law is received, inverters do not need other control orders from the central entity to operate. Many papers[23] propose a decentralized control scheme of the reactive power to control the voltage and/or minimize system losses. Based on voltage sensitivity analysis, Di Fazio *et al.* suggest a control scheme to achieve the same goals, while Ahn[24] aims at minimizing the losses inside a microgrid.

1.2 Contributions of this paper

In this paper, a *distributed* control scheme that changes the active and reactive power injected by PV units in LV DNs is proposed. The objective of the control algorithm is to mitigate overvoltage problems by directing PV units to consume reactive power and, if necessary, to curtail active power generation. The distributed controllers are implemented on the PV inverters with five modes of operation.

First, if there are no overvoltage problems in the LV feeder, the controllers act preventively adjusting the PV units' reactive power to avert their occurrence while performing Maximum Power Point Tracking (MPPT) for active power. In this mode of operation, only local measurements are used and no communication among the controllers is needed. Second, if overvoltage problems occur, the controllers make use of limited communication for coordinating their reactive power consumption, within each DN LV feeder. Third, if the overvoltage persists even after all PV units have utilized their maximum reactive capabilities, the controllers switch to active power curtailment. Finally, the fourth and fifth modes of operation restore active and reactive power production to normal operation. The proposed scheme is model-free, as the model and parameters of the DN are not required by the controllers.

The behavior of the proposed control scheme is studied first on a single LV feeder and then on a larger network composed of 14 low-voltage feeders. Following, it is compared to a centralized OPF-based scheme applied to the same systems. Their performance is assessed based on their ability to alleviate the overvoltage problems and the amount of active power curtailed to achieve this. Both control schemes manage to ensure the security of the

[17] Kakimoto, Piao, and Ito ("Voltage control of photovoltaic generator in combination with series reactor")

[18] H. Xin et al. ("Cooperative control strategy for multiple photovoltaic generators in distribution networks") and Huanhai Xin et al. ("A self-organizing strategy for power flow control of photovoltaic generators in a distribution network")

[19] H. E. Farag, El-Saadany, and Seethapathy ("A two ways communication-based bistributed control for voltage regulation in smart distribution feeders")

[20] Varma, Khadkikar, and Seethapathy ("Nighttime application of PV solar farm as STATCOM to regulate grid voltage")

[21] Vovos et al. ("Centralized and distributed voltage control: impact on distributed generation penetration")

[22] Kiprakis and Wallace ("Maximising energy capture from distributed generators in weak networks")

[23] Samadi et al. ("Coordinated active power-dependent voltage regulation in distribution grids with PV systems"), Tanaka et al. ("Decentralised control of voltage in distribution systems by distributed generators"), Cagnano et al. ("Online optimal reactive power control strategy of PV inverters"), and Calderaro et al. ("Optimal decentralized voltage control for distribution systems with inverter-based distributed generators")

[24] Ahn and Peng ("Decentralized voltage control to minimize distribution power loss of microgrids")

system. However, the OPF-based does this with a little less curtailed active power.

On the other hand, the main contribution of the proposed scheme is that it does not require a network model or remote measurements, thus making it easy to deploy. Moreover, the use of communication is limited to a distress signal that can be easily implemented with the use of Power-Line Communication (PLC).[25] This technology has been exploited for several decades to provide low-cost remote switching capabilities to utilities; one such example is the day/night tariff signal used in some countries.[26]

Additionally, it does not require a centralized entity to collect measurements, compute set-points and dispatch the PV units. Thus, it is more robust to component failures. Finally, when active power curtailment is required to secure the system, the proposed scheme is designed to proportionally share the burden among all the PV units in the problematic feeder.

The chapter is organized as follows. In Section 2 the dynamic model of the PV units is presented. In Section 3 the proposed distributed control is detailed. Section 4 introduces the centralized OPF-based scheme used for performance evaluation. The test systems and simulation results are reported in Section 6. Finally, conclusions and future work are presented in Section 6.

[25] Galli, Scaglione, and Z. Wang ("Power line communications and the smart grid")

[26] Carcelle (*Power Line Communications in Practice*)

2 PV unit dynamic model

The dynamic PV unit model selected for this study is detailed in Figure 3.1. The closed-loop voltage regulator and the DC dynamics have been neglected for simplicity.[27] The model reflects active power priority with the active current command (I_{pcmd}) limited by the rating of the inverter (I_{max}). P_{set} and Q_{set} are the active and reactive power set-points computed by the controller detailed in the next section. When $P_{set} \geq P_{MPP}$, the unit operates in MPPT mode. T_g ($\sim$20 ms) and T_m ($\sim$50 ms) are the inverter current and voltage measurement time constants, respectively. Finally, the limits on the reactive current command (I_{qcmd}) are calculated from the unit's reactive power capability curve using I_{pcmd}.

[27] WECC REMTF (*Generic solar photovoltaic system dynamic simulation model specification*)

The capability curves of PV inverters are usually defined by the standards and grid codes of the country. In the recent years, numerous international and national standards and guidelines have been published, by different types of organizations (e.g. CENELEC, VDE, CEI), to introduce the new concept that LV active users have to provide some sort of ancillary services to the grid by adjusting their reactive power exchanged.[28]

For example, the German standard VDE-AR-N 4105[29] dictates that all DGs connected to LV grids should apply a Power Factor (PF) adjustment in order to contribute to the network voltage regulation according to their nominal power. For units smaller than 3.68 kVA a PF between $0.95_{leading}$ to $0.95_{lagging}$ is necessary according to DIN EN 50438,[30] for units between 3.68 and 13.8 kVA a characteristic curve provided by the network operator within $PF = 0.95_{leading}$ to $PF = 0.95_{lagging}$ and for larger units a characteristic curve provided by the network operator within $PF = 0.90_{leading}$ to $PF = 0.90_{lagging}$.

[28] Caldon, Coppo, and Turri ("Coordinated voltage control in MV and LV distribution networks with inverter-interfaced users"), Kotsampopoulos et al. ("Review, analysis and recommendations on recent guidelines for the provision of ancillary services by distributed Generation"), and Caldon, Coppo, and Turri ("Distributed voltage control strategy for LV networks with inverter-interfaced generators")

[29] VDE (*VDE-AR-N 4105: Generators connected to the LV distribution network - Technical requirements for the connection to and parallel operation with low-voltage distribution networks*)

[30] Kotsampopoulos et al. ("Review, analysis and recommendations on recent guidelines for the provision of ancillary services by distributed Generation")

The latter PF ranges are visualized by the triangular shaded area in Figure 3.2 and the reactive current limit can be computed as:

$$I_{qmax} = I_{pcmd} \tan\left(\cos^{-1} PF_{\min}\right) \tag{3.1}$$

It can be seen that PV inverters must be oversized to comply with the PF requirements even at full active power generation (Curve A, Figure 3.2).

Likewise, the recent Italian standard CEI 0–2[31] states similar operating conditions for DGs connected to LV networks. For units between 3 and 6 kVA, an adjustable PF of $0.95_{leading}$ to $0.95_{lagging}$ is necessary, while for larger units a rectangular capability is required as sketched in Figure 3.2.[32]

Finally, while it is technically possible to design PV inverters to provide reactive support at smaller PF values or even if solar input is zero, much like a STATCOM, this functionality is not standard in the industry. The standards dictate the upper and lower bounds on reactive power with respect to active power and rating of the inverter, and manufacturers must comply with these. However, such functionality might be provided by future PV inverters. Thus, in Section 5.2 the benefits arising from this operation are discussed.

[31] CEI (*CEI 0–21: Reference technical rules for connecting active and passive users to networks low-voltage electricity networks of energy providers*)

[32] Caldon, Coppo, and Turri ("Distributed voltage control strategy for LV networks with inverter-interfaced generators")

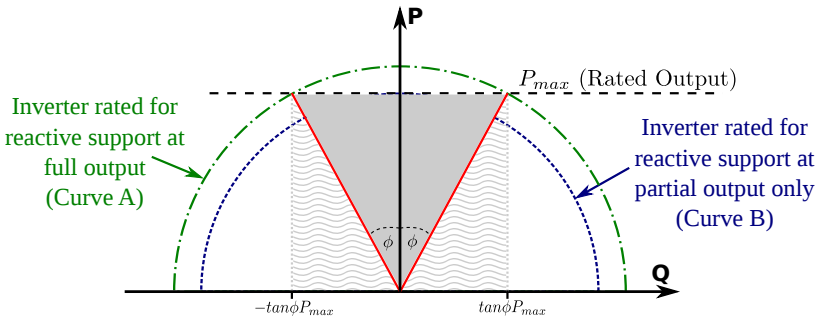


Figure 3.2: PV capability curve under constant voltage

3 *Distributed control scheme*

The proposed controller is implemented at the inverter level, taking as input the locally measured terminal voltage (V_{tm}) and setting the active and reactive power set-points (P_{set}, Q_{set}). Each controller is implemented as a discrete device, updating the control actions with a period T_{upd}, that is the n-th action takes place at time $t_n = nT_{upd}$. This discrete nature is due to the use of embedded microcontrollers and the communication, measuring and computing delays involved in the procedure.

As summarized in the Introduction, the controllers have five modes of operation. They are shown in Figure 3.3, and detailed in the remaining of this section.

3.1 *Mode A: Normal operating conditions*

During normal operating conditions, when all PV terminal voltages in the LV feeder are below a predefined maximum V_4, the PV units follow an MPPT logic for active power set-point and adjust the reactive power as a function of their terminal voltage, as shown in Figure 3.4, inspired Dugan's work,[33] and firstly introduced by Bollen.[34] The reactive power adjustment is aimed

[33] Dugan, Sunderman, and Seal ("Advanced inverter controls for distributed resources")

[34] M.H.J. Bollen and Sannino ("Voltage control with inverter-based distributed generation")

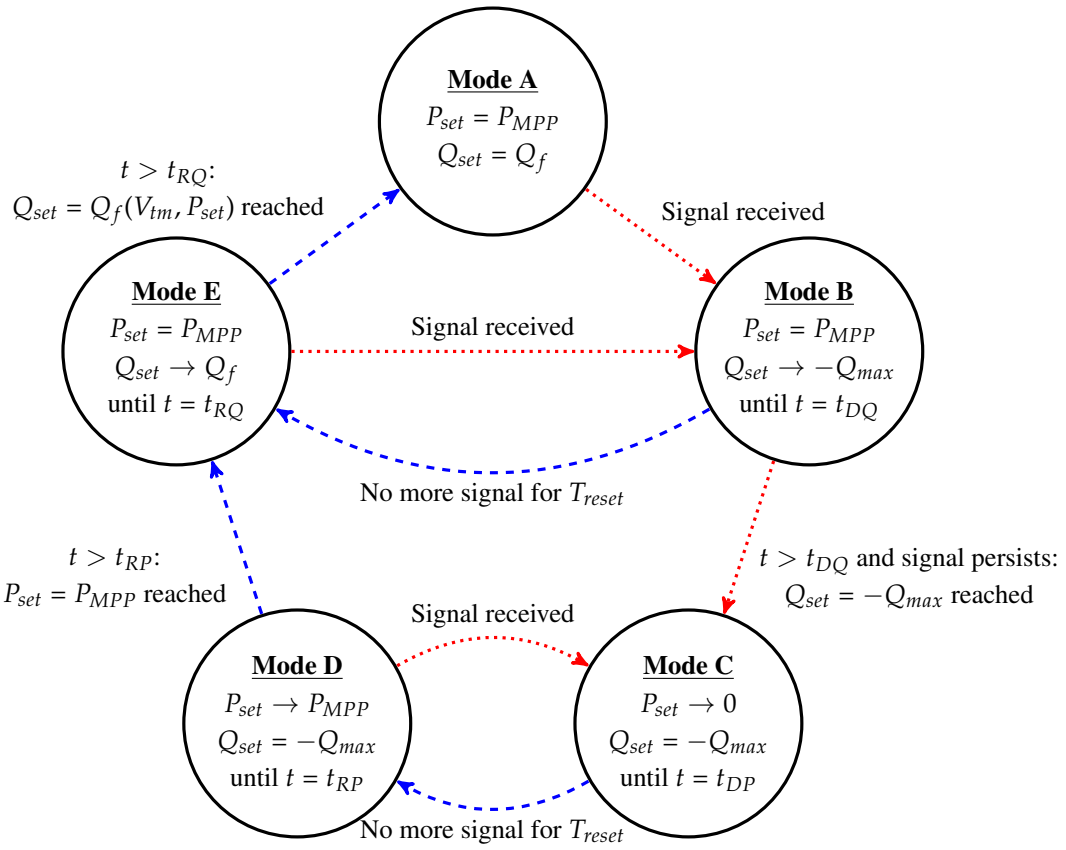


Figure 3.3: State transition diagram of the distributed control scheme. The red dotted lines are the emergency control transitions while blue dashed lines are the restoring ones. t_{DQ} (resp. t_{DP}) is the time needed in Mode B (resp. Mode C) to use all available reactive (resp. active) controls. T_{reset} is the elapsed time without emergency signal for the controller to start restoring active/reactive power. t_{RP} (resp. t_{RQ}) is the time needed in Mode D (resp. Mode E) to restore active (resp. reactive) power to the set-point values of Mode A. P_{set} and Q_{set} are the active and reactive power set-points of the controller. P_{MPP} is the maximum available active power of the PV module and depends on the solar irradiation. Q_{max} is the maximum available reactive power; it varies according to the capability curve (e.g. Figure 3.2) as a function of the active power output.

at counteracting overvoltages when V_{tm} exceeds V_3. The symmetrical part of the reference figure is used in undervoltage situations, which are not considered in this study.

In this mode of operation only local measurements are used and no communication among the controllers is needed.

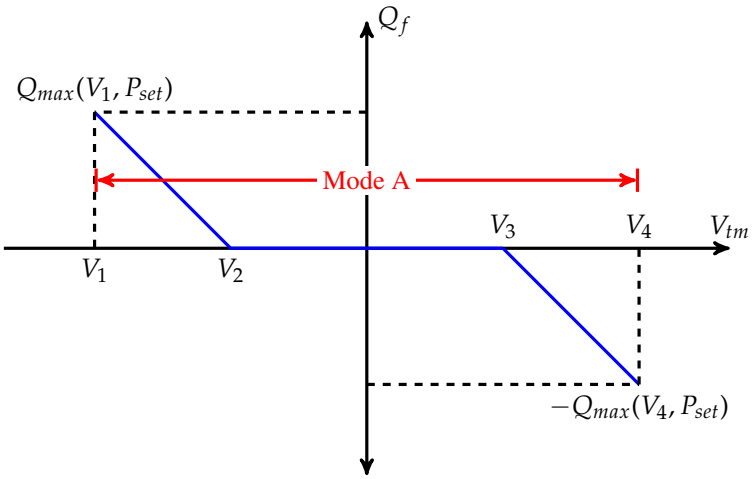


Figure 3.4: $Q_f(V_{tm}, P_{set})$ function for Mode A. For $V_{tm} \geq V_4$ an emergency signal is issued and the controller moves to Mode B.

3.2 *Mode B: Coordinated reactive power adjustment*

A PV unit whose terminal voltage has reached V_4, has already made full use of its reactive power adjustment capability and cannot mitigate the overvoltage by itself without proceeding to active power curtailment. At this moment and for as long as the overvoltage persists, a repeating distress signal is sent to all PV controllers in the same feeder as sketched in Figure 3.5.

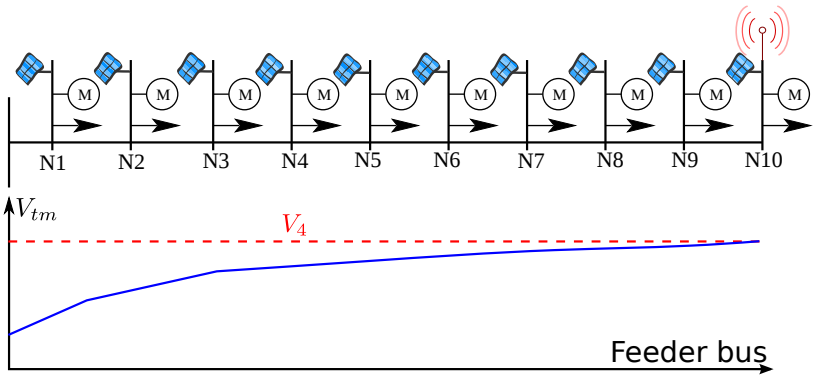


Figure 3.5: Feeder emergency signal

Upon receiving this signal, all PV controllers in the same feeder start adjusting their reactive power consumptions to decrease the terminal voltage of the distressed PV(s). During this mode of operation, the PV units follow an MPPT logic for active power set-point while increasingly consuming more reactive power with the target of reaching their $-Q_{max}$ values at $t = t_{DQ}$. More specifically, Q_{set} is adjusted as follows:

$$Q_{set}[t_n] = Q_{set}[t_{n-1}] + (-Q_{max}[t_n] - Q_{set}[t_{n-1}]) \frac{t_n - t_{n-1}}{t_{DQ} - t_{n-1}} \quad (3.2)$$

where the value of $Q_{max}[t_n] = Q_{max}(V_{tm}[t_n], P_{set}[t_n])$ is updated at each step based on the capability curve (Figure 3.2) and t_{DQ} is calculated based on the first distress signal received. An example of this adjustment is presented in Figure 3.6.

This mode of operation has two possible outcomes. If the overvoltage problem persists after t_{DQ}, the PV controllers proceed to coordinated active power curtailment (Mode C) while having reached $Q_{set} = -Q_{max}$. On the other hand, if the problem has been resolved only with the use of reactive power adjustments, the PV controllers freeze their reactive power set-points while continuing to follow the MPPT logic for active power. After a time T_{reset} without new distress signals, the PV units move to Mode E, where they try to restore Q_{set} to the value given in Figure 3.4 for normal operating conditions (Mode A).

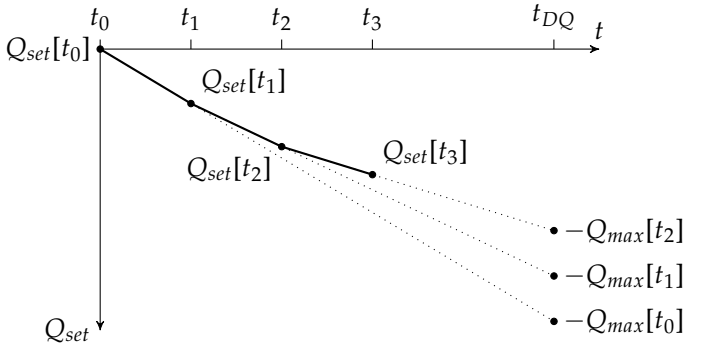


Figure 3.6: Mode B: Q_{set} adjustment

3.3 Mode C: Active power curtailment

In this mode, the PV controllers stop applying MPPT and start curtailing active power according to:

$$P_{set}[t_n] = P_{out}[t_0]\left(1 - \frac{t_n - t_{n-1}}{t_{DP} - t_{n-1}}\right) \tag{3.3}$$

where $P_{out}[t_0]$ is the PV unit's active power output at the moment t_0 of entering Mode C and t_{DP} is the time by which all PV units need to have curtailed all their active power outputs, calculated based on the time of entering Mode C.

In this mode, the reactive power set-point keeps being adjusted to $-Q_{max}(V_{tm}[t_n], P_{set}[t_n])$, with the latter being updated due to the variations of the active power and the change of terminal voltage, as dictated by the capability curve. That is, for the capability curve sketched in Figure 3.2, Q_{max} will reach zero at t_{DP} as the active power reaches zero.

If the overvoltage problem is resolved, that is no more emergency signals are received, the PV controllers freeze their power set-points and timers. Then, after a period of T_{reset} without new distress signals, the controllers move to Mode D.

However, if the PV units in the distressed feeder have exhausted all their possible controls and the overvoltage problem persists, then the problem was not created by the PV generation. At the end of Mode C, the active power

output of all PV units is at zero. Any overvoltages in such passive feeders is a consequence of MV voltages or MV/LV transformer set-points. In such a particular case, the controllers remain in Mode C so as not to further aggravate the problem.

3.4 Mode D: Restoring active power generation

The PV controllers reach this mode after having curtailed active power generation (Mode C) and not receiving a distress signal over a period T_{reset}. The purpose of this mode is to smoothly and uniformly restore the active power generation of the PV units to the MPP values, without creating new overvoltage problems. Thus, the active power set-point is modified as:

$$P_{set}[t_n] = P_{max}\frac{t_n - t_{n-1}}{t_{RP} - t_{n-1}} \tag{3.4}$$

where P_{max} is the PV unit's nominal active power output and t_{RP} is the time by which all PV units need to have restored to their MPP value, calculated based on the time of entering Mode D.

During this active power restoration mode, the reactive power set-point is fixed to $-Q_{max}(V_{tm}[t_n], P_{set}[t_n])$, with the latter being updated due to the variations of the active power and the change of terminal voltage, as dictated by the capability curve (Figure 3.2).

If the problem reoccurs (a new distress signal is received) during the active power increase, the PV units move back to Mode C. Otherwise, at time $t = t_{RP}$, all the PV units have reached $P_{set} = P_{max}$, hence MPP output, and they move to Mode E.

3.5 Mode E: Restoring reactive power to Mode A set-point

The PV units in a feeder reach this mode either from Mode D or directly from Mode B after T_{reset} time without a distress signal. The purpose of this mode is to smoothly adjust the reactive power set-points Q_{set} of the PV units according to Figure 3.4 while keeping active power output to MPP values. The reactive power set-point is modified as:

$$Q_{set}[t_n] = Q_{set}[t_{n-1}] + \left(Q_f[t_n] - Q_{set}[t_{n-1}]\right)\frac{t_n - t_{n-1}}{t_{RQ} - t_{n-1}} \tag{3.5}$$

where the value of $Q_f[t_n] = Q_f(V_{tm}[t_n], P_{MPP}[t_n])$ and t_{RQ} is the time by which all PV units need to have restored their reactive powers, calculated based on the time of entering Mode E.

If an overvoltage problem re-occurs (a new distress signal is received) during the reactive power restoration, the PV units move back to Mode B. Otherwise, at time $t = t_{RQ}$, all the PV units have reached $Q_{set} = Q_f(V_{tm}, P_{MPP})$ and, hence, they have come back to Mode A.

3.6 General considerations

The proposed distributed scheme makes no use of the network model or parameters. Moreover, it does not need information on the position of each PV inside the LV feeder. In normal operating conditions (Mode A), there is no

exchange of information among the controllers. Limited communication, in the form of a distress signal, is needed for Modes B and C.

The proposed scheme is designed so that at the end of Mode C all PV units in the distressed feeder will have curtailed the same percentage of their active power at normal operating conditions ($P_{out}[t_0]$). In this way, the financial loss from curtailing active power is shared evenly amongst all the PV units. In reality though, as will be shown in Section 6, some minor differences between them exist due to communication delays and the discrete and asynchronous nature of embedded controllers.

In this work it is assumed that the voltages without the PV injections (only loads) are within the acceptable limits. If this assumption does not hold, it is possible to meet situations where in the same LV feeder exist both buses with over- and under-voltage conditions. If this scenario occurs, the units with undervoltage problems will not contribute to the coordinated reduction but will remain in Mode A, that is producing P_{MPP} and adjusting their reactive power according to Figure 3.4. Such cases, however, are not considered in this study.

4 Performance evaluation

To assess its performance, the proposed scheme is compared to a centralized one whose objective is to minimize the active power curtailment while satisfying the network voltage constraints. This type of optimization problem belongs to the general class of OPF problems.

4.1 Optimal Power Flow formulation and solution

Let $\mathcal{B}$ be the set of buses of the network. Each bus is characterized by the magnitude V and the phase angle θ of its voltage. Let $\mathcal{PV}$ be the subset of buses of $\mathcal{B}$ to which PV units are connected. A PV unit is characterized by five elements: P_{PV}, the active power supplied by the inverter to the grid, Q_{PV}, the reactive power supplied to the grid, P_{PV}^{MPP}, the maximum active power the PV panels can produce given the current sunlight and temperature, $I_{PV}^{\max}$, the maximum current of the inverter, and $PF_{PV}^{\min}$, the minimal power factor under which the inverter can operate.

The minimization of the total curtailed active power in the whole network is taken as objective:

$$\min_{P_{PV},Q_{PV},V,\theta} \quad \sum_{j\in\mathcal{PV}} P_{PV_j}^{MPP} - P_{PV_j} \tag{3.6}$$

subject to

$$h(P_{PV}, Q_{PV}, P_{inj}, Q_{inj}, V, \theta) = 0 \tag{3.7}$$

$$V^{\min} \leq V_j \leq V^{\max}, \quad j \in \mathcal{PV} \tag{3.8}$$

$$0 \leq P_{PV_j} \leq P_{PV_j}^{MPP}, \quad j \in \mathcal{PV} \tag{3.9}$$

$$\sqrt{(P_{PV_j})^2 + (Q_{PV_j})^2} \leq I_{PV_j}^{\max} \cdot V_j, \quad j \in \mathcal{PV} \tag{3.10}$$

$$\frac{P_{PV_j}}{\sqrt{(P_{PV_j})^2 + (Q_{PV_j})^2}} \geq PF_{PV_j}^{\min}, \quad j \in \mathcal{PV} \tag{3.11}$$

- equation (3.7) is a compact notation for the power flow equations, P_{inj} (resp. Q_{inj}) is a vector of active (resp. reactive) powers injected at each bus by other equipment than the PV installations (loads are considered to inject negative power);
- inequality (3.8) forces voltages to stay within their limits;
- inequality (3.9) stresses that the active power injected by the inverter should be positive and no larger than the maximal power that the PV panel can produce;
- inequality (3.10) sets a limit on the current the inverter can supply;
- inequality (3.11) puts a limit on the power factor.

It should be noted that the constraints defined here are the same as the one directly or indirectly implemented in the distributed scheme. The problem is solved with MATLAB optimization toolbox, using interior point method.

4.2 Comparison with a centralized scheme

To assess the performance of the distributed scheme, the following procedure was used.

A step change from zero to a chosen value is applied to P_{MPP}, aiming to cause overvoltage problems. The response of the system with the use of distributed controllers is simulated until an equilibrium point is reached. The total amount of active power produced or curtailed and the reactive power absorbed by the PV units are then extracted.

Similarly, the system's response to the same step change of P_{MPP} is simulated when the OPF problem (5.21)–(3.11) is solved to acquire the PV units' set-points. The corresponding values of the total active power produced or curtailed, the reactive power absorbed by the PV units and the network losses are used to evaluate the performance of the distributed scheme.

5 Simulation results

The performance of the proposed scheme has been tested using dynamic simulations, first on a single LV feeder, then on a larger MV/LV network including 14 LV feeders. The simulations, in phasor mode, have been performed using the RAMSES [35] software developed at the University of Liège.

[35] Aristidou, Fabozzi, and Van Cutsem ("Dynamic simulation of large-scale power systems using a parallel schur-complement-based decomposition method")

5.1 Single feeder example

The single LV feeder in Figure 3.7 is used to further illustrate the various modes of operation and the behavior of the distributed controllers. Each bus includes a PV unit, an equivalent motor and a voltage dependent load. For the sake of simplicity the same fraction of motor load is considered for all loads, namely 30%. Each equivalent induction motor has a 6-kVA rated apparent power and its model accounts for the presence of a double-cage rotor.

The lines have a X/R ratio of 0.89. The X/R ratio dictates how much the voltages are affected by active power or reactive power changes. If the X/R ratio is high, then reactive power dominates the voltage changes. On

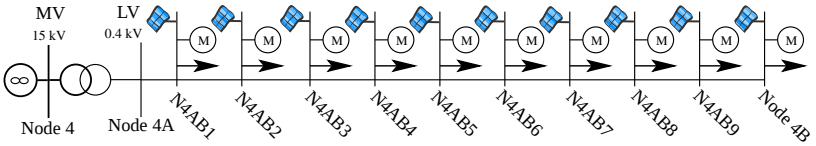


Figure 3.7: Single feeder model

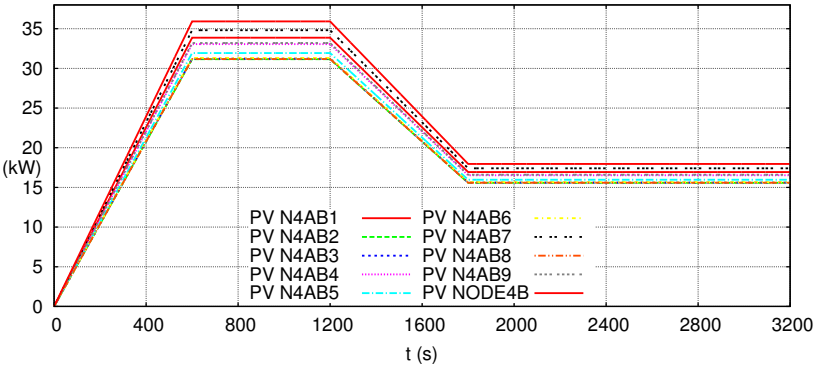


Figure 3.8: Single feeder input power (P_{MPP})

the other hand, distribution networks typically have a low X/R ratio. Hence, reactive support may not be sufficient to control voltages and active power changes are needed.[36] However, the distributed algorithm controls first the reactive power as it is considered to be a cheap resource from the producer point of view, and then active power. The outcome does not depend on the X/R ratio but it affects how much active power will be curtailed to prevent overvoltages, if any.

[36] Morren, Haan, and J. A. Ferreira ("Contribution of DG units to voltage control: active and reactive power limitations")

The nominal power of each PV unit was randomly selected between 30.5 and 37.5 kW. The model follows the capability curve in Figure 3.2 and is equipped with the proposed controller with the activation period T_{upd} randomly drawn between 22 and 30 seconds. Finally, the control parameters were chosen at $V_4 = 1.07$ pu, $V_3 = 1.02$ pu, $PF_{min} = 0.95$, $T_{DQ} = 300$ s, $T_{DP} = 300$ s, $T_{RP} = 600$ s, $T_{RQ} = 600$ s and $T_{reset} = 300$ s.

The smooth solar power (P_{MPP}) variation in Figure 3.8 is considered over a period of 3200 s. P_{MPP} starts from zero and reaches a randomly chosen value between 90 and 100% of the PV unit's nominal power after 600 s. From 1200 s on, P_{MPP} smoothly decreases to reach 50% of its maximal value. The loads vary only with their voltages during this period of time.

Initially, all PV units are at zero generation and the LV feeder is importing its whole active power from the MV system. Moreover, as all terminal voltages are lower than V_4, the distributed controllers are in Mode A.

As PV generation increases, due to the increase of P_{MPP}, the imported active power decreases at the profit of locally generated one. This can be seen in Figure 3.9, depicting the active power transfer through the transformer. From the same figure, it can be seen that at $t \approx 200$ s, the LV feeder starts exporting active power to the MV network.

Figures 3.10 to 3.13 show the various state transitions of the controllers, the voltages at the connection points of the PV units, the active and reactive power output of the inverters, respectively.

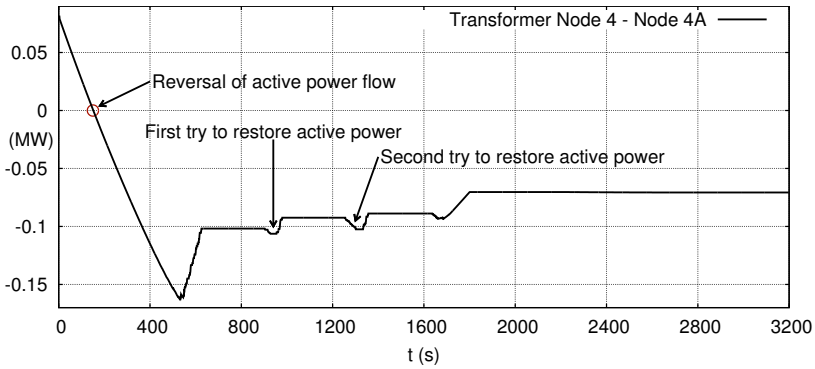


Figure 3.9: Active power flow in the MV/LV transformer of the feeder (negative values mean exportation of power to the MV level)

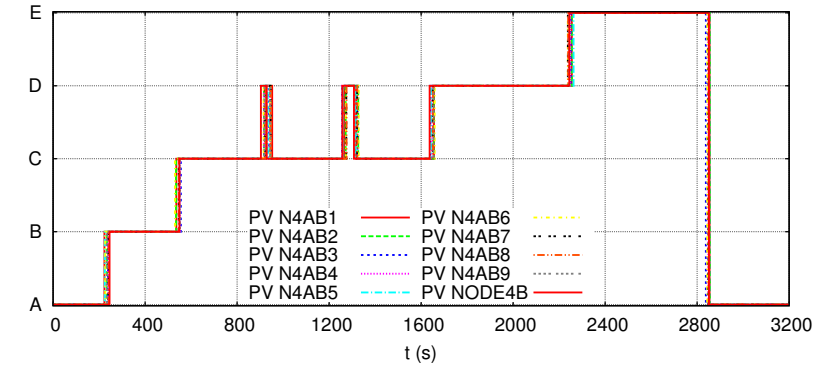


Figure 3.10: Single feeder state transitions

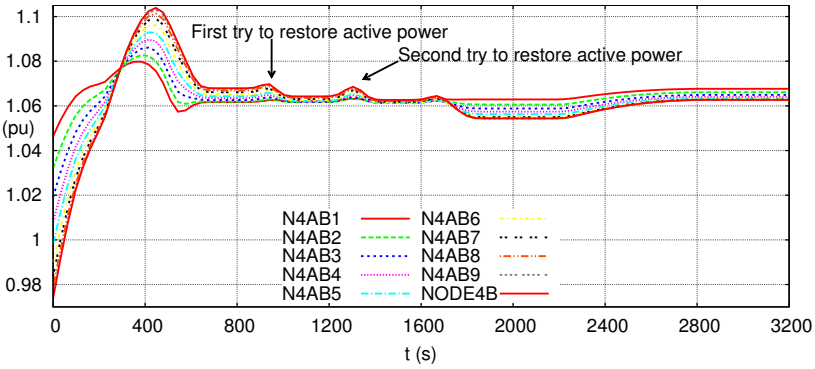


Figure 3.11: Single feeder voltages

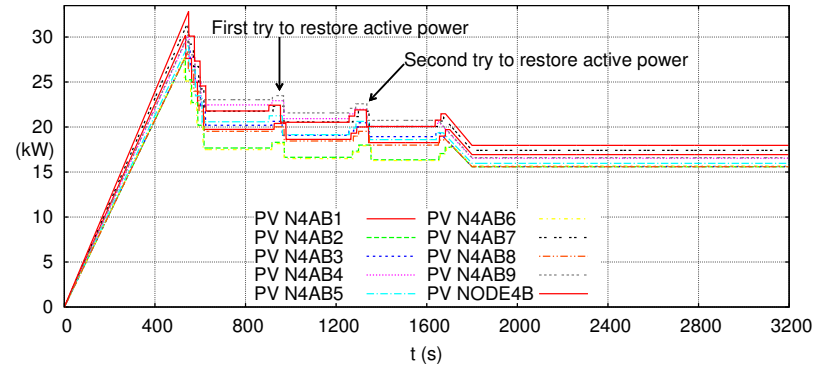


Figure 3.12: Single feeder PV active power

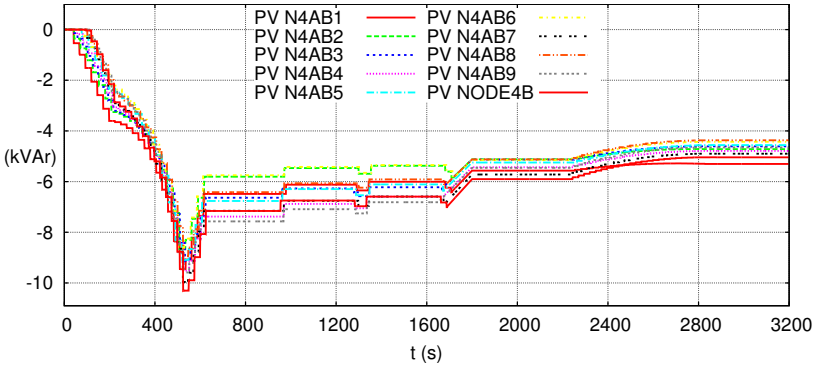


Figure 3.13: Single feeder PV reactive power

Along with the PV active power increase, the LV bus voltages rise and the reactive power is adjusted according to Figure 3.4. This leads, in turn, to a smooth increase of the reactive power absorption by the inverters as seen in Fig.3.13.

At $t \approx 250$ s, an overvoltage takes place and a distress signal is sent by PV unit *N4AB1* (see Figure 3.7) through the feeder. Consequently, all controllers move to Mode B and gradually absorb more reactive power targeting to reach their maximum after $T_{DQ} = 300$ s, i.e. $t = t_{DQ} \approx 550$ s.

Since the attempt to deal with the overvoltage by means of reactive power adjustments fails, the controllers go to Mode C and start curtailing active power at $t \approx 550$ s. As can be seen in Figure 3.12, active power is curtailed until all voltages decrease under $V_4 = 1.07$ pu at $t \approx 600$ s.

Following, after a period T_{reset} with no other overvoltage alarm, at $t \approx 900$ s the controllers enter the restoration phase and try to increase the active power production (Mode D). However, this increase results in overvoltage and the controllers move back to Mode C. This attempt is reiterated after a second T_{reset} period, at $t \approx 1250$ s, with the same result.

The smooth decrease of P_{MPP} after $t = 1200$ s limits the active power production. Thus, during the third attempt in Mode D, at $t \approx 1620$ s, active power restoration to P_{MPP} is successful.

In the last part of simulation, the controllers move to Mode E and slowly decrease their reactive power absorption. Once reaching the values dictated by the curve in Figure 3.4, the controllers proceed to Mode A and the system returns to its final state.

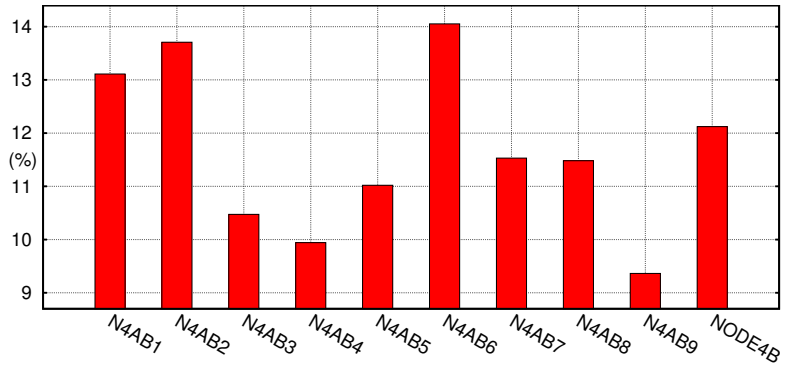


Figure 3.14: Single feeder energy lost by every PV ($T_{upd} = 22 \sim 30$ s)

Figure 3.14 shows the percentage of electrical energy lost by each PV unit due to curtailment, computed as:

$$Energy\ Lost_{PV_j} = \frac{\int_0^{3200} P_{PV_j}^{MPP} dt - \int_0^{3200} P_{PV_j}^{out} dt}{\int_0^{3200} P_{PV_j}^{MPP} dt} \tag{3.12}$$

It can be seen that all PV units have similar values, with the variations between them caused by the asynchronous and discrete nature of the controllers. To verify this behavior, the simulation is repeated with $T_{upd} = 1$ s and $T_{upd} = 0$ s respectively.

In the first case, it is assumed that the measurement, control and communication delays are 1 s and the energy lost is seen in Figure 3.15. It can be seen that the variation between the different PV units has been significantly decreased. In the second case, it is assumed that the controllers are infinitely fast without any measurement or communication delays. The energy lost in this case is the same for all PV units and equal to 12.2%.

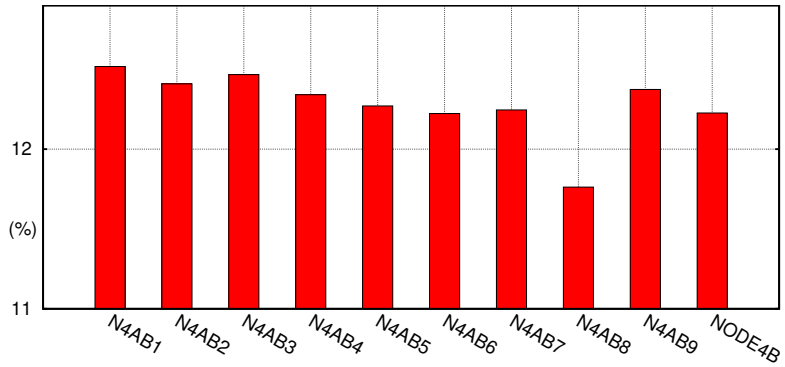


Figure 3.15: Single feeder energy lost by every PV ($T_{upd} = 1$ s)

The performance of the proposed control is compared to that of an OPF-based centralized scheme, as explained in Section 4.2. A step change of P_{MPP} is applied causing an overvoltage problem. Figure 3.16 displays the active power output by the PV units for both the distributed and the centralized scheme at their final equilibrium point.

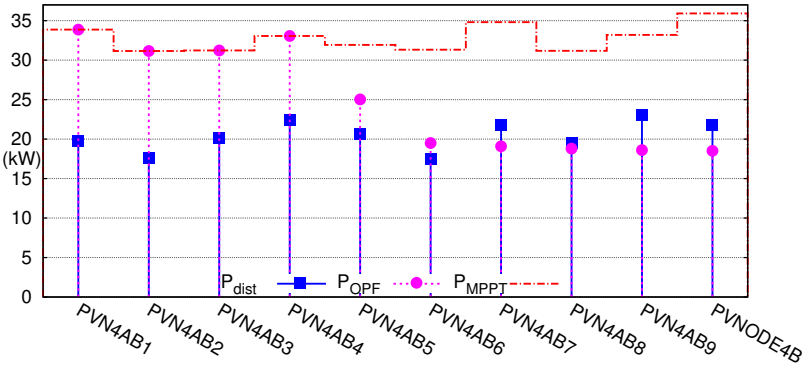


Figure 3.16: Active power production by each PV unit, with squares for the distributed and circles for the centralized scheme, at the end of Mode C. The maximum power that could be produced by each PV unit is shown by the upper dash-dotted line. The PV units closest to the distribution transformer are on the left

It is observed that the centralized scheme prioritizes curtailing active power of PV units further away from the distribution transformer, where the overvoltage problem is more prominent and the curtailment of active power is

more effective. On the other hand, the distributed scheme leads to a more uniformly shared active power curtailment among the PV units.

Table 3.1 shows the total active power production and curtailment, the total reactive power absorption and the network losses in the feeder at the equilibrium points. As expected, the distributed scheme curtails more active power than the centralized, 37.6% versus 24.0%. This is the cost associated with uniformly sharing the active power curtailment throughout the feeder and not prioritizing the curtailment on the problematic buses. Moreover, based on the capability curve (see Figure 3.2), a higher active power output allows for more reactive power to be absorbed. Thus, the OPF-based approach allows for less active power curtailment and at the same time higher reactive power absorption. Nevertheless, the network losses are higher for this approach because of the larger amount of active and reactive power transiting through the network.

This better performance of the OPF-based scheme is anticipated as the latter can perform a centralized optimization, knowing both the system model and the remote measurements from all the nodes and inverters. As mentioned in the Introduction, the main advantage of the proposed scheme is that it provides a slightly suboptimal solution with far less information and communication requirements.

	Single feeder		14-feeder	
	Dist.	Centr.	Dist.	Centr.
Active power generated (kW)	204	249	1460	1475
Active power curtailed (kW)	123	78	18.4	0
Reactive power absorbed (kVar)	67.1	81.8	367	365
Losses in the network (kW)	20.5	27.3	123	125

Table 3.1: Comparison between active and reactive power generated and curtailed by PV units and network losses with the distributed and centralized scheme for the two test-systems

5.2 Selection of the capability curve

As discussed in Section 2 the capability curve of Figure 3.2 is used in the PV model to calculate the available reactive power. This is implemented through Eq. 3.1. However, it is technically possible for a PV inverter to operate as a STATCOM device. In this case, the capability curve is defined by the entire semi-circular area of Figure 3.2 and can be implemented by redefining Eq. 3.1 as:

$$I_{qmax} = \sqrt{I_{max}^2 - I_{pcmd}^2} \tag{3.13}$$

Comparing the two capability curves, it can be clearly seen that PV units operating as STATCOM devices offer higher flexibility in adjusting reactive power, especially when the amount of active power produced is small. Figure 3.17 shows the state transition of the single feeder example using the STATCOM capability curve. In this case, the controllers never reach Mode C and do not curtail any active power; the reactive power adjustments are enough to secure the system.

The proposed control scheme does not make any assumptions on the form of the capability curve; whether this is a more restricted curve like Figure 3.2 due to the grid code requirements, a more permissive STATCOM-type or

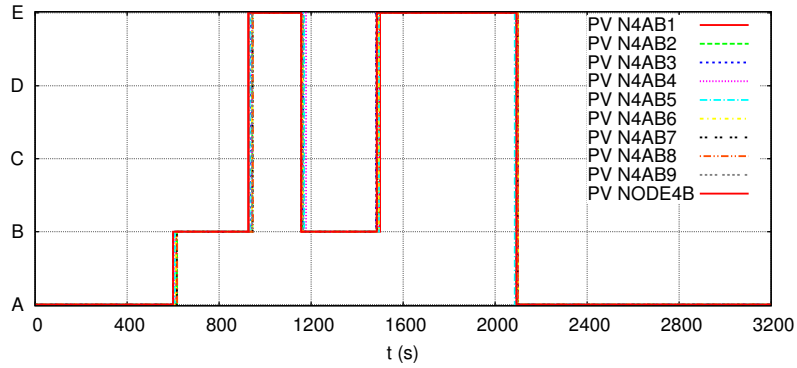


Figure 3.17: Single feeder state transitions (STATCOM operation)

even a combination of the two in the same LV feeder. The scheme uses first the available reactive power control and then proceeds to active power curtailment, if needed.

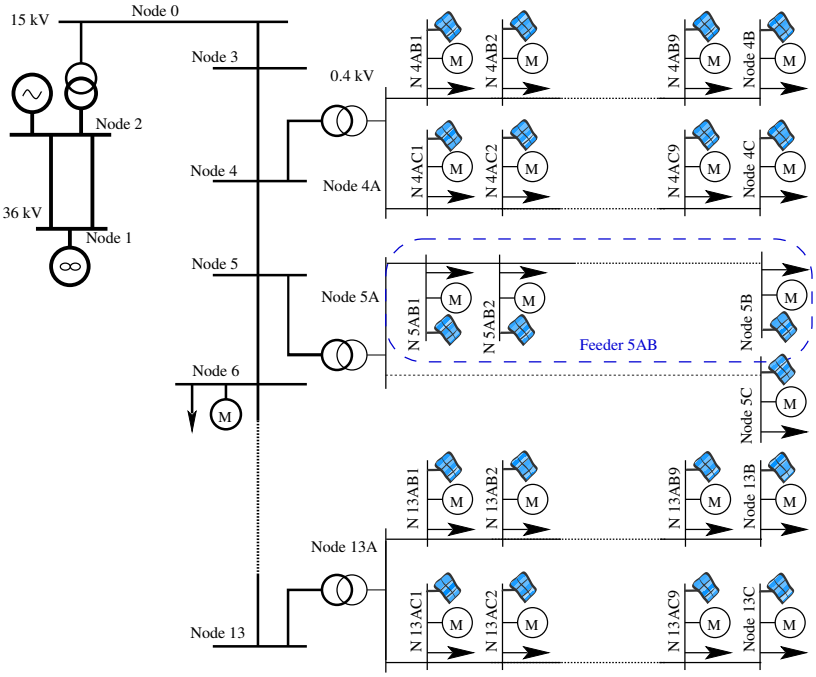


Figure 3.18: 14-feeder test-system model

5.3 *14-feeder test-system*

In this subsection, the MV/LV distribution system presented in [37] is used, modified for the purpose of demonstrating the performance of the method in a demanding situation. Its one-line diagram is shown partially in Figure 3.18. The system includes 14 LV feeders similar to the one considered in the previous subsection. The feeders are connected in pairs to the MV buses 4, 5, 7, 8, 10, 11 and 13. Larger motors are connected to the MV buses 6, 9 and 12. Furthermore, a synchronous machine with detailed model is connected to Node 2.

The same variation of P_{MPP}, as in the previous system, is considered for each feeder. The nominal power of the PV units in this case was randomly

[37] Otomega and Van Cutsem ("Distributed load interruption and shedding against voltage delayed recovery or instability")

chosen between 9.5 and 12.5 kW. The control parameters were chosen at $V_4 = 1.10$ pu, $V_3 = 1.05$ pu, $PF_{min} = 0.95$, $T_{DQ} = 300$ s, $T_{DP} = 600$ s, $T_{RP} = 600$ s, $T_{RQ} = 300$ s and $T_{reset} = 100$ s.

From Figure 3.19, it can be observed that the distributed scheme successfully manages to keep voltages below 1.1 p.u. after a temporary excess that starts at $t = 200$ s. The controller state transitions are presented in Figure 3.20. It can be seen that controllers switch to Mode B when an overvoltage takes place at $t \approx 200$ s. As reactive power support is not sufficient to clear the overvoltage situation, they proceed to Mode C and start curtailing active power. Controllers then enter the restorative phase of the algorithm until they reach Mode E where an overvoltage occurs. Next, they switch back and forth between Mode B and E. At $t \approx 2500$ s, the controllers manage to reach the normal mode of operation (Mode A) without creating an overvoltage situation.

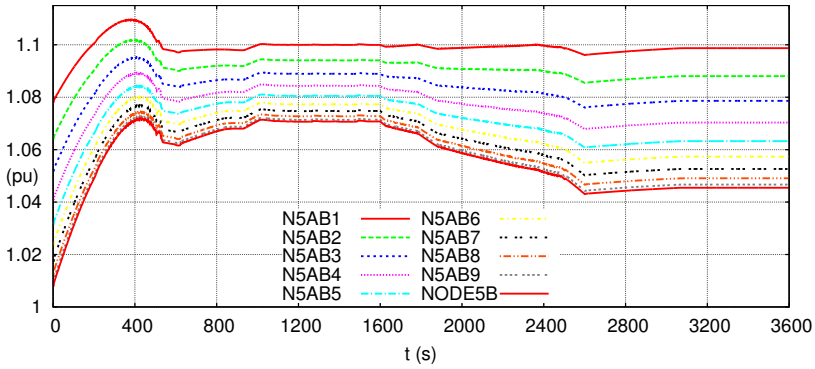


Figure 3.19: 14-feeder test-system: voltage evolution for feeder 5AB

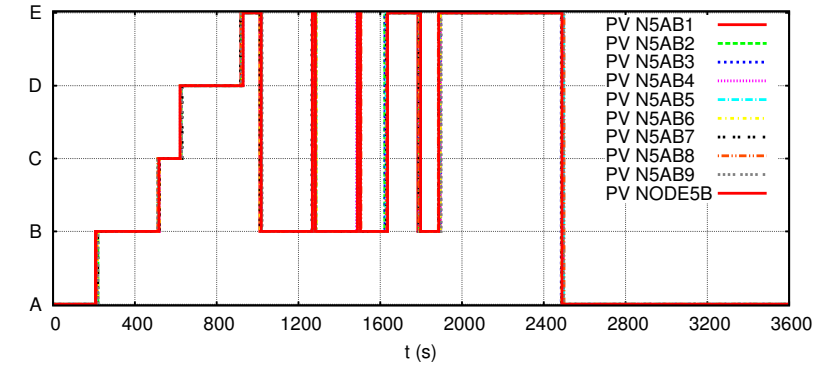


Figure 3.20: 14-feeder test-system: state transitions for feeder 5AB

Table 3.2 shows the total active power curtailment (as a percentage of P_{MPP}) and the reactive power absorbed (as a percentage of Q_{max}) for each feeder at the end of Mode C ($t \approx 620$ s). It can be seen that LV feeders located closer to the HV/MV transformer have to curtail some active power while the more remote ones rely only on reactive power adjustments. This is expected as the voltage at the MV buses further from the HV/MV transformer are lower and, hence, the voltage rise along the LV feeder does not result in excessive voltages. Moreover, there is no communication between the feeders, thus the control actions rely only on local measurements within each feeder.

Feeder	P_{curt}	Q_{abs}	Feeder	P_{curt}	Q_{abs}
4AB	1.89%	100.0%	8AC	0.00%	65.7%
4AC	2.76%	100.0%	10AB	0.00%	67.2%
5AB	5.61%	100.0%	10AC	0.00%	64.4%
5AC	6.78%	100.0%	11AB	0.00%	59.2%
7AB	0.37%	100.0%	11AC	0.00%	60.5%
7AC	0.00%	100.0%	13AB	0.00%	75.5%
8AB	0.00%	65.2%	13AC	0.00%	78.1%

Table 3.2: Percentage of curtailed active power (P_{curt}) and absorption of reactive power (Q_{abs}) in percentage of the total power per feeder at the end of Mode C ($t \approx$ 620 s).

Finally, Table 3.1 shows the total active power production and curtailment, the total reactive power absorption and the network losses in the feeder at the equilibrium points after a step change of P_{MPP}. It can be seen that the proposed scheme performs almost as well as the OPF-based one and that the system losses are almost equal in both schemes. Nevertheless, the former resorts to some active power curtailment whereas the latter utilizes only reactive power adjustments to resolve the overvoltage problem. This is to be expected as the OPF-based controller optimizes the entire MV/LV network and can provide support between separate feeders. On the contrary, the distributed scheme coordinates only PV units within the same feeder and does not know about the other LV feeders.

6 Conclusion

As an alternative to network reinforcement, a distributed control scheme that alleviates the voltage problems caused by a large penetration of PV units in LV DNs has been proposed. This scheme makes use of little communication between the controllers to activate the inverters' reactive support and minimize active power curtailment. Through dynamic simulations, the behavior of the scheme has been analyzed and it was shown that its performance is comparable to a centralized, OPF-based, control scheme which would be more expensive to deploy and complex to operate (for instance, when the DN topology changes).

Several research directions exist for this distributed control scheme. First, it would be interesting to extend it to an unbalanced three-phase network. Unbalances can be generally neglected in MV networks in Europe but unbalances are often present in LV DNs since house appliances and PV units are usually single phase. This implies taking into account the coupling between the phases.

Second, rather than relying on active power curtailment, the use of local storage systems or flexible loads could be considered. This would change fundamentally the control problem considered and would raise new research questions. For example, the existence of local energy storage would couple the control actions at different times, since accumulating active power into storage at a specific time, influences the amount of energy available in the future.

Finally, a cost analysis could be performed to evaluate the economical profitability of deploying such a distributed scheme compared to a reinforcement of the network.

Chapter 4
Distributed control of photovoltaic units in unbalanced LV distribution networks to prevent overvoltages

As more and more photovoltaic units are being installed, some LV networks have already attained their maximum hosting capacity, i.e. the maximum amount of distributed energy resources that they can accommodate during regular operations without suffering problems, such as overvoltages. As an alternative to network reinforcement, active network management (ANM) can, to a certain extent, increase their hosting capacity by controlling the power flows. In the framework of ANM, we presented in Chapter 3 a distributed control scheme, which makes use of a distress signal sent by each participating unit, when its terminal voltage is higher than 1.1 p.u. All units then proceed to absorb the maximum reactive power available. If the problem is not resolved, the units proceed to active power curtailment. This paper extends this control scheme to the case of unbalanced three-phase four-wire distribution networks with single- and/or three-phase inverters. The control scheme works by first partitioning the inverters into four groups, three for the single-phase inverters (one for each phase), and one for the three-phase converters. Each group then independently applies a distributed algorithm similar to the one presented in Chapter 3. Their performance are then compared to those of two reference schemes, an on-off algorithm that models the default behaviour of PV inverters when there is an overvoltage, and the other one based on an unbalanced OPF. Its resulting total curtailed energy always lies between the two, with the on-off algorithm presenting the poorest performance, and the proposed algorithm losing its edge when the network is strongly unbalanced.

1 Introduction

More and more customers are reporting to Distribution System Operators (DSOs) that their PhotoVoltaic (PV) units are disconnecting from the network. This phenomenon has several causes, among which are occurrences of overvoltages in the low-voltage (LV) distribution feeder they are connected to. Indeed, as more PVs are being installed, some LV networks have already attained their maximum hosting capacity, i.e. the maximum amount of distributed energy resources (DER) that they can accommodate during regular

operations without causing problems, such as violations of operational constraints (overvoltages and congestions). Norms impose the disconnection of a PV when the ten-minute average voltage at its connection point is higher than 110% of the nominal voltage, or when the instantaneous voltage surpasses 115% of the nominal voltage (cf. EN 50160). Theses disconnections cause a loss in renewable energy production and, henceforth, a loss of earnings for the PV owners.

To solve this problem, there are two alternative possibilities: on the one hand, reinforce the network or, on the other hand, control the power flows inside it by performing active network management (ANM), the focus of this paper. In the framework of ANM, several schemes to control the power generated by DER have been proposed in the literature, whether centralised or distributed. Since Chapter 3 provides a detailed literature review of the algorithms to control the PV units' output, we limit ourselves to references presenting control algorithm specifically designed for three-phase unbalanced networks. Su[1] presented a centralized controller based on a multi-objective optimal power flow (OPF) problem that can simultaneously improve voltage magnitude and balance profiles, while minimizing network losses and generation costs. Dall'Anese[2] considers the economic dispatch distributed generation units with a semidefinite relaxation of the OPF. Chua[3] proposed an energy storage system for mitigating voltage unbalance as well as improving the efficiency of the network, using a local controller targeting specified voltage values. In particular, we presented a distributed control scheme[4] which makes use of a distress signal sent by each participating unit when its terminal voltage is higher than 1.1 p.u. When this happens, all units proceed to absorb the maximum reactive power available. If the problem is not resolved, the units proceed to active power curtailment. This ensures a fair and coordinated use of reactive and active power.

The goal of the control scheme is to increase the hosting capacity of the network by controlling the power flows inside it. Incidentally, it maximises the active power production, or minimize the active power curtailed, which means minimising the occurrence of overvoltages.

This paper extends the algorithm previously developed to the case of unbalanced three-phase four-wire distribution networks with single- and/or three-phase inverters. Its key principles are inherited from the original algorithm and are reiterated as follows: (i) it first makes use of reactive power, (ii) it uses active power curtailment as a last resort, (iii) it only needs communication in the form of a distress signal sent throughout the feeder to pool available resources, and (iv) it does not require a detailed model of the network.

The proposed algorithm is explained, illustrated and compared to two other algorithms: an on-off algorithm – that models the default behaviour of PV inverters when there is an overvoltage – and an unbalanced optimal power flow. Their performance is compared in different PV connection scenarios, using three metrics: curtailed energy, reactive power usage and ohmic losses in the network.

The paper is organized as follows: Section 2 presents the modelling of the distribution system and its component throughout the entire paper. Section 3 presents the extension of the previous distributed algorithm,[5] and Section 4

[1] Su, M. A. Masoum, and P. Wolfs ("Comprehensive optimal photovoltaic inverter control strategy in unbalanced three-phase four-wire low voltage distribution networks") and Su, M. A. S. Masoum, and Peter J. Wolfs ("Optimal PV inverter reactive power control and real power curtailment to improve performance of unbalanced four-wire LV distribution networks")

[2] Dall'Anese, Giannakis, and Wollenberg ("Optimization of unbalanced power distribution networks via semidefinite relaxation")

[3] Chua et al. ("Energy storage system for mitigating voltage unbalance on low-voltage networks with photovoltaic systems")

[4] Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

[5] Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

details the different algorithms to which it will be compared. Sections 5 and 6 explain the scenarios for the numerical simulations and include an analysis of the results. Finally, Section 7 concludes.

2 Modelling

To simulate the behaviour of the distribution system, the time will be discretised into multiple time steps for which successive steady-state power flows will be solved. This is motivated by the low inertia of the few rotating machines in LV networks, as well as the fast response of PV inverters. The following section defines the different modelling choices, as well as the different notations used in this paper. Throughout this section, the subscript t will refer to the current time step of simulation.

2.1 Buses

A node (a.k.a. bus) represents a connection point in the distribution system, e.g. a junction or a supply point. The set of node indices is $\mathcal{N}$. Throughout the network, there are four conductors, one for each phase identified, independently of the nodes, by their index set $\mathcal{P} = \{c_a, c_b, c_c\}$. The neutral conductor is referred to by c_n. Let $\mathcal{C} = \mathcal{P} \cup \{c_n\}$. Furthermore, the connection point of two or more conductors at a node is called a terminal. Double indexing is used to identify terminals, the first index referring to the node and the second one to the conductor. Using that convention, each terminal has a complex voltage $V_t^{(n,c)}\ \forall n \in \mathcal{N}, c \in \mathcal{C}$. The reference for this voltage is the common ground of the electrical system.

For proper operations of the network, the magnitude of the phase-to-neutral voltage should always lie between a lower bound $V_{\min}$ and an upper bound $V_{\max}$.

$$V_{\min} \leq \left| V_t^{(n,p)} - V_t^{(n,c_n)} \right| \leq V_{\max},\ \forall n \in \mathcal{N}, p \in \mathcal{P}. \tag{4.1}$$

2.2 Electrical lines

Electrical lines are called branches whether overhead lines or underground cables, and are identified by their index set $\mathcal{B}$. They are defined by a square impedance matrix $Z^{(b)}\ \forall b \in \mathcal{B}$, whose size is equal to the number of conductors, and which represents the resistance of the lines and the self- and mutual reactance of the conductors.[6] The shunt capacitances are neglected, a choice commonly made[7] given the small length of LV lines and their small charging currents. The topology of the network is defined by associating to each branch a sending node n_s and a receiving node n_r. The currents flowing in each conductor, $I_t^{(b,c)}\ \forall b \in \mathcal{B}, c \in \mathcal{C}$, are considered to be positive if they flow from the sending node to the receiving node.

The voltage drop along the line is calculated with (4.2).

$$V_t^{(n_s,c)} - V_t^{(n_r,c)} = \sum_{c' \in \mathcal{C}} Z^{(b)}(c, c') I_t^{(b,c')}\ \forall b \in \mathcal{B}, c \in \mathcal{C} \tag{4.2}$$

[6] William H Kersting (*Distribution System Modeling and Analysis*) and Urquhart ("Accuracy of Low Voltage Distribution Network Modelling")

[7] D. Das, Nagi, and Kothari ("Novel method for solving radial distribution networks"), Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach"), and Sunderland and Conlon ("4-Wire load flow analysis of a representative urban network incoprating SSEG")

Finally, when currents flow in a branch b, they generate losses $L_t^{(b)}$ according to Ohm's law:

$$L_t^{(b)} = \sum_{c \in \mathcal{C}} \sum_{c' \in \mathcal{C}} Z^{(b)}(c, c') \left(I_t^{(b,c')} \right)^2 \ \forall b \in \mathcal{B} \tag{4.3}$$

2.3 Loads and photovoltaic units

The photovoltaic units (resp. the loads), whose index set is $\mathcal{G}$ (resp. $\mathcal{L}$), are defined by their injection of active $P_t^{(i,p)}$ and reactive power $Q_t^{(i,p)}$ $\forall p \in \mathcal{P}, i \in \mathcal{G}$ (resp. $\forall i \in \mathcal{L}$), and the node to which they are connected, identified thanks to the function $N(\cdot)$ that links the index of a PV or load to the index of the node to which it is connected. $P_t^{(i,p)}$ and $Q_t^{(i,p)}$ are considered positive when the power is injected into the network.

$P_t^{(i)}$ and $Q_t^{(i)}$ refer to the total active and reactive power injected by PV or load i. In the case of a three-phase PV unit i, $P_t^{(i,p)} = P_t^{(i)}/3$ and $Q_t^{(i,p)} = Q_t^{(i)}/3$, $\forall p \in \mathcal{P}$. In the case of a single-phase PV unit i connected to phase p', $P_t^{(i,p')} = P_t^{(i)}$ and $Q_t^{(i,p')} = Q_t^{(i)}$, and $P_t^{(i,p)} = 0$ and $Q_t^{(i,p)} = 0$, $\forall p \in \mathcal{P} \setminus p'$.

The currents $I_t^{(i,p)}$ injected by the PV or load i in the phase conductor p of node $N(i)$ are calculated according to (4.4).

$$I_t^{(i,p)} = \left(\frac{P_t^{(i,p)} + jQ_t^{(i,p)}}{V_t^{(N(i),p)} - V_t^{(N(i),c_n)}} \right)^* \ \forall i \in \mathcal{L} \cup \mathcal{G}, p \in \mathcal{P} \tag{4.4}$$

where $(\cdot)^*$ denotes the complex conjugate.

The current $I_t^{(i,c_n)}$ injected by the PV or load i in the neutral conductor c_n is equal to the sum of the phase currents, as in Equation (4.5).

$$I_t^{(i,c_n)} = - \sum_{p \in \mathcal{P}} I_t^{(i,p)} \ \forall i \in \mathcal{G} \cup \mathcal{L} \tag{4.5}$$

The next three subsections describe the constraints limiting the production of PV units.

Maximum power point

Maximum power point tracking (MPPT) algorithms are implemented to maximise the output of the PV unit based on their I-V curve. In practice, this maximum production is not a priori known, as it depends on solar irradiance and the temperature of panels. For the purpose of this study, it is assumed that this point is known and equal to $P_{MPP,t}^{(i)}$ $\forall i \in \mathcal{G}$, and that it limits the active power production according to (4.6).

$$P_t^{(i)} \leq P_{MPP,t}^{(i)} \ \forall i \in \mathcal{G} \tag{4.6}$$

Capability curve

The active power and reactive power production of a PV unit are limited by the capability curve of their inverter, i.e. they are limited by the maximum apparent power $S_{\max}^{(i)}$, $\forall i \in \mathcal{G}$ (4.7) and the minimum power factor $PF_{\min} = \cos \varphi_{\min}$ (4.8).

$$|P_t^{(i)} + jQ_t^{(i)}| \leq S_{\max}^{(i)} \ \forall i \in \mathcal{G} \tag{4.7}$$

$$-\tan\varphi_{\min} P_t^{(i)} \leq Q_t^{(i)} \leq \tan\varphi_{\min} P_t^{(i)} \ \forall i \in \mathcal{G} \tag{4.8}$$

As in Chapter 3, the one used in this study is coming from the German and Italian standards VDE-AR-N 405 and CEI 0-2. Those standards impose that the inverters are sized so that they can produce and absorb reactive power at the minimum power factor when the active power production is maximal. Henceforth, condition (4.7) will always be considered satisfied. For convenience, the function $Q_{\max}^{(i)}(P_t^{(i)})$ gives the maximum amount of reactive power that PV unit i can produce or absorb, given its current active power production $P_t^{(i)}$.

Set points

$P_{set,t}^{(i)}$ (resp. $Q_{set,t}^{(i)}$) is the set point for active (resp. reactive) power of PV unit i. They are values that the control schemes can modify and optimise. If $P_{set,t}^{(i)}$ and $Q_{set,t}^{(i)}$ satisfy conditions (4.6) and (4.8), $P_t^{(i)} = P_{set,t}^{(i)}$ and $Q_t^{(i)} = Q_{set,t}^{(i)}$. If they do not, then they take the maximum values satisfying (4.6) and (4.8).

2.4 Slack bus

The equivalent of the external network is modelled as a slack bus whose voltages vary during the day according to on-site phase-to-neutral voltage measurements V_{SL}. The phase-to-neutral voltage of the slack bus (n_{SL}) is given by (4.9).

$$V_t^{(n_{SL},p)} - V_t^{(n_{SL},c_n)} = V_{SL,t}^{(p)}, \ \forall p \in \mathcal{P} \tag{4.9}$$

2.5 Kirchhoff's Current Law

Letting $\mathcal{B}_{in}^{(n)}$ (resp. $\mathcal{B}_{out}^{(n)}$) be the subset of lines whose receiving end (resp. sending end) corresponds to bus n, and $\mathcal{G}^{(n)}$ (resp. $\mathcal{L}^{(n)}$) be the subset of PV units (resp. loads) connected to node n, equation (4.10) implements Kirchhoff's current law.

$\forall n \in \mathcal{N}$,

$$\begin{gathered} I_t^{(b_{in},c)} + I_t^{(g,c)} + I_t^{(l,c)} = I_t^{(b_{out},c)} \\ \forall b_{in} \in \mathcal{B}_{in}^{(n)}, b_{out} \in \mathcal{B}_{out}^{(n)}, g \in \mathcal{G}^{(n)}, l \in \mathcal{L}^{(n)}, c \in \mathcal{C} \end{gathered} \tag{4.10}$$

3 Distributed control scheme

To create the distributed control scheme, the inverters are first partitioned into four groups. The first three gather the single-phase inverters according to the phase to which they are connected, and the last group is composed of the three-phase inverters. Each group acts independently of the other ones, according to a distributed control scheme which is the direct extension of the one proposed in Chapter 3. Its seven modes of operation as well as its state transition diagram are described in Section 3.2 and 3.3.

Thanks to a distress signal (further explained in Section 3.1), each phase group will react only if it records a voltage problem on its phase, and the three-phase group will react to all voltage violations. Doing so ensures that

the actions of single-phase inverters do not create or worsen overvoltage situations, or even generate undervoltages. This also prevents the curtailment of active power in a phase where there is no overvoltage.

3.1 Distress signal

As a logical evolution from the control scheme of Chapter 3, where only one distress signal is sent when an overvoltage occurs, there are now three distress signals, one for each phase. They are repeatedly sent by the inverters recording an overvoltage in the phase to which they are connected. Three-phase inverters can thus send multiple distress signals if more than one phase suffer from an overvoltage.

3.2 Modes of operation

The controller is composed of seven modes of operations: one normal, two depleting, two waiting and two restoring ones. A detailed explanation of the modes is available in Chapter 3, and a short summary can be found in the next paragraphs. All modes are implemented as timers, i.e. once the inverter enters these modes, it will definitely exit them once their timer has elapsed. The normal mode of operation $Q_{loc}^{\downarrow\uparrow}(V)$ and the active power curtailment mode $P^{\downarrow}$ are two exceptions. Finally, all modes (except $Q_{loc}^{\downarrow\uparrow}(V)$) have a target for active and reactive power to reach at the end of the timer.

- Mode $Q_{loc}^{\downarrow\uparrow}(V)$ – normal mode of operation: The inverters maximise the active power with a Maximum Power Point Tracking (MPPT) algorithm. They absorb reactive power as a function of the local voltage (resp. maximal local voltage) for single-phase inverters (resp. for three-phase inverters). The function is illustrated in Figure 4.1.

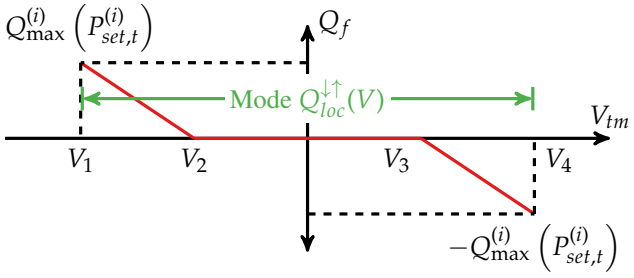


Figure 4.1: $Q_f(V_{tm}, P_{set,t})$ function of PV unit i for Mode $Q_{loc}^{\downarrow\uparrow}(V)$, where V_{tm} is the voltage (resp. maximum voltage of the three-phase) at the connection point of the single-phase (resp. three-phase) PV unit. For $V_{tm} \geq V_4$ an emergency signal is issued, and the controller moves to Mode $Q^{\downarrow}$.

- Mode $Q^{\downarrow}$ – depleting reactive power: the active power is at its maximum and the inverters gradually absorb reactive power until they reach the maximum at the end of a timer of t_{DQ} minutes. Once the time is up, the inverter moves to the next mode.
- Mode $P^{\downarrow}$ – curtailing active power: the reactive power absorption is maximal, and the active power production is gradually lowered to reach zero at the end of a timer of t_{DP} minutes. When the timer ends, the inverter remains in this mode until there is no further distress signal.
- Mode $Q^{\rightarrow}$ – waiting to start increasing reactive power: once there is no further distress signal, the inverter enters a waiting mode of t_{reset} minutes

to avoid rapid oscillations between depleting and restoring active and reactive power. In this mode, the active and reactive power set points remain constant.

- Mode $P^{\rightarrow}$ – waiting to start increasing active power: the active power set point remains constant for t_{reset} minutes and the reactive power absorption is maximal.
- Mode $P^{\uparrow}$ – restoring active power: the active power is gradually restored to its maximum. Since $P_{MPP,t}$ is not known a priori, the target for the active power is 110% of the installed capacity at the end of a timer of t_{RP} minutes. The reactive power absorption is maximal.
- Mode $Q^{\uparrow}$ – restoring reactive power: the inverters produce active power at its maximum. They gradually absorb less reactive power until they attain the value dictated by the function $Q_f(V)$ of Mode $Q_{loc}^{\downarrow\uparrow}(V)$ after t_{RQ} minutes.

3.3 State diagram

Figure 4.2 represents the state transition diagram of the controller. The red dotted lines are the emergency control transitions, i.e. when the PV unit receives a distress signal. The green dash-dotted lines represent the transition when there are no more distress signals. Finally, the blue dashed lines are the transition associated with the end of a timer. t_{DQ} (resp. t_{DP}) is the time needed in Mode $Q^{\downarrow}$ (resp. Mode $P^{\downarrow}$) to use all available reactive (resp. active) controls. t_{reset} is the elapsed time in Modes $Q^{\rightarrow}$ and $P^{\rightarrow}$ without an emergency signal for the controller to start restoring active/reactive power. t_{RP} (resp. t_{RQ}) is the time needed in Mode $P^{\uparrow}$ (resp. Mode $Q^{\uparrow}$) to restore active (resp. reactive) power to the set-point values of Mode $Q_{loc}^{\downarrow\uparrow}(V)$.

The four groups of inverters can be found in different locations of the state diagram, but all the inverters of the same group will be in the same mode. Indeed, they move from one mode to the other, thanks either to the timers or the presence or absence of a distress signal. Since three-phase PV units receive the distress signals from the three phases, they are more likely to stay in a depleting mode than a single-phase one.

3.4 Fairness

When curtailment of active power is inevitable, it is desirable for the control scheme to evenly distribute the burden among the PV units, a property known as 'fairness'.

However, the default behaviour of PV units that disconnect when there is an overvoltage is not fair, because this behaviour will likely curtail the PV units located further away from the beginning of the feeder, where the voltage sensitivity is higher, i.e. the ratio of voltage variations to the active and reactive power productions is larger.

By contrast, the initial algorithm[8] was designed with fairness in mind as the curtailment was proportional to the installed capacity. The adaptation to three-phases is no longer fair in the same sense as it is now only fair group-wise.

[8] Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

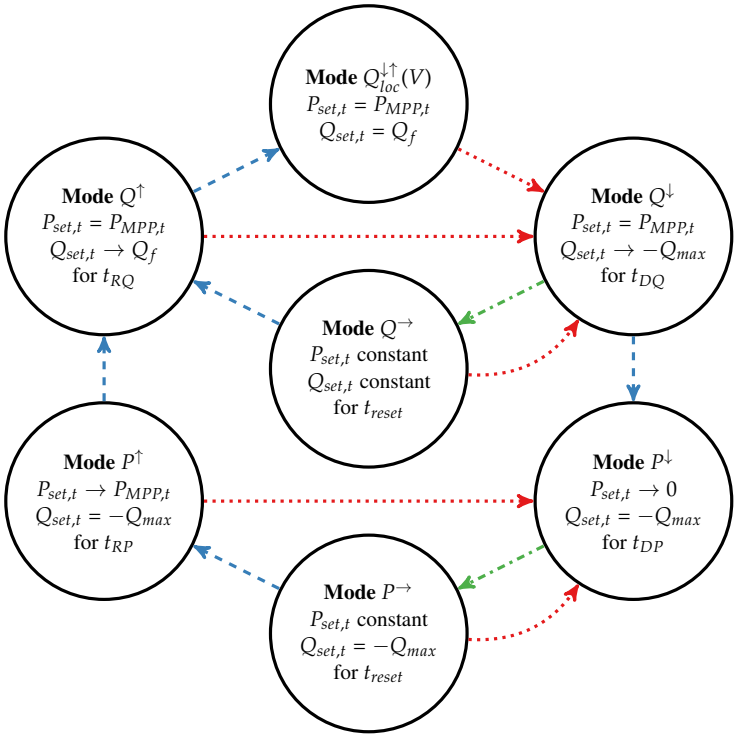


Figure 4.2: State transition diagram of the distributed control scheme. The red dotted lines are the emergency control transitions, i.e. when the PV unit receives a distress signal. The green dash-dotted lines represent the transition when there are no more distress signals, and the blue dashed lines are the transition associated with the end of a timer. $P_{set,t}$ and $Q_{set,t}$ are the active and reactive power set-points of the controller. $P_{MPP,t}$ is the maximum available active power of the PV module and depends on the level of solar irradiation. Q_{max} is the maximum available reactive power; it varies according to the capability curve as a function of the active power output.

3.5 Practical implementation

The controller can be implemented as a microcomputer. It can communicate with the inverter it controls thanks to RS-485 and the use of a proprietary communication protocol. It reads the voltage measurements in the appropriate registers of the micro-controller of the inverter, compute the active and reactive power set points according to the algorithm, and then write them in the appropriate registers.

Regarding the distress signal, if it is sent using Power Line Communication (PLC), only the inverter connected to the same phase will receive it, thus ensuring the right grouping of the inverters. However, if the distress signal is sent using another technology (e.g. GPRS), the phase of each inverter must first be identified, for example using voltage correlation.[9]

[9] Olivier, Sutera, et al. ("Phase Identification of Smart Meters by Clustering Voltage Measurements")

4 Comparison control schemes

The performance of the distributed algorithm is compared to those of two other algorithms. The first one is fully distributed and does not require a model of the network. It is the on-off algorithm already implemented in PV units. The second one is a more elaborate, centralized control scheme based on an optimal power flow and a detailed model of the network.

4.1 On-off control algorithm

According to EN 50160, the inverters must disconnect if the ten-minute moving average of the voltage at their connection point is higher than 1.1 p.u. Moreover, they must disconnect immediately if the instantaneous voltage is higher than 1.15 p.u.

Controller modes

The on-off control algorithm has two modes: an MPPT mode where the active power production is maximised, and an off mode, where the production is shut down. If there are no overvoltages, the PV units are in MPPT mode. If the voltage at their connection point is higher than 1.1 p.u., they switch to the off mode. Let $\mathcal{G}^{(V_{\max})}$ be the subset of PV units observing an overvoltage for the time step $t-1$.

$$P_{set,t}^{(i)} = 0, \forall i \in \mathcal{G}^{(V_{\max})} \tag{4.11}$$

The production of active power stays at zero for one minute and the PV units then switch back to the MPPT mode. If an overvoltage occurs anew, they switch back to the off mode.

With this algorithm, the production or absorption of reactive power is always equal to zero.

$$Q_{set,t}^{(i)} = 0, \ \forall i \in \mathcal{G} \tag{4.12}$$

Practical implementation

This algorithm is the one currently implemented in commercial inverters. It only relies on local measurements and controls. The inverters constantly monitor the voltage at their terminal and if an overvoltage occurs, they shut down the production.

4.2 Unbalanced Optimal Power Flow Control Algorithm

A centralised solution to optimize the active and reactive power production of the inverters can take the form of a three-phase Optimal Power Flow (OPF) (See [10] for more information on unbalanced OPF).

[10] Bruno et al. ("Unbalanced three-phase optimal power flow for smart grids")

Optimisation variables

Strictly speaking, optimisation variables are the active and reactive power produced by the PV units: $P_t^{(g)}$ and $Q_t^{(g)}, \ \forall g \in \mathcal{G}$. However, the voltages, currents and line losses are included in the optimisation variables, since they are not explicitly linked to the injected powers. This simplifies the definition of the constraints, e.g. the voltage constraints. The optimisation variables are thus:

$$P_t^{(g)}, Q_t^{(g)}, \quad \forall g \in \mathcal{G} \tag{4.13}$$
$$I_t^{(i,c)}, \quad \forall i \in \mathcal{G} \cup \mathcal{B} \cup \mathcal{L}, c \in \mathcal{C} \tag{4.14}$$
$$L_t^{(b,c)}, \quad \forall b \in \mathcal{B}, c \in \mathcal{C} \tag{4.15}$$
$$V_t^{(n,c)}, \quad \forall n \in \mathcal{N}, c \in \mathcal{C} \tag{4.16}$$

Constraints

All the equations and inequalities presented in Section 2 are included as constraints of the OPF.

Optimisation criterion

The optimisation criterion can be defined as:

$$\max \sum_{g \in \mathcal{G}} P_t^{(g)} - \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} L_t^{(b,c)} \tag{4.17}$$

With that formulation, the flow of active power exported by the LV network through the LV/MV transformer is maximised. On a physical point of view, applying the solution of this optimisation problem to the inverters ensures two desirable behaviours: (i) the increase in PV production will not be at the expense of an increase in network losses to a point where the increment in the losses is larger than the increment in the PV power; (ii) if the maximum limit of PV production is reached and some operational margins in reactive power remain, the optimisation will minimise the losses.

Practical implementation

A centralised control scheme comprises three different parts. The first one is composed of the infrastructure necessary to evaluate the state of the system it controls. The second part is the controller itself. It computes the control actions from the (previous) information. The third and last part is the infrastructure used for sending and applying its control actions. A centralised scheme based on an OPF would require a detailed model of the network that must be kept up-to-date, and an extended communication infrastructure.

5 Numerical simulations

This section presents the methods used to solve the power flow and the optimal power flow. It also describes the input of the simulations, i.e. the test network, the loads, the different scenarios, as well as the various numerical values chosen for the parameters.

5.1 Power flow

In the case of LV networks, two methods are commonly used to solve the power flow equations: (i) the Newton-Raphson method applied to the power flow equations expressed in terms of current injections [11], and (ii) the Backward Forward Sweep algorithm [12]. Given its simple implementation and the radial nature of LV networks, we have chosen the latter and implemented it in Python.

[11] Garcia et al. ("Three-phase power flow calculations using the current injection method"), M. J. Alam, Muttaqi, and Sutanto ("A three-phase power flow approach for integrated 3-wire MV and 4-wire multi-grounded LV networks with rooftop solar PV"), and Penido et al. ("Four wire newton-rapshon power flow based on the current injection method")

[12] C. Cheng and Shirmohammadi ("A three-phase power flow method for real-time distribution system analysis") and Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

5.2 Time sequence of actions

The simulations have a time step of one minute. They are first initialized with a power flow. Then, for each time step, the sequence of actions is as follows: (i) Based on the voltages from the previous time step, all controllers compute and apply their control actions, i.e. they specify the active and reactive power set points of the PV unit they control. (ii) A power flow is solved taking into consideration the current energy consumption and production. The process is repeated until the end of the simulations.

5.3 Three-phase OPF

The optimisation model is created with Pyomo and solved using the interior point method with IPOPT. The optimisation problem is expressed in rectangular coordinates, i.e. the complex variables and equations are separated into their real and imaginary parts. The initial feasible solution is coming from power flow calculation assuming that the PV production is zero. Since the power flow equations and some constraints are non-linear, there is no guarantee of reaching a global optimum.

5.4 Test network

The test network used for this study is a single feeder from an existing Belgian LV distribution network with a star configuration 400V/230V and an ungrounded neutral. Detailed unbalanced three-phase four-wire modelling of the network has been used according to [13] based on the data provided by the Distribution System Operator (DSO) (topology, line length, cable type, etc.). Figure 4.3 shows a graphical representation of the test network. The overhead lines are BAXB cables (Aluminium 3x70 mm^2 1x54 mm^2), and the connexion to the house uses BXB cables (Aluminium 4x10 mm^2). The impedance matrices $Z^{(b)}, \forall b \in \mathcal{B}$ are computed with PowerFactory.

[13] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

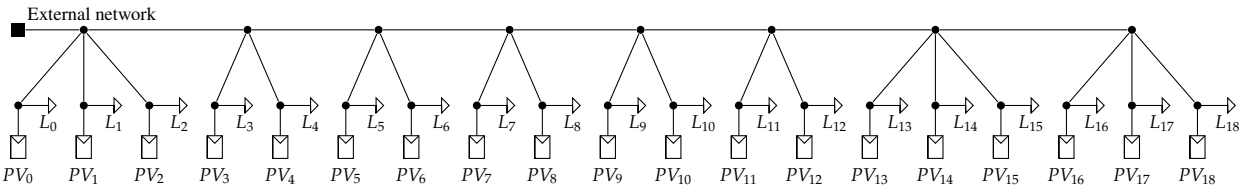


Figure 4.3: Graphical representation of the test network.

5.5 PV connection scenarios

As can be seen in Figure 4.3, each house is equipped with a PV unit. There are 17 5 kWp single-phase PV units, and two three-phase units of 7 kWp each (PV_9 and PV_{14}). They will be used to exhibit the behaviour of the three-phase group of inverters. According to the German and Italian standard, all PV units have a minimum power factor $PF_{\min}$ of 0.95.

Different scenarios for the connection of the single-phase PV units are studied. In all of them, the peak power of the PV unit is unchanged, only their connection to the three phases changes. The two three-phase units remain unchanged throughout each scenario.

- *3P*: The single-phase units are converted to three-phase ones. The production is thus balanced.
- *ABC*: PV_0 is connected to Phase A, PV_1 to Phase B, PV_2 to Phase C. This pattern is repeated until the end of the feeder.
- *AABC*: PV_0 is connected to Phase A, PV_1 to Phase A, PV_2 to Phase B, PV_3 to Phase C. This pattern is repeated until the end of the feeder.

- *AAABC*: PV_0 is connected to Phase A, PV_1 to Phase A, PV_2 to Phase A, PV_3 to Phase B, PV_3 to Phase C. This pattern is repeated until the end of the feeder.
- *AAA*: All single-phase inverters are connected to Phase A.

5.6 Loads

The consumption of the loads is defined by profiles created using a three-phase unbalanced version of [14]. The example of the active power consumed by load L_{18} is presented in Figure 4.4. The loads are assumed to have a constant power factor of 0.95 lagging.

[14] Richardson et al. ("Domestic electricity use: a high-resolution energy demand model")

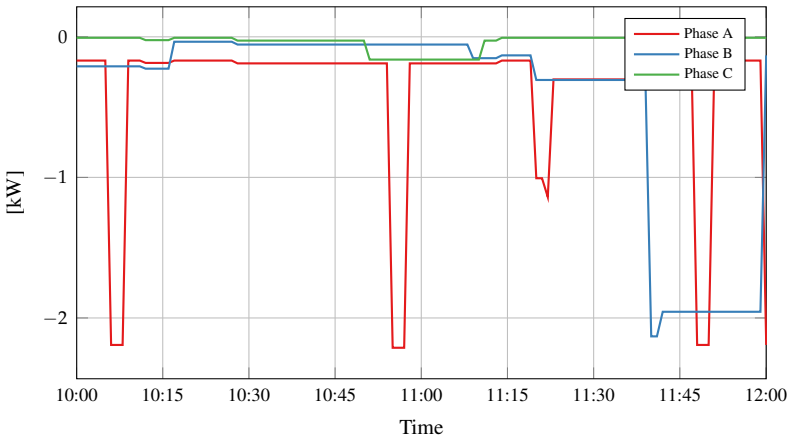


Figure 4.4: Active power consumption of L_{18}.

5.7 Numerical values of the parameters

The voltage should lie within $\pm 10\%$ of the nominal voltage, hence $V_{\min}$ = 0.9 p.u. and $V_{\max}$ = 1.1 p.u. The different values for the timers of the distributed control scheme are $t_{DQ} = t_{DP}$ = 5 min., t_{reset} = 5 min., and $t_{RQ} = t_{RP}$ = 10 min. In mode $Q_{loc}^{\downarrow\uparrow}(V)$, $V_1 = V_{\min}$ = 0.9 p.u., V_2 = 0.92 p.u., V_3 = 1.08 p.u. and $V_4 = V_{\max}$ = 1.1 p.u.

6 Simulation results

6.1 Synthetic production profile

This section demonstrates the behaviour of the distributed algorithm in the case of a synthetic production profile (Figure 4.5), specially designed to bring to light the different modes of the distributed control algorithm. The connection scenario used for the PV units is *AABC* to have a reasonably unbalanced network.

Phase B Group

Figure 4.6 shows the phase-to-neutral voltage of the single-phase PV unit PV_{11}, the one connected to Phase B which exhibits the highest voltage. As can be seen, there are no voltage problems for Phase B. Thus, no reactive

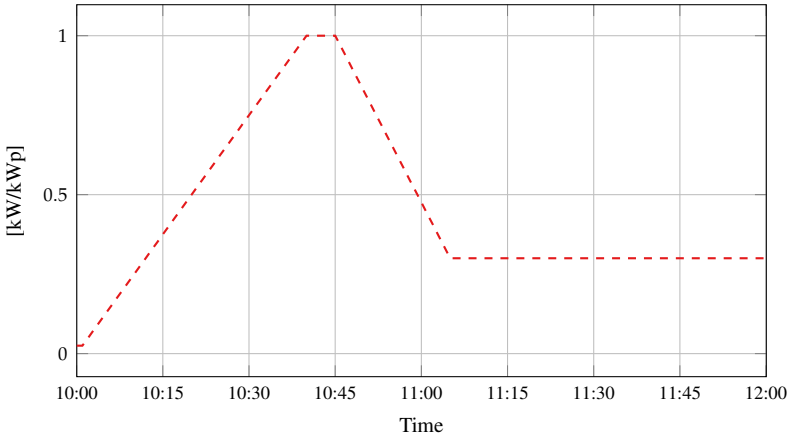


Figure 4.5: Maximum power point for the PV units in kW/kWp.

power is absorbed by PV units belonging to the Phase B group, nor is active power curtailed.

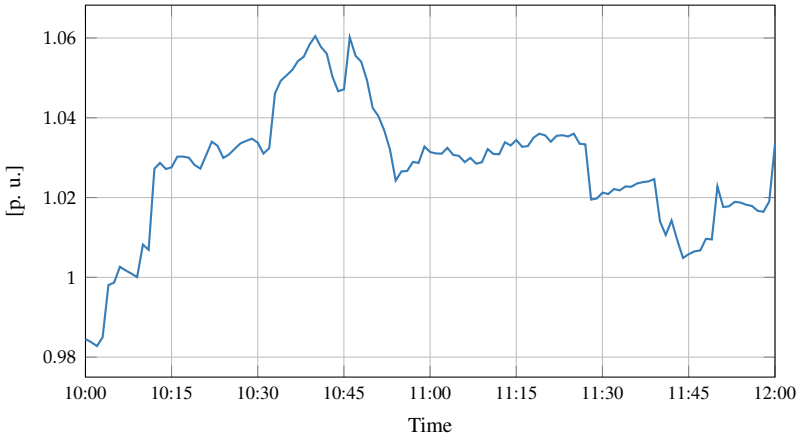


Figure 4.6: Phase-to-neutral voltage for PV_{11} (connected to Phase B).

Phase C Group

Figure 4.7 displays the actions of the controllers for the Phase C, Phase A and three-phase groups. Present n this figure are the PV units of each group which recorded the highest voltages at their point of connection, i.e. PV_{12} (Phase C), PV_{13} (Phase A) and PV_{14} (three-phase). The first line of graph shows the phase-to-neutral voltages, the second one the mode in which the controllers are, the third one corresponds to the active power produced, and the final one corresponds to the reactive power absorption.

Focusing on the first column of Figure 4.7, one can see that the voltage is close to 1.1 p.u., the overvoltage limit. The voltage exceeds this limit two times: one time at 10:41 and the other one at 10:46. At those moments, the PV units send a distress signal to all PV units connected to phase C, including the three-phase ones, and they switch to mode $Q^{\downarrow}$, where they gradually absorb more reactive power. Once the overvoltage has cleared, the PV units first move to Mode $Q^{\rightarrow}$, where they wait until they can attain Mode $Q^{\uparrow}$, where they decrease the absorption of reactive power. Finally, since there

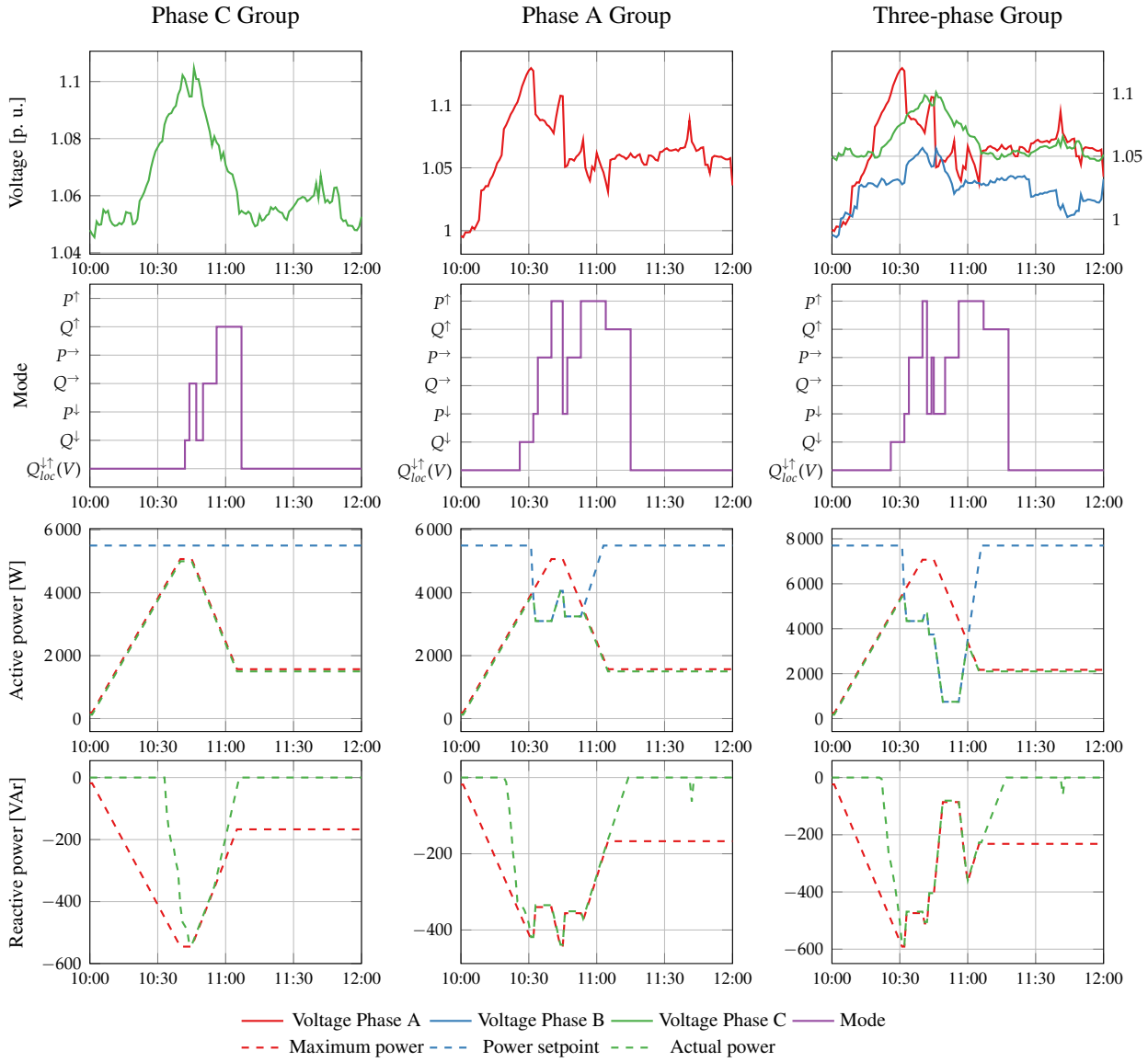


Figure 4.7: Active and reactive power production and absorption, controller mode and phase-to-neutral voltage of PV_{12} (Phase C group), PV_{13} (Phase A group) and PV_{14} (three-phase group).

is no new overvoltage, they proceed to Mode $Q_{loc}^{\downarrow\uparrow}(V)$, the normal mode of operation.

The PV units of the Phase C group never resort to active power curtailment, and only use reactive power to mitigate the voltage problems.

Phase A Group

The voltage of Phase A is the highest of the three phases. This is related to the fact that the majority of PV units are connected to it. Two overvoltages occur from 10:24 to 10:32, and from 10:44 to 10:45. During the first one, the PV units move to Mode $Q^{\downarrow}$, where they increase their absorption of reactive power. Unfortunately, the use of reactive power to lower the voltage is not sufficiently effective to mitigate the overvoltage, and after t_{DQ} minutes (5 min), they switch to the active power curtailment mode (Mode $P^{\downarrow}$). The latter has two effects on the inverters: the first one is the curtailment of active power, and the second one is the induced reduction of the maximum amount of reactive power that they can absorb owing to their capability curve. At 10:32, when the overvoltage is solved, the inverters move to waiting Mode $P^{\rightarrow}$, where they keep the active power constant for t_{reset} minutes, until they try to increase the active power, resulting in another overvoltage, forcing them to return to Mode $P^{\downarrow}$. The second time the inverters increase the active power, the maximum power point limit is sufficiently low for them to reach the limit without generating an overvoltage. They then proceed to Mode $Q^{\downarrow}$, where they gradually limit the absorption of reactive power to reach the value dictated by Mode $Q_{loc}^{\downarrow\uparrow}(V)$. Finally, reactive power absorption spikes at 11:42 due to the sudden increase in the voltage, a behaviour generated by the $Q_f(V)$ curve of Mode $Q_{loc}^{\downarrow\uparrow}(V)$.

Three-phase group

The three-phase group is the most impacted by the voltages given that it records the overvoltages in the three phases. Indeed, its actions are dictated by the union of the distress signals in all the phases. The main difference between the Phase A group and the three-phase group is the increased curtailment from 10:45 to 10:49 due to the overvoltage in Phase C, that is not yet relieved, forcing the three-phase inverters to proceed to additional curtailment.

External network

Figure 4.8 shows the power flowing in the line at the beginning of the feeder. First, there is no PV production and the feeder imports power from the external network. At 10:00, 10:07, 10:10, reverse power flows occur respectively in Phase C, A and B, which will last for the length of the simulations. The graph shows that power can flow in opposite directions at the same time. It also clearly shows that Phase A hosts the most PV production, and one can observe the two spikes around 10:30 and 10:45 due to active power curtailment.

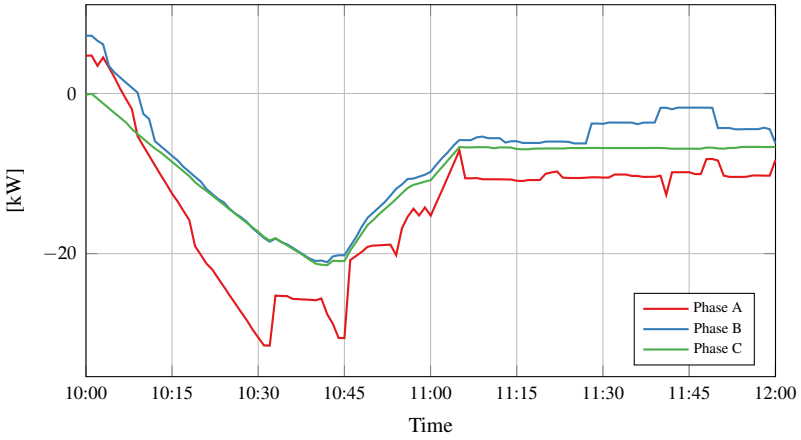


Figure 4.8: Distribution transformer: active power in each phase.

6.2 *Comparison between algorithms*

The three algorithms are compared in the setting of a sunny day where the maximum PV power ($P_{MPP,t}$) follows the profile shown in Figure 4.9. The maximum power is scaled according the peak power of the PV unit.

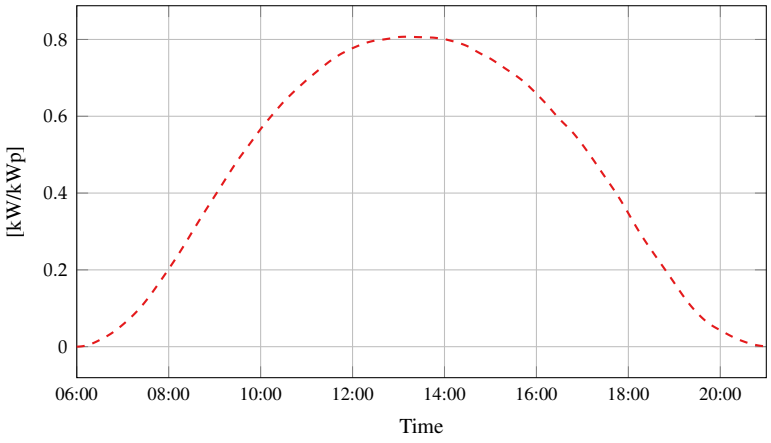


Figure 4.9: Maximum power for the PV units of 5 kWp

Curtailed energy

Figure 4.10 shows the total energy curtailed during the simulation. For each scenario, the relative performance of the algorithms is the same, with the OPF outperforming the distributed algorithm and the on-off algorithm. This is to be expected since the OPF performs an optimisation over the entire system with a detailed model of this one. The gap of performance between the three algorithms reduces when Phase A is increasingly loaded.

Of course, the curtailed energy increases when more PV units are connected to the same phase as imbalance reduces the hosting capacity of the network [15]. Indeed, the voltage rise in one phase depends on the amount of power transferred by the conductor.

Finally, the distributed algorithm and the OPF have poorer performances in scenario *3P* than in scenario *ABC*, because single-phase inverters are better suited to minimize curtailment, when the consumption of the houses is not balanced. Given that they can operate independently of the other phases,

[15] Bletterie et al. (“Understanding the effects of unsymmetrical infeed on the voltage rise for the design of suitable voltage control algorithms with PV”)

single-phase inverters only curtail power in the phase(s) with a voltage problem, whereas three-phase inverters curtail power in the three phases, even if only one of them suffers from an overvoltage.

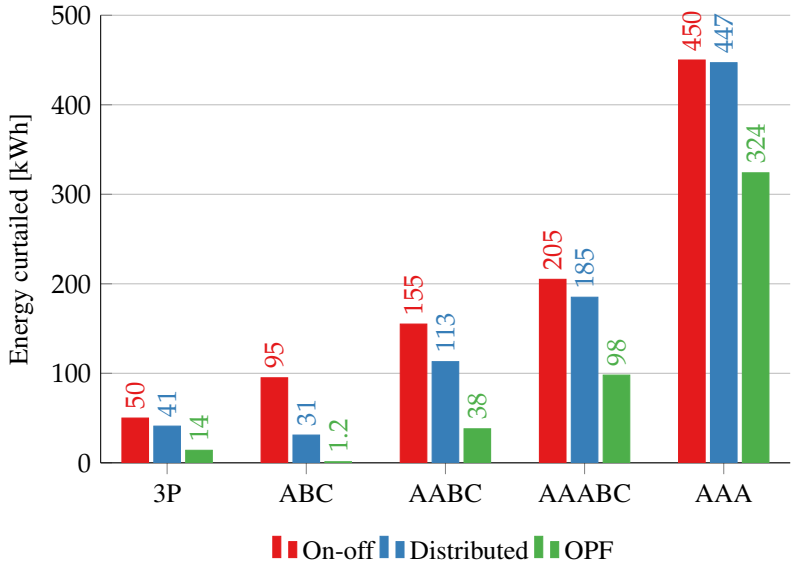


Figure 4.10: Comparison between the total energy curtailed by each algorithm in the different PV connection scenarios.

Reactive power usage

Figure 4.11 shows the use of reactive energy during the simulations. It is expressed in kVArh and counted regardless of its sign (absorption or production). Of course, the on-off algorithm makes no use of reactive power

The available reactive power is strongly dependent on the active power production, and thus the curtailed power, as it reduces the maximum amount of reactive power than the PV unit can produce or absorb (cf. Section 2.3). This explains the smaller use of reactive energy in the last configuration, as it is the configuration where power is most curtailed.

The distributed algorithm uses reactive power in a constant manner regardless of the scenario. Indeed, given that the maximum power that one phase can accommodate should remain almost identical between the scenarios, the reactive power energy use is relatively constant. However, the OPF produces and absorbs more reactive power because it optimises the three phases at the same time, leading to the production of reactive power in some phases to produce more active power in the others, whereas with the distributed algorithm, reactive power is used independently of the other phases.

Losses

The losses for the OPF are almost always larger than for the other algorithms. It increases the losses in the network to allow the production of more photovoltaic energy. This is linked to the two previous figures. However, the objective of the OPF ensures that the losses are not increased to a point where they would become prohibitive in regards to the marginal produced energy.

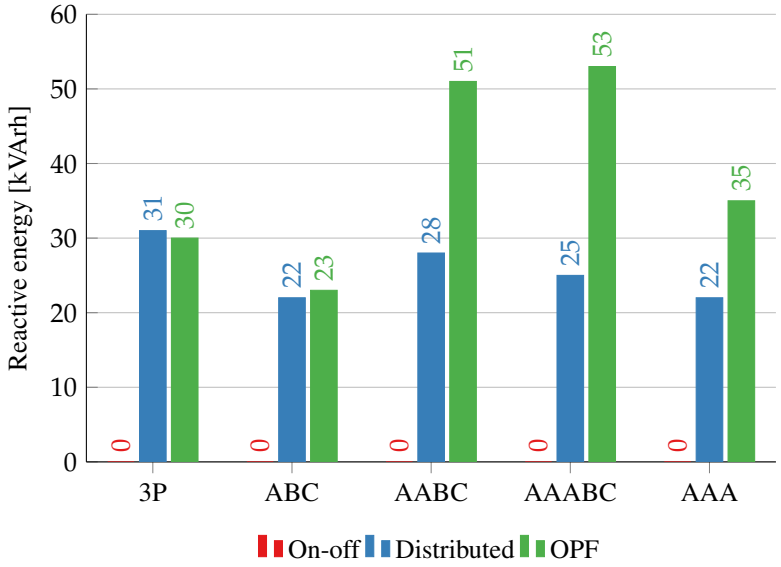


Figure 4.11: Comparison between the total reactive energy used by each algorithm in the different PV connection scenarios.

Moreover, the losses decrease from scenario *ABC* to scenario *AAA* because of the increased curtailed energy which lowers the currents in the network.

The losses with the on-off algorithm – especially in configuration *AAA* – are relatively high compared to the ones with the distributed algorithm that uses reactive power as a support. One explanation is the higher variations of current with the on-off algorithm. With the rapid disconnection and reconnection of the PV units, large temporary currents occur in the network. Since the losses are proportional to the square of the currents, if the variations of currents are higher, even if they have the same mean value, the losses will be larger. Furthermore, the rapid variations of current lead to voltage flicker, as shown in Figure 4.13, an effect undesirable for the proper network operations.

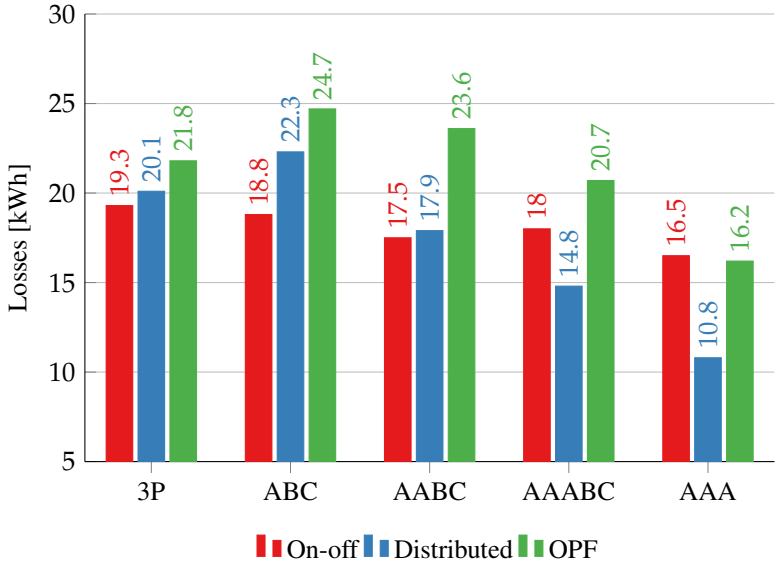


Figure 4.12: Comparison between the total losses in the network for each algorithm in the different PV connection scenarios.

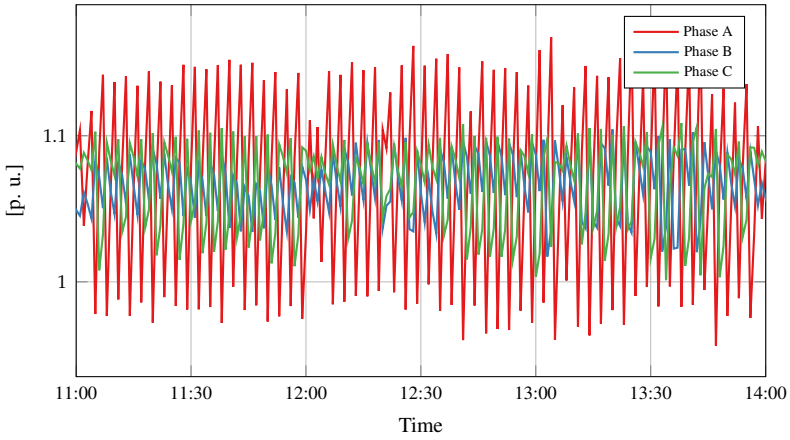


Figure 4.13: NDum15: Phase-to-neutral voltage (on-off – *AABC*). The voltage flicker is due to the rapid connection and disconnection of the PV units.

7 *Conclusion*

In this paper, we have presented the extension of a distributed control scheme, which controls the active and reactive power production of PV units, to prevent overvoltages. It does not require a detailed model of the LV network, and it only relies on limited communication in the form a distress signal when there is an overvoltage, in order to pool the resources.

The control scheme works by first partitioning the inverters into four groups, three for the single-phase inverters (one for each phase), and one for the three-phase ones. Each group then applies independently a distributed algorithm similar to the one presented in Chapter 3. The single distress signal is replaced by three distress signals, one for each phase. Since a group only reacts to the distress signal of the phase to which it is associated, single-phase inverters do not curtail power in a phase without voltage problems.

The behaviour of the proposed scheme is first illustrated and explained. Its performance are then compared to those of two reference schemes, one based on the on-off algorithm and the other one based on an unbalanced OPF. Its resulting total curtailed energy always lies between the two, with the on-off algorithm presenting the poorest performance, and the proposed algorithm losing its edge when the network is strongly unbalanced. It performs indeed better when the connection of the single-phase PV units is balanced between the phases. Finally, it generates less losses in the network than the other schemes, but at the expense of the produced energy, compared to the OPF.

Future works could include the detailed analysis of the reactive power produced by the OPF. Indeed, the use of reactive power in the phases that do not have overvoltage problem is changed to allow more PV energy to be produced in the problematic phases. This needs to be further studied to replicate this behaviour in a distributed way.

Chapter 5
Forseeing new control challenges in electricity prosumer communities

This paper is dedicated to electricity prosumer communities, which are groups of people producing, sharing and consuming electricity locally. This paper focuses on building a rigorous mathematical framework in order to formalise sequential decision making problems that may soon be encountered within electricity prosumer communities. After introducing our formalism, we propose a set of optimisation problems reflecting several types of theoretically optimal behaviours for energy exchanges between prosumers. We then discuss the advantages and disadvantages of centralised and decentralised schemes and provide illustrations of decision making strategies, allowing a prosumer community to generate more distributed electricity (compared to commonly applied strategies) by mitigating overvoltages over a low-voltage feeder. We finally investigate how to design distributed control schemes that may contribute reaching (at least partially) the objectives of the community, by resort in to machine learning techniques to extract, from centralised solution(s), decision making patterns to be applied locally. First empirical results show that, even after a post-processing phase so that it satisfies physical constraints, the learning approach still performs better than predetermined strategies targeting safety first, then cost minimisation.

In this Chapter, the values related to active power will be denoted by a subscript P and the ones related to reactive power by a subscript Q. This is to further emphasize the dual modelling of the exchanges between prosumers of the community both in terms of active and reactive power.

1 Introduction

This paper is dedicated to Electricity Prosumer Communities (EPCs), i.e. groups of people producing, sharing and consuming electricity locally. One of the main triggers of the emergence of the concept of energy communities is distributed electricity generation. By distributed electricity generation, we mainly mean PhotoVoltaic (PV) units, small wind turbines and Combined Heat and Power (CHP) that may be installed close to consumers. A cost-drop has been observed over past recent years, especially in the cost of producing PV panels. In addition to this, promises raised by recent advances made in the field of Electric Vehicles (EVs) and batteries may also emphasise in the coming years the metamorphosis of the electricity production, distri-

bution and consumption landscape that is already happening. In addition to electricity production and storage technology improvements, one should also mention the emergence of information technologies facilitating interactions between prosumers.[1] One should also note the existence of projects related to the use of distributed ledgers for managing energy exchanges[2] between microgrids[3].

Our goal is to propose a rigorous mathematical framework for studying energy prosumer communities. We first propose a mathematical framework for modelling the interactions between several prosumers. We then formalise a few optimisation problems targeting several different objectives (e.g., maximising "green" production, taking losses into account, optimising costs and revenues, etc). We address two ways to target these objectives: centralised and distributed control schemes, and we provide examples for each. In the centralised approach, we propose to design a community strategy dedicated to the maximisation of the local renewable energy production by formalising it as an Optimal Power Flow (OPF) problem. Then, in the context where we want to minimise community costs, we propose to design a distributed strategy that may still approach community optimality. To do so, we build upon Fortenbacher'work[4] who proposes a centralised optimal strategy named Forward-Backward Sweep OPF (FBS-OPF). We use time series provided by FBS-OPF solutions to build learning sets in the form $(input, output)$, where $input$ contains only local measurements related to one single prosumer, and where $output$ contains an optimal value that was output by the FBS-OPF for this input configuration. The learning sets are processed by machine learning techniques in order to build regressors able to compute suggestions for any input configuration.

The remainder of the paper is structured as follows: Section 2 describes EPCs and the main drivers for their emergence. Section 3 details our EPC mathematical formalisation, and Section 4 specifically focuses on formalising possible objectives for EPCs. Sections 5, 6 and 7 describe control strategies, from centralised to decentralised approaches, to target community objectives, also including numerical results. Section 8 discusses how community control could be extended to unbalanced three-phase networks, and Section 9 concludes.

[1] J. Wang, Costa, and Cisse ("From distribution feeder to microgrid: an insight on opportunities and challenges")

[2] Beitollahi and Deconinck ("Peer-to-peer networks applied to power grid")

[3] See for instance the Brooklyn Microgrid project.

[4] Fortenbacher, Zellner, and Andersson ("Optimal sizing and placement of distributed storage in low voltage networks")

2 *The Electricity Prosumer Community*

2.1 *Definition*

An Electricity Prosumer Community (EPC) is a group of electricity consumers and producers who decide to unite in order to reach a specific goal. In many ways, it is similar to a microgrid: it is composed of Distributed Energy Resources (DERs), loads (shiftable or not), possibly storage and EVs. However, it has a key distinction: it is not designed to operate off-the-grid. To reach this specific goal, it uses the flexibility of its assets to optimise its behaviour. Furthermore, we will assume that EPCs operate at a local level. It has been suggested to cluster consumers and producers in communities based on their production and consumption profiles regardless of their geographical location. This would be done in order to optimise their relations to

energy markets.[5] However, one driver for the community is to increase the hosting capacity of electrical networks and thus, the operational constraints should be taken into account and the impact of the community on the network studied. In this paper, we consider that the energy community is composed of prosumers that are connected to the same low-voltage distribution network and that there is only one point of connection between the community and the power system, which is called the root connection of the community. This means that the power exchanges between the prosumers do not transit through distribution transformers. Examples of such communities include companies in an industrial park, houses in suburban areas, co-ownerships of PV panels installed on rooftops of apartment buildings, etc.

[5] Rathnayaka et al. ("Goal-oriented prosumer community groups for the smart grid")

2.2 Drivers

On the one hand, the drivers for energy communities are clear when an infrastructure is shared between its members, be it DER, storage, CHP or EV charging stations. The installation of such infrastructure is conditioned by the existence itself of a community.

On the other hand, the drivers for a community where all members own their production units or storage are less trivial because they arise from the structure of the energy sector and the behaviour of electrical networks. If we consider communities that exist solely for financial purposes, it is certain that the benefits of the community have to surpass the sum of the benefits of the prosumers if they were isolated, taking into consideration the costs for operating the community. These benefits originate from two sides: network operations and energy markets.

Regarding the first, at the level of a single prosumer, self-consumption by consuming produced power locally allows the reduction of the network losses and the overall use of the networks, and a reduction of the electricity bill. Communities extend the perimeter of self-consumption from one prosumer to several to pool production and flexibility means. This pooling allows for an increased self-consumption of the community, for more optimised power flows, and an increased hosting capacity.

Considering energy markets, by pooling production and flexibility, communities could reach a size where they could offer services to the grid and negotiate their electricity prices.

3 Formalising an Energy Prosumer Community

3.1 The Prosumers

We consider a set of $N \in \mathbb{N}$ prosumers dynamically interacting with each other over a time horizon $T \in \mathbb{N}$. We first consider a discrete time setting: $t \in \{0, \dots, T-1\}$. In the following, power related variables are average values over a time window Δ, corresponding to the time interval between two time steps. At every time-step $t \in \{0, \dots, T-1\}$, each prosumer is characterised by active (resp. reactive[6]) power variables subscripted by P (resp. Q): production variables $P^{(i)}_{P,t}$ and $P^{(i)}_{Q,t}$, (note that $P^{(i)}_{Q,t}$ is positive when producing reactive power and negative when consuming it), a power injected into a storing device $S^{(i)}_t$, the level of charge of the storage device $\lambda^{(i)}_t$, loads

[6] Considering reactive power is important as it allows greater flexibility in the management of the networks and can allow the community to have more leverage on the network constraints.

(or consumptions) $L_{P,t}^{(i)}$ and $L_{Q,t}^{(i)}$, and powers injected into the distribution network $D_{P,t}^{(i)}$ and $D_{Q,t}^{(i)}$.

We assume that all prosumers may interact with each other. In particular, we denote by $\theta_t^{(i\to j)}$ the (positive) power that is transferred at time t from prosumer (i) to prosumer (j), $(i,j) \in \{1,\dots,N\}^2$. In the same time, prosumer (j) receives a (positive) power from prosumer (i) denoted by $\theta_t^{(j\leftarrow i)}$. By definition, we have:

$$\forall (i,j) \in \{1,\dots,N\}^2, \forall t \in \{0,\dots,T-1\}, \qquad \theta_t^{(i\to j)} = \theta_t^{(j\leftarrow i)} \tag{5.1}$$

with the convention that $\theta_t^{(i\to i)} = 0, \theta_t^{(i\leftarrow i)} = 0 \ \forall i,t$.

At every time step, the active power that is produced by prosumer (i) must satisfy the following relationship:

$$\forall t,i,\ P_{P,t}^{(i)} = L_{P,t}^{(i)} + D_{P,t}^{(i)} + S_{P,t}^{(i)} \tag{5.2}$$

$$\forall t,i,\ D_{P,t}^{(i)} = \sum_{j=1}^{N} \left(\theta_t^{(i\to j)} - \theta_t^{(i\leftarrow j)}\right) + \delta D_{P,t}^{(i)} \tag{5.3}$$

where $\delta D_{P,t}^{(i)}$ is the difference between the power injected into the distribution network and the sum of active power exchanges between the members of the community. Note that, in the case where the local production $P_{P,t}^{(i)}$ is not high enough to cover the load $L_{P,t}^{(i)}$, the variable $D_{P,t}^{(i)}$ may take some negative values (depending also on the amount of power that can be taken from the storage device).

The conservation of reactive power at the prosumer's location induces the following:

$$\forall t,i,\ P_{Q,t}^{(i)} = L_{Q,t}^{(i)} + D_{Q,t}^{(i)} \tag{5.4}$$

In this paper, we focus on energy exchanges among prosumers. For this reason, we choose to neglect electricity losses that are not directly associated with energy exchanges between prosumers.

Prosumers may not always be able to produce electricity at its maximal potential (for instance, PV production may be curtailed when the local storage device is fully recharged, no exchanges between prosumers are possible, and overvoltages are observed on the distribution network because many prosumers are injecting electricity simultaneously: such a situation may appear on sunny days[7]). Thus, for every prosumer, for every time-step, we define the maximal production potential which depends on hardware and weather data:

$$\forall t,i,\ P_{P,t}^{(i)} \leq P_{P,t}^{(i),\max} \tag{5.5}$$

$P_t^{(i),\max}$ may be influenced by several parameters, in particular, weather conditions.

The reactive power is limited by the capability curve of the distributed energy resources. It depends on the minimum power factor, the maximum apparent power, and the current active power production.

$$\forall t,i, \qquad \left|P_{Q,t}^{(i)}\right| \leq P_Q^{(i),\max}\left(P_{P,t}^{(i)}\right) \tag{5.6}$$

[7] Walling et al. ("Summary of distributed resources impact on power delivery systems") and Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

The injected power into the storage device is capped by a factor that mainly depends on the characteristics of the storage device and its current level of charge:

$$\forall t, i, \qquad \left|S_t^{(i)}\right| \leq S^{(i),\max}\left(\lambda_t^{(i)}\right) \tag{5.7}$$

The injected power into the distribution network is also capped, depending on the load and local production, characteristics and level of charge of the battery, and also additional (stochastic) variables, such as weather, that may influence the voltage of the low-voltage feeder (e.g. inability to inject power into the network if the voltage is higher than 1.1 p.u.), thus potentially preventing from power injection:

$$\forall t, i, \qquad \left|D_{P,t}^{(i)}\right| \leq D_P^{(i),\max}\left(P_{P,t}^{(i)}, P_{Q,t}^{(i)}, L_{P,t}^{(i)}, L_{Q,t}^{(i)}, S_t^{(i)}, D_{P,t}^{(j\neq i)}, D_{Q,t}^{(j\neq i)}\right) \tag{5.8}$$

$$\forall t, i, \qquad \left|D_{Q,t}^{(i)}\right| \leq D_Q^{(i),\max}\left(P_{P,t}^{(i)}, P_{Q,t}^{(i)}, L_{P,t}^{(i)}, L_{Q,t}^{(i)}, S_t^{(i)}, D_{P,t}^{(j\neq i)}, D_{Q,t}^{(j\neq i)}\right) \tag{5.9}$$

The level of charge of the storage capacity is also bounded:

$$\forall t, i \quad 0 \leq \lambda_t^{(i)} \leq \lambda^{(i),\max} \tag{5.10}$$

3.2 The Community

Everything that is not produced by the community has to come from the distribution network through the root of the community. By measuring the active and reactive power transfer at the root, and by comparing the measured powers to those measured at the prosumers' location, we can deduce the losses and the import of reactive power:

$$\forall t \quad \Lambda_{P,t}^{(c)} = D_{P,t}^{(c)} - \sum_{i=1}^{N} D_{P,t}^{(i)} \tag{5.11}$$

$$\forall t \quad \Lambda_{Q,t}^{(c)} = D_{Q,t}^{(c)} - \sum_{i=1}^{N} D_{Q,t}^{(i)} \tag{5.12}$$

where $\Lambda_{P,t}^{(c)}$ (resp. $\Lambda_{Q,t}^{(c)}$) denotes the overall losses inside the electrical network of the community (resp. reactive power absorbed by the community network lines), $D_{P,t}^{(c)}$ (resp. $D_{P,t}^{(c)}$) is the active (resp. reactive) power measured at the root of the community.

3.3 Costs and Revenues

At every time-step, we define a set of price variables, expressed in €/kWh. First, each prosumer (i) may purchase electricity from its retailer at a specific price $Pr_t^{(D\to i)}$. Also, each prosumer (i) may buy electricity from prosumer (j) $(j \neq i)$ at a price $Pr_t^{(j\to i)}$. Each prosumer (i) may also sell electricity to the (distribution) network at a price $Pr_t^{(i\to D)}$, and to other prosumers at a price $Pr_t^{(i\to j)}$. By convention, we assume that all prices considered in the paper are

positive. From time t to $t+1$, a prosumer (i) will incur the following cost:

$$c_t^{(i)} = \Delta\left(\max\left(-\delta D_t^{(i)}, 0\right) Pr_t^{(D\to i)} + \sum_{j=1}^{N} \max\left(\theta_t^{(i\leftarrow j)}, 0\right) Pr_t^{(j\to i)} \right) \tag{5.13}$$

At the same time, they will also receive the following revenues:

$$r_t^{(i)} = \Delta\left(\max\left(\delta D_t^{(i)}, 0\right) Pr_t^{(i\to D)} + \sum_{j=1}^{N} \max\left(\theta_t^{(j\leftarrow i)}, 0\right) Pr_t^{(j\to i)} \right) \tag{5.14}$$

3.4 Community Dynamics

The variables dynamically evolve over time, also suffering some stochasticity. We define a state vector Ξ_t as being the collection of all (measurable) variables related with the physical characteristics of the system, and a price vector Φ_t gathering all prices : $\forall t \in \{0, \dots, T-1\}$,

$$\Xi_t = \begin{pmatrix} P_{P,t}^{(1)} & P_{Q,t}^{(1)} \\ P_{P,t}^{(1),\max} & P_{Q,t}^{(1),\max} \\ \vdots & \vdots \\ P_{P,t}^{(N)} & P_{Q,t}^{(N)} \\ P_{P,t}^{(N),\max} & P_{Q,t}^{(N),\max} \\ S_t^{(1)} & \lambda_t^{(1)} \\ \vdots & \vdots \\ S_t^{(N)} & \lambda_t^{(N)} \\ L_{P,t}^{(1)} & L_{Q,t}^{(1)} \\ \vdots & \vdots \\ L_{P,t}^{(N)} & L_{Q,t}^{(N)} \\ D_{P,t}^{(1)} & D_{Q,t}^{(1)} \\ \vdots & \vdots \\ D_{P,t}^{(N)} & D_{Q,t}^{(N)} \end{pmatrix}, \quad \Phi_t = \begin{pmatrix} Pr_t^{(D\to 1)} \\ Pr_t^{(1\to D)} \\ \vdots \\ Pr_t^{(D\to N)} \\ Pr_t^{(N\to D)} \\ Pr_t^{(1\to 2)} \\ Pr_t^{(2\to 1)} \\ \vdots \\ Pr_t^{(1\to N)} \\ Pr_t^{(N\to 1)} \\ \vdots \\ Pr_t^{(N-1\to N)} \\ Pr_t^{(N\to N-1)} \end{pmatrix} \tag{5.15}$$

We also define two series of matrices. The first series $\Theta_t^{\rightarrow}$ is related to energy exchanges between prosumers according the producer's point of view, whereas the second series $\Theta_t^{\leftarrow}$ is written according to the receiver's (or consumer's) point of view:

$$\Theta_t^{\rightarrow} = \left(\theta_t^{(i\to j)}\right)_{i,j}, \qquad \Theta_t^{\leftarrow} = \left(\theta_t^{(i\leftarrow j)}\right)_{i,j} \tag{5.16}$$

Since it may not be easy to assess whether the system defined through the previously described state vectors is Markovian or not, we have : $\forall t \in \{0, \dots, T-1\}$,

$$\Xi_{t+1} = F(\Xi_t, \Phi_t, \Theta_t^{\rightarrow}, \Theta_t^{\leftarrow} \dots, \Xi_0, \Phi_0, \Theta_0^{\rightarrow}, \Theta_0^{\leftarrow}, \omega_t) \tag{5.17}$$

where $\omega_t \in \Omega$ is an exogenous random variable drawn according to an exogenous, time-dependent probability distribution $\omega_t \sim P_t(\cdot)$.

4 New Control Challenges

In this paper, we focus on the formalisation of decision making problems within a community of energy prosumers. Many control algorithms have already been proposed in literature, however, without specifically approaching it from a community angle.[8]) By decision making, we mean that at every time step, prosumers have the opportunity to make several decisions:

1. adapting their level of production and/or consumption,
2. buying/selling to other prosumers,
3. buying /selling to the retailer.

In the following, we detail a few optimisation criteria that may be considered when optimising a community of prosumers.

[8] Ahn and Peng ("Decentralized voltage control to minimize distribution power loss of microgrids"), Tonkoski, L. A. C. Lopes, and El-Fouly ("Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention"), and L. F. Ochoa and Harrison ("Minimizing energy losses: optimal accommodation and smart operation of renewable distributed generation")

4.1 Maximising the distributed production

As briefly discussed previously, it may occur that decentralised production may by curtailed, mainly because load, storage and the distribution network may not be able to host it on some sunny days. It may make sense to investigate control strategies dedicated to maximising decentralised production. More formally, one may seek to optimise, over the time horizon T, the production of decentralised electricity:

$$\max_{\substack{P_{P,t}^{(i)},P_{Q,t}^{(i)},L_{P,t}^{(i)},L_{Q,t}^{(i)},S_t^{(i)},\Theta_t^{\rightarrow},\Theta_t^{\leftarrow} \\ t\in\{0,\dots,T-1\} \\ i\in\{1,\dots,N\}}} \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{i=1}^{N} P_{P,t}^{(i)}\right] \tag{5.18}$$

while satisfying all constraints and time coupling between time steps.

Another optimisation criterion that may be of interest is to optimise distributed production while also limiting losses due to energy exchanges:

$$\max_{\substack{P_{P,t}^{(i)},P_{Q,t}^{(i)},L_{P,t}^{(i)},L_{Q,t}^{(i)},S_t^{(i)}, \\ \Theta_t^{\rightarrow},\Theta_t^{\leftarrow} \\ t\in\{0,\dots,T-1\} \\ i\in\{1,\dots,N\}}} \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{i=1}^{N} P_{P,t}^{(i)} - \Lambda_{P,t}^{(c)}\right] \tag{5.19}$$

while satisfying all constraints and time coupling between time-steps.

4.2 Optimising overall costs and revenues

Costs and revenues may be globally optimised by optimising the overall costs and revenues of the prosumer community:

$$\max_{\substack{P_{P,t}^{(i)},P_{Q,t}^{(i)},L_{P,t}^{(i)},L_{Q,t}^{(i)},S_t^{(i)}, \\ \Theta_t^{\rightarrow},\Theta_t^{\leftarrow} \\ t\in\{0,\dots,T-1\} \\ i\in\{1,\dots,N\}}} \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{i=1}^{N} r_t^{(i)} - c_t^{(i)}\right] \tag{5.20}$$

while satisfying all constraints and coupling between time-steps.

5 Control strategies

To achieve the objectives formalized in the previous section, two classes of control strategies compete with each other: a centralised one and a distributed one. All control strategies require controllable inverters, batteries, charging stations for active and reactive power. Controllable loads can also be considered. More specifically, in a centralised scheme, the modulation orders are computed by a centralised entity responsible for gathering the data, computing the orders and sending them to the prosumers. In a distributed scheme, all actors compute their own actions based on local objectives and measurements. The choice for a control strategy depends on several assumptions regarding the available information on the network (a detailed electrical model, estimation of the distance between the prosumer and the distribution transformer, etc.), the presence of communication (GPRS, PLC, Broadband, etc.), the presence of storage or, a central controller.

6 Centralised schemes

6.1 Technical challenges for building the centralised scheme

A centralised control scheme comprises three different parts. The first part is all the elements on which it relies for acquiring information about the system it controls. The second part is the "brain" of the scheme, something that is usually called the controller in the control literature. It computes, from the (history of) information, control actions. The third and last part is the infrastructure used for sending and applying its control actions. In the next subsections, we discuss the main elements of infrastructure that need to be put in place to build a centralised control scheme.

Information gathering

This part is typically composed of sensors used for measuring physical values, and of a communication infrastructure for sending them to the controller.

A centralised control scheme needs a full knowledge of the system. Therefore, the infrastructure needs to have: (i) Sensors able to measure the power consumed by the loads, the current state of charge of the batteries, estimation of the maximum production of DERs, etc. and (ii) communication channels able to transfer these measurements from the houses to the centralised controller. As communication channels, different technologies exist. For example, internet connections or General Packet Radio Service (GPRS) connections can be used. We could also think about using Power Line Communication (PLC) that carries data on the AC line. A mix of several communication technologies could also be used. For example, the data from the houses could be transmitted using a PLC-based technology to the nearest substation, from which GPRS technology would be used for transferring them to the centralised controller.

The centralised controller for processing the information

The second part of the infrastructure is related to the machinery needed for storing the information gathered about the system, processing this information, and computing the control actions based on measurements.

From computational results to applied actions

Once the actions have been computed by the centralised controller, they need to be applied to the system. This implies having a communication channel between the centralised controller and the inverters. This also implies having inverters which are able to modify, upon request, the amount of active and reactive power injected into the network.

6.2 Local Energy Markets

As a way to target the goals of an EPC, it has been suggested creating local energy markets to generate *incentives that boost investment in DER while at the same time creating enticements for containing and balancing the renewable energy produced*.[9] In this paper, the authors propose a combined market model for energy, flexibility and services at the neighbourhood level. The market is managed by an SESP (Smart Energy Service Provider) which can operate as a broker when local trades are peer-to-peer, as a retailer for over-the-counter sales with bilateral contracts, or as a market maker when a call auction is necessary.

[9] Bremdal et al. ("Creating a local energy market")

6.3 Optimal Power Flow

Another scheme to solve the control challenges in a centralised fashion is to use Optimal Power Flow techniques, where the objectives and the constraints are the ones in Section 3. In addition to those constraints, power flow constraints are added to link the powers injected at the node of the network to the voltages. Several methods[10] exist to solve such problems. A method of particular interest is the one developed by Fortenbacher et al.[11] where they adapt the Forward-Backward Sweep algorithm to OPF by linearising the power flow equations, given common assumptions that can be made in low-voltage distribution networks such as high R/X ratio, small angles deviation, etc.

[10] Gabash and Li ("Active-reactive optimal power flow in distribution networks with embedded generation and battery storage")

[11] Fortenbacher, Zellner, and Andersson ("Optimal sizing and placement of distributed storage in low voltage networks")

6.4 A first illustration: optimising PV injection over the network without storage

In this section, we assume that the low-voltage feeder gathers N houses, each of them being provided with a photovoltaic installation. We provide an illustration of the network in Figure 5.1. Presuming a deterministic setting, the following experiments show how to control the active power injected into the distribution network by each inverter for each time step in order to maximise the overall injected power while avoiding over-voltages, subject to operational constraints.:

$$\forall t \ \left(P_{P,t}^{(1),*}, \ldots, P_{P,t}^{(N),*}\right) \in \underset{P_{P,t}^{(1)},\ldots,P_{P,t}^{(N)}}{arg\,max} \quad \sum_{i=1}^{N} P_{P,t}^{(i)} \tag{5.21}$$

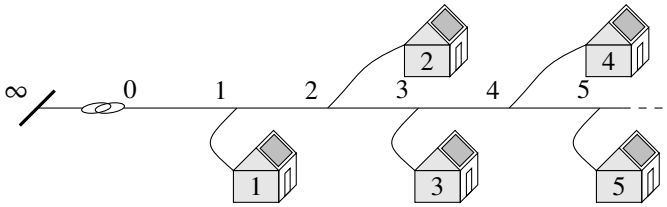


Figure 5.1: Graphic representation of the test network.

In the following, we assume that:

- The electrical distances between two neighbouring houses are the same and all electrical cables have the same electric properties,
- The line resistance is $0.24\ \Omega\ \mathrm{km}$,
- The line reactance is $0.1\ \Omega\ \mathrm{km}$,
- The distance between houses $50\ \mathrm{m}$,
- The nominal voltage of the network is 400 V,
- The value of the impedance of the Thévenin equivalent Y_{Th} is equal to $0.0059 + j0.0094\ \Omega$,
- The value of the Thévenin voltage is equal to 420 V.

As a consequence, for having a fully defined energy-based prosumer community; we just need to define the four following quantities:

- The number of houses is set to $N = 18$,
- The impedance between two neighbouring houses,
- The load profile, for every house and every time-step.

In Figure 5.2, we provide a graph of the evolution of the PV energy production for all the houses of the feeder. In Figure 5.3, we provide a graph of the evolution of the voltage for all buses of the feeder. One can observe that the production of houses located at the end of the feeder (i.e., far from the transformer) is modulated in order to avoid over-voltages. Even if the community still suffers partial curtailment, it has to be compared with the complete disconnection of PV units when overvoltages are observed. In the centralised community strategy, the total curtailment was 21.38 kWh, whereas the complete disconnection of inverters observing an overvoltage at their bus would lead to a curtailment of 31.63 kWh.

7 Distributed Schemes

In the previous section, we have proposed a centralised control scheme to suppress overvoltages in the community. As specified in Section 6.1, one of the main shortcomings of centralised controllers is their cost of implementation and maintenance. They indeed require building and maintaining a costly communication infrastructure between the houses and the centralised controllers. They also require a detailed model of the low-voltage network that may be expensive to obtain. Therefore, it would be interesting to design other types of control schemes that would be much cheaper. Ideally,

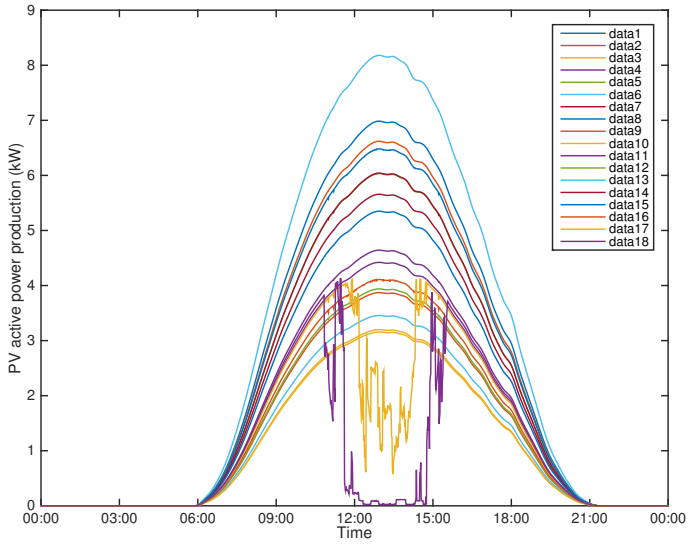


Figure 5.2: Illustration of the PV production of the 3 last houses of the low-voltage feeder over 24 hours.

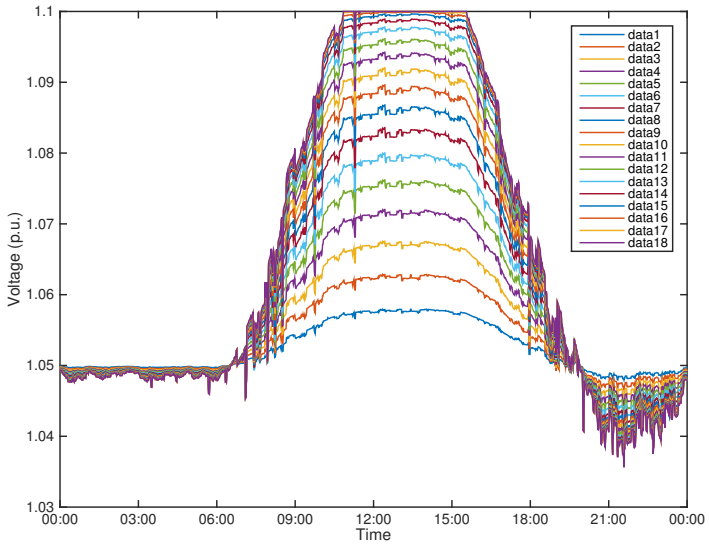


Figure 5.3: Illustration of the voltage potential over 24 hours.

these schemes should not rely on an expensive communication architecture and should be able to work even without knowing a detailed model of the low-voltage network.

In this section, we investigate how to design distributed control schemes that may contribute to reaching (at least partially) the objectives of the community. Our strategy is to resort to machine learning techniques that may extract, from centralised solution(s), decision making patterns to be applied locally, i.e. by only measuring features about the (local) prosumer. Our machine learning approach is an imitative learning-type approach where we learn four different regressors from data. These four regressors are dedicated

to learning the optimal levels of active power production, reactive power production, power injected into the storage device and power drawn from the storage device to be applied by prosumer (i).

7.1 Generating data for solving machine learning tasks

First, for each prosumer $i \in \{1, \ldots, N\}$, we generate a set of load profiles $L^{(i)}_{P,t}, t \in \{1, \ldots, T\}$ and maximal production potentials $P^{(i),max}_{P,t}, t \in \{1, \ldots, T\}$, associated with a time series of price vectors $\phi_t, t \in \{1, \ldots, T\}$. We consider the optimisation criterion described in Equation 5.20. This power flow problem can be solved using, for instance, the FBS-OPF algorithm.[12] Solving one such problem outputs a time series of data, corresponding to the evolution of all the indicators over the time horizon:

[12] Fortenbacher, Zellner, and Andersson ("Optimal sizing and placement of distributed storage in low voltage networks")

$$\left[\Xi^*_0, \ldots, \Xi^*_{T-1}\right] \tag{5.22}$$

From this time series of data, one can extract a series of local data, i.e. relative to one single prosumer (i):

$$\left[\Xi^{(i),*}_0, \ldots, \Xi^{(i),*}_{T-1}\right] \tag{5.23}$$

where $\forall t \in \{0, \ldots, T-1\}, \forall i \in \{1, \ldots, N\}$,

$$\Xi^{(i),*}_t = \begin{pmatrix} P^{(i),*}_{P,t} & P^{(i),*}_{Q,t} \\ P^{(i),\max}_{P,t} & P^{(i),\max}_{Q,t} \\ S^{(i),*}_t & \lambda^{(i),*}_t \\ L^{(i)}_{P,t} & L^{(i)}_{Q,t} \\ D^{(i),*}_{P,t} & D^{(i),*}_{Q,t} \end{pmatrix}, \tag{5.24}$$

From these extractions, we generate the following learning sets:

- For generating a learning set dedicated to learning how to optimize the level of active power production, we process the variables $\Xi^{(i),*}_t$ into the following set of (input, output) pairs:

$$\mathcal{L}^P = \left\{ \left(in^{i,t}_P, out^{i,t}_P \right) \right\}_{i=1,t=0}^{i=N,t=T-1} \tag{5.25}$$

where, $\forall t \in \{0, \ldots, T-1\}, \forall i \in \{1, \ldots, N\}$,

$$in^{i,t}_P = \left(i,\ |\overline{\mathbf{v}}^{(i)}_t|,\ arg(\overline{\mathbf{v}}^{(i),*}_t),\ \phi_t,\ \lambda^{(i),*}_t,\ L^{(i)}_{P,t},\ P^{(i),max}_{P,t} \right) \tag{5.26}$$

$$out^{i,t}_P = P^{(i),*}_{P,t} \tag{5.27}$$

where:

- i : id number of the bus
- $|\overline{\mathbf{v}}^{(i)}_t|$: magnitude of the voltage at bus i at time step t
- $arg(\overline{\mathbf{v}}^{(i)}_t)$: phase of the voltage at bus i at time step t
- ϕ_t : electricity price at time step t, considered as being unique in the whole feeder

- $\lambda_t^{(i)}$: level of charge of the storage of bus i at time step t
- $L_{P,t}^{(i)}$: load consumption at bus i at time step t
- $P_{P,t}^{(i),max}$: maximal production potential at bus i at time step t

- For generating a learning set dedicated to learning how to optimize the level of reactive power production, we process the whole variables $\Xi_t^{(i),*}$ into the following set of (input, output) pairs:

$$\mathcal{L}^Q = \left\{ \left(in_Q^{i,t}, out_Q^{i,t} \right) \right\}_{i=1,t=0}^{i=N,t=T-1} \tag{5.28}$$

where, $\forall t \in \{0, \ldots, T-1\}, \forall i \in \{1, \ldots, N\}$,

$$\begin{aligned} in_Q^{i,t} &= &= in_P^{i,t} \\ out_Q^{i,t} &= &= P_{Q,t}^{(i),*} \end{aligned}$$

- For generating a learning set dedicated to learning how to optimize the level of power injected into the battery, we process the whole variables $\Xi_t^{(i),*}$ into the following set of (input, output) pairs:

$$\mathcal{L}^C = \left\{ \left(in_C^{i,t}, out_C^{i,t} \right) \right\}_{i=1,t=0}^{i=N,t=T-1} \tag{5.29}$$

where, $\forall t \in \{0, \ldots, T-1\}, \forall i \in \{1, \ldots, N\}$,

$$\begin{aligned} in_C^{i,t} &= in_P^{i,t} \\ out_C^{i,t} &= \max\left(S_t^{(i),*}, 0 \right) \end{aligned}$$

- For generating a learning set dedicated to learning how to optimize the level of power injected into the battery, we process the whole variables $\Xi_t^{(i),*}$ into the following set of (input, output) pairs:

$$\mathcal{L}^D = \left\{ \left(in_D^{i,t}, out_D^{i,t} \right) \right\}_{i=1,t=0}^{i=N,t=T-1} \tag{5.30}$$

where, $\forall t \in \{0, \ldots, T-1\}, \forall i \in \{1, \ldots, N\}$,

$$\begin{aligned} in_D^{i,t} &= in_P^{i,t} \\ out_D^{i,t} &= \max\left(-S_t^{(i),*}, 0 \right) \end{aligned}$$

The machine learning task is performed using Extremely Randomized Trees[13] using the Scikit-learn library.[14]

[13] Geurts, Ernst, and Wehenkel ("Extremely randomized trees")

[14] Pedregosa et al. ("Scikit-learn: machine learning in Python")

7.2 Post-processing the predictions

When the regressors learned from data are used to set the value of a decision variable inside the community, their output needs to be post-processed, otherwise it could create a violation of physical constraints (e.g. the predicted value of power drawn from the storage is greater than the power that the storage can offer). In that case, the value is corrected and set equal to the limit that it crossed (e.g. the power drawn from the storage becomes equal to the maximum power that the storage can offer).

7.3 Applying the learned strategies in different load, solar production and prices configurations

A new set of load profile time series $L^{(i)}_{P,t}, t \in \{1, \dots, T\}$ and maximal production potentials time series $P^{(i),max}_{P,t}, t \in \{1, \dots, T\}$, associated with a new time series of price vectors $\phi_t, t \in \{1, \dots, T\}$ are generated for each prosumer $i \in \{1, \dots, N\}$. Starting from the initial time step, at every t, the required inputs are passed to the regressors for each prosumer and the outputs (after a post-processing step) are used to set the value of their actions. The power flow problem is solved every time to check the voltages, the net power exchanged with the main grid and respect of the physical constraints.

7.4 Empirical illustration

In this section, we compare the performance of the learned strategies in a deterministic setting with two other strategies: (i) the centralised optimised strategy as defined by Fortenbacher,[15] and (ii) another decentralised strategy relying on a predetermined, thresholds-based, decision rule. This second decentralised strategy is designed so that it ensures the safety of the system, and then, tries to restrain the overall costs of the community. The first point of this second decentralised algorithm is, thus, to check if there is a risk of overvoltages or undervoltages at the bus and, in this case, to orient the actions of that prosumer to avoid it (fully charging/discharging the storage and maximising/minimising the power production). In the case where the safety of the grid seems ensured, the decisions are imposed based on the price of the electricity at that time step (when it is above/under a predetermined price, impose a predetermined prosumer's action). It is certainly simplistic, but it has the merit of providing a comparison base. Details about this decision rule can be found in Appendix.

[15] Fortenbacher, Zellner, and Andersson ("Optimal sizing and placement of distributed storage in low voltage networks")

As a comparison metric, we consider the overall costs that the community incurs (in the same overall environment, i.e. same loads, solar production, PV and batteries sizes, prices) exchanging power with the main grid during an entire year ($T = 8760$, one time-step per hour during one year). The comparison is made in an environment where loads, solar production and prices are not the same as the one from which the learned strategies were built. As expected, the centralised model is able to achieve the lowest costs, equal to 641.70 €. If we adopt the predicted actions made by the learned regressors, the community meets a total cost of 1549.70 €, a result that seems expensive when compared to the centralised model one, but it becomes remarkable when we consider that the "rule of thumb" algorithm produces an expense equal to 3276 €.

8 One step further: taking into account the three phases

The mathematical formalisation presented in this paper considers a balanced operation of the network. Indeed, the power exchanges between the prosumers do not take into consideration the phase to which they are connected. It considers only one value for active and reactive power per dwelling. However, low-voltage distribution networks are intrinsically unbalanced because

even if a prosumer has a three-phase connection to the grid, house appliances are mainly single phase. Our concern is the relevance of exchanging powers between members of the electricity community, that are not connected to the same phase. Physically, current from the DER would flow to the distribution transformer and out of the community while current to supply the load would flow from the distribution transformer. While this may reduce the losses to some extent because power does not flow from the transmission network, controlling the community in this fashion could further unbalance the network, and result in the violation of voltage constraints and a reduction of the hosting capacity of the network. One solution would be to divide the community into three groups: one per phase and prosumers that have a three-phase connection would belong to the three groups. This would allow the application of the same formalism and would ensure that exchange of powers takes place solely between prosumers that are connected to the same phase.

9 Conclusion

This paper proposes a mathematical framework for modelling energy exchanges between prosumers. In particular, it allows formalising a family of optimisation problems, depending on the target objective (maximising green production, minimising losses, optimising revenues), and also on the structure of the control strategies (centralised or decentralised). In particular, we have proposed a machine learning approach whose objective is to mimic, at an individual level (i.e., using local measurements only), a behaviour that is optimal at the community level. First empirical results show that, even after a post-processing phase so that it satisfies physical constraints, the learning approach still performs better than predetermined strategies targeting safety first, then cost minimisation.

Future work includes designing decentralized strategies relying on other data-based approaches, in particular, we think Reinforcement Learning (RL)[16] could be a powerful paradigm to learn decentralised strategies, mainly because RL agents have the characteristic to self-improve over time with new data acquisition.

[16] Glavic, Fonteneau, and Ernst ("Reinforcement learning for electric power system decision and control: past considerations and perspectives")

Appendix

We define $\eta_c^{(i)}$ as the charge efficiency and $\eta_d^{(i)}$ as the discharge efficiency of the storage of prosumer $i \in \{1, \dots, N\}$. We set arbitrary values for ϕ^- and ϕ^+, that are thresholds, respectively, for what can be considered a low price and an high price for electricity. At every time-step $t \in \{0, \dots, T-1\}$ and for each prosumer $i \in \{1, \dots, N\}$ the "rule of thumb" algorithm used in Section 7 is structured as follows:

if $|\overline{\mathbf{v}}_t^{(i)}| \leqslant 0.91 pu$

$P_{P,t}^{(i)} = P_{P,t}^{(i),max}$

$P_{Q,t}^{(i)} = P_{Q,t}^{(i),max}$

$S_t^{(i)} = -\lambda_t^{(i)} \eta_d^{(i)}$

else if $|\overline{\mathbf{v}}_t^{(i)}| \geqslant 1.09 pu$

$P_{P,t}^{(i)} = 0$

$P_{Q,t}^{(i)} = -P_{Q,t}^{(i),max}$

$S_t^{(i)} = \frac{\lambda_t^{(i),max} - \lambda_t^{(i)}}{\eta_c^{(i)}}$

else

$P_{P,t}^{(i)} = P_{P,t}^{(i),max}$

$P_{Q,t}^{(i)} = 0$

if $\phi_t \geqslant \phi^+$

if $P_{P,t}^{(i)} \geqslant L_{P,t}^{(i)}$

if $\lambda_t^{(i)} \geqslant 0.3\ \lambda_t^{(i),max}$

$S_t^{(i)} = -\left(\lambda_t^{(i)} - 0.3\ \lambda_t^{(i),max}\right) \eta_d^{(i)}$

else

$S_t^{(i)} = 0$

else

$S_t^{(i)} = -\lambda_t^{(i)}\ \eta_d^{(i)}$

else if $\phi_t \leqslant \phi^-$

if $P_{P,t}^{(i)} \geqslant L_{P,t}^{(i)}$

if $P_{P,t}^{(i)} - L_{P,t}^{(i)} \leqslant \left(\lambda_t^{(i),max} - \lambda_t^{(i)}\right) \eta_c^{(i)}$

$S_t^{(i)} = \frac{P_{P,t}^{(i)} - L_{P,t}^{(i)}}{\eta_c^{(i)}}$

else

$S_t^{(i)} = \frac{\lambda_t^{(i),max} - \lambda_t^{(i)}}{\eta_c^{(i)}}$

else

if $\lambda_t^{(i)} \geqslant 0.3\ \lambda_t^{(i),max}$

$S_t^{(i)} = -\left(\lambda_t^{(i)} - 0.3\ \lambda_t^{(i),max}\right) \eta_d^{(i)}$

else

$S_t^{(i)} = \frac{0.3\ \lambda_t^{(i),max} - \lambda_t^{(i)}}{\eta_c^{(i)}}$

if $S_t^{(i)} > S_t^{(i),max}$

$S_t^{(i)} = S_t^{(i),max}$

if $S_t^{(i)} < -S_t^{(i),max}$

$S_t^{(i)} = -S_t^{(i),max}$

General conclusion

With the decrease in the price of photovoltaic panels and electric cars, problems of an electrical nature appear on the distribution networks, in particular low-voltage ones. When their electricity production level is high, photovoltaic units create overvoltage problems in the low-voltage grid. These overvoltages can lead to the malfunction of electrical appliances designed to operate in a specific voltage range, as well as the disconnection of PV units that are no longer producing electricity when the voltage is too high. This represents a financial loss for the owner of these units, and has the potential to considerably slow down the deployment of PV units should those problems remain unsolved. Moreover, as the distribution networks are three phase, these problems are often aggravated by the imbalance between the phases. Since not every phase of the network carries the same power and current, the number of PV installations or maximum number of charging stations that can be installed in the network without creating problems is more limited, i.e. the hosting capacity of the network is reduced.

To increase the hosting capacity of the network, we have investigated two solutions as an alternative to network reinforcement. The first one is the balancing of the production on the three phases, which implies first identifying the phases, a problem that we have addressed in this thesis. The second one is the control of the active and reactive power injection into the network. To do so, we proposed, on the one hand, a distributed control scheme for the PV units alone, and on the other hand, the framework of Energy Prosumer Communities to control the injections, whether they were induced by PV units, storage or demand response, and to reach the common goal of the community while satisfying the operational constraints of the network.

In the following sections, we summarise our contributions and detail future research directions related to them.

1 Identifying the phases of the smart meters

At the root of more balanced low-voltage distribution networks is the identification of the phases of the smart meters. If the smart meters use GPRS or RF technologies to repatriate the measurements, the DSOs do not know to which phase(s) of the network these measurements correspond. Taking advantage of the correlation between the voltages of the same phase at different points of the network, we have developed algorithms to group the time series of voltage measurements phase by phase to resolve the identification of the phases.

This part provided a creative and innovative response to a problem that will become increasingly preponderant with the development of communicating meters.

We introduced a novel method to identify the phases of smart meters in LV distribution networks by clustering the voltage measurements using graph theory and the correlation between measurements. The algorithm, named constrained multi-tree clustering, successfully manages to identify the phases of smart meters based on real voltage measurements from an LV network in Belgium. It takes, as input, a root smart meter upon which the clustering process is done. Computations show that a good choice for the root is the meter at the distribution transformer. This is mainly due to the central position played by the distribution transformer in LV networks. The performance of the algorithm in various settings is compared to those of constrained k-means clustering. The constrained multi-tree method performs better regardless of the ratio between single-phase and three-phase smart meters, or the increasing number of houses that are not monitored, i.e. a very small set of measurements. Moreover, it performs very well with a very small set of measurements. Finally, another one of its advantages is its capacity to properly handle voltages that are averaged over a longer period of time without wrongly identifying smart meters.

In this thesis, the algorithms have been tested on voltage measurements from a three-phase four-wire network with an ungrounded neutral. Future works could include tests on measurements from other network configurations, such as: (i) three-phase four-wire with a grounded neutral, (ii) three-phase three-wire (3x230V) without a neutral. The proposed algorithm can be applied as such to those networks since the identification is based on the fact that voltage measurements between the same phases are better correlated than measurements between different phases. On the one hand, the neutral voltage is kept low in configuration (i) thanks to the grounding of the neutral. Henceforth, it will have a limited impact on the phase-to-neutral voltages, thus decreasing the correlation between the phases compared to the ungrounded configuration. The performance is thus expected to increase. On the other hand, the chance of misidentification should increase in configuration (ii) because the correlation between two phase-to-phase voltage measurements should be higher than the one between two phase-to-neutral voltage measurements because they share the voltage of the common phase.

We could also improve the method to handle which cluster is growing and at what pace, to avoid the growth of one cluster at the expense of the others. Finally, it could be interesting to use this novel method to infer network topology, especially since the tree-structured assumption seems very well adapted to distribution networks.

2 Distributed control algorithm

Obtaining an accurate model of a low-voltage network is complicated and time-consuming, as the connection points are so numerous and some information is incomplete. We therefore developed a distributed scheme to control the PV energy production that does not require a network model to alleviate the voltage problems caused by a large penetration of PV units in LV DNs.

It avoids overvoltages in the network thanks to a coordinated response of the PV within a feeder, via a distress signal during overvoltages. This distress signal allows the reactive support and curtailment to be distributed over the various PV units of the feeder. By pooling the resources, it prevents the production of electricity of a specific group of PV units at the expense of another.

Through balanced three-phase dynamic simulations, the behaviour of the scheme has been analysed and it was shown that its performance is comparable to a centralized, OPF-based control scheme which would be more expensive to deploy and complex to operate (for instance, when the DN topology changes).

We have also presented the extension of the distributed scheme, so that it could be used in unbalanced three-phase networks with single-phase inverters. Instead of a single one, there are now three distress signals, one for each phase. The extension gathers the inverters in four groups, one per phase and one for the three-phase inverters, in order to ensure that there is no curtailment in a phase without voltage problems. Indeed, they only react to the distress signal of the phase to which they are connected.

The proposed control scheme performs well compared to two reference schemes, one based on the on-off algorithm and the other one based on an unbalanced OPF. Its resulting total curtailed energy always lies between the two, with the on-off presenting the poorest performance, and the proposed algorithm losing its edge when the network is strongly unbalanced. Indeed, it performs better when the connection of the single-phase PV units is balanced between the phases. Furthermore, unbalanced simulations show that the hosting capacity of the network is indeed reduced when the PV units are not well balanced over the different phases.

Several research directions exist for this distributed control scheme. First, rather than relying on active power curtailment, the use of local storage systems or flexible loads could be considered. This would fundamentally change the control problem considered and would raise new research questions. For example, the existence of local energy storage would couple the control actions at different times, since accumulating active power into storage at a specific time influences the amount of energy available in the future.

Then, a cost analysis could be performed to evaluate the economical profitability of deploying such a distributed scheme compared to a reinforcement of the network.

Finally, future works could include the detailed analysis of the reactive power produced by the OPF. Indeed, the use of reactive power in the phases that do not have overvoltage problem is changed to allow more PV energy to be produced in the problematic phases. This needs to be further studied to replicate this behaviour in a distributed way.

3 Local energy communities

Finally, we have proposed a mathematical framework to model energy exchanges between prosumers within a community. In particular, it allows formalising a family of optimisation problems, depending on the target objective (maximising green production, minimising losses, optimising rev-

enues), and also on the structure of the control strategies (centralised or decentralised). In particular, we have proposed a machine learning approach whose objective is to mimic, at an individual level (i.e., using local measurements only), a behaviour that is optimal at the community level. First empirical results show that, even after a post-processing phase so that it satisfies physical constraints, the learning approach still performs better than predetermined strategies targeting safety first, then cost minimisation.

Future work includes designing decentralized strategies relying on other data-based approaches. In particular, we think Reinforcement Learning (RL)[17] could be a powerful paradigm to learn decentralised strategies, mainly because RL agents have the characteristic to self-improve over time with new data acquisition.

[17] Glavic, Fonteneau, and Ernst ("Reinforcement learning for electric power system decision and control: past considerations and perspectives")

4 Going one step further – From Research to Industry

In addition to distributed generation, electric vehicles will have a strong impact of electricity distribution. Charging electric cars can severely congest an electrical network, i.e. cause too much current to pass through the cables and the distribution transformer, which could damage them and cause undervoltages.

Network operators therefore wish to improve knowledge understanding, and management of their LV networks, in particular by studying the impact of PV production and EV charging. This will require the creation of an electrical model of the network, and the use of load-flow simulations, a practice not yet widespread for LV networks. Moreover, it is also important to know the penetration rate of PV units and electric cars from which these problems will arise. After this identification stage, it is also necessary to develop preventive actions to eliminate problems before they occur, and corrective actions when limiting the charging power of electric vehicles is necessary to ensure the safety of the network.

From a practical point of view, we should develop the knowledge, methodologies and algorithms necessary to enable DSOs to better manage their unbalanced LV networks in order to avoid heavy network upgrading investments, by answering the following questions:

1. How to automatically reconstruct the structure of LV networks from the data available to the DSO, in particular the measurements collected by the smart meters. This model is needed to assess the impact of PV panels and electric vehicles on the grid.

2. How to identify the LV networks most at risk in terms of voltage, congestion and imbalance problems, i.e. how to evaluate the hosting capacity of the networks and the maximum penetration rate.

3. How to increase the hosting capacity of networks in a preventive way, for example, by rebalancing phases, by choosing connection points for public charging stations, etc.

4. How to control, in a corrective way, electric car charging or distributed generation to avoid grid congestion or overvoltage.

These four questions provide four directions to transform the work done in this thesis into a tool that could be used by Distribution System Operators to operate their low-voltage distribution networks and increase their hosting capacity.

Appendix

Appendix A
Modelling of three-phase four-wire low-voltage cables taking into account the neutral connection to the earth

Local energy communities (LECs) usually occur at the level of low-voltage distribution networks, which are inherently unbalanced due to single-phase household appliances and distributed generation. To simulate and optimise the behaviour of an LEC, the three phases and neutral must be modelled explicitly. This paper aims at numerically assessing the influence of the modelling of the earth and the connection between the neutral and the earth, in terms of voltages and currents. The simulations are performed on an existing Belgian low-voltage feeder supplying 19 houses, which are all equipped with a smart meter measuring the mean voltage, current, and active and reactive power every minute for each phase. The simulations show that the explicit modelling of the earth using Carson's equations has a moderate effect on the simulation results. In particular, it creates differences in the simulations that are around ten-times smaller than the errors between simulations and measurements.

1 Introduction

Local energy communities (LEC) usually occur at the level of low-voltage distribution networks, which are inherently unbalanced due to single-phase household appliances and distributed generation. The common assumption made in high-voltage and medium-voltage networks – that the three phases are balanced – can no longer be made. To simulate and optimise the behaviour of an LEC, the three phases and neutral must be modelled explicitly. Indeed, in some cases, the current in the neutral is of the same order of magnitude as the phase current, which can have two effects: on the one hand, if the neutral conductor is not grounded, the neutral current creates a neutral voltage, which modifies the voltage of the phases. This phenomenon is called neutral point shifting.[1] On the other hand, if the neutral conductor is grounded, there is an unchecked current flowing through the earth. In either case, this should be taken into account.

[1] Degroote et al. ("Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks")

This paper aims at reviewing the main assumptions made to model a three-phase electrical line, and more specifically at assessing numerically the influence of the modelling of the earth. This is done in the setting of an existing Belgian low-voltage (LV) distribution feeder. Each house has been equipped with a smart meter measuring the one-minute average voltage, current, and

active and reactive power for each phase. One minute has been chosen to clearly exhibit the imbalance in the network and the variability of the loads. A longer averaging period would smooth out those characteristics[2] that emphasize the impact of the neutral and earth modelling.

The paper is organized as follows. The first section will review the common modelling hypotheses and simplifications for LV electrical lines. The second section will numerically assess the impact of Carson's equations to model the earth return path, and finally, the third section will compare the simulation results with the actual measurements from the smart meters.

[2] Wright and Firth ("The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations")

2 *Electrical Line Modelling*

In this study, the electrical lines are composed of four conductors (three phases and one neutral). In the simulations, they are modelled as PI equivalents with an impedance matrix, which corresponds to the resistance, self- and mutual impedance of the conductors, and a shunt admittance matrix, representing the capacitance between the conductors. This admittance matrix is divided in two halves, one at the sending end of the line and the other at the receiving end. Both matrices are computed using PowerFactory, by entering the physical constants of the conductors (material, section, etc.). This implies that the impedance does not depend on the length of the cable, i.e. that the leakage flux at the end of the lines can be neglected.[3] Finally, since the aim of this study is to model LEC to simulate their operations and optimize their behaviour, only power-flows in steady-state will be considered (no harmonics or fault studies).

[3] Urquhart ("Accuracy of Low Voltage Distribution Network Modelling")

2.1 *Carson's equations and earth return path*

PowerFactory uses Carson's equations[4] to update the impedance matrix to take into consideration the earth return path. Let Z_4 be the 4x4 impedance matrix with only the resistance of the conductors and their self- and mutual impedance. Carson's equations provide a way to modify Z_4 to take into account the earth return path.[5] Let $\hat{Z}_4$ be the updated version of Z_4. Thanks to the work of Kersting,[6] the earth impedance can be extracted from $\hat{Z}_4$ to create Z_5, a 5x5 matrix where the earth return path is modelled explicitly as an extra-conductor. This formulation has been used to solve power flow in Ciric's paper.[7] However, Carson's equations were originally designed for overhead lines with widely spaced conductors, which is not the case in LV networks. Moreover, even if it has been shown that they can be applied to underground cables for the mains frequency,[8] they use image theory to compute the earth return path, and thus assume the latter is parallel to the conductors, a hypothesis which may not hold in an LV distribution system presenting a complex topology and a proximity to underground pipes, which can provide a path with a smaller impedance.

[4] Carson ("Wave propagation in overhead wires with ground return")

[5] W. H. Kersting and Green ("The application of Carson's equation to the steady-state analysis of distribution feeders")

[6] W. Kersting ("The computation of neutral and dirt currents and power losses")

[7] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

[8] Srivallipuranandan ("Series Impedance and Shunt Admittance of an Underground Cable System")

2.2 *Kron's reduction*

By assuming that all buses have a neutral conductor perfectly connected to the ground, one can reduce the size of the impedance matrix from 4 to 3. Indeed, the neutral voltage difference at the beginning and the end of the line

is zero. As explained by Kersting,[9]

[9] William H Kersting (*Distribution System Modeling and Analysis*)

$$\Delta V = Z_4 I \tag{A.1}$$

$$\begin{pmatrix} \Delta V_{abc} \\ \Delta V_n \end{pmatrix} = \begin{pmatrix} Z_{abc} & Z'_{abcn} \\ Z_{abcn} & Z_n \end{pmatrix} \begin{pmatrix} I_{abc} \\ I_n \end{pmatrix} \tag{A.2}$$

$$\Delta V_n = 0 \tag{A.3}$$

$$\Delta V_{abc} = \left(Z_{abc} - Z'_{abcn} Z_n^{-1} Z_{abcn} \right) I_{abc} = Z_3 I_{abc} \tag{A.4}$$

where ΔV represents the voltage differences along the line for the four conductors and I is the current through them. With Kron's reduction, the impedance matrix is split into four submatrices, one related to the phases Z_{abc}, another related to the neutral Z_n, and the last ones linking the two, Z_{abcn}.

Kron's reduction can be applied to both Z_4 and $\hat{Z}_4$, depending on the use of Carson's equations. If all nodes are grounded, Kron's reduction does not modify the simulation results. However, care must be taken if this reduction is used for networks with an ungrounded neutral, as it will lead to erroneous neutral voltages.

2.3 Sequence impedance matrix

Instead of using phase impedance matrices to represent the lines or cables, one could use sequence impedance matrices. If the lines are considered to be transposed regularly, the sequence impedance matrix will be diagonal. However, zero sequence impedance is not commonly given by manufacturers as it depends on the network earthing. The positive sequence impedance is commonly multiplied by a factor between 2.5 and 3.5 for lines without a ground return.[10] However, several papers[11] have shown that approximating the zero sequence impedance with the positive one can lead to large errors in the voltages, justifying the use of phase impedances.

[10] Urquhart ("Accuracy of Low Voltage Distribution Network Modelling")

[11] Frame, Ault, and Huang ("The uncertainties of probabilistic LV network analysis") and W. Kersting and Phillips ("Distribution feeder line models")

2.4 Shunt admittance matrix

Finally, given that the shunt capacitances of LV networks are small, even for underground cables, as in several papers,[12] the admittance matrices will be neglected.

[12] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach"), D. Das, Nagi, and Kothari ("Novel method for solving radial distribution networks"), and Sunderland and Conlon ("4-Wire load flow analysis of a representative urban network incoprating SSEG")

3 Test cases

Four network configurations have been selected to analyse the influence of the earth modelling.

1. No buses are grounded, except the neutral point of the distribution transformer which is perfectly grounded.
2. The neutral of specific buses is grounded through a 0.5-Ω impedance.
3. The neutral of specific buses is perfectly grounded.
4. The neutral of all buses is perfectly grounded.

4 Test system and measurement campaign

The test network used for this study is an existing Belgian LV distribution network with a star configuration 400V/230V. It is composed of one feeder made with underground cables of the EVAVB-F2 3x95 + 1x50 type. Detailed, unbalanced three-phase four-wire modelling of the network has been used according to Ciric[13] based on the data provided by the Distribution System Operator (DSO) (topology, line length, cable type, etc.). The actual configuration of the network is the first one (ungrounded neutral). In configurations 2 and 3, the neutral is assumed to be grounded at the buses marked with a red square in Figure A.1.

[13] Ciric, Feltrin, and L. Ochoa ("Power flow in four-wire distribution networks-general approach")

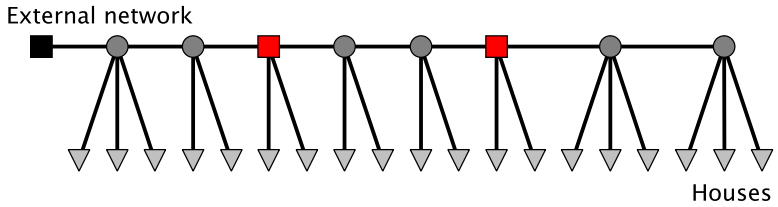


Figure A.1: Graphic representation of the test network.

The network supplies 19 houses, all equipped with a three-phase smart meter. Each smart meter provides phase-to-neutral voltage, current, active and reactive power measurements for each phase, as a one-minute average. The beginning of the feeder is also equipped with a smart meter. It is important to note that the phase identification of all smart meters is known thanks to this algorithm presented in Chapter 2.

5 Results

Simulations have been performed on the test network for each network configuration, on the one hand without using Carson's equations (i.e. using Z_4) and on the other hand using them to represent the earth return path (i.e. using Z_5). A Newton-Raphson algorithm has been applied to solve the power flow equations expressed in rectangular form. The consumption and production profiles of the different houses are the ones from the smart meters, thus reflecting the true operations and imbalance of the network.

The difference between the simulation results from the two earth modelling techniques is assessed with two measures:

1. The mean absolute difference (MeanAD):

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\left|V_i - V_i'\right|}{230} \tag{A.5}$$

2. The maximum absolute difference (MaxAD):

$$\max_i \frac{\left|V_i - V_i'\right|}{230} \tag{A.6}$$

where V_i and V_i' are two voltage time series, and n is their length. All results are represented with box plots, where the samples are the measures of all

time series. For example, since there are 19 three-phase buses, the boxplot will represent the distribution of the mean absolute error for the 84 voltage time series. On each box, the central mark is the median, while the edges of the box are the 25th and 75th percentiles. Outliers are identified and plotted individually, while the whiskers extend to the most extreme data points not considered as outliers.

5.1 Modelling the earth return path

Phase to neutral voltages

As can be seen in Figure A.2, the median of the MeanAD is small, approximately $4 \cdot 10^{-4}$ p. u. As expected, the MeanAD is equal to 0 when the neutral is not grounded. Indeed, no currents flow through the earth, and since the sum of the currents in the four conductors is equal to zero, the earth modelling has no influence on the simulation results. The MeanAD increases incrementally in Configuration 2 where specific buses are grounded through an impedance, 3 where specific buses are perfectly grounded, and 4 where all buses are perfectly grounded, the earth's influence being most prominent in that case.

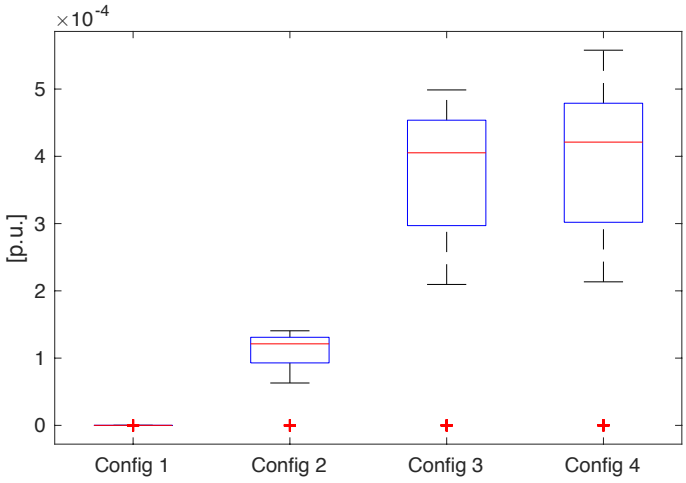


Figure A.2: Boxplot of the mean absolute differences between the phase-to-neutral voltages when the impedance of the earth is not modelled, and when Carson's equations are used.

The same tendencies can be observed for the MaxAD, with median of $2.5 \cdot 10^{-3}$ in Configuration 4, as shown in Figure A.3.

Currents

In the case of phase currents, the MeanAD is taken relatively to the current intensity at the same time, i.e.

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\left|I_i - I_i'\right|}{I_i} \tag{A.7}$$

where I is the magnitude of the current.

As Figure A.4 shows, the differences between the phase currents when Carson's equations are used is negligible, a difference of approximately 0.05

Kron's reduction

Kron's reduction can only be used in Configuration 4 without leading to simulation errors. In that case, simulations using Z_3 and Z_4 are completely equivalent. The same can be said for simulations using $\hat{Z}_3$, $\hat{Z}_4$ and Z_5. The only difference is the computation burden as the size of the network matrix is uselessly larger.

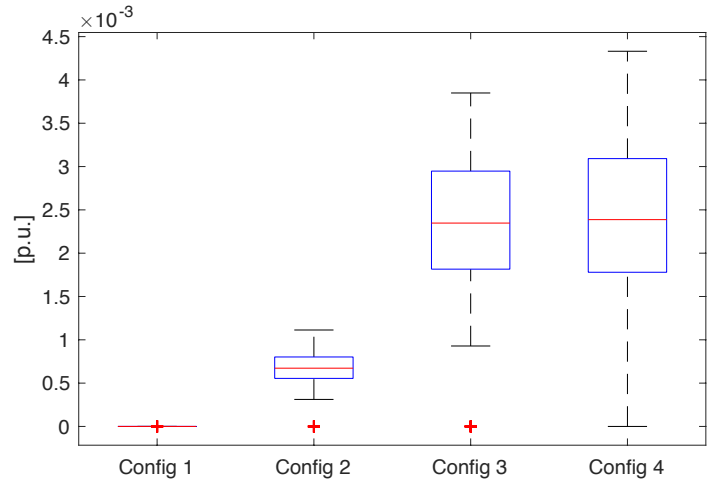


Figure A.3: Boxplot of the maximum absolute differences between the phase-to-neutral voltages when the impedance of the earth is not modelled, and when Carson's equations are used.

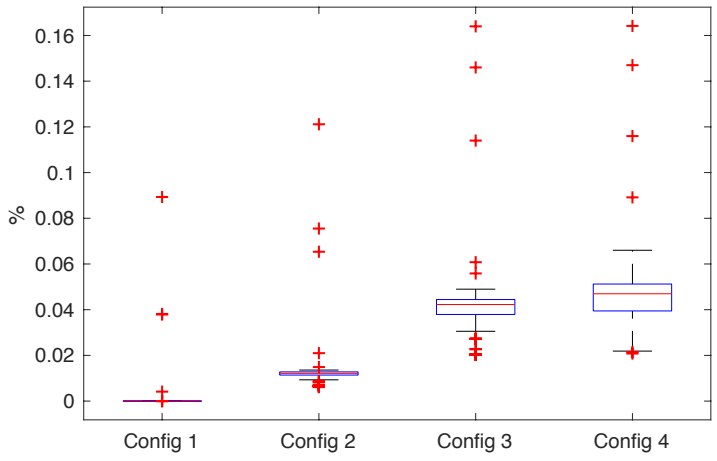


Figure A.4: Boxplot of the mean absolute difference between the currents in the phase conductors. In this case, the MeanAD is taken relatively to the current magnitude.

Main observations from these results

Given the uncertainty in the applicability of Carson's hypotheses to LV networks, the increased computation burden to use 5x5 matrices to represent the lines, and the small value of the MeanAD and MaxAD, we advise the use of 4x4 matrices, and to consider the earth as a single electrical point, i.e. the only ground of the electrical equivalent circuit.

5.2 Comparison between simulations and measurements

Taking into consideration the conclusions from the previous section, Figure A.5 and Figure A.6 show the error between the phase-to-neutral voltage measurements and the voltages simulated in the exact same conditions (i.e. same network topology, same cables, same consumption and production in the houses), when the earth is not modelled with Carson's equations. It is expected that Configuration 1 leads to the smallest error, as the neutral of the test network is ungrounded. With a median of approximately 0.001 p.u., Figure A.5 highlights a reasonable error, indicating that the modelling of the test LV network was done properly. Compared to the previous sections, the MeanAD and MaxAD are approximately ten-times larger, further showing that the impact of the earth modelling is small, and can be neglected in front of the errors between simulations and measurements. These errors can be as a result of: (i) The network topology in the Geographical Information System (GIS) of the DSO does not reflect reality. (ii) There is noise is the measurements and perhaps unmonitored consumption. (iii) The line impedances are calculated theoretically and not measured on site, etc.

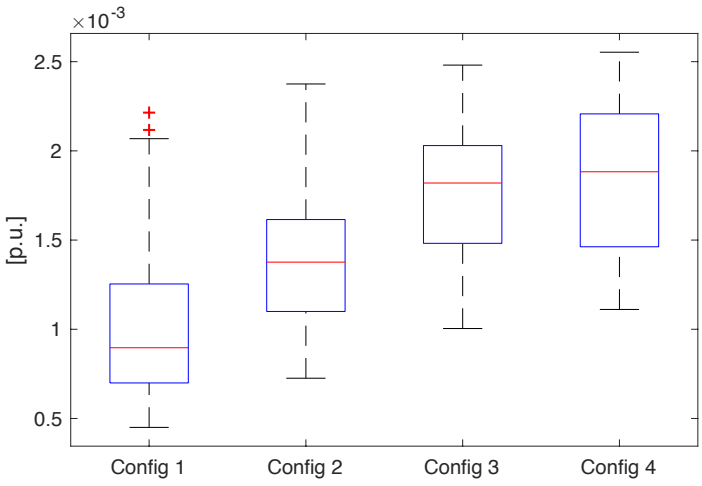


Figure A.5: Boxplot of mean absolute errors between the phase-to-neutral voltages that are simulated and measured.

6 Conclusion

This paper provides guidelines to help modelling LV electrical lines. The first one is that Kron's reduction should only be used when all the buses of the network are perfectly grounded, a case which is rarely met on LV networks. Second, Carson's equations should be used with caution as their hypotheses may not hold in LV networks. For the sake of simplicity, we advise one to consider the earth as a single electrical point in the network to which all the grounded neutral points are connected, with or without a grounding impedance. Indeed, the voltage differences brought by Carson's equations are much smaller than the global error between simulations and measurements.

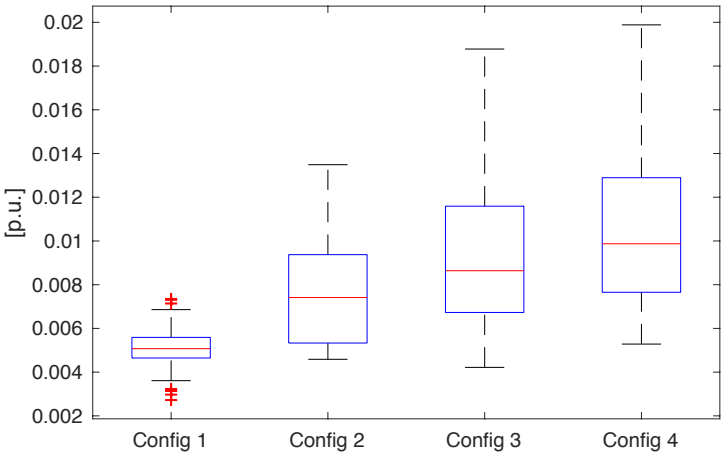


Figure A.6: Boxplot of the maximum absolute error between the phase-to-neutral voltages that are simulated and measured.

Appendix B
Modelling and emulation of an unbalanced LV feeder with photovoltaic inverters

In this paper, the penetration of grid-connected photovoltaic systems is studied, experimentally tested and compared to simulation results. In particular, how the inverse current flow and unbalance situations affect the voltage in the low-voltage grid. Thus, a test platform has been developed for obtaining experimental results with grid-tied commercial inverters. Photovoltaic arrays are emulated and subjected to different irradiance profiles and the inverters are controlled to produce at different power conditions. A model has been developed in order to reproduce the same operating conditions and working environment. Simulations are performed with the software PowerFactory and the results compared to the experimental ones.

1 Introduction

The operation of grid-tied photovoltaic (PV) units is characterized by several uncertainties due to the number of currently operating units, the points where they are connected, and the delivered power. Due to the historical design of low-voltage (LV) feeders, PV generation (PVG) can have adverse effects and cause voltage deviations due to reversed power flow.[1] Some authors[2] have also studied the overvoltage in low-voltage grids based in probabilistic models. In this context, a scheme that controls the active and reactive power of inverters was previously designed and numerically tested.[3] The first step toward the validation of the control scheme in a real setting is to reproduce in a laboratory the behavior of a LV feeder under different voltage and unbalance conditions.

In previous research, the impact of the penetration of PV units in the low-voltage network with dispersed loads was investigated.[4] The former presents modeling and field measurements and addresses the problems of increasing PV installations in these networks. In the latter, a benchmark LV microgrid network is presented, which is suitable for steady state and transient simulations. In validation, laboratory reproduction allows time-, cost- effective and repeatable tests of overvoltage conditions of the grid without need of deploying expensive equipment close to large concentrations of PV systems.

In this paper, we present the test platform (Figure B.1) emulating a LV feeder, and the voltage and power variations for several operation scenarios. The effect of active and reactive power variations on the voltage profile

[1] Barker and Mello ("Determining the impact of distributed generation on power systems. I. Radial distribution systems")

[2] Klonari, Vallee, Durieux, and Lobry ("Probabilistic tool based on smart meters data to evaluate the impact of distributed photovoltaic generation on voltage profiles in low voltage grids") and Klonari, Vallee, Durieux, De Greve, et al. ("Probabilistic modeling of short term fluctuations of photovoltaic power injection for the evaluation of overvoltage risk in low voltage grids")

[3] Olivier, Aristidou, et al. ("Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units")

[4] Papaioannou et al. ("Modeling and field measurements of photovoltaic units connected to LV grid. Study of penetration scenarios") and Papathanassiou, Hatziargyriou, and Strunz ("A benchmark low voltage microgrid network")

is studied. The lab feeder is modeled numerically and simulations are performed for each operation scenarios. The results from the measurements and the simulations are then compared.

2 *Test platform and model in PowerFactory*

A LV feeder is composed of several elements: the main ones are the external grid, the distribution transformer, the cables, the PV systems and the loads which represent houses consumptions. In our study, loads are not considered. Since they would be connected at the same point as the PV systems, they would reduce the power injected into the network and thus decrease the voltage variations, a behavior that we specifically want to exhibit in our simulations. In this section, every part of the test platform in Figure B.1 is described and its lab implementation explained. The network reproduced in the lab is modeled in PowerFactory, a power system analysis software, with the electrical and hardware characteristics of the external grid, cables and inverters to allow a meaningful comparison.

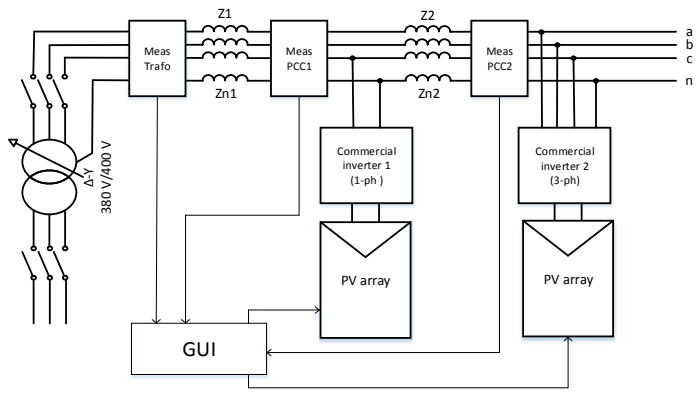


Figure B.1: Laboratory test bench for the study of voltage fluctuations and unbalanced conditions

2.1 *External grid*

In a typical setting, the external grid would be the medium voltage network at a voltage close to 15 kV and the distribution transformer would bring this voltage down to 400 V. In the lab setting, the external grid is the low voltage network. In PowerFactory, it is modeled as a slack bus with a constant voltage.

2.2 *Distribution transformer*

The transformer that feeds the reproduced LV-grid has the following characteristics: Δ-Y configuration, S_{trafo} = 100 kVA, the Primary (380 V, three-phase, delta) and the Secondary (400 V, three-phase, four-wire). In particular, abnormal grid conditions can be emulated with e.g. undervoltage or overvoltage in order to test the behaviour and the compensation features of the inverters. The impedance is estimated considering the standardized short-circuit voltage for distribution transformers (where $\omega = 2 \cdot \pi \cdot f$ is the angular frequency and f = 50 Hz),[5] as follows:

[5] de Metz-Noblat, Dumas, and Poulain (*Cahier technique no 158 – Calculation of short-circuit currents*)

$$V_{CC} = 4\%\ V_{ph-n} = 0.04 \cdot 230V = 9.2\ \text{V} \tag{B.1}$$

Where the short-circuit current is

$$I_{CC} = \frac{S_{trafo}}{\sqrt{3}\ V_{Nom}} = \frac{100\ \text{kVA}}{\sqrt{3}\ 400} = 144\ \text{A}, \tag{B.2}$$

and the value of the reactance and inductance is calculated:

$$X_{L,trafo} = \frac{V_{CC}}{I_{CC}} = 64\ \text{m}\Omega \tag{B.3}$$

$$L_{trafo} = \frac{X_{L,trafo}}{\omega} = 204\ \mu\text{H} \tag{B.4}$$

In PowerFactory, the transformer is modeled considering its short-circuit voltage, copper losses, iron losses and magnetizing impedance.

Table B.1: Cable parameters

Symbol	Description	Value	Units
R	Phase line resistance	0.1	Ω
X	Phase line reactance	65.97	mΩ
R_N	Neutral line resistance	0.165	Ω
X_N	Neutral line reactance	370.7	mΩ

2.3 Cables

The impedance characteristics of the feeder are important as they will determine the actual voltage variations inside the feeder (see section III). The cables are reproduced in the test bench placing impedances between the transformer and the points of connections of the inverters (PCC) as in Figure B.1. The values of the impedances are gathered in Table B.1.

2.4 PV systems

PV array emulator

The PV array emulator allows the reproduction of the characteristics of a standard PV installation in a flexible manner.[6] It uses the single-exponential model of the solar cells[7] with an adjustable number of panels in parallel and series in function of the output characteristics required and takes into account the influence of the in-plane irradiance and the PV cell temperature. The characteristic parameters of the PV panels used for the PV array emulation

[6] Geury and Gyselinck ("Emulation of Photovoltaic arrays with shading effect for testing of grid-connected inverters")

[7] Villalva, Gazoli, and Filho ("Comprehensive approach to modeling and simulation of photovoltaic arrays")

Table B.2: Parameters of one PV panel

Symbol	Description	Value	Units
V_{OC}	Open-circuit voltage	37.3	V
I_{SC}	Short-circuit current	8.52	V
V_{MPP}	MPP voltage	30.5	V
I_{MPP}	MPP current	8.04	V
P_{MPP}	MPP power	245.22	W

Array	N. of panels	Vmax (V)	Peak power (W)
Array 1	11	410.3	2697.42
Array 2	20	746	4904.4

Table B.3: Configuration of the PV arrays

are grouped in Table B.2. The emulator is able to reproduce realistic atmospheric conditions either with the clear-sky model or actual recorded data. In addition, shading can be easily set in the Graphical User Interface (GUI) of the emulator in order to test the Maximum Power Point Tracking (MPPT) capabilities of the PV inverters under these conditions.[8] In particular, shading results in several local maxima on the instantaneous Power-Voltage (P-V) curve of the PV array, which requires an appropriate algorithm for proper MPPT.

[8] Geury and Gyselinck ("Emulation of Photovoltaic arrays with shading effect for testing of grid-connected inverters") and Balasubramanian, Ilango Ganesan, and Chilakapati ("Impact of partial shading on the output power of PV systems under partial shading conditions")

Commercial inverter

The inverter, through which the PV array is connected to the LV grid, is commonly designed to comply with the latest grid standards. For this reason, compensation features to help support the grid are more and more often implemented by the manufacturers. In commercial inverters, this includes mainly power regulation through active power reduction (w.r.t. the default maximum power production) and reactive power compensation. The former is introduced in order to limit the power delivered for a specific section of the grid and the latter to correct locally some power quality issues such as voltage fluctuations.

Depending on the inverter, the reactive power compensation can be set in different ways; it can be adapted to the needs of the system (so-called static or dynamic $\cos\varphi$ setpoint) or to the country grid codes. For example, the German grid codes require the reactive power setpoint to be either fixed or adjustable by a signal from the network operator. The setpoint value is either a fixed displacement factor (static $\cos\varphi$), a variable displacement factor depending on the active power ($\cos\varphi(P)$), a fixed reactive power value in VAr (dynamic $\cos\varphi$) or a variable reactive power depending on the voltage $Q(U)$.[9] These features have been previously tested and the voltage compensation capabilities of the inverters assessed.[10] In order to adjust the power setpoints, it is necessary to allow the communication between the user (e.g. network operator) and the inverter. This is done using the RS485 communication protocol, setting absolute values of the active and reactive powers or the $\cos\varphi$ parameter.

[9] Troester ("New German grid codes for connecting PV systems to the medium voltage power grid")

[10] Lopez-Erauskin et al. ("Testing the enhanced functionalities of commercial PV inverters under realistic atmospheric and abnormal grid conditions")

Symbol	Description	Value	Units
V_{DC-MPP}	DC-voltage MPPT range	350 - 600	V
P_{DC-Max}	Maximum DC input power	3200	W
V_{AC}	AC rms voltage (ph-N)	230	V
P_{AC}	Nominal AC power	2600	VA
$\cos\varphi$	Power factor	-0.8..1..0.8	-

Table B.4: Parameters of the single-phase inverter

Symbol	Description	Value	Units
V_{DC-MPP}	DC-voltage MPPT range	245 - 800	V
P_{DC-Max}	Maximum DC input power	5150	W
V_{AC}	AC rms voltage (ph-ph)	400	V
P_{AC}	Nominal AC power	5000	VA
$\cos\varphi$	Power factor	-0.8..1..0.8	-

Table B.5: Parameters of the three-phase inverter

The characteristics such as the configuration of the PV arrays and the peak power for the single- and three-phase inverters are specified in Table B.3, B.4 and B.5.

2.5 The Graphical User Interface

ControlDesk software is used together with the dSPACE ds1104 platform for the GUI. The user can observe the relevant system variables, such as the grid voltages and currents, and the DC-side voltage and current. Also, the instantaneous characteristic curves of the PV array for the adjustable meteorological conditions set are displayed, so that the evolution of the working point can be observed. This is especially useful for evaluating the MPPT capability of the inverter.

3 Voltage fluctuation

The validation of the results obtained in the simulation are of relevant importance in order to ensure the reliability of a simulation model. For that purpose, the power exchange between the PVGs and the LV feeder will be tested here to see how it affects to the grid voltage. The relation between power exchange and voltage fluctuation is discussed hereafter.[11]

As shown in Figure B.1, any current flow will generate a voltage drop and a phase shift between two arbitrary points on the feeder. In LV feeders, to which distributed generation units are commonly connected, the inductive and resistive components have to be considered.[12] Considering no load connected at any PCC, the complex power flowing through that section is the one coming from the PV inverters: $S_{inv} = P_{inv} + jQ_{inv}$. The voltage at the PCC1 is here considered as reference with $V_1 = V_1\angle 0°$ while the one at the transformer $V_{tr} = V_{tr}\angle\delta$ and the grid current $I = I\angle\varphi$ are phase-shifted by angles δ and φ, respectively. The complex power at the PCC1 is, therefore, the sum of the powers delivered by the inverters and expressed as $S_{PCC-1} = P_{PCC-1} + jQ_{PCC-1}$.

In unbalanced condition for this LV feeder configuration, the existing neutral impedance displaces the neutral voltage V_{N1} from the one at the transformer (V_N). Represented in Figure B.2, V_A, V_B and V_C (in dark blue) are the line voltages at PCC1. The single-phase inverter is connected between phase C and the neutral point, where the inverter voltage is in phase with V_C (Figure B.3).

The voltage drop at the impedance between the transformer and the CC1 is:

[11] Mastromauro, Liserre, and Dell'Aquila ("Single-phase grid-connected photovoltaic systems with power quality conditioner functionality")

[12] Mastromauro, Liserre, and Dell'Aquila ("Single-phase grid-connected photovoltaic systems with power quality conditioner functionality")

$$\underline{\Delta V}_1 = \underline{Z}_1 \cdot \underline{I} \tag{B.5}$$

considering the feeder impedance $\underline{Z}_1 = R_1 + jX_1$. The phase A voltage at the transformer is, therefore:

$$\underline{V}_{AN} = \underline{V}_A - \underline{\Delta V}_1 \tag{B.6}$$

and its neutral point voltage:

$$\underline{V}_N = \underline{V}_{N1} + \underline{\Delta V}_{LN} \tag{B.7}$$

where $\underline{V}_{N1}$ is the neutral voltage at PCC1 and $\underline{\Delta V}_{LN}$ is the voltage drop at the neutral impedance Z_{N1}, caused by the current $\underline{I}_N$ flowing through the neutral line:

$$\underline{\Delta V}_{LN} = \underline{I}_N \cdot \underline{Z}_{N1} \tag{B.8}$$

considering the neutral impedance $\underline{Z}_{N1} = R_{N1} + jX_{N1}$. The neutral voltage at PCC-1 satisfies the following expression:

$$\underline{V}_{N1} = \underline{V}_N - \underline{\Delta V}_{LN} \tag{B.9}$$

and the phase A to neutral voltage at PCC-1 is:

$$\underline{V}_{AN1} = \underline{V}_A - \underline{V}_{N1} \tag{B.10}$$

The displacement of $\underline{V}_{N1}$ from the neutral point of the transformer $\underline{V}_N$ causes a decrease in the magnitude of the vector in phase A and the resulting effect on the voltages (in light blue) at PCC1.

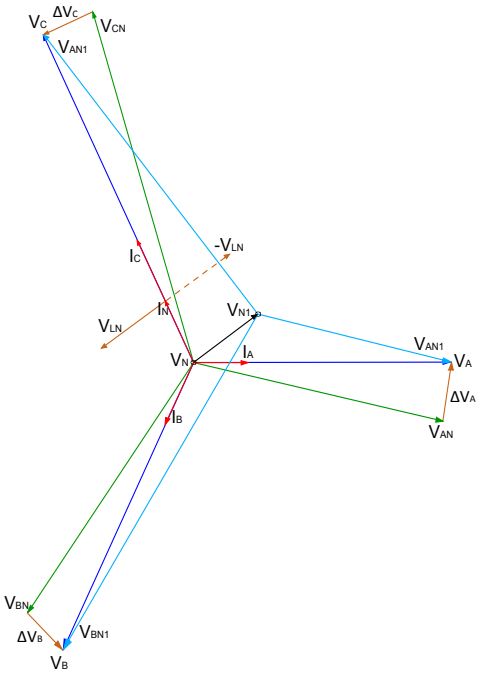


Figure B.2: Voltage phasors of the LV feeder at PCC-1

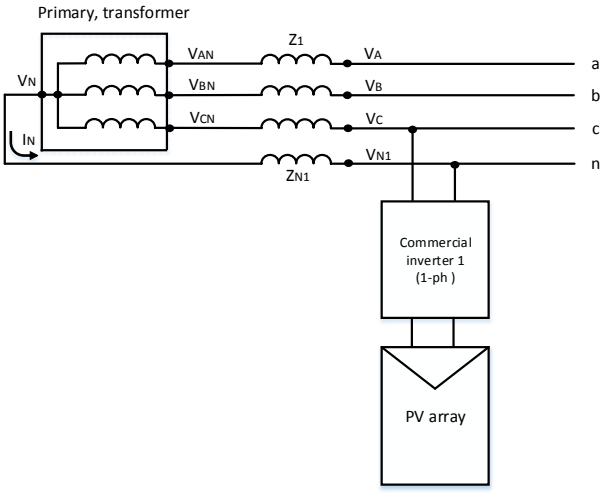


Figure B.3: Part of the laboratory test platform and the neutral current flow direction

4 *Results and comparison*

The testing is done for single- and three-phase inverters, focusing on their power regulation functionality and the resulting voltage compensation capability. The general parameters of the platform are shown in Table B.1.

The simulations are performed in two different phases:

4.1 Static production values

The different operating scenarios are:

1. Operating scenario 1 (OP1): all PV inverters produces their maximum active power.
2. Operating scenario 2 (OP2): all PV inverters produce half their maximum power.
3. Operating scenario 3 (OP3): all PV inverters produce half their maximum active power and absorb the maximum reactive power. This last operating scenario allows us to observe the influence of reactive power on the voltages at the LV level.

	Single-phase inverter		Three-phase inverter	
	P (W)	Q (var)	P (W)	Q (var)
OP1	2317	0	4810	0
OP2	1200	0	2460	0
OP3	1200	-1700	2460	-2500

Table B.6: Active and reactive power produced by the inverters in each scenario

The three OPs show the voltage fluctuation in a LV feeder in presence of dispersed generation, where the consumption is null. The goal of these simulations is to put in evidence the phenomenon by which a overvoltage can occur. In that regard, power consumption was not considered as it would

reduce the power injected in the different PCCs and diminish the effect of power on voltage.

Figs. B.4-B.6 show the voltage profile evolution for simulation (dashed lines) and experimental (solid lines) results for the different scenarios and the impact of PVG and reactive power on voltage.

OP1

In Figure B.4, all the inverters are producing at their maximum power, without reactive power compensation. The figure shows the influence of the injected power on the grid voltage for each PCC. The furthest the PVG is from the LV feeder transformer, the highest is the voltage level at the PCC for V_{bn} and V_{cn}. In this situation, the generators linked to it are the first to disconnect from the grid if an overvoltage occurs. However, the same does not apply to V_{an}, affected by the displacement of the neutral voltage V_{N1} from the neutral point of the transformer V_N, as explained in section III.

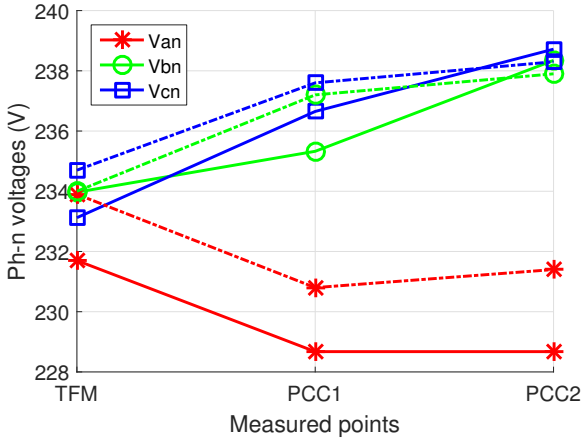


Figure B.4: Voltage profile of the LV feeder, PV inverters producing at full power. Simulation (dashed lines) and experimental (solid lines) results.

OP2

In Figure B.5 the inverters are working at half their maximum active power with a lower influence in the voltage increment.

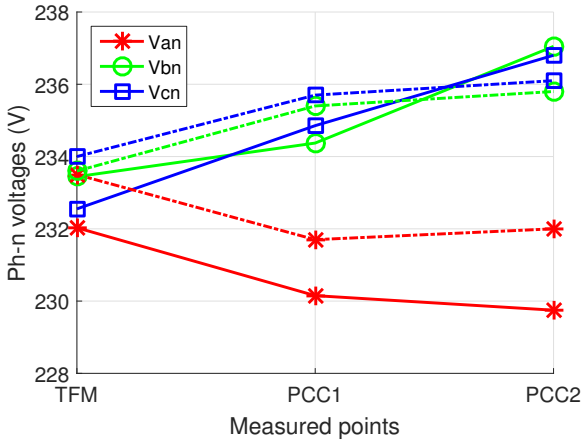


Figure B.5: Voltage profile of the LV feeder, PV inverters producing at half maximum power. Simulation (dashed lines) and experimental (solid lines) results.

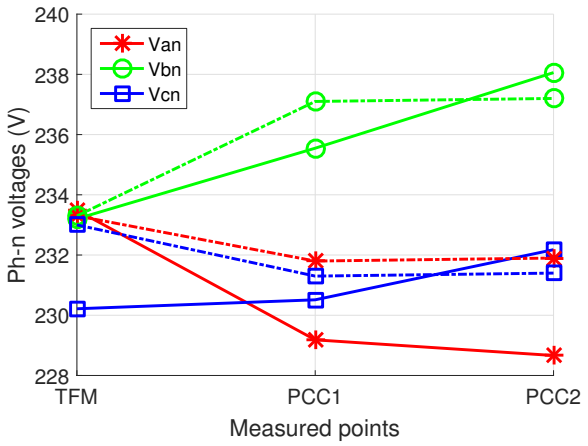


Figure B.6: Voltage profile of the LV feeder, PV inverters producing at half maximum power and absorbing maximum reactive power. Simulation (dashed lines) and experimental (solid lines) results.

OP3

For this scenario, the voltage at phase C decreases more notably than at the others due to the two inverters' reactive power consumption.

According to (B.5)-(B.10), the voltage drop (ΔV) at the line impedance depends on the feeder characteristics, the direction of the current and the amount of this current. The behavior of the voltage profile along the different points of the feeder is, in most of the cases, the same for both simulation and experimental results. The differences between the numerical results and the measured ones (e.g. measured V_{an} = 231.7 V and simulated V_{an} = 233.9 V in Figure B.4) are due to the voltage differences in the real transformer, not reproducible in software reliably. However, these divergences need to be studied more deeply to bring the simulations and the real environment measurements closer.

Table B.7 gathers the values of Figs. B.4-B.6 for the three different operation scenarios: OP1, OP2, and OP3, where the added letter "L" means "Laboratory" and "S" "Simulation".

Table B.7: Voltages measured at the LV feeder

Operation scenario	OP1L	OP1S	OP2L	OP2S	OP3L	OP3S
$V_{an,Tr}$ (V)	231.7	233.9	232	233.5	233.5	233.3
$V_{bn,Tr}$ (V)	233.9	234	233.4	233.6	233.2	233.3
$V_{cn,Tr}$ (V)	233.1	234.7	232.5	234	230.2	233
$V_{an,PCC1}$ (V)	228.7	230.8	230	231.7	229	231.8
$V_{bn,PCC1}$ (V)	235.3	237.2	234.4	235.4	235.5	237.1
$V_{cn,PCC1}$ (V)	236.7	237.6	234.9	235.7	230.5	231.3
$V_{an,PCC2}$ (V)	228.7	231	229.8	232	228.7	231.8
$V_{bn,PCC2}$ (V)	238.4	237.9	237	235.8	238.1	237.2
$V_{cn,PCC2}$ (V)	238.7	238.3	236.8	236.1	232.2	231.4

4.2 Dynamic production values

The PV array emulators follow predefined irradiance profiles to simulate realistic behavior of the inverters during a day. The irradiance profile used for

the simulations is the one of a sunny day in Belgium and its acquisition time is 1 minute. The variation of the irradiance is reproduced in the laboratory every 2 seconds, accelerating the execution 30 times. This way, faster experimental results than in field measurements can be obtained, as can be seen in Figure B.8. The variation of voltages, active and reactive power flows are studied and compared to the numerical results to validate them thanks to the real hardware emulation.

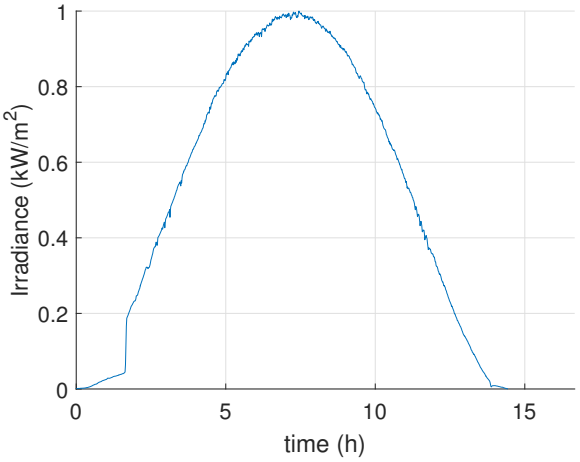


Figure B.7: Irradiation profile during a sunny day

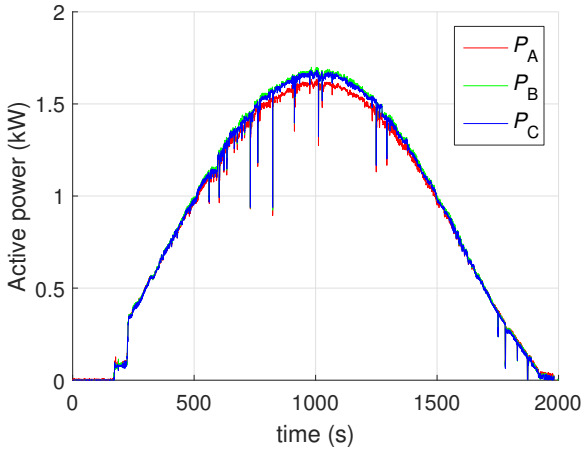


Figure B.8: Time variation of the active power at PCC2 during a sunny day

Figure B.9 shows the evolution of the output voltages of the inverters, as well as the active and reactive power flows that corresponds to the irradiance profile for a sunny day (Figure B.7). The output behavior of the inverters is analyzed in the PowerFactory model, as for the static production values, with the same predefined parameters as in the test platform. The results obtained also present a similar evolution, although the voltage differences between phases cannot be reproduced with PowerFactory due to an small initial unbalance in the test platform.

5 *Conclusion*

In this manuscript, an unbalanced LV feeder caused by the grid-tied inverters has been studied for different values of active and reactive power production. A test bench that reproduces an specific LV feeder has been designed for this

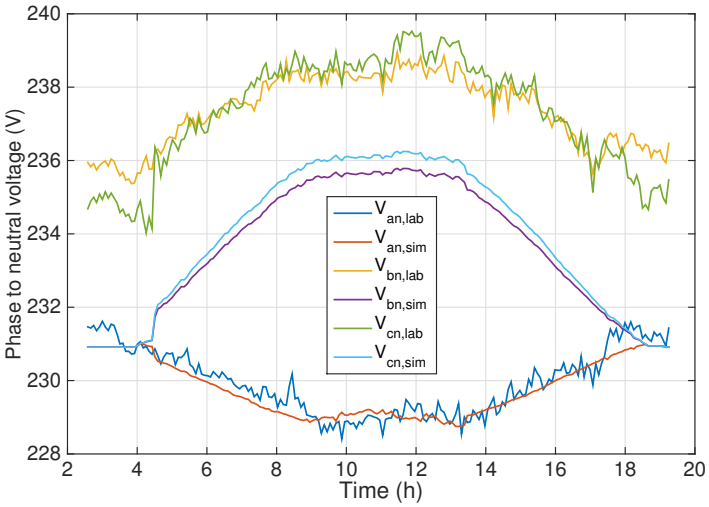


Figure B.9: Time variation of the phase-to-neutral voltages at PCC2 during a sunny day

study and its behavior compared to its numerical model. This comparison has illustrated the reproduction of a similar behavior between the experimental work and the simulated model but also the difficulties of obtaining reliable results in simulations, due to the lack of information for some of the parameters of the system. In addition, the effect of local neutral point displacement has been exhibited and explained. It changes the shape of the phase-to-neutral voltages in magnitude and phase and can aggravate the unbalance situation.

Appendix C
Effect of voltage constraints on the exchange of flexibility services in distribution networks

Many possibilities exist to organise exchanges of flexibility within a distribution system. In this paper, we call such a possibility an interaction model. The DSIMA (Distribution System Interaction Model Analysis) testbed allows one to compare quantitatively candidate interaction models by simulating a distribution system with actors taking decisions to maximise their own profit or minimise their costs. The original testbed focused on establishing the procedures to exchange information between actors and used a network flow model considering only active power. This paper extends DSIMA with a linear approximation of the power flow equations, the line limits and the voltage constraints. This linear flow model is compared to a network flow model by simulating three different interaction models governing the exchange of flexibility services within a Belgian distribution system. Results show that changing the network model may significantly impact the quantitative results obtained from the simulations.

1 Introduction

Distribution systems were typically sized to supply the empirically observed peak consumption, and to require few preventive or corrective control actions, by making the assumptions that the grid users were only *consuming* power in a *non-coordinated* way. However, the development of distributed generation, renewable or not, and demand-side management is likely to lead to many issues in operation by making those assumptions less realistic. Two main options are available to overcome these issues: network reinforcement and *active network management*.[1] The latter involves increasing the efficiency of distribution systems operating the system using all control means available, and the flexibility of the grid users, which can be offered as flexibility services taking the form of modulations with respect to their intended production or consumption.

[1] Gemine et al. ("Active network management: planning under uncertainty for exploiting load modulation")

This change of practices in distribution systems calls for a revision of the *interaction model*, that is, the set of rules guiding the interactions between all the parties of the system. A new interaction model is usually chosen after a discussion between affected actors, based on qualitative analyses making simplistic assumptions. Using only qualitative analyses can lead to unexpected behaviours of the actors and a lower social welfare. To avoid such

consequences, one should rely on quantitative analyses, for example with agent-based modelling,[2] to determine the financial impact of changing the interaction model. Since actors are modelled as individuals, every exchange of information and financial transaction has to be modelled explicitly. Therefore, implementing a given interaction model into an agent-based model provides insights on the difficulties to implement the interaction model from a technical perspective, and highlights the necessary exchanges of information that a qualitative analysis could have missed.

[2] Tesfatsion and Judd (*Handbook of Computational Economics: Agent-Based Computational Economics*)

To compare alternative interaction models relying on active network management, an agent-based system has been devised and published as an open-source testbed: DSIMA (Distribution System Interaction Model Analysis).[3] Within this framework, each agent optimises its individual objective following the rules guided by the interaction model. This article extends the original network flow model used in optimisation problems solved by the Distribution System Operator (DSO), with a flow model that considers reactive power flow and voltage constraints. In order to keep the simulation computationally tractable, the updated flow model is a linear version of the classic optimal power flow (OPF) formulation based on Bolognani's work,[4] which provides a linear relation between the voltages and the injected power at each bus, and bounds the approximation error. Simulations are performed with DSIMA to quantify the effect of the flow models under different interaction models.

[3] Mathieu, Louveaux, et al. ("DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks")

[4] Bolognani and Zampieri ("On the existence and linear approximation of the power flow solution in power distribution networks")

The paper is organised as follows. The literature relevant to the main methodological addition to our previous work[5] is reviewed in Section 2. Section 3 describes the original DSIMA testbed in more detail and presents the evaluated interaction models. The linear flow model and its implementation are proposed in Sections 4 and 5. Section 6 discusses the simulation results, and Section 7 concludes.

[5] Mathieu, Louveaux, et al. ("DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks")

2 *Literature review*

The linearised direct current (DC) power flow approximation is often used to solve large-scale systems[6] but is not well suited to represent distribution networks since the model does not accurately capture under- or over-voltages. Alternative linear models have been proposed to overcome this issue. For instance, the power flow equations can be linearly approximated by considering their perturbations with respect to deviations from references of active and reactive power injection in each bus.[7] Another linear approximation is proposed by Coffrin[8] and compared to the DC power flow approximation on experimental test systems. On a more theoretical perspective, Bolognani[9] provides a linear relation between the voltages and the injected power at each bus, and bound the approximation error. The latter formulation is used in our analysis to model the power flows in the distribution networks.

[6] Stott, Jardim, and Alsaç ("DC power flow revisited")

[7] A. Farag, Al-Baiyat, and T. Cheng ("Economic load dispatch multiobjective optimization procedures using linear programming techniques")

[8] Coffrin and Van Hentenryck ("A linear-programming approximation of AC power flows")

[9] Bolognani and Zampieri ("On the existence and linear approximation of the power flow solution in power distribution networks")

3 *Definition of the interaction models and their implementation in DSIMA*

This section recalls the main elements of the system modelled in the open-source testbed DSIMA.[10] We consider the short-term flexibility exchanges

[10] For more details, see Mathieu, Louveaux, et al. ("DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks")

in an operational planning phase of a medium-voltage network. The actors simulated are producers and retailers, which are called Grid Users (GUs), the DSO and the Transmission System Operator (TSO). A grid user is responsible of its imbalance and may use or provide flexibility services to other grid users as well as the DSO and TSO.

3.1 Access bounds and access contracts

We assume that each GU has an access contract with the DSO for each bus it has access to. This contract specifies a time-invariant full-access range, in which the GU can produce or consume without any restriction. Each GU first requests a specific access range to the DSO for each bus it wants to access. Based on that, the DSO computes the safe access range, which is lower or equal to the requested range and ensures that no congestion occurs if every GU accesses to the network within these limits. One method to perform this computation is explained in our previous work.[11]

[11] Mathieu, Louveaux, et al. ("DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks")

Three access types are considered, respectively named: *restricted*, *unrestricted* and *dynamic*. In the first access type, the GU production or consumption must always remain within the safe access range, i.e. its full access range is equal to the safe access range. For the second, no restrictions are applied to GUs, and the DSO relies on flexibility services to operate its network. Stakeholders are free to exchange flexibility services among them, and in particular with the DSO. For the third, the DSO can restrict the access of the GU to a dynamic range, which changes for each time period, e.g. 15 minutes, and constrained to be larger than the full access range.

3.2 Interaction models

The interaction models implemented in DSIMA follow the procedure below and summarised in Figure C.1, where some steps may be skipped. (i) GUs provide baseline proposals. (ii) Based on these, the DSO computes the dynamic ranges so that its network is secure and communicates them to the GUs. (iii) After the clearing of the day-ahead energy market, GUs submit their (new) baselines, constrained to lie within the dynamic ranges. They are used as references for the provision of flexibility services. If the realisation of a GU violates the dynamic ranges, the GU is penalised at a regulated tariff. (iv) The DSO assesses the state of the system and they and other GU announces their flexibility needs. (v) GUs provide flexibility offers sequentially to the other actors, the first being the DSO. (vi) GUs request and/or buy some proposed flexibility offers. (vii) Closer to real-time, activation requests of flexibility services are communicated to the GUs. (viii) Right before real-time, each GU optimises its realisation, taking into account the requests. (ix) The distribution network is operated using these realisations and takes last resort actions if necessary, such as shedding buses. If such actions are needed and GUs did not provide their service to the DSO, they are penalised at a regulated price.

Given the previous definitions, this paper focuses on three interaction models:

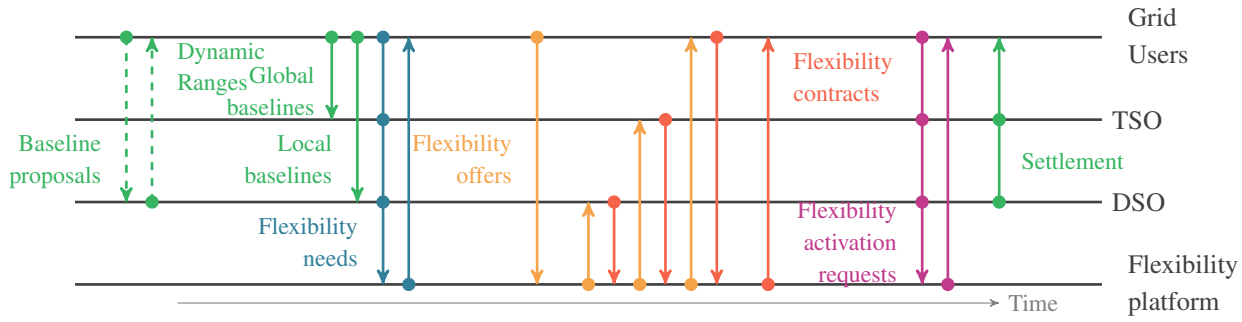


Figure C.1: Interactions between the agents where each type of actor is represented by a horizontal line. Each vertical arrow represents an interaction between two types of actor.

- *Model 0*, where the access type is restricted, and there is no financial compensation for flexibility,
- *Model 1*, where the access type is dynamic, and the DSO does not use flexibility to operate its network.
- *Model 2*, where the access is unrestricted and the DSO pays a reservation and activation cost and pays for the resulting imbalance caused by the activation of flexibility services.

The analysis could be extended to others, such as the remaining three models presented in previous work.[12]

[12] Mathieu, Louveaux, et al. ("DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks") and Mathieu, Ernst, and Cornélusse ("Agent-based analysis of dynamic access ranges to the distribution network")

4 *Linear power flow*

The traditional power flow equations use the admittance of the lines to evaluate the voltages and obtain the losses. They can be written as:

$\forall n \in \mathcal{N}, n' \in \mathcal{N}(n), t \in \mathcal{T}$,

$$s_{n,t} = \sum_{n' \in \mathcal{N}(n)} S_{n,n',t} \tag{C.1a}$$

$$V_n^{min} \leq ||V_n||_2 \leq V_n^{max} \tag{C.1b}$$

$$\bar{S}_{n,n',t} = \bar{V}_{n,t} Y_{n,n'} (V_{n,t} - V_{n',t}) \tag{C.1c}$$

$$||S_{n,n',t}||_2 \leq C_{n,n'} \tag{C.1d}$$

where $\mathcal{N}$ is the set of buses, $\mathcal{N}(n)$ the set of buses connected to bus n, $\mathcal{T}$ the set of time periods, t the time period, $V_{n,t} = e_{n,t} + jf_{n,t}$ the complex voltage at bus n, $S_{n,n',t} = p_{n,n',t} + jq_{n,n',t}$ the complex power flowing from bus n to bus n', and $s_{n,t}$ the complex power injection at bus n. The conjugate of a complex number x is denoted $\bar{x}$ and its two-norm $||x||_2$. The main complexity of this formulation originates from the nonlinearity of the quadratic term in the equality constraint (C.1c) and the concave minimum bound in (C.1b).

In this study, (C.1c) is approximated using the Bolognani's work.[13] The authors show that the voltages in each buses, in a column vector, can be approximated by

[13] Bolognani and Zampieri ("On the existence and linear approximation of the power flow solution in power distribution networks")

$$\hat{\mathbf{V}}_t \simeq V_0 \left(\mathbb{1} + \frac{1}{||V_0||_2^2} Z \bar{\mathbf{S}}_t \right) \tag{C.2}$$

where $\mathbf{S}_t$ is the column vector with the complex power injections and V_0 is the complex voltage of the slack bus. The matrix Z is defined as the inverse of the nodal admittance matrix and $\mathbb{1}$ is a column vector of ones. With our

notations, by developing (C.2) and neglecting the shunt admittances, one can obtain:

$$\bar{S}_{n,n',t} = \bar{V}_0 Y_{n,n'}(V_{n,t} - V_{n',t}) \tag{C.3}$$

The necessary data are the resistance and reactance of all lines, and the active and reactive power injections at each bus. Note that this linear approximation yields $S_{n,n',t} = -S_{n',n,t}$ and therefore the power flows are conserved. As a result, no losses occur in the system and only the voltage differences are approximated.

As shown in Figure C.2, a convex circled-shape constraint can be approximated by a conservative feasible set of cuts. This approximation can be used for the constraint on the flow in each line and the constraint on the maximum voltage. If I is the number of cut points by quadrant of the complex plane, $\mathcal{I} = \{1, \ldots, I\}$ and $\Phi(I)$ is an ordered set with the corresponding angles:

$$\Phi(I) = \left\{0, \frac{\pi}{2I}, \frac{2\pi}{2I}, \frac{3\pi}{2I}, \ldots, \frac{\pi}{2}\right\}, \tag{C.4}$$

the set of constraints defining the boundary in the north-east quadrant of the complex plane of $S_{n,n',t} = p_{n,n',t} + jq_{n,n',t}$ is given by, $\forall i \in \mathcal{I}$

$$A_i p_{n,n',t} + B_i q_{n,n',t} \leq A_i C_{n,n'}, \tag{C.5}$$

where $A_i = \sin\Phi_i - \sin\Phi_{i-1}$, $B_i = \cos\Phi_i - \cos\Phi_{i-1}$.

The constraints bounding the three other quadrants can be obtained by changing the signs of the $p_{n,n',t}$ and $q_{n,n',t}$ terms in (C.5). These constraints are too conservative and the maximal error margin can be easily computed by trigonometry:

$$\frac{1}{2}\sqrt{\left(\cos\frac{\pi}{2I} + 1\right)^2 + \sin^2\frac{\pi}{2I}} \tag{C.6}$$

which for $I = 4$ leads to a conservative error of 1.92%.

The same method may be used to approximate the bound on the maximum voltage amplitude, but not the one on the minimum voltage amplitude, which is concave. This work chooses to conservatively approximate the feasible domain of the voltage with the constraint $e_{n,t} \geq V_n^{min}$, as shown in Figure C.2. This approximation is acceptable for angle deviations that are relatively small, which is often the case in practice. Denoting $\theta_{n,t}$ the angle of the voltage of the bus n where $e_{n,t} = V_n^{min}$, the conservative error margin is given by $1/\cos\theta_{n,t}$ (e.g. 3.5% when $\theta_{n,t} = 15°$).

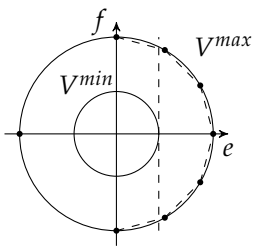


Figure C.2: Linear approximation of the constraint $V^{min} \leq ||e + jf||_2 \leq V^{max}$ with three cut points.

5 *Integration in the DSO operations*

In DSIMA, the linear power flow model is used in all optimisation problems solved to define the behaviour of the DSO, from the definition of the full ac-

cess ranges to the real-time operation of the network. This section describes how the testbed is updated so that the DSO uses the linear flow models.

Since a design choice of the testbed is to never obtain an infeasible optimisation problem in order to always have results to analyse, the voltage constraints are converted into soft constraints, where $\zeta_{n,t}$ and $\nu_{n,t}$ are slack variables respectively on the minimum and maximum voltage bound. These slack needs to be penalised in the objective function of the DSO's optimisation problems and this cost is added to the shedding cost to form a *protections cost*.

The complete linear flow model is given by (C.7).
$\forall n \in \mathcal{N}, n' \in \mathcal{N}(n), t \in \mathcal{T}, i \in \mathcal{I}$,

$$s_{n,t} = \sum_{n' \in \mathcal{N}(n)} S_{n,n',t} \tag{C.7a}$$

$$S_{n,n',t} = p_{n,n',t} + jq_{n,n',t} \tag{C.7b}$$

$$\bar{S}_{n,n',t} = \bar{V}_0 Y_{n,n'}(V_{n,t} - V_{n',t}) \tag{C.7c}$$

$$A_i p_{n,n',t} + B_i q_{n,n',t} \leq A_i C_{n,n'} \tag{C.7d}$$

$$-A_i p_{n,n',t} + B_i q_{n,n',t} \leq A_i C_{n,n'} \tag{C.7e}$$

$$A_i p_{n,n',t} - B_i q_{n,n',t} \leq A_i C_{n,n'} \tag{C.7f}$$

$$-A_i p_{n,n',t} - B_i q_{n,n',t} \leq A_i C_{n,n'} \tag{C.7g}$$

$$A_i e_{n,t} + B_i f_{n,t} \leq A_i (V_n^{max} + \nu_{n,t}) \tag{C.7h}$$

$$A_i e_{n,t} - B_i f_{n,t} \leq A_i (V_n^{max} + \nu_{n,t}) \tag{C.7i}$$

$$V_n^{min} - \zeta_{n,t} \leq e_{n,t} \tag{C.7j}$$

$$\zeta_{n,t}; \nu_{n,t} \geq 0 \tag{C.7k}$$

To run its optimisation problems, the DSO assumes constant power factor in order to estimate injections of reactive power.

6 Test system and results

6.1 Test system

The interaction models are evaluated with the network flow (*Netflow*) and the linear power flow (*LinOPF*) model. The test system is called Ylpic and originates from a real distribution system that is operated by ORES, a Belgian DSO. It was slightly modified to anonymise sensitive information and to update generation and consumption devices to illustrate what the grid could face in 2020.[14] This test system contains 328 MV buses in a radial configuration. They are simulated under the same conditions over 12 days, divided in 24 periods and simulated independently. These days are selected to be representative of one year,[15] based on the time series of aggregated active and reactive production and consumption of the distribution system. To simulate the tap changer of the HV/MV transformer, the voltage at the MV side is considered constant. To some extent, other voltage regulators could help to alleviate voltage violations, but their presence in the test network is very limited. In order to keep the computations tractable, the discrete behaviour of capacitor banks is not optimised. Simulating the twelve days on one interaction model takes respectively 13 and 14 hours of simulation time using the

[14] Cornélusse et al. ("A process to address electricity distribution sector challenges: the GREDOR project approach")

[15] Mathieu (*DaysXtractor*) and Poncelet et al. ("Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems")

network flow model and the linear power flow model. The experiments are run on a computer equipped with an Intel Core i7 at 3.40GHz with 32GB of RAM, and optimisation problems are solved with SCIP.

To aggregate the large amount of results obtained from simulating a complex system, a simple measure of the welfare is defined as the sum of the benefits and costs of each actor with their sign, and a protection cost. The latter includes the penalty for under and over-voltages as well as the cost of shedding, due to the tripping of a protection (e.g. a selector or a circuit breaker resulting in the complete disconnection of the components behind the protection). This last part adds a notion of quality of service in the welfare, which directly impacts the final consumer of electricity. With this definition, one way to increase the welfare is to increase the amount of renewable generation that is produced within the distribution network, because it directly increases the benefits of the corresponding producers. Moreover, an interaction model, which does not ensure the safe operation of the network, will cause protections to trip and decrease the welfare.

6.2 Results

Figure C.3 shows the computed welfare of each interaction model, which are all equal when the network flow model is used and hence will not be further analysed in this paper. Integrating voltage constraints highlights the weaknesses of Model 0 and 2, and thus clearly motivates the use of the linear power flow model to assess the performance of interaction models. The detailed contributions to the welfare are given in Table C.1. The linear power flow model decreases by 20% the welfare associated with Model 0. This is directly explained by the same decrease in renewable production. In two of the representative days, congestion of ten lines trigger protections leading to consumption shedding. These two days are problematic in all the interaction models, implying that the flexibility available within the system is not sufficient to solve the issue. Therefore, either more flexibility should be made available, or the network should be reinforced, for example, by increasing the capacity of the ten lines identified by this simulation (in red in Figure C.4).

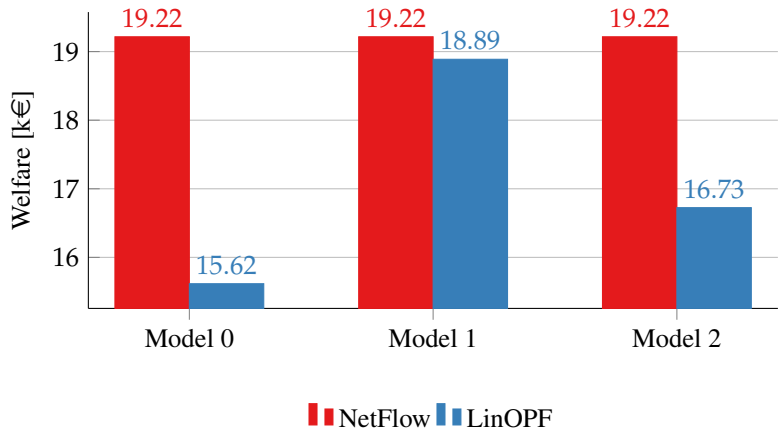


Figure C.3: Welfare of the three interaction models with the network flow model and the linear optimal power flow model.

The welfare increase of Model 1–2 with respect to Model 0 can be first explained by the 26% increase in the renewable energy production. The welfare of Model 2 is significantly less than the others due to its protection costs,

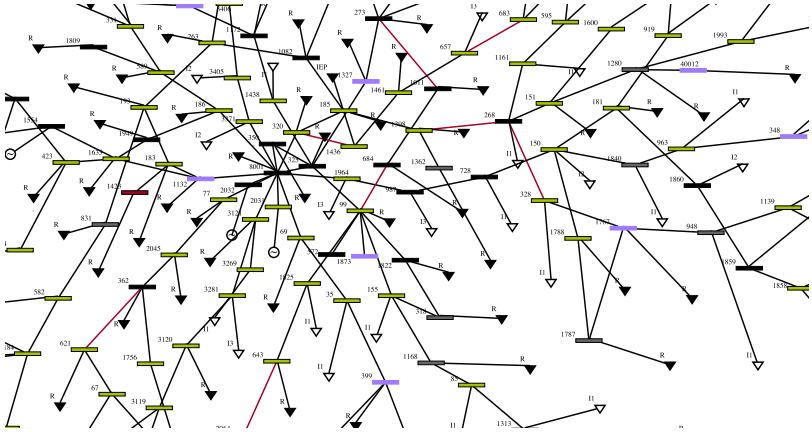


Figure C.4: Summary of the state of one part the Ylpic network for one critical 15-minute period in Model 0. Lines in red are congested. Buses in green contain flexibility activation. The ones in purple suffer under voltages, while the ones in red are shed in real time.

coming from consumption shedding in the fourth representative day during the evening consumption peak. Indeed, Model 2 is based on unrestricted access to the network, and therefore does not force GU to provide flexibility services to the DSO. As a result, too much consumption flows throughout the network and the DSO does not have access to enough flexibility to alleviate the congestion of the lines. Some congestion results from flexibility activation from the TSO within the distribution network. This lack of coordination between the DSO and TSO is solved in Model 1 since grid users were restricted, preventing them from proposing their service to the TSO.[16] Previous works have shown that Model 1 outclasses the other tested interaction models using a network flow representation. These results which consider voltage constraints strengthen the latter conclusion.

[16] For more details, see Mathieu, Ernst, and Cornélusse ("Agent-based analysis of dynamic access ranges to the distribution network").

Table C.1: Impact of the flow model on the interaction models for the Ylpic system.

	Model 0		**Model 1**		**Model 2**		
	Netflow	*LinOPF*	*Netflow*	*LinOPF*	*Netflow*	*LinOPF*	
Welfare	19216	15616	19216	18888	19216	16725	€
Protections cost	0	318	0	283	0	2393	€
DSOs costs	0	0	0	0	0	0	€
TSOs surplus	-2349	-2492	-2349	-2376	-2349	-2349	€
Producers surplus	19010.6	15891.8	19010.6	19007.7	19010.6	19059.7	€
Retailers surplus	2554.06	2534.31	2554.06	2535.32	2554.06	2407.57	€
Total production	633.14	501.1	633.14	624.22	633.14	629.05	MWh
Total consumption	-1142.74	-1141.4	-1142.74	-1141.53	-1142.74	-1133.44	MWh
Total imbalance	0	0.32	0	0.28	0	1.58	MWh
Max. imbalance	0	0.03	0	0.03	0	0.29	MW
Total usage of flex.	29.63	26.82	29.63	27.42	29.63	29.63	MWh
Total energy shed	0	-0.32	0	-0.28	0	-1.3	MWh

7 Conclusion

This paper extends the DSIMA testbed with a linear approximation of the power flows, integrating reactive powers and voltage constraints. It is based on Bologanani'swork[17] which provides a linear relation between the voltages and the injected power at each bus. This linear flow model is compared to the network flow model by simulating three different interaction models governing the exchange of flexibility services within a realistic distribution system. Results show that changing the network does impact the quantitative results obtained from the simulations when voltage constraint are binding. Indeed, using a network flow model provides similar welfares, whereas using a linear power flow model allows one to decide which interaction model to choose. The simulations suggest that the operation of the tested distribution network should rely on dynamic access ranges.

This work can be extended along several lines. One could refine the modelling of the reactive power injections by developing a dedicated prediction for the DSO and add uncertainty to these injections in its optimisation problems. The exchange of reactive power flexibility services would be worth implementing to assess its impact on the network operation costs. Finally, new interaction models that target the reactive power support could be quantitatively assessed.

[17] Bolognani and Zampieri ("On the existence and linear approximation of the power flow solution in power distribution networks")

Appendix D
An app-based algorithmic approach for harvesting local and renewable energy using electric vehicles

The emergence of electric vehicles (EVs), combined with the rise of renewable energy production capacities, will strongly impact the way electricity is produced, distributed and consumed in the very near future. In this paper, we focus on the problem of dispatching a fleet of EVs in order to match at best the harvesting of local, renewable energy. The combination of (mobile) EVs and (static) electricity sources is modeled as a multi-agent system, where mobile agents (the EVs) should gather at best renewable energy under travel distances and electricity production fluctuations constraints. This position paper exposes how a mobile application may offer an efficient solution for addressing this problem. Such an app would play two main roles: Firstly, it would incite and help people to play a more active role in the energy sector by allowing PV owners to sell their electrical production directly to consumers. Secondly, it would help distribution system operators (DSO) to modulate more efficiently the load due to the EV deployment by allowing them to influence EV charging behavior in real time. Finally, the present paper advocates for the introduction of a two-sided market-type model between EV drivers and electricity producers.

1 Introduction

The past decade has seen a steep rise in the development of renewable energy production capacities, mainly driven by the willingness to (i) reduce pollution and greenhouse gas effects emissions and (ii) limite the dependency on fossil fuels. In addition to this, electric vehicles (EVs) are now very likely to emerge. This dual emergence may be an opportunity to set the basis of a join optimization approach for charging EVs at best regarding the fluctuation of renewable energy production. This has already been the topic of academic research in the past decade.[1] In particular, it has been shown that EVs may be efficiently used in order to balance the load in within distribution networks.[2] O'Connell et al.[3] show how vehicle charging can be influenced by dynamically adapt tariffs. Even though these results are very important, none of them gave a convenient way to apply them to real situations.

In this paper, we focus on the subproblem of optimizing the charge of a fleet of EVs in order to match at best a set of local and renewable sources of electricity. In this context, we argue that a mobile application would be an

[1] For a comprehensive review, see Palensky et al. ("Modeling intelligent energy systems: co-simulation platform for validating flexible-demand EV charging management")

[2] Caramanis and Foster ("Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion")

[3] O'Connell et al. ("Electric vehicle (EV) charging management with dynamic distribution system tariff")

efficient and elegant way to dispatch EVs to electricity sources. We first describe how such an app should be designed. Then, based on recent scientific developments, we explain why such an app would be a suitable solution to address the problem of both charging EVs and managing the fluctuations of renewables. Finally, we conclude with possible future works that could be derived from the present work.

sectionAPP DESCRIPTION The app, as an interface between EV drivers and the electrical network, must be useful to both of them to be successful and fruitful. The app lays on a shared economy model.

1.1 Two-sided Market

One main role of the app is to create a convenient way for PV owners to share (and sell) their electricity to EV drivers. Indeed, the electrical network has not been designed to absorb electricity production from private houses, so, during PV production peak, panels are turned off to avoid overload on the network. Neither has the electric network been designed to supply an amount of power sufficient to fully recharge fleet of EV in one night in a neighborhood.

Consequently, on one side we have PV owners that will may want to find a way to profitably exploit their PV panels, and on the other side EV driver who may be constrained to charge their vehicles while they are occupied by other tasks. The app is aimed at satisfying both requirements and make it easy for each side to benefit of the other needs. Here under, we show how this can be accomplished and we demonstrate the feasibility of the solution.

1.2 Driver Services

To correctly carry out this work, the app must be attractive to EV drivers. For this purpose, a booking service will be provided to guarantee a certain charge for the drivers. Moreover, the app intelligence will be able to propose a flexible and suitable choice of charging point depending on the user needs.

One way to implement a convenient proposition service would be to split journey type into two categories. The drivers would have the choice between one-way ride or A-B-C ride. In the one-way option, the app would compute the optimal stations at which the driver should stop to go from place A to B, knowing the level of charge, consumption and autonomy of the vehicle. However, if the user wants to be more foreseeing, he could choose the second option. In this case, the app would also take into account the time spent at location B, during which a possible recharge could be made, and the return trip to optimize the total journey. Taking all these parameters into account, it would give the drivers the best stations to stop during the day. Obviously, stations proposed should also be chosen to globally optimize a function, for example to minimize the energy curtailed.

As the app tries also to benefit the network by balancing the load, it could propose non-optimal solutions for certain users. A reward for EV drivers providing their car to balance load could therefore be used to compensate this non-optimality. It could consist in money but also in other advantages like free charge or booking priority among the other users of the app.

1.3 Producer Services

Producers can be divided into two categories, professional producers and the rest. Professional producers can either be large renewable energy source owners, DSOs or big charging point owners whereas the rest consists mainly of private owners of PV panels, representing a power of less than 10kWc, who want to sell their electricity.

As small producers can only furnish a small amount of power, it is not interesting for car owners to recharge directly at this type of charging point. Indeed, the amount of energy produced in one day couldn't exceed 20kWh for the biggest installations and would be more usually around 10kWh. Keeping in mind that the consumption of an electric car is more or less 20kWh/100km, the energy possibly recharged wouldn't benefit the driver. A configuration that would be more interesting for both PV owners and EV drivers would be a unique charging point in a neighborhood. As wanted, this solution would help to avoid curtailment and would be an attractive charging point for EV drivers. This type of structure would mitigate curtailment while bringing acceptable power to recharge EV.

If we consider this configuration for small producers, there is no reason to keep the division previously cited and we can deploy a unique set of services. This services would include booking system, detailled in a subsequent section, but also data based analysis to manage prices and attractiveness of the charging points. More precisely, the app will allow to efficiently retrieve data about those stations. From this data, analysis about how pricing can influence the attractiveness of charging points could be extracted. Finally, based on these analysis, the app should play a role of managing tool for producers, who will also be able to practice a more accurate dynamic pricing to adapt demand to production. Finally, other advantages like use of smaller batteries due to supply and demand adequation will indirectly result of the app.

2 Main Strength

Although a mobile application is a simple concept, it may surprisingly be a game changer in the world of electricity production and consumption, mainly because of its capacity to make people (and in particular, consumers) react quickly.

2.1 Booking Service

In the short-term, congestion problems are going to take place near charging points as a complete charge take at least 30 minutes with superchargers, and more generally many hours with classic charger. As a result, we can see that knowing if a station is free or not will probably be one of the main problems for EV drivers. For now, no tangible solution has emerged on the emerging market of charging points although most of the EV drivers have already experienced this disagreement.

The app's booking service is a solution to this problem. The booking will work as follow. Firstly, stations will only be bookable 24-hours in advance, mainly for a fairness reason between EVs. Indeed, if someone could book a station for an entire year, this would lead to a lack of attractiveness of the

app for new users. Secondly, the app intelligence will be responsible for organizing a fair distribution of charging stations among users to avoid the same type of problem. Indeed, the user that is far-sighted will have a larger choice of stations but the app must arrange reservations in a way to let free attractive stations for less far-sighted users. Actually, there are two reasons for doing that. On the one hand, the app must be attractive and thus it has to provide each type of user a good solution. On the other hand, one of the missions of the app is to bring flexibility in demand, so the purpose is to keep some hand on the demand as long as possible.

We propose two solutions to keep flexibility on demand while allowing chargers booking.

- The first solution is to play with electricity pricing. By choosing the charging station suggested by the app just before the charge, the driver gives some service to the network and thus could be rewarded for that service by a reduction of the price of the charge. On the opposite the early bird driver is constraining the production (or the network) to take account of his charge, thus this user could pay for this demand.
- The second solution is based on blockchain technology.[4] In this second solution, the booking transactions could be stored in blocks which would only be accepted after a certain time allowing more flexibility in terms of charging station allocation. Moreover, smart contracts[5] could be use to determine the conditions of the transactions without passing by a third party, adding more dynamism and autonomy to the app.

[4] Nakamoto (*Bitcoin: a peer-to-peer electronic cash system*)

[5] Smart contracts (also called self-executing contracts, blockchain contracts, or digital contracts) are simply computer programs that act as agreements where the terms of the agreement can be coded in advance with the ability to self-execute and self-enforce itself. This code defines the rules and consequences in the same way that a traditional legal document would, stating the obligations, benefits and penalties which may be due to either party in various different circumstances. This code can then be automatically executed by a distributed ledger system.

A problem which rapidly comes to mind when exposing the idea of proposing a global booking service is the difficulty to control whether bookings are respected. Indeed, except if you own every single station that you propose to drivers, you cannot guarantee that a station booked by one driver will not be occupied by another driver that is not using the app. The solution we propose amounts in creating a global booking service, which wouldn't be part of the app as it should be available for other app creators. It would retain all the booking data about each station. Station owners would just have to carry on this tool to manage the booking around their stations. This tool could probably have some specific features and propose some additional services but this is not the subject we want to explore here.

2.2 Load Management and Dynamic Pricing

Managing the load is probably the greatest innovation that may be offered by the app. By maintaining a certain flexibility on the booking, the app allows to better locally correlate consumption with production by making EVs charge at local energy sources. A dynamic pricing system[6] could allow the app to boost the demand in order to follow more precisely the production. Moreover, with smart chargers (that may automatically change the amount of power they output) and good weather forecasts, we can hope to exactly fit the (over)production to the EV charging by allocating each vehicle to an appropriate station for a precise amount of time.

[6] O'Connell et al. ("Electric vehicle (EV) charging management with dynamic distribution system tariff")

To allow such a load management, a convenient dynamic pricing system must be implemented. It therefore has to allow producers to keep hand on

the demand simply by raising or lowering charging prices. For instance, they could lower prices at overproduction times to rise up demand and do the opposite when domestic consumption is at a peak. This will all be done through the app that should help charging point owners to understand the influence of these pricing choices. Indeed, the app should provide a forecast on the demand in function of the price chosen for one charging point.

2.3 Data Mining

Another strength of the app is the flow of data passing through it. This data consists mainly of information about drivers trips (travel time, distance, ...) or the performance of EV vehicles (mean consumption, charge level variations, ...). More generally, we can summarize this information as when, where, at which frequency do EV drivers charge their cars and how does it influence the vehicles' performances, the vehicles fleet movements and the network load. This information is susceptible to interest a number of actors in the automobile and energetic communities. Here is a short list of some of the actors and how they could use the app's data.

- **Drivers:** The drivers themselves would be the first to benefit from this data through a self-learning improvement of the app's dispatching strategy based on previous results.
- **Car manufacturers:** More practically, EV real-time performances could help manufacturers identify key points for future improvements.
- **Producers:** Forecast on the density of traffic at different locations, at different times could help producers to apply an appropriate dynamic pricing systems.
- **Electricity Networks:** These same forecasts combined to renewable energy production predictions can be used by DSO to find methods to balance the load and prevent overloads. These methods can take two opposite points of view: either, we can influence the drivers to change their behavior by playing on the charging prices or proposing them some incentives or, we can move the load towards these density spots in order to maximize the energy distributed.

In a word, the app has a true database potential and could be of interest for various actors. Moreover, as explained here under, this database could be commercialize and become one of the main source of income of the app.

2.4 Shared Economy Model

EV drivers can already use a panel of apps that can help them to find information about charging stations that they could possibly use (PlugShare®, ChargeBump®, NextCharge®). However, all of them work with a one-view driver-oriented system. An advantage of our app would be its multi-view organization trying to profit to every actors concerned. This view, supported by a shared economy model, would benefit drivers by optimizing their travel and charging time but also the producers by allowing them to implement dynamic pricing method to modulate the attractiveness of their product.

In addition, shared economy models are becoming more and more popular and are literally revolutionizing every branch of capitalism.[7] The success of Uber®or Airbnb®prove that people are eager for a change of economic model. Easy adequation of supply and demand, efficiency, low cost, fast scaling are all different reasons why these models are so successful and are all properties that would be beneficial for an EV driver app. As a matter of fact, it would seem more convenient nowadays to directly book a station to its owner than having to solicit the intervention of an external organization.

[7] Rifkin ("Capitalism is making way for the age of free")

3 Incentives

The choice of an app as a platform for the services that were described before is driven by several incentives, financial as well as practical.

3.1 Economical Solution

The creation and maintenance of an app are operations that come at very low cost (range from 1000$ to 100,000$ depending on the type of app). These costs are negligible if we compare them to the price of alternative solutions like batteries. Indeed, if the app was not used, batteries could be used to balance the overload on the network while allowing PV owners not to curtail their electricity production. However, installing enough batteries to cover the risks of curtailment in a country like Belgium would cost around 1,000,000,000$. In addition, even if to the costs of the app you may have to add the creation of a database and its maintenance, they are still quite low compared to these costs (500GB of cloud for one year costs approximately 1000$ on Google Cloud or 1 PB costs around 1 million$ on Hadoop according to Forbes[8]).

[8] See http://www.forbes.com/sites/ciocentral/2012/04/16/the-big-cost-of-big-data

Moreover, the investment in such a database is an opportunity to commercialize its content. Indeed, as mentioned before, data over drivers' habits, how prices influences their choices and how it influences the network in a whole are of great interests for several electrical market actors and can generate large incomes as concluded in the abovementioned Forbes article. Other techniques can be used to make some profits. For instance, commissions can be taken from the payments from drivers to producers or another income could consist in special in-apps purchases offering advantages to users.

3.2 Permittivity

Another advantage of an app is its fast flexibility due to the absence of heavy infrastructure that usually slow down the launch of a new product. Indeed, once the program is finished, the product can directly be launched on the market. We could therefore imagine that after 3-6 months of development, the system would be operational. This is a crucial feature since the EV market is likely going to skyrocket in the coming years: the app market will probably also get matured accordingly.

3.3 Reactivity

Nowadays, as the number of EVs and charging stations increase, there is a growing need for a real time information platform. An app is currently one of the most reactive platform and a true answer to this kind of problem. In particular, the booking can be easily and dynamically managed by connecting together a series of drivers and producers's phones. More than a connection between people, it would also offer a direct connection to smart electric cars, collecting immediately all the data needed to predict when and where the next charge should be done. Moreover, the app would open the door to more connected payment method like virtual money or phone payment which would add some more fluidity in the system.

Finally, an app adds a layer of reactivity over all these features due to its proximity to people. More than 2 billions people have a smartphone[9] and they accompany us wherever we go. Therefore, apps are also always with us allowing them to inform us of any important update like price modifications, overproduction risks or reservation problems directly when they happen so that we react in a minimum amount of time.

[9] according to the website statista.com: http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

4 Conclusion and Future Work

This article shows how an app may be an appropriate solution for tackling several upcoming obstacles associated with the rise of EVs and renewable energy production capacities, combined with the management of the electrical network load. This app may be seen as an interface between drivers and producers, providing services to both of them and having a series of assets for making it an appropriate solution. This paper also details some incentives to prove that such an app may also be economically viable.

Some additional points, not detailed in this paper, may be investigated in future works. For example, we have briefly mentioned the idea of developing a booking platform that would be dissociated from the app. This alternative solution would allow a broader collaboration between competing apps allowing to optimize the booking system and, at the same time, also providing the platform itself with a technical advantage over competitors. Methods to ensure that booking choices made by drivers are respected and the sanctions that should be applied or inversely the rewards given to "good" drivers or producers remain to be defined. Finally, paying methods, using maybe the blockchain system, should be thoroughly considered in order to find a convenient option adaptable to different cases.

Bibliography

Agalgaonkar, Yashodhan P., Bikash C. Pal, and Rabih A. Jabr. "Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators". In: *IEEE Transactions on Power Systems* 29.1 (2014), pp. 182–192.

Aghatehrani, R. and A. Golnas. "Reactive power control of photovoltaic systems based on the voltage sensitivity analysis". In: *Proc. of IEEE PES 2012 General Meeting*. 2012.

Ahn, Changsun and Huei Peng. "Decentralized voltage control to minimize distribution power loss of microgrids". In: *IEEE Transactions on Smart Grid* 4.3 (2013), pp. 1297–1304.

Alam, M. J. E., K. M. Muttaqi, and D. Sutanto. "Mitigation of rooftop solar PV impacts and evening peak support by managing available capacity of distributed energy storage systems". In: *IEEE Transactions on Power Systems* 28.4 (2013), pp. 3874–3884.

Alam, M. J.E., K. M. Muttaqi, and D. Sutanto. "A three-phase power flow approach for integrated 3-wire MV and 4-wire multigrounded LV networks with rooftop solar PV". In: *IEEE Transactions on Power Systems* 28.2 (2013), pp. 1728–1737.

Aristidou, Petros, Davide Fabozzi, and Thierry Van Cutsem. "Dynamic simulation of large-scale power systems using a parallel schur-complement-based decomposition method". In: *IEEE Transactions on Parallel and Distributed Systems* 25 (2013), pp. 2561–2570.

Balasubramanian, Indu Rani, Saravana Ilango Ganesan, and Nagamani Chilakapati. "Impact of partial shading on the output power of PV systems under partial shading conditions". In: *IET Power Electronics* 7.3 (2014), pp. 657–666.

Barker, PP and R.W. De Mello. "Determining the impact of distributed generation on power systems. I. Radial distribution systems". In: *Proc. of IEEE PES 2000 Summer Meeting*. Vol. 3. 2000, pp. 1645–1656.

Bazilian, Morgan et al. "Re-considering the economics of photovoltaic power". In: *Renewable Energy* 53 (2013), pp. 329–338.

Beitollahi, Hakem and Geert Deconinck. "Peer-to-peer networks applied to power grid". In: *Proc. of the International conference on Risks and Security of Internet and Systems (CRiSIS)*. 2007.

Bletterie, B et al. "Understanding the effects of unsymmetrical infeed on the voltage rise for the design of suitable voltage control algorithms with PV". In: *Proc. of the 26th European Photovoltaic Solar Energy Conference and Exhibition*. Hamburg, 2011.

Bollen, M.H.J. and A. Sannino. “Voltage control with inverter-based distributed generation”. In: *IEEE Transactions on Power Delivery* 20.1 (2005), pp. 519–520.

Bolognani, S. and S. Zampieri. “On the existence and linear approximation of the power flow solution in power distribution networks”. In: *Power Systems, IEEE Transactions on* 99 (2015).

Bonfiglio, Andrea et al. “Optimal control and operation of grid-connected photovoltaic production units for voltage support in medium-voltage networks”. In: *IEEE Transactions on Sustainable Energy* 5.1 (2014), pp. 254–263.

Bremdal, Bernt A. et al. “Creating a local energy market”. In: *Proc. of the International Conference on Electricity Distribution (CIRED)*. 2017.

Bruno, Sergio et al. “Unbalanced three-phase optimal power flow for smart grids”. In: *IEEE Transactions on Industrial Electronics* 58.10 (2011), pp. 4504–4513.

Cagnano, A. et al. “Online optimal reactive power control strategy of PV inverters”. In: *IEEE Transactions on Industrial Electronics* 58.10 (2011), pp. 4549–4558.

Calderaro, Vito et al. “Optimal decentralized voltage control for distribution systems with inverter-based distributed generators”. In: *IEEE Transactions on Power Systems* 29.1 (2014), pp. 230–241.

Caldon, R., M. Coppo, and R. Turri. “Coordinated voltage control in MV and LV distribution networks with inverter-interfaced users”. In: *Proc. of IEEE PES 2013 Grenoble PowerTech Conference*. 2013.

– “Distributed voltage control strategy for LV networks with inverter-interfaced generators”. In: *Electric Power Systems Research* 107 (2014), pp. 85–92.

Caramanis, Michael and Justin M Foster. “Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion”. In: *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*. Ieee. 2009, pp. 4717–4722.

Carcelle, X. *Power Line Communications in Practice*. Artech House Telecommunications Library. Artech House, Incorporated, 2009.

Carson, John R. “Wave propagation in overhead wires with ground return”. In: *Bell System Technical Journal* 5.4 (1926), pp. 539–554.

Carvalho, P.M.S., P.F. Correia, and L.A.F. Ferreira. “Distributed reactive power generation control for voltage rise mitigation in distribution networks”. In: *IEEE Transactions on Power Systems* 23.2 (2008), pp. 766–772.

CEI. *CEI 0–21: Reference technical rules for connecting active and passive users to networks low-voltage electricity networks of energy providers*. 2011.

Cheng, C.S. and D. Shirmohammadi. “A three-phase power flow method for real-time distribution system analysis”. In: *IEEE Transactions on Power Systems* 10.2 (1995), pp. 671–679.

Chua, K. H. et al. “Energy storage system for mitigating voltage unbalance on low-voltage networks with photovoltaic systems”. In: *IEEE Transactions on Power Delivery* 27.4 (2012), pp. 1783–1790.

Ciric, R.M., A.P. Feltrin, and L.F. Ochoa. “Power flow in four-wire distribution networks-general approach”. In: *IEEE Transactions on Power Systems* 18.4 (2003), pp. 1283–1290.

Coffrin, Carleton and Pascal Van Hentenryck. “A linear-programming approximation of AC power flows”. In: *INFORMS Journal on Computing* 26.4 (2014), pp. 718–734.

Cornélusse, Bertrand et al. “A process to address electricity distribution sector challenges: the GREDOR project approach”. In: *Proceedings of the International Conference on Electricity Distribution, CIRED 2015*. 2015.

Cundeva, S., M. Bollen, and D. Schwanz. “Hosting capacity of the grid for wind generators set by voltage magnitude and distortion levels”. In: *Proc. of 2016 Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower)*. 2016.

Dall’Anese, Emiliano, Georgios B. Giannakis, and Bruce F. Wollenberg. “Optimization of unbalanced power distribution networks via semidefinite relaxation”. In: *Proc. of the 2012 North American Power Symposium (NAPS)*. 2012.

Das, D., H.S. Nagi, and D.P. Kothari. “Novel method for solving radial distribution networks”. In: *IEE Proceedings - Generation, Transmission and Distribution* 141.4 (1994), p. 291.

de Metz-Noblat, B., F. Dumas, and C. Poulain. *Cahier technique no 158 – Calculation of short-circuit currents*. 2005.

Degroote, Lieven et al. “Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks”. English. In: *Proc. of 2009 IEEE Bucharest PowerTech*. IEEE, 2009.

Demirok, Erhan et al. “Local reactive power control methods for overvoltage prevention of distributed solar inverters in low-voltage grids”. In: *IEEE Journal of Photovoltaics* 1.2 (2011), pp. 174–182.

Dubey, Anamika, Surya Santoso, and Arindam Maitra. “Understanding photovoltaic hosting capacity of distribution circuits”. In: *Proc. of 2015 IEEE Power & Energy Society General Meeting*. 2015.

Dugan, R., W. Sunderman, and B. Seal. “Advanced inverter controls for distributed resources”. English. In: *Proc. of CIRED 2013 22nd International Conference and Exhibition on Electricity Distribution*. 2013, pp. 1221–1221.

EPIA. *Connecting the Sun: Solar photovoltaics on the road to large-scale grid integration*. 2012.

Farag, Ahmed, Samir Al-Baiyat, and TC Cheng. “Economic load dispatch multiobjective optimization procedures using linear programming techniques”. In: *Power Systems, IEEE Transactions on* 10.2 (1995), pp. 731–738.

Farag, Hany E., Ehab F. El-Saadany, and Ravi Seethapathy. “A two ways communication-based bistributed control for voltage regulation in smart distribution feeders”. In: *IEEE Transactions on Smart Grid* 3.1 (2012), pp. 271–281.

Farivar, M. et al. “Optimal inverter VAR control in distribution systems with high PV penetration”. In: *Proc. of IEEE PES 2012 General Meeting*. 2012.

Fortenbacher, Philipp, Martin Zellner, and Goran Andersson. “Optimal sizing and placement of distributed storage in low voltage networks”. In: *Proc. of the Power Systems Computation Conference (PSCC)*. 2016.

Frame, D. F., G. W. Ault, and S. Huang. “The uncertainties of probabilistic LV network analysis”. In: *Proc. of 2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012.

Gabash, Aouss and Pu Li. “Active-reactive optimal power flow in distribution networks with embedded generation and battery storage”. In: *IEEE Transactions on Power Systems* 27.4 (2012), pp. 2026–2035.

Galli, Stefano, Anna Scaglione, and Zhifang Wang. “Power line communications and the smart grid”. English. In: *Proc. of IEEE 2010 First International Conference on Smart Grid Communications*. 2010, pp. 303–308.

Garcia, P.A.N. et al. “Three-phase power flow calculations using the current injection method”. In: *IEEE Transactions on Power Systems* 15.2 (2000), pp. 508–514.

Garcia-Gracia, Miguel et al. “Integrated control technique for compliance of solar photovoltaic installation grid codes”. In: *IEEE Transactions on Energy Conversion* 27.3 (2012), pp. 792–798.

Gemine, Quentin et al. “Active network management: planning under uncertainty for exploiting load modulation”. In: *2013 Bulk Power System Dynamics and Control (IREP) Symposium*. IEEE. 2013, pp. 1–9.

Geurts, Pierre, Damien Ernst, and Louis Wehenkel. “Extremely randomized trees”. In: *Machine learning* 63.1 (2006), pp. 3–42.

Geury, Thomas and Johan Gyselinck. “Emulation of Photovoltaic arrays with shading effect for testing of grid-connected inverters”. In: *2013 15th European Conference on Power Electronics and Applications, EPE 2013*. IEEE, 2013, pp. 1–9.

Glavic, Mevludin, Raphaël Fonteneau, and Damien Ernst. “Reinforcement learning for electric power system decision and control: past considerations and perspectives”. In: *Proc. of the 20th World Congress of the International Federation of Automatic Control*. 2017.

IEEE. *IEEE 1547 standard for interconnecting distributed resources with electric power systems IEEE Std 1547-2003*. 2003.

Kakimoto, Naoto, Qin-Yun Piao, and Hiroo Ito. “Voltage control of photovoltaic generator in combination with series reactor”. In: *IEEE Transactions on Sustainable Energy* 2.4 (2011), pp. 374–382.

Kersting, W. H. and R. K. Green. “The application of Carson’s equation to the steady-state analysis of distribution feeders”. In: *Proc. of 2011 PES Power Systems Conference and Exposition*. 2011.

Kersting, W.H. “The computation of neutral and dirt currents and power losses”. In: *Proc. of 2004 PES Power Systems Conference and Exposition*. IEEE, 2004.

Kersting, W.H. and W.H. Phillips. “Distribution feeder line models”. In: *IEEE Transactions on Industry Applications* 31.4 (1995), pp. 715–720.

Kersting, William H. *Distribution System Modeling and Analysis*. Boca Raton, 2006.

Kiprakis, A.E. and A.R. Wallace. “Maximising energy capture from distributed generators in weak networks”. In: *IEE Proceedings - Generation, Transmission and Distribution* 151.5 (2004), pp. 611–618.

Klonari, V., F. Vallee, O. Durieux, Z. De Greve, et al. “Probabilistic modeling of short term fluctuations of photovoltaic power injection for the evaluation of overvoltage risk in low voltage grids”. In: *Proc. of IEEE International Energy Conference (ENERGYCON)*. 2014, pp. 897–903.

Klonari, V., F. Vallee, O. Durieux, and J. Lobry. “Probabilistic tool based on smart meters data to evaluate the impact of distributed photovoltaic generation on voltage profiles in low voltage grids”. In: *Proc. of 2013 Solar Integraton Workshop*. 2013.

Kotsampopoulos, Panos et al. “Review, analysis and recommendations on recent guidelines for the provision of ancillary services by distributed Generation”. English. In: *Proc. of IEEE 2013 International Workshop on Inteligent Energy Systems (IWIES)*. 2013, pp. 185–190.

Liew, S.N. and G. Strbac. “Maximising penetration of wind generation in existing distribution networks”. In: *IEE Proceedings – Generation, Transmission and Distribution* 149.3 (2002), pp. 256–262.

Lloyd, Stuart. “Least squares quantization in PCM”. In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 129–137.

Lopes, J.A. Peças et al. “Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities”. In: *Electric Power Systems Research* 77.9 (2007), pp. 1189–1203.

Lopez-Erauskin, Ramon et al. “Testing the enhanced functionalities of commercial PV inverters under realistic atmospheric and abnormal grid conditions”. In: *Solar Integraton Workshop*. 2015.

Malekpour, Ahmad Reza, Anil Pahwa, and Sanjoy Das. “Inverter-based var control in low voltage distribution systems with rooftop solar PV”. In: *2013 North American Power Symposium (NAPS)*. IEEE, 2013, pp. 1–5.

Masters, C.L. “Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines”. English. In: *Power Engineering Journal* 16.1 (2002), pp. 5–12.

Mastromauro, R A, M Liserre, and A Dell’Aquila. “Single-phase grid-connected photovoltaic systems with power quality conditioner functionality”. In: *Power Electronics and Applications, 2007 European Conference on*. IEEE. 2007, pp. 1–11.

Mathieu, Sébastien. *DaysXtractor*. 2017.

Mathieu, Sébastien, Damien Ernst, and Bertrand Cornélusse. “Agent-based analysis of dynamic access ranges to the distribution network”. In: *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2016 6th IEEE/PES*. IEEE, 2016.

Mathieu, Sébastien, Quentin Louveaux, et al. “DSIMA: a testbed for the quantitative analysis of interaction models within distribution networks”. In: *Sustainable Energy, Grids and Networks* 5 (2016), pp. 78–93.

Maycock, P. *The future of energy summit 2013. bloomberg new energy finance*. Available: htt://bnef.com/Presentations/download/136 [Accessed January 2014]. 2013.

Mitra, Rajendu et al. “Voltage correlations in smart meter data”. In: *Proc. of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2015.

Morren, Johan, Sjoerd W. H. de Haan, and J. A. Ferreira. “Contribution of DG units to voltage control: active and reactive power limitations”. English. In: *Proc. of IEEE PES 2005 Russia PowerTech Conference*. 2005.

Nakamoto, Satoshi. *Bitcoin: a peer-to-peer electronic cash system*. 2008.

Ochoa, Luis F. and Gareth P. Harrison. “Minimizing energy losses: optimal accommodation and smart operation of renewable distributed generation”. In: *IEEE Transactions on Power Systems* 26.1 (2011), pp. 198–205.

O’Connell, Niamh et al. “Electric vehicle (EV) charging management with dynamic distribution system tariff”. In: *Innovative Smart Grid Technologies (ISGT Europe), 2011 2nd IEEE PES International Conference and Exhibition on*. IEEE. 2011, pp. 1–7.

Olivier, Frédéric, Petros Aristidou, et al. “Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units”. In: *IEEE Transactions on Smart Grid* 7.2 (2016), pp. 926–936.

Olivier, Frédéric, Damien Ernst, and Raphaël Fonteneau. “Automatic phase identification of smart meter measurement data”. In: *Proc. of the 24th International Conference on Electricity Distribution (CIRED)*. 2017.

Olivier, Frédéric, Antonio Sutera, et al. “Phase Identification of Smart Meters by Clustering Voltage Measurements”. In: *Proc. of XXth Power System Computation Conference (PSCC)*. Dublin, 2018.

ORES. “Les coulisses du photovoltaïques”. In: *ORES Mag* (2012), pp. 6–7.

Otomega, Bogdan and Thierry Van Cutsem. “Distributed load interruption and shedding against voltage delayed recovery or instability”. In: *Proc. of IEEE PES 2013 Grenoble PowerTech Conference*. 2013.

Palensky, Peter et al. “Modeling intelligent energy systems: co-simulation platform for validating flexible-demand EV charging management”. In: *IEEE Transactions on Smart Grid* 4.4 (2013), pp. 1939–1947.

Papaioannou, Ioulia T. et al. “Modeling and field measurements of photovoltaic units connected to LV grid. Study of penetration scenarios”. In: *IEEE Transactions on Power Delivery* 26.2 (2011), pp. 979–987.

Papathanassiou, S., N. Hatziargyriou, and K. Strunz. “A benchmark low voltage microgrid network”. In: *Proc. of CIGRE Symposium*. 2005.

Pappu, Satya Jayadev et al. “Identifying topology of low voltage (LV) distribution networks based on smart meter data”. In: *IEEE Transactions on Smart Grid* (2017).

Pedregosa, Fabian et al. “Scikit-learn: machine learning in Python”. In: *Journal of Machine Learning Research* 12.Oct (2011), pp. 2825–2830.

Penido, D.R.R. et al. “Four wire newton-rapshon power flow based on the current injection method”. In: *Proc. of IEEE PES Power Systems Conference and Exposition*. New York, 2004.

Pezeshki, Houman and Peter J Wolfs. “Consumer phase identification in a three phase unbalanced LV distribution network”. In: *Proc. of 2012 Innovative Smart Grid Technologies (ISGT Europe)*. 2012.

Poncelet, Kris et al. “Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems”. In: *IEEE Transactions on Power Systems* (2016).

Prim, R. C. “Shortest connection networks and some generalizations”. In: *Bell System Technical Journal* 36.6 (1957), pp. 1389–1401.

Rajeswaran, Aravind, Nirav P Bhatt, Ramkrishna Pasumarthy, et al. "A novel approach for phase identification in smart grids using graph theory and principal component analysis". In: *Proc. of 2016 IEEE American Control Conference (ACC)*. 2016.

Rathnayaka, A.J. Dinusha et al. "Goal-oriented prosumer community groups for the smart grid". In: *IEEE Technology and Society Magazine* 33.1 (2014), pp. 41–48.

Richardson, Ian et al. "Domestic electricity use: a high-resolution energy demand model". In: *Energy and Buildings* 42.10 (2010), pp. 1878–1887.

Rifkin, Jeremy. "Capitalism is making way for the age of free". In: *The Guardian* 31 (2014).

Samadi, Afshin et al. "Coordinated active power-dependent voltage regulation in distribution grids with PV systems". In: *IEEE Transactions on Power Delivery* 29.3 (2014), pp. 1454–1464.

Schleicher-Tappeser, Ruggero. "How renewables will change electricity markets in the next five years". In: *Energy Policy* 48 (2012), pp. 64–75.

Schwaegerl, C. et al. "Voltage control in distribution systems as a limitation of the hosting capacity for distributed energy resources". English. In: *Proc. of CIRED 2005 18th International Conference and Exhibition on Electricity Distribution*. 2005.

Schwanz, Daphne et al. "Stochastic assessment of voltage unbalance due to single-phase-connected solar power". In: *Proc. of 2016 International Conference on Harmonics and Quality of Power (ICHQP)*. 2016.

Srivallipuranandan, Navaratnam. "Series Impedance and Shunt Admittance of an Underground Cable System". PhD thesis. University of British Columbia, 1986, p. 125.

Stott, Brian, Jorge Jardim, and Ongun Alsaç. "DC power flow revisited". In: *Power Systems, IEEE Transactions on* 24.3 (2009), pp. 1290–1300.

Su, Xiangjing, Mohammad A. S. Masoum, and Peter J. Wolfs. "Optimal PV inverter reactive power control and real power curtailment to improve performance of unbalanced four-wire LV distribution networks". In: *IEEE Transactions on Sustainable Energy* 5.3 (2014), pp. 967–977.

Su, Xiangjing, Mohammad A.S. Masoum, and Peter Wolfs. "Comprehensive optimal photovoltaic inverter control strategy in unbalanced three-phase four-wire low voltage distribution networks". In: *IET Generation, Transmission & Distribution* 8.11 (2014), pp. 1848–1859.

Sunderland, Keith M. and Michael. F. Conlon. "4-Wire load flow analysis of a representative urban network incoprating SSEG". In: *Proc. of 2012 47th International Universities Power Engineering Conference (UPEC)*. IEEE, 2012.

Tanaka, K. et al. "Decentralised control of voltage in distribution systems by distributed generators". In: *IET Generation, Transmission and Distribution* 4.11 (2010), pp. 1251–1260.

Tant, Jeroen et al. "Multiobjective battery storage to improve PV integration in residential distribution grids". In: *IEEE Transactions on Sustainable Energy* 4.1 (2013), pp. 182–191.

Tesfatsion, Leigh and Kenneth L Judd. *Handbook of Computational Economics: Agent-Based Computational Economics*. Vol. 2. Elsevier, 2006.

Tonkoski, Reinaldo, Luiz A. C. Lopes, and Tarek H. M. El-Fouly. “Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention”. In: *IEEE Transactions on Sustainable Energy* 2.2 (2011), pp. 139–147.

Troester, Eckehard. “New German grid codes for connecting PV systems to the medium voltage power grid”. In: *2nd International workshop on concentrating photovoltaic power plants: optical design, production, grid connection*. 2009, pp. 9–10.

Turitsyn, Konstantin et al. “Options for control of reactive power by distributed photovoltaic generatorss”. In: *Proceedings of the IEEE* 99.6 (2011), pp. 1063–1073.

Urquhart, Andrew J. “Accuracy of Low Voltage Distribution Network Modelling”. PhD thesis. Loughborough University, 2016, p. 355.

Varma, Rajiv K., V. Khadkikar, and Ravi Seethapathy. “Nighttime application of PV solar farm as STATCOM to regulate grid voltage”. In: *IEEE Transactions on Energy Conversion* 24.4 (2009), pp. 983–985.

VDE. *VDE-AR-N 4105: Generators connected to the LV distribution network - Technical requirements for the connection to and parallel operation with low-voltage distribution networks*. 2011.

Villalva, Marcelo Gradella, Jonas Rafael Gazoli, and Ernesto Ruppert Filho. “Comprehensive approach to modeling and simulation of photovoltaic arrays”. In: *IEEE Transactions on Power Electronics* 24.5 (2009), pp. 1198–1208.

Vovos, Panagis N. et al. “Centralized and distributed voltage control: impact on distributed generation penetration”. In: *IEEE Transactions on Power Systems* 22.1 (2007), pp. 476–483.

Wagstaff, Kiri et al. “Constrained k-means clustering with background knowledge”. In: *Proc. of International Conference on Machine Learning*. 2001, pp. 577–584.

Walling, R.A. et al. “Summary of distributed resources impact on power delivery systems”. English. In: *IEEE Transactions on Power Delivery* 23.3 (2008), pp. 1636–1644.

Wang, Jing, Luis Miguel Costa, and Bouna Mohamed Cisse. “From distribution feeder to microgrid: an insight on opportunities and challenges”. In: *Proc. of the IEEE International Conference on Power System Technology (POWERCON)*. 2016.

Wang, Wenyu, Nanpeng Yu, Brandon Foggo, et al. “Phase identification in electric power distribution systems by clustering of smart meter data”. In: *Proc. of 2016 IEEE Machine Learning and Applications (ICMLA)*. 2016.

Wang, Wenyu, Nanpeng Yu, and Zhouyu Lu. “Advanced metering infrastructure data driven phase identification in smart grid”. In: *GREEN 2017 Forward* (2017), p. 22.

Warren Liao, T. “Clustering of time series data—a survey”. In: *Pattern Recognition* 38.11 (2005), pp. 1857–1874.

Watson, Jeremy D, John Welch, and Neville R Watson. “Use of smart-meter data to determine distribution system topology”. In: *IET Journal of Engineering* (2016).

WECC REMTF. *Generic solar photovoltaic system dynamic simulation model specification*. Available: http://www.wecc.biz/ [Accessed January 2014]. 2012.

Widén, Joakim and Ewa Wäckelgård. "A high-resolution stochastic model of domestic activity patterns and electricity demand". In: *Applied Energy* 87.6 (2010), pp. 1880–1892.

Wright, Andrew and Steven Firth. "The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations". In: *Applied Energy* 84.4 (2007), pp. 389–403.

Xin, Huanhai et al. "A self-organizing strategy for power flow control of photovoltaic generators in a distribution network". In: *IEEE Transactions on Power Systems* 26.3 (2011), pp. 1462–1473.

Xin, H. et al. "Cooperative control strategy for multiple photovoltaic generators in distribution networks". In: *IET Control Theory and Applications* 5.14 (2011), pp. 1617–1629.

Yeh, Hen-Geul, Dennice F. Gayme, and Steven H. Low. "Adaptive VAR control for distribution circuits with photovoltaic generators". In: *IEEE Transactions on Power Systems* 27.3 (2012), pp. 1656–1663.