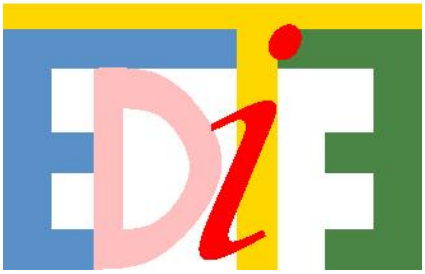
Modeling and optimizing a distributed power network : a complex system approach of the "prosumer" management in the smart grid

Nicolas Gensollen



Spécialité

Informatique

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Nicolas GENSOLLEN

Pour obtenir le grade de

DOCTEUR DE TELECOM SUDPARIS

Sujet de la thèse :

Modeling and optimizing a distributed power network: A complex system approach of the *'prosumer'* management in the smart grid

soutenue le 07 Octobre 2016

devant le jury composé de :

Prof. André-Luc BEYLOT	Rapporteur
Prof. Jean-Guy CAPUTO	Rapporteur
Prof. Matthieu LATAPY	Examinateur
Dr. Frederica DAREMA	Examinateur
M. Slawomir PIETRASZ	Examinateur
Associate Prof. Vincent GAUTHIER	Encadrant
Prof. Michel MAROT	Directeur de thèse
Prof. Monique BECKER	Directeur de thèse

Thèse n° 2016TELE0019

Declaration of Authorship

I, Nicolas GENSOLLEN, declare that this thesis titled, "Modeling and optimizing a distributed power network: A complex system approach of the '*prosumer*' management in the smart grid" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSITY PIERRE ET MARIE CURIE

Abstract

Doctor

Modeling and optimizing a distributed power network: A complex system approach of the *'prosumer'* management in the smart grid

by Nicolas GENSOLLEN

A french version of the abstract is available in the appendix B.

In recent years, it appears that autonomous operation of electrical systems is required. As technologies have evolved significantly since the deployment of electrical networks, their operation and methods have not evolved commensurately. Progress on communication networks and data management could open the doors to a new generation of power systems. The flexibility that can be enabled with these new technologies could radically change the current operation of power grids. Renewable energies, often considered unreliable, also bring their share of complications. Indeed, how to ensure system stability if production and consumption are stochastic? Within unidirectional and rigid systems, this sounds complicated. By considering the end user no longer as a fixed load, but rather as an involved actor, it creates the possibility that new solutions could emerge.

This thesis is devoted to the study of agents called prosumers because they can, via generators based on renewable energy, both produce and consume electricity. It is conceivable that, if their productions are above their own needs, prosumers will seek to sell their surplus. Of course, these surpluses are extremely volatile because they depend on many factors such as the weather at a given time, but also on the consumption of these agents. Since it does not exist, for the moment, accurate real data on the dynamics of these surpluses, the beginning of the thesis was devoted to the modeling and simulation of these prosumers. For this purpose, we used real weather data as inputs. These simulations enabled us to obtain time series describing how the excess productions (positive or negative) evolve for different agents.

The study of these series highlighted nontrivial spatiotemporal correlations which are of great importance for aggregators. As their name suggests, these agents form portfolios of generators, loads, and storage means, in order to sell services to the network operator. These services are intended to maintain the stability of the system, either by frequency regulation or by balancing production and consumption. By aggregating various prosumers, an aggregator may seek to stabilize its capacity. This is important because an aggregator, bound by a contract with the operator, may be subject to penalties if it is unable to fulfill its role. We show in this thesis that the correlation structure between the prosumers has a direct and decisive impact on the stability of aggregations, and therefore the risk assumed by aggregators. Thus, we propose an algorithm to minimize the risk of a set of aggregations, while maximizing their expected gain. The greedy method is based on graph theory and allows significant gains compared to a more random approach.

If the energy market and the interconnections between countries are supposed to help stabilize the system, it is still highly likely that a large increase in storage devices will be required. The placement of these devices in a network where generators and loads are dynamic and stochastic appears as a challenge. Indeed, the nonlinear dynamics of the electricity grid and the multitude of possible disturbances make the existence of a globally dominant placement rather unlikely. It seems more appropriate to seek a strategy with a good performance on average of possible situations. Given the size of the state space, a non-exhaustive method with warranty

appears to be necessary. We propose to answer this question with an approach based on control theory in networks. We model the electrical system by a network of coupled oscillators, which phase angles dynamic is an approximation of the actual dynamic of the system. The goal is to find the subset of the graph nodes that, during a disturbance of the system, would control it back to the equilibrium if the right signals are injected. We show that these signals can be interpreted as the power to which the storage devices inject or absorb. We then seek the placement that allows minimizing the average control energy. We thus propose a scalable algorithm, based on sub-modular functions, for finding a placement close to the optimum. We also show that there is a limit that forces the worst possible case returned by the algorithm.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Associate Prof. Vincent Gauthier and my two directors Prof. Michel Marot and Prof. Monique Becker for the continuous support of my Ph.D study and related research. I really enjoyed my Ph.D time especially because of the pleasant working atmosphere that we had. They have always supported me even in periods of doubts. Their guidance and experience helped me a lot in all the work and publications born out of this thesis. I could not have imagined having better advisor and directors for my Ph.D study.

I also want to thank Prof. Hossam Afifi for his numerous advices as well as my fellow labmates for the stimulating discussions and for the fun that we had during our Ph.D time.

Besides, I would like to thank the members of the jury: Prof. Jean-Guy Caputo, Prof. Andre-Luc Beylot, Prof. Matthieu Latapy, Dr. Frederica Darema, and M. Slawomir Pietrasz, for their insightful comments that improved the present work.

Last but not the least, I would like to thank my girlfriend and family for supporting me spiritually throughout writing this thesis and my life in general.

Contents

List of Figures

List of Abbreviations

AC	**A**lternating **C**urrent
BA	**B**arabàsi **A**lbert
DC	**D**irect **C**urrent
DER	**D**istributed **E**nergy **R**esources
DP	**D**ynamic **P**ricing
DSM	**D**emand **S**ide **M**anagment
EDF	**E**lectricité **D**e **F**rance
ER	**E**rdös **R**ényi
EV	**E**lectric **V**ehicle (s)
HVDC	**H**igh-**V**oltage **D**irect-**C**urrent
OSM	**O**pen **S**treet **M**ap
PV	**P**hoto**V**oltaic panel
RTE	**R**éseau de **T**ransport d'**E**lectricité
SM	**S**mall **W**orld
UPMC	**U**niversité **P**ierre et **M**arie **C**urie
V2G	**V**ehicle to **G**rid
VPP	**V**irtual **P**ower **P**lant
WT	**W**ind **T**urbine
WWW	**W**orld **W**ide **W**eb

List of Symbols

Prosumer model notations :

$P_i(t)$	Available extra-production of agent i	W
$P_i^P(t)$	Total production of agent i	W
$P_i^D(t)$	Total consumption of agent i	W
Z_i	zone of agent i	-

Wind turbine model notations :

$\nu_i(t)$	wind speed at agent i location	$\mathrm{m\,s^{-1}}$
ν_{in}	cut-in-speed	$\mathrm{m\,s^{-1}}$
ν_{out}	rated output speed	$\mathrm{m\,s^{-1}}$
ν_{cut}	cut-out speed	$\mathrm{m\,s^{-1}}$
P_m	rated output power	W
$\mathcal{F}_{WT}$	wind turbine power curve	-

Solar array model notations :

$\Psi_i(t)$	solar radiance at agent i location	$\mathrm{W/m^2}$
$\Psi_i^\star(t)$	perfect solar radiance at agent i location	$\mathrm{W/m^2}$
$\eta_i(t)$	cloudiness at agent i location	octat
ρ	array efficiency	-
S	array surface	$\mathrm{m^2}$
$\mathcal{F}_{PV}$	solar array power curve	-

Consumption model notations :

$\tau_i(t)$	outside temperature at agent i location	K
T_i	desired inside temperature for agent i	K
R_i	thermal resistance for agent i	$\mathrm{K\,W^{-1}}$
B_i	surface of thermal exchanges for agent i	$\mathrm{m^2}$
Ω_i	maximum consumption possible for agent i	W
ω_i	vector of the average fraction of Ω_i used for each hour	-
ϵ	noise term	-
$\mathcal{F}_i^{heat}$	heating power curve of agent i	-
$\mathcal{F}_i^{elec}$	appliance power curve of agent i	-

Market model notations :

$\mathcal{A}$	set of prosumers	-
N	number of prosumers	-
$P_S(t)$	available extra-production of coalition S	W
P_S^{CRCT}	contract value for coalition S	W

$P_S^{CRCT\star}$	best possible contract value for coalition S	W
$\mathcal{N}(\mu, \sigma)$	normal distribution with mean μ and standard deviation σ	-
ρ_{ij}	pearson's correlation coefficient between P_i and P_j	-
$Pr[\cdot]$	probability of some event	-
$erf(\cdot)$	error function	-
ϕ	reliability threshold	-
P^{MIN}	minimum production threshold	W
$\mathcal{U}(\cdot)$	utility function	-
α	coalition size parameter	-
P^{MAX}	normalizing factor for utility function	W
$\delta_S(i)$	marginal contribution of agent i to coalition S	-
N_{COAL}	number of coalitions	-
$\bar{N}$	average number of agent in a coalition	-
M	correlation matrix	-
$G(V, E)$	graph G with vertex set V and edge set E	-
$\mathcal{N}_i$	neighborhood of node i	-
d_{ij}	distance between i and j	-
ϵ	filtering threshold	-
$\epsilon^\star$	optimum filtering threshold	-
$\Theta_k(G)$	set of non overlapping cliques of size k in G	-
$\mathcal{R}_S$	resilience of coalition S	-
ψ	fraction of failed agents	-

Power grid dynamic notations :

θ_i	phase angle of node i (or deviation from Ω)	rad
δ_i	phase angle of node i	rad
ω_i	natural frequency of node i	Hz
K	coupling strength	-
γ	phase cohesiveness	rad
$L^\dagger$	moore-penrose pseudo inverse of laplacian L	-
$\|\cdot\|_{\infty,E}$	worst case dissimilarity over edge set E	-
Ω	main frequency	Hz
K_D	dissipation constant	$kg.m^2.s^{-1}$
I	inertia	$kg.m^2$
P_{ij}^{MAX}	maximum capacity of line (i, j)	W
$P_{S,i}$	ideal power source/sink at node i	W
α	dissipation term	s^{-1}

Control notations :

$u(t)$	input vector of control signals	-
$u^\star(t)$	optimum input vector of control signals	-
A	dynamic of the system	-
$X(t)$	state of the system	-
B	control matrix	-
T	control time	s
C	controllability matrix	-
I	identity matrix	-
N_D	number of drivers	-
n_D	fraction of drivers in the system	-

$\mathcal{E}$	control energy	-
$W(t)$	Gramian matrix	-

Storage model notations :

r	maximum charge/discharge rate	W
Λ_{MAX}	Maximum capacity	Wh
$\Lambda(t)$	Energy stored at time t	Wh

Chapter 1

Introduction

1.1 Energetic transition

Since the industrial revolution the amount of energy consumed increased continuously, as shown in figure 1.1. Since the world population has also grown in the mean time, figure 1.2 shows what is happening at the level of the individual (data of figure 1.1 are divided by world population estimates). With the advance of technology, more and more energy is required to secure our living standards, and it seems quite unlikely that this curve will plateau all of a sudden.

During the 20th century, countries built impressive systems for delivering power, gaz, and water to the populations. Production techniques evolved such that different sources can be combined. Nowadays, a large portion of the energy that we consume comes from nuclear plants, especially in France. But a non negligible share is still provided by coal or oil based plants (see chapter 2 for more details).

Although these systems work with decent reliability, the necessity for improvements is becoming more and more obvious. There are several well known reasons for this, ecological concerns, progresses in renewables and storage efficiencies, the necessity to renovate old installations, or security to name a few [2] [77]. Indeed, most power systems were built before the information revolution that occurred at the end of the 20th century. The whole architecture was thus not initially thought of as to incorporate information, at least not in the way we use it nowadays.

Power systems are not the same as they were at their early stage though. They indeed evolved with the technology advances so that the real benefits of re-thinking the whole architecture versus the cost of such a huge operation is dividing the community. In a period of budget cuts, should we invest in the renovation of old fossil plants or in the development of renewables and new techniques ?

One of the most known problems of the smart grid vision is perhaps the will of a large penetration of renewables [92] [77]. These generators use natural and free resources such as wind, tides, or solar power instead of fossil ones. Most of these resources are known to be hardly predictable and not necessarily correlated with the consumption curves [89]. In a world where the service rate should be as close as 100% as possible, a system relying only on such stochastic generators seems quite unrealistic.

A reliable power system functioning with distributed renewable energy sources might indeed be hard to imagine when we start thinking about it in more depth. But we might also feel this way because we are so used to the current system that conceiving something completely different is beyond

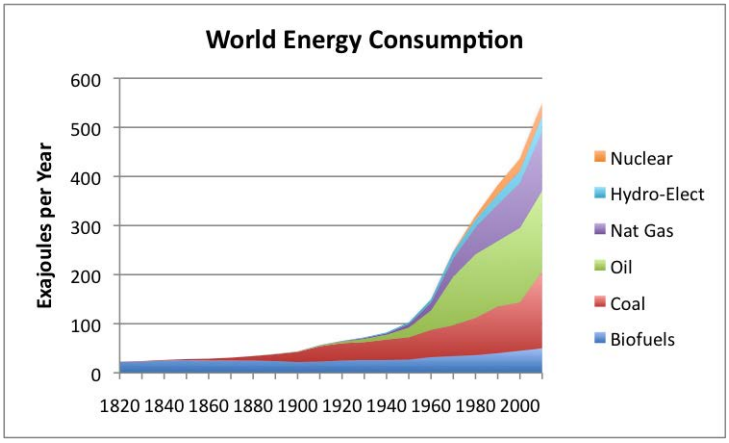


FIGURE 1.1: World Energy Consumption by Source (from [68]).

our reach. Not taking the availability of power for granted and organize the consumption in function of the production can be hard to accept when our whole lives where spent the other way around. Is the transition to smart grids a simple upgrade in the technologies, or is it rather a deeper mutation down to the way users perceive the energy ?

This thesis focuses on the end users of a smart grid system. We assume that they are equipped with smart meters, storage equipments, and that they own small renewable generators that supply energy to their home appliances. These agents will be called prosumers because they have both the ability to consume and produce power.

1.2 Structure of the thesis

This thesis is structured in two main parts that both deal with prosumers, but with a different perspective. Both parts are organized with the same logic. The first chapters for each part (chapter 3 for Part 1 and chapter 6 for Part 2) present the notions and formalism required for the understanding of the rest of the work. Their role is also to locate our contributions in the related literature.

The core chapters for each part (chapter 4 for Part 1 and chapter 7 for Part 2) explain our contributions to the smart grid literature. Hence, the works presented in these chapters are the fruits of our labor unless it is clearly stated otherwise. Although we tried to split clearly our contributions from the state of the art in different chapters, it was sometimes more logical and practical to explain some concepts directly within the contribution chapters. In this situation the references are clearly stated such that no confusion should be possible.

The last chapter of Part 1 (chapter 5) presents currently ongoing work that we chose to include here because it constitutes a natural extension to the contributions made in chapter 4. It relies on concepts introduced in chapter 3 and 4 but requires more tools which are reviewed within chapter 5 itself. Again, the distinctions between the existing literature and our contributions are clearly stated.

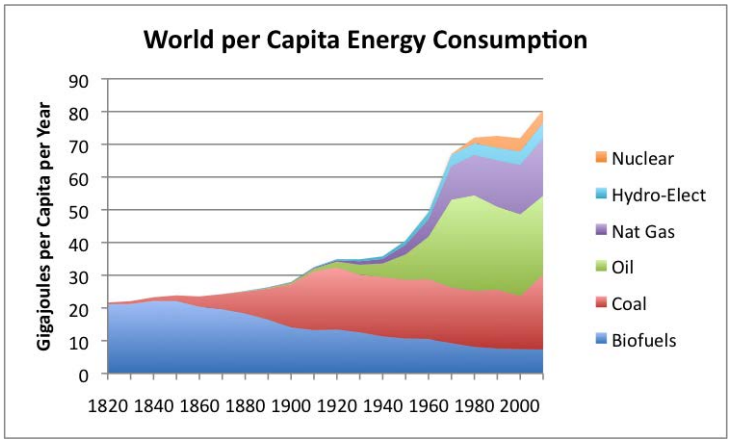


FIGURE 1.2: Per capita world energy consumption, calculated by dividing world energy consumption shown in Figure 1.1 by population estimates (from [68]).

The two parts of the thesis both require more basic knowledge about power grids, graph theory, complex systems, and smart grids. These are introduced in chapter 2 that follows this introduction. Obviously, no personal contribution is made here, and this chapter should be read as a refresher on these topics. The reader comfortable with these concepts can skip chapter 2 and go directly to Part 1.

1.3 Discussed subjects

Understanding smart grid systems requires the study of traditional power grids. Chapter 2 will thus explain the basic architecture of power systems from generation to end-users. However, we will rapidly turn to a complex system framework for their description. Some basic notions of graph theory and complex systems will be presented and later used to better understand the topology of power grids. The last part of chapter 2 introduces smart grid key challenges and its global architecture.

1.3.1 Part 1

As explained briefly in this introduction, the focus of this thesis is on a special kind of end-users called prosumers. Part 1 is devoted to the study of prosumers as they try to sell their production on electricity markets. Chapter 3 define more precisely these prosumers, their importance in the smart grid ecosystem, and the challenges they pose. One of these challenges consists in enabling prosumers to sell their surplus of production. In this thesis, we chose to study such possibility through a market environment, where energy is traded between market participants with the idea of maximizing the benefits while keeping a stable system. Energy markets are already parts of power systems but they are expected to become more significant in the future. They are introduced in the second section of chapter 3.

We will see that, although simple prosumers cannot reasonably enter these markets, aggregations controlled by a central entity could very well be considered as participants. Forming these aggregations with the aim of

joining the market and make profits is the task of aggregators (see chapter 3). These aggregators sells a service such as a production capacity for a given period and are remunerated according to the quality and reliability of this service. Eventually, if an aggregator fails to provide the service, some penalties could be applied. Maximizing the benefits necessitates thus to pay attention to the compositions of the aggregations.

Chapter 4 starts by explaining how we simulated prosumer productions and consumptions based on real weather data. The chapter then presents the algorithm that we developed as to form stable aggregations of prosumers for electricity markets. The idea behind this work is to optimize the expected production to risk ratios for the aggregators. This algorithm relies on a special kind of graphs, called correlation graphs, that describe the correlation relationships between a set of variables. The work presented in chapter 4 was done on static correlation graphs although correlation can be seen as a time dependent measure. In chapter 5, we take this into account and study correlation graphs as temporal graphs whose topologies evolve along time. Relying on recent results based on tensor factorization, we study time dependent clusters of correlated variables with some real examples related to energy consumption in power grids.

1.3.2 Part 2

In all chapters of Part 1, the network topologies and power dynamics were abstracted so that we could only focus on the clustering of the prosumers without the constraints of the network connecting them together. The goal of Part 2 is to include these constraints by explicitly considering the grid topology and the power dynamics (in a simplified version). More specifically, we assume in Part 2 a pool of prosumers interconnected by electrical lines. As will become clear later on, this kind of network is quite unstable because of the stochasticity of the prosumers productions and consumptions. The central question of Part 2 is to decide where to place storage elements in the network such that we have the ability to control the system at low energetic costs.

Chapter 6 presents the tools used in chapter 7. It is divided in two main sections about dynamics on networks and control of networks. More precisely, chapter 6 starts by explaining dynamics on networks at large before turning to the power grids. It turns out that the complex and non linear power grid dynamics can be represented, under a few assumptions, by a synchronization model. Chapter 6 explains how this abstraction can be done, and presents a few examples for simple networks. The emergence of synchronization in complex systems has been studied lately and fascinating results have been obtained and will be of great help in chapter 7.

Synchronization of the electrical machines to a common frequency in the power grid is a necessary condition for stability (see chapter 2). If the frequency deviates from the reference value, it can potentially damage elements in the system. The synchronization model described in the first section of chapter 6 can be used to show that such problems can happen when there is an imbalance between production and consumption. Intuitively, a smart grid system with stochastic production and consumption seems quite likely to experience thess kinds of situations, such that some sort of control should be implemented. These control actions, called frequency regulation,

are already implemented in current power systems. Nevertheless, a large penetration of renewables calls for new optimized methods.

In Part 2, we study the frequency regulation problem in the case where users are prosumers and the control actions are taken at storage elements distributed in the network. That is, we consider a network of renewable generators and loads whose outputs might change over time. It can very well be the case that generators at $t-1$ become loads at t and vice versa. The generators and loads of such networks are thus susceptible to change. The problem we address is to find beforehand the best placement for storage elements in this situation. The second section of Chapter 6 provides tools related to control and optimal control theory. Since this is a vast subject, we particularly focus on recent advances on control in networks.

In chapter 7, we use the results of chapter 6 to tackle the storage placement problem. For a power system abstracted in an oscillator network, we develop an algorithm relying on optimal control theory and the optimization of submodular set functions. We show that this algorithm enables us to find placements that require less energy for frequency regulation, even in this stochastic context.

Finally, the last chapter concludes the present thesis by a summary of the contributions as well as an extension to ongoing and future work.

Chapter 2

Smart grids as complex systems

A smart grid can be thought of as an electrical grid coupled with an information system [90]. Such information system is utilized to organize and insure the proper functioning of the global system. It thus takes information from measurement units (such as smart meters), sends signals to smart appliances, and composes with uncertain renewable energy resources. Obviously, all these operations should be made within stable regions of the power grid state space. Actually, understanding the smart grid challenges necessitates the good comprehension of what power grids are, what they look like, and how they behave. This chapter first presents power grids in a very general manner and introduces concepts of graph theory that will help us understand what they are. Finally, we will introduce in more depth smart grids and related challenges.

2.1 Power grids

A power grid is an interconnected network for delivering electricity from suppliers to consumers. Power is produced at generating stations that may be located near a fuel source, at a dam site, or somewhere to take advantage of renewable energy sources. Generating stations are usually located away from heavily populated areas and are usually quite large to take advantage of the economies of scale. This means that the generated power needs to be transported from these generation sites to the wholesale customer, which is usually the company that owns the local electric power distribution network. To decrease the losses in the transmission network, the bulk power is stepped up to high voltage through transformers. At a substation, the power is stepped down from a transmission level voltage to a distribution level voltage. And finally, upon arrival at the service location, the power is stepped down again from the distribution voltage to the required service voltage (see figure 2.1).

Power grids are networks that do not seem random at first glance [71]. Indeed, strong geographical and budget constraints are weighting on the construction of power grids. Generation sites as well as cities, where most of the end users live, are not positioned randomly over a given territory and impact the topology of the grid. Since deploying a substation, a transformer, or an electrical line is expensive, we expect the grid to be the result of some kind of optimization process, even unconscious. For example, it is well known that tree structures are the cheapest topologies still providing connectivity. But at the same time, since redundancy is absent from such structures, they are extremely fragile against failures. Such redundancy

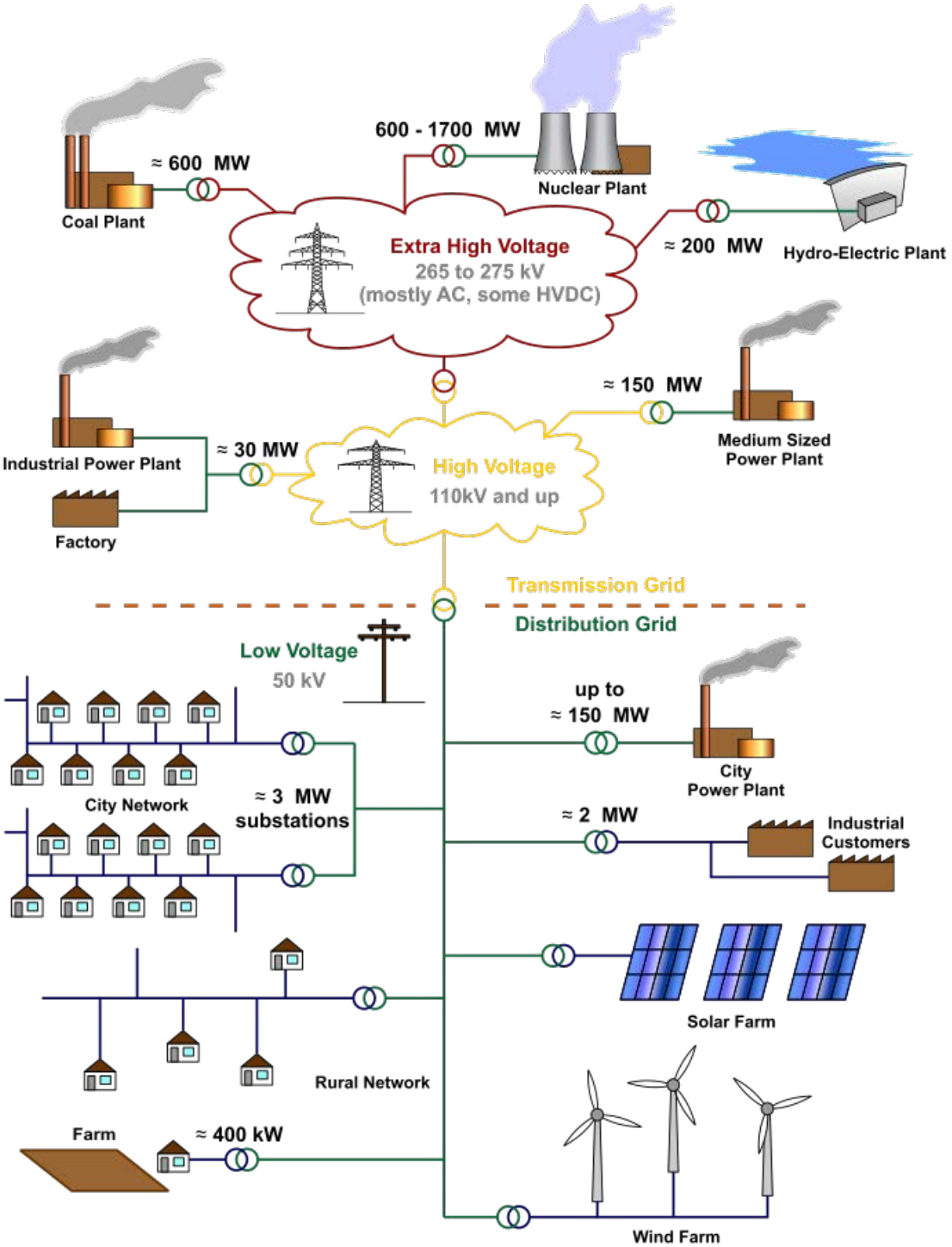


FIGURE 2.1: Power grid architecture. Source : Wikipedia

over cost tradeoffs are actually typical of engineered systems. Another particularity that might be visible by the eye is that each country is responsible for developing its own network, meaning that there might be large variability across countries, and under-developed interconnections between them.

We clearly need a more rigorous framework to study the topology of power grids. In the next section, we introduce notions of graph theory and complex system theory that will be used throughout this thesis.

2.2 Graphs and complexity

2.2.1 Graphs and their representations

A graph G can be defined as a tuple (V, E), where V is the vertex set and E is the edge set. It is also very common to refer to vertices as nodes and to edges as links. In this thesis we use these words alternately without any distinction. In this section, we denote by $n = |V|$ the number of nodes in the graph, and by $m = |E|$ the number of edges. These edges can be directed or undirected depending on the kind of relationships we are trying to model. Directed edges can represent people following each other on Twitter for instance while undirected edges can be used for facebook friendships. Indeed, Facebook requires that both parties agree on being friends while one can follow anybody on Twitter without any approval. Along this thesis, we will restrict ourselves to simple graphs, that is, graphs that do not exhibit multiple edges or self-loops. Unless stated otherwise, a graph will always mean a simple undirected graph.

If a graph has no edges, i.e $m = 0$, it is called an empty graph. On the contrary, a graph where all possible edges are present, i.e $m = \frac{n(n-1)}{2}$ for undirected and $m = n(n-1)$ for directed graphs, is called a complete graph. In between these two extremes, a graph is said to be connected if any node can be reached by any other node. In other words, a graph is connected if there exists a path between any pair of nodes. A path is a special kind a walk (sequence of vertices and edges, with both endpoints of an edge appearing adjacent to it in the sequence) where there is no repetition of vertices and edges. A graph that is not connected can be split in a set of connected components, the smallest possible connected component being a single node.

The most basic way of representing a graph is to store the list of nodes and the list of edges. Because most of real world networks are sparse, this method has the advantage of requiring small amount of memory to store a graph. However, a matrix representation, the so-called adjacency matrix, is often used because several graph properties can be deduced from it. Let A be the $n \times n$ adjacency matrix of G such that $a_{ij} = 1$ if node i is connected to node j and $a_{ij} = 0$ otherwise. Obviously, adjacency matrices of undirected graphs are symmetric and therefore have real eigenvalues and an orthogonal eigenvector basis. This set of eigenvalues is often called the spectrum of the graph and is known to display several interesting properties. Besides, the element (i, j) of A^k gives the number of (directed or undirected) walks of length k from vertex i to vertex j. A direct and pretty useful consequence is that the number of triangles N_Δ of an undirected graph G can be calculated by:

$$N_\Delta = \frac{Tr[A^3]}{6} \tag{2.1}$$

,where $Tr[\cdot]$ is the trace operator (see [84] for more information).

For an undirected simple graph, a node i can be connected to $k_i \in [0, n-1]$ other nodes, where k_i is called the degree of node i. For directed graphs, each node i has an in-degree k_i^{in} of links pointing to it, and an out-degree k_i^{out} of edges departing from it.

That being said, another important matrix representation of a graph is the Laplacian matrix $L = D - A$, where D is the $n \times n$ degree matrix ($d_{ij} = k_i$ if $i = j$, and $d_{ij} = 0$ otherwise). The Laplacian matrix, is symmetric and positive-semidefinite such that it has a positive real spectrum. The smallest eigenvalue of the Laplacian will always be zero, and the number of times zero appears as an eigenvalue is the number of connected components in the graph. This matrix has a long list of fascinating properties [54] and is often used for studying complex phenomena in networks [13] [61]. It will be pretty useful later on in this thesis when the subject of synchronization in networks will be tackled.

2.2.2 Characterizing graphs

Although they might seem pretty abstract to some readers, graphs can be used to represent an outstanding number of systems or phenomena. From airports interconnections [34] to social networks [1], world wide web topology [19], or power grids [72] [79], the network-oriented approach is becoming more and more popular. But once the system has been abstracted into a nice and beautiful graph, what kind of information have we obtained ? For very small graphs, a simple plot might give some interesting information, but when the number of nodes increases above, say, a hundred, one usually obtain a hairy and messy ball that does not really teach us anything. In this context, how can we say that two graphs are similar to each other?

For such systems, a possible solution is to build metrics that captures some property of the network. Then, by comparing these metrics across networks, one is theoretically able to assess whether two networks can be considered as close or not and on which criterion. By just looking at the number of nodes and edges, one can come up with the density measure that captures the fraction of possible edges that are indeed present in the graph:

$$density(G) = \frac{2m}{n(n-1)} \tag{2.2}$$

If we consider connected graphs, we know that we can reach any node j from any other node i. Most likely, there will exists multiple walks from i to j, but among them, the one that requires the minimum number of edges is called the geodesic distance between i and j (d_{ij}). This enables to look at two other natural measures :

$$diameter(G) = max_{i \in V} e(i) \tag{2.3}$$

$$radius(G) = min_{i \in V} e(i) \tag{2.4}$$

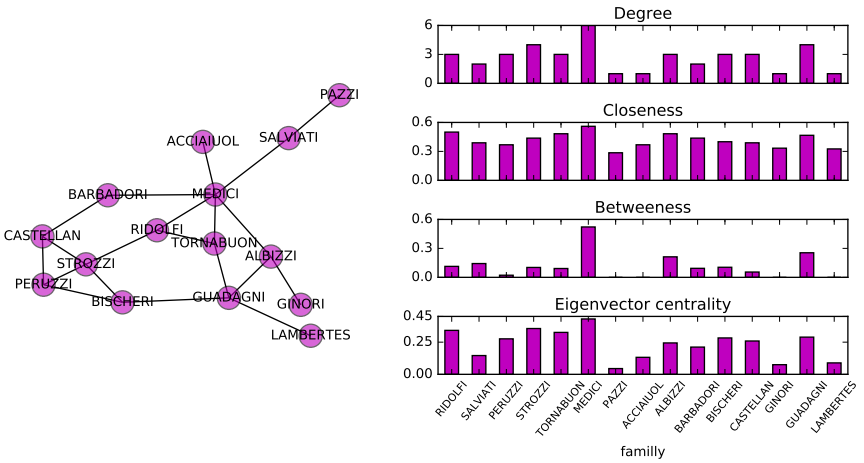


FIGURE 2.2: Network of medicis

,where $e(i)$ is the eccentricity of node i, that is, the greatest geodesic distance between node i and any other node. Basically, these measures capture how spread out the graph is. These three measures give a rough idea of the graph structure, but in most cases this is not sufficiently accurate and more precise quantities need to be computed.

2.2.3 Centrality measures

Instead of considering the whole graph, let focus on its parts, i.e its nodes and links. Are every node and every edge equally "*important*" within a graph ? And what do we mean by "*important*" ?

A centrality measure attributes a real value to each node and/or edges depending on some property that we wish to capture. The most basic node centrality measure, the degree, has already been introduced in the previous sections. The measure enables to rank nodes according to their number of connections. Although simple, this number is in many networks an indicator of nodes' importance.

Figure 2.2 shows the network of social relations among Renaissance Florentine families. Nodes represent families and a link is present between two families if they have business relationships (note that another network with marriage relationships also exists). Among these families, the Medicis were known to be extremely powerful at the time. Some studies used network representations to understand why the medicis where so strong [69]. Here, we are not particularly interested by understanding the shape of the network, but rather the relative importance of its nodes. The right panels of figure 2.2 display different centrality measures as to compare the families. Here, the degree captures the number of business relations that a given family has created. We can see that the Medicis have the greatest degree, but also that the difference with the others is not that large, meaning that there are perhaps other explanations for the Medicis domination.
The closeness centrality is defined as :

$$C_G(i) = \frac{1}{\sum_{j \in V} d_{ij}} \tag{2.5}$$

, where d_{ij} is the distance between nodes i and j. This metric quantifies how close to all the other nodes a given node is. We can see on figure 2.2 that this metric does not seem to separate clearly the Medicis from the other families.
The betweeness centrality is defined as :

$$B_G(i) = \frac{2}{(n-1)(n-2)} \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{2.6}$$

, where σ_{st} is total number of shortest paths from node s to node t, and $\sigma_{st}(v)$ is the number of those paths that pass through v. Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. By looking at figure 2.2, we can see that the betweenness centrality of the Medicis is much larger than for the other families. This means that, in order to communicate efficiently, the families had to pass through the Medicis more frequently than any other family. It has been suggested that this could be an explanation for the Medicis' domination [69].

There exists many more centrality measures for nodes or edges that aim at capturing a specific effect [15] [19]. We only present here another measure which might feel less obvious to the reader. The eigenvector centrality is a measure of the influence of a node in a network. The key idea is that, for a given node i, a connection to another node j with a high centrality should contribute more to the centrality of i than other links to low centrality nodes. It is possible to think of it as a popularity measure, that is, my popularity should depend on the popularity of the nodes I am connected to. Indeed, the famous Google's PageRank is a variant of the eigenvector centrality measure. The Eigenvector centrality of node i can be obtained by looking at the i^{th} component of the eigenvector corresponding to the largest eigenvalue of the adjacency matrix.

2.2.4 Random models

As for now, we looked at graphs in a static way, meaning that we consider relations within a population at a given time and we abstract these into a graph. Nevertheless, we might be interested in understanding how such topologies arised. Indeed, most networks do not appear suddenly with all their properties, they are usually the result of a long process of additions and removals (nodes and/or links). Take a social network for instance. At the beginning, only very few people are present, the so-called early adopters, but as time goes on, more and more people subscribe and begin to form ties with already present persons.

A key point in graph theory was the discovery of random graph models by the physicists Paul Erdös and Alfréd Rényi in 1959 [20]. Their model builds a network from a set of nodes by assuming that all nodes are equally likely to link together. More precisely, one considers all possible pairs of nodes and link them with a given probability p. Obviously, if p tends to zero the graph tends to an empty graph, and if p tends to one, the graph tends to a complete graph. In between, Erdös and Rényi discovered the different values at which different phenomena occur. Figure 2.3 shows a small example with $n = 50$ nodes for different values of p. It has been

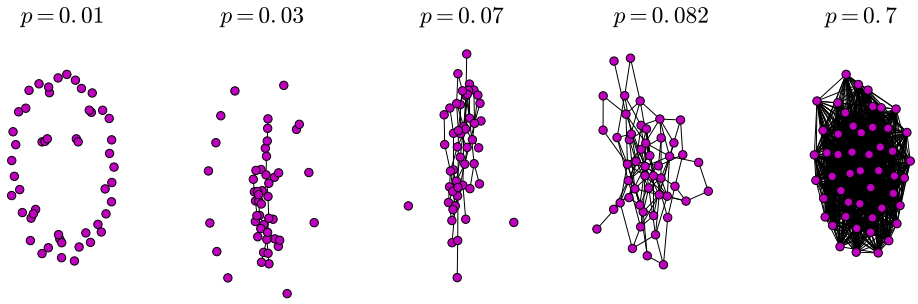


FIGURE 2.3: Erdös-Rényi networks of $n = 50$ nodes for different values of p.

shown that if $p > 1/n$ then the graph will contain almost surely a giant component. For smaller values of p, the network will only contains isolated nodes and small components as visible on the two left most subplots of figure 2.3. Moreover, if $p > \frac{ln(n)}{n}$, then the graph will almost surely be connected (see subplots 3 and 4 of figure 2.3).

This model is extremely important in complex systems theory since it gives a null hypothesis. That is, if a graph is shown to be close to an erdos-renyi graph, one can presume that the underlying process that gives rise to this graph is mostly due to randomness. A natural question that we can ask now is: are real networks random ? Very interestingly, the answer tends to be almost always negative [62], meaning that real world networks are not the results of pure noise, but rather the outcomes of underlying hidden processes.

An erdos-renyi random graph has on average $p\binom{n}{2}$ edges and the distribution of the degree of any vertex is binomial :

$$P[k_i = k] = \binom{n-1}{k} p^k (1-p)^{n-1-k} \tag{2.7}$$

which tends to a Poisson distribution when $n \to \infty$ and $np = Cste$. This means that nodes degrees are quite homogeneous and that large deviations from the mean degree are extremely unlikely. This is a first point which can let us think that most real networks are not random [62], because they tend to have very heterogeneous degrees [6]. A very few nodes have often very large degrees while the rest has relatively small numbers of connections. The most famous networks exhibiting this kind of property are the internet, the world-wide-web, or the airport network. More precisely, it is known that their degree distributions can be put as a power-laws [6] :

$$P[k_i = k] \sim k^{-\gamma} \tag{2.8}$$

where γ is a parameter whose value typically ranges between 2 and 3, and are usually called scale-free networks. Power-law distributions are not specific to networks and can be found in many diverse domains. For instance, it has been shown that the wealth of populations tends to follow power-laws. There exist very nice models that explain and reproduce the scale-free effect. They usually rely on the "rich get richer" principle as to mimic the large inequalities that come from power-laws. For graph theory, the

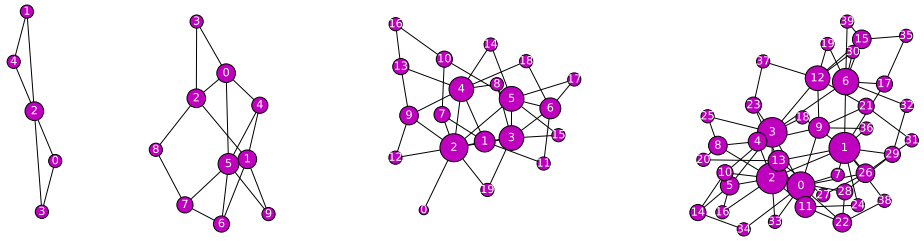


FIGURE 2.4: Barabasi-Albert preferential attachment model. Nodes sizes are proportional to their degrees.

Barabási–Albert model is one of several proposed models that generates scale-free networks [6]. It incorporates two important general concepts: growth and preferential attachment. Basically, nodes are added one by one and the probability that a new node j connects to an older node i is proportional to the degree of that node $p_i = \frac{k_i}{\sum_v k_v}$ (the sum goes over all existing nodes). Because of preferential attachment, older nodes tend to accumulate more connections than younger ones, forming hubs. Figure 2.4 shows a small example of a growing network with preferential attachment. Every time a node is added, it connects to δ old nodes ($\delta = 2$ in figure 2.4) with a biased probability toward high degree nodes. It is clearly visible that old nodes, e.g nodes with small labels, tend to have more connections that new nodes.

By looking at social networks, it seems clear that neither ER or BA models are able to capture their true underlying structures. The BA model captures well the degree distribution, but misses more complex structural properties. Indeed, nodes of real networks do not necessarily connect only based on degree. For instance people do not necessarily become friends with the most popular person in the network if they do not know each other. They rather tend to connect to people they know. And since it is likely that two of my friends have met and are friends themselves, local triadic structures are extremely common in real networks, especially social networks. The clustering coefficient is a way of measuring the importance of triangles within a graph and can be computed either globally or locally :

$$C_{GLOB} = \frac{3 \times N_\Delta}{N_V} \tag{2.9}$$

where N_Δ is the number of triangles in the graph, and N_V is the number of connected triplets. C_{GLOB} basically computes the portion of closed triplets over the total number of triplets. The local clustering coefficient of a vertex i in a graph $G(V, E)$ quantifies how close its neighbors are to being a clique :

$$C_i = \frac{|\{e_{jk} : v_j, v_k \in \mathcal{N}_i,\ e_{jk} \in E\}|}{k_i(k_i - 1)} \tag{2.10}$$

where e_{jk} is the edge from node j to node k, and k_i is the degree of node i. This measure was introduced in 1998 by Duncan Watts and Steven Strogatz who observed that some real networks (especially social networks)

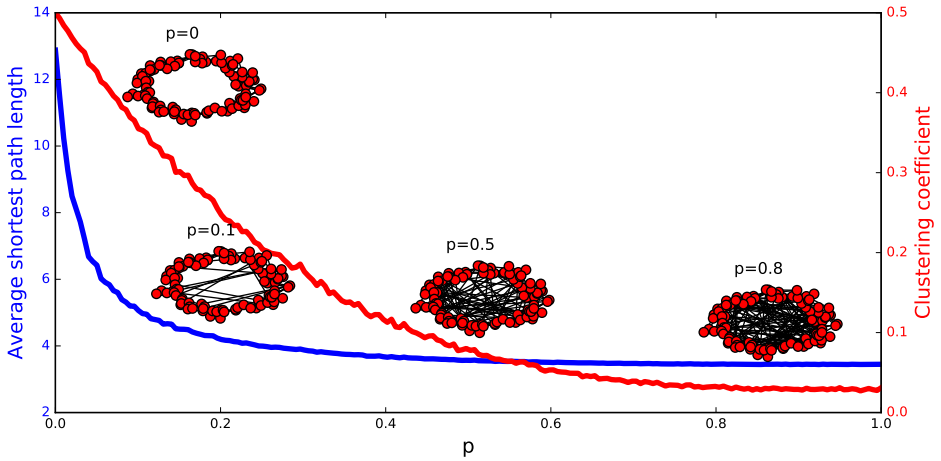


FIGURE 2.5: Small World model.

exhibit both a small average shortest path length and a high clustering coefficient [91]. This observation is at odds with ER networks that do not have such clustering. Watts and Strogatz then proposed a novel graph model, often called the small-world model, that reproduces these characteristics [91]. The idea of the model is to start from a large world such as a lattice of locally well connected nodes. This starting point has a high clustering coefficient but a large average shortest path (see figure 2.5). The model then re-wires a fraction p of the links to some randomly chosen nodes. Clearly, if $p = 0$ we still have the lattice, and if $p \longrightarrow 1$ we tend to obtain a random ER network (see figure 2.5). Nevertheless, in between these two extremes, we see on figure 2.5 that the average shortest path length decreases much more rapidly when p increases above zero than the clustering coefficient. This means that for a region of relatively small p values, we obtain a graph which has both a low average shortest path length as well as a high clustering coefficient.

2.2.5 Community structures

The small-world model captures the small distance, yet, highly clustered topologies of some real networks such as social networks. However, when comparing, even by eye inspection, small-world networks to real social networks, it seems that something is missing. Indeed, social networks seem to have much more complicated underlying structures with regions of high density of links and regions with very few links. These clusters of well connected nodes are often called communities in complex system theory, and a whole branch is devoted to uncover them in large graphs.

The first idea that might come into mind for detecting communities is to minimize a graph cut function that basically indicates how many edges lie in between groups. However, methods relying on cut functions were shown not to perform well. Indeed, the basic cut placing all nodes in one group and none in the other always minimizes the cut function but is of little use. There has been multiple suggestions in order to penalize this result and force the algorithm to perform non trivial cuts. But, most of these methods require special knowledge about the partitioning such as the sizes

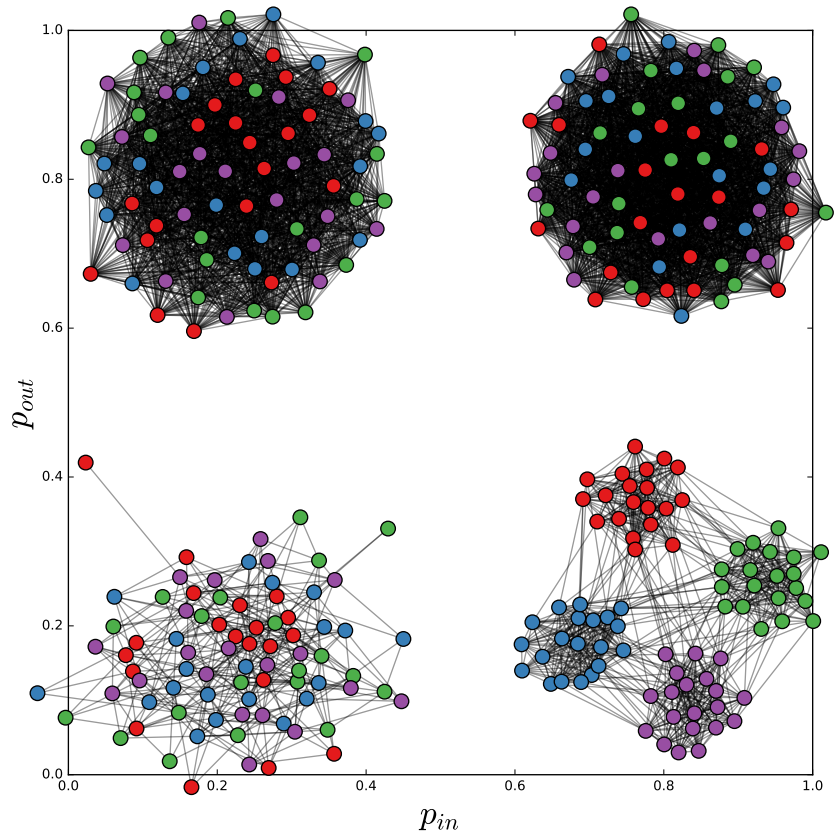


FIGURE 2.6: Planted partition model.

of the communities. It turns out that cut functions might be not the most suited functions for community detection :

> *The fundamental problem with all of these methods is that cut sizes are simply not the right thing to optimize because they don't accurately reflect our intuitive concept of network communities. A good division of a network into communities is not merely one in which the number of edges running between groups is small. Rather, it is one in which the number of edges between groups is smaller than expected.*
>
> Mark Newman in [59] (2006)

This is a significantly different approach to community detection that led Newman to propose another objective function called the modularity [59], which is widely used nowadays. Modularity is the fraction of the edges that fall within the assumed communities minus the expected such fraction if the edges were distributed at random :

$$Q = \frac{1}{(2m)} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{(2m)} \right] \delta(c_v, c_w) \tag{2.11}$$

where $\delta(c_v, c_w) = 1$ if node v and w belong to the same community and $\delta(c_v, c_w) = 0$ otherwise. Finding both the number of communities and

partition of the nodes requires an algorithm for optimizing Q. One of the mostly used methods is the Louvain algorithm, but several other methods have been developed (Label propagation [76], clique percolation [73], Walktrap [75], spectral methods [60], maximum likelihood estimation [93]...).

In order to evaluate the performance of community detection algorithms, one need to compare the results with the true structure. In most networks, this ground truth might not be accessible such that models were developed as to reproduce these graphs from a known community structure. A widely used model for this purpose is the planted partition model, where nodes are assigned to communities and links are added afterwards. In this model, there are two different probabilities :

- p_{in} : probability of having a link between two nodes within the same community
- p_{out} : probability of having a link between two nodes in different communities.

Figure 2.6 shows the resulting graphs for a small example with 4 communities of 20 nodes each. When p_{in} is similar to p_{out}, the resulting graph is mostly random with no real community structure. When p_{in} is larger than p_{out}, communities start to appear and become more and more separated as p_{in} increases. The case where $p_{out} > p_{in}$ basically means that nodes within the same community are less likely to connect that nodes in different communities. However this might not make much sense at first glance, anti-communities have received more attention lately [95] [59].

These definitions and models still have some limitations. For example, we made the implicit assumption that a node could only belong to a single community although this might not be true in the reality. Actually, in social networks people most often belong to multiple communities (family, friends, work colleagues, associations, and so on...) [93]. Obviously, if two nodes share multiple communities they become much more likely to be connected. Figure 2.7 shows an example of such overlapping community structure. The network of figure 2.7 was constructed with 3 different communities of different sizes (blue, red, and green), but nodes can belong to more than one group such that we observe smaller but more densely connected clusters at the intersections of the communities (nodes in white are involved in all communities). Detecting overlapping community structures has attracted much attention lately, and multiple solutions have been proposed [73] [93].

This section was devoted to give to the non specialist reader the necessary information to read the rest of the manuscript and is not intended to be a complete state of the art of the complex system theory, a task that is way out of the scope of this thesis. The interested reader could refer to very complete and well-written survey on the subject [61] [15]. The next section will use some of these complex system concepts to study power grid topologies.

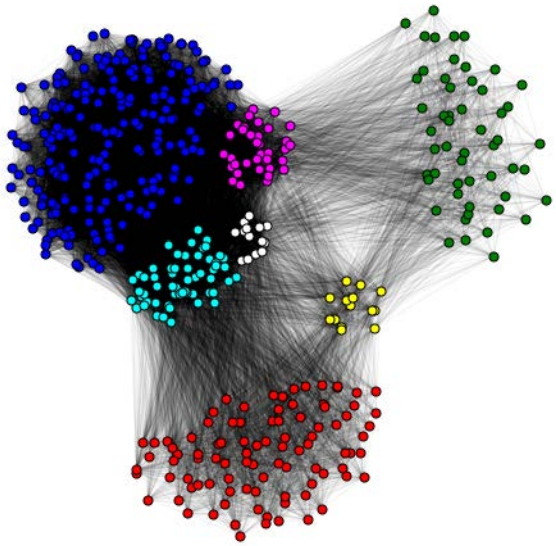


FIGURE 2.7: Overlaping community structure.

2.3 Power grids as complex networks

2.3.1 Abstracting the topology

As explained in the introduction part of this chapter, power grids are usually divided in 4 different parts (see figure 2.8) :

- **Generation :** Installations where power is generated. In order to be sent over long distances, power is elevated to high voltages through transformers and injected in the transmission network.
- **Transmission :** Most transmission lines are high-voltage three-phase alternating current (AC). High-voltage direct-current (HVDC) technology is used for greater efficiency over very long distances, for submarine power cables, or for interchange of power between grids that are not mutually synchronized. Electricity is transmitted at high voltages (115 kV or above) to reduce the energy loss which occurs in long-distance transmission.
- **Distribution :** Carries electricity from the transmission system to individual consumers. Distribution substations connect to the transmission system and lower the transmission voltage to medium voltage ranging between 2 kV and 35 kV with the use of transformers. Primary distribution lines carry this medium voltage power to distribution transformers located near the customer's premises. Distribution transformers again lower the voltage to the utilization voltage of household appliances and typically feed several customers through secondary distribution lines at this voltage.
- **Customer :** The end-user of the electricity. Residential and most commercial customers have classic needs and are connected to the secondary distribution lines. However, some professional customers whose business requires larger amount of power may be connected directly to the primary distribution level or the transmission level.

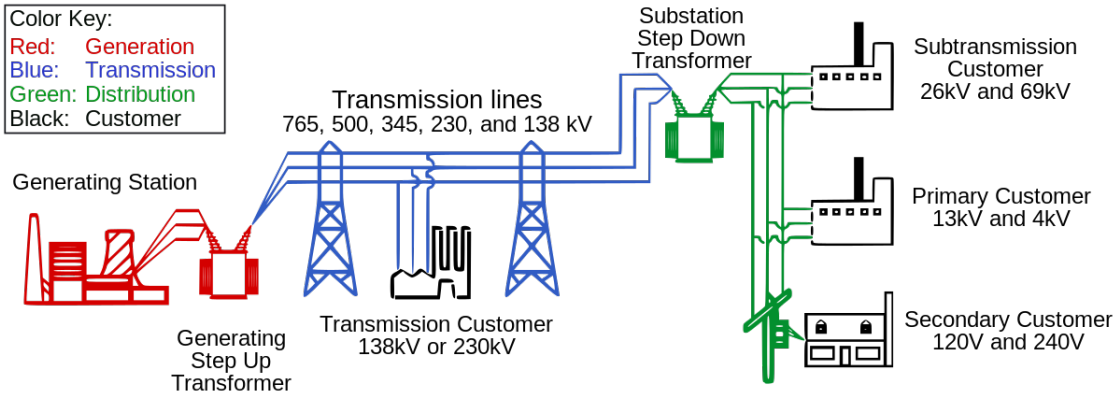


FIGURE 2.8: Structure of the power grid. Source : Wikipedia

Although power grids include heterogeneous structures, it seems appealing to represent them as abstract graphs where edges represent electrical lines and nodes any kind of elements that may be at the end of a line (generators, transformers, substation, loads...). With this approach, it was shown that transmission networks exhibit very different topologies from distribution ones [33] [71] [32]. While the later tend to be tree-like (or locally tree-like) with only few redundant links, the former exhibit a mesh structure. Intuitively, because power grids are hierarchical structures, we expect that a failure in the transmission grid would result in far worse damages than a failure in one of distribution systems. Operators have therefore invested more money in increasing the redundancy within transmission than distribution systems.

Such an important difference between transmission and distribution led the scientists to study them apart. Although more and more data are getting public these days, it should be noted that for a very long time very few reliable data sources were available concerning power grids. A consequence is that most topological studies focus on the transmission network which is often easier to obtain than distribution local networks.

2.3.2 Are power grids similar to other networks ?

A first question that one usually tries to answer when studying a graph is to understand whether it is the result of some underlying process or the pure product of randomness. In the case where the network is not random, finding other well-studied networks that exhibit similarities could potentially lead to a better understanding. The following results may not be true for every possible sample of power grids since there exists a large amount of them and some may have specific characteristics. We review here the tendency that emerges when compiling the literature on the subject.

Several measures indicate that power grids (at any voltage level) have little in common with Erdös-Rényi random graphs. This point is quite clear and universal across the literature [70]. For example, the diameters of power grids tend to be much larger than for ER networks. It was also shown that the degree distributions of power grids tend to be exponential. That is, the node degrees within power grids tend to be quite homogeneous.

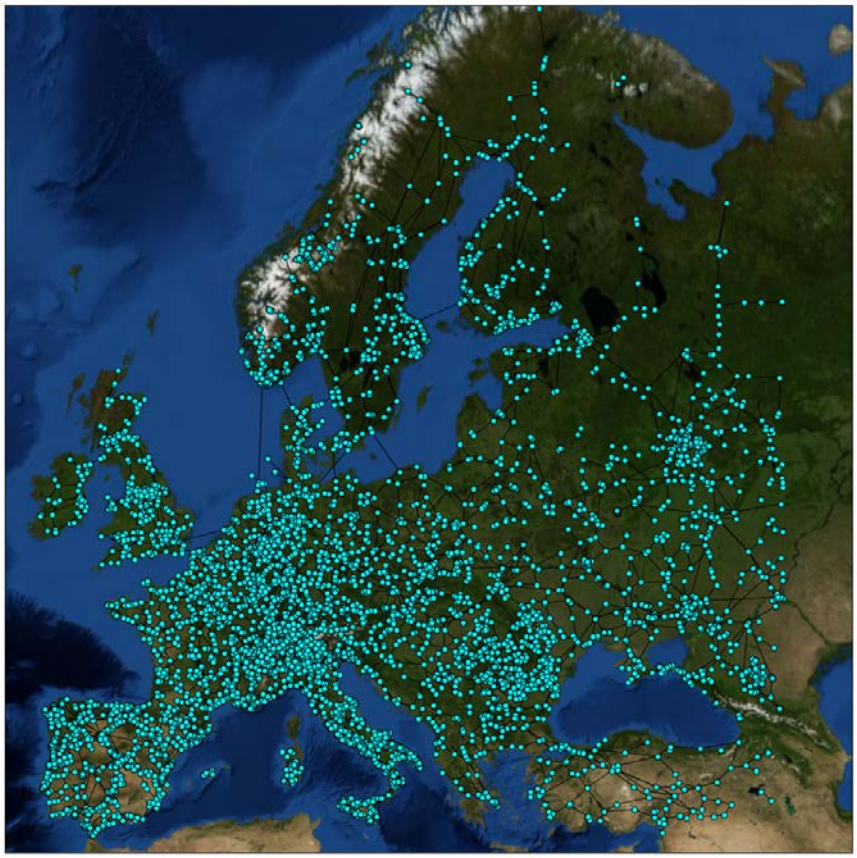


FIGURE 2.9: European high voltage power grid extracted with Gridkit (13871 nodes and 17277 edges).

This particularity is in strong opposition with scale-free networks such as the internet or the WWW, where some nodes have very large degrees.

If power grids are neither random nor scale-free, researchers then tried to check whether they were small-world. In general, the various studies tend not to have a common answer for this question [70]. It is indeed very specific to the samples analyzed and no general conclusion can be drawn. This is especially true for the High Voltage grids, while the Medium and Low Voltage networks seem far from being small-world networks.

2.3.3 Are power grids robust ?

Robustness is the ability of a system to tolerate failures or attacks. Failures are usually supposed more or less random while attacks are the results of a strategy aiming at maximum damages. These two concepts are very different and some networks, such as scale-free networks, tolerate random failures very well while they are extremely vulnerable against hub-oriented attacks. For power grids, the same behavior has been observed, they tend to bear random failures while being vulnerable to high degree or high betweeness oriented attacks.

That being said, there are multiple ways to account for robustness. Static approaches remove a certain fraction of the nodes or edges and look at the size of the giant connected component [88] [97]. As nodes are removed, the

size of the giant component goes to zero at some point. The nature of this transition is of particular interest since it has been shown that, depending on the topology and the attack, a discontinuous (first order) or a continuous (second order) transition can occur.

Percolation in complex networks [14] appears as a convenient framework for studying cascades. Although several works have considered such approaches for power grid robustness, it has been suggested that they were not capturing the reality really well [33]. Indeed, power grids have static topologies but dynamic behaviors occur within this topology since electricity is flowing along the edges (see chapter 6 for dynamics on networks). Not taking this into account may lead to inaccurate conclusions.

Although most power systems are built to sustain the loss of any single component (the so-called "N-1" rule), large blackouts are well-known phenomenon that occur once in a while. Usually, blackouts do not happen due to the loss of some fraction of edges or nodes. They rather are the result of a cascading process when the failure of one component leads to the failure of one or more other components and so on. When a line fails (because of a tree fall, heat, or overcapacity for instance), the power that was flowing along this line is redistributed on adjacent lines according to Kirchoff's law. These lines could in turn become overloaded and fail and so on. Depending on the initial failure, the topology, and the utilization of the lines, a cascade could die out quickly or spread to a non negligible portion of the network. Eventually, the cascade brings the whole network down. Multiple studies have investigated the robustness of power grid with this approach by using betweeness centrality or more electricity-specific measures [79] [72] [85].

2.4 Smart grids

2.4.1 Toward a renewable production

One of the most important goals of the 21st century consists in changing radically the way our society produces energy. The traditional way based on fossil energy is becoming tenuous as fossil deposits are becoming scarcer and scarcer while the demand keeps increasing over the years. Increasing the penetration of renewables into the production portfolio has clearly been displayed as a crucial objective by politicians and scientists. Nevertheless, recent facts from Germany have shown that using more renewables could increase the amount of CO2 released in the atmosphere. As contradictory as it may seem, fast ramping coal plants had to be used in order to compensate for wind power fluctuations.

Renewable energies such as wind or sun power, despite of being "free and clean", pose indeed numerous challenges [77]. First of all, generators based on these energies cannot be scheduled like traditional power plants. They produce when the resource is active, which might not coincide with the moment when consumers desire to use energy (see chapter 3). Furthermore, fluctuations are stochastic and hardly predictable such that a strong increase in flexibility for future power systems is required [35].

The first idea that might come to mind is to store the electricity when it is abundant for discharging it later when it becomes scarce. Although a useful idea, it has been argued that this approach as a large scale solution would not be viable. Electricity is indeed a quantity that does not store really well,

and most storage devices are far from perfect and have a finite number of life cycle. In addition to the huge losses, the costs (installation and maintenance) for such a large scale storage system would be prohibitive.

2.4.2 Demand-Side Management and Dynamic pricing

It appears that changing the way we produce requires that we change both the way we handle electricity and the way we consume it. These parts can no longer be (almost) independent of each other, but rather linked through information. Using sensors deployed across the system for monitoring its state is nothing new, but the information flows from the base to the top while electricity moves in the opposite direction. Imagine that both could flow in both directions : operators could send real time information to end-user and see their reactions, while bi-directional power-flows could open the doors to local generation-consumption cycles with less losses.

The feasibility of such bi-directional systems is still under investigation, especially for large scale systems. Nevertheless, smart meters, which provide the interface between the end-user and the system, are deployed at an accelerating rate. In addition to the fine grain records of the consumptions, these units will be connected to the smart appliances of the house and be able to control them to some extent. The idea is to incorporate the end-user in the system by no longer considering it as a dead load, but rather as some signal-responding entity. End users have usually different kinds of loads. Some appliances, such as TV, lights, or oven for instance should clearly be available on demand. One can indeed hardly imagine the TV going off in the middle of a film because of the grid operation conditions... However, other kind of appliances like washing machines, air conditioners, or heaters for example are delayable to some extent. That is, users might not care at the exact time the machine is started as long as it finishes at a given time. For air conditioners, users are usually only concerned with the temperature of their home when they are present.

It has been suggested that having a dynamic price (DP) for electricity could help reverberate the production conditions on the end users [10] [50] [81] [26]. That is, periods when the expected production is supposed to be larger than the consumption, users could face a small price and conversely for peak consumption periods. Smart meters receive these upcoming variable rates and schedule delayable loads as to take advantage of them. Demand-Side Management (DSM) is currently an active research area where multiple data analysis and machine learning methods are implemented as to make the meters more efficient and autonomous [17] [94].

2.4.3 Electric vehicles

An important challenge comes from the fact that the portion of Electric Vehicles (EV) is supposed to grow in the upcoming years [77]. If, in the future, the whole population is equipped with EV, the induced load on the system will be huge and poses great challenges. For example, it is known that most people commute in a correlated fashion. They go to work in the morning and come home in the evening. It means that, if nothing is done, the amplitude of the already expensive and problematic morning and evening peaks

will increase even more. Furthermore, it has been pointed out that unexpected correlated behaviors (football match, clearance sales...) could create bottlenecks in the grid. That is, some unexpected crowd drives to some specific location and put their vehicles in charge, thus creating a load that the system might not be able to handle.

If EVs seem to create problems, they could also provide solutions. Because an EV is basically a moving battery, it could be used as an emergency reserve. This can be the case for the owner of the EV that faces an outage and still want to consume meanwhile. But this could also be used for grid stability through the Vehicle to Grid (V2G) technology. V2G consists in allowing EV, against remuneration to the owners, to discharge into the grid when there is an emergency [39]. Although a beautiful idea, it has been argued that rebalancing production and consumption with V2G was not realistic. However, V2G could very well provide a mean for frequency regulation.

2.4.4 Toward a more distributed system

Actual bulk power generation sites are located far away from consumption areas yielding important losses during the transport. They are also large and not many such that they constitute strategic points whose failure would be catastrophic. By exploiting renewable energy sources closer to the consumption, losses could be reduced. These Distributed Energy Resources (DER) could also provide more flexibility and resilience to the system because the loss of a single DER would have much less impact than the loss of a plant.

Microgrids are small subsystems composed of DER, loads, and storage units that are connected to the main grid through the point of common coupling (PCC) [98] [21]. The idea behind microgrids is that they could disconnect from the main grid and operate in islanding mode [41]. Such islands could then continue to operate on their own without pumping power from the endangered grid. These islandings could decrease the load on the main grid meanwhile the operator restores a stable state. There has been a lot of studies on the possibilities that come with microgrids. Super-islanding methods, for example, assume that microgrids are able to disconnect in groups and share energy together [72] [21]. In this manner, microgrids which have an excess of production could feed other microgrids where there is a production deficit.

Electric vehicles, microgrids, or prosumers are representative concepts of the emergence of a new decentralized system. Studying these systems where divergent interests come into play has been tackled with success by using game theory [80] [63]. Coalitional game theory for example can be used to study whether multiple microgrids would have a common interest in forming a super-island. The core idea is that, by forming coalitions, agents achieve better payoffs than when playing alone. However, the mechanism by which a coalition re-distributes the profits to the agents is paramount for its stability. Indeed, if agents realize that another organization would increase their gains, they will probably leave the current coalition. Finding the stable coalition structure that maximizes the social welfare is often the objective although it is known to be a hard problem [5]. Assessing the stability of a given game is one of the main strengths of this

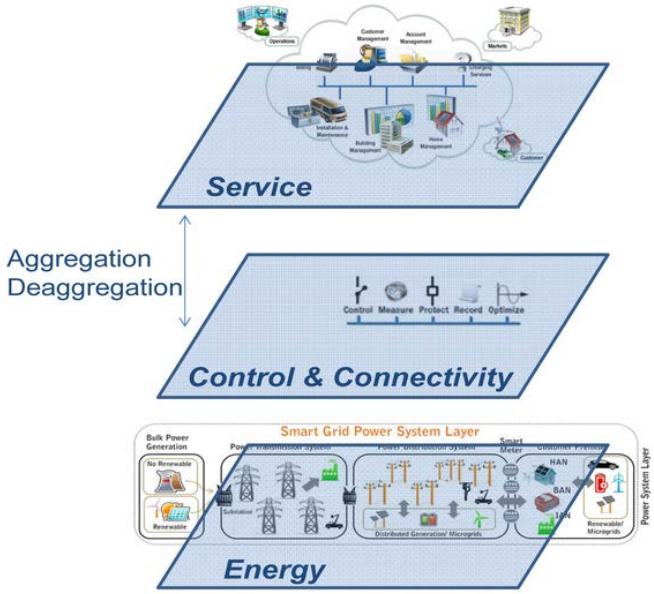


FIGURE 2.10: Smart grid multilayer achitecture. Source: Wikipedia.

framework, and may be one of the reasons for its popularity in the smart grid literature.

2.4.5 Smart grid architecture

Although smart grids are still new systems under active development, a general layered architecture seems to be accepted by the community (see figure 2.10). The bottom layer contains all physical components that generate, exchange, and consume power. On top of this, a control layer uses the information acquired by sensors such as Phasor Measurement Units (PMU) to maintain a stable system. The top layer implements the services such as energy markets. This schema is of course very general and each layer can be sub-divided in multiple sublayers with different time scales.

This very general architecture gives a rough idea about how the different components are organized. However, it does not tell us how the different subsystems should be built. Although the power grids will surely need to adapt to this new system, the communication network has received a lot of attention lately [90]. Its architecture [58], supported technologies [92], and even specific protocols [49] are currently developed and improved.

Smart grids appear as complicated systems that have multiple dependencies. Understanding whether such systems are possible, viable, or robust is one of the main questions in the field and will probably require extensive use of models and experimentation to find answers. Because human beings are part of this system, complex multi-disciplinary models mixing power engineering, complex system theory, economics, game theory, sociology, psychology, and so on, need to be developed. The most basic inputs might indeed not have the expected large scale effect. If attention is not paid to understanding the system in depth, unexpected and devastating effects may occur. Dynamic pricing for example provides some sort of control on

the consumption, but the mechanism by which individuals respond to such prices is way more complex than for typical dynamic systems. Wanting to delete an evening peak by increasing the prices and scheduling lower rates later in the night, can very well result in delayed consumption peaks in the middle of the night when the production is at its lowest.

2.5 Discussion

We can see that the smart grid, in order to be a viable solution, would require several important changes to the actual power grids. As explained in the introduction, we focus in this thesis on the end-user who is expected to change radically. The user of the future will posses small DER units, storage equipments (EV, batteries...), as well as smart appliances optimizing their consumption according to variable rates sent by the operator. Obviously, these agents will behave in specific and still unknown ways.

This chapter presented important notions for understanding the rest of the thesis. Due to space constraints, we had to make a drastic selection on what to include in this opening chapter, such that it only gives a high-level overview of the fields. The interested reader is refered to the references to get a more precise idea about these topics.

We now turn to the first part of this thesis related to prosumers and electricity markets. In chapter 3 of Part 1 we discuss more extensively the concept of prosumers and provide the formalisms and methods to support such environments.

Part I

Formation of Stable Coalitions for Electricity Markets

Structure of Part 1

The first part of this thesis is devoted to study the formation of coalitions of prosumers in the smart grid. The objective is to enable prosumers to sell their extra-production on electricity markets. We propose in this part an aggregation method based on the correlation between the prosumers. The main idea consists in grouping uncorrelated agents together in such a way that the obtained coalitions have a high productivity with a relatively low risk.

This part is organized in three chapters. Chapter 3 presents the concepts of prosumers, energy markets, aggregators, and correlation graphs that we will use extensively later on. Chapter 4 constitutes the core of this part and presents our work on the simulation of prosumers and the formation of coalitions. This chapter is basically a summary of two publications that we made in 2015 [28] and 2016 [29]. Finally, chapter 5 presents ongoing work on dynamic correlation graphs. Although it is not published yet, we chose to include this more recent work in this part because it is a natural extension of the work done in chapter 4 on static correlation graphs.

Chapter 3

Prosumers and Market Aggregators

In this chapter, we introduce the required concepts that we will use in chapter 4. We explain in section 3.1 what prosumers are and their role in the smart grid environment [31]. We will show that prosumers might be able to sell their own production, when available, on electricity markets [74]. Section 3.2 explains how these specific environments work. A point that we argue in this thesis is that, because of their size and stochasticity, prosumers alone might not be allowed to enter these markets. We wish to show in this first part of the thesis that proper aggregations of prosumers [58] are preferable both in terms of production and reliability. Therefore, we give in section 3.3 some notions about market aggregators and portfolios. We specifically review in details the Markowitz's portfolio theory that exhibits similarities with our approach. We also point out limitations of the portfolio theory that motivated us to find alternative approaches. In section 3.4 we introduce correlation graphs as a way to represent the correlation structure of a pool of variables by using graph theory related tools. These types of graphs will be used in our work presented in chapter 4.

3.1 The Prosumers

Nowadays, power grids are essentially top-down hierarchical architectures with few centralized power plants. This is especially true in France where a large majority of the production comes from nuclear power plants (see figure 3.1). In such a system, the consumption for the next days is predicted and production is scheduled accordingly with safety margins. The end user is then completely passive and consumes when he needs to. Because people tend to consume more when they are at home, this leads to well-known consumption peaks occurring in the evening and, to a lesser extent, in the morning. This has numerous consequences among which the fact that the plants and the networks are dimensioned according to these peaks, which is expensive.

One of the major goals of the smart grid is to increase significantly the share of renewables in the production portfolio [92]. In addition to their stochastic natures, DER usually cannot be scheduled and only produce when the resource is active, which may not coincide with the consumption peaks. This is striking when thinking about solar power. The production peak (expected around 12 am to 3 pm) lies right in-between the morning and evening consumption peaks [64] (see figure 3.2). At this point, there

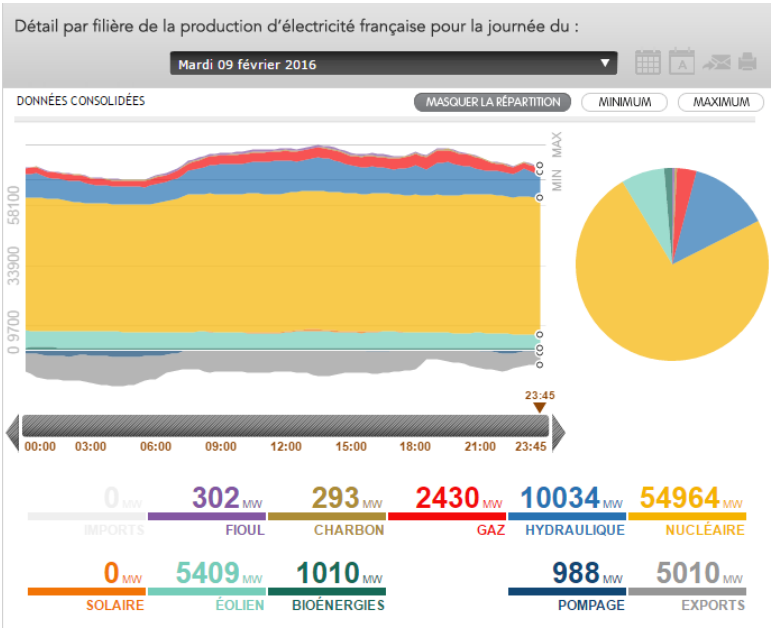


FIGURE 3.1: Screenshot from RTE website. Repartition of the production for a random day.

are two ways of thinking: either acting on the production side by deploying storage or increasing the capacities and the interconnection, or acting on the consumption side.

Thanks to the ability of the smart grid to vehicle information, the grid operators have the possibility to transfer the production conditions on the end users. Dynamic pricing and demand-side-management are indeed technologies that could sensitize consumers to the real time state of the system. Elastic loads could hence be delayed to off-peak periods saving money on both sides [10] [50] [81].

But the idea goes even further. It is indeed likely that, at some point, people will invest in small personal or residential DER. At the time this thesis is written, it is already an ongoing process with solar arrays deployed on roof tops every day. The idea is simply to reduce the electricity bills by taking advantage of the free resources, that is, using one's own production to satisfy their consumption. Here, we will refer to these agents, that both consume and produce, as *prosumers* [78]. If a prosumer produces more than he consumes, there are a few possibilities :

- Consuming more by delaying some loads initially scheduled further in time.
- Storing the surplus of production for using it later.
- Selling the surplus.
- Spill the surplus (turning off some DER).

Obviously, the last possibility is the one that a prosumer will use in last resort since it is wasting resources. The first two are totally legitimate solutions although they might not always be possible. A prosumer might

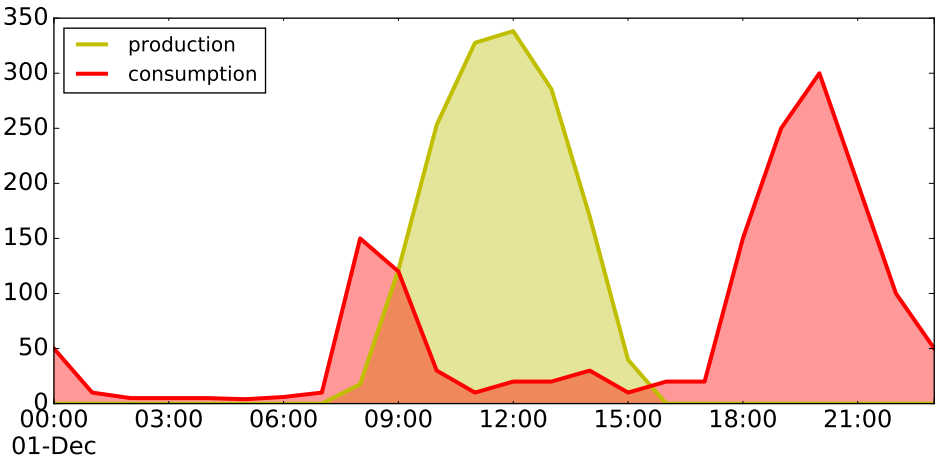


FIGURE 3.2: Solar production (yellow) and consumption of user (red). Peaks of production do not coincide with consumption peaks.

indeed have no load to delay and a full battery. Additionally, since storages are not perfect, a prosumer might be willing to sell power even if its storage is not full, depending on the current selling price.

There is a clear necessity to design comprehensive information systems such that the prosumer operations are supported, simplified, and automated as much as possible. It is also crucial that such agents do not endanger the system stability by tacking crazy actions or using it unexpectedly. Building models and running simulations could help preventing such problems.

3.2 Electricity markets

In this section we study how prosumers could sell their production surplus. In a competitive environment, such operations happen on a market where sellers post their offers and buyers place bids on the offers that they want to purchase [74] [47]. Although markets are common systems of our society, the emergence of energy markets, and more precisely electricity markets, is recent [36].

A possible reason is that electricity is, by nature, difficult to handle. Indeed, it is difficult to store and route properly. This means that electricity has to be available on demand and that measuring precisely who sells what to who might be difficult. Furthermore, since demand and supply vary continuously, the system operator need to coordinate the dispatch of generating units to meet the expected demand of the system. If there is a mismatch between supply and demand the generators speed up or slow down causing the system frequency to increase or decrease. In most power systems all units should be synchronized to a common frequency (usually $50Hz$ or $60Hz$). If the frequency deviates too much from the reference value, severe damage could occur to the equipment. In such a situation, the system operator will act to add or remove either generation or load as to get back to the

stable state. The fact that electricity behaves in such complex ways makes electricity markets very specific environments.

A first distinction should be made between retail electricity markets and wholesale markets. A wholesale electricity market exists when competing energy producers offer their electricity output to retailers. The retailers then re-price the electricity and take it to the retail market. The retail electricity market enables end-users to choose their supplier from competing electricity retailers. It is common that end users of the retail markets face fixed prices contrariwise to the wholesale markets. In these cases, customers have no incentive to take into account, through their consumption, the production conditions that impact the wholesale markets. Since we are studying how prosumer could sell power to markets, we consider here wholesale electricity markets.

The price in the day-ahead market is generally determined by matching the offers to the bids as to obtain a classic supply and demand equilibrium price for all hourly intervals. Obviously, such a market relies on predictions and participants are therefore exposed to risk. This risk can be either related to the price volatility, or the volumes exchanged. At times of peak demand or supply shortages, the wholesale electricity prices can have indeed very high volatility. The particular characteristics of this price risk are highly dependent on the physical fundamentals of the market such as the mix of types of generation plant and relationship between demand and weather patterns.

Volume risk is related to the fact that participants have uncertain volumes of production and consumption. This becomes especially true when portfolios contain large shares of renewables. Furthermore, market price risk and volume risk are not independent, and one can usually observe a correlation between extreme prices and volume events. This is easily understandable since outages, which can lead to price spikes, are more likely to happen during peak periods when the network operates in tense conditions.

Because of risk, market participants tend to establish contracts in sophisticated electricity markets. The idea behind contracts is that participants try to share the financial risks in a pre-agreed manner. These contracts can take multiple forms such as simple fixed price forward contracts for physical delivery and contracts for differences where the parties agree a strike price for defined time periods. The contract for differences specifies a strike price, such that, in a given period, the wholesale price index can either be higher, lower, or equal to the strike price. In cases when it is higher, the generator will refund the difference between the strike price and the actual price for that period. Likewise, a retailer will refund the difference to the generator when the actual price is less than the strike price. Many other hedging arrangements, such as swing contracts, Virtual Bidding, Financial Transmission Rights, call options and put options are traded in modern electricity markets.

3.3 Aggregators and portfolios

3.3.1 Aggregators in the wholesale market

The wholesale market could enable prosumers to sell their production surplus to other entities. Nowadays, there exist already numerous participants to such markets. But, in a smart grid system, where a large number of end users seek to sell or buy in random ways, a stable market seems pretty unlikely. Managing the interactions and financial transactions of so many small and unpredictable entities is a challenge.

Aggregators are market participants that are responsible for portfolios of generators, loads, and storages. An aggregator can be viewed, from the market point of view, as a single dispachable generation unit. The profits and losses made by the aggregator are then shared in a pre-agreed manner between the entities within the portfolio. On the one hand, aggregators provide a hierarchical organization of the market participants, and on the other hand, they reduce the market volatility. Indeed, having a large and diverse pool of resources is known to provide stability. Although a key point for wholesale electricity market aggregators, this is not specific to electricity and was studied in other contexts such as finance.

3.3.2 Markowitz portfolio theory

The optimization of expected returns to risk is a traditional goal in finance, and a wide literature exists on this topic. It is well-known for instance, that the more risk one is willing to accept, the higher his potential gains. On the contrary, when investing exclusively on low risk assets, one should expect relatively small gains. This trade-off is formalized in the Markowitz' portfolio theory [53] in 1952 for which he was later awarded a Nobel price in economics. More precisely, given a set of assets for which we have some historic data of returns, the objective is to find a linear combination of these assets (the so-called portfolio) which maximizes the expected value while minimizing the variance of the portfolio's return [56].

Under this model, the return of a portfolio is the weighted combination of the constituent assets' returns, while the portfolio volatility is a function of the correlations between the component assets [100]. The fact that risk depends on the correlation between the assets is at the root of the diversification principle. An investor can indeed reduce the risk of a portfolio simply by holding combinations of elements that are not perfectly positively correlated.

A first assumption from Markowitz is that investors are risk averse, meaning that given two portfolios that offer the same expected return, investors will prefer the less risky one. The exact trade-off between expected returns and risk is the same for all investors, but different investors will evaluate the trade-off differently based on individual risk aversion characteristics.

It is possible to have a visual representation by plotting the assets as dots in the risk-expected return space (see figure 3.3). The collection of all such possible portfolios defines a region in this space which left boundary is a hyperbola. In the absence of a risk-free asset, the upper edge of this region

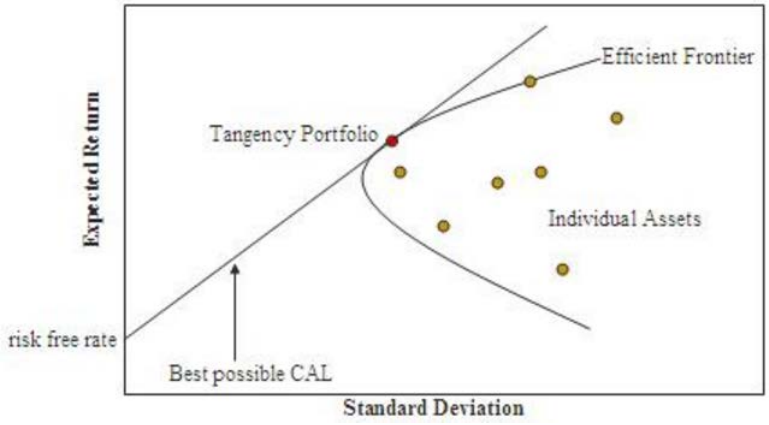


FIGURE 3.3: Efficient frontier given by the Markowitz theory of portfolios. (source: Wikipedia)

is called the efficient frontier (also called "the Markowitz bullet"). Combinations along the efficient frontier represent portfolios for which there is lowest risk for a given level of expected return. This is basically the answer of Markowitz to the problem: if one is able to put a number on his risk acceptance or on the target expected return, the corresponding efficient portfolio is a priori the best option.

Although this model is a cornerstone in finance and risk management, it was criticized for diverse reasons. First, risk is estimated by looking at past market data and quantified by a probability of losses. But the reasons for losses are not explained by the model such that there is no guarantee that the estimated parameters would give good results for predicting the future. This is a major difference compared to many engineering approaches to risk management. Indeed, structural models focus on the relationships between the constituents and Monte-carlo simulations are used to assess the probability of a given outcome. These models can thus be used to compute rare events probabilities, which is only possible with past data if there were several occurrences of these events.

Furthermore, the portfolio theory uses variance to quantify risk, which might be justified under the assumption of elliptically distributed returns. This is the case in the modern portfolio theory where returns are assumed to follow a Gaussian distribution. But this assumption has been criticized by many economists because of its inadequacy with real market returns [12] :

> *"Simply, if you remove their Gaussian assumptions and treat prices as scalable, you are left with hot air."*
>
> Nassim Nicholas Taleb

3.4 Correlation graphs

As shown by Markowitz portfolios theory (see section 3.3), the correlation between the assets impacts the variance of the expected returns. The gaussian assumption of portfolio theory enables analytic results in the form of

an efficient frontier. However, in the case of the prosumers production surplus this assumption might not hold such that we need to find other ways to study the correlations.

Since the work of [52], an interesting approach consists in computing a distance metric based on the correlation coefficients in order to organize the series in a correlation graph. Nodes represent the series considered while the edges are weighted by the metric. Because the metric can be computed for all pairs, these graphs are complete and of little use as is. Historically, the approach used by [52] was to compute a minimum spanning tree as to obtain a hierarchical clustering of the series.

Later on, it was pointed out that, by definition, a spanning tree could not capture the underlying clustering structure hidden in the correlation graph. In this paper, we use another classical filtering technique called ϵ-graph [25]. It consists in selecting a threshold ϵ, and filtering out edges with smaller weights. As we will see further in this paper, this approach has the advantage of preserving clusters of correlated series.

3.5 Discussion

This chapter introduced the concepts that will be used in chapter 4. We started by defining and explaining why prosumers are so important in the smart grid systems. These agents both consume and produce such that their extra-production might be positive or negative depending on multiple criteria such as weather conditions or consumption habits. Because they use renewables, prosumers tend to have stochastic extra-production that are difficult to predict on the individual scale. A gigantic market allowing these unreliable units to sell and buy on short time-scale seems quite unstable and difficult to manage. This is especially true considering the strong stability constraints that weight on power grids.

We thus believe, and we will show in the next chapter, that forming aggregations controlled by a central entity could be a solution for having clearer and more stable markets. Aggregators select prosumers and form groups that they control. They are responsible for all the market operations and are rewarded with a share of the financial gains of the coalition. This approach is similar to the portfolio theory presented in this chapter where the goal is to find a linear combination of assets that maximizes the expected return for a given risk acceptance. Nevertheless, we will see in the next chapter that production distributions are usually not gaussian, such that we propose another formation algorithm that will use correlation graphs.

Chapter 4

Forming Stable Aggregations of Prosumers for Electricity Markets

This chapter uses the concepts introduced in chapter 3 and explains the formation of stable aggregations of prosumers for electricity markets. Section 4.1 gives the details about how we simulated the production and consumption of these agents. Section 4.2 introduces the market model that we used, and section 4.3 explains the algorithm that we built for forming aggregations. In section 4.4, we give some examples and study the performance of the algorithm. The work presented in this chapter was published in [28] and [29].

4.1 Simulation of prosumers

In this first section we wish to simulate the production and consumption of the prosumer agents presented in chapter 3. Using pure random distributions is a direct and simple possibility but is probably not very realistic. In order to get a better understanding about prosumers, we should start by collecting and analyzing data on these agents.

4.1.1 Existing production/consumption data sources

An essential component of the smart grid is the smart meter which makes the interface between the end user and the rest of the system. Smart meters coupled with sensors measure quantities of interest (like instantaneous consumption), receive information from the grid (electricity prices for instance), and take actions accordingly (demand side management programs). Smart meters are currently and gradually deployed, and will probably provide interesting datasets to work on. Unfortunately, at the time this thesis was started, production and consumption data for prosumers over a large region were not yet available to our knowledge. Some interesting experiments are nonetheless being conducted and data are progressively made public [38] [8].

4.1.2 Weather data

In this section, we use weather quantities like wind speed or solar radiance as alternative data for generating realistic production and consumption series. Fortunately, these kinds of data are easier to find, and since the development of small personal weather stations, their geographical granularity keeps increasing. Since these quantities depend both on time and location, we discretise time into slots and space into zones in the following (see block 1 of fig. 4.1). A zone is simply a portion of the considered region of study for which we sampled data. Therefore, if prosumers i and j are positioned on the same zone, they are exposed to the same weather (see figure 4.2). Adding some intra-zone noise can easily be done though not considered in this paper. We denote by $P_i(t)$ the instantaneous available extra-production of agent i at time t :

$$P_i(t) = P_i^P(t) - P_i^D(t) \tag{4.1}$$

Where $P_i^P(t)$ represents the total production of agent i at time t and $P_i^D(t)$ its consumption at time t. In other words, $P_i(t)$ represents the instantaneous surplus of power that agent i is willing to sell at time t. We simulated these traces by considering separately P_i^P and P_i^D. For a prosumer i, it is possible to write both quantities as a sum over the distributed energy resources (DER_i) and loads ($load_i$) of i :

$$P_i^P(t) = \sum_{k \in DER_i} P_k(t) \tag{4.2}$$

$$P_i^D(t) = \sum_{k \in load_i} P_k(t) \tag{4.3}$$

For simplicity, in this paper we only consider wind-turbines (WT) and photovoltaic panels (PV) as possible DERs for the agents ($DER_i = WT_i \cup PV_i$):

$$P_i^D(t) = \sum_{k \in WT_i} P_k(t) + \sum_{k \in PV_i} P_k(t) \tag{4.4}$$

We denote by $\nu_i(t)$ and $\Psi_i(t)$ the wind speed (in $m.s^{-1}$) and the solar radiance (in $W.m^{-2}$) at the location of agent i and at time t, so that :

$$P_i^P(t) = \sum_{k \in WT_i} \mathcal{F}_{WT}(\nu_i(t)) + \sum_{k \in PV_i} \mathcal{F}_{PV}(\Psi_i(t)) \tag{4.5}$$

Where $\mathcal{F}_{WT}$ (resp. $\mathcal{F}_{PV}$) is the power curve for the wind-turbines (resp. photovoltaic panels).

4.1.3 DER models

Wind Turbine model

As explained above, wind turbines are incorporated through their power curve $\mathcal{F}_{WT}$ that maps an input wind speed to a produced power quantity. We are thus looking, in this section, for such a function. First of all, it is known that, within functioning range, the output power P is a function of the cubed wind speed ν^3. This means that a small increase of wind speed

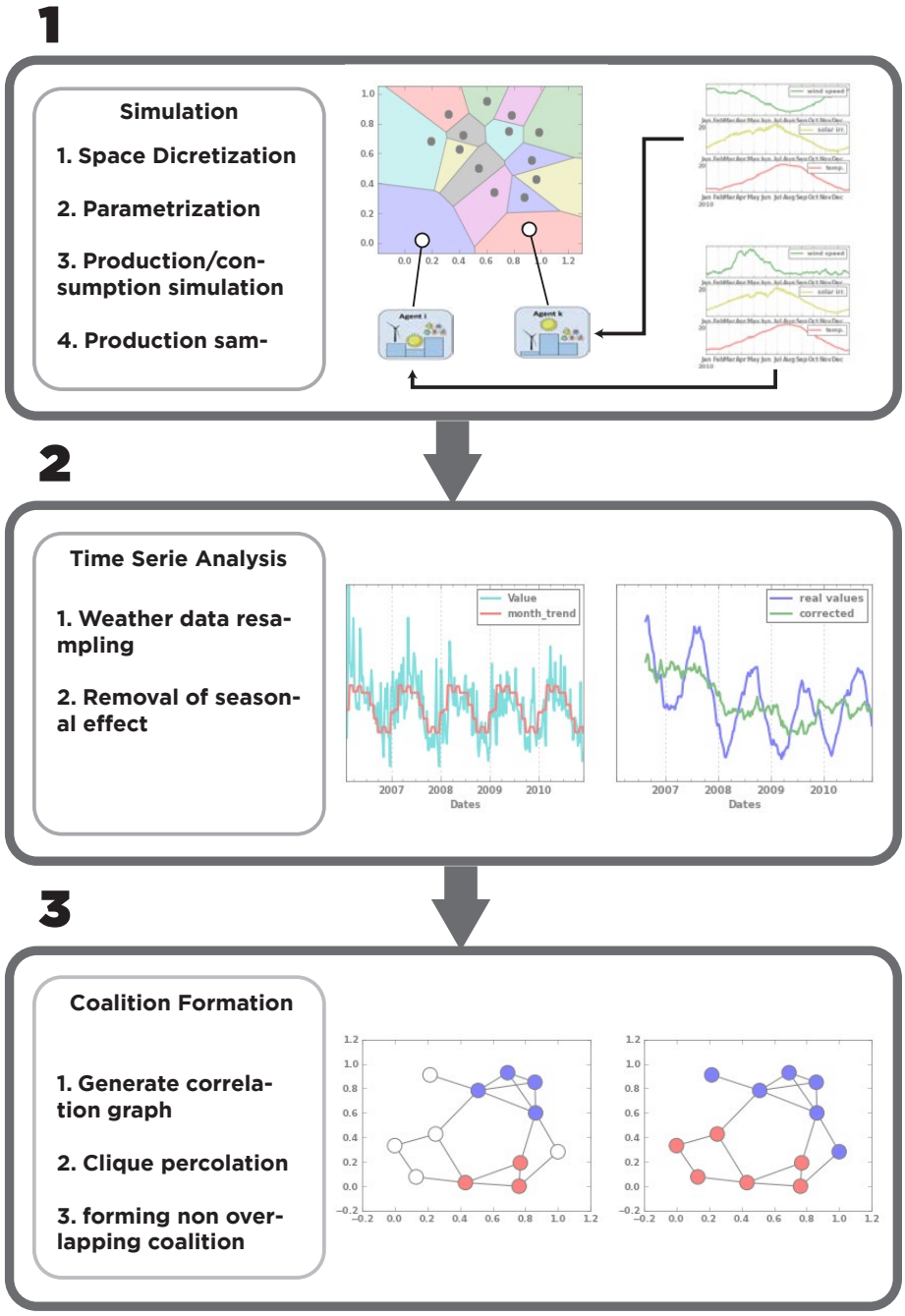


FIGURE 4.1: Simulation process scheme.

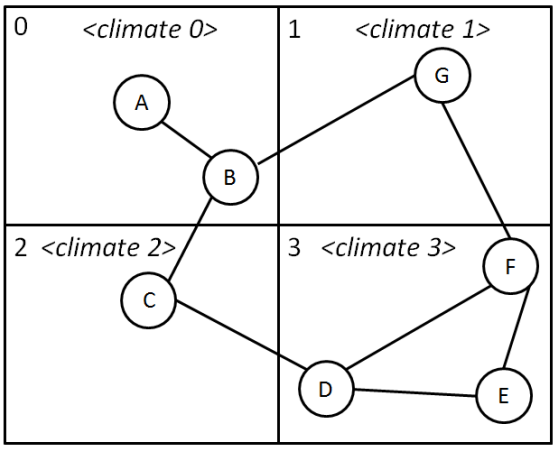


FIGURE 4.2: Schema of the space discretization. Agents A and B for example are both located in zone 0, and thus exposed to the same weather.

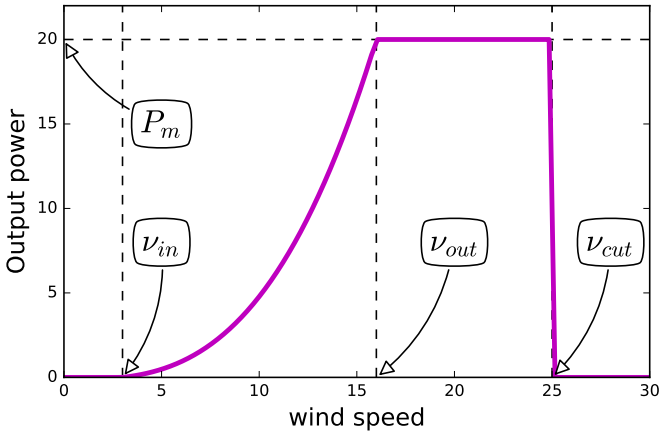


FIGURE 4.3: Wind power curve model.

yields a large increase of output power. However, it is the fact that not all wind turbines are identical and that there exist multiple designs with different characteristics such as the height, the length of the blades and so on. Since wind speed is usually correlated with altitude, higher turbines tend to produce more than smaller ones. In the same manner, the rotor swept area as well as the air density are parameters of interest for designing a precise power curve.

Furthermore, the orientation of the turbine according to the orientation of the wind also impacts the production. There are already too many parameters and information to feed in the model. Indeed, since our focus is not on precise DER modeling but more general, we seek more simple and scalable models with few parameters. The turbines were thus modeled by the following parameters (see figure 4.3 for a graphic representation) :

- **Cut-in speed** : At very low wind speeds, the force exerted by the wind on the blades is not sufficient to make them rotate. But, as the

speed increases, they begin to rotate and generate power. The speed at which the turbine first starts to rotate and generate power is called the cut-in speed. This parameter is denoted ν_{in} and typically ranges between $3ms^{-1}$ and $4ms^{-1}$.

- **Rated output power** : Above the cut-in speed, if the wind speed continues to increase, the level of output power rises rapidly. However, there is a point where the output power reaches the limit that the electrical generator is capable of. This limit to the generator output is called the rated output power and denoted by P_m.
- **Rated output speed** : The wind speed at which the rated output power is reached. Above the rated output speed, the design of the turbine is arranged to limit the power to maximum possible level and there is no further rise in the output power. This parameter is denoted by ν_{out} and typically ranges between $12ms^{-1}$ and $17ms^{-1}$.
- **Cut-out speed** : As the speed increases above the rated output speed, the forces on the turbine structure continue to rise and, at some point, there is a risk of damage to the rotor. As a consequence, a braking system is employed to bring the rotor to a standstill. This is called the cut-out speed. This parameter is denoted by ν_{cut} and is typically around $25ms^{-1}$.

In this thesis, we set all speed limits for all turbines to be the same, and only use the rated output power as a parameter representing the turbine maximum production capacity. We now have to decide an equation relating the output power to the wind speed when the latter is in-between the cut-in speed and the rated output speed. We choose the simple form (see figure 4.3) :

$$\mathcal{F}_{WT}(\nu) = \begin{cases} 0 \quad if\ \nu < \nu_{in} \\ \frac{P_m}{\nu_{out}^3 - \nu_{in}^3}\nu^3 - \frac{P_m \nu_{in}^3}{\nu_{out}^3 - \nu_{in}^3} \quad if\ \nu \in [\nu_{in}, \nu_{out}] \\ P_m \quad if\ \nu \in [\nu_{out}, \nu_{cut}[\\ 0 \quad if\ \nu > \nu_{cut} \end{cases} \tag{4.6}$$

Solar Array power curve

As for wind turbines, we need to define a power curve for solar arrays. Such a function would take as input the solar radiance and output a production power. Obviously, if we succeed in writing the input radiance as a power density ($W.m^{-2}$), then almost all the work would be done. The problem is that weather stations extremely rarely record such a quantity, and for the datasets that we used, the best that we could obtain at the time was time series of cloudiness. Cloudiness is a well-known measures ranging between 0 (no clouds) to 8 (sky completely cloudy).

The approach that we used was to consider the cloudiness traces as degradation factors on some perfect solar radiance profile that we need to generate. Therefore, we want that the perfect profile reproduce sun power basic properties like daily periodicity and seasonality. To do so, we used a model called Helios that can be found in [16]. The whole model is not reproduced here for space concerns but we explain the main idea. Basically,

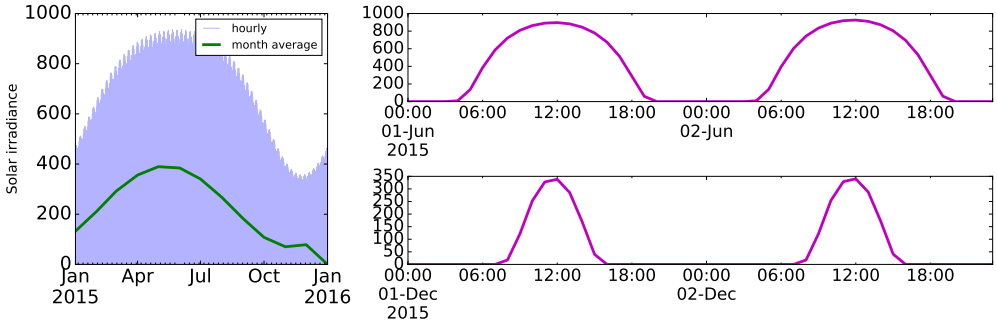


FIGURE 4.4: Solar irradiance at Paris during 2015 obtained with the Helios model.

the Helios model takes as input the location on earth defined as the triplet (latitude, longitude, altitude) as well as a timestamp, and computes the fraction of sun power that arrives at this point at that specific time if there are no clouds. Figure 4.4 shows an example for the city of Paris during 2015. We can see on the left panel that seasonality is reproduced, that is, the solar radiance is stronger during the summer than in the winter. Furthermore, we can observe on the two right panels that daily periodicity and logic are respected (no sun at night), and that the period of sunshine is larger in the summer than in the winter.

The cloudiness can be incorporated to the Helios model through the following equation :

$$\Psi(t) = \Psi^{\star}(t)\left(1 - \frac{3}{4}\left(\frac{\eta(t)}{8}\right)^{3.4}\right) \tag{4.7}$$

where $\Psi(t)$ (resp. $\Psi^{\star}(t)$) is the real (resp. ideal) radiation at time t, and $\eta(t)$ is the cloudiness at time t. If we know the location of a prosumer (latitude, longitude, altitude), and if we have the corresponding cloudiness traces, we can compute the density of solar power (in Wm^{-2}) at this location.

As for wind turbines, modeling solar arrays is a domain on its own. There are indeed multiple parameters to take into account like the efficiency, the angle of the array towards the sun light and so on. Here, we use the following simplified power curve :

$$\mathcal{F}_{PV}(\Psi) = \rho S \Psi \tag{4.8}$$

where S is the total surface of the arrays, and ρ is the array efficiency (around 20%).

Please note that these power curves are very simple models and that more complex and detailed model can be found for example in [51].

4.1.4 Consumption Model

At this point, if we have the right data, we are able to model the production of prosumers using wind turbines, solar arrays, or a combination of both. However, we still have to account for the consumption of these prosumers. Real consumption data can be found in [38] or [] for example. Figure 4.5 shows two examples of meter records for electricity consumption [38]. But,

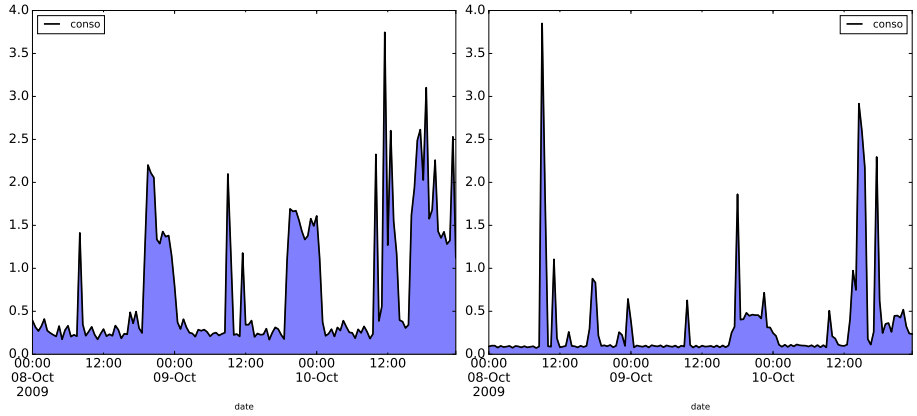


FIGURE 4.5: Two examples of electicity consumption at the user level (smart meter records from the ISSDA dataset).

since these data were not available in the beginning of the present thesis, we also design a model to simulate consumption traces.

When looking at figure 4.5 it appears that electricity consumption tends to be bursty. There is some base load and large peaks at specific hours. Usually, these peaks are in the morning and evening, but this is not always the case. The amplitudes of the base load and peaks differ from user to user and are related to the habits and installation of each user (the house may be heated with electricity or gaz which changes the consumption significantly).

Modeling electric consumption has already been widely tackled in the literature. Models can be basically divided into two main categories : Top-down and bottom-up approaches. Top-down techniques take aggregated consumption data as inputs and try to estimate individual consumption patterns while bottom-up methods use a fine modeling of users consumptions as to obtain realistic aggregated consumption curves. In this paper, we used a bottom-up model since the end user, or relatively small aggregations of end users, are in our interest. The main objective was to capture both daily patterns and seasonal variations of the consumptions. We assumed an additive model where the consumption of an agent is the sum of a seasonal heating term that depends on the outside temperature and an electronic consumption term that depends on the hour of the day. By denoting $\tau(t)$ the outside temperature at timestamp t, we can express the consumption $P_i^D(t)$ of agent i at time t :

$$P_i^D(t) = \mathcal{F}_i^{heat}(\tau(t), t) + \mathcal{F}_i^{elec}(t) \tag{4.9}$$

where $\mathcal{F}_i^{heat}(\tau(t), t)$ is the power curve that maps the temperature to a heating consumption, and $\mathcal{F}_i^{elec}(t)$ computes the consumption of agent i (other than heating) at a given hour of the day. In the simulation, all agents have a desired inside temperature T_i, supposed to be a constant for simplification. By using thermodynamic laws $\mathcal{F}_i^{heat}(\tau(t), t)$ can be approximated by :

$$\mathcal{F}_i^{heat}(\tau(t), t) = \frac{B_i}{R_i}\left[T_i - \tau(t)\right] \tag{4.10}$$

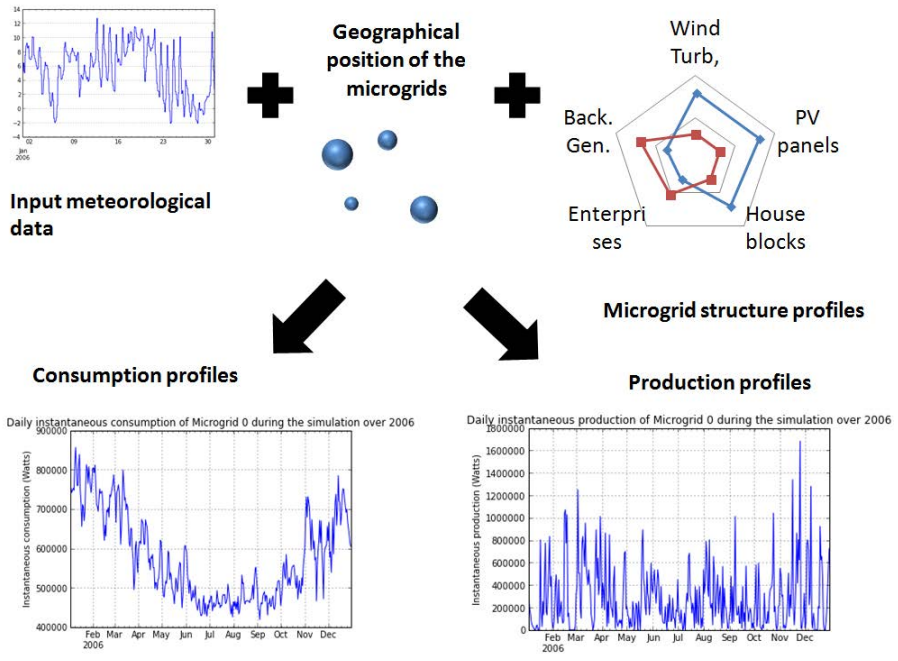


FIGURE 4.6: The proposed model can also be used to model electricity production and consumption for microgrids. Depending on the configuration, an agent can be used to represent DER(s), load(s), prosumer(s), or a microgrid.

where B_i is the surface of thermal exchanges for agent i and R_i is their thermal resistance.

We denote by Ω_i the maximum consumption possible for agent i, which is basically the sum of all its appliances powers. We also denote by $\omega_i(t) = \{\omega_i(t_0), ..., \omega_i(t_{24})\}$ the vector of the average fraction of Ω_i used for each hour. We can therefore write :

$$\mathcal{F}_i^{elec}(t) = \Omega_i(\omega_i(t) + \epsilon) \tag{4.11}$$

where ϵ is a noise term. The vector $\omega_i(t)$ enables us to easily differentiate agent consumption behaviors. Business or residential areas for instance can be easily distinguished with this kind of model.

4.1.5 Results of simulations

Data used

We collected two main datasets for modeling the prosumers [37] [57]. For [37], the data start in January 2006 and end in December 2012, with a sampling frequency of three hours. Note that such a low sampling rate is far from perfect since short time event are very unlikely to be reported (see [4] for more information). Furthermore, the data available are average wind speed and temperature, but the cloudiness is not reported. The data are collected are weather stations located in the main french agglomerations, which gives a quite sparse sampling of the french territory.

For [57], data were available for the year 2010 and contains, among other things, the hourly average wind speed, temperature, and cloudiness collected at US airports. This gives us between 3 and 10 stations per states.

Configuration of the prosumers

In the model, a prosumer i is instantiated by specifying :

- Z_i : The zone on the map where prosumer i lives.
- DER_i The DER set of i that comprises :
  - WT_i : The set of wind turbines owned by i. A wind turbine is configured by fixing its rated output power.
  - PV_i : The set of solar arrays owned by i. A solar array is configured by fixing its surface (all array have the same efficiency for simplicity).
- $Load_i$: The load set of i that comprises :
  - Heating habits computed from the desired temperature T_i and the surface B_i.
  - Other consumptions

Therefore, the notion of prosumer is quite loose here, the agents can be configured as to represent anything from a single wind-turbine for instance ($DER_i = \{WT_0\}$ and $load_i = \emptyset$) to a pure load ($DER_i = \emptyset$ and $load_i = \{L_0\}$) through more complex combinations. This enables us to control precisely the profile of the agent population that we create. Of course, random configurations of the agents is possible if there is no prior on their profiles.

Examples

Figure 4.7 shows the normalized net productions of 3 different prosumers obtained from the model developed above. The curves span years 2007 and 2008 and are daily averaged for clarity. As visible, the combination of the location, DER, and consumption yields very different net production patterns. Some prosumers with low consumption and solar production exhibit very bursty production without visible seasonality (see top subplot of figure 4.7). While others have a clear seasonal pattern with a net production much higher in the summer than in the winter (see bottom subplot of figure 4.7). These prosumers might have a majority of solar arrays as DER and/or a strong heating consumption.

The examples of figure 4.7 show prosumers which almost always produce because they were configured with much more DER than loads. The opposite case, or a more balanced situation where the net production can be either positive or negative are also possible.

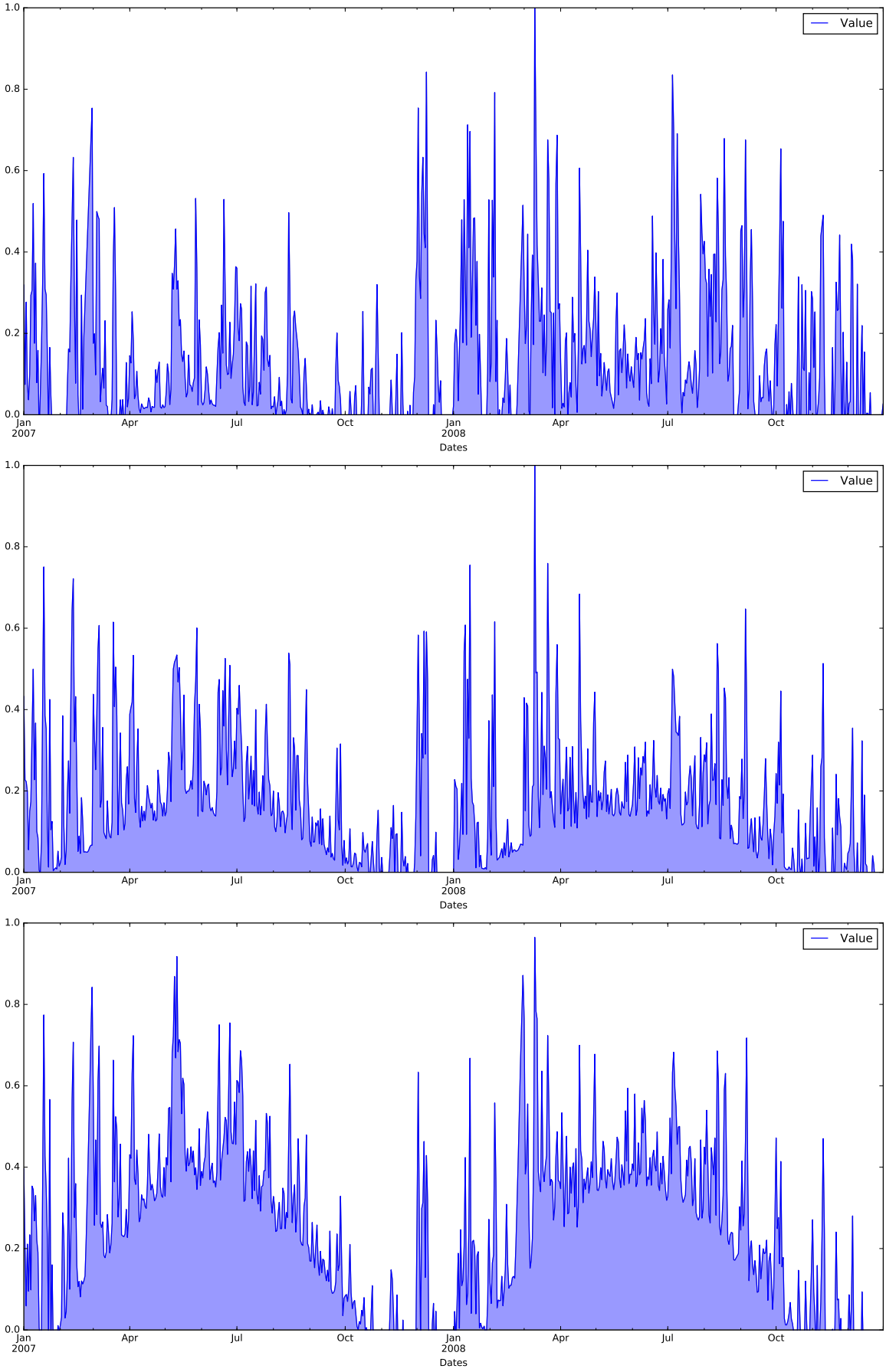


FIGURE 4.7: Normalized net production for 3 different prosumers over 2007 and 2008 (daily average values).

4.1.6 Limitations

In this section we aimed at simulating the production and consumption of prosumers. More precisely, we were interested in being able to generate timeseries of production and consumption given a distribution of prosumers on a territory. We believe that the originality of our approach was to use weather data as inputs of the model rather than standard distributions that would have been less realistic. Depending on weather and time, agents can have positive or negative net productions that we can control with good granularity thanks to this work.

We want to stress out that multiple improvements could easily be done to this model for people having both the time and the will to do so. The number of types of DER could be increased and more realistic power curves could be designed for instance. Besides, we neglected, for simplicity, several factors such as the wind orientation, and the angle of the sun towards the arrays. A higher frequency for the weather data could also lead to interesting work concerning short-time perturbations.

4.2 Market model

In the previous section we explained how we could simulate production and consumption for different prosumers by using weather data. The purpose of this section was to obtain timeseries of extra-production for a pool of prosumers. Depending on the DER they own as well as their personal behavior, these series might look very different (see figure 4.7). In this section, we study how prosumers could enter electricity markets and on what criteria.

We consider a set $\mathcal{A} = \{a_1, a_2, ..., a_N\}$ of N prosumers of the distribution network. Each agent is configured randomly and we simulate its extra-production $P_i(t),\ \forall i \in \mathcal{A}$ for a given period (see chapter 3). Based on these historical values, our objective is now to form groups of prosumers so that the global power production resulting from the superposition of individual's extra-productions be both sufficiently high and predictable. Let $P_S(t) = \sum_{i \in S} P_i(t)$ be the available extra-production of coalition S at time t.

Suppose now that coalition S has to suggest a production value P_S^{CRCT} to enter the market. This means that, during the time S is on the market, it will have to inject in the grid exactly P_S^{CRCT} at any time t and will be rewarded proportionally to this amount, with penalties if it deviates. Obviously, the actual extra-production will not be constant at this value and will oscillate due to intermittences in the production and consumption (see figure 4.8). If S has an available production always greater than P_S^{CRCT}, it is losing some gains since it could have announced a higher contract value. If the production oscillates around P_S^{CRCT}, by using batteries or demand side management techniques, S could be able to maintain its production to the contract value at any time. Nevertheless, if the oscillations are too important compared to the available storage capacity, S will probably break the contract and pay penalties. We can see that there is a return over risk tradeoff here. Coalitions should thus find the right balance between announcing too low and losing some potential gains, and claiming too high and paying penalties.

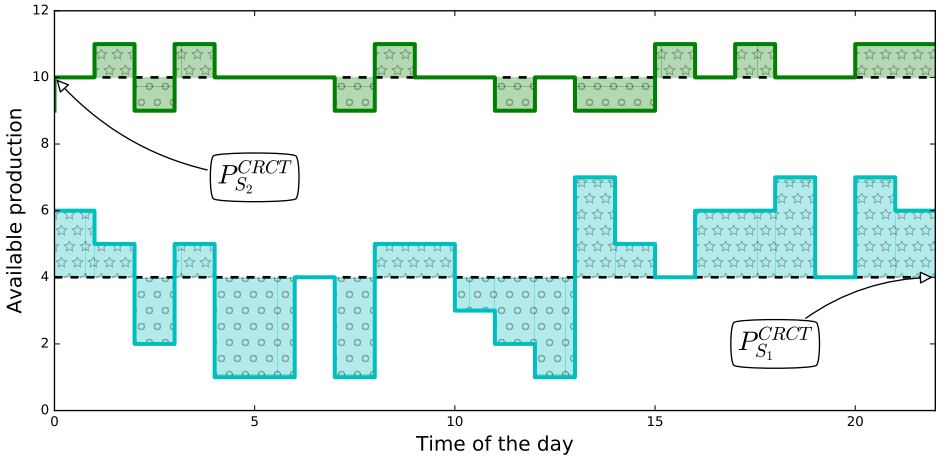


FIGURE 4.8: Available production (straight lines) and contract values (dashed lines) for two coalitions S_1 and S_2. The areas filled with stars represent periods where storage should be charged, while areas filled with circles represent periods where storage should be discharged.

Let us illustrate the rest of the notations and concepts with a simple example. We consider only two agents i and j such that the distribution of their extra-production can be approximated by normal distributions : $P_i \sim \mathcal{N}(\mu_i, \sigma_i)$ and $P_j \sim \mathcal{N}(\mu_j, \sigma_j)$. This is only for explanation purposes as it is of course rather unrealistic in real situations where the distributions are skewed. We can write the distribution of the coalition $S = \{i, j\}$ as $P_{\{i,j\}} \sim \mathcal{N}(\mu_{ij}, \sigma_{ij})$ (see figure 4.9), where :

$$\begin{cases} \mu_{ij} = \mu_i + \mu_j \\ \sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2 + \rho_{ij}\sigma_i\sigma_j} \end{cases} \tag{4.12}$$

, ρ_{ij} being the Pearson's correlation coefficient between P_i and P_j. If the coalition $\{i, j\}$ proposes a contract value P_S^{CRCT}, all instants when $\{i, j\}$ will produce less than P_S^{CRCT} is critical. Indeed, in this kind of situations, $\{i, j\}$ will either have to discharge batteries to keep up with its contract, or pay penalties to the grid. The probability that $\{i, j\}$ is under-producing compared to the contract : $Pr[P_{i,j} \leq P^{CRCT}]$ is thus an important indicator of the coalition's quality. A well-known result for normal distributions is that the cumulative distribution function can be written as :

$$Pr[P_{ij} \leq P_S^{CRCT}] = \frac{1}{2}\left[1 + erf\left(\frac{P_S^{CRCT} - \mu_{ij}}{\sigma_{ij}\sqrt{2}}\right)\right] \tag{4.13}$$

, where erf is the error function : $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

The contract a given coalition is willing to take depends on its capacity to compensate for under-producing (using batteries, backup generators...), and its risk acceptance. Selecting the right contract value appears thus as an interesting problem on its own that we plan to investigate in future works. In order to keep the present paper in a reasonable length, we simplify the contract value selection problem by giving some responsibilities to a third party named the market operator. Their role is to constrain the market entry

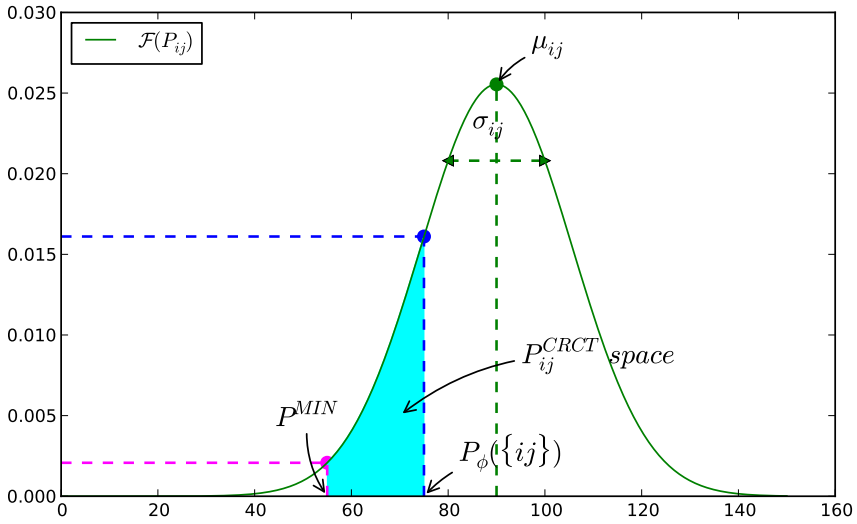


FIGURE 4.9: Probability distribution for the production of a simple gaussian example.

to coalitions able to propose both sufficiently high and sufficiently credible contract values. More formally, let $\phi \in [0, 1]$ be the reliability threshold fixed by the market operator as a maximum value for the probability of under-producing. The highest contract value that a coalition can propose is thus $P_S^{CRCT\star}$ such that $Pr[P_{ij} \leq P_S^{CRCT\star}] = \phi$. In the Gaussian example, it implies that coalition $\{i, j\}$ is announcing (see figure 4.9) :

$$P_S^{CRCT\star} = \mu_{ij} - \sqrt{2}\sigma_{ij} erf^{-1}(1 - 2\phi) \tag{4.14}$$

This is the best contract value that the coalition S can afford for a given stability policy ϕ of the market operator. Figure 4.10 shows how $P_S^{CRCT\star}$ evolves according to the reliability parameter ϕ. For illustration, the range of ϕ values is shown from 0 to 1, but in practice, only small values of ϕ really make sense : $\phi = 1$ for instance means that coalitions can announce absolutely anything since the probability of producing less than any contract value is necessarily less than one by trivial definition of a probability. As shown on figure 4.10, coalitions with high expected productions but presenting a high unpredictability are penalized and can only afford small contracts. The market operator also specifies a lower bound P^{MIN} on the contract values as not to overload the market with unrealistic small coalitions. We thus characterized a valid coalition as one satisfying the two conditions :

$$\begin{cases} Pr[P_{ij} \leq P_S^{CRCT}] \leq \phi \\ P_S^{CRCT} \geq P^{MIN} \end{cases} \tag{4.15}$$

On figure 4.10, P^{MIN} is fixed to 2 units for illustration purpose. For $\phi = 0.1$, only blue triangles and cyan circles coalitions are valid while red diamonds coalition is not.

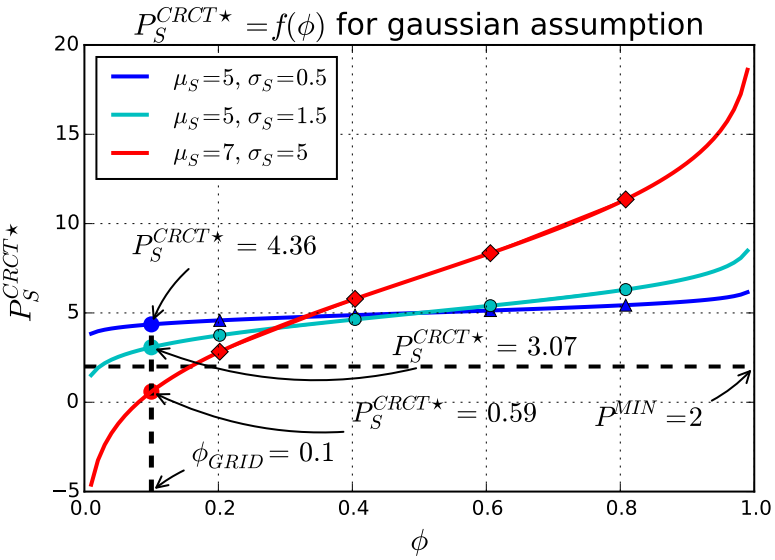


FIGURE 4.10: $P_S^{CRCT\star}$ depending on reliability parameter ϕ for Gaussian distributions (see equation 4.14). Blue curve with triangles stands for a coalition S with an expected production of 5 units and a standard deviation of 0.5. Under a grid policy of $\phi = 0.1$, it is able to announce a contract value of $P_S^{CRCT\star} = 4.36$. The same coalition in term of expected production ($\mu = 5$), but with a higher variance ($\sigma = 1.5$, cyan curve with circles) can only afford a smaller contract value of $P_S^{CRCT\star} = 3.07$. The red curve with diamonds stands for a coalition with a higher expected production ($\mu = 7$), but with a very high unpredictability ($\sigma = 5$). For low values of ϕ, this coalition is thus heavily penalized and can only afford a contract of 0.59 units. Under grid policy ($\phi = 0.1, P^{MIN} = 2$), this last coalition is thus not allowed to enter the market (red dot below the horizontal dashed line).

The Gaussian assumption of this small example is convenient as it allows us to write $P_S^{CRCT\star}$ analytically. Nevertheless, such an assumption is rather unrealistic in practice. In the following, we keep the same framework but release this Gaussian assumption unless the contrary is specified (see eq. 4.20).

4.3 Aggregation method

In this section we study the formation of prosumer aggregations such as to maximize the available production on the market. That is, from the pool of agents and a given number of aggregators that want to participate in the market transactions, we wish to form the corresponding number of coalitions that optimize the expected production to risk ratio. For this purpose, we introduce a utility function based on the contract value explained above.

4.3.1 Utility function

Here, we use the notions of contract values and valid coalitions developed in section 4.2 in order to design a proper utility function. The contract basically indicates the rate at which a coalition has to inject power in the grid. It seems then natural that coalitions are remunerated proportionally to their contract values $gain(S) \propto P_S^{CRCT\star}$. More precisely, if λ is the unitary price rate for electricity, a coalition S injecting $P_S^{CRCT\star}$ in the grid during a period $[t_0, t_k]$ earns :

$$gain(S) = \int_{t_0}^{t_k} \lambda P_S^{CRCT\star} dt = P_S^{CRCT\star} \int_{t_0}^{t_k} \lambda dt \tag{4.16}$$

since $P_S^{CRCT\star}$ is supposed to be a constant rate over the contracted period). Unfortunately $gain(S)$ is not a concave function of the coalitions' sizes, meaning that coalitions can grow as large as the number of agents allows it, without any counterbalance effects. Such a model, that virtually allows infinitely large coalitions and contract values, is in practice not realistic. There are indeed costs (communication costs for instance) that increase with the coalitions sizes. We take this observation into account by rescaling the utility of a coalition S by its size in term of number of agents ($|S|$):

$$\mathcal{U}(S) = \begin{cases} \frac{1}{|S|^\alpha} \frac{P_S^{CRCT\star}}{P^{MAX}}, \ if\ S\ is\ valid, \\ 0, \ if\ S\ is\ not\ valid \end{cases} \tag{4.17}$$

where parameter α controls to what extent the size of a coalition impacts its utility, and P^{MAX} is a normalizing factor. Based on $\mathcal{U}$, the marginal contribution of an agent i can be expressed as $\delta_S(i) = \mathcal{U}(S + \{i\}) - \mathcal{U}(S)$. A coalition S has thus an interest in adding an additional agent i if :

$$\delta_S(i) \geq 0 \quad \Leftrightarrow \quad P_{S+\{i\}}^{CRCT\star} \geq P_S^{CRCT\star} \left(\frac{|S|+1}{|S|} \right)^\alpha \tag{4.18}$$

If $\alpha = 0$, agents are added as long as they increase the contract value of the coalition. If $\alpha > 0$, additional agents have to increase the contract value by some factor.

In real situations, the shape of such utility function might be complex with numerous local maximums. We choose in this paper to favor the exploration of the space around balanced regions, i.e regions where coalitions sizes are relatively balanced. Therefore, we use α as a proxy for having both a concave utility function and preferences in the search space. Given a situation with N prosumers and N_{COAL} desired coalitions, we wish to select α such that the utility function tend to favor coalitions of approximately $\bar{N} = N/N_{COAL}$ agents.

In the general case where distributions and correlation structures have no special form, finding an analytical expression for α appears complicated. To overcome this problem, we seek an approximation for α in a simplified situation where all power distributions are approximated by normal distributions and all prosumers are considered equivalent to a mean agent : $\forall i \in \mathcal{A},\ P_i \sim \mathcal{N}(\bar{\mu}, \bar{\sigma})$, where $\bar{\mu} = \frac{1}{N}\sum_{i\in\mathcal{A}} \mu_i$, $\bar{\sigma} = \frac{1}{N}\sum_{i\in\mathcal{A}} \sigma_i$, and $\forall i, j,\ \rho_{ij} = \bar{\rho}$. In this simplified case, we can express the utility as a function of $\bar{\mu}$, $\bar{\sigma}$, $\bar{\rho}$, $\bar{N}$, and α (see Appendix). We then select $\alpha^\star$ such that :

$$\begin{cases} \left[\frac{\partial U}{\partial |S|}\right]_{|S|=\bar{N}} = 0 \\ \left[\frac{\partial^2 U}{\partial |S|^2}\right]_{\alpha=\alpha^\star} \leq 0 \end{cases} \tag{4.19}$$

Which leads to :

$$\alpha^\star_{\bar{N}} = \frac{0.7\bar{\sigma}(\bar{\rho}-1)erf^{-1}(2\phi-1)}{\bar{\mu}\sqrt{\bar{N}(\bar{\rho}\bar{N}-\bar{\rho}+1)} + 1.4\bar{\sigma}erf^{-1}(2\phi-1)(\bar{\rho}\bar{N}-\bar{\rho}+1)} \tag{4.20}$$

Figure 4.11 shows how $\alpha^\star$ and the utility function evolves according to the mean size of the coalitions $\bar{N}$. Given a policy (ϕ, P^{MIN}) and a pool of N prosumers, we are now able to quantify the quality of any coalition S.

4.3.2 De-correlation graphs

We aim at forming N_{COAL} coalitions $\{S_1; S_2; ...; S_{N_{COAL}}\}$ such that the global utility $U = \sum_{i=1}^{N_{COAL}} \mathcal{U}(S_i)$ is maximized. Since achieving such a goal is a well-known hard problem due to the combinatorial explosion in the number of possible aggregations, we propose in this section a greedy heuristic using decorrelation graphs to solve the problem.

4.3.3 Representing the correlation structure

As seen in section 4.2, the variance of the aggregated production impacts directly the contract values, and depends on the covariances between the agents productions. We argue here that, by having some representation of the correlation structure between the agents, the search landscape for high utility coalitions could be reduced, such that good coalitions are more likely to be found quickly. Usually, this correlation structure is formalized with a covariance matrix or a correlation matrix that contains all the correlation coefficients between the agents : $M = (\rho_{ij})_{\forall i,j\in\mathcal{A}^2}$. By using a metric to map this matrix in a weighted adjacency matrix, it is possible to obtain a graph representation of the correlation relationships between the agents. In the following, we use two opposite distance metrics :

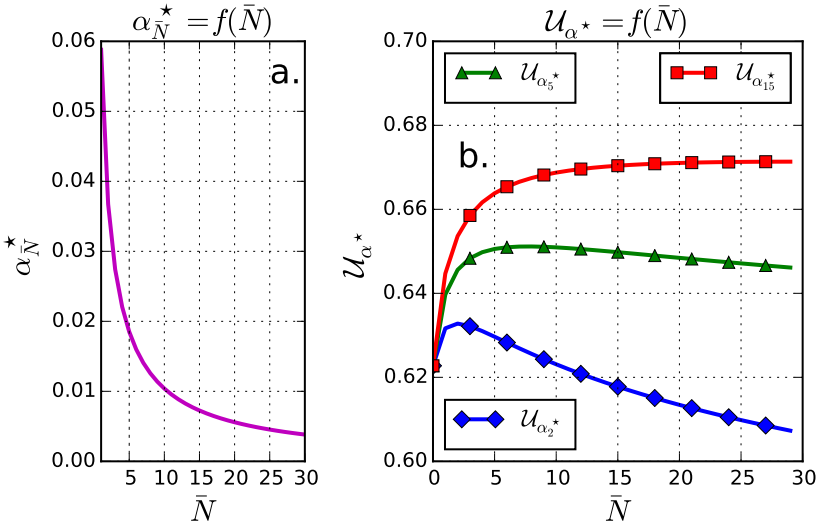


FIGURE 4.11: Gaussian mean approximation ($N = 100$, $\phi = 0.1$, $\bar{\mu} = 5$, $\bar{\sigma} = .5$, $\bar{\rho} = 0.2$). Subplot **a** shows how the parameter α of the utility function should be chosen in function of the mean size of the coalitions (see equation 4.20). Subplot **b** displays the corresponding utility functions for different values of α. Blue curve with diamonds favors small coalitions of 2 agents while the green one with triangles favors 5 agents coalitions. Finally, the red curve with squares has an optimal size of 15 agents.

$$\begin{cases} d^1_{ij} = 1 - \rho^2_{ij} \\ d^2_{ij} = \rho^2_{ij} = 1 - d^1_{ij} \end{cases} \tag{4.21}$$

Clearly, d^1 (resp. d^2) maps two correlated series as close points (resp. distant) while two uncorrelated series are distant (resp. close). These metrics enable us to compute a correlation graph $G_1 = (\mathcal{A}, E_1)$ and a "de-correlation" graph $G_2 = (\mathcal{A}, E_2)$. For any i and j, the weight of the edge e_{ij} is d^1_{ij} in G_1 and d^2_{ij} in G_2. In both cases, we want to keep only the edges which weights are located in the lower tail of the distance distributions. In other words, we want to compute the ϵ-graphs of G_1 and G_2 such that only meaningful edges remain. Selecting the filter ϵ is an important point affecting the landscape search for the coalition formation. Unfortunately, there seems to be no clear consensus in the literature on how to select such a threshold. We will see later in this section that cliques in G_2 are potential seeds for the coalitions. Since we want to generate N_{COAL} coalitions, we need at least N_{COAL} cliques of a given size to start. Besides, since we consider coalitions as disjoint, the starting cliques should be non overlapping. We select our optimal threshold for G_2 as :

$$\epsilon^\star = min_{\epsilon \in [0,1]} \{\epsilon \; s.t. \; |\Theta_k(G^\epsilon_2)| \geq N_{COAL}\} \tag{4.22}$$

where G^ϵ_2 is the de-correlation graph G_2 filtered by ϵ, and $\Theta_k(G)$ is the set of non overlapping cliques of size k in a given graph G. In other words we select $\epsilon^\star$ as the smallest threshold possible such that the filtered de-correlation

graph contains at least N_{COAL} non overlapping cliques of size k. In practice, since finding cliques requires exponential time, we use triangles [84] ($k = 3$ in eq. 4.22) rather than cliques as soon as the number of nodes is not small. Note also that the existence of $\epsilon^\star$ as defined in equation 4.22 is not guaranteed.

4.3.4 Cliques

In [25] the structural roles of weak and strong links on financial correlation graphs is investigated. The author shows that strong links, accounting for strong correlation relationships, are responsible for the clustering, while weak links provide the connectivity between clusters. Indeed, if we consider three items, say a, b, and c such that a and b are strongly correlated and b and c are also strongly correlated, then it is likely that a and c are also strongly correlated. It can be easily shown using the cosine addition formula[1], that if $\rho_{ab} > x$ and $\rho_{bc} > x$ with $x > 0$, then $\rho_{ac} > 2x^2 - 1$). Correlation graphs capture this weak transitivity notion through clusters of correlated series.

Nevertheless de-correlation seems like a more complex concept than correlation in the sense that there is not even a partial notion of transitivity when it comes to it. Therefore, the clustering coefficients of G_1^ϵ is much higher than the one of G_2^ϵ. This can be seen as another formulation of [25] on the structural roles of weak and strong links on financial correlation graphs. Strong links, accounting for strong correlation relationships, are responsible for the clustering, while weak links provide the connectivity between clusters. Searching for clusters in G_2^ϵ and hoping that this strategy will provide a nice coalition structure of internally uncorrelated coalitions seems thus pointless. Consider now a clique in G_2^ϵ, which is a complete subgraph of G_2^ϵ. Since there is a link for every pairs of nodes, we know, by construction, that a clique has a mean correlation and a maximum correlation less than ϵ.

Figure 4.12 shows the distributions of the utility values for cliques of size 3 (triangles) in $G_2^{\epsilon^\star}$ and for all the other possible triplets of agents. It is clearly visible that cliques tend to exhibit higher utilities because of their de-correlation property. Choosing cliques in $G_2^{\epsilon^\star}$ as coalitions seems therefore appealing. Nevertheless, the quality of the results seems to decrease as the sizes of the cliques increase. Indeed, the larger the desired cliques, the more dense $G_2^{\epsilon^\star}$ becomes (see equation 4.22). There is a point where cliques result more from noisy edges than true de-correlation, which decreases the quality of the results. Directly mapping cliques to coalitions by this de-correlation oriented approach is thus not sufficient. It is indeed possible that adding agents to these cliques has the combined effect of increasing the expected production while decreasing its stability. The question revolves around measuring the benefits of this production surplus compared to the disadvantage of having coalition with high volatility. This can be quantified by the marginal benefit in equation 4.18.

4.3.5 Clique percolation

The algorithm takes inputs from :

[1] $cos(a + b) = cos(a)cos(b) - sin(a)sin(b)$

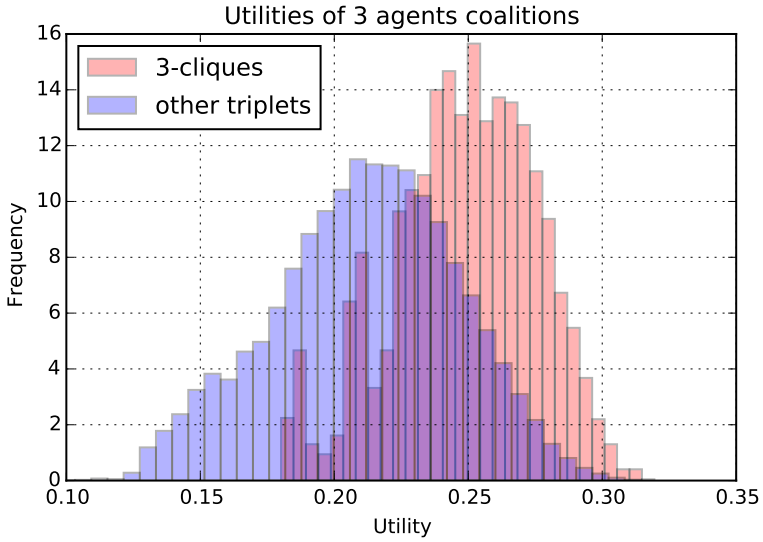


FIGURE 4.12: Histograms of utility values for coalitions of size 3 ($\bar{N} = 3$) in a $N = 200$ prosumers example ($\phi = 0.1$, $\alpha^\star = 0.08$, $\bar{\mu} = 3.9\ MW$, $\bar{\sigma} = 1.9MW$, $\bar{\rho} = 0.69$). Red bars stand for cliques in the decorrelation graph, and blue bars for all the other possible triples. Cliques tend to exhibit higher utilities than random coalitions.

- **The agents :** historical series of available productions P_i,
- **The market operator :** market entrance policy (P^{MIN}, ϕ),
- **The "user" :** Number of desired coalitions N_{COAL} and size of starting cliques k.

The first steps consist in computing the de-correlation graph G_2 as well as the optimal threshold $\epsilon^\star$. Cliques of size k in $G_2^{\epsilon^\star}$ are considered as coalition seeds. The next step is a local greedy improvement over the landscape represented by $G_2^{\epsilon^\star}$. Cliques add alternatively the node $i^\star$ in their neighborhood that yields the best marginal benefit $MAX_{i \in N(clique)} \delta_{clique}(i)$ where $N(clique)$ is the neighborhood of a given clique. See algorithms section in the appendix.

4.4 Results

The algorithm presented in the previous section is supposed to generate a given number of coalitions that have good utilities. As it comprises mainly of a greedy optimization based on local improvements, there is no guarantee that the algorithm finds the global optimum. Since there is, to our knowledge, no state of the art algorithm that aggregates uncorrelated agents in an optimum way, we compare the results with :

- **Random sampling of coalitions :** Coalitions are formed randomly without any other constraint than the desired size. This enables us to have an idea about the distributions of utility values for coalitions of a given size.

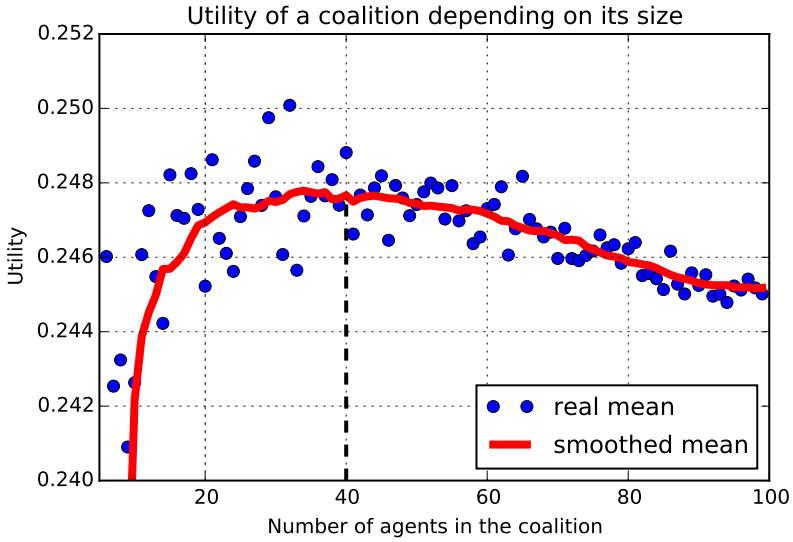


FIGURE 4.13: Utility of random coalitions depending on their size in a $N = 200$ prosumers example ($\phi = .1$, $\bar{\mu} = 3.9\ MW$, $\bar{\sigma} = 1.9MW$, $\bar{\rho} = 0.69$). Blue dots show real mean utility values and the thick red curve its smoothed version by applying a Savitzky-Golay filter. On this plot, $\alpha^\star = 0.006$, and was selected according to eq. 4.20 in order to favor $\bar{N} = 40$ agents coalitions.

- **Random sampling of coalition structures :** Coalition structures are sampled randomly by shuffling and random divisions of the agents. This algorithm (see appendices) uses such a sampling and returns the highest utility coalition structure sampled. We refer to this algorithm as "*random*".

- **Correlated :** Opposite version of our algorithm. It performs a community detection on the correlation graph G_1. The resulting coalitions have thus very high internal correlations. We thus expect this algorithm to perform very bad compared to the others. See appendices.

Before running the algorithms, we need to calibrate the utility function by choosing the value of the α parameter. Recall that the purpose of this parameter is to take into account some constraints on the coalition's sizes if needed. In this paper, we do not have any technical constraints on coalition sizes though we designed the utility such that these could be taken into account. We select the desired size as being $\lfloor N/N_{COAL} \rfloor$ (where $\lfloor . \rfloor$ means floor). Figure 4.13 shows how the mean utility of a coalition evolves with its size when the optimum size is set to 40 agents.

Figure 4.14 shows the coalitions formed with the considered algorithms in the contract value / volatility space. The color map in the background indicates regions where we expect high utilities (red) and the ones where we expect very poor utility values (blue). The bottom right corner, with high contract values and low volatilities, is therefore the region where we wish to form our coalitions. A single coalition is represented by a marker and the color and shape of a marker indicates by which algorithm the coalition has been formed. Besides, the sizes of the coalitions are indicated on

the markers, and the marker size is also proportional to the coalition size. We can see that the utility function results in approximately balanced coalitions. Small yellow markers indicates the gravity centers of their respective coalition structures. The coalitions of correlated agents (green squares) are clearly of poor quality according to our criteria since they can only afford small production contracts, and with a very high volatility. The decorrelated coalitions (blue dots) are closer to the bottom right corner indicating a much better quality in term of productivity over volatility ratio. The black dotted line indicates the mean values for the random coalitions sampling technique. Each small dot stands for the mean position of all sampled coalitions of this given size. Variances are not indicated for readability, but are usually quite large since this sampling only takes the size as a constraint. We can see that as coalitions get larger, they tend to increase on average their contract values, but at the price of a higher volatility. The results of the random coalition structure sampling are shown with the red ellipses that represent the distribution of the gravity centers of the sampled structures. Since the center of the ellipses stands for the mean and each ellipse adds one standard deviation, more than 99% of the sampled gravity centers are within the largest ellipse. The small yellow dot below the ellipses indicates the gravity center of our solution. It is thus visible that our greedy graph based algorithm is able to find a quite good coalition structure in terms of volatility and contract values.

A key point for the coalitions, besides stability and productivity, is their resilience. The resilience of a system refers to its ability to perform its tasks when subject to failures of its components. We consider here the case of random failures of the power electronics of some agents that has the consequence of preventing them from participating in the market. Therefore, the notion of resilience we will use in the following can be seen as the ability of the coalition structures to inject stable power in the grid when some of its internal agents are removed. According to our model, the market operator specified two thresholds (P^{MIN} and ϕ) such that the power injected by every coalition is constrained : $P_S^{CRCT} \in [P^{MIN}, P_S^{CRCT\star}]$. As long as a coalition can propose a contract value higher than P^{MIN}, it is valid and allowed to enter the energy market. We define the resilience of a coalition S as the probability that S produces more than the P^{MIN} threshold : $\mathcal{R}_S = Pr[P_S >= P^{MIN}] = 1 - Pr[P_S < P^{MIN}]$. And we extend this measure to the coalition structures :

$$\mathcal{R}_{CS} = \prod_{S \in CS} \Big(1 - Pr[P_S < P^{MIN}]\Big) \tag{4.23}$$

We consider that prosumers fail randomly, and we denote by $\psi \in [0, 1]$ the fraction of agents that failed. Figure 4.15 exhibits how the resilience of the coalition structures evolves according to ψ (see appendix section *Resilience algorithm* for more details). On the top subplot, P^{MIN} was voluntarily selected relatively low such that the resiliences of the three structures fit on the same figure. When the P^{MIN} requirement increases, the differences between the algorithms also increase as visible on the bottom subplot of figure 4.15. The decorrelated coalitions achieve a more resilient production on the market in the sense that they sustain a higher fraction of node failures.

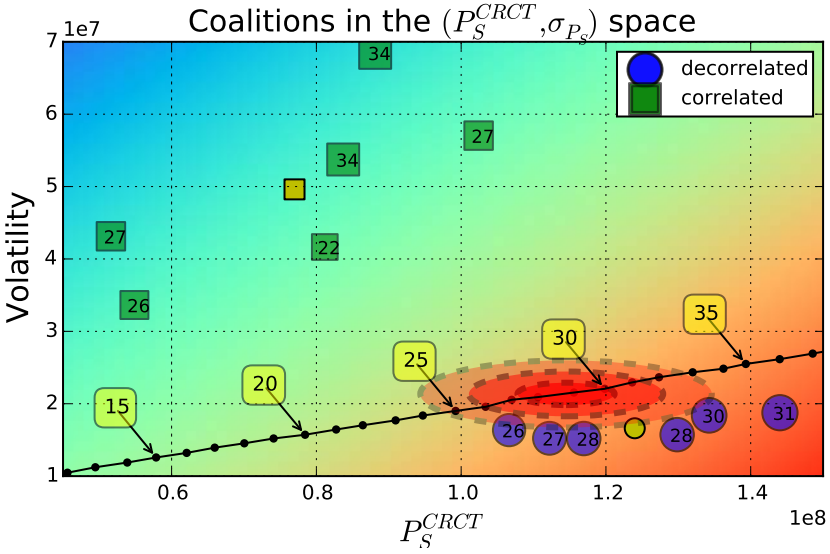


FIGURE 4.14: Coalitions formed in the (contract value, volatility) space. The color map indicates the underlying quality. The closer to red the better (high contract values with small volatility). On the opposite, blue areas show poor quality (small contract values with high volatility). Blue dots stand for the decorrelated coalitions that we formed while green squares show correlated coalitions. The smaller yellow markers stand for the gravity centers of the coalition structures. The black dotted line shows how contract values and volatility evolve when the size of the coalitions increases (a subset of the points are labeled by the size of the coalition they represent). Each point is the average over 10^5 unconstrained draws of a random coalition. We also show the distribution of the gravity centers of random coalition structures with red ellipses (center is the mean, and each ellipse corresponds to one standard deviation).

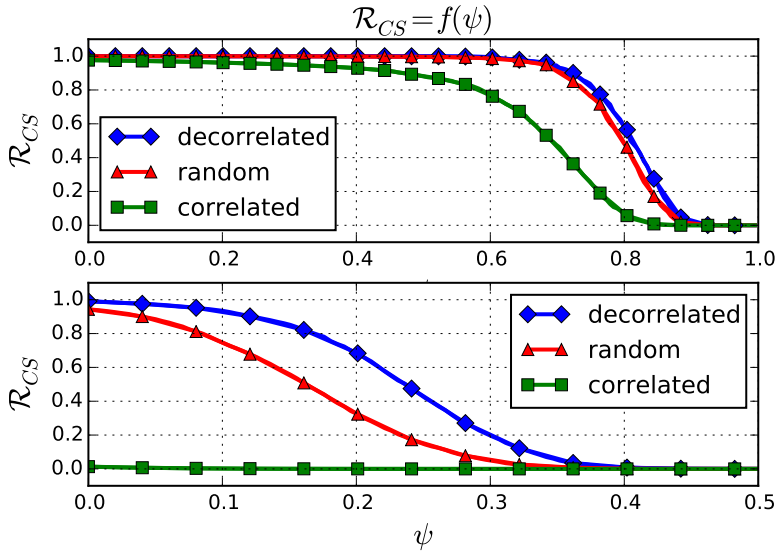


FIGURE 4.15: Resilience of the coalition structures when nodes fail randomly (see equation 4.23) for $P^{MIN} = 10MW$ (top subplot) and $P^{MIN} = 80MW$ (bottom subplot)

4.5 Discussion

Chapter 4 presented a method for energy market aggregators. By seeking diversification in the assets they own, we show that the expected productivity to risk ratios of the portfolios are improved. In this framework, this translates to higher contract values for the coalitions, and therefore higher utilities. We particularly studied the case where a given number N_{COAL} of aggregators were trying to build optimized portfolios. We showed that the variance of the production of a coalition was related to the inner-correlations between the prosumers inside the coalition. Since these agents produce and consume in non-trivial ways, these correlations are far from obvious. Furthermore, because of the diversity of the possible distributions (usually not gaussian) for the productions and consumptions of the agents, we could not use a Markowitz framework for our purposes. Instead, we chose a graph based approach where correlation (and decorrelation) graphs were used to find and expand uncorrelated cliques, which is one of the main contribution of this work. We showed that our algorithm was performing better than other aggregation methods.

Although this approach gave interesting results, we believe that correlation graphs could be used in a deeper manner. In this chapter, (de)correlation graphs were static representations of the (de)correlation structures between the agents. Nevertheless, the correlations between the agents are susceptible to change depending on both the time at which they are computed, and the size of the window on which they are computed. Intuitively, we feel that this lack of dynamics may hide some information, and even worse, may lead to false conclusions. We will see in the next chapter, devoted to a dynamic approach of the correlation graphs, why this could be the case and what are the improvements that we can obtain.

Chapter 5

Correlation graphs dynamics

In chapter 4, correlation graphs were introduced and used to form coalitions of prosumers. In this chapter, we aim at studying these graphs in more depth. The work done in chapter 4 was considering static correlation graphs although correlation can be seen as a dynamic measure. We show here that using the dynamics of correlation can be useful for a deeper understanding of the data.

This chapter starts with a small section (section 5.2) on the Pearson correlation coefficient and how we could use this measure to define temporal correlation graphs. Following the approach developed in [27], we use three dimensional tensors to represent these temporal graphs. We thus give some notions about tensors and tensor decompositions in section 5.3. The idea developed in [27] consists in using tensor decomposition in order to find a dynamic clustering of the nodes. We propose in this chapter to use this method for studying dynamic correlation graphs. After explaining the idea in more depth, we show some results on toy models and real data in section 5.4. Nevertheless, at the time this thesis is written, this work is still ongoing and not published yet.

5.1 Limitations of static methods

The correlation graph based methods proposed in [52] or [66] were only applied to static graphs. That is, either single snapshots of dynamic graphs, or graphs based on correlation computed over the total samples lengths. Although both methods give interesting results, the former focuses on a single short period while the latter might overlook short time-scale events. For correlation graphs, we could very well imagine two nodes i and j that behave most of the time in a de-correlated fashion, but over a short period, these two nodes could exhibit a strongly correlated behavior. By plotting $\rho_{ij}(t)$ we could easily spot this peak, but when we are dealing with a complex system with a large number of interacting components, things can get more complicated.

In [44], the authors study the time dependent cross correlations of stock returns at the NYSES with the idea that the return value of one stock at a given time may influence that of another stock at a different time. This can be done by studying correlation between shifted time series. This approach enables [44] to build a directed network of influence where a is connected to b if a "pulls" b. However, [44] indicates that this effect cannot be strong and the shift in the correlation coefficients should be small on financial data. Indeed, a strong effect could be used for arbitrage purposes, which is excluded from an efficient market.

[7] considers the case of correlation networks which structures might change at specific time points. Economic recession for financial data for instance might cause such transformation. The idea of [7] is to use a dichotomic approach : when a statistically significant change point is found then the correlation graph is split into two distinct graphs, one for each period. The method is then repeated on each graphs until no change point is found.

Here we consider the correlation graphs as dynamic, meaning that the edges present at time t depends on the correlation coefficients computed over a time window $[t - \Delta t, t]$. Therefore, the topology evolves according to how the correlation structure changes.

5.2 Dynamic correlation graphs

We assume that we have a set $\mathcal{X}$ of N measurements over a given period $[t_0, t_f]$ that we call the samples. We further suppose that the sampling rate is the same for all the samples $x_i \in \mathcal{X}$, and that all measurements are aligned.

As explained in chapter 4, the Pearson's correlation coefficient between two variables X and Y is the covariance of X and Y ($cov(X, Y)$) divided by the product of their standard deviations :

$$\rho_{XY} = \frac{cov(X, Y)}{\sigma_X \sigma_Y} \tag{5.1}$$

where σ_X is the standard deviation of variable X. When working with samples ($X = \{x_{t_0}, ..., x_{t_f}\}$ and $Y = \{y_{t_0}, ..., y_{t_f}\}$), this coefficient can be written in another form that might be more convenient :

$$\rho_{XY} = \frac{\sum_{t=t_0}^{t_f}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=t_0}^{t_f}(x_t - \bar{x})^2}\sqrt{\sum_{t=t_0}^{t_f}(y_t - \bar{y})^2}} \tag{5.2}$$

This coefficient varies between -1 and 1 and gives an intuition about the linear dependency between the two variables X and Y that produced the samples $\{x_{t_0}, ..., x_{t_f}\}$ and $\{y_{t_0}, ..., y_{t_f}\}$ (see 5.1). A value close to 1 or -1 (resp. 0) indicates a strong (resp. weak) linear dependency. Note however, that a null correlation coefficient does not mean independence. Some structure, different than linear dependence, might very well be present (see figure 5.1).

It can be the case that X and Y represent processes with some kind of temporal patterns such that the correlation between X and Y might change. In such a situation, we might be interested in observing the evolution of this correlation. We denote by :

$$\rho^t_{XY} = \frac{\sum_{t=t-\Delta_t}^{t}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=t-\Delta_t}^{t}(x_t - \bar{x})^2}\sqrt{\sum_{t=t-\Delta_t}^{t}(y_t - \bar{y})^2}} \tag{5.3}$$

the correlation coefficient between X and Y at time t over the time window Δ_t. That is, we compute the correlation only on the sufficiently recent measurements that we have. For a given resolution Δt, we can plot $\rho^t_{X,Y}$ as a function of time in order to gain insight about the evolution of the dependencies between variables X and Y.

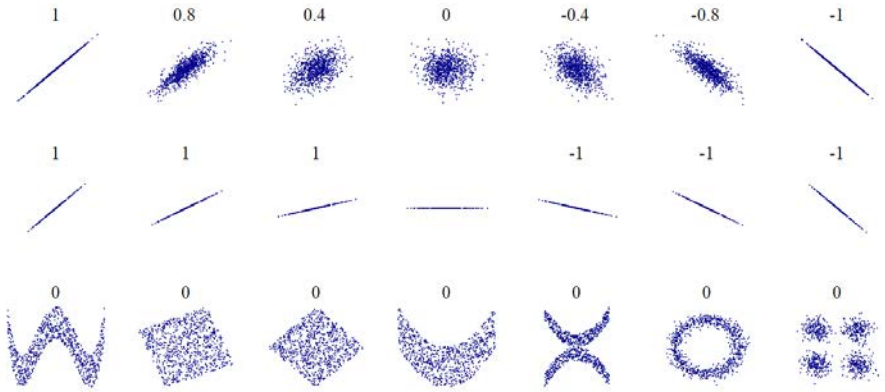


FIGURE 5.1: Sets of point and their Pearson's correlation coefficient. Source : Wikipedia.

By using the same methods as described in chapter 4, we can define G_ϵ^t as the filtered correlation graph at time t. It is then possible to build a sequence of snapshots $\left\{G_\epsilon^{t_0}, G_\epsilon^{t_1}, ..., G_\epsilon^T\right\}$ of the temporal graph G_ϵ.

5.3 Tensors

5.3.1 Temporal graphs as three-ways tensors

Temporal graphs are receiving more and more attention by the scientific community lately. Such graphs are indeed an interesting framework to study the evolution of some kind of relationship within a community. As will be explained in chapter 6, the dynamics of graphs describe how the topology evolves with time, while dynamics on graphs focus on describing a given process over a static topology. In other words, dynamics on graphs affect the states of the nodes or edges while dynamics of graphs affect the whole topology.

As explained in chapter 2, one of the most basic way to account for the topology of a graph is the adjacency matrix that indicates the presence of links between any two nodes. We can therefore represent the sequence of snapshots $\left\{G_\epsilon^{t_0}, G_\epsilon^{t_1}, ..., G_\epsilon^T\right\}$ by a succession of matrices $\{A_0, A_1, ..., A_{T-1}\}$. By stacking these matrices in a third dimension, we obtain a three dimensional array or a three ways tensor $\mathcal{T} \in \Re^{N\times N\times T}$ (see figure 5.2). In other words, the element t_{ijk} of $\mathcal{T}$ indicates whether or not node i and node j are connected at time k.

Although simple, this way of representing the temporal network hardly gives information about the structure of the data. For regular human beings, it is usually difficult to see the structure of the data within a multi-dimensional object. Dimensionality reduction techniques have been used for decades in data science, among which the famous Principal Component Analysis (PCA). Roughly speaking, PCA projects the data over the directions of the space that yield the maximum variance. Inevitably, this results in throwing away information for gaining clarity.

Tensor decomposition has been used recently in order to study temporal graphs [82] [27]. Tensor decomposition can also be seen as a dimension

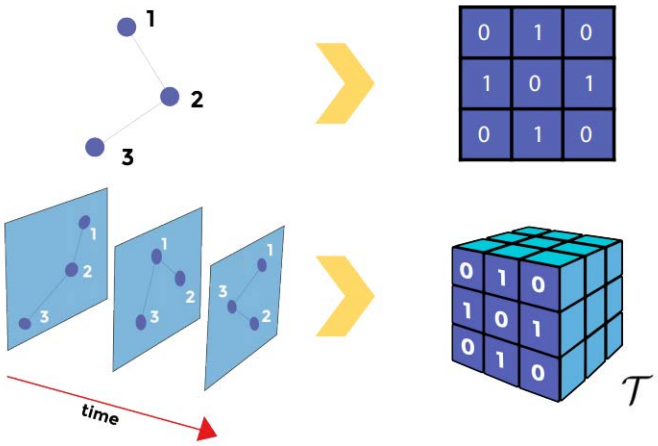


FIGURE 5.2: From temporal networks to three dimensional tensors. Source : Andre Panisson at PyData, 2016

reduction technique from multi-dimensional arrays to matrices. First recall that the rank of a matrix is the minimum number of column vectors needed to span the range of the matrix. A matrix thus has rank one if it can be written as an outer product of two nonzero vectors: $A = u \circ v = uv^T$. More generally, the rank of a matrix is the smallest number of such outer products that can be summed to produce it:

$$A = u_1 v_1^T + u_2 v_2^T + ... + u_k v_k^T \tag{5.4}$$

Similarly, a N-way tensor of rank one is a tensor that can be written as an outer product of N vectors:

$$\mathcal{T} = u_1 \circ u_2 \circ ... \circ u_N \tag{5.5}$$

It has been shown that every tensor could be expressed as a sum of rank 1 tensors [3]. The rank of a tensor $\mathcal{T}$ is thus defined as the minimum number of rank 1 tensors with which it is possible to express $\mathcal{T}$ as a sum. This means that a three dimensional tensor $\mathcal{T}$ of rank R can be decomposed as a sum of R rank 1 tensors:

$$\mathcal{T} = \sum_{r=1}^{R} a_r \circ b_r \circ c_r = \{A, B, C\} \tag{5.6}$$

where $A = (a_0, a_1, ..., a_R)$, $B = (b_0, b_1, ..., b_R)$, and $C = (c_0, c_1, ..., c_R)$ are the loading matrices of $\mathcal{T}$.

5.3.2 Tensor factorization

Finding matrices A, B, and C of equation 5.6 given $\mathcal{T}$ is not trivial at all and is a problem on its own. The problem can be relaxed in the following way [3] :

$$\mathcal{T} = \{A, B, C\} + E \tag{5.7}$$

where E denotes the residuals. Obviously, we seek the matrices A, B, and C such that we minimize the residuals, or equivalently, the difference between the data $\mathcal{T}$ and the model $\{A, B, C\}$. The most commonly used model to factorize tensor is the PARAFAC model :

$$x_{ijk} = \sum_{l=1}^{R} a_{il} b_{jl} c_{kl} + e_{ijk} \tag{5.8}$$

The idea behind PARAFAC is to use an alternative least square algorithm (ALS) that computes iteratively the difference between the data and the model, update the model accordingly, and loop until convergence is reached. There are multiple difficulties when fitting a PARAFAC model to data among which the slow convergence or the necessary initialization of the model that can impact the end results as well as the convergence speed. The PARAFAC model tries to find a decomposition that fits the data as well as possible, but the rank of the tensor that we are looking for is unknown. That is, the number of components is unknown and has to be specified by the user.

This again looks similar to PCA where the user has to provide the number of components. However, the components for PCA do not change when the model is fitted with a different number of components. That is, the axes are always the same and ranked in the same order no matter how many are computed. This means that one can simply select the number of components that gives the desired complexity / interpretability ratio, and discard the others. This is not the case for PARAFAC which components completely change when fitting with different numbers of components. The point is that, by selecting too many components, the model will probably overfit the data, and underfit with too few components.

Finding the right value for the number of components is again far from trivial and can be done in different ways. The most common technique is called the core consistency check [9] [3]. Explaining all the details about core consistency is out of the scope of this paper, however we present here the general idea such that the reader with no tensor decomposition knowledge is not lost. As a matter of fact, PARAFAC can be seen as a restricted version of the more general Tucker3 model [3] (see figure 5.3) :

$$x_{ijk} = \sum_{l=1}^{w_1} \sum_{m=1}^{w_2} \sum_{n=1}^{w_3} a_{il} b_{jm} c_{kn} g_{lmn} + e_{ijk} \tag{5.9}$$

The decomposition provided by the Tucker3 model contains a core 3 dimensional array $G = (g_{ijk})$. Basically, the outer product between the ith factor from A, the jth factor from B and the kth factor from C is weighted by the g_{ijk} element of the core G. PARAFAC is a restricted version of Tucker3 in the sense that it can be seen as a Tucker3 with a core containing only zeros except on the diagonal which contains only ones.

Imagine now that we want to fit our data with a three components PARAFAC and Tucker3 model. If a three component PARAFAC model is appropriate, it means that the core of the Tucker3 model will be close to a tensor with 3 ones on the diagonal and 24 zeros elsewhere (denoted by $\mathcal{H}$ in the following). We can compute a metric that basically indicates if the core G that we found is close to $\mathcal{H}$. The most commonly used metric is the core consistency :

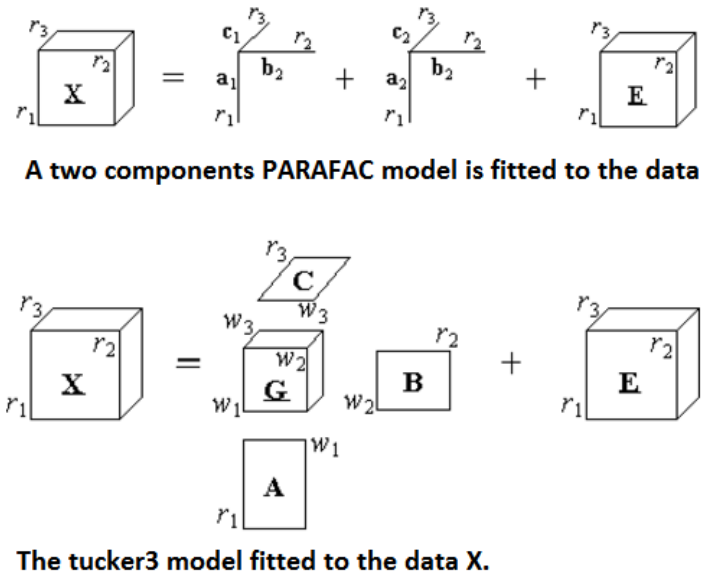


FIGURE 5.3: PARAFAC and Tucker3 models.

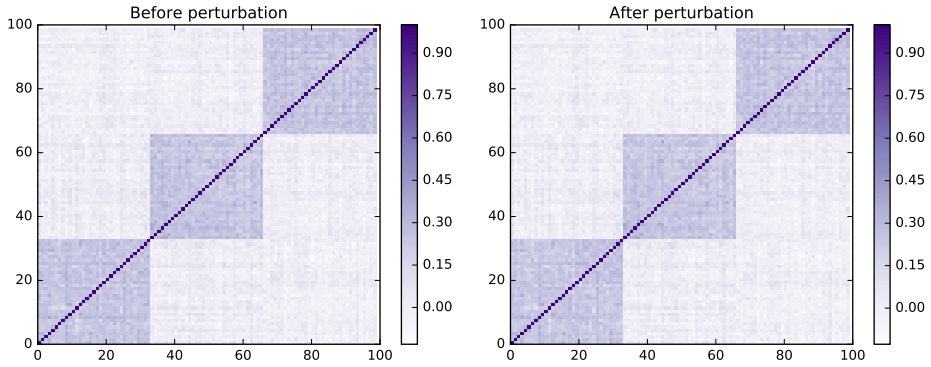


FIGURE 5.4: Correlation matrix based on total sample length before and after perturbation.

$$Core\ Consistency\ = 100\left(\frac{1-\sum_{i=1}^{F}\sum_{j=1}^{F}\sum_{k=1}^{F}\left(g_{ijk}-h_{ijk}\right)^2}{\sum_{i=1}^{F}\sum_{j=1}^{F}\sum_{k=1}^{F}t_{ijk}^2}\right) \tag{5.10}$$

A core consistency between 100% and 80% indicates good fit while much lower values or even negative values indicate that the fit is poor. Generally, the core consistency is quite high for small numbers of components, but the variance explained by the model is quite low. When we increase the number of components, the core consistency decreases while the variance explained increases. Hopefully, if PARAFAC is appropriate to model the data, there might be a point where the core consistency drops from an acceptable value to a much smaller value, and might indicate what number of components to use.

The question of how to interpret the decomposition remains [82] [27].

Let suppose that we are studying what people write (on Twitter for example). For a set of users, we can collect the tweets that he/she published as well as the time at which they were published. By splitting these tweets into words, we obtain, for each user, a list of terms that were used at different times. If we wish to represent the data as a temporal network, we could use a three way tensor $\mathcal{X} \in \Re^{N \times M \times T}$ such that x_{nmt} indicates if the term m was used by user n at time t. If we decompose $\mathcal{X} = \{A, B, C\}$ with K components, we have :

- $A^{N \times K}$ is the association of each user n to a factor k
- $B^{M \times K}$ is the association of each term m to a factor k
- $C^{T \times K}$ shows the time activity of each factor

In the case of dynamic correlation graphs, the tensor stacks a sequence of adjacency matrices in a third dimension, such that the frontal slices of the tensor are symmetric. In this particular case, the matrices A and B are equivalent and provides the same information. This means that, for temporal undirected networks, matrices A and B associate each user to the factors, and C provides the time activity of each factor. We show in the next section examples of the PARAFAC decomposition.

5.4 Results

5.4.1 An illustrative example

We build here an illustrative example to show how the decomposition of dynamic correlation networks works. We create artificially correlated clusters of timeseries with the following procedure. We draw $N_{clusters}$ random signals (here $N_{clusters} = 3$) of length L (here $L = 500$) such that the different signals are uncorrelated. We then generate N samples (here $N = 99$) of length L by adding noise to the cluster signals. This way, we have a simple structure of 3 clusters of 33 series that are strongly correlated within a cluster, and poorly across clusters.

The correlation matrix computed over the whole length of the sample is visible on figure 5.4 (left panel). As expected, the three clusters structure of the data is clearly visible. This random and stationary example is a case where simple static correlation graphs completely render the correlation structure. But most data are not that simple. Most of the time they are the results of some complex and non stationary processes.

Let us now introduce a perturbation in the data. Lets imagine that the timeseries represent stock values on financial markets for instance (even if this is pretty unrealistic). The clusters could be the economic sectors, since the stocks within a same sector tend to be more correlated than across sectors. During economic crisis, people tend to lose their self control and often act in a very correlated fashion. We introduce a small perturbation in the data supposed to represent a crisis impacting some, but not all, sectors. Besides, only a fraction of nodes within the concerned sectors will be impacted. More precisely, between $t = 50$ and $t = 100$, we assume that the nodes $\{20, 21, 22, ..., 39, 40\}$ correlate quite strongly. That is, a fraction of

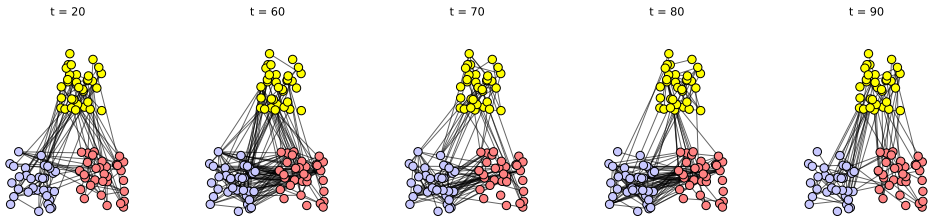


FIGURE 5.5: Snapshots of the dynamic correlation graph.

cluster 1 and a fraction of cluster 2 will be way more correlated than before during this period, but afterwards everything will be back to "normal".

By looking at the new correlation matrix on figure 5.4 (right panel), we still see the global structure of three correlated clusters, but we completely miss the perturbation that we introduced. Actually, it is very hard to spot any difference between the before and after perturbation correlation matrices. Even with the nodes ordered properly and with the knowledge of the impacted nodes, we can hardly see a difference between the two correlation matrices. The reason is that these matrices are obtained by computing the correlation over the total length of the samples, averaging out the perturbation introduced above.

Figure 5.5 shows snapshots of the filtered correlation graph, where the nodes are colored and positioned according to their cluster. The first and last snapshots show the topology at times outside the perturbation window, and the three middle snapshots show the topology within the perturbation. For such a simple example, the perturbation is clearly visible since there is a net increase of connectivity between blue and red nodes in the three middle snapshots. Nevertheless, for realistic situations with more nodes and more complex perturbations, recovering the structure by looking at snapshots of the topology might become complicated.

Another simple way for studying this graph would consist of plotting diverse metrics versus time. For example, plotting the number of edges at the different times clearly shows an increase during $t \in [50, 100]$. Because the situation here is extremely simple, we would be able to detect the perturbation with such a direct approach. However, in real situations this would probably not be feasible. Actually, it is not hard to build an example where two perturbations occurring at the same time have opposite consequences on a global metric such as the number of edges. In this case, the perturbations would be undetectable by this approach.

Our objective is thus to use tensor decomposition of the temporal graph in order to recover both the clusters and the perturbation. Figure 5.6 shows the results when using PARAFAC tensor decomposition on the present simple illustrative model. Recall that we need to find the number of components of the model that we want to fit. For this purpose, we use the core consistency check introduced in the previous section. The top left subplot of figure 5.6 shows the core consistency of the decomposition. Clearly, there is a sharp drop between 4 and 5 factors. This means that using 5 or more components for the model will probably overfit the data, while using 3 or less components will probably underfit. In other words, this tells us that

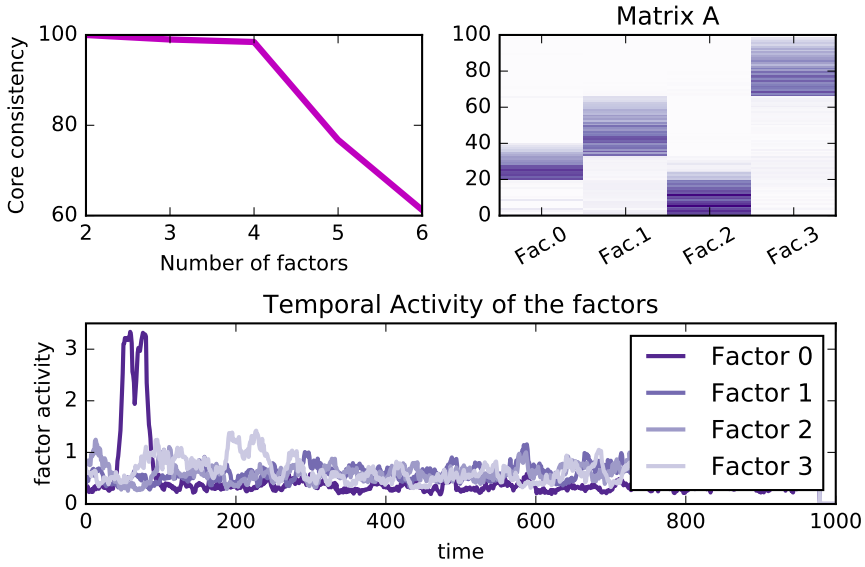


FIGURE 5.6: Tensor decomposition.

we should use 4 factors for the decomposition although this can seem surprising since we know that $N_{clusters} = 3$.

The top right plot of figure 5.6 shows the matrix A of the decomposition (matrix B is not shown because it exhibits the same information since the frontal slices of the tensor are symmetric). We can see first that each cluster is recovered: nodes $\{0, 1, ..., 32\}$ are mostly represented by factor 2, nodes $\{33, 34, ..., 65\}$ by factor 1, and nodes $\{66, 67, ..., 98\}$ by factor 3. Additionally, nodes $\{20, 21, ..., 40\}$ are also represented by factor 0. The bottom subplot of figure 5.6 shows the temporal activity of the 4 factors. Factors 1, 2, and 3 look pretty close to noise (because they are, due to the model), but factor 0 has a clear peak of activity during the period $t \in [50; 100]$. This means that the nodes $\{20, 21, ..., 40\}$, associated with this factor, tend to have way more links during this period than during the rest of the simulation. The PARAFAC model was thus able to capture both the cluster structure and the short time perturbation that we introduced.

5.4.2 Using tensor decomposition on real data

We now investigate how this method works on real data. As a an example, we use individual electricity consumption data that can be obtained from the Irish Social Science Data Archive (ISSDA)[38]. The Commission for Energy Regulation (CER) is the regulator for the electricity and natural gas sectors in Ireland. They initiated a smart metering project in 2007 with the purpose of undertaking trials to assess the performance of smart meters, their impact on consumers' energy consumption and the economic case for a wider national rollout. The project took place during 2009 and 2010 with over 5000 Irish homes and businesses participating. The customers, who participated in the trials had an electricity smart meter installed in their homes/premises and agreed to take part in the research project.

The dataset that we use contains the electricity consumption of these anonymized users. The records start around the 15th July 2009 and end the 31 December 2010 with a sampling frequency of one record every 30 minutes. We select a random sample of a hundred users and only look at the correlation over August 2009 (for example). Because this tensor already contains $100 \times 100 \times 31 \times 24 \times 2 = 14,88.10^6$ integers, increasing either the population size or the time window quickly triggers memory issues. There are multiples ways this can be improved. First of all, since the frontal slices are symmetric, we have redundant information that we could get rid of by considering only the upper or lower triangle matrices. Secondly, since we consider adjacency matrices of filtered correlation graphs, we have the possibility to "sparsify" the graphs by tuning the filter threshold. By doing so, we should obtain sparse three dimensional tensors that would theoretically require less memory to be stored. Of course, this is not "free" since at some point, a higher filter threshold would throw away meaningful information. At the time this chapter is written, this is still ongoing work, such that we only present here the result of tensor decomposition on the reduced example (100 users during August 2009).

Figure 5.7 shows the result of tensor decomposition on this reduced example. By using the core consistency check, we found that a 6 component PARAFAC model may be a good choice. The left panel of figure 5.7 shows the matrix A that maps the users to the factors. The right panel shows the activity of each factor during the month of August. By looking at factor 0 (at the bottom) we clearly see some periodic pattern. Indeed, nodes attributed to this component tend to correlate during the week and are uncorrelated on weekends. It means that these nodes tend to consume power in a correlated fashion during the week while this is much less marked in the weekends. The other factors are more noisy but also exhibit some periodic behaviors. For example, factor 1 tends to have a spike of activity every Monday.

Clearly, both node attribution and temporal activity of the factors are much more noisy than the ones obtained in the toy model. Although this had to be expected due to the presence of noise in real data, there may be some improvements that could be done to the method that we wish to look at in a near future. Furthermore the dataset also provides information about the users, such that checking what kind of users are clustered together could give insights about the quality of the results.

5.5 Discussion

Chapter 5 presented dynamic correlation graphs studied through a tensor decomposition method. This method uses the PARAFAC model that fits the data tensor with a decomposition that can be represented with three matrices. When fitting the model, one has to provide the number of components. On the contrary to PCA, this number has to be selected carefully such that there is no underfitting or overfitting. A method for estimating this number is the core consistency check that, ideally, exhibit a sudden drop when the number of components is too large.

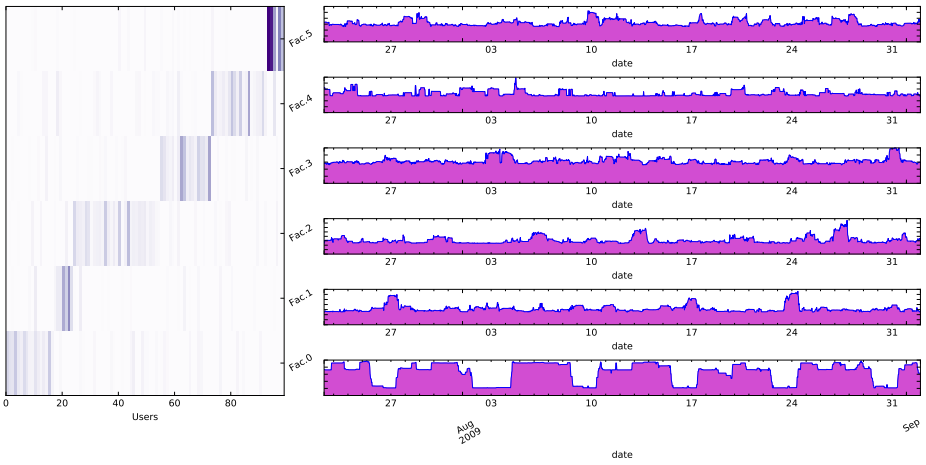


FIGURE 5.7: Tensor decomposition for smart meters records.

We showed in this chapter that this method could be used with success for studying dynamic correlation graphs. The illustrative model, studied in section 5.4.1, showed that we were able to find both the underlying structure and short time events within the data. We then proposed a real world example in section 5.4.2 where we studied a correlation graph of the electricity consumption of Irish homes and businesses. We showed that periodic pattern of correlation could be extracted and nodes classified according to these patterns.

We want to stress out the fact that the work presented in this chapter is still ongoing and had to be paused due to thesis report obligations. Hopefully, this will be improved and completed in a near future.

The three following chapters also deal with prosumer agents within the smart grid, but a different approach is explored. Until now, the electricity network was not explicitly considered, that is, we did not assume any topology nor physical constraints. In the following chapter, we introduce dynamics on networks and explain how the power grid could be represented in a tractable model.

Part II

Storage Placement in Prosumer Networks

Structure of Part 2

In the first part of this thesis we studied prosumers on a high level, such that network topologies and power dynamics were abstracted. In this second part, the electrical grid and the physical constraints of the system will be considered explicitly. We assume here a network of prosumers interconnected with electrical lines. Some of these agents can be viewed as generators while others can be viewed as pure loads. As we saw in the first part, prosumers extra-productions are likely to change for diverse reasons such that these generators and loads might not be fixed on the contrary to current power grid. We will see that the equilibrium between production and consumption is a necessary condition for stability. But, in this stochastic context, it seems quite unlikely to hold without interventions of the operator.

There exists multiple ways to adjust production and consumption (see chapter 2), but we focus here on storage devices such as batteries. The main question that we answer in this second part is to find the set of locations, within a prosumer network, for installing these storages. Obviously, the more storage we install, the easier the control of the system, the extreme case being one storage device per node. For budget reasons, it seems quite likely that only a fraction of the nodes will be equipped with storage devices, at least at the beginning. As one might expect, we show in this part that all sets are not equally efficient, even when they are of the same size. We propose here to quantify this efficiency notion and to tackle the problem from a controllability perspective. More precisely, we wish to find the subset of nodes in the graph that would require the minimum energy to control the system on average of the possible perturbations, given that generators and loads might change over time.

This second part is organized in two chapters. Chapter 6 gives the required notions on dynamics on networks and control of networks. Chapter 7 explains how we answered the storage location problem. We indeed propose an algorithm, based on submodular set functions, that takes the physical constraints of the system into account and outputs a set of locations that enables the control with low energy. The work presented in chapter 7 was submitted to the SmartGridComm 2016 conference and is currently under review.

Chapter 6

Power Grid Dynamics and Control

This chapter is divided in two main sections. Section 6.1 is related to the dynamics in networks and aims to explain how the non-linear power grid dynamics can be expressed through a concise synchronization model. We start by giving some notions about dynamics on networks and how they can be studied. We focus then on synchronization, and especially on the Kuramoto model that will be used later on to simplify the power grid dynamics. We also provide, at the end of this first section, some examples as to give to the non specialist reader some intuition for the model.

Section 6.2 introduces control theory through the special case of networks. More precisely, we consider the case of a dynamic on a given network that we wish to control. We first explain what is control and what are the state-of-the-art tools to discuss the controllability of networks. We then focus on optimal control which aims at finding the right inputs such that a specific cost function is minimized.

6.1 Power grid dynamics

6.1.1 Dynamical systems

Dynamical systems theory considers how systems change over time. A dynamical system is a system whose state is uniquely specified by a set of variables and whose behavior is described by predefined rules [83]. Dynamical systems are pretty much everywhere and can be very different, the growth of populations, the motions of objects, the swinging pendulum, or even the behavior of "rational" individuals can all be represented as dynamical systems. There are two classic ways for studying their dynamics: the discrete time steps approach or the continuous time approach. Their general mathematical formulations are as follows:

- Discrete time dynamic :

$$x_t = F(x_{t-1}, t) \tag{6.1}$$

- Continuous time dynamic :

$$\frac{dx}{dt} = F(x, t) \tag{6.2}$$

where x_t is the state vector of the system at time t. That is, x_t contains the state variables of the system at time t. F is the function that encodes

the dynamics of the system and can take various forms from simple linear functions to much more complicated non-linear ones.

For "simple" systems, although the dynamics might be quite complicated, the number of variables involved is usually relatively small. For these systems it might be possible to observe their behavior with standard methods such as time series, phase space plots, or cobweb plots [83]. However, when it comes to complex systems, the number of variables involved in the model jumps from just a few to thousands or tens of thousands. In addition, all dynamical components may interact with each other in non-trivial ways. This is exactly what makes complex systems so difficult to study, but at the same time so fascinating. Key concepts, such as emergence and self-organization, all derive from the fact that such a system is made of a massive amount of interactive components. Studying these properties at various scales and understanding how these properties are linked across scales is the basic approach for studying complex systems.

6.1.2 Networks and dynamics

Although dynamical (complex) systems are a very interesting topic, it is out of the scope of the present thesis. We will focus here on the special case of networks for which we can discern three different kinds of dynamics:

- **Dynamics on networks:** Usually, models of this category consider how the states of the nodes change over time through the interactions represented by the links between them. A key point is that the topology of the network is fixed over time, only the states of the nodes are evolving.
- **Dynamics of networks:** These models consider dynamical changes of the network topology over time. The reasons might be to understand the mechanisms that give birth to particular network topologies (preferential attachment for example), to evaluate robustness and vulnerability of networks, or to understand how to improve certain properties of networks.
- **Adaptive networks:** These models describe the co-evolution of dynamics on and of networks, where node states and network topologies dynamically change adaptively to each other.

The power grid dynamic clearly belongs to the first class, since the desired model should explain how current is exchanged between interconnected fixed units. The topology of the grid might change with the addition of new lines, transformers, generators and so on, but the time-scale at which such modifications occur is much larger than the time-scale of the power dynamic, such that we can consider that the grid topology is fixed over time.

The general idea of dynamics on networks is that each node i is associated to one (or more) variable(s) x_i. Each node is connected to other nodes, i.e its neighbors, that impact the value of x_i. In turn, i also impacts the values of its neighbors' variables $x_j,\ j \in \mathcal{N}_i$. The rules that describe how these modifications occur are the essence of the model. For example, epidemic models try to explain the diffusion of diseases or information in a network.

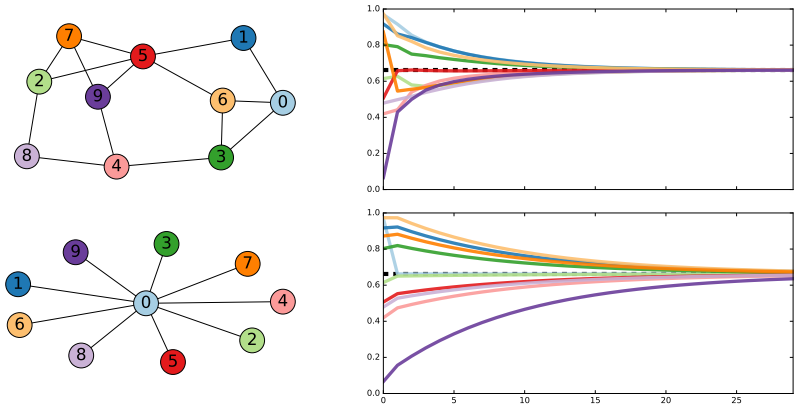


FIGURE 6.1: Average consensus in networks. Topologies have a direct impact on the convergence of the algorithm.

Roughly speaking, the variables x_i says whether the node is infected or susceptible to become infected, and the links account for the presence of contacts between individuals. The disease then spreads along links from infected individuals to susceptible ones with some probability. Of course, numerous variants have been developed, and these models are used in unexpected domains such as viral marketing for instance.

Other models focus on explaining the emergence of global information when individuals all have different opinions at the beginning. In this case, the variable x_i represents the belief of node i and its value is updated by taking into consideration the beliefs of its neighbors. Imagine that we wish to compute the mean belief $x_{AVG} = \frac{1}{N}\sum_i x_i$ in such a distributed way :

$$X(t+1) = WX(t) \tag{6.3}$$

where $X(t) = \{x_0(t), x_1(t), ..., x_N(t)\}$ is the vector of beliefs at time t, and W is the following matrix :

$$W_{ij} = \begin{cases} \frac{1}{max\{k_i,k_j\}+1} & if\ (i,j) \in E \\ 1 - \sum_{l \in \mathcal{N}_i} \frac{1}{max\{k_i,k_l\}} & if\ i = j \\ 0 & otherwise \end{cases} \tag{6.4}$$

where k_i is the degree of node i. Figure 6.1 shows an example with simple network topologies. After convergence of the algorithm, all nodes have the same average value. However, the speed of convergence is clearly impacted by the topology of the network. Such consensus models can be used to recover true information from several noisy measurements of sensors for instance. Because of their distributed nature they also provide an interesting framework for smart grid systems. [99] for example uses distributed consensus algorithms in order to re-dispatch production among generators after a power imbalance.

The literature on dynamical models on networks is very rich and is a trendy subject. In the next section, we will focus on synchronization models that share some characteristics with consensus algorithms.

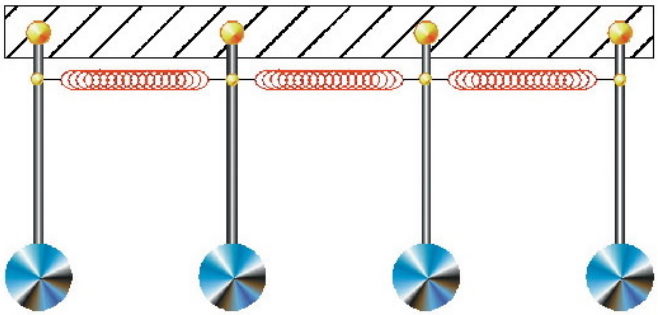


FIGURE 6.2: Schema of coupled oscillators.

6.1.3 Coupled oscillators

In 1665 Christian Huygens noticed that the two clocks placed on his mantelpiece had acquired an opposing motion. That is, their pendulums were beating in unison but in the opposite direction. Regardless of how the two clocks were started, he found that they would return to this state. This is the first recorded observation of a coupled oscillator system. The reason for such a behavior was that the two pendulums were affecting each other through slight motions of the supporting mantelpiece. He reported the results by letter to the Royal Society and it was referred to as "*an odd kind of sympathy*" in the Society's minutes. This concept is now known as entrainment.

Since the discovery of Christian Huygens, coupled oscillators have been extensively studied in the physics community. Amusing experiments with metronomes placed on a wooden board, itself placed on wheels are visible on the internet (see for instance Kuramoto model). Yet, synchronization arises in various natural systems where the coupling seems less obvious than the wooden board. For instance, it is known that fireflies clouds achieve synchronization of their flash while no entity gives the global frequency. That is, fireflies do not look at a single bandmaster on which they try to synchronize, they rather have a small local knowledge of the neighboring bugs. How is it possible to explain the emergence of synchronization in such a complex system ?

6.1.4 Simple Kuramoto model

Although coupled oscillators were known for centuries, a complex system approach of coupled oscillators had to wait until 1975, year in which the Kuramoto model was published [45]. This model was first proposed by Yoshiki Kuramoto and is a mathematical model used to describe synchronization. In the most popular version, each of the oscillators is considered to have its own intrinsic natural frequency ω_i, and each is coupled equally to all other oscillators. In this situation, the dynamic is often written as :

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \qquad i = 1 \dots N \tag{6.5}$$

This is clearly a nonlinear model but, as $N \to \infty$, this model can be solved exactly [86]. Let us introduce the order parameters r and ψ as :

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} \tag{6.6}$$

In equation 6.6, r represents the phase-coherence of the oscillators, while ψ indicates the average phase. By combining equations 6.5 and 6.6, we can re-write the dynamic as :

$$\frac{d\theta_i}{dt} = \omega_i + Kr\sin(\psi - \theta_i) \tag{6.7}$$

Which can further be simplified by working on a rotating frame (thus setting $\psi = 0$) :

$$\frac{d\theta_i}{dt} = \omega_i - Kr\sin(\theta_i) \tag{6.8}$$

In this situation, by using the order parameter, we are thus able to decouple the equations.

Most of the time however, the coupling is not general and a given oscillator is only coupled to a set of neighbors (the right-hand side sum of equation 6.5 goes only on the neighbors of node i). In this context, a graph representation is often used such that nodes represent oscillators and link weights indicate the strength of the coupling between two nodes. A more general form for the dynamic of the system can be written as :

$$D_i\dot{\theta}_i = \omega_i - \sum_{j=1}^{N} a_{ij} sin\left(\theta_i - \theta_j\right) \tag{6.9}$$

where a_{ij} is the "row i, column j" term of adjacency matrix A, and D_i is the damping coefficient of node i.

6.1.5 Synchronization

The purpose of the Kuramoto model is to study and explain how synchronization occurs in a system of interacting units. One of the most fascinating thing about oscillators networks is that, in certain conditions, nodes that have only local knowledge are able to synchronize at the system level to a main frequency. Phrased in this way, synchronization seems close to consensus, and there are indeed strong similarities between the two phenomena. The key question is whether we can predict if a given network will synchronize or not. Intuitively, there are multiple parameters that impact the synchronization of a network. For example, we expect both the distribution of the natural frequencies $g(\omega)$ and the topology of the underlying network to have an impact on the synchronization.

The most obvious definition of synchronization is that all frequencies are equal to a same value : $\omega_i = \omega_{sync}$, $\forall i$. However, another important notion is the phase cohesiveness γ that quantifies the spread of the phase angles. That is, we say that a solution has phase cohesiveness γ if every pair of connected oscillators has phase distance smaller than some angle $\gamma \in [0, \pi/2[: |\theta_i - \theta_j| \leq \gamma,\ \forall (i,j) \in E$.

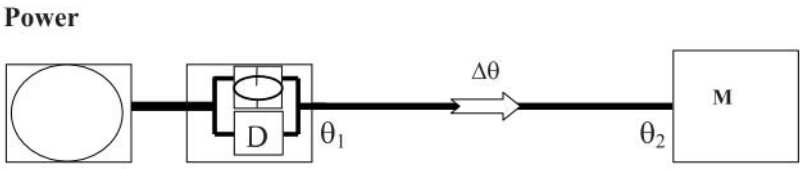


Fig. 1. Equivalent diagram of generator and machine connected by a transmission line. The turbine consists of a flywheel and dissipation D.

FIGURE 6.3: Electrical line representation. (source: [23])

A very interesting work by Florian Dörfler [18] showed that the synchronization condition could be written in a surprisingly concise way. For a network of coupled oscillators with adjacency matrix A, Laplacian matrix L, and vector of natural frequencies ω, the synchronization condition can be written as :

$$\|L^{\dagger}\omega\|_{\infty,E} \leq sin(\gamma) \tag{6.10}$$

where $L^{\dagger}$ is the Moore-Penrose pseudo inverse of L, and

$$\|x\|_{\infty,E} = max_{\{i,j\}\in E}\,|x_i - x_j| \tag{6.11}$$

is the worst case dissimilarity for a vector $x = (x_1, x_2, ..., x_n)$ over the edge set E. This means that, if this condition is respected for a given network, it will synchronize with phase cohesiveness γ to a common frequency :

$$\omega_{sync} = \frac{\sum_{k=1}^{N}\omega_k}{\sum_{k=1}^{N}D_k} \tag{6.12}$$

For the classic Kuramoto oscillator model coupled in a complete graph with uniform weights $a_{ij} = K/N$ (see equation 6.5), the synchronization condition 6.10 reduces to the condition $K > max_{i,j\in\{1,...,N\}}\,|\omega_i - \omega_j|$. That is, the coupling strength has to be larger than the maximum natural frequencies dissimilarity in order for the system to synchronize.

6.1.6 From oscillators to the power grid

At first glance the link between oscillators and the power grid might not seem obvious. In this section we explain briefly why a kuramoto-like model is an interesting choice for modeling the power grid dynamic.

The power grid can be roughly seen as generators linked to loads through electrical lines (see figure 6.3). A generator is an electric machine that converts mechanical energy to electrical energy. Other kind of electric machines are motors that convert electrical energy to mechanical energy, and transformers that change the voltage level of an alternating current. Electromagnetic generators can be divided in two broad categories: dynamos and alternators. The formers generate direct current while the latter generate alternating current.

Electric machines are composed of a rotating part called the rotor, and a static part called the stator. The rotor generates a moving magnetic field

around the stator, which induces a voltage difference between the windings of the stator. This produces the alternating current (AC) output of the generator. In simple words, the rotating speed of the rotor impacts directly the electric production of the generator, such that it could be possible to model the dynamic of the power grid by considering the phase angles and frequencies of the rotating parts of the elements in the grid.

6.1.7 Modeling the power grid dynamic

In this section, we introduce the coupled oscillators network model used to simplify the power grid dynamic. More details can be found in [23].

The objective is to achieve synchronization of the grid at the main frequency $\Omega = 50Hz$. Each oscillator i has a phase angle δ_i and a frequency $\dot{\delta}_i$. Therefore, we seek an equilibrium of the form :

$$\dot{\delta}_i = \Omega, \ \forall i \tag{6.13}$$

For convenience, we express the dynamic of the oscillators in terms of the deviations from the main frequency : $\delta_i(t) = \Omega t + \theta_i(t)$. Let $\omega_i = \dot{\theta}_i$, such that $\dot{\delta}_i = \Omega + \omega_i$. The equilibrium, in terms of the deviations dynamics, is : $\forall i, \ \omega_i = 0$

The next step consists in translating the dynamics of the generators and machines into equations involving the phase angles θ_i and the frequencies ω_i. Generators and machines are composed of a turbine that dissipates energy at a rate proportional to the square of the angular velocity :

$$P_{diss,i}(t) = K_{Di}(\dot{\delta}_i(t))^2 \tag{6.14}$$

where K_{Di} is the dissipation constant of entity i. Furthermore, it also accumulates kinetic energy at a rate :

$$P_{acc,i}(t) = \frac{1}{2} I_i \frac{d}{dt} \left(\dot{\delta}_i(t)^2 \right) \tag{6.15}$$

where I_i is the moment of inertia of entity i. For simplicity, we consider that all entities have the same dissipation constants(K_D) and moment of inertia (I).

The condition for the power transmission between i and j is that the two devices do not operate in phase. The phase difference between i and j is : $\delta_j(t) - \delta_i(t) = \Omega t + \theta_j(t) - \Omega t - \theta_i(t) = \theta_j(t) - \theta_i(t)$. The transmitted power along the line can be written as :

$$P_{transmitted} = -P_{ij}^{MAX} sin(\theta_j - \theta_i) \tag{6.16}$$

with P_{ij}^{MAX} being the maximum capacity of the line (i, j). Each entity i is then described by a power balance equation of the type (see figure 6.3) :

$$P_{S,i} = P_{diss,i} + P_{acc,i} + P_{transmitted,i}, \tag{6.17}$$

where $P_{S,i}$ is the power of an ideal source or sink at node i. By substituting equations 6.14, 6.15, and 6.16 into equation 6.17 and re-arranging the terms, we obtain the following non-linear coupled system of equations :

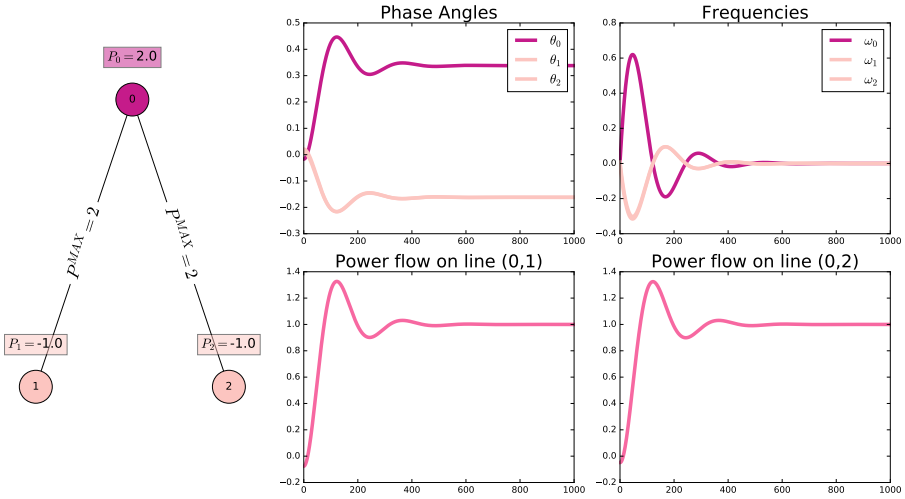


FIGURE 6.4: Simple example of Kuramoto-like model for the power grid dynamic.

$$P_{S,i} = I\Omega\ddot{\theta}_i + \left[I\ddot{\theta}_i + 2K_D\Omega\right]\dot{\theta}_i + K_D\Omega^2 + K_D\dot{\theta}_i{}^2 - \sum_{j\in\mathcal{N}_i} P_{ij}^{MAX} sin\left[\theta_j - \theta_i\right] \tag{6.18}$$

We now use simplifications based on the fact that we consider small deviations from the main frequency : $\dot{\delta}_i \sim \Omega$ which means that $\omega_i = \dot{\theta}_i << \Omega$, such that the squared term $K_D\dot{\theta}_i^2$ can be neglected. Moreover, we assume that the rate at which energy is stored in the kinetic term is much less than the rate at which energy is dissipated by friction : $\ddot{\theta}_i << \frac{2K_D}{I}$ (see [23] for more details). Equation 6.18 becomes :

$$\ddot{\theta}_i \sim \psi_i - \alpha\dot{\theta}_i - \sum_{j\neq i} K_{ij} sin\left[\theta_j - \theta_i\right] \tag{6.19}$$

where $\alpha = \frac{2K_D}{I}$ is the dissipation term, $K_{ij} = \frac{P_{ij}^{MAX}}{I\Omega}$ are the coupling strengths, and $\psi_i = \left[\frac{P_{S,i}}{I\Omega} - \frac{K_D\Omega}{I}\right]$ encodes the power distribution of the elements. In order not to overload the equations in the following, we simplify the constant term $\frac{K_D\Omega}{I}$ by working in a rotating frame such that $\psi_i = \frac{P_{S,i}}{I\Omega}$.

6.1.8 Examples

In this section we show some examples where we use the model 6.19 to simulate the grid dynamic. We first consider a very simple network with a generator feeding two machines (see left panel of figure 6.4). We assume that the generator produces 2 units of power and that machines consume one unit each, such that the production matches the consumption. Moreover, the maximum capacities of the lines are set to 2 units and we take the following values for the parameters : $I = 10^{-1}$, $K_D = 10^{-1}$, and $\Omega = 50Hz$. We select the initial state randomly inside some hypersphere centered on the stable state. This will result in a transient phase before synchronization that is visible on the right plots of figure 6.4. We also show the power that flows on each electrical line. As visible, the power oscillates during the

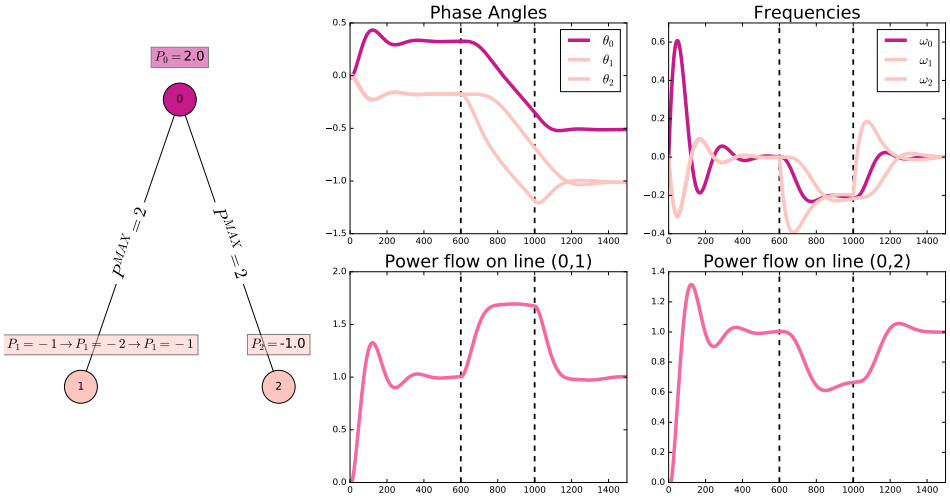


FIGURE 6.5: Simple example of the effect of a power mismatch in the Kuramoto-like model for the power grid dynamic.

transient but stabilizes to one unit on each line. We also see that during, the transient, no line is overloaded ($P_{ij}(t) < P^{MAX}\ \forall t$).

It is interesting to observe how the model reacts when there is a mismatch between production and consumption. For this simulation we start in the same conditions as the previous one, but between $t = 600$ and $t = 1000$, we introduce a perturbation. During this period, one of the machine will consume 2 units of power instead of one. This results in a power imbalance that steers the frequencies away from the synchronized state as visible on figure 6.5. Indeed, after $t = 600$, the frequency drops significantly before reaching the synchronized state after $t = 1000$. Although such a mismatch does not seem to cause much problem in the model, this could result in dramatic damages on real power systems.

6.2 Control in networks

6.2.1 What is control ?

In this section, we introduce some key notions related to the controllability of complex networks. Although we are concerned with control in networks here, control theory has a much larger spectrum of applications. Control theory asks how to influence the behavior of a dynamical system with appropriately chosen inputs so that the system's output follows a desired trajectory or ends in some chosen final state. One of the most fundamental notion of control theory is the feedback. Basically, the difference between the actual and the desired final state is applied as a feedback into the system input. By acting so, the system's output converges to the desired value.

Obviously, we need a way of representing the state of the system and the equations that govern its evolution. The concept of state was introduced into control theory by Rudolf Kalman in 1960s, and was inspired by the phase space concept used in physics. Any state of a dynamical system can then be represented as a vector in the state space where the axes correspond

to the state variables. The equations governing the dynamic are usually written in the form of a set of differential equations :

$$\begin{cases} \dot{x}(t) = F(t, x(t), u(t); \Theta) \\ y(t) = H(t, x(t), u(t); \Theta) \end{cases} \tag{6.20}$$

In equation 6.20, $x(t) \in \Re^N$ is the state vector of the system at time t, and $u(t) \in \Re^M$ is the input vector of control signals injected at time t. Θ is the set of system parameters, and $y(t)$ is the observed state of the system. Obviously, the study of system 6.20 highly depends on the functions F and H. There exists two main categories for these functions : linear and non-linear.

A large part of control theory focuses on the study of linear systems because non-linear systems are known to be much more difficult to study. A classic approach that partially overcomes the difficulties due to non-linearity is to linearize the systems around their equilibrium points. It is then possible to look at the controllability of the linearized version of the system.

For complex systems, things become quite difficult since they comprise many interacting units on which we often lack full knowledge. It means that we may not know all the details about equations 6.20. Nevertheless, there has been extensive work on understanding control of these systems, such that we are able to discuss the controllability of a system without the precise and complete knowledge of all the parameters.

In the rest of the section, we focus on control in networks. We consider a graph $G = (V, E)$ of N nodes, where V is the vertex set and E is the edge set. The topology of this network can be encoded through the adjacency matrix M, which element $m_{ij} \in \{0, 1\}$ indicates whether i and j are connected. If the edges are weighted we can set $m_{ij} = w_{ij}$ where w_{ij} is the weight of edge(i, j), and zero if there is no edge. Graphs are often used to model complex interactions inside a population. Let us attribute to each node i the variable x_i that represents the state of node i. Depending on the model, this state can represent an opinion, a physical quantity, a probability and so on. If there is an edge between i and j, we might expect some interaction that would eventually change the values of x_i and x_j. Such a dynamic is usually described with a system of differential equations. If we assume a linear first order dynamic, this can be written as $\dot{X}(t) = AX(t)$, where $X(t) = \{x_0(t), ..., x_{N-1}(t)\}$ is the vector of node states at time t, and matrix A encodes both the dynamics and the topology of the system (A can be, among other, the weighted adjacency matrix of the underlying graph).

If there is no action from the outside, the system state will evolve according to $X(t) = X_0 e^{At}$ where X_0 is the initial state of the system. Imagine now, that this dynamic could be somehow influenced by injecting some signals. More precisely, let us assume that we can inject these signals in a subset of the nodes, called drivers, such that the dynamics becomes $\dot{X}(t) = AX(t) + Bu(t)$, where the matrix B indicates which nodes receive the signals and $u(t)$ is the vector of inputs at time t. Assume that the system is in some initial state X_0 and we aim at bringing it to some final state X_T in some amount of time T. Control can be seen as finding the sequence of inputs $\{u(t_0), ...u(t_T)\}$ that would do it given the dynamics A. Furthermore, we say that the system is controllable in T steps if it can be steered from

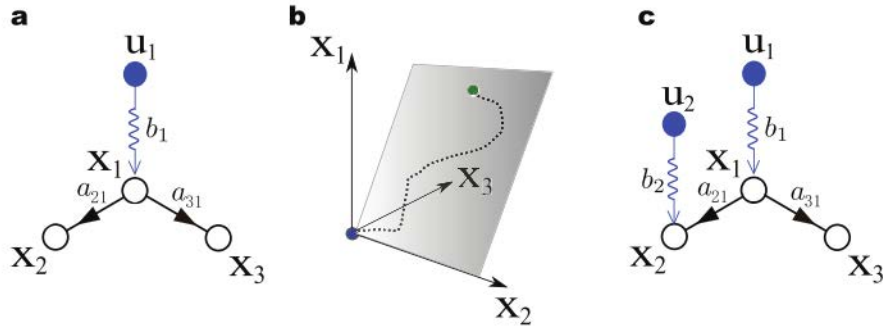


FIGURE 6.6: Controllability of networks. Source : [48]

any initial state X_0 to any final state X_T through a sequence of control inputs. This means that drivers may or may not enable the full control of the system depending on the dynamic, the topology, and of course the selected drivers.

Figure 6.6 presents a very small example of a three node network. On subplot a of figure 6.6 we apply a single control input on node 1. Although it might seem surprising to some reader, the system is not fully controlled by only one input on node 1 [48]. Indeed, by changing the input $u_1(t)$, $x_2(t)$ and $x_3(t)$ can only evolve in a correlated fashion. This can be verified by writing the dynamic equations where the result is that the system state is stuck in some hyperplan $a_{31}x_2(t) = a_{21}x_3(t)$ as shown on subplot b. As a matter of fact, it is impossible to fully control this small network with a single input, and one has to add another driver (either node 2 or node 3) as shown on subplot c of figure 6.6.

The example of figure 6.6 was simple enough so that we were able to determine if a given set of drivers controlled the system, and what was the smallest driver set that could control the system. In practice, for larger systems, these two questions are far from trivial and are at the heart of several fascinating works.

6.2.2 Kalman rank criterion

A classical result for quantifying the ability of driver sets to control a system was proposed by Kalman. He showed that a linear system (A, B) is controllable if the controllability matrix

$$C = [B, AB, A^2B, ..., A^{N-1}B] \tag{6.21}$$

has full rank [65]. That is :

$$rank(C) = N \tag{6.22}$$

For the small example of figure 6.6 (subplot a) we have :

$$C = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & a_{21}b_1 & 0 \\ 0 & a_{31}b_1 & 0 \end{pmatrix} \tag{6.23}$$

which does not have full rank, hence the system of figure 6.6.a is not controllable. For the subplot c, we have :

$$C = \begin{pmatrix} b_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_2 & a_{21}b_1 & 0 & 0 & 0 \\ 0 & 0 & a_{31}b_1 & 0 & 0 & 0 \end{pmatrix} \tag{6.24}$$

which, in this case, has full rank, meaning the system of figure 6.6.c is controllable.

Kalman's rank criterion appears as a nice and concise condition to test the controllability of a given system. Nevertheless, this approach quickly becomes complicated when studying the controllability of non trivial networks. Indeed, in such cases matrix C becomes huge and requires large amount of memory. In addition, computing the rank of very large matrix can be time consuming, especially if such operation has to be done repetitively. But even more problematic is the numeric precision of the machine on which the matrix is computed. If the system is stable, we have that $A^k \longrightarrow 0$ when $k \longrightarrow \infty$. Since the matrix A is taken at large power exponents for these real systems, differentiating zeros from non zeros quickly becomes an issue, and yields inaccurate rank computations [48]. This means, that other methods have to be employed to determine the controllability of a system.

6.2.3 PHB test

The Popov-Belevitch-Hautus (PBH) controllability test [11] states that the system (A, B) is controllable if and only if :

$$rank\,[sI - A, B] = N,\ \forall s \in \mathcal{C} \tag{6.25}$$

Since the $sI - A$ block of the matrix $[sI - A, B]$ has full rank whenever s is not an eigenvalue of A, we only need to check each eigenvalue of A for the PBH test. Although the PBH test links controllability to the eigenvalues and eigenvectors of matrix A, it can be shown to be equivalent to the Kalman's rank condition.

The PBH test can also be used to determine the minimum number of drivers required to fully control the system. First recall that for an eigenvalue λ, its algebraic multiplicity is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) = det(A - \lambda I)$. The geometric multiplicity of λ is the maximal number of linearly independent eigenvectors corresponding to this eigenvalue. It can be shown that the number of control inputs must be greater than or equal to the largest geometric multiplicity of the eigenvalues of A :

$$N_D = max_i\,\{\mu(\lambda_i)\} \tag{6.26}$$

where $\mu(\lambda_i) = dim(V_{\lambda_i}) = N - rank(\lambda_i IN - A)$ is the geometric multiplicity of eigenvalue λ_i (which is also the dimension of its eigenspace). Note that for undirected networks the adjacency matrix is symmetric, such that the algebraic multiplicity is equivalent to the geometric multiplicity.

Based on the PBH test and the spectral study of A, [96] built an algorithm for finding the minimum driver set to control a system. One of the

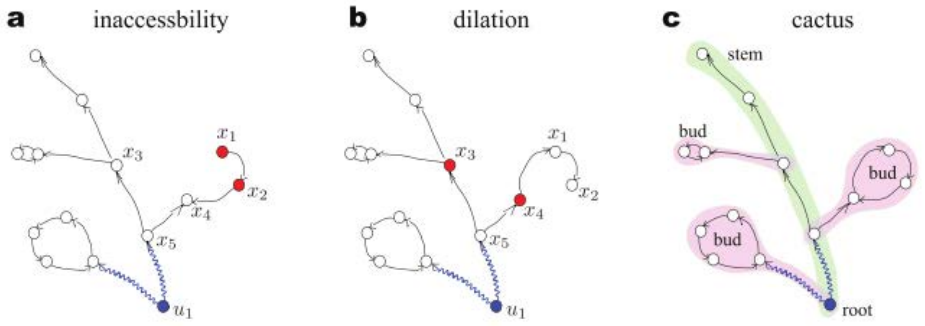


FIGURE 6.7: Structural controllability. Source : [48]

advantage of this method is that it works with directed and undirected networks.

The PBH approach is a pretty useful tool for control in networks and can be applied to much larger networks than the Kalman's rank criterion. The complexity of this approach is roughly equivalent to the diagonalization problem for which efficient algorithms exist. However, diagonalizing matrices is still a time and resource consuming operation. For very large networks, this approach turns into memory and CPU related issues, such that other methods are needed. Furthermore, the approach developed in [96] supposes that we have the exact anf full knowledge of the system parameters, which might not be always the case.

6.2.4 Lin structural controllability

A different approach to study controllability in complex network was introduced by Lin in 1974 [46]. This approach is particularly interesting for large systems where the exact values for the parameters are not known. More precisely, we call (A, B) a structured system if the elements in A and B are either fixed zeros or independent free parameters. In this case, matrices A and B are called structured matrices. The system (A, B) is then structurally controllable if the nonzero elements in A and B can be selected such that the resulting system is controllable ($rank(C) = N$). Since the nonzero elements of A and B are free, we can build a set of of all possible systems that share the same zero-nonzero connectivity structure as (A, B). It has been shown that almost all systems of this set are controllable except for some specific cases with Lebesgue measure zero. This means that if a system is structurally controllable, then it is controllable for almost all possible parameter realizations. Structural controllability requires thus the knowledge of the links between the elements, but not the exact weights of these links.

This framework enables to study the controllability of a system by looking only at its topology, avoiding matrices operations. It has been shown by Lin that the system (A, B) is not structurally controllable if and only if it has inaccessible nodes or dilations. Inaccessible nodes are simply nodes for which there is no path starting at an input vertex (see subplot a of figure 6.7). Obviously, these nodes cannot be influenced by the actions of the drivers. A dilation occurs when a small subset of nodes tries to control a

larger subset of nodes. More precisely, we say that a digraph $G(A, B)$ contains a dilation if there exists a subset of nodes such that its neighborhood set has fewer node (see subplot b of figure 6.7). The minimal structure that contains neither inaccessible nodes nor dilation is called a cactus. Cacti contain two kinds of structures : a stem, that is a path originating from an input vertex, and buds that are cycles with an extra-edge ending in one of the vertices of the cycle. An example of a cactus is shown on subplot c of figure 6.7. By definition, if we add a single edge to a cactus then the system is not controllable anymore unless we add one or more control signals. This leads to Lin's structural controllability theorem:

Theorem 1. *An LTI system (A,B) is structurally controllable if and only if G(A,B) is spanned by cacti.*

Although an interesting theorem, this does not tell us how to find the minimum set of drivers that control a given system. Yet, an elegant and purely graph theoretical method has been developed from this theorem. This approach relies on the maximum matching problem, which is a widely studied problem in graph theory. For undirected graphs, a matching is a set of edges without common vertices, while for directed graphs, a matching is a set of directed edges that do not share common start or end vertices. A maximum matching is then a matching of the largest possible size. A matching splits the nodes in two categories : the matched nodes located at the end of a matching edge, and the unmatched nodes. Furthermore, we say that a maximum matching is perfect if all nodes are matched. This leads to the minimum input theorem of Lin :

Theorem 2. *To fully control a directed network, the minimum number of inputs, or equivalently the minimum number of driver nodes, is* $N_D = max\left\{N - |M^\star|, 1\right\}$, *where* $|M^\star|$ *is the size of the maximum matching.*

If we can find a maximum matching for a given graph, then, by applying a control input to every unmatched node we are able to control the system. This means that, in this context, finding the minimum set of driver for a given system is equivalent to finding the maximum matching of the corresponding digraph. The real value of this theorem comes from the fact that the maximum matching problem in a digraph in not NP-hard, but can be solved in polynomial time by using the Hopcroft-Karp algorithm, which runs in $O(\sqrt{V}E)$ time.

Structural control is therefore an extremely powerful method since it allows to discuss (large) network control easily even if the exact weights of the edges are unknown. Note however that the independence of the parameters is a key assumption. If A is the adjacency matrix of an undirected network for instance, the independence of the nonzero parameters is not verified since A is symmetric.

6.2.5 Optimal control

In the previous section we gave some notions about control, and more specifically, control in networks. We reviewed different methods for determining if a given system is controllable or not, and how to select the minimum set of drivers to obtain full control. If we have the control of the system, we are able, by definition, to move it from any initial state x_0 to any

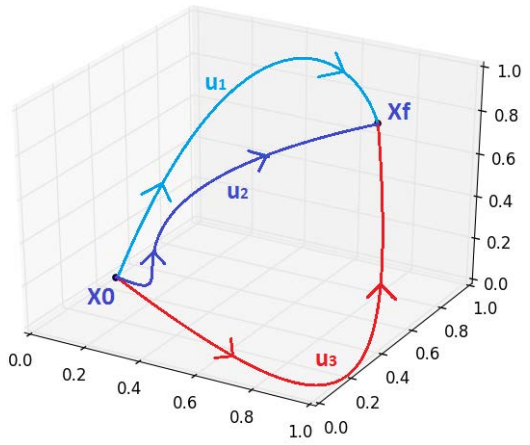


FIGURE 6.8: Different control input sequences can drive the system from initial to final state.

final state x_T by injecting a sequence of inputs $u(t)$. In this case however, there might be more than one sequence of control inputs $u(t)$ that could drive it from x_0 to x_T (see figure 6.8). Although all these sequences have the same result on the state of the system, they may not incur the same cost to the controller or the environment. Depending on our definition of the cost, some sequences of inputs will be less expensive than others. Among all these possibilities, optimal control is devoted to find the sequence that minimizes some cost function J [42]. In the general case, the cost can be written as a function of the states of the system $x(t)$, the control inputs $u(t)$, and the final state x_T:

$$J = \Phi\left[x_0, t_0, x_T, t_T\right] + \int_{t_0}^{t_T} \mathcal{L}\left[x(t), u(t), t\right] dt \tag{6.27}$$

where Φ is the endpoint cost, and $\mathcal{L}$ is the Lagrangian.

A special case of the general nonlinear optimal control problem is the linear quadratic (LQ) optimal control problem [42]. The LQ problem consists in minimizing the quadratic continuous-time cost function :

$$J = \frac{1}{2}\mathbf{x}^T(t_f)\mathbf{S}_f\mathbf{x}(t_f) + \frac{1}{2}\int_{t_0}^{t_f}\left[\,\mathbf{x}^T(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}(t)\mathbf{u}(t)\,\right] dt \tag{6.28}$$

Subject to the linear first-order dynamic constraints :

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \tag{6.29}$$

and the initial condition :

$$\mathbf{x}(t_0) = \mathbf{x}_0 \tag{6.30}$$

Control of LQ systems, and a fortiori control of systems, is out of the scope of this thesis and numerous articles and books are available on the subject [42] [24]. In this thesis, we will focus on a very specific case of cost function. Indeed, we will see in chapter 7 that we will use control theory for modeling the actions of storage devices on the system. For these devices,

we are concerned with the amount of energy that they inject or withdraw from the grid. This energy $\mathcal{E}$ can be written in terms of the control inputs $u(t)$ during the control phase [87] :

$$\mathcal{E} = \int_0^T \|u(t)\|^2 dt \tag{6.31}$$

This control energy can be seen as a specific case of the cost function 6.28 where $\mathbf{S}_f = 0$, $\mathbf{Q}(t) = 0,\ \forall t$, and $\mathbf{R}(t) = I,\ \forall t$. It can be shown [48] that the control inputs that minimize $\mathcal{E}$ can be written as :

$$u^\star(t) = B^T e^{A^T(T-t)} W^{-1}(T) v_f \tag{6.32}$$

where $\nu_f = X_T - X_0 e^{AT}$ is the difference between the desired final state X_T and the final free state $X_0 e^{AT}$, and

$$W(t) = \int_0^t e^{A\tau} B B^T e^{A^T \tau} d\tau \tag{6.33}$$

is called the Gramian matrix of the system [48]. It can also be shown that the system is controllable if and only if the Gramian is not singular, and that its rank indicates the dimension of the controllable subspace [87]. Besides, the minimum control energy associated with the inputs $u^\star(t)$ can be written as [87] :

$$\mathcal{E}_{min} = \nu_f^T W^{-1}(T) \nu_f \tag{6.34}$$

In cases where W is not invertible, the pseudo-inverse $W^\dagger$ can be used to obtain similar information in the controllable subspace.

6.3 Discussion

In this chapter we first explained how the dynamics on networks could be studied. We reviewed different types of dynamics on networks, and studied in more depth the Kuramoto model for synchronization. This model was indeed shown to provide a good simplification of the power grid dynamics.

Starting from a network of generators and loads inter-connected with electrical lines, we explained how the non linear dynamics of such a system could be simplified to a Kuramoto-like model. The obtained system of equations 6.19 can be used to understand how the parameters and the topology influence the dynamics of the system. We saw that synchronization is indeed not the objective in all networks and that it occurs only if certain conditions are satisfied (see equation 6.10).

In the examples of section 6.1 we looked at the effects that a small power imbalance has on the frequency at which the system synchronizes. Roughly speaking, if the production is lower than the consumption, the frequency tends to decrease and vice versa if the production is higher than the consumption. The inertia of the electrical machines has a direct impact on the repercussion of a power imbalance on the frequency. Unfortunately, DERs are known to have smaller inertia than power plants, meaning that the frequency will be impacted faster in case of perturbations. In addition, we saw that DER cannot be scheduled and tend to be hardly predictable. In a

power system with a large share of renewables, maintaining the frequency at a constant value seems thus even more complicated than it is today. Is is likely that different methods should be used in the future for frequency regulation as to cope with these kinds of problems.

In section 6.2 we presented control in networks. We considered some kind of dynamical process occurring on a given topology. This process changes the system state in a deterministic manner such that, if we know the initial state and the dynamic we can compute the successive states the system will be in. In the case where this evolution does not coincide with the desired one, we assumed that we were able to influence the dynamics at some specific nodes, by injecting signals. These nodes, called the drivers, and these signals, called the control inputs, are related in complex ways. That is, for a given topology and dynamic, the control signals depend on the chosen drivers. But this can be understood the other way around : the drivers that we choose might depend on the control signals that they require for controlling the system.

First of all, we saw that the ability to move the system in the state space without restriction was given by some driver sets but not necessarily by all of them, such that the question of finding the smallest one is paramount. Although a lot of contributions focus on this minimum driver set problem, we argued that controlling the dynamics was maybe not all we need. In some cases the signals that are injected do not incur a cost, such that the theoretical ability of control is sufficient. However, as we will see in chapter 7, we are particularly interested in these signals. More precisely, we would like to be able to control the system with a small amount of control energy (see 6.34) possibly at the expense of the addition of a few more drivers. Optimal control theory provides useful tools for this kind of problems. Actually, we consider here a specific case of optimal control sufficiently simple for the existence of an analytical expression for the optimum control inputs (see equation 6.32).

In the next chapter, we use the results developed here to study the storage placement in power systems composed of prosumers using renewables.

Chapter 7

Optimal storage placement for prosumer networks

This chapter tackles the problem of optimum storage placement in a network of prosumers. It uses the concepts introduced in chapter 6 on dynamics and control. We first explain the problem in more depth and explain what we mean by prosumer networks. We will see that control theory can be used in order to model the actions of the storage devices in the modified Kuramoto model developed in chapter 6. In section 7.2 we derive the multiple constraints that weight on the grid and storage devices. Our objective, as we will see, consists in finding the optimum subset of nodes for installing storage.

Due to the size of the systems considered, we argue that clever optimization methods are preferable to complete-enumeration ones. Relying on [87], we use the submodularity of some control based functions in order to build an optimization algorithm. Submodular set functions are introduced in section 7.3.2 and the algorithm developed is presented in section 7.4. The last section of this chapter provides some results on the performance of the algorithm and how the number of storages evolves with the topology of the network.

7.1 Prosumer networks

In this chapter, we explore the use of storage in a network of prosumers. These agents, both producer and consumer of electricity, can behave as generators or loads depending on complex external conditions (see figure 7.1). As we saw in the previous chapters, weather conditions and the time of the day have strong impact on both the consumption and production of the prosumers. A consequence of the extensive use of DER is that perturbations are very likely, and may endanger the system stability.

In chapter 4, we studied a market oriented solution where aggregations of prosumers were forming coalitions tied to produce a pre-defined amount of power with penalties if they deviate. Nevertheless, since penalties will not prevent coalitions to fail if they must fail, it seems important for the grid operator to anticipate this kind of situations. Mismatches between production and consumption may indeed cause desynchronization of the system, which in turn, may damage the equipment and lead to blackouts.

There are multiple solutions for providing this kind of safety. Usually, frequency regulation is provided by some tier with a specific contract. The goal is to inject or withdraw large amount of power on a relatively short time in order to maintain the frequency of the system within acceptable

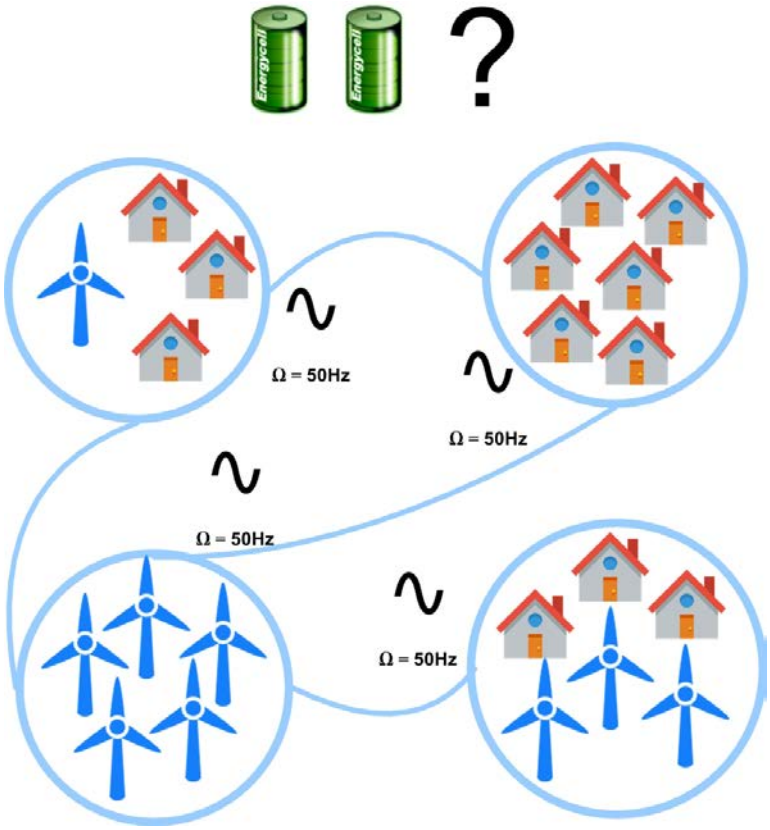


FIGURE 7.1: Storage placement in prosumer networks.

bounds. During this period, actions should be undertaken by the operator so that the required backup generation is brought online. In addition to this kind of schema, storage devices scattered across the network are a possible safety reinforcement.

If rebalancing the production and the consumption is necessary, we should make sure that it is done without violation of the system's physical constraints. Electrical lines have indeed maximum capacities, and storage is also limited with bounded rates of charge/discharge. In other words, we must have the capability of controlling the grid's dynamic within a physically possible portion of the state space.

Storage placement has often been studied in the smart grid literature with an optimization framework. Authors consider a set of possible locations as well as a budget and they maximize some utility function. For example [30] studies the placement of storage devices within microgrids in a collaborative scheme. The authors solve an interesting optimization problem mixing placement-dimensioning of storage with utilization. Moreover, [30] uses a Nash bargaining framework to determine how prosumers should share the costs and benefits of the storages. However, [30] does not take into account the grid dynamic as well as its controllability.

Control of smart grid systems has been recently studied in [22], where the authors use linear-quadratic optimal control theory in order to compute DER outputs. Control inputs rely on information collected by sensors and phasor measurement units and a communication network is used for sharing the information. The performance is then studied under practical

limitations such as latency, sampling rate, or signal-to-noise ratio. The objective of [22] is the control of the system when perturbations occur. The control signals are used for adjusting the productions of the DER to some extent, but storage is not considered within this work. The present work differs from [30] and [22] in that we study the placement of the storages from a controllability perspective in a stochastic context due to DER. Moreover, we aim at models and methods that could be scalable to some extent.

We assume that we are given a prosumer network and we have to find both the number and the locations of the storage devices that should be deployed. Once this choice is made, we have to stick with it whatever perturbations may occur. In addition, we suppose that the storage devices are controlled in a centralized way. Since installing these devices is expensive, we would like to minimize their number. In parallel, the amount of energy that flows out or from the devices is also a quantity of interest. If this energy is large, it means that we are charging and discharging a lot, which may result in reduced life time for example. The objective is then to find the smallest set of storages, as well as their locations, that will require small control energy (see chapter 7).

We saw in chapter 7 that the control energy depends on the initial and final states of the system. Since we do not have any prior on the future perturbations that may occur, we chose to minimize the average control energy required to move the system around the state space.

7.2 Dynamic and constraints

Recall from chapter 6 that an approximation of the grid's dynamic can be written in the following way :

$$\ddot{\theta}_i \sim \psi_i - \alpha\dot{\theta}_i - \sum_{j\neq i} K_{ij} sin\left[\theta_j - \theta_i\right], \tag{7.1}$$

where $\alpha = \frac{2K_D}{I}$ is the dissipation term, $K_{ij} = \frac{P_{ij}^{MAX}}{I\Omega}$ are the coupling strengths, and $\psi_i = \left[\frac{P_{S,i}}{I\Omega} - \frac{K_D\Omega}{I}\right]$ encodes the power distribution of the prosumers. In order not to overload the equations, we simplify the constant term $\frac{K_D\Omega}{I}$ by working in a rotating frame such that $\psi_i = \frac{P_{S,i}}{I\Omega}$.

The dynamic is still non linear because of the sine coupling. Therefore, we also assume that the phase angle differences are small such that $sin\left[\theta_j - \theta_i\right] \sim \theta_j - \theta_i$. By using vector notations, the dynamic can be written in the following form :

$$\ddot{\theta} = \Psi - \alpha\dot{\theta} - (K \circ L)\theta \tag{7.2}$$

Where $A \circ B$ represents the Hadamard product between matrices A and B, and L is the Laplacian matrix of the underlying topology ($L_{ij} = k_i$ if $i = j$ and $L_{ij} = -m_{ij}$ otherwise, see chapter 2). Equation 7.2 is a continuous time second order linear system of N equations.

By introducing $X = \begin{pmatrix} \theta \\ \dot{\theta} \\ 1 \end{pmatrix}$, we transform this into a first order linear system of $2N + 1$ equations, which is discretized with time step Δt, leading to :

$$X(t + \Delta t) = AX(t) \tag{7.3}$$

With transition matrix A :

$$A = \begin{pmatrix} I & I\Delta t & 0 \\ -(K \circ L)\Delta t & (1 - \alpha\Delta t)I & \Psi\Delta t \\ 0 & 0 & 1 \end{pmatrix} \tag{7.4}$$

Note that the transition matrix A encodes all the system parameters, topology, and power distribution. So far we have expressed the grid dynamic as a coupled oscillators network, but we did not incorporate the different physical constraints on the network.

7.2.1 Flow Constraints

Recall from chapter 6 that a condition for synchronization in a coupled oscillators network is $\|L^{\dagger}\omega\|_{\infty,E} \leq sin(\gamma)$ [18], where $L^{\dagger}$ is the Moore-Penrose pseudo inverse of the Laplacian matrix of the network, ω is the vector of the natural frequencies of the oscillators, and $\|x\|_{\infty,E} = max_{(i,j)\in E} |x_i - x_j|$. If this condition is satisfied, the oscillators synchronize at the common frequency ω_{SYNC} with the phase lock $|\theta_i - \theta_j| \leq \gamma \in [0, \frac{\pi}{2}]$, $\forall (i,j) \in E$.

On the other hand, the power that flows on line $(i,j) \in E$ can be written as $P_{i \longrightarrow j} = -P_{ij}^{MAX} sin\left(\theta_j(t) - \theta_i(t)\right)$. If $\theta_j(t) - \theta_i(t)$ is small, the flow constraints at time t are straightforward :

$$\forall (i,j) \in E, \ | \ \theta_j(t) - \theta_i(t) \ | \leq 1 \tag{7.5}$$

If these constraints are verified for all instants t during the control phase then no line gets overloaded by the action of the control inputs. Writing the synchronization condition in our settings thus gives :

$$\| \left(L \circ K\right)^{\dagger} \Psi \|_{\infty} \leq sin(1) \tag{7.6}$$

If constraint 7.6 is satisfied, the oscillators synchronize to a common frequency $\omega_{SYNC} = \frac{\sum_{k=1}^{N} \Psi_k}{\sum_{k=1}^{N} \alpha_k} = \frac{\sum_{k=1}^{N} P_{S,k}}{I\Omega N\alpha}$.

Since the dynamics is expressed in term of deviations from Ω, synchronization at Ω is achieved if $\omega_{SYNC} = 0$. Which gives the production consumption balance constraint :

$$\sum_{k=1}^{N} P_{S,k} = 0 \tag{7.7}$$

Constraints 7.5 7.6 and 7.7 ensure that the system is able to synchronize at Ω without overloading lines.

7.2.2 Battery Constraints

A storage i has a maximum charge/discharge rate r_i, some amount of energy stored $\Lambda_i(t)$ at time t, and a maximum capacity $\Lambda_{i,MAX}$ (see figure 7.2). We assume that all maximum rates (r) and all maximum capacities Λ_{Max} are the same across the storage equipments. We also denote by $\Lambda(t)$ the vector of energy level at time t. Since $u^{\star}(t)$ specifies the control inputs,

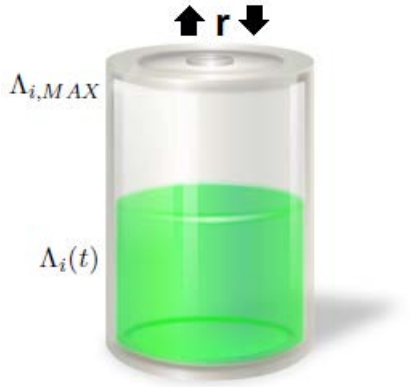


FIGURE 7.2: Simple storage model.

the energy level dynamic of the batteries is $\Lambda(t + \Delta t) = \Lambda(t) - u^\star(t)I\Omega$. Which can also be written as :

$$\Lambda(t) = \Lambda(0) - I\Omega \sum_{k=0}^{t} u^\star(k) \tag{7.8}$$

where $\Lambda(0)$ is the vector of initial levels in the batteries. Obviously, $\Lambda(t)$ has to stay within possible bounds : $\forall t,\ \ 0 \leq \Lambda(t) \leq \Lambda_{MAX}$. Which can be written in terms of the control inputs :

$$\forall t, -\frac{\Lambda(0)}{I\Omega} \leq -\sum_{k=0}^{t} u^\star(k) \leq \frac{\Lambda_{MAX} - \Lambda(0)}{I\Omega} \tag{7.9}$$

During each time slot, the battery cannot charge or discharge at a rate higher than r (see figure 7.2) :

$$\forall t,\ \ | u^\star(t)I\Omega | \leq r \tag{7.10}$$

Given the dynamic of equation 7.3 and the constraints of equations 7.5, 7.6, 7.7, 7.9, and 7.10, how can we select the driver nodes such that we use low control energy on average ?

7.3 Submodular set functions

In this section, we explain the method that we used in order to find the driver nodes for the grid's dynamics (eq. 7.3).

7.3.1 Gramian based optimization

Recall that $\mathcal{E}_{min}$ (see eq. 6.34) depends on the initial and final states as well as on the inverse of the gramian matrix W, which only depends on A and B. W can thus be used to obtain information about the average control energy required to move the system in the state space. In the prosumer network that we consider, generators and loads (Ψ) are susceptible to change, meaning that initial (X_0) and final (X_f) states might also vary. In such a scenario, we prefer to aim at good performance on average, rather than very good

performance in few specific cases and bad performance in all other situations.

Since the control energy is related to W^{-1}, systems with "large" W will tend to be controlled with low energy. Still, this notion of "large" W is not very accurate and we need to quantify it by defining some metric based on W. We already mentioned the rank of W as being the dimension of the controllable subspace, but it is also known that the trace of W and W^{-1} give information about average controllability and average control energy [87]. In such a situation, we can build a set function that, given a set of driver nodes S, returns the value of one of these metrics. Indeed, given A and S, we can obtain the system dynamics (A, B_S) and compute the gramian W_S. In other word, such a function would quantify the ability of a given set of nodes to control the system on average. The objective would then be to find the set $S_k^\star$ of size k that maximizes this function.
A key point, demonstrated in [87], is that the set function

$$F : S \longrightarrow Tr[W_S] \tag{7.11}$$

is modular and the two functions

$$\begin{cases} F : S \longrightarrow Tr[W_S^{-1}] \\ F : S \longrightarrow rank[W_S] \end{cases} \tag{7.12}$$

are submodular. As we will see in the next section, this nice result enables us to look for the driver node set that optimizes the average control energy with a simple greedy heuristic that also provides a worst case guarantee.

7.3.2 Submodularity

We introduce here submodular set functions and explain how their maximization can be achieved in reasonable time. More information about submodularity can be found in [43].

A set function $F : 2^V \longrightarrow \Re$ defined over a finite set V is said to be submodular if for all sets $X, Y \in V$, such that $X \subseteq Y$ and for all element $x \in V \setminus Y$, we have :

$$F(X \cup \{x\}) - F(X) \geq F(Y \cup \{x\}) - F(Y) \tag{7.13}$$

A very intuitive example is the optimum placement of sensors in an area (see figure 7.3). Sensors can be placed on a grid of locations and function F computes the surface of the space that is being sensed. Note first that adding a new sensor i to the current set S cannot decrease the value of F : $F(S + \{i\}) >= F(S)$. Furthermore, if we add sensor i to a small set S_1 we tend to get larger improvements than if we add i to a larger set $S_2 \supset S_1$.

This basically means that submodular functions exhibit a diminishing return property which makes them particularly interesting for optimization. Generally speaking, finding the set $S_k^\star$ of size k that maximizes a set function F is a difficult problem because the number of sets grows exponentially with the number of nodes. Therefore complete enumeration and evaluation is only feasible on very small examples. Nevertheless, if the set function is submodular, a simple greedy heuristic returns a solution

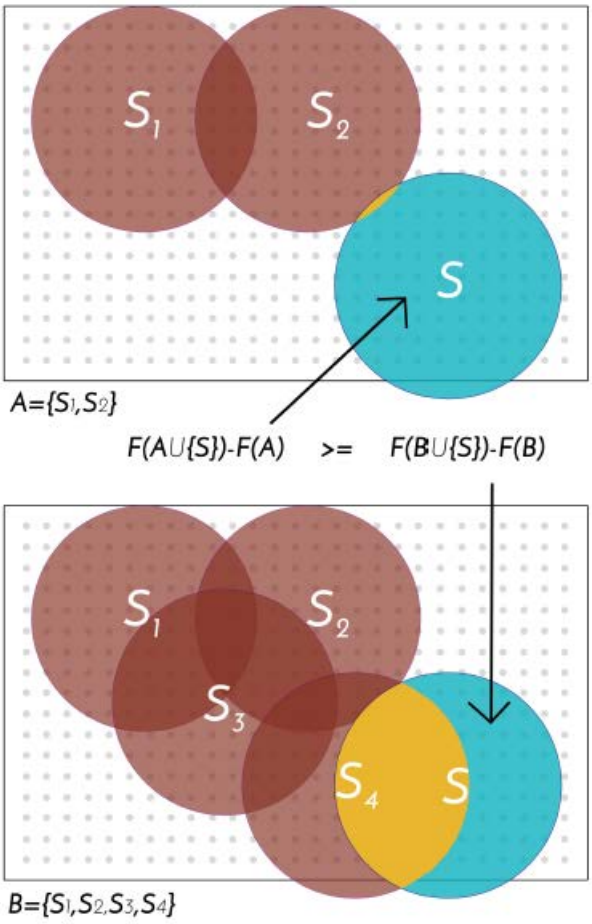


FIGURE 7.3: Submodularity example with sensor placements.

$S_k^\star$ such that, in the worst case, $\frac{F(S_k^\star)}{F(S_k^{OPT})} \sim 63\%$ (where S_k^{OPT} is the optimum set of size k) [43]. This heuristic starts with a set S (possibly empty) and iteratively adds the element i that exhibits the highest marginal gain : $F(S \cup \{i\}) \geq F(S \cup \{j\}) \, \forall j$.

For a ground set of N elements, this heuristic computes F $\frac{k(2N-k+1)}{2}$ times. Since the evaluation of F can be costly, a well-known lazy-greedy variation has been proposed by [55]. This smart implementation uses the submodular structure of the marginal gains in order to reduce the number of calls to F. This requires to maintain a sorted table of marginal gains for all elements. When looking for new element i to add to set S, the top one is selected and the new marginal gain $F(S \cup \{i\}) - F(S)$ is computed. If this gain is larger than the gain of the second element in the table, then i is added to S. Otherwise i is inserted back in the table with its updated gain and the same treatment is applied to the element that is now on the top of the table. Because of the submodularity of F, this method performs as well as the original one, but can result in speedups of several orders of magnitude.

Theoretically, we are now able, for a given prosumer network, to compute A, B, and W. For some trajectory in the state space, we can use equation 6.32 to find the optimum control inputs. However, these inputs do not take into account the physical constraints on the possible trajectories due to lines and batteries finite capacities.

7.4 Finding the optimal storage placement in prosumer networks

7.4.1 Performance evaluation

Finding the driver set S_k of size k using the gramian W_{S_k} of the system (A, B_{S_k}) does not require initial and final states. Besides, W_{S_k} does not depend on the modified power distribution Ψ of the nodes. We are indeed looking for k controllers that perform well on average over all possible situations. On the contrary, the constraints derived above (eq. 7.5, 7.6, 7.9 and 7.10) are bounded to a particular trajectory in the state space. We indeed check whether the power grid can sustain the controlled dynamics when transitioning from a given initial state to a target final state. As we do not know what these states could be, we could use multiple scenarios and test whether S_k can control the system without violation of constraints. In this paper, we use a slightly different approach. We consider that the system is initially at equilibrium (all elements are synchronized at Ω) and that control will be necessary if a perturbation (power imbalance) occurs and takes the system out of equilibrium (see figure 7.4). Depending on how much time we need to start the control phase, the state of the system (the initial state for control) might be "somewhere around" the synchronized state. In order to test whether a set S_k can control the system without violation of constraints, we sample initial states within some hypersphere centered on the synchronized state and check all the constraints. If for all trajectories, the constraints are respected, then we consider that S_k enables the control of the system.

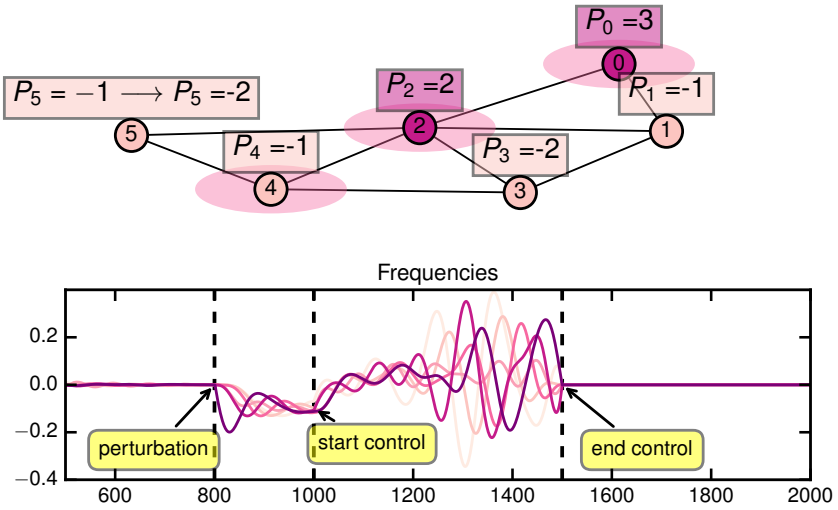


FIGURE 7.4: At $t = 800$, P_5 goes from -1 to -2 (power imbalance). Consequently, the frequencies $\dot{\theta}_i$ deviate from the synchronized state. At $t = 1000$, optimal control inputs (see equation 6.32) are injected at nodes 0, 2, and 4 (nodes with ellipses) such that the system is brought to the synchronized state at time $t = 1500$ (control time $T = 500$).

Increasing k means that we deploy more storage which increases the costs but tends to lower the energy required as we will see in the next section. Using the submodularity of the set functions introduced above, we can build a sequence of growing sets $S_1 \subset S_2 \subset ... \subset S_k$ and stop as soon as S_k enables the control of the system for some k.

Algorithm 1 Optimization with grid constraints

$k = 0$
$S_k = \emptyset$
while Not Constrained control **do**
  $k \leftarrow k + 1$
  $S_k = S_{k-1} \cup argmax_{i \in N \setminus S_{k-1}} F(S_{k-1} \cup \{i\})$
end while

The purpose of selecting the drivers according to a Gramian related metric is to minimize, on average, the amount of control energy needed. Conversely, if we select the drivers randomly, we expect to need, on average, more energy to control the system. In figure 7.5, we compare the average control energy $\mathcal{E}$ in function of the proportion of drivers among the total number of nodes $n_D = N_D/N_{nodes}$, for a set of drivers selected thanks to our algorithm with randomly selected drivers (in both cases overloading and battery limits constraints are satisfied). We draw 10^4 scale-free topologies with 50 nodes, and random power distributions and line capacities. For each system we select a random number of drivers $N_D \sim \mathcal{U}(1, N_{nodes})$ and we find two driver sets of size N_D. One is chosen randomly and the other is found with algorithm 1. For both sets, if the control is possible, we draw a random initial state Y_i and a random final state Y_f, and we compute the control energy required for driving the system from Y_i to Y_f. Since the

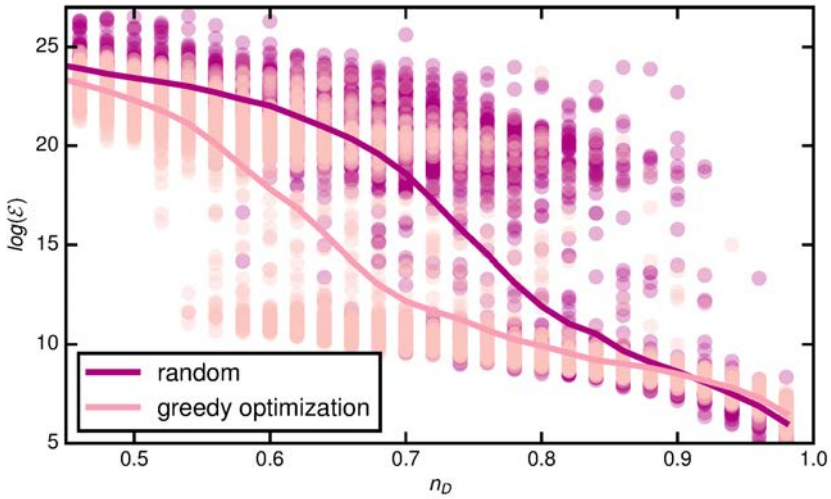


FIGURE 7.5: $log(\mathcal{E})$ against n_D for random and optimized driver sets ($N_{nodes} = 50$).

algorithm is based on $Tr[W^{-1}]$, we could plot $Tr[W^{-1}]$ in function of n_D but we prefer to validate our method by computing the actual average energy. As expected, both curves decrease as the number of drivers increases, meaning that, as the number of drivers grows, the average control energy tends to decrease, but we tend to use less energy when the drivers are selected with our algorithm. Note that this difference tends to zero when n_D tends to one, because almost all nodes are then selected, yielding little flexibility for optimization.

7.4.2 Topological effect on n_D

We investigate next how the topology of the grid and the physical constraints affect the minimum size of the driver set. n_D is no longer selected randomly, but minimized. We consider the simple case of an Erdös-Rényi topology with probability of connection p. We show on Figure 7.6 how the minimum size of the driver set evolves with p for different Gramian based metrics and for two levels of constraints :

- Level 1 : full control and grid constraints. The system is controllable (i.e it can be moved from any point to any other point) and under the constraints (i.e it can be moved without overloading any line or breaking any battery limits).
- Level 2 : full control only : the system can be moved from any point to any other point of the state space, without considering any overloading or battery limit constraints.

When $p \sim 0$, nodes tend to be very poorly connected such that we need to control almost all nodes in the grid. As p increases, the connectivity of the graph rises and the number of drivers decreases. At some point, the connectivity of the graph starts to harm its controllability, and more drivers are needed (this effect is in accordance with the literature). As expected,

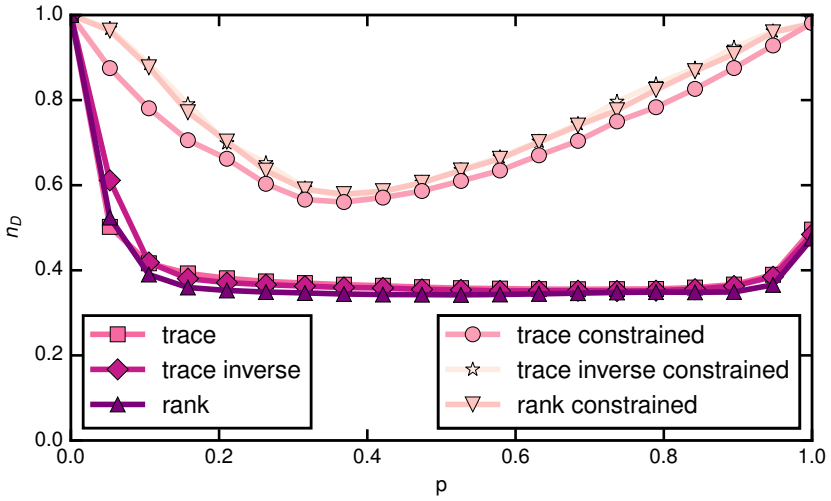


FIGURE 7.6: n_D against link probability p for erdos-renyi topologies ($N = 100$). Curves are averaged over 100 measurements. The top three curves show the results for the three metrics taken into consideration when all constraints are considered (see Level 1 in the main text). The bottom three curves exhibit the results for the same metrics when only the full controllability constraint is considered (see Level 2).

the driver sets for the level 1 of constraints are larger than for level 2 (for all metrics) because we impose far more constraints on the control inputs.

Erdos-Renyi or scale-free networks are known to provide poor approximations for real world power grids, which tend to have homogeneous degrees as well as spatial properties (see chapter 2). Indeed, power grids aim at routing power from generation sites to the consumers, which are both far from uniformly distributed over the different countries. In addition, power grids were built by governments or companies within each country without a global cross border thinking, which often results in well connected subgraphs within countries, and few connections across borders. These constraints lead to topologies exhibiting community structures. Therefore, we study the control of the prosumers in graphs with communities. More precisely, we use a block model $(N, N_{clusters}, p_{in}, p_{out})$ to generate random topologies where N is the number of nodes, $N_{clusters}$ is the number of clusters (or communities), p_{in} is the probability that two nodes within the same cluster are linked, and p_{out} is the probability that two nodes in two distinct clusters are connected (see top panel of figure 7.7).

The bottom left panel of figure 7.7 displays how the number of drivers evolves with the number of clusters for different values of p_{in} when p_{out} is fixed ($p_{out} = 0.1$). For small values of $p_{in} \sim p_{out}$ clusters are poorly marked, and the connectivity is low. The number of drivers in these conditions is large and increases slowly when the number of clusters augments. As p_{in} increases, the clusters becomes more densely connected such that within a cluster less drivers are required in order to control it. As the number of clusters grows, more drivers are required. For large values of p_{in}, the behaviour is more complex. We see indeed that for $p_{in} = 0.9$ and $N_{clusters} = 2$, almost

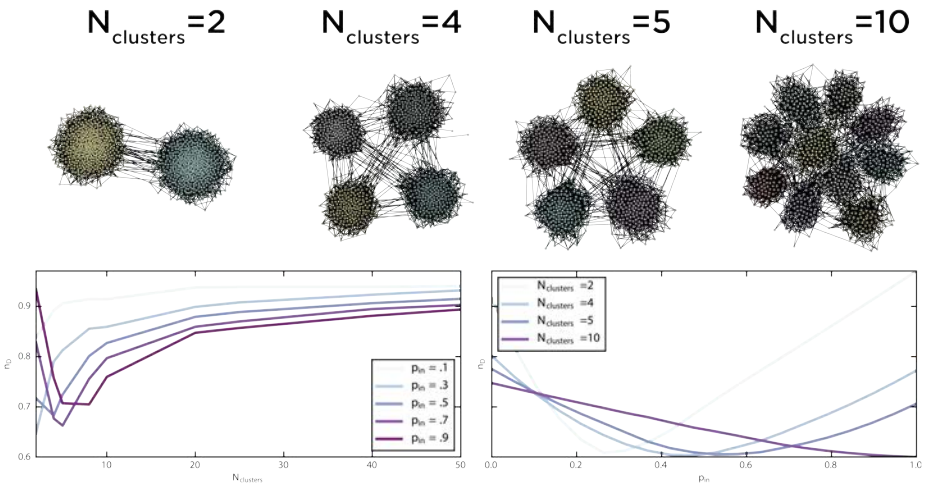


FIGURE 7.7: left panel : n_D vs $N_{clusters}$ for different values of p_{in} and $p_{out} = 0.1$. $N = 200$ and curves are averaged over 100 realizations. right panel : n_D against p_{in} for systems of $N = 200$, $p_{out} = 0.1$, and for different values of $N_{clusters}$. The curves are averaged over 100 realizations.

94% of the nodes are needed for control, but this quantity first decreases with the addition of a few more clusters (71% for $N_{clusters} = 5$), before increasing. This behavior can be explained by the fact that controlling a very dense network requires, in our settings, a very large portion of nodes to be drivers. Controlling one cluster alone thus requires to control almost all nodes within this cluster. Nevertheless, when a relatively small (compared to the number of nodes in the graph) number of clusters are interconnected with a few links, nodes in one cluster are able to control nodes in other clusters to which they are connected. As the number of clusters grows, the global connectivity increases rapidly, such that the number of drivers also rises. We show this behavior in more details on the bottom right panel of figure 7.7, where the number of drivers is plotted against p_{in} for $N = 200$ and $N_{clusters} \in \{2, 4, 5, 10\}$. For $N_{clusters} = 2$ and $p_{in} = 0.1$ the number of drivers is large since the graph is poorly connected. When p_{in} increases, n_D decreases until a minimum value $n_D \sim 0.6$ at $p_{in} \sim 0.25$. After this point, n_D augments with p_{in}. For $N_{clusters} = 10$, we do not see this behavior : n_D decreases continuously as p_{in} increases as expected from the curves of the bottom left panel of figure 7.7.

7.4.3 Effect of storage capacities

Until now, we have considered all batteries to be equal, i.e they have the same capacities and charge/discharge rates. Although simple, this assumption might not be true in practice where different types of batteries exist. Optimizing both the locations and cost over reliability tradeoff of different types of storage is out of the scope of this work. Nevertheless, we studied how the number of drivers behaves when the batteries have different capacities. More precisely, we draw the capacities of the batteries from a normal distribution $\mathcal{N}(\mu_\lambda, \sigma_\lambda)$ and keep a simple erdos-renyi topology for the power grid. Note that each node is assigned a capacity regardless of its

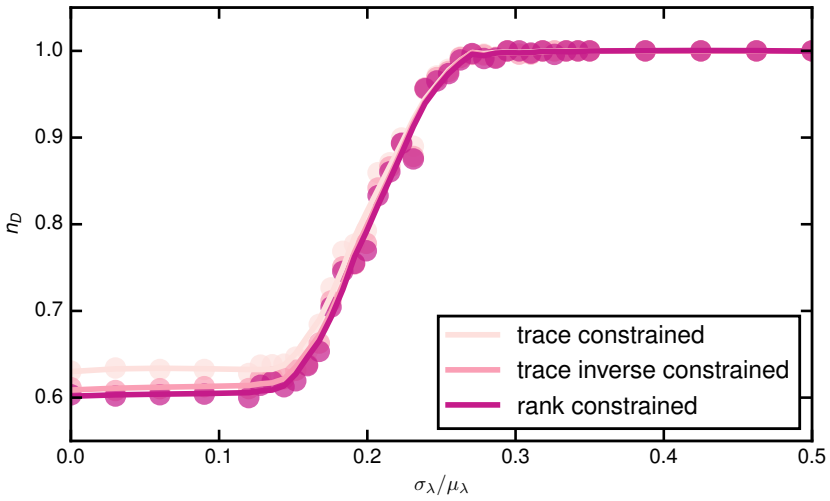


FIGURE 7.8: n_D vs the relative standard deviation of the capacity distribution of the batteries $\frac{\sigma_\lambda}{\mu_\lambda}$. The dots are averages over 100 measurements and the curves are obtained using a Savitzky-Golay filter, each color shows the results for a given metric. The graphs are erdos-renyi with $N = 100$ and $p = 0.3$, and $\mu_\lambda = 100$ units.

characteristics (degree, betweeness, and so on...). In figure 7.8, we plot n_D as a function of the relative standard deviation of the battery capacity distribution $\sigma_\lambda/\mu_\lambda$. We see that n_D increases with the variance of the capacity distribution for all three gramian based metrics considered. Moreover, this rise is abrupt since n_D goes from 60% for $\frac{\sigma_\lambda}{\mu_\lambda} \sim 0.15$ to 100% for $\frac{\sigma_\lambda}{\mu_\lambda} \sim 0.27$.

7.4.4 Real topologies

Since real power grids topologies are known to be far from random, we consider here, as a real case, the European transmission power grid which topology was obtained from [40]. The network contains 1494 nodes and 2156 edges and spans 25 European countries. A representation of this graph can be found in figure 7.9 where the nodes are colored according to their country. We select the power distribution and the line capacities randomly, and use the method described above.

We use the trace of the Gramian matrix for the optimization. The top panel of figure 7.10 shows that the rank of the Gramian increases with the size of the controllers until reaching full rank ($rank[W_S] = 2N_{nodes}$) for $N_D = 1244$. Above this point, we compute the average control energy needed to drive the system from an initial random state to a random final state. The curves on the bottom panel of figure 7.10 show the results for the optimized driver sets and for random driver sets. As can be seen on figure 7.10, as the size of the driver set increases, the control energy tend to decrease. But, we need less energy with the optimized driver set than with randomly sampled sets.

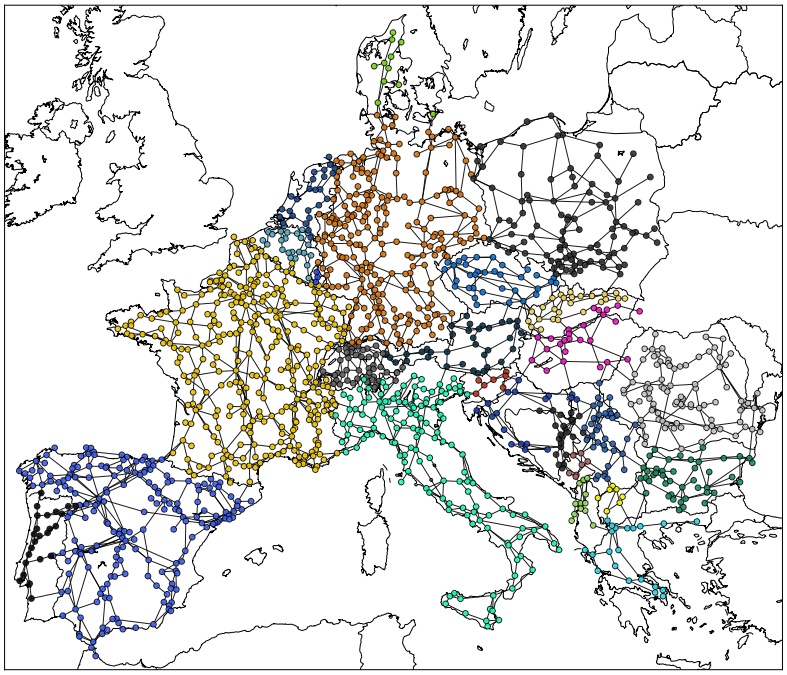


FIGURE 7.9: European transmission power grid [40]. Nodes are colored according to their county.

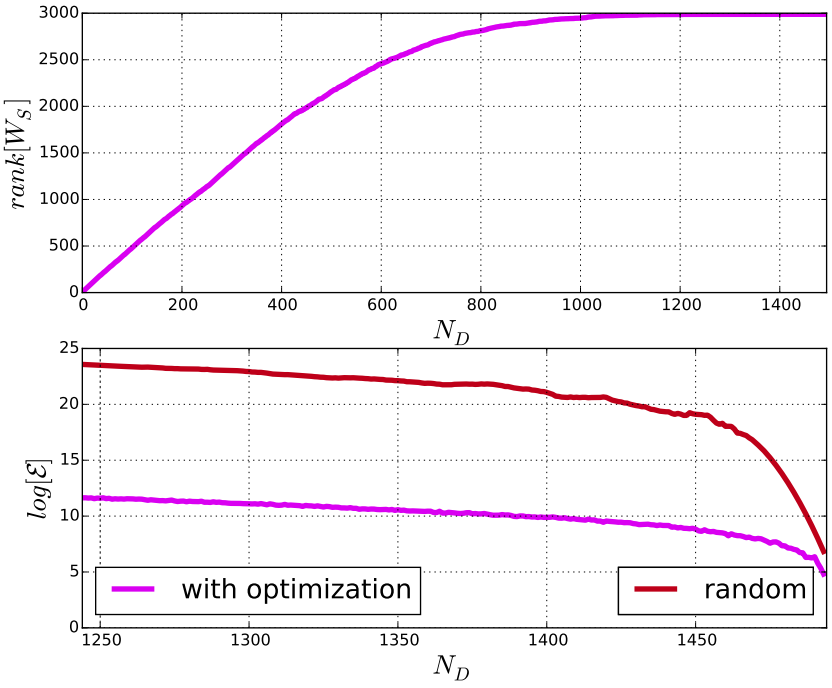


FIGURE 7.10: Gramian rank and control energy evolution for European transmission power grid (see figure 7.9).

7.5 Discussion

In this chapter, we considered networks of prosumers which can behave as generators or loads depending on weather conditions. In this stochastic context, we proposed an optimization method of the storage placement for frequency regulation through optimal control theory (see chapter 7). The modeling approach developed in chapter 6 enabled us to simplify the grid's dynamic such that the framework of optimal control could be used. Besides, the physical constraints on the network and storages were taken into account in the process.

Because we do not consider any prior on the perturbations that may occur, we choose an approach that aims at minimizing the control energy in any direction of the state space. The optimization method relies on submodular set functions (see section 7.3) and the gramian matrix (see chapter 7). We show in this chapter that, by using our algorithm, we are able to find a storage placement that, on average, minimizes the energy that should be injected or absorbed by the storage devices while complying with the physical constraints of the grid and storages.

We believe that interesting work could be done by combining this model with real production and consumption data. There are indeed complex spatial and temporal correlations that impact these distributions [28]. In contrast, we considered that the control has to be in any direction of the state space but confronting this assumption with real traces could lead to interesting further research.

The next chapter first gives a short summary of the present thesis, and then addresses the ongoing works. Finally, the main ideas of some future works that are still at an early stage of development are presented.

Chapter 8

Conclusion

This thesis was devoted to study smart grids with a complex system approach. Power grid systems already have a significant array of control tools to operate in a stable and reliable way. In the current unidirectional and static systems, these tools are able to maintain stability. But in more complex, interconnected, and distributed environments, there is need for smart and automated actions. Real time data could be obtained by sensors scattered across the network. Distributed system architectures for storing and processing these large amount of data could be combined with machine learning algorithms that would predict events and reinforce the control of the system.

In such multi-scale and interdependent environments it might be difficult to understand how the whole system behaves. Complex system theory addresses these kinds of problems, where some collective phenomena appear from individual behaviors. While a large amount of work focuses on very specific components of the smart grids that could be tested in isolation, this work used relatively simple models and assumptions for the individuals and studied their repercussions at a larger scale.

This thesis was divided in two main sections. The first one focuses on modeling energy markets with aggregation of prosumers as participants, while the second part addresses the optimization of storage placement in order to achieve frequency regulation in a network of prosumers.

The main contributions for the first part were the development of a prosumer model based on real weather data that enabled us to capture the spatio-temporal correlations between these agents. A model for the energy market was also proposed and used for studying how market aggregators could group prosumers in such a way that they maximize their expected gains at low risk. We argued that diversification was a key concept in this process, such that we studied a de-correlation based clustering. We proposed a greedy optimization that starts with uncorrelated cliques of agents, and increases their sizes as long as their utilities are improved. We then compared the results against other methods and showed that the proposed algorithm was performing better than the others. We believe that methods based on this work could be used by aggregators on energy markets as to maximize their expected gains for a given risk acceptance. Using real prosumers data when they will become available could help improve the proposed method. Furthermore, we suggest in chapter 5 a possible direction to gain insight on the market and improve the method. We indeed proposed to use tensor decomposition to study dynamic correlation graphs as opposed to static versions of these graphs. Being able to understand the patterns in the dynamics of such graphs would surely provide precious

information to the market aggregators.

Concerning the second part, we believe that the novel contribution was first to combine the power grid dynamic with optimal control as to answer the question of storage placement in a stochastic context. In addition, we incorporate the physical constraints of the power grid and energy storage devices in the process. We also explored an optimization method based on submodular set functions that is easy to implement and has a worst case guarantee. The proposed algorithm is thus able to produce an efficient placement strategy given a topology for the power grid. It seems unavoidable that large amounts of storage devices would have to be installed in future smart grid systems. Optimizing their locations beforehand could therefore lead to important savings. This question has been tackled in the literature before, but never, to the best of our knowledge, by this optimal control and submodularity approach. We believe that this work brings a different perspective on this question, and we hope that more work would be undertaken in this direction.

In terms of publications, the research performed for the first part resulted in a conference paper published at the IEEE ICC 2015 as well as a journal paper published in IEEE Transactions on Smart Grid 2016, both as first author (see appendix A for a complete list of publications). For the second part, a conference paper has been published at IEEE SmartGridComm 2016. The work presented in chapter 6 on dynamic correlation graphs is still incomplete because it was undertaken short before the redaction of the present thesis. This work is currently being completed and will hopefully be submitted around September 2016.

My current work also includes the mentoring of M. Lester Padilla, a master level intern, who works on extracting power grid topologies from OpenStreetMap (OSM) data. This internship was created after an exciting seminar organized in our offices in March 2016 and ended in September 2016. The idea is based on the SciGRID project developed at *NEXT ENERGY, EWE-Forschungszentrum für Energietechnologie* in Germany. The internship aims at extracting, from the OSM data, a power grid topology with generators and loads. With additional information like the type of the generators, their capacities, the capacities of the electrical lines, and so on, we believe that this work will help researchers to obtain realistic data to test their research on.

On a more distant perspective, because of my ongoing work on temporal graphs, I developed interest for studying the detectability of evolving communities in temporal graphs. This interest is still at a very early stage and I hope to have time in a near future to focus on it. Another future work might also be related to the *Grid4Earth* project on microgrids.

Besides the work done as part of my thesis, I also gave courses (monitorat) at the Pierre et Marie Curie University (UPMC) during these three years. In the beginning, this was the opportunity for me to experience teaching because I was not sure, right from the start, that it would suit me. In the end, it confirmed to me that this is something that I particularly enjoy. Looking back on these last years, I feel very fortunate to have had the opportunity to learn and experience a lot of fascinating things. It is also frustrating to have so many things on the "to-do" or "to-look-at" lists and so little time to explore them in depth. Although it was in the back of my head

since the beginning of my PhD, now, I know for sure that I wish to work in academia as a professor in the future.

Appendix A

List of publications

- **Conferences :**
  - *Coalition formation algorithm of prosumers in a smart grid environment*, Nicolas Gensollen, Monique Becker, Vincent Gauthier, Michel Marot, 2015 IEEE International Conference on Communications (ICC). DOI: 10.1109/ICC.2015.7249262. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7249262
  - *Submodular Optimization for Control of Prosumer Networks*, Nicolas Gensollen, Vincent Gauthier, Michel Marot, Monique Becker, 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), submitted awaiting decision.
- **Journals :**
  - *Stability and Performance of Coalitions of Prosumers Through Diversification in the Smart Grid*, Nicolas Gensollen, Vincent Gauthier, Monique Becker, Michel Marot, IEEE Transactions on Smart Grid (Issue 99). DOI: 10.1109/TSG.2016.2572302. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7479542

Appendix B

Thesis abstract in french

Au cours de ces dernières années, il semble que la modernisation des systèmes électriques soit devenue une nécessité. Alors que les technologies ont évolué de façon significative depuis le déploiement des réseaux électriques, leur fonctionnement et les méthodes sont restés plus ou moins les mêmes. Les progrès sur les réseaux de communication et de gestion des données pourraient ouvrir les portes à une nouvelle génération de systèmes de gestion de l'énergie. La souplesse procurée par ces nouvelles technologies pourrait changer radicalement la manière dont les réseaux électriques sont pilotés. Les énergies renouvelables, souvent considérées comme peu fiables, apportent également leur part de complications. En effet, comment assurer la stabilité du système si la production et la consommation sont stochastiques? Au sein de systèmes unidirectionnels et rigides, cela peut sembler compliqué. Toutefois, en considérant l'utilisateur final non plus comme une charge fixe, mais plutôt comme un acteur impliqué, il semble que des solutions pourraient émerger.

Cette thèse est consacrée à l'étude d'agents appelés prosumers parce qu'ils peuvent, par l'intermédiaire de générateurs basés sur des énergies renouvelables, à la fois produire et consommer de l'électricité. Il est concevable que, si leurs productions excèdent leurs propres besoins, les prosumers chercheront à vendre leur surplus. Bien entendu, ces excédents sont extrêmement volatiles, car ils dépendent à la fois des conditions météorologiques à un instant donné, mais aussi de la consommation des agents. Étant donné qu'il n'existe pas, pour le moment, de données réelles précises sur la dynamique de ces excédents, le début de cette thèse a été consacré à la modélisation et à la simulation de ces prosumers. Dans un souci de réalisme, nous nous sommes basés sur des données météorologiques réelles plutôt que sur des distributions aléatoires classiques. Ces simulations nous ont permis d'obtenir des séries temporelles décrivant comment les excès de productions (positifs ou négatifs) évoluent pour différents agents.

L'étude de ces séries a mis en évidence des corrélations spatio-temporelles non triviales qui sont d'une grande importance pour les agrégateurs. Comme leur nom l'indique, ces agents forment des portefeuilles de générateurs, de charges électriques, et de moyens de stockage afin de vendre des services à l'opérateur du réseau. Ces services sont destinés à maintenir la stabilité du système, que ce soit par de la régulation de fréquence ou en équilibrant production et consommation. En regroupant divers prosumers, un agrégateur peut chercher à stabiliser sa production. Cela est primordial car un agrégateur, lié par un contrat avec l'opérateur, peut faire l'objet de sanctions s'il est incapable de remplir son rôle. Nous montrons dans cette thèse que la structure de corrélation entre les prosumers a un impact direct et décisif

sur la stabilité des agrégats, et donc sur le risque encouru par les agrégateurs. Ainsi, nous proposons un algorithme pour minimiser le risque d'un ensemble d'agrégations, tout en maximisant leur gain attendu. Cette méthode gourmande est basée sur la théorie des graphes et permet des gains significatifs par rapport à des approches plus aléatoires.

Si le marché de l'énergie et les interconnexions entre les pays sont censés aider à stabiliser le système, il est encore très probable qu'une forte augmentation de la capacité de stockage sera nécessaire. La mise en place de dispositifs de stockage dans un réseau où les générateurs et les charges sont dynamiques et stochastiques apparaît comme un véritable défi. En effet, la dynamique non linéaire du réseau électrique et la multitude de perturbations possibles rendent l'existence d'un placement globalement dominant plutôt improbable. Il semble plus approprié de rechercher une stratégie avec une bonne performance en moyenne des situations possibles. Compte tenu de la taille de l'espace d'état, une méthode non-exhaustive avec une garantie semble être nécessaire. Nous proposons de répondre à cette question avec une approche basée sur la théorie du contrôle dans les réseaux. Nous modélisons le système électrique par un réseau d'oscillateurs couplés, dont la dynamique des angles de phase est une approximation de la dynamique réelle du système. Le but est de trouver le sous-ensemble des nœuds du graphe qui, lors d'une perturbation du système, permettrait le retour à l'équilibre si les bons signaux y étaient injectés. Nous montrons que ces signaux peuvent être interprétés comme la puissance à laquelle les dispositifs de stockage doivent injecter ou absorber de l'électricité. Nous cherchons ensuite le placement qui permet de minimiser l'énergie moyenne de contrôle. Nous proposons un algorithme, basé sur l'optimisation de fonctions sous-modulaires, pour trouver un placement proche de l'optimum. Nous montrons également qu'il existe une limite qui contraint le résultat de l'algorithme, de telle sorte que nous avons une garantie d'être proche de l'optimum.

Appendix C

Modéliser et optimiser un réseau électrique distribué, une approche systèmes complexes des prosumers dans le smart grid

C.1 Introduction

C.1.1 Vers un nouveau système énergétique

Au cours de ces dernières années, il semble que la modernisation des systèmes électriques soit devenue une nécessité. Les installations crées au cours du XXème siècle atteignent actuellement la fin de leur durée de vie théorique et devraient être, en toute logique, réparées et modernisées. Toutefois, depuis leur déploiement initial, de vastes progrès ont été réalisés au sein des technologies de l'information. Internet, les réseaux mobiles, ou encore les objets connectés pour ne citer qu'eux, sont le fruit d'un développement effréné du monde le l'informatique au sens large.

Parallèlement à ce boom, les systèmes électriques n'ont que peu changé depuis leur déploiement. Bien entendu, la manière dont ces systèmes sont gérés a évolué, en utilisant largement les possibilités offertes par ces nouvelles technologies. Toutefois, l'architecture de base, c'est-à-dire la philosophie initiale avec laquelle ces systèmes ont été établis est restée globalement la même. De façon simplifiée, la production d'imposantes centrales réparties sur le territoire est planifiée suivant la prévision de la consommation des utilisateurs. Cette production de base est ensuite affinée le jour J grâce à des générateurs pouvant être allumé ou éteint beaucoup plus facilement et rapidement. Durant ces dernières décennies, ce mode de fonctionnement a fait ses preuves et a l'avantage d'être relativement simple et d'apporter une stabilité satisfaisante au système. Pourquoi vouloir alors révolutionner un système qui fonctionne ?

Il existe un certain nombre de raisons plus ou moins profondes pour envisager une transition vers un nouveau type de système. Premièrement, les centrales et générateurs évoqués jusqu'à maintenant fonctionnent grâce à des sources d'énergies fossiles (nucléaire, charbon, pétrole…). Ce type de générateur possède l'avantage de pouvoir être programmé à l'avance. Il est alors relativement aisé de s'adapter à la consommation des utilisateurs du réseau. Cependant, il est désormais bien connu que ces sources

fossiles ne sont pas inépuisables et qu'elles ont été fortement surexploitées ces dernières décennies. Par ailleurs, il a été démontré que leur exploitation engendrait un certain nombre de problèmes environnementaux. Nous sommes donc arrivés à un moment clé de l'histoire des systèmes énergétiques puisqu'il apparait essentiel d'investir pour leur modernisation. Le point crucial étant de savoir si ce budget doit supporter la réinstallation de dispositifs dont les sources s'amenuisent et tendent à aggraver la situation climatique, ou s'il doit soutenir des projets alternatifs innovants.

La question est donc de déterminer s'il existe des alternatives crédibles à nos systèmes actuels. On souhaiterait idéalement pouvoir jouir des mêmes services sans en avoir les conséquences problématiques. Comme la source du problème semble provenir de la génération, on pourrait imaginer le résoudre en remplaçant les sources d'énergies utilisées. Il est en effet connu depuis longtemps que certaines énergies peuvent être considérées, à notre échelle temporelle, comme inépuisable. L'énergie solaire, le vent, le courant des fleuves, les vagues et marrées, ou la géothermie ne sont que les exemples les plus évidents. Ces énergies, dites renouvelables, ont l'avantage d'être non polluantes, gratuites, et inépuisables de telle sorte qu'on pourrait même s'étonner de ne pas avoir construit dès le début nos systèmes énergétiques autour de celles-ci.

En réalité, bâtir un système énergétique de grande taille en utilisant uniquement des sources renouvelables est un problème qui attend encore, au jour d'aujourd'hui, une solution. En effet, un important défaut de ces énergies renouvelables est qu'elles sont généralement intermittentes et stochastiques. Contrairement aux générateurs fossiles que l'on peut programmer à l'avance, les générateurs renouvelables ne produisent que lorsqu'ils sont soumis à la ressource en question. Hors il est bien connu que l'intensité de la plupart de ces ressources est tout sauf constantes. Le rayonnement du soleil par exemple est nul durant la nuit, et fortement variable durant la journée car il dépend à la fois de l'heure, de l'orientation, de la nébulosité, etc. Il en va de même avec le vent qui apparait comme une ressource extrêmement volatile.

Les consommateurs, qui sont la raison d'être des systèmes énergétiques, n'ont généralement que peu conscience de ces problématiques. Dès lors que ceux-ci payent leur abonnement, ils considèrent la disponibilité de l'énergie comme acquise, et ce quelle que soit l'heure, le jour, ou la saison. Il semble donc extrêmement compliqué de fournir ces utilisateurs avec des moyens de production aussi peu fiables étant donné l'architecture de nos systèmes actuels. C'est l'une des problématiques clé de la transition énergétique, et c'est précisément là que les technologies de l'information ont beaucoup à apporter pour provoquer une mutation en profondeur des réseaux énergétiques.

C.1.2 L'utilité de l'information dans le smart grid

L'architecture des réseaux électriques actuels est hiérarchisée et relativement rigide. Les générateurs ont été construits à proximité des ressources qu'ils exploitent, c'est-à-dire souvent loin des utilisateurs finaux, si bien qu'il est nécessaire d'acheminer l'électricité depuis ces sites distants jusqu'aux clients. Afin de transporter l'électricité sur de longues distances, il est judicieux d'élever sa tension par l'intermédiaire de transformateurs, car cela

a pour conséquence de fortement réduire les pertes par effet Joules lors du transport. Les réseaux très haute tension maillent généralement les pays en fonction de la répartition des populations et des ressources et constituent les squelettes des réseaux électriques. L'opération inverse est ensuite effectuée pour réduire la tension jusqu'aux niveaux d'utilisation des clients. Ce type d'architecture « top-down », où l'électricité ne circule que dans un sens, possède un certain nombre de défauts que nous évoquerons plus loin. Par ailleurs, il n'existe que très peu d'échanges d'information entre les entités opérant le réseau et les utilisateurs finaux. Au mieux, les compteurs électriques individuels ne sont relevés que tout les six mois afin d'ajuster les montants des factures.

Pourtant, au niveau des utilisateurs, la consommation n'est pas constante et dépend de nombreux paramètres comme le jour, la saison, la température, les habitudes des clients, etc... Il existe néanmoins des tendances générales bien connues. Les pics de consommation du matin et du soir sont notamment des phénomènes émanant de la corrélation des comportements. Si ces pics sont prévisibles, ils n'en restent pas moins problématiques puisqu'ils obligent les opérateurs à dimensionner le réseau pour y répondre alors même que leurs durées sont assez brèves. Lisser ces pics est un objectif déjà ancien au sein du monde de l'électricité, mais cela apparait encore plus important pour un système basé sur des sources renouvelables. L'exemple de la production photovoltaïque est extrêmement parlant puisque le pic de production, situé entre midi et deux heures de l'après-midi, est justement placé au milieu des deux pics de consommation. Autrement dit, si rien n'est fait, le système est constamment en déficit ou en excédent de production, ce qui ne peut être acceptable pour des raisons évidentes de stabilité.

Contrairement aux générateurs fossiles, il n'est généralement pas possible d'agir directement sur la production renouvelable pour la décaler dans le temps, si bien que d'autres solutions moins directes doivent être envisagées. L'une des plus évidentes consisterait à constamment stocker et déstocker les surplus de production pour s'ajuster à la demande. Même si le stockage de l'électricité est un des axes clé des réseaux électriques futurs, il est probablement inconcevable de se baser uniquement sur cela. En effet, l'électricité est une grandeur qui se stocke mal et dont les équipements ont généralement des durées de vie assez faibles (batteries). Le coût d'investissement et d'entretien d'un tel parc de dispositifs de stockage apparait donc comme prohibitif. Néanmoins, il est assez clair désormais que le stockage de l'électricité, dans des proportions beaucoup plus importantes qu'aujourd'hui, sera nécessaire dans le futur.

Une des idées fondamentale du smart grid est donc d'associer l'utilisateur final à l'équilibre du réseau, en ne le considérant, non plus comme une « charge morte », mais comme un acteur à part entière. Cela suppose un échange constant d'information entre les divers acteurs du système que seul un réseau de communication peut fournir (certaines études proposent aussi d'utiliser le réseau électrique grâce au CPL). L'idée repose sur le déploiement massif de compteurs intelligents qui font l'interface entre le réseau et l'utilisateur. Ces dispositifs ont pour rôle, entre autres, de mesurer régulièrement la consommation et de la transmettre en amont sur le réseau. Cela permet à l'opérateur, par agrégation, de connaitre en quasi temps réel la distribution de la consommation dans le système.

Par ailleurs, l'opérateur peut également envoyer des informations aux compteurs afin de les avertir des conditions sur le réseau. L'un des proxys envisagés pour répercuter ces conditions sur l'utilisateur est un prix de l'électricité dynamique. En fonction de l'état du système et de ses prévisions à court termes, l'opérateur fixe les prix pour les périodes futures et les diffuse sur le réseau. Grâce à ces informations, les compteurs intelligents sont en mesure de planifier les divers appareils afin d'optimiser une fonction d'utilité propre à chaque utilisateur. En effet, il apparait que certains appareils électroménagers comme les machines à laver, lave-vaisselles, ou climatisations peuvent être reporté dans une certaine mesure. A contrario, d'autres appareils comme la télévision, ou l'éclairage doivent être opérationnels à la demande pour des raisons évidentes de confort. Il est ainsi envisageable de déclencher certaines consommations uniquement lors de périodes où les prix sont faibles si le gain monétaire réalisé dépasse la perte utilitaire engendrée. C'est globalement l'idée du « Demand Side Management (DSM) », implémenté sur le compteur, qui tire profit de la dynamicité des prix de l'électricité.

Demand Side Management, prix dynamiques, et stockage distribué de l'électricité sont des concepts de bases des réseaux électriques du futur. Cependant, une grande partie de la littérature consacrée au smart grid va encore plus loin dans le rapprochement entre la production et la consommation.

C.1.3 Une production décentralisée et un réseau plus résilient

Comme évoqué plus haut, la production et la consommation sont aujourd'hui clairement séparées, tant au niveau géographique que topologique. Par ailleurs, cette production est assez centralisée, notamment en France où la majeure partie provient des centrales nucléaires. Sous ses apparences de stabilité, cette organisation est en réalité extrêmement fragile et vulnérable puisque le fonctionnement du système entier repose sur quelques points et liens très haute tension.

L'étude de la topologie des réseaux électriques a connu un regain d'intérêt ces dernières années grâce au développement des systèmes complexes. Le réseau électrique peut en effet être facilement abstrait en un graphe où les arêtes représentent les lignes électriques, et les nœuds les divers équipements (générateurs, sous-stations, transformateurs...). Analyser ensuite les propriétés structurelles de ces graphes permet de mieux comprendre comment ceux-ci tendent à se comporter. Prenons le « World Wide Web » comme exemple, où les pages internet représentent les nœuds, et les liens hypertexte forment les arrêtes. Intuitivement, on imagine qu'il doit exister une sorte de diversité du nombre de liens par pages. Sous la forme de graphe, cela revient à étudier la distribution des degrés des nœuds du graphe. Un résultat connu est que le Web possède une distribution des degrés en loi de puissance, ce qui signifie qu'une grande majorité des nœuds ont des degrés assez petits mais qu'il existe des nœuds, appelés « hubs », avec des degrés beaucoup plus grands. Pour des graphes purement aléatoires (Erdös-Rényi) la probabilité de trouver ce genre de nœuds est quasiment nulle, de telle sorte que la présence des ces hubs n'est en aucun cas le fruit du hasard, mais le résultat d'un processus sous-jacent. Une des conséquences de cette distribution en loi de puissance est que le Web est très résistant aux

pannes aléatoires, mais extrêmement vulnérable aux attaques ciblées sur ces fameux hubs. Pour s'en convaincre peut-être faut-il s'imaginer surfer sur une toile dépourvue de Google, Yahoo, ou Facebook...

L'architecture un peu particulière des réseaux électriques rend leur classification moins évidente que pour d'autres réseaux réels. Par ailleurs, cette approche purement topologique ne prend pas en compte la dynamique de l'électricité sur le réseau, si bien que d'autres approches ont du être mises au point spécialement pour étudier la résilience des réseaux électriques. Ces travaux se focalisent généralement sur les phénomènes de cascades qui représentent les répercutions d'une perturbation initiale. Il a ainsi été montré que la chute de seulement quelques lignes pouvait totalement fragmenter le réseau jusqu'à le rendre non fonctionnel. Ces « black-out », aux conséquences parfois catastrophiques, ne nous sont pas étrangers et se produisent de temps à autre. La tendance étant par ailleurs à l'aggravation avec la complexification et le vieillissement des systèmes électriques. Une question cruciale est ainsi d'estimer la résilience d'un réseau électrique intelligent dont la production est en grande majorité renouvelable. Toutefois, afin de pouvoir parler de sa résilience encore faut-il avoir une idée de sa topologie, des emplacements de ses générateurs, etc...

Si de larges centrales renouvelables raccordées au réseau très haute tension seront surement construites (c'est d'ailleurs déjà le cas actuellement) afin de tirer profit des économies d'échelles ou de ressources éloignées (éolien offshore par exemple), de nombreux articles prônent le développement parallèle d'une production locale. En incluant des moyens de production au plus proche des consommateurs on diminue fortement le besoin de transport, et donc les pertes que celui-ci engendre. Par ailleurs, un réseau dont la production est décentralisée apparait comme beaucoup plus résilient car, même fragmenté, il peut éventuellement continuer à fonctionner. Bien entendu ce genre de réseaux est encore au stade de la recherche, mais des applications moins ambitieuses, comme les micro-grilles, sont d'ores et déjà en train de voir le jour. Ces micro réseaux électriques sont composés de générateurs, charges électriques, et moyens de stockage, et sont reliés au réseau principal par un point unique (« Point of Common Coupling », PCC). Ils ont ainsi la possibilité de se déconnecter du réseau pour opérer en autosuffisance pendant une période donnée. L'intérêt pour ce genre d'opérations peut être, par exemple, de soulager le réseau principal lorsqu'il opère proche de sa limite, permettant ainsi de se prémunir contre d'éventuels black-out.

C.1.4 Les prosumers

En poussant ce concept de production locale encore plus loin, on pourrait imaginer que l'utilisateur lui-même produise une partie de son électricité. C'est d'ailleurs déjà le cas actuellement avec l'augmentation des installations photovoltaïques sur les toits des lotissements. Ces utilisateurs, souvent appelés « prosumers », ont donc la possibilité de consommer « gratuitement » ce qu'ils produisent. Comme expliqué plus haut, il est néanmoins fort probable que leurs activités ne coïncident pas vraiment avec leur production, si bien qu'à moins de posséder également une batterie, cette production est perdue. Bien entendu, il existe un certain nombre de parades

comme retarder certains appareils (DSM), ou utiliser l'énergie pour chauffer l'eau par exemple. Toutefois, il est envisagé que ces prosumers puissent revendre leur production lorsque celle-ci est en excédent. Cela est loin d'être évident à mettre en place et doit faire l'objet d'importantes études tant sur le plan électrique que sur le plan économique. Contrairement aux flux électriques « top-down » décrits plus haut, on envisage ici des flux bidirectionnels lourds de conséquences sur l'architecture et la stabilité du système.

Cette thèse est centrée sur ces utilisateurs et leur intégration au sein du système électrique intelligent. Nous abordons ces prosumers suivant deux angles d'attaques qui donnent naissance aux deux parties principales de cette thèse. Nous commençons par étudier comment ces-derniers peuvent parvenir à revendre leurs surplus de production sur des marchés de l'électricité. Dans un second temps, nous nous focaliserons sur la stabilité du réseau électrique en présence de ces prosumers. Nous chercherons alors à installer des dispositifs de stockage à des points stratégiques du réseau afin de maintenir l'équilibre du système lors de fluctuations liées à la stochasticité des énergies renouvelables utilisées. Tout au long de la thèse des concepts liés à la théorie des graphes et des systèmes complexes seront utilisés. Ces concepts sont expliqués dans le chapitre 2 de la thèse, et de nombreuses références bibliographiques y sont fournies.

C.2 Formation de coalitions stables pour les marchés de l'électricité

C.2.1 Modélisation des prosumers

Comme expliqué précédemment, un prosumer possède un ou plusieurs petits générateurs renouvelables, des appareils consommant de l'électricité, et éventuellement un moyen de stockage (qui est considéré dans certaines études comme la batterie d'un véhicule électrique). Intuitivement, on comprend qu'un certain nombre de paramètres, comme la météo, la saison, le jour de la semaine, l'heure de la journée, ou le comportement du prosumer influent sur le surplus de production de ce dernier. Comprendre comment ce surplus évolue dans le temps pour divers prosumers est au centre de notre problématique de revente sur les marchés. Toutefois, étant donné que les prosumers sont encore des projections théoriques de ce que pourrait être les futurs consommateurs, il n'existe que très peu de données les concernant. Même si cela tend à s'améliorer au fil des années, nous n'avions pas, au début de cette thèse, de données sur lesquelles nous baser.

Une partie non négligeable de cette thèse fut ainsi consacrée à modéliser et simuler ces agents par l'intermédiaire d'un autre type de données plus facilement accessibles : les données météorologiques. Outre l'accessibilité de celles-ci, nous avons fait ce choix car la vitesse du vent et l'ensoleillement apparaissent comme des proxys raisonnables pour modéliser des productions éoliennes et photovoltaïques. Par ailleurs, les stations météorologiques forment généralement un maillage assez bon des territoires que nous avons considéré, ce qui nous permet d'avoir une discrétisation de l'espace acceptable. On effectue ainsi un découpage de Voronoï du territoire considéré autour des stations météorologiques pour lesquelles nous avons récolté

des données. On dispose ensuite de manière aléatoire un ensemble de N prosumers et on associe chaque agent au profil météorologique de la cellule sur laquelle il se trouve. Pour n'importe quel instant t de la simulation on peut donc avoir accès, pour un agent i, à la vitesse du vent ($\nu_i(t)$), l'ensoleillement ($\Psi_i(t)$), ou la température ($\tau_i(t)$) auxquels il est soumis.

En notant $P_i(t) = P_i^P(t) - P_i^D(t)$ le surplus de production d'un agent i (sa production moins sa consommation à l'instant t), l'idée de base de notre approche est de décomposer $P_i^P(t)$ et $P_i^D(t)$ en fonction des équipements et préférences de i. Plus précisément, dans un cas où seuls des éoliennes et panneaux photovoltaïques sont disponibles, cela revient à écrire :

$$P_i^P(t) = \sum_{k \in WT_i} \mathcal{F}_{WT}(\nu_i(t)) + \sum_{k \in PV_i} \mathcal{F}_{PV}(\Psi_i(t)) \tag{C.1}$$

Où WT_i (resp. PV_i) représente les éoliennes (resp. panneaux photovoltaïques) de i. De plus, il nous faut un moyen de convertir nos entrées (vitesse du vent, ensoleillement) en production pour les DER considérés. Cela est réalisé grâce aux deux fonctions $\mathcal{F}_{WT}$ et $\mathcal{F}_{PV}$ qui sont des modèles simplifiés de ces générateurs.

En ce qui concerne la consommation, notre objectif était de parvenir à reproduire les effets saisonniers comme quotidiens grâce à un modèle « bottom-up ». Nous avons ainsi divisé la consommation en deux contributions : l'une ($\mathcal{F}_i heat$) s'interprétant comme un terme de chauffage, et l'autre ($\mathcal{F}_i elec$) comme une consommation d'équipements :

$$P_i^D(t) = \mathcal{F}_i heat(\tau(t), t) + \mathcal{F}_i elec(t) \tag{C.2}$$

Nous ne rentrerons pas plus en détails ici sur les fonctions $\mathcal{F}_i$ qui sont détaillées dans le chapitre 3 de la thèse. Les jeux de données principaux ont été obtenus à partir de [37] et [57] et couvrent la France de 2006 à 2012 avec une fréquence d'échantillonnage de 3h, et les Etats-Unis durant 2010 avec une fréquence d'une heure.

C.2.2 Les agrégateurs

Grâce à cette modélisation, nous sommes en mesure de paramétrer avec précision un large panel de prosumers, et de simuler l'évolution de leurs surplus de production. En étudiant ces quantités, nous avons remarqué une certaine diversité des motifs due à la multiplicité des profils sous-jacents. La question qui nous intéresse est donc de savoir si cette diversité peut servir à l'intégration des prosumers sur les marchés de l'électricité. En effet, de par leur faible production et l'instabilité de celle-ci, il semble assez improbable que des prosumers seuls puissent proposer des contrats sur les marchés. Il a ainsi été proposé qu'un autre type d'agent, appelé agrégateur, fasse l'intermédiaire entre les marchés et les prosumers. Un aggrégateur représente un ensemble de prosumers (appelé agrégation ou coalition dans cette thèse) et s'occupe de prendre les décisions stratégiques et économiques pour ces derniers. En d'autres termes, il forme un portfolio de générateurs, charges, et moyens de stockage qu'il est à même de contrôler et s'occupe de passer des contrats avec d'autres entités. Contrairement à

un prosumer seul, un agrégateur possède une capacité beaucoup plus importante qu'il peut en plus moduler avec une granularité beaucoup plus fine.

Toutefois, le point que nous soulevons dans cette thèse est que l'étape d'agrégation est primordiale pour la pérennité de la coalition. En effet, lorsqu'un contrat est passé entre deux entités, l'une des deux s'engage à injecter sur le réseau une certaine puissance durant une certaine période contre une rémunération proportionnelle à cette quantité et cette durée. Toutefois, si pour une raison quelconque, l'agrégateur ne parvenait pas à honorer son contrat, il subirait des pénalités financières proportionnelles à l'écart de production et à la durée de ces écarts. Tout le problème est que la production qu'offre un agrégateur est en réalité une somme de surplus de production de prosumers, réalisés grâce à des générateurs renouvelables.

L'agrégateur est donc face à un certain nombre de problèmes. Tout d'abord, il ne peut pas programmer sa production mais il doit l'estimer afin de proposer un contrat qu'il pense réalisable. Par ailleurs, s'il possède un certain contrôle sur les prosumers, il ne peut raisonnablement pas les forcer à suivre toutes ses directives. Un prosumer est en effet libre de consommer ce qu'il souhaite lorsqu'il le souhaite. En d'autres termes, un agrégateur se base sur des prévisions pour proposer un contrat, mais sa décision dépend aussi de son rapport au risque. En effet, plus le contrat est élevé, plus les gains peuvent être importants. Mais la probabilité de ne pas produire suffisamment, i.e le risque, augmente aussi. Cette probabilité est donc une quantité cruciale dans les décisions stratégiques et économiques d'un agrégateur, si bien que former sa coalition en fonction de celle-ci peut être une stratégie payante.

C.2.3 Agrégation et risque

C'est ce point ci que nous adressons dans le chapitre 4 de cette thèse. Etant donné un ensemble de N prosumers pour lesquels on possède un historique de leurs surplus de production, on cherche à former M coalitions optimisant le rapport entre l'espérance de leur production et le risque de sous produire par rapport aux contrats annoncés. Plus formellement, soit $P_S(t)$ la production d'une coalition S à un instant t. Par définition, on peut écrire $P_S(t) = \sum_{i \in S} P_i(t)$ comme la somme des surplus de productions ($P_i(t)$) réalisés par les prosumers appartenant à la coalition S. Comme expliqué précédemment, l'agrégateur gérant S doit proposer un contrat P_S^{CRCT} sur le marché pour une période Δt. Durant cette période, si $P_S(t) \geq P_S^{CRCT} \forall t$ alors l'agrégateur respecte constamment son contrat et la coalition sera rémunérée $\int_t^{t+\Delta t} \rho^+(t) P_S^{CRCT} dt$, où $\rho^+(t)$ est le prix unitaire de revente de l'électricité. A l'inverse, si $P_S(t) \leq P_S^{CRCT} \forall t$ alors l'agrégateur est toujours en défaut de production et la coalition sera pénalisée de $\int_t^{t+\Delta t} \rho^-(t) P_S^{CRCT} dt$, où $\rho^-(t)$ est la pénalité unitaire. Ainsi, lorsque l'agrégateur propose un contrat P_S^{CRCT}, son espérance de gain peut s'écrire comme :

$$E[Gains_S] = \int_t^{t+\Delta t} [1 - \lambda_S(t)] \rho^+(t) P_S^{CRCT} - \lambda_S(t) \rho^-(t) [P_S^{CRCT} \backsim P_S(t)] dt \tag{C.3}$$

Où $\lambda_S(t) = P[P_S(t) < P_S^{CRCT}]$. $\rho^+(t)$ et $\rho^-(t)$ sont des paramètres fixés par l'opérateur de marché sur lesquels l'agrégateur n'a aucun contrôle. Ainsi, le but de ce dernier est de proposer une valeur de contrat qui maximise $E[Gains_S]$. Bien qu'étant un problème intéressant en soit, le travail réalisé ici se place en amont de ce choix. En effet, l'agrégateur possède un autre degré de liberté, il peut choisir la composition de sa coalition parmi un ensemble de prosumers afin d'agir directement sur le terme $P_S(t)$ de l'équation précédente. La difficulté de cette opération est que l'agrégateur ne possède que des informations sur le passé des prosumers alors qu'il souhaite calculer ses gains futurs. On comprend toutefois qu'un agrégateur cherche à maximiser l'espérance de sa production future $E[P_S(t)]$ tout en minimisant sa variance $Var[P_S(t)]$.

Cela n'est pas sans rappeler certaines problématiques clé des marchés financiers. Plus précisément, la théorie des portfolios de Markowitz adresse presque exactement ce problème. Un trader cherche une combinaison linéaire des actions disponibles telles que le portfolio soit optimum. Markowitz montre qu'il existe un ensemble de portfolios, plus ou moins risqués, ayant cette propriété. Bien entendu, plus un portfolio optimum est risqué, plus les gains attendus sont importants. Cette théorie souffre néanmoins d'un défaut de taille. En effet, son fonctionnement repose sur l'hypothèse que les retours des actions suivent des lois normales (en réalité, on peut étendre les résultats de Markowitz à un ensemble un peu plus grand mais toujours très limité). Il a depuis été montré que les marchés financiers réels ne satisfaisaient généralement pas cette hypothèse.

Les agrégateurs ont également ce problème car les distributions des productions des prosumers ne sont généralement pas normales. On peut tout de même écrire :

$$\begin{cases} E[P_S(t)] = \sum_{i \in S} E[P_i(t)] \\ \sigma(P_S(t)) = \sqrt{\sum_{i \in S} \sigma_i^2 + \sum_i \sum_j \rho_{ij} \sigma_i \sigma_j} \end{cases} \tag{C.4}$$

Où $\sigma_i = \sqrt{Var[P_i(t)]}$ et $\rho_{ij} \in [-1, 1]$ est le coefficient de corrélation de Pearson entre P_i et P_j. Les termes croisés faisant intervenir les corrélations entre prosumers ont donc une grande importance dans la variance de la production de la coalition. Autrement dit, la structure de corrélation entre les prosumers d'une même coalition impacte directement la stabilité de sa production.

C.2.4 Former les coalitions maximisant les gains attendus

Etant donné que les agrégateurs ont accès aux historiques des $P_i(t)$, ils peuvent estimer les espérances, variances, et corrélations qui en découlent. Ils peuvent ainsi estimer, pour une coalition donnée, leur espérance de gains. Revenons toutefois à notre problème où M agrégateurs cherchent à former M coalitions à partir de N prosumers. Dans ces travaux, nous n'avons pas considéré de compétition entre les agrégateurs même si cela doit surement mener à des études très intéressantes. Notre approche fut de réaliser cette tâche tel que les gains soient maximisés globalement. Plus précisément, on cherche une répartition des prosumers en M groupes tels que la somme des gains soit maximum.

Bien que l'on soit en mesure d'estimer les gains pour n'importe quelle coalition, une recherche exhaustive est impensable dès lors que N dépasse quelques dizaines d'agents. Nous proposons donc une approche alternative se basant sur la théorie des graphes. Nous représentons la structure de corrélation entre les prosumers comme un graphe où les nœuds représentent ces derniers tandis que les arêtes ont un poids dépendant de la corrélation entre les deux extrémités. Cette approche fut proposée à la base par Mantegna en 1999 [52] pour étudier la structure des marchés financiers, et fut depuis améliorée par différents travaux de Onella [67] [25]. La principale contribution de ces travaux fut de montrer qu'un filtrage sur les arêtes du graphe permettait de mettre en évidence des clusters fortement corrélés. En d'autres termes, si on est en mesure de construire ce graphe et de le filtrer correctement, on peut retrouver la structure de corrélation par l'intermédiaire d'un algorithme de détection de communautés. Appliquer cette méthode stricto sensu à notre problème génère des coalitions fortement corrélées. En d'autres termes, on regroupe ainsi des agents produisant de façon similaire, ce qui, nous le verrons plus en avant, résulte en des coalitions extrêmement instables.

Nous proposons ainsi de former un graphe tel que l'absence de corrélation soit un facteur rapprochant les nœuds. Ainsi, plus deux prosumers tendent à produire de façon décorrélée, plus le poids du lien les connectant sera important. Si bien que le filtrage du graphe est censé mettre en évidence les regroupements que nous considérons comme stables. Bien qu'un bon plan de prime abord, la réalité est plus compliquée car le graphe obtenu ne possède pas de structure bien définie comme son dual. Nous montrons toutefois que les cliques de ce type de graphe ont une corrélation interne assez faible de telle sorte qu'elles apparaissent comme des structures stables selon nos critères. Nous proposons donc d'utiliser ces cliques comme des points de départ pour les coalitions. Nous utilisons ensuite un algorithme inspiré de la détection de communautés par expansion de clique afin d'augmenter la taille de ces dernières. Cette opération gloutonne est contrôlée par l'intermédiaire d'une fonction d'utilité que nous cherchons à maximiser. Des prosumers sont ainsi ajoutés et retirés des cliques si ces opérations augmentent l'utilité des cliques, et ce jusqu'à ce qu'un maximum soit atteint. Bien entendu, rien dans cet algorithme glouton ne garantie l'optimalité globale de ce maximum. Nous ajoutons ainsi une part d'aléatoire dans cet algorithme qui peut être vue comme une sorte de recuit simulé. Avec une certaine probabilité diminuant dans le temps, un prosumer n'augmentant pas l'utilité d'une coalition peut être tout de même ajouté. En effectuant un certain nombre de passes de cet algorithme, nous montrons que l'on est en mesure d'obtenir des résultats proches de l'optimum (lorsque celui-ci peut être trouvé).

La contribution principale de nos travaux [28] et [29] fut de montrer qu'en formant les coalitions de prosumers en fonction de la dé-corrélation, les agrégateurs étaient en mesure de mieux prédire la production de leur coalition, et ainsi de générer des profits plus importants que s'ils avaient formé leurs coalitions au hasard. Nous montrons également que former ses coalitions au hasard est loin d'être la pire des stratégies possibles. En effet, la diversité des profils que permet d'obtenir un choix aléatoire est assez bonne et permet à cette stratégie de générer des gains honorables. Former des coalitions de prosumers fortement corrélés conduit par contre à des

productions fortement instables et difficilement prédictibles. Nous montrons ainsi que les agrégateurs ayant opté pour cette stratégie se retrouvent criblés de dettes dues aux pénalités.

Ces travaux se focalisent sur les surplus de production des prosumers et leur revente sur les marchés. Nous avons ainsi été amenés à faire un certain nombre d'hypothèses et de simplifications. La plus visible peut-être est l'absence de topologie et de dynamique pour le système électrique. On ne prend en effet jamais en compte dans ces travaux les contraintes que cela induit sur les coalitions. Par exemple, lors de l'étape d'agrégation, il se peut que certaines coalitions ne puissent être formées à cause de la topologie du réseau, ou que certaines coalitions aient des contraintes plus complexes sur leur production dues aux capacités finies des lignes électriques. Une autre simplification concerne les graphes de corrélation que nous avons utilisés pour former les coalitions. Ces graphes sont des versions statiques d'une structure dynamique plus complexe. En effet, la corrélation entre deux séries temporelles peut être vue comme une mesure dynamique de la dépendance linéaire entre ces séries. Plus précisément, il est possible de calculer la corrélation sur une fenêtre temporelle glissante dont la taille est un paramètre. En fonction de la taille de cette fenêtre il est possible de mettre en évidence des événements à l'échelle de temps plus ou moins grande, tandis que la dynamicité de la mesure permet de capter d'éventuels motifs saisonniers.

L'étude des graphes dynamiques est plus complexe que celle des graphes statiques, si bien que pendant longtemps la parade consistait à agréger les relations pour les ramener à une version statique. De récents progrès ont permis toutefois à l'étude de ces graphes de prendre un nouvel essor. Dans le chapitre 4 de cette thèse nous présentons des travaux encore en cours dont l'objectif est l'étude des graphes de corrélation dynamiques. Pour ce faire, nous nous basons sur une approche développée par [27] et qui utilise la factorisation de tenseurs pour mettre en évidence la structure spatio-temporelle de ces graphes. Un graphe dynamique peut être vu comme une suite temporelle de matrices d'adjacence qui encodent les snapshots successifs. En organisant ces matrices selon un troisième axe, on obtient un tenseur en trois dimensions contenant toute l'information du graphe dynamique. Toutefois, il est compliqué pour un esprit humain normalement constitué d'extraire des informations structurelles condensées à partir de cette représentation tridimensionnelle. La factorisation non négative de tenseurs est alors une des techniques permettant de condenser l'information. Nous montrons qu'en appliquant cette méthode aux graphes de corrélation, on est en mesure d'observer des motifs récurrents ainsi que des pics de corrélation courts et intenses qui sont généralement la traduction d'événements particuliers (crises financières par exemple). Nous proposons ensuite une application directe de cette méthode pour étudier la consommation électrique de foyers irlandais (données obtenues par l'intermédiaire de [38]). Ces travaux sont encore en cours et seront probablement publiés après la soutenance de thèse, si bien que nous ne nous attarderons pas plus sur ce sujet.

C.3 Placement de stockages dans un réseau de prosumers

C.3.1 Problématique générale

La seconde partie de cette thèse est toujours consacrée aux prosumers, mais propose une approche légèrement différente. Nous proposons ainsi de prendre en compte explicitement la topologie et la dynamique du réseau électrique. Notre objectif n'est alors plus la revente de surplus de production, mais l'installation de dispositifs de stockage permettant de maintenir la stabilité du système. Plus précisément, nous considérons un réseau de prosumers interconnectés par des lignes électriques. Certains de ces prosumers se comportent comme des générateurs car ils sont en excédant de production, tandis que d'autres peuvent être vus comme des charges car ils consomment plus qu'ils ne produisent. Comme nous l'avons vu dans la partie I, cet état n'est pas gravé dans la pierre, il dépend principalement des conditions météorologiques et des comportements des prosumers à un instant donné. Si bien que cette organisation entre générateurs et charges est fortement susceptible d'évoluer dans le temps. Nous sommes donc face à un réseau où générateurs et charges ne sont pas fixes, et dont les contributions sont stochastiques.

Il est assez connu que le réseau électrique doit être synchronisé sur une fréquence commune Ω pour fonctionner (en Europe $\Omega = 50Hz$). Plus précisément, la fréquence à laquelle le réseau est synchronisé doit être située dans un intervalle de valeurs acceptables $[\Omega - \epsilon, \Omega + \epsilon]$. Une condition nécessaire pour que cette synchronisation puisse avoir lieu est que la production et la consommation soient équilibrées. Si l'une est plus importante que l'autre, la fréquence du système dévie de Ω. Si rien n'est fait pour rééquilibrer la production et la consommation, la fréquence peut éventuellement sortir de l'intervalle de stabilité et provoquer des dommages aux divers équipements.

De ce point de vue, la stabilité d'un réseau de prosumers semble difficile à garantir puisque générateurs et charges changent constamment d'emplacements et de puissances. Nous proposons de maintenir la stabilité du système grâce à des dispositifs de stockage disséminés dans le réseau. L'idée, assez simple, consiste à stocker l'électricité lorsque la production excède la consommation, et à décharger celle-ci lors du cas inverse. Comme nous l'avons expliqué précédemment, cette méthode seule ne pourra probablement pas permettre de piloter le système, tout du moins avec les connaissances sur le stockage que nous avons actuellement. Néanmoins, activer des batteries est beaucoup plus rapide que d'allumer ou éteindre des générateurs de secours. En d'autres termes, le stockage peut être vu comme un moyen d'action rapide préalable à une réorganisation plus profonde du portfolio de production. Ce genre d'action est communément appelée régulation de fréquence.

La question que nous adressons ici est de déterminer à la fois le nombre et l'emplacement de ces dispositifs tels que le système puisse être contrôlé. Comme ces dispositifs sont couteux, nous cherchons à minimiser leur nombre (on fait l'hypothèse que les coûts d'installation sont les mêmes sur tout les nœuds du graphe). Nous montrerons dans ce qui suit que ces ensembles d'emplacements sont loin d'être tous équivalents. Nous développons

un critère énergétique afin de comparer les performances de plusieurs ensembles, et nous montrons comment trouver celui qui fournit les meilleures performances en moyenne.

C.3.2 Modéliser la dynamique du réseau électrique

Pour cette étude nous avons besoin de mettre la dynamique du réseau électrique en équations. Toutefois, comme nous le verrons dans les parties suivantes, ce modèle doit être suffisamment simple pour pouvoir l'étudier par l'intermédiaire de la théorie du contrôle. Nous avons ainsi opté pour l'approche utilisée par [23] qui modélise cette dynamique par un réseau d'oscillateurs couplés. Ces entités possèdent tous une fréquence propre à laquelle ils oscillent s'ils sont isolés. Par l'intermédiaire du réseau, un oscillateur est couplé à un ensemble d'autres oscillateurs voisins. Ainsi, la fréquence d'un oscillateur est impactée par celles de ses voisins et vice versa. On peut montrer que sous certaines conditions sur la topologie et la distribution des fréquences propres, le réseau se synchronise à une fréquence commune.

Le lien entre le réseau électrique et les oscillateurs couplés n'est pas forcément évident de prime abord. Ce processus ainsi que les hypothèses simplificatrices nécessaires sont détaillés dans [23] et dans le chapitre 6 de la thèse. L'idée principale est que la dynamique des angles de phase de ces oscillateurs est une approximation de la dynamique réelle du système. Même dans cette forme simplifiée, cette dynamique n'est toujours pas linéaire car les équations sont couplées au travers d'un terme sinusoïdal. Nous faisons alors une hypothèse classique consistant à linéariser ces termes au premier ordre. On obtient alors une dynamique pouvant s'écrire sous la forme $\dot{X} = AX$, où X est l'état du système à un instant donné, et A est la matrice gouvernant la dynamique du système. Cette dernière encode les paramètres du système comme la topologie du réseau, les coefficients de d'amortissement ou d'inertie des éléments par exemple.

C.3.3 Contrôle et réseaux

Sans aucune intervention extérieure, une dynamique du type $\dot{X} = AX$ a pour résultat une trajectoire dans l'espace d'état donnée par $X(t) = e^{At}$. Cette trajectoire peut être ce que l'on désire, mais on peut également avoir le besoin d'influencer la dynamique de manière à suivre une trajectoire différente. Dans notre cas d'étude, supposons qu'un prosumer se mette soudainement à consommer plus qu'avant. Si tous les autres conservent la même attitude, le système est perturbé car la production n'est plus à la hauteur de la consommation. Ceci modifie l'état du système et place ce dernier sur une trajectoire qui peut potentiellement l'amener dans une région instable. Notre objectif est donc de ramener le système sur une trajectoire correcte en injectant au niveau de certains nœuds des signaux judicieusement choisis. Plus formellement, on peut réécrire la dynamique du système comme étant $\dot{X} = AX + Bu$, où la matrice B encode les nœuds au niveau desquels sont injectés les signaux u(t). Toute la question est alors de déterminer B et $u(t)$ en fonction de A, X_0, et X_T, où X_0 est l'état initial et X_T l'état final désiré.

La théorie du contrôle permet de répondre à ce genre de question. Plus précisément, le système (A, B) est dit contrôlable s'il est possible de le piloter de n'importe quel état initial à n'importe quel état final au moyen d'une suite finie de signaux $u(t)$ injectés au niveau des nœuds encodés par B. Déterminer si un système est contrôlable est un objectif classique de la théorie du contrôle. Pour ce faire il existe un certain nombre de méthodes, dont l'une des plus anciennes et connues est certainement le critère de Kalman, qui consiste à vérifier que la matrice $C = [B, AB, A^2B, \ldots, A^{N-1}B]$ possède un rang égal à N. Cela fonctionne très bien sur de petits systèmes mais devient rapidement problématique pour des systèmes plus complexes.

C.3.4 Contrôle optimal

Pour certaines applications, les signaux $u(t)$ n'ont pas de signification physique réelle si bien que de grandes variations d'amplitude n'ont pas vraiment d'importance. Pour d'autres applications toutefois, ces signaux représentent une action qui peut être couteuse. Dans notre cas de système électrique, on peut interpréter ces signaux comme des taux de charge et de décharge des équipements de stockage. Bien entendu, charger et décharger des batteries possède un coût même dans le cas où l'énergie provient de générateurs renouvelables, si bien que l'on souhaiterait contrôler le système en effectuant un minimum de ces opérations. Cela est d'autant plus important qu'il peut exister plusieurs séquences de contrôle permettant d'amener un système donné d'un même état initial à un même état final. Trouver la « meilleure » séquence est au cœur de la théorie du contrôle optimal. Plus précisément, pour un système (A, B), un état initial X_0, et un état final X_T, le but est de trouver une séquence $u^\star(t)$ qui minimise une certaine fonction de coût $J(u, X)$. Cette fonction de coût peut dépendre des signaux $u(t)$ mais aussi des états dans lesquels le système va se trouver durant la phase de contrôle.

Pour le cas qui nous intéresse, nous choisissons une forme particulière de fonction de coût : $J(u) = \int_{t_0}^{t_f} \|u(t)\|^2 dt = \mathcal{E}$ appelée énergie de contrôle. Celle-ci quantifie indirectement la quantité d'énergie physique nécessaire pour le contrôle, et pénalise les fortes charges et décharges successives des batteries. Pour ce cas particulier de fonction de coût, il est possible de déterminer $u^\star(t)$:

$$u^\star(t) = B^T e A^T (T - t) W^{-1}(T) \nu_f \tag{C.5}$$

Où $\nu_f = X_T \breve{} X_0 e^{AT}$ est la différence entre l'état final désiré et celui atteint sans contrôle, et $W(t) = \int_0^t e^{A\tau} B B^T e^{A^T \tau} d\tau$ est la matrice gramienne du système. L'énergie de contrôle pour amener le système (A, B) de X_0 à X_T en un temps T grâce aux signaux $u^\star(t)$ est alors minimum et s'écrit :

$$\mathcal{E}_{min} = \nu_f^T W^{-1}(T) \nu_f \tag{C.6}$$

Assez logiquement, l'énergie de contrôle minimum dépend des matrices A et B, du temps T, ainsi que des états initial et final. Cela signifie que pour un même triplet (A, B, T), mais pour des couples (X_0, X_T) différents, l'énergie de contrôle peut varier fortement. Etant donné que les prosumers qui composent le réseau sont difficilement prévisibles, la nature des perturbations peut être très diverse. Autrement dit, l'état initial X_0 dans lequel se

trouve le système peut varier fortement en fonction du type de perturbation à laquelle nous devons répondre. Nous avons ainsi besoin de trouver une matrice B de nœuds contrôlés telle que l'énergie de contrôle soit minimum en moyenne. Pour ce faire, remarquons que la matrice W intervient dans l'expression de l'énergie de contrôle, mais ne s'exprime qu'en fonction des matrices A et B. En d'autres termes, W ne dépend pas des états initial et final, mais joue tout de même un rôle dans le calcul de l'énergie de contrôle. Ainsi, W peut être utilisé pour obtenir des informations sur l'énergie moyenne de contrôle.

C.3.5 Minimiser l'énergie moyenne de contrôle

Comme $\mathcal{E}_{min} \sim W^{-1}(T)$, nous cherchons à sélectionner un ensemble de contrôleurs tel que W soit « grand ». Afin de formaliser cette notion, nous utilisons la trace de l'inverse de W ($Tr[W^{-1}]$) comme proxy. En effet, il a été montré que cette quantité était directement liée à l'énergie moyenne de contrôle, si bien que nous souhaitons minimiser cette quantité en sélectionnant l'ensemble optimum $S^\star$ de contrôleurs. En notant B_S la matrice B obtenue en sélectionnant le sous-ensemble S de contrôleurs, W_S la matrice gramienne du système (A, B_S), et $F : S \longrightarrow -Tr[W_S^{-1}]$, le problème devient :

$$MAX_{S \in \mathcal{N}} F(S) \tag{C.7}$$

Pour une fonction F quelconque ce problème d'optimisation est difficile à résoudre, et il est souvent nécessaire d'utiliser des méthodes heuristiques sans garantie de performances. Cependant, des travaux récents sur le contrôle optimal ont montré que certaines métriques liées à la matrice gramienne possèdent une structure spéciale. Il a ainsi été démontré que la fonction F évoquée précédemment possède une structure sous modulaire, ce qui signifie que :

$$F(X \cup \{x\}) \breve{} F(X) \geq F(Y \cup \{x\}) \breve{} F(Y), \forall X, Y,\ tq\ X \subset Y\ et\ \forall x \in \mathcal{N} \setminus Y \tag{C.8}$$

Ce résultat nous permet d'utiliser un algorithme glouton relativement simple qui ajoute des éléments un à un en fonction des gains marginaux engendrés. Pour une fonction quelconque il n'y a aucune garantie qu'un algorithme aussi simple donne de bons résultats. Si la fonction d'ensemble est sous-modulaire, il a été montré que le résultat obtenu était égal, dans le pire des cas, à 63% de l'optimum. En pratique, cet algorithme fonctionne très bien et retourne des résultats souvent beaucoup plus proches de l'optimum que de la borne inférieure.

C.3.6 Intégrer les contraintes physiques du système

Notre méthode commence ainsi à émerger. A partir de la dynamique simplifiée (matrice A), nous pouvons utiliser l'algorithme glouton évoqué ci-dessus pour trouver le plus petit ensemble de nœuds du graphe tels que nous ayons le contrôle du système et que notre énergie de contrôle moyenne soit minimum. Cette approche, bien que valide en théorie, ne prend pas

en compte les diverses contraintes physiques qui pèsent sur le réseau électrique et sur les dispositifs de stockage. En effet, les lignes électriques possèdent des capacités maximum, tout comme les batteries dont les capacités de charge et décharge sont limitées. L'état du système et la trajectoire de celui-ci lors de la phase de contrôle sont donc contraints. Il est ainsi nécessaire de vérifier que ces contraintes ne seront pas violées lors de la sélection des contrôleurs.

Une des difficultés est que ces contraintes ne peuvent être vérifiées que pour une trajectoire donnée, alors même que nous cherchons un placement efficace en moyenne. L'état initial de la phase de contrôle, et donc l'énergie de contrôle, dépend de la perturbation à laquelle le système a été soumis. Afin de déterminer si un placement S respecte les contraintes physiques du système, nous tirons aléatoirement des états initiaux dans une hyper-sphère centrée sur l'état d'équilibre, et nous vérifions les contraintes pour chacune des trajectoires qui en découlent. Si toutes ces trajectoires respectent les contraintes, alors nous considérons que le placement S permet le contrôle du système à la fois du point de vue théorique et physique.

C.3.7 Résultats

Afin de valider notre algorithme, nous comparons ses performances par rapport à une approche plus aléatoire. Pour un grand nombre de topologies aléatoires, nous tirons des ensembles de contrôleurs aléatoirement et nous comparons l'énergie moyenne de contrôle résultante avec celle obtenue via notre méthode. Cela nous permet de montrer que notre algorithme performe, en moyenne, toujours mieux que l'approche aléatoire, mais que la différence de performance est variable en fonction de la taille de l'ensemble des contrôleurs. Plus précisément, si les contrôleurs sont très peu nombreux par rapport au nombre de nœuds dans le graphe, alors les deux méthodes donnent des énergies de contrôle extrêmement élevées avec une différence assez faible entre les deux. A l'opposé, si le nombre de contrôleurs est quasiment égal au nombre de nœuds dans le graphe, alors tout est contrôlé et notre optimisation n'a pas de réelle utilité. Les deux énergies moyennes sont donc similaires. Pour des cas médians, nous montrons que notre méthode performe beaucoup mieux que l'approche aléatoire.

C.4 Conclusion

Cette thèse a été consacrée à l'étude des réseaux électriques intelligents avec une approche basée sur les systèmes complexes. Bien que les réseaux électriques actuels possèdent déjà un important panel d'outils de contrôle qui permet un fonctionnement fiable et stable, l'émergence de systèmes plus complexes, interconnectés, et distribués, nécessite des actions intelligentes et automatisées. Des données mesurées en temps réel par des capteurs disséminés sur les réseaux pourraient être combinées avec des algorithmes d'apprentissage automatique qui permettraient de prédire les événements et renforcer le contrôle du système.

Dans ce genre d'environnements interdépendants et multi-échelle, il pourrait être difficile de comprendre comment l'ensemble du système se

comporte. La théorie des systèmes complexes aborde ce genre de problématiques où certains phénomènes collectifs émergent à partir de comportements individuels. Alors qu'une grande partie des travaux liés au smart grid se concentrent sur des composants très spécifiques des réseaux intelligents qui peuvent être testés en isolement, ce travail utilise des hypothèses et des modèles relativement simples pour les individus et étudie leurs répercussions à plus grande échelle.

Cette thèse est divisée en deux sections principales. La première se concentre sur la modélisation de marchés de l'énergie avec des agrégations de prosumers peu fiables en tant que participants, tandis que la seconde partie aborde l'optimisation du placement de batteries afin d'assurer une régulation de fréquence dans un réseau de prosumers.

La principale contribution pour la première partie est le développement d'un modèle de prosumer basé sur des données météorologiques réelles qui nous a permis de capturer les corrélations spatio-temporelles entre ces agents. Un modèle pour le marché de l'énergie a également été proposé et utilisé pour étudier comment les agrégateurs de marché pouvaient former des coalitions de prosumers de telle sorte qu'ils maximisent leurs gains avec de faibles risques. Nous avons montré que la diversification était un concept clé dans ce processus, de telle sorte que nous avons étudié une stratégie de regroupement sur la base de la corrélation. Nous avons proposé une optimisation gourmande qui commence par des cliques d'agents peu corrélés, et augmente leur taille aussi longtemps que leurs utilités sont améliorées. Nous avons ensuite comparé les résultats par rapport à d'autres méthodes et montré que l'algorithme proposé était plus performant que les autres. Nous pensons que les méthodes basées sur ce travail pourraient être utilisés par les agrégateurs sur les marchés de l'énergie afin de maximiser leurs gains attendus pour un risque donné. La méthode proposée pourrait probablement être améliorée en utilisant des données réelles de prosumers lorsque celles-ci seront disponibles. En outre, nous proposons dans le chapitre 4 une direction possible pour mieux comprendre la structure de corrélation entre les prosumers et possiblement améliorer la méthode. Nous avons en effet proposé d'utiliser une décomposition de tenseurs pour étudier les graphes de corrélation dynamiques au lieu de se contenter de leurs versions statiques. Être capable de comprendre la dynamique de ces graphes pourrait certainement fournir des informations précieuses aux agrégateurs de marché.

Nous pensons qu'une des contributions principales de la seconde partie fut de combiner la dynamique du réseau électrique avec la théorie du contrôle optimal afin d'étudier le positionnement de batteries dans un contexte stochastique. Par ailleurs, nous avons intégré les contraintes physiques du réseau électrique et des batteries dans ce processus. Nous avons également utilisé une méthode d'optimisation basée sur les fonctions sous-modulaires qui est facile à mettre en œuvre et qui possède une garantie de performance. L'algorithme proposé est donc en mesure de produire une stratégie de placement efficace pour une topologie de réseau électrique donnée. Il semble inévitable que de grandes quantités de batteries devront être installées dans les futurs systèmes électriques. Optimiser leur emplacement à l'avance pourrait donc conduire à des économies importantes. Cette question a été abordée dans la littérature, mais jamais, à notre connaissance, par la théorie du contrôle optimal et la sous-modularité. Nous pensons que

ce travail apporte une perspective différente sur cette question, et nous espérons que d'autres travaux seront entrepris dans cette direction.

En termes de publications, le travail réalisé pour la première partie a donné naissance à un article présenté lors de la conférence IEEE ICC 2015, ainsi qu'à un article de journal publié dans IEEE Transactions on Smart Grid 2016, tous deux en tant que premier auteur (voir annexe ref AnnexeA pour obtenir une liste complète des publications). Pour la deuxième partie, un article sera présenté à la conférence IEEE SmartGridComm 2016 en novembre prochain. Un papier journal est également en cours de rédaction et sera probablement soumis au cours de l'été 2016. Les travaux présentés dans le chapitre 5 sur les graphes de corrélation dynamiques sont encore incomplets, car ils ont été entrepris peu avant la rédaction de la présente thèse. Ce travail est en cours d'achèvement et nous espérons qu'il sera soumis lors de l'automne 2016.

Mon travail actuel comprend également le mentorat de M. Lester Padilla, stagiaire de Master, qui travaille sur l'extraction de topologies de réseaux électriques à partir d'OpenStreetMap (OSM). Ce stage fut décidé après un séminaire passionnant organisé dans nos bureaux en Mars 2016 et devrait se terminer en Septembre 2016. L'idée est basée sur le projet SciGRID développé à *NEXT ENERGY, EWE-Forschungszentrum für Energietechnologie* en Allemagne. Le stage vise à extraire, à partir des données OSM, une topologie de réseau électrique contenant les générateurs et les charges. Avec des informations supplémentaires comme le type des générateurs, leurs capacités, les capacités des lignes électriques, et ainsi de suite, nous pensons que ce travail aidera les chercheurs à obtenir des données réalistes pour tester leurs recherches.

Dans une perspective plus lointaine, à cause de mon travail actuel sur les graphes temporels, j'ai développé une attirance pour l'étude de la détectabilité et l'évolution des communautés dans les graphes temporels. Cet intérêt est encore à un stade très précoce et j'espère avoir du temps à y consacrer dans un avenir proche. Un autre travail futur pourrait également être lié au projet Grid4Earth sur les microgrids.

Outre le travail effectué dans le cadre de ma thèse, j'ai également donné des cours (monitorat) à l'université Pierre et Marie Curie (UPMC) au cours de ces trois années. Au début, ce fut l'occasion pour moi de faire l'expérience de l'enseignement, car je ne savais pas, dès le départ, que cela me conviendrait. En fin de compte, le monitorat m'a confirmé que l'enseignement est quelque chose que j'apprécie particulièrement. Rétrospectivement, je me sens très chanceux d'avoir eu l'occasion d'apprendre et de découvrir beaucoup de choses fascinantes. Il est également frustrant d'avoir tant de choses à faire ou étudier et si peu de temps pour les explorer en profondeur. Bien que c'était dans un coin de ma tête depuis le début de ma thèse, je sais maintenant que je souhaite travailler dans le milieu universitaire en tant que professeur à l'avenir.

Bibliography

[1] Yong-Yeol Ahn et al. "Analysis of Topological Characteristics of Huge Online Social Networking Services". In: *Proceedings of the 16th International Conference on World Wide Web* (2007), pp. 835–844. ISSN: 08963207. DOI: 10.1145/1242572.1242685.

[2] M. Ajmone-Marsan et al. "The emerging energy web". In: *The European Physical Journal Special Topics* 214.1 (2012), pp. 547–569. ISSN: 1951-6355. DOI: 10.1140/epjst/e2012-01705-1. URL: http://www.springerlink.com/index/10.1140/epjst/e2012-01705-1.

[3] Anh Huy Phan Andrzej Cichocki Rafal Zdunek and Shun-Ichi Amari. *Nonnegative Matrix and Tensor Factorizations*. ISBN: 9780470746660.

[4] M. Anvari et al. *Suppressing the non-Gaussian statistics of Renewable Power from Wind and Solar*. 2015. eprint: arXiv:1505.01638.

[5] Haris Aziz and Bart de Keijzer. "Complexity of Coalition Structure Generation". In: *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2011)*. Ed. by Kagan Tumer et al. International Foundation for Autonomous Agents and Multi-Agent Systems (IFAAMAS), 2011, pp. 383–390.

[6] Albert-László Barabási and Réka Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (1999), pp. 509–512. ISSN: 0036-8075. DOI: 10.1126/science.286.5439.509. eprint: http://science.sciencemag.org/content/286/5439/509.full.pdf. URL: http://science.sciencemag.org/content/286/5439/509.

[7] Ian Barnett and Jukka-Pekka Onnela. "Change Point Detection in Correlation Networks". In: *Nature Publishing Group* 2 (2014), p. 23. ISSN: 20452322. DOI: 10.1038/srep18893. arXiv: 1410.0761. URL: http://arxiv.org/abs/1410.0761.

[8] Andrey Bogomolov et al. "Energy consumption prediction using people dynamics derived from cellular network data". In: *EPJ Data Sci.* (2016). DOI: 10.1140/epjds/s13688-016-0075-3. URL: http://dx.doi.org/10.1140/epjds/s13688-016-0075-3.

[9] Rasmus Bro and Henk A. L. Kiers. "A new efficient method for determining the number of components in PARAFAC models". In: *Journal of Chemometrics* 17.5 (2003), pp. 274–286. ISSN: 1099-128X. DOI: 10.1002/cem.801. URL: http://dx.doi.org/10.1002/cem.801.

[10] Shengrong Bu, F. Richard Yu, and Peter X. Liu. "Dynamic pricing for demand-side management in the smart grid". In: *2011 IEEE Online Conference on Green Communications* (2011), pp. 47–51. DOI: 10.1109/GreenCom.2011.6082506. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6082506.

[11] M. Kanat Camlibel. "Popov-Belevitch-Hautus type controllability tests for linear complementarity systems". In: *Systems and Control Letters* 56.5 (2007), pp. 381–387. ISSN: 01676911. DOI: 10.1016/j.sysconle.2006.10.023.

[12] Remy Chicheportiche and Jean-Philippe Bouchaud. "The Joint Distribution of Stock Returns is Not Elliptical". In: *International Journal of Theoretical and Applied Finance* 15.03 (2012), p. 1250019. DOI: 10.1142/S0219024912500197. eprint: http://www.worldscientific.com/doi/pdf/10.1142/S0219024912500197.

[13] F.R.K. Chung. *Spectral Graph Theory*. CBMS Regional Conference Series 92. Conference Board of the Mathematical Sciences. ISBN: 9780821889367.

[14] Reuven Cohen. "Complex Networks". In: (2007).

[15] Luciano Fontoura Costa et al. "Characterization of complex networks: A survey of measurements". In: *Advances in Physics* 56.1 (2007), pp. 167–242. ISSN: 0001-8732. arXiv: 0505185 [cond-mat].

[16] G Dans and Forest Science. "Rayonnement solaire". In: 360 (2007).

[17] Panagiotis D. Diamantoulakis, Vasileios M. Kapinas, and George K. Karagiannidis. "Big Data Analytics for Dynamic Energy Management in Smart Grids". In: *Big Data Research* 2.3 (2015). Big Data, Analytics, and High-Performance Computing, pp. 94 –101. ISSN: 2214-5796. DOI: http://dx.doi.org/10.1016/j.bdr.2015.03.003. URL: http://www.sciencedirect.com/science/article/pii/S2214579615000283.

[18] Florian Dörfler, Michael Chertkov, and Francesco Bullo. "Synchronization in complex oscillator networks and smart grids." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.6 (2013), pp. 2005–10. ISSN: 1091-6490. DOI: 10.1073/pnas.1212134110. URL: http://www.ncbi.nlm.nih.gov/pubmed/23319658.

[19] S N Dorogovtsev and J F F Mendes. "Evolution of Networks: From Biological Nets to the Internet and WWW". In: *Booksgooglecom* 57 (2003), p. 280. ISSN: 00319228. DOI: 10.1063/1.1825279. arXiv: 0106144 [cond-mat].

[20] P ERDdS and A R&WI. "On random graphs I." In: *Publ. Math. Debrecen* 6 (1959), pp. 290–297.

[21] M. Erol-Kantarci, B. Kantarci, and H. T. Mouftah. "Reliable overlay topology design for the smart microgrid network". In: *IEEE Network* 25.5 (2011), pp. 38–43. ISSN: 0890-8044. DOI: 10.1109/MNET.2011.6033034.

[22] Abdallah Farraj, Eman Hammad, and Deepa Kundur. "On Using Distributed Energy Resources to Reshape the Dynamics of Power Systems During Transients". In: *2015 IEEE International Conference on Smart Grid Communications, SmartGridComm 2015* (2015), pp. 756–761.

[23] G. Filatrella, a. H. Nielsen, and N. F. Pedersen. "Analysis of a power grid using a Kuramoto-like model". In: *The European Physical Journal B* 61.4 (2008), pp. 485–491. ISSN: 1434-6028. DOI: 10.1140/epjb/e2008-00098-8. URL: http://www.springerlink.com/index/10.1140/epjb/e2008-00098-8.

[24] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control*. Stochastic Modelling and Applied Probability. Springer New York, 2012. ISBN: 9781461263807. URL: https://books.google.fr/books?id=qJDbBwAAQBAJ.

[25] A. Garas et al. "The structural role of weak and strong links in a financial market network". English. In: *The European Physical Journal B* 63.2 (2008), pp. 265–271. ISSN: 1434-6028. DOI: 10.1140/epjb/e2008-00237-3.

[26] Nicolas Gabriel Gast, Jean-Yves Le Boudec, and Dan-Cristian Tomozei. "Impact of Demand-Response on the Efficiency and Prices in Real-Time Electricity Markets". 2014.

[27] Laetitia Gauvin, André Panisson, and Ciro Cattuto. "Detecting the Community Structure and Activity Patterns of Temporal Networks: A Non-Negative Tensor Factorization Approach". In: *PLoS ONE* 9.1 (Jan. 2014), pp. 1–13. DOI: 10.1371/journal.pone.0086028. URL: http://dx.doi.org/10.1371%2Fjournal.pone.0086028.

[28] Nicolas Gensollen et al. "Coalition formation algorithm of prosumers in a smart grid environment". In: *2015 IEEE International Conference on Communications (ICC)*. Institute of Electrical & Electronics Engineers (IEEE), 2015. DOI: 10.1109/icc.2015.7249262. URL: http://dx.doi.org/10.1109/ICC.2015.7249262.

[29] Nicolas Gensollen et al. "Stability and Performance of Coalitions of Prosumers Through Diversification in the Smart Grid". In: *IEEE Trans. Smart Grid* (2016), pp. 1–1. DOI: 10.1109/tsg.2016.2572302. URL: http://dx.doi.org/10.1109/TSG.2016.2572302.

[30] Lazaros Gkatzikis et al. "Collaborative placement and sharing of storage resources in the Smart Grid". In: *2014 IEEE International Conference on Smart Grid Communications, SmartGridComm 2014* (2015), pp. 103–108. DOI: 10.1109/SmartGridComm.2014.7007630.

[31] S. Grijalva and M. U. Tariq. "Prosumer-based smart grid architecture enables a flat, sustainable electricity industry". In: *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*. 2011, pp. 1–6. DOI: 10.1109/ISGT.2011.5759167.

[32] P. Hines et al. "The topological and electrical structure of power grids". In: *Proceedings of the Annual Hawaii International Conference on System Sciences* (2010). ISSN: 15301605. DOI: 10.1109/HICSS.2010.398.

[33] Paul Hines, Eduardo Cotilla-Sanchez, and Seth Blumsack. "Do topological models provide good information about electricity infrastructure vulnerability?" In: *Chaos (Woodbury, N.Y.)* 20.3 (2010), p. 033122. ISSN: 1089-7682. DOI: 10.1063/1.3489887.

[34] Murad Hossain et al. "Australian Airport Network Robustness Analysis : A Complex Network Approach". In: *Australasian Transport Research Forum 2013 Proceedings* October (2013), pp. 1–21.

[35] Matthias Huber, Desislava Dimkova, and Thomas Hamacher. "Integration of wind and solar power in Europe : Assessment of fl exibility requirements". In: *Energy* 69 (2014), pp. 236–246. ISSN: 0360-5442. DOI: 10.1016/j.energy.2014.02.109. URL: http://dx.doi.org/10.1016/j.energy.2014.02.109.

[36] Frede Hvelplund. "Renewable energy and the need for local energy markets". In: *Energy* 31.13 (2006). Double Special Issue: 2nd Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems/PRES 03 and {PRES} 2004 Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction2nd Dubrovnik Conference on Sustainable Energy, Water and Environment Systems/PRES 03 and {PRES} 2004 Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction, pp. 2293 –2302. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2006.01.016. URL: http://www.sciencedirect.com/science/article/pii/S0360544206000417.

[37] "Infoclimat". In: (). URL: http://www.infoclimat.fr.

[38] "Irish Social Science Data Archive". In: (). URL: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/.

[39] B. Jansen et al. "Architecture and Communication of an Electric Vehicle Virtual Power Plant". In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. 2010, pp. 149–154. DOI: 10.1109/SMARTGRID.2010.5622033.

[40] Tue V. Jensen et al. *The RE-Europe data set*. Dec. 2015. DOI: 10.5281/zenodo.35177. URL: http://dx.doi.org/10.5281/zenodo.35177.

[41] F D Kanellos et al. "Micro-Grid Simulation during Grid-Connected and Islanded Modes of Operation". In: ().

[42] D.E. Kirk. *Optimal Control Theory: An Introduction*. Dover Books on Electrical Engineering. Dover Publications, 2012. ISBN: 9780486135076. URL: https://books.google.fr/books?id=onuH0PnZwV4C.

[43] Andreas Krause and Daniel Golovin. "Submodular function maximization". In: *Tractability: Practical Approaches to Hard Problems* 3 (2014), pp. 71–104. DOI: 10.1017/CBO9781139177801.004.

[44] L. Kullmann, J. Kertész, and K. Kaski. "Time-dependent cross-correlations between different stock returns: A directed network of influence". In: *Phys. Rev. E* 66 (2 2002), p. 026125. DOI: 10.1103/PhysRevE.66.026125. URL: http://link.aps.org/doi/10.1103/PhysRevE.66.026125.

[45] Yoshiki Kuramoto and H. Araki. In: *Lecture Notes in Physics, International Symposium on Mathematical Problems in Theoretical Physics* (1975), p. 420.

[46] Ching-Tai Lin. "Structural controllability". In: *IEEE Transactions on Automatic Control* 19.3 (1974), pp. 201–208. ISSN: 0018-9286. DOI: 10.1109/TAC.1974.1100557.

[47] T. Linnenberg et al. "A market-based multi-agent-system for decentralized power and grid control". In: *Emerging Technologies Factory Automation (ETFA), 2011 IEEE 16th Conference on*. 2011, pp. 1–8. DOI: 10.1109/ETFA.2011.6059126.

[48] Yang-Yu Liu and Albert-Laszló Barabási. "Control Principles of Complex Networks". In: (2015), pp. 1–55. arXiv: 1508.05384. URL: http://arxiv.org/abs/1508.05384.

[49] J. Lloret et al. "A Group-Based Protocol for Improving Energy Distribution in Smart Grids". In: *2011 IEEE International Conference on Communications (ICC)*. 2011, pp. 1–6. DOI: 10.1109/icc.2011.5962877.

[50] Thillainathan Logenthiran, Dipti Srinivasan, and Tan Zong Shun. "Demand side management in smart grid using heuristic optimization". In: *IEEE Transactions on Smart Grid* 3.3 (2012), pp. 1244–1252. ISSN: 19493053. DOI: 10.1109/TSG.2012.2195686.

[51] M. Lydia et al. "A comprehensive review on wind turbine power curve modeling techniques". In: *Renewable and Sustainable Energy Reviews* 30 (2014), pp. 452–460. ISSN: 13640321. DOI: 10.1016/j.rser.2013.10.030.

[52] R.N. Mantegna. "Hierarchical structure in financial markets". English. In: *The European Physical Journal B* 11.1 (1999), pp. 193–197. ISSN: 1434-6028. DOI: 10.1007/s100510050929.

[53] Harry M Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, 1959.

[54] Russell Merris. "Laplacian matrices of graphs: a survey". In: *Linear Algebra and its Applications* 197 (1994), pp. 143 –176. ISSN: 0024-3795. DOI: http://dx.doi.org/10.1016/0024-3795(94)90486-3. URL: http://www.sciencedirect.com/science/article/pii/0024379594904863.

[55] Michel Minoux. "Accelerated greedy algorithms for maximizing submodular set functions". English. In: *Optimization Techniques*. Ed. by J. Stoer. Vol. 7. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 1978, pp. 234–243. ISBN: 978-3-540-08708-3. DOI: 10.1007/BFb0006528. URL: http://dx.doi.org/10.1007/BFb0006528.

[56] "Modern Portfolio Theory". In: (). URL: https://en.wikipedia.org/wiki/Modernportfoliotheory.

[57] "National Centers For Environmental Information". In: (). URL: http://www.ncdc.noaa.gov/.

[58] Ebisa Negeri and Nico Baken. "Architecting the Smart Grid As a Holarchy". In: *Proceedings of the 1st International Conference on Smart Grids and Green IT Systems* (2012), pp. 73–78. DOI: 10.5220/0003952500730078. URL: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0003952500730078.

[59] M E J Newman. "Finding community structure in networks using the eigenvectors of matrices". In: (). arXiv: 0605087v3 [arXiv:physics].

[60] M. E. J. Newman. "Spectral methods for community detection and graph partitioning". In: *Phys. Rev. E* 88 (4 2013), p. 042822. DOI: 10.1103/PhysRevE.88.042822. URL: http://link.aps.org/doi/10.1103/PhysRevE.88.042822.

[61] Mark Newman. *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc., 2010. ISBN: 0199206651, 9780199206650.

[62] Mark E. J. Newman. "Random graphs as models of networks". In: *Handbook of Graphs and Networks: From the Genome to the Internet* 1 (2002), pp. 34–68. ISSN: 03788733. DOI: 10.1002/3527602755.ch2. arXiv: 0202208 [cond-mat].

[63] P. H. Nguyen, W. L. Kling, and P. F. Ribeiro. "A game theory strategy to integrate distributed agent-based functions in smart grids". In: *IEEE Transactions on Smart Grid* 4.1 (2013), pp. 568–576. ISSN: 19493053. DOI: 10.1109/TSG.2012.2236657.

[64] "Nice Grid". In: (). URL: http://nicegrid.fr/.

[65] K. Ogata. *Modern Control Engineering*. Instrumentation and controls series. Prentice Hall, 2010. ISBN: 9780136156734. URL: https://books.google.com.et/books?id=Wu5GpNAelzkC.

[66] J.-P. Onnela, K. Kaski, and J. Kertesz. "Clustering and information in correlation based financial networks". In: *The European Physical Journal B - Condensed Matter* 38.2 (2004), pp. 353–362. ISSN: 1434-6028. DOI: 10.1140/epjb/e2004-00128-7. URL: http://www.springerlink.com/openurl.asp?genre=article{\&}id=doi:10.1140/epjb/e2004-00128-7.

[67] J. P. Onnela et al. "Dynamics of market correlations: Taxonomy and portfolio analysis". In: (2003). DOI: 10.1103/PhysRevE.68.056110. eprint: arXiv:cond-mat/0302546.

[68] "Our Finite World". In: (). URL: https://ourfiniteworld.com/2012/03/12/world-energy-consumption-since-1820-in-charts/.

[69] John F Padgett and Christopher K Ansell. "7. Robust Action and the Rise of the Medici, 1400-1434". In: *American Journal of Sociology* 98.6 (1993), pp. 1259–1319. ISSN: 0002-9602. DOI: 10.1086/230190.

[70] Giuliano Andrea Pagani and Marco Aiello. "The Power Grid as a complex network: A survey". In: *Physica A: Statistical Mechanics and its Applications* 392.11 (2013), pp. 2688 –2700. ISSN: 0378-4371. DOI: http://dx.doi.org/10.1016/j.physa.2013.01.023.

[71] Giuliano Andrea Pagani and Marco Aiello. "Towards Decentralized Trading : A Topological Investigation of the Dutch Medium and Low Voltage Grids". In: (2011), pp. 1–50. arXiv: arXiv:1101.1118v4.

[72] S Pahwa, C Scoglio, and N Schulz. "Topological Analysis and Mitigation Strategies for Cascading Failures in Power Grid Networks". In: (). arXiv: arXiv:1212.5620v1.

[73] Gergely Palla et al. "Uncovering the overlapping community structure of complex networks in nature and society - Supplementary". In: *Nature* 435.7043 (2005), pp. 814–818. ISSN: 1476-4687. DOI: 10.1038/nature03607. arXiv: 0506133 [physics].

[74] Yael Parag and Benjamin K. Sovacool. "Electricity market design for the prosumer era". In: *Nature Energy* March (2016), p. 16032. ISSN: 2058-7546. DOI: 10.1038/nenergy.2016.32. URL: http://www.nature.com/articles/nenergy201632.

[75] Pascal Pons and Matthieu Latapy. *Computing communities in large networks using random walks (long version)*. 2005. eprint: arXiv:physics/0512106.

[76] Usha Nandini Raghavan, Reka Albert, and Soundar Kumara. "Near linear time algorithm to detect community structures in large-scale networks". In: (2007). DOI: 10.1103/PhysRevE.76.036106. eprint: arXiv:0709.2938.

[77] Sarvapali D. Ramchurn et al. "Putting the 'Smarts' into the Smart Grid: A Grand Challenge for Artificial Intelligence". In: *Commun. ACM* 55.4 (Apr. 2012), pp. 86–97. ISSN: 0001-0782. DOI: 10.1145/2133806.2133825. URL: http://doi.acm.org/10.1145/2133806.2133825.

[78] A. J. Dinusha Rathnayaka et al. "Prosumer Management in Sociotechnical Smart Grid". In: *Proceedings of the CUBE International Information Technology Conference*. Pune, India, 2012, pp. 483–489. ISBN: 978-1-4503-1185-4. DOI: 10.1145/2381716.2381808.

[79] Martí Rosas-Casals, Sergi Valverde, and Ricard V. Solé. "Topological Vulnerability of the European Power Grid Under Errors and Attacks". In: *International Journal of Bifurcation and Chaos* 17.07 (2007), pp. 2465–2475. ISSN: 0218-1274. DOI: 10.1142/S0218127407018531.

[80] Walid Saad et al. *Game Theoretic Methods for the Smart Grid*. 2012. eprint: arXiv:1202.0452.

[81] P. Samadi et al. "Advanced Demand Side Management for the Future Smart Grid Using Mechanism Design". In: *IEEE Transactions on Smart Grid* 3.3 (2012), pp. 1170–1180. ISSN: 1949-3053. DOI: 10.1109/TSG.2012.2203341.

[82] A. Sapienza et al. "Detecting Anomalies in Time-Varying Networks Using Tensor Decomposition". In: *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*. 2015, pp. 516–523. DOI: 10.1109/ICDMW.2015.128.

[83] Sayama. *Modeling Complex systems*. Vol. 53. 9. 2013, pp. 1689–1699. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[84] Thomas Schank and Dorothea Wagner. "Finding, Counting and Listing all Triangles in Large Graphs, An Experimental Study". In: *Framework* 001907 (2001), pp. 3–6. ISSN: 03029743. DOI: 10.1007/b136461.

[85] John W. Simpson-Porco, Florian Dorfler, and Francesco Bullo. "Voltage collapse in complex power grids". In: *Nat Commun* 7 (2016). Article. URL: http://dx.doi.org/10.1038/ncomms10790.

[86] Steven H. Strogatz. "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators". In: *Physica D: Nonlinear Phenomena* 143.1–4 (2000), pp. 1 –20. ISSN: 0167-2789. DOI: http://dx.doi.org/10.1016/S0167-2789(00)00094-4. URL: http://www.sciencedirect.com/science/article/pii/S0167278900000944.

[87] Tyler H. Summers, Fabrizio L. Cortesi, and John Lygeros. "On Submodularity and Controllability in Complex Dynamical Networks". In: *Ieee Transactions on Control of Network Systems* 3.1 (2016), pp. 91–101.

[88] Osman Ya ğan. "Robustness of power systems under a democratic-fiber-bundle-like model". In: *Phys. Rev. E* 91 (6 2015), p. 062811. DOI: 10.1103/PhysRevE.91.062811. URL: http://link.aps.org/doi/10.1103/PhysRevE.91.062811.

[89] Frans Van Hulle et al. "Integrating Wind, European project report". In: (2009).

[90] Wenye Wang, Yi Xu, and Mohit Khanna. "A survey on the communication architectures in smart grid". In: *Computer Networks* 55.15 (2011), pp. 3604–3629. ISSN: 13891286. DOI: 10.1016/j.comnet.2011.07.010. URL: http://linkinghub.elsevier.com/retrieve/pii/S138912861100260X.

[91] D J Watts and S H Strogatz. "Collective dynamics of 'small-world' networks." In: *Nature* 393.6684 (1998), pp. 440–2. ISSN: 0028-0836. DOI: 10.1038/30918.

[92] Ye Yan et al. "A Survey on Smart Grid Communication Infrastructures: Motivations, Requirements and Challenges". In: *IEEE Communications Surveys & Tutorials* 15.1 (2013), pp. 5–20. ISSN: 1553-877X. DOI: 10.1109/SURV.2012.021312.00034. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6157575.

[93] J. Yang and J. Leskovec. "Overlapping Communities Explain Core Periphery Organization of Networks". In: *Proceedings of the IEEE* 102.12 (2014), pp. 1892–1902. ISSN: 0018-9219. DOI: 10.1109/JPROC.2014.2364018.

[94] Daniel Young, Mike Mcgaraghan, and Nate Dewart. "Leveraging Big Data to Develop Next Generation Demand Side Management Programs and Energy Regulations Recent Research Methods Focusing on Dynamic Market Analysis". In: (2014), pp. 332–347.

[95] Qiang Yu and Ling Chen. "A New Method for Detecting Anti-community Structures in Complex Networks". In: *Journal of Physics: Conference Series* 410.1 (2013), p. 012103. ISSN: 1742-6596. DOI: 10.1088/1742-6596/410/1/012103.

[96] Zhengzhong Yuan et al. "Exact controllability of complex networks." In: *Nature communications* 4 (2013), p. 2447. ISSN: 2041-1723. DOI: 10.1038/ncomms3447. arXiv: 1310.5806. URL: http://www.nature.com/ncomms/2013/130912/ncomms3447/full/ncomms3447.html.

[97] Yingrui Zhang and Osman Yagan. *Optimizing the robustness of electrical power systems against cascading failures*. 2016. eprint: arXiv:1602.07763.

[98] Yu Zhang, Nikolaos Gatsis, and Georgios B. Giannakis. "Robust energy management for microgrids with high-penetration renewables". In: *IEEE Transactions on Sustainable Energy* 4.4 (2013), pp. 944–953. ISSN: 19493029. DOI: 10.1109/TSTE.2013.2255135. arXiv: arXiv:1207.4831v3.

[99] Z. Zhang and M. Y. Chow. "Incremental cost consensus algorithm in a smart grid environment". In: *2011 IEEE Power and Energy Society General Meeting*. 2011, pp. 1–6. DOI: 10.1109/PES.2011.6039422.

[100] Eric Zivot. "Chapter 1 Portfolio Theory with Matrix Algebra". In: (2013).