HAL open science

Optimal control of energy flexibilities in a stochastic environment

Maxime Grangereau

To cite this version:

Maxime Grangereau. Optimal control of energy flexibilities in a stochastic environment. Optimization and Control [math.OC]. Institut Polytechnique de Paris, 2021. English. NNT : 2021IPPAX010 . tel-03221066

HAL Id: tel-03221066

https://theses.hal.science/tel-03221066

Submitted on 7 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAX010

Thèse de doctorat

Contrôle optimal de flexibilités énergétiques en contexte incertain

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'Ecole polytechnique

École doctorale n°574 Ecole Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat: Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 23 mars 2021, par

Maxime Grangereau

Composition du Jury :

Agnès Sulem Directeur de recherche, INRIA Paris	Présidente
François Delarue Professeur, Université Côte d'Azur (Laboratoire J.A. Dieudonné)	Rapporteur
Roland Malhamé Professeur, Polytechnique Montréal (GERAD)	Rapporteur
Miguel Anjos Professeur, University of Edinburgh (School of Mathematics)	Examinateur
Nizar Touzi Professeur, Ecole polytechnique (CMAP)	Examinateur
Emmanuel Gobet Professeur, Ecole polytechnique (CMAP)	Directeur de thèse
Stéphane Gaubert Directeur de recherche, INRIA and Ecole polytechnique (CMAP)	Co-directeur de thèse
Wim van Ackooij Chercheur expert, EDF R&D (Département Osiris, groupe R36)	Encadrant industriel

Remerciements

Je remercie tout d'abord François Delarue et Roland Malhamé d'avoir accepté de rapporter ma thèse, ainsi que Nizar Touzi, Agnès Sulem et Miguel Anjos de prendre part à mon jury de soutenance, ce qui constitue pour moi une vraie fierté au vu de leurs contributions scientifiques.

Je tiens à remercier tout particulièrement mes directeurs et encadrants Emmanuel Gobet, Stéphane Gaubert et Wim van Ackooij, qui m'ont accompagné sur ce chemin sinueux qu'est une thèse.

Merci Emmanuel de m'avoir suivi tout au long de mon stage, puis de ma thèse, avec toujours le juste dosage entre accompagnement, incitation à me dépasser et conseils pour me développer professionnellement. Tu peux être fier de toi: tu auras réussi à me convertir à la régression empirique (que tu m'enseignais sur les bancs de l'école quelques années plus tôt), au contrôle stochastique et autres noms barbares qui amusent mes proches quand ils m'entendent les prononcer!

Merci Stéphane de m'avoir communiqué ta passion pour la recherche opérationnelle, ce qui m'a faire un projet d'EA, un stage de recherche puis un double diplôme à Munich. Je te suis très reconnaissant de tes conseils et ta bienveillance au cours de ces dernières années.

Merci Wim de m'avoir fait confiance et de m'avoir offert une grande liberté dans la construction de mon sujet de thèse. Et merci de m'avoir supporté dans ton bureau au cours des 3 dernières années. On en aura passé du temps, à blaguer sur les méthodes heuristiques!

Je remercie également Ghislain Hallez et Julien Loron pour avoir créé un climat d'écoute et de confiance, ce que je considère comme essentiel pour s'épanouir dans son travail.

Je suis également très reconnaissant envers les personnes qui m'auront aidé lors de ces 3 ans de thèse à Osiris. Je remercie sincèrement Clémence Alasseur, que j'ai rencontré lors des séminaires FIME (au cours desquels j'étais le plus souvent complètement perdu au début) et qui a joué le rôle de vrai catalyseur pour que je puisse faire une thèse à Osiris. Merci à Nadia Oudjane pour nos échanges, ainsi que d'avoir collaboré avec Wim et moi sur un projet sur les méthodes particulaires. Je remercie Riadh Zorgati pour ces nombreuses discussions sur l'OPF ainsi que sur des problèmes métier ou maths, et pour tous les bons plans concernant les summer/winter schools (avec une mention spéciale pour le ski en Norvège). Je remercie également Mathieu Caujolle, Héloïse Dutrieux Baraffe et Bhargav Prasanna Swaminathan qui m'ont aidé à y voir plus clair sur l'épineux problème d'Optimal Power Flow.

Je remercie également les équipes administratives d'EDF, Sylvie Malézieux, Christelle Roger et Audrey Cathala en particulier, et du CMAP: Nasséra Naar et Alexandra Noiret pour leur bonne humeur et leur patience lorsque j'avais des questions pour mes démarches (et il y en a dans une thèse!).

Je suis reconnaissant envers Tom Siebel et sa fondation pour le soutien et les perspectives qu'ils m'ont offerts. Je reste marqué par la conférence sur les réseaux sociaux à Chicago, accompagnée de son inoubliable concert du groupe Maroon 5 dans le Musée d'Histoire Naturelle ! Tout mon entourage en aura entendu parler !

Une mention particulière à la team de stagiaires/doctorants (anciens ou actuels) EDF: Cécile Rottner pour être si bon public même à mes blagues ratées, Paulin Jacquot: on se suit depuis notre entrée à l'X (gendarmerie, EDF, référent des doctorants Osiris, grand frère de thèse...), Michael Richard, mon autre modèle en tant que doctorant référent Osiris, Rodolphe Griset, et son légendaire côté bon-vivant, Adèle Pass-Lanneau qui aura soutenu une semaine avant moi alors qu'elle est arrivée 6 mois plus tard que moi à EDF, Adrien Séguret, qui me doit un ordinateur portable! Mention également à Margarita Veshchezerova, dont j'ai dû vérifier l'orthographe du nom en l'écrivant mais

dont je ne suis toujours pas sûr, Ksucha Syrtseva, que j'ai rencontré lors du cours sur les MDP puis nous a rejoint à Osiris, Benoît Sohet, militant inconditionnel de l'investissement responsable sur le PEG, Maximilien Germain, l'étoile montante qui écrit des papiers plus vite que son ombre, Margaux Brégère, qui organise des séminaires plus vite que son ombre, Bruno Barracosa, notre modèle de sagesse, Emma Hubert, toujours bon public et qui égaye les couloirs d'Osiris. Petite pensée à Laura Tinsi et David Obst qui ont organisé la journée des doctorants avec Adèle, après que j'en ai été exempté car j'étais référent des doctorants Osiris. Sans oublier Paul Javal et Irene Danti-Lopez!

Je remercie également mes collègues de R36 pour tous les moments d'échange et de convivialité passés ensemble, en réunion, autour d'un café, d'un verre aux Grands Voisins ou au Shannon pour les plus téméraires. Merci donc à Rebecca, Olivier & Olivier, Sandrine, Sandy, Saad, Hugo, Marc, Jayakumar, Thomas, Jean-Yves, Cheng, Pierre, Fatine, Andrea, Guilhem, Fabienne, Raphaël, Arnaud, Jean-Armand, Thibault, Pascal, Pascale, Beyram (et j'espère n'oublier personne). Merci plus largement à mes collègues d'Osiris ou d'autres départements avec qui j'ai échangé lors de ma thèse. Les citer tous me demanderait un an de veille bibliographique pour m'assurer de n'avoir oublié personne et ferait passer mon manuscrit de 200 à 400 pages: je serais alors accusé de gonfler artificiellement ma thèse.

Je remercie également les ingénieurs, chercheurs, doctorants et post-doctorants du CMAP avec qui j'ai passé de bons moments même si j'ai souvent délaissé mon bureau du CMAP pour celui d'EDF, juste à côté. Merci en particulier à Heythem, Frédéric, Corentin, Othmane, Valentina, Luca, Julian, Franchesco, Lionel, Francisco, Anne-Claire. Mention particulière à mes frères et soeurs de thèse: Florian Bourgey pour les moments passés lors de ce super voyage en Australie rythmé par les restaurants asiatiques, Linda Chamakh, ma voisine de séminaire PGMO, Marin Boyer, de la Team Clémenceau, Paulin Jacquot encore une fois, Isaque Pimentel, Margaux Faucher, Quentin Jacquet, Omar Saadi, Michael Allouche, Manon Rivoire.

Enfin je tiens à remercier mes proches pour leur soutien et pour avoir égayé mon parcours au cours de ces trois dernières année. Je remercie en particulier mes amis, même si la crise sanitaire a fait que nous nous sommes moins vus ces derniers temps.

Je remercie mes parents, Dominique et Sylvia, ainsi que mon frère et ma soeur, Thomas et Aurore pour leur soutien au cours de mes études puis pendant cette aventure. Merci de m'avoir toujours encouragé à me dépasser.

Enfin, je remercie Stéphanie d'avoir été là et de m'avoir soutenu. Merci en particulier pour ta patience et d'avoir fait avec mes longues heures de télétravail dans notre petit appartement au cours de cette dernière année.

List of publications

Here is a list of articles, all submitted, that were written during this thesis:

- Gobet, E. & Grangereau, M. (2021). *Extended McKean-Vlasov optimal stochastic control applied to smart grid management*. Submitted to ESAIM: Control, Optimisation and Calculus of Variations (ESAIM: COCV).
  Available at https://hal.archives-ouvertes.fr/hal-02181227/

- Gobet, E. & Grangereau, M. (2021). *Federated stochastic control of numerous heterogeneous energy storage systems*. Submitted to Journal of Optimization Theory and Applications (JOTA).
  Available at https://hal.archives-ouvertes.fr/hal-03108611/document

- Gobet, E. & Grangereau, M. (2021). *Newton method for stochastic control problems*. Submitted to SIAM Journal on Control and Optimization (SICON).
  Available at https://hal.archives-ouvertes.fr/hal-03108627/document

- M. Grangereau, S. Gaubert & W. van Ackooij (2021). *Multi-stage Stochastic Alternating Current Optimal Power Flow with Storage: Bounding the relaxation gap*. Submitted to Electric Power Systems Research (EPSR).

Contents

1 Introduction **11**
1.1 Climate change and renewable energy development . . . 11
1.2 Historical context in energy management . . . 11
1.2.1 Historical features of electricity production, consumption and transmission . . . 11
1.2.2 Supply-demand balance . . . 12
1.2.3 Electricity networks . . . 14
1.3 Change of paradigm due to the rising share of renewable energy sources . . . 14
1.3.1 Supply-demand balance . . . 14
1.3.2 Network management . . . 15
1.4 New opportunities and challenges . . . 16
1.4.1 Decentralized energy storage systems . . . 16
1.4.2 Demand response . . . 17
1.4.3 Leverage flexibility from multi-generation and combined operation of multi-energy networks . . . 17
1.4.4 Local controls for network operation . . . 18
1.5 Challenges in energy management . . . 18
1.6 Objectives and contents of the thesis . . . 19
1.7 Contributions of the chapters . . . 20
1.8 State-of-the-art . . . 23
1.8.1 Stochastic control of micro-grids (without network constraints modeling) . . . 23
1.8.2 State-of-the-art for optimal control of electricity networks with network constraints modeling . . . 25

2 Introduction (en français) **27**
2.1 Changement climatique et développement des énergies renouvelables . . . 27
2.2 Contexte historique en gestion de l'énergie . . . 27
2.2.1 Spécificités historiques de la production, de la consommation et du transport de l'électricité . . . 27
2.2.2 Équilibre offre-demande . . . 28
2.2.3 Réseaux d'électricité . . . 30
2.3 Changement de paradigme lié à l'essor des renouvelables . . . 30
2.3.1 Équilibre offre-demande . . . 30
2.3.2 Gestion des réseaux . . . 32
2.4 Nouvelles opportunités et nouveaux défis . . . 33
2.4.1 Systèmes de stockage décentralisés . . . 33
2.4.2 Demand response . . . 33
2.4.3 Multi-génération et pilotage conjoint de réseaux multi-énergies . . . 34
2.4.4 Moyens de contrôle locaux pour piloter les réseaux . . . 34
2.5 Défis en gestion de l'énergie . . . 35
2.6 Objectifs et contenu de la thèse . . . 35
2.7 Contributions des chapitres . . . 37
2.8 État de l'art . . . 39

2.8.1 Contrôle de micro-grids en contexte incertain (sans prise en compte des contraintes physiques du réseau) . . . 40
2.8.2 État de l'art pour le contrôle optimal de réseaux électriques prenant en compte les contraintes physiques . . . 42

I Optimal control of micro-grid in a stochastic environment **45**

3 Day-ahead Commitment **47**
3.1 Introduction . . . 47
3.2 Stochastic control and MKV-FBSDEs . . . 51
3.2.1 Stochastic model and smart grid framework . . . 51
3.2.2 Standing assumptions . . . 52
3.2.3 Necessary condition for optimality . . . 53
3.2.4 Solvability of the MKV Forward-Backward SDE . . . 54
3.2.5 Existence and uniqueness of critical point do not necessarily imply existence of a minimum . . . 55
3.2.6 Existence of an optimal control . . . 56
3.2.7 Sufficient condition for optimality . . . 56
3.3 Effective computation and approximation of battery control . . . 58
3.3.1 Model/Context . . . 58
3.3.2 Computation of first order expansion . . . 59
3.3.3 Effective simulation of first order expansion of optimal control . . . 61
3.4 Proofs . . . 67
3.4.1 Proof of Theorem 3.2.2 . . . 67
3.4.2 Proof of Theorem 3.2.4 . . . 70
3.4.3 Proof of Proposition 3.2.6 . . . 72
3.4.4 Proof of Proposition 3.3.2 . . . 73
3.4.5 Proof of Proposition 3.3.6 . . . 74
3.4.6 Boundedness of solutions to linear ODE with bounded stochastic coefficient . . . 75
3.5 Conclusion . . . 75

4 Decentralized stochastic control of energy storage systems **77**
4.1 Introduction . . . 77
4.2 Model, assumptions and first properties . . . 81
4.2.1 Assumptions . . . 81
4.2.2 Differentiability, convexity, characterization of optimality . . . 82
4.3 Decomposition of the problem and equivalent representation as a stochastic differential game . . . 83
4.3.1 The coordination problem . . . 83
4.3.2 The individual problems . . . 84
4.4 Approximate decentralized control architecture preserving privacy . . . 85
4.4.1 Convergence of the coordination problem for large populations . . . 86
4.4.2 Convergence of the individual problems for large populations . . . 88
4.5 Numerical experiments . . . 89
4.5.1 The model . . . 89
4.5.2 Solution method of affine-linear FBSDEs . . . 91
4.5.3 Numerical simulations and results . . . 95
4.6 Online decentralized control scheme with minimal telecommunication . . . 97
4.6.1 Discussion on the decentralized control scheme . . . 99
4.6.2 On the privacy of individual users habits . . . 100
4.7 Proofs . . . 100
4.7.1 Proof of Theorem 4.2.1 . . . 100
4.7.2 Proof of Proposition 4.3.1 . . . 101

4.7.3 Proof of Proposition 4.4.3 . . . 102
4.7.4 Proof of Theorem 4.4.5 . . . 102
4.7.5 Proof of Proposition 4.4.6 . . . 103
4.7.6 Proof of Corollary 4.4.8 . . . 104
4.7.7 Proof of Theorem 4.4.12 . . . 104
4.8 Conclusion . . . 105

5 Newton method for stochastic control problems **107**
5.1 Introduction . . . 107
5.2 Control problem: setting, assumptions and preliminary results . . . 111
5.2.1 Setting and assumptions . . . 111
5.2.2 Well-posedness, existence and uniqueness of an optimal control . . . 112
5.2.3 First-order necessary and sufficient optimality conditions . . . 112
5.2.4 Second-order differentiability . . . 114
5.3 The Newton method for stochastic control problems . . . 115
5.3.1 Definition and interpretation of the Newton step . . . 115
5.3.2 Solution of affine-linear FBSDEs and computation of the Newton step . . . 117
5.3.3 Global convergence of Newton's method with an adapted line-search method . . . 119
5.4 Application: energy storage system control for power balancing . . . 127
5.4.1 Problem setting . . . 127
5.4.2 Solving the stochastic control problem . . . 130
5.4.3 Practical implementation . . . 132
5.4.4 Analysis of the numerical performance . . . 137
5.4.5 Over-fitting, under-fitting, and automatic tuning of regression parameters by cross-validation . . . 137
5.4.6 Analysis of the results from the application point of view . . . 138
5.5 Proofs . . . 139
5.5.1 Proof of Lemma 5.3.5 . . . 139
5.5.2 Proof of Lemma 5.3.6 . . . 140
5.5.3 Proof of Theorem 5.3.7 . . . 141
5.5.4 Proof of Corollary 5.3.8 . . . 142
5.6 Conclusion . . . 142

II The Alternating Current Optimal Power Flow problem and its extension to the multi-stage stochastic setting **145**

6 Introduction to the Optimal Power Flow problem **147**
6.1 General presentation of the OPF problem . . . 147
6.2 Surveys . . . 147
6.3 The Bus Injection Model (BIM): formulation and equivalence . . . 147
6.4 The Branch Flow Model (BFM) . . . 149
6.5 Equivalence between BIM and BFM formulations . . . 151
6.6 Difficulty of AC OPF problem . . . 151
6.7 Approximation/Linearization of the AC OPF problem . . . 151
6.7.1 Single-phase networks . . . 151
6.7.2 Unbalanced multi-phase networks . . . 152
6.8 Non-linear programming . . . 152
6.9 Convex relaxations of the OPF problem . . . 153
6.9.1 Semi-definite, chordal and second-order cone relaxations . . . 153
6.9.2 Exactness conditions of convex relaxations of OPF problem . . . 153
6.9.3 Advanced convex relaxations of the OPF problem . . . 155
6.9.4 Unbalanced multiphase networks . . . 156

6.10 Optimal Power Flow in dynamic setting . . . 156
6.11 Optimal Power Flow in stochastic environment . . . 156
6.11.1 Chance-Constraints OPF . . . 156
6.11.2 Multistage stochastic OPF . . . 157

7 Bounding the duality gap of the multi-stage stochastic AC OPF problem **159**
7.1 Introduction . . . 160
7.1.1 Optimal Power Flow Problem . . . 161
7.1.2 Existing work on convex relaxations of AC OPF . . . 161
7.1.3 Our contributions . . . 162
7.1.4 Related work . . . 162
7.1.5 Outline of the chapter . . . 163
7.2 The multistage stochastic AC OPF model . . . 163
7.2.1 Formulation of the problem . . . 163
7.2.2 On the formulation of the non-anticipativity constraints . . . 164
7.2.3 Second-Order Cone relaxation of the problem . . . 166
7.3 Restriction of the feasible set ensuring a vanishing relaxation gap . . . 166
7.3.1 Presentation of the problem with restricted feasible set . . . 166
7.3.2 Second-order cone relaxation of the problem with restricted feasible set . . . 167
7.3.3 Conditions ensuring equality of the feasible sets of the original and restricted problems . . . 167
7.3.4 Vanishing relaxation gap for the problem with restricted feasible set . . . 169
7.4 Numerical illustration . . . 173
7.4.1 General framework . . . 174
7.4.2 Network model . . . 174
7.4.3 Generation of i.i.d. scenarios for solar power production . . . 177
7.4.4 Choice of scenario tree structure . . . 179
7.4.5 Numerical results . . . 179
7.5 Explicit bounds on duality gap of multi-stage stochastic problems in energy management . . . 181
7.5.1 Shapley-Folkman Theorem and related results . . . 181
7.5.2 On node formulation of multi-stage stochastic programming problems . . . 183
7.5.3 Motivation of the problem . . . 184
7.5.4 Bounds of the duality gap . . . 185
7.6 Conclusion . . . 188

A Acronyms and Abbreviations **191**

Chapter 1

Introduction

1.1 Climate change and renewable energy development

In most industrialized countries, there is a rising concern about climate change and strong supports to solutions allowing to reduce greenhouse gas emissions. Substantial efforts are planned to limit the temperature increase well below $2°C$ by 2100, as targeted by the 2015 Paris agreement. Conventional energy production units such as coal, oil, and gas power plants are considered to be one the main sources of greenhouse gas emissions. Some studies show that they correspond to as much as 25% of the worldwide CO2 emissions, see [Cha+14]. To reduce the carbon footprint of electricity production, renewable energy sources, such as solar panels and wind turbines, are deemed viable solution. Recently, the European Commission has set an objective of a share 32% for renewable energy sources in the energy mix by 2030 [Com19]. Research on these energy sources has allowed to drastically reduce their production costs. For instance, the price of Photovoltaic cells dropped by 99.6 % from 1976 to 2018, from $79.3/W to $0.3/W, see [WB19]. However, the aforementioned renewable energy sources suffer from major drawbacks which require particular care for their integration in the energy mix. They are non-dispatchable (i.e., non-controllable) intermittent sources [Suc+20] and the amount of energy they produce is random, due to the inaccuracy of weather forecasts. These issues raise challenging problems in the field of energy management. To understand the impact of the rise of renewable energy on energy management, we first need to briefly outline the historical context of this field.

1.2 Historical context in energy management

1.2.1 Historical features of electricity production, consumption and transmission

The majority of electricity production is done centrally by dispatchable power plants, which include coal, fuel, gas or nuclear power plants. The electricity produced is sent to the end users by electricity wires through the transmission then the distribution networks. The transmission network is a high voltage network (225 to 400 kV) with meshed architecture (i.e., with cycles), with a single phase. It is operated by Réseau de Transport d'Electricité (RTE) in France, which is a regulated monopoly. The distribution network is a low (from 220 to 430 V) to medium voltage (1 kV to 52 kV) network. It is traditionally operated radially (i.e., without cycles) by Enedis in France, which is a regulated monopoly as well. The global historical architecture of power networks allowing to deliver power from power plants to consumers is represented in Figure 1.1[1].

[1] Icons made by Freepik and Smashicons from www.flaticon.com

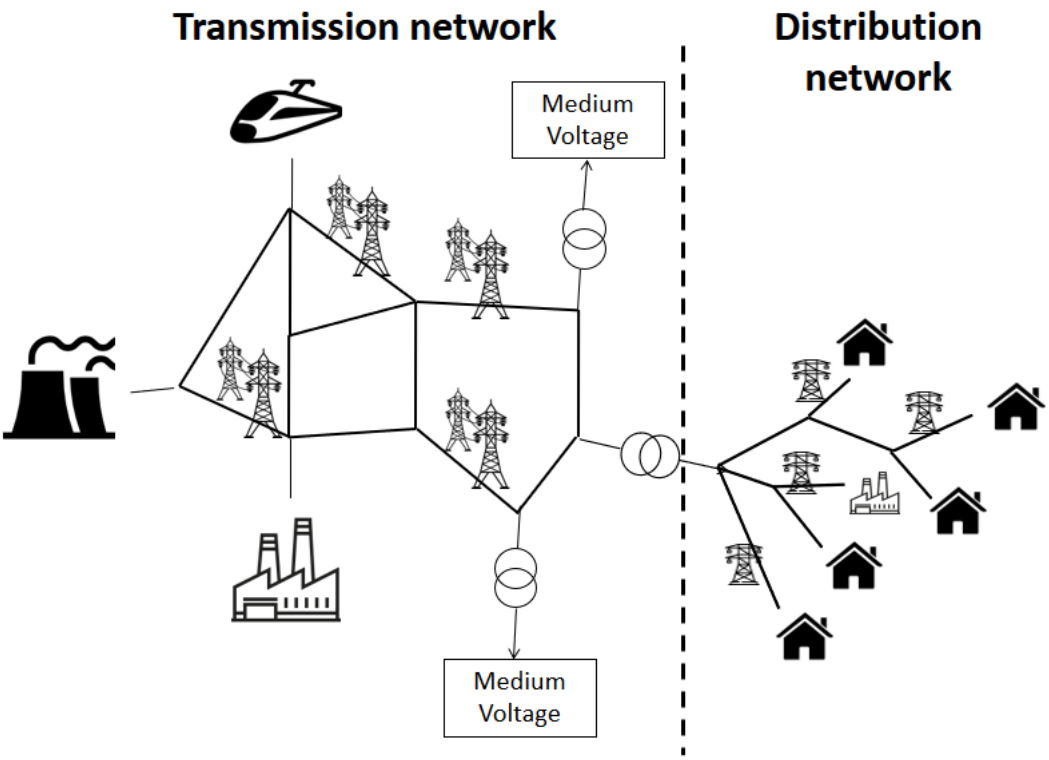


Figure 1.1: Historical structure of electricity networks

1.2.2 Supply-demand balance

Generalities

A particularity of electricity is that, unlike many other goods, it can only be stored marginally, due to the high storage costs. In France, the Pumped Hydroelectric Energy Storage systems are the main tool to store energy. Due to their installation costs and requirements, they can only absorb 3% of the yearly consumption and allow to move around 1.5% of the nationwide yearly consumption.

For this reason, supply-demand balance for electrical energy has to be ensured at all time. Many actors are responsible for the balance of the network: energy suppliers, traders, aggregators, producers. All these actors need to ensure that, on their perimeters, production and consumption are equal. For instance, EDF, as electricity producer and supplier is responsible for the balance between its production (which includes electricity generation by its power plants, energy bought on markets...) and the consumption of its customers. EDF plans its production to match the forecast consumption of its customers (up to margins). In other words, power balance is traditionally done by adjusting the production to the forecast consumption.

Demand forecast

Demand on a specific day depend on several factors, like the type of day (week day, public holiday, strike day), specific events (sport competitions) and meteorological conditions, like the temperature. Temperature is one of the biggest explanatory factor for consumption differences between days [PMV02], due to heating and cooling. However, buildings have thermal inertia, and users are usually indifferent with respect to small temperature variations. For this reason, it is mostly an averaged over time version of the temperature which is important to forecast accurately power demand [Jov+15]. Therefore, instantaneous weather conditions play a smaller role than their average over time, which can be more accurately forecast than real-time conditions. This explains why demand can be forecast accurately at short term, in a context of low renewable energy penetration. Although weather forecasts can exhibit

errors, the weather trends (average over a time window for instance) can be predicted rather accurately at short term.

Production management

To lower production costs and penalties due to imbalance between their production and the consumption of their customers, producers need to adjust their production to the demand forecasts. This is done from the long term to the short term before the period considered:

- From 20 to 5 years ahead, decisions to build new power plants, to establish strategic partnerships, or to extend existing networks are taken.
- From 3 years to 2 weeks ahead, the maintenance planning of nuclear power plants, hydro power plants are designed. Resources are optimized, like Uranium reserves or water in hydro-reservoirs.
- From 2 weeks to 1 day ahead, more accurate forecasts of the demand are available, and adjustments in plannings of flexible production units are made.
- One day ahead, the forecast of aggregated demand and the production plannings are transmitted to Balancing Authority (RTE in France).
- During the day, some last-minute adjustments are made, by activating bilateral contracts with third parties or modifying the operating point of some flexible power plants.
- In real time, power-balance is ensured by RTE at the scale of the transmission network by activating some reserve mechanisms.

Energy is bought throughout the process (long term to intra-day). The degree of uncertainty taken into account as well as the physical accuracy of models required in energy management problems, depending on the planning horizon, is represented in Figure 1.2[2].

Real-time adjustment mechanisms

In real-time, unforeseen events may happen, and corrections need to be made to ensure power balance. In France, these real-time adjustments are ensured by RTE (Réseau de Transport d'Electricité), the regulated monopoly in charge of regulating the Transmission Network. The real-time adjustments are based on 3 control layers called Frequency reserves.

1. The Primary reserve (which can be activated from 1 second after a specific event until 30 seconds after) aims at stopping deviation of the frequency of the network from its nominal value (50 Hz in Europe). It is automatically deployed.
2. The Secondary reserve (activated 30 seconds to 30 minutes after an event) aims at restoring the frequency at the scale of the network. It is automatic, just like the Primary reserve.
3. The tertiary reserve (activated 30 minutes after an event or later) aims at adjusting the production unit set-points to restore margins for reserves. It is activated manually.

Frequency reserves are activated by RTE and currently, they are supplied mainly by generators, and to a lesser extent by load-shedding. When participating to the capacity mechanisms, generators allocate part of their capacities to primary and secondary reserves.

[2] Icons made by Freepik and Smashicons from www.flaticon.com

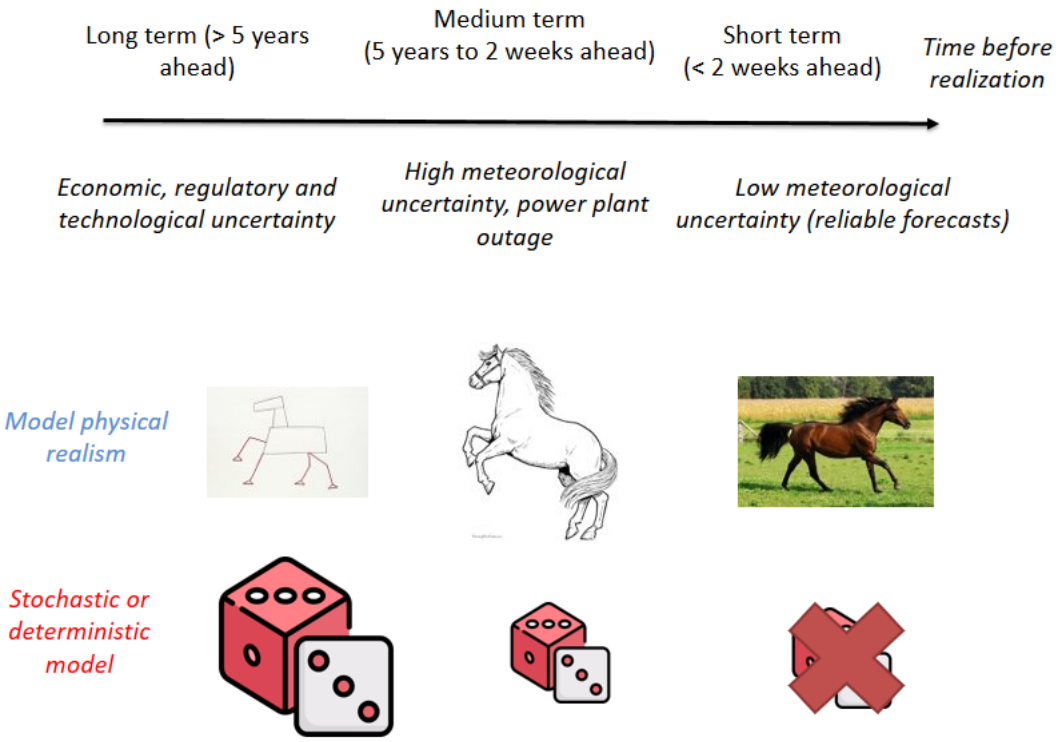


Figure 1.2: Historical structure of decision-making models for energy management

1.2.3 Electricity networks

Network values (voltage magnitude, intensity magnitude) have to be maintained in certain operational ranges. As power flows traditionally from large centralized power plants towards the loads (consumers), the power flow is traditionally unidirectional for distribution networks, which are operated radially. Therefore, voltage magnitude at the buses of the network is traditionally monotone decreasing from the root (substation) to the leaves of the tree formed by the distribution network. Hence, to provide corrective voltage control, it is traditionally enough to measure and modify the voltage at the root, using a transformer. This is similar to traffic regulating: for one-way streets, to remove congestion in a line, it suffices to modulate vehicles arrivals at the root. This metaphor is represented in Figure 1.3[3].

1.3 Change of paradigm due to the rising share of renewable energy sources

1.3.1 Supply-demand balance

Generalities

The intermittency of renewable energy sources like solar panels and wind turbines can create rapid power generation variations on the network [Suc+20]. It makes supply-demand balance more challenging as it requires more frequency reserves, with increased ramping capabilities, defined as the ability to follow steep variations in global supply-demand balance. On the other hand, the uncertainty of production can create errors in the data of problems allowing to design operation plannings, if we use traditional deterministic models.

[3] Icons made by Freepik and Smashicons from www.flaticon.com

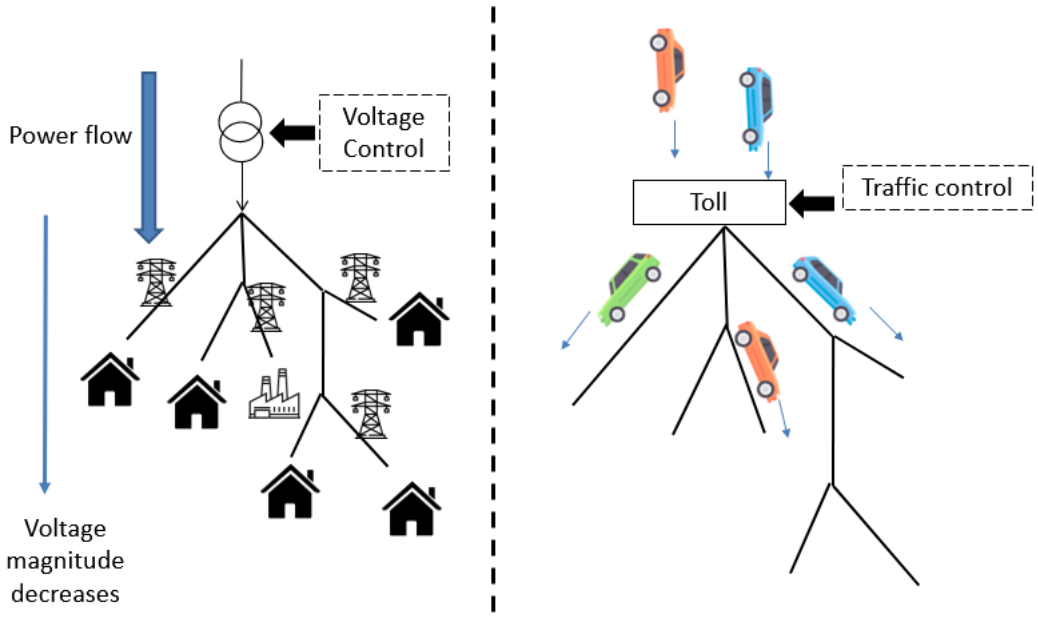


Figure 1.3: Historical network regulation - Comparison with traffic routing

Demand and production forecast

Due to their very nature, solar and wind power production depend on instantaneous meteorological conditions, which exhibit temporal and geographical correlations [ADS99; ZDK16]. Therefore, the higher the share of renewable energy sources, the more the production depends on real-time weather conditions. With this respect, stochastic models of solar [Bad+18; SG16] and wind production [Pin+09; ADS99] can bring significant value [MA+16] by decreasing electricity generation costs. They can be used as inputs of stochastic models used in energy management.

Production management

Production planning needs to take the uncertainty induced by renewable energy sources into account: stochastic models need to be used. For instance, short-term decision problems which are traditionally formulated in a deterministic framework need to be extended to the stochastic case, see Figure 1.4[4], like the Unit Commitment problem [Håb19; Ack+18] for instance.

Adjustment mechanisms

In a context of high penetration of renewable energy sources, the intermittency of these sources induces high forecast errors for the production, and therefore, more ancillary services are required. Power plants could provide these additional reserves, but this solution would be costly, as they would not work at their full capacity. Besides, operational constraints on power plants can prevent them to follow high ramps generated by the intermittent energy sources. Other mechanisms, like Demand Response or combined operation of multi-energy networks (presented later) seem like viable alternatives, provided some optimization and control tools are developed.

1.3.2 Network management

Traditionally, production was centralized, but it becomes more and more distributed: many consumers are now equipped with solar panels, and small wind/solar plants are now directly connected to distribution networks. Distributed generation can create local physical violations in the network. On the other hand, renewable energy sources

[4] Icons made by Freepik and Smashicons from www.flaticon.com

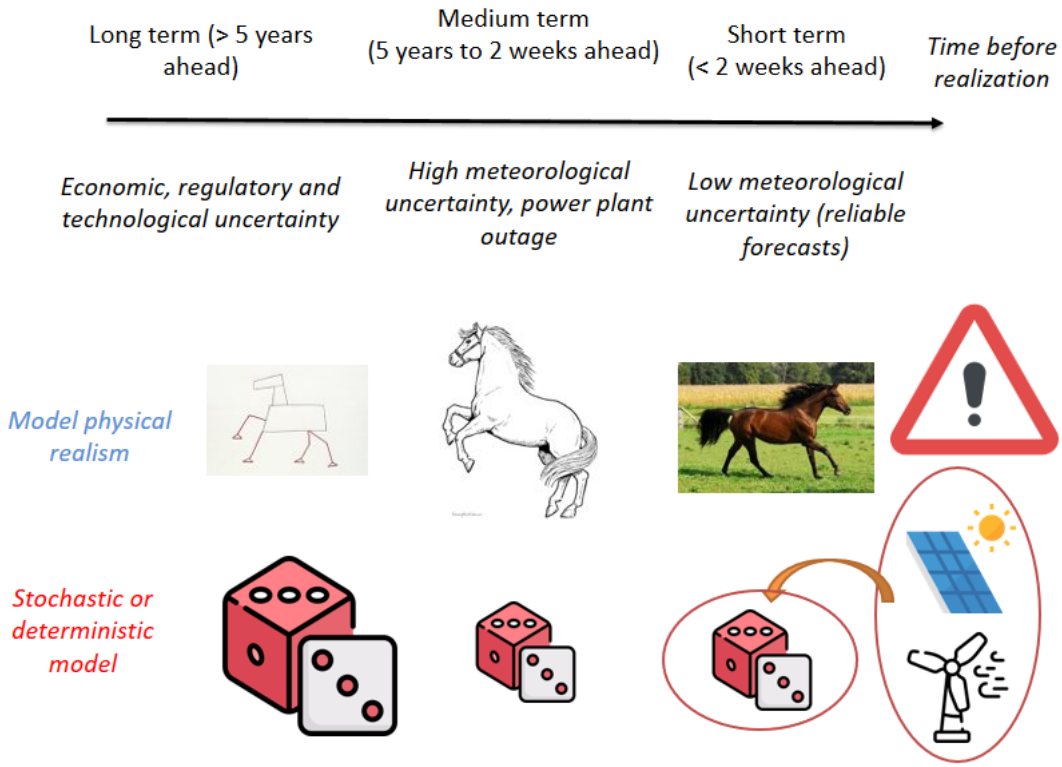


Figure 1.4: Structure of decision-making models for energy management with high renewable energy penetration

often exhibit a stochastic behavior. As a consequence, the network operating point is subject to uncertainty as well, which is problematic for network operation planning [ST14].

Besides, distributed generation creates two potential problems: reverse power flows and non-monotony of voltage along lines with respect to the distance to the substation [NT15]. The first issue relates to the fact that, with high renewable penetration, specific weather conditions may cause local production to exceed local consumption. This can cause power to flow from distribution networks to transmission networks, which were not designed for this purpose. The non-monotony of voltage along the tree formed by the distribution network makes voltage regulation more complex than before, as is illustrated in Figure 1.5[5]. We can go on with the metaphor of traffic regulation, and imagine a situation where we go from a situation with one-way streets with traffic control at a toll, to a situation where the streets become two-way. One can easily imagine that, for two-way streets, controlling the traffic only at the toll may be insufficient.

1.4 New opportunities and challenges

1.4.1 Decentralized energy storage systems

Ancillary services could be supplied by energy storage systems located at various points of electricity networks. Such projects already exist, such as the McHenry installation in Illinois, with total installed capacity of 20 MW, operational since 2015. This type of solution may be expensive, as it requires installation of power systems only for the purpose of integrating renewable sources. Other solutions could be found by leveraging already existing energy flexibilities: demand response, which aims at leveraging the diffuse flexibility of a large number of consumers, and combined operations of multi-energy networks.

[5] Icons made by Freepik and Smashicons from www.flaticon.com

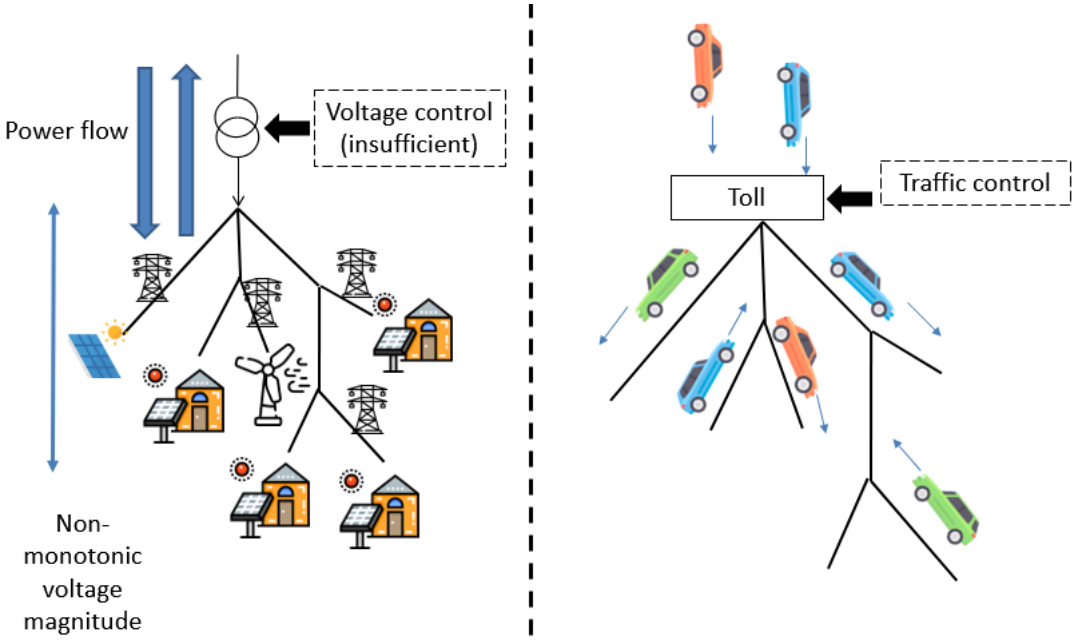


Figure 1.5: Network operation with high renewable penetration - Comparison with traffic routing

1.4.2 Demand response

Demand response can be defined as the set of techniques allowing to modify the electricity consumption pattern of end users [AES08]. It can be done using price signals, or load-shedding signals in case a (usually large) consumer has a specific contract with a regulating authority. Several electricity markets have created programs to use loads to provide ancillary services, which usually focus on large industrial sites. The aggregated flexibility brought by the simultaneous control of small residential loads could however become more valuable [Cal11], but the small amount of flexible consumption in residential households usually does not justify a direct participation in electricity reserves markets. However, new actors are emerging. Among them, aggregators aim at gathering a large number of small flexible loads offered by consumers and value the obtained aggregated flexibility on a specific market, to provide ancillary services for instance [CH10; Mat+12]. Virtual energy storage systems which can be leveraged to provide ancillary services can be obtained by controlling a fleet of electrical vehicles [TRY16], whose share is expected to increase significantly in the next few years [Gna+18]. It has also been demonstrated that controlling an aggregation of Thermostatically Controlled Loads (TCLs), which include fridges, Air Conditioning (AC), heat pumps, water heaters, pool pumps,... may have significant value [Mat+12] to contribute to frequency reserves mechanisms. The specificity of these devices is that they aim at maintaining a certain temperature within an acceptable range, and due to thermal inertia, their consumption can be modulated without jeopardizing quality of service. It is shown in [Cam+18b] that the control of a large number of TCLs allows to replicate a large virtual battery, but at lower cost.

1.4.3 Leverage flexibility from multi-generation and combined operation of multi-energy networks

Other energy networks, like gas networks [Koc+15] or heat/cold networks could be leveraged in order to bring additional flexibilities to electricity networks, required in a context of uncertain renewable production. See for instance [Ord+17] for a combined market design for electricity and gas. Optimal planning of energy networks expansion is studied in [UV+10] for electricity and gas, [VBGS15; VB+17] for general multi-energy networks, to decide when to expand existing networks, when to install conversion devices (power-to-gas and power-to-heat) or storage systems. Combined operation of distributed multi-generation units, such as co-generation units, is also a promising lead [CM09a; CM09b] to bring additional recourse to uncertain systems. Other references and approaches are further discussed in the subsequent Section 1.5.

1.4.4 Local controls for network operation

Distribution Network Operators are installing local control loops and power electronics to deal with network constraints violations. They include components able to help maintain voltage stability, such as voltage regulation transformers (VRT), static var compensators (SVC), static synchronous compensators (STATCOM), and shunt capacitor banks [Liu+17]. Energy storage systems can be used to avoid line congestion phenomenon or backward flows from distribution network to transmission networks, and therefore help renewable energy sources integration [Das+18]. The comparison of network regulation with traffic regulation can be further developed by comparing these local control loops to signalization or traffic lights, see Figure 1.6[6].

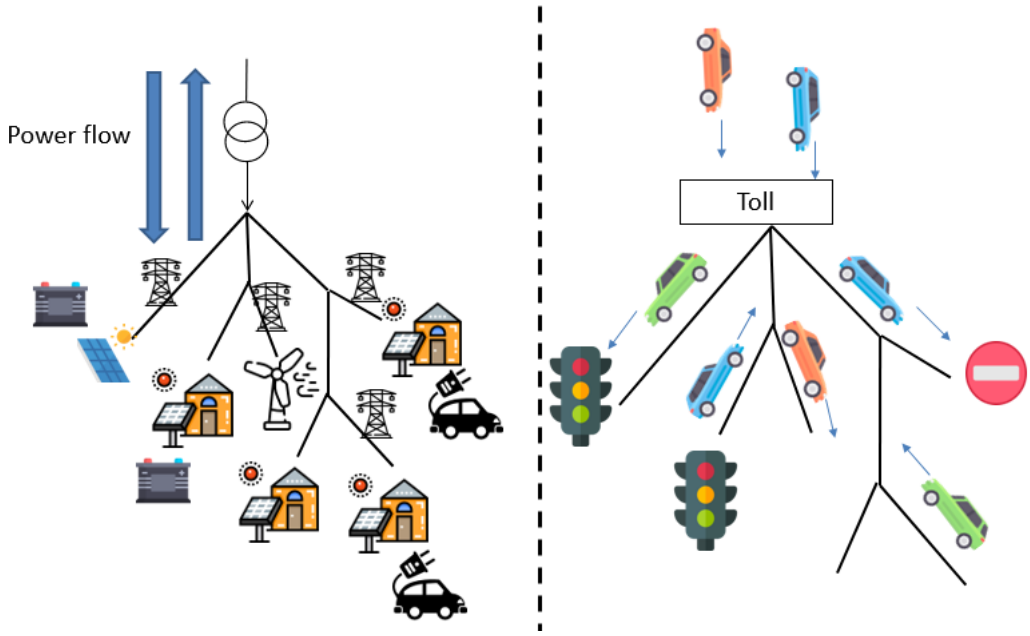


Figure 1.6: Smart network operation with high renewable penetration - Comparison with traffic routing

1.5 Challenges in energy management

Integration of renewable energy sources can be enhanced using market designs, which give incentives to micro-grids to reduce the impact of their renewable production units by controlling storage devices. The optimization of such micro-grids amounts to control in an optimal way one or several storage systems in an uncertain environment, brought by the stochastic renewable production and the random market behavior. For optimal control problems of many distributed storage systems providing ancillary services (Demand response), a curse of dimensionality phenomenon may rapidly occur, when the number of state variables grows. Besides, some particular requirements are needed for practical implementation. First, the privacy of data of individual consumers need to be guaranteed. Second, energy storage systems controllers should be local, to maintain quality of service in case of telecommunication disruption. Third, due to bandwidth and privacy constraints, optimization should be done with limited communication between agents.

The intermittent energy sources and their stochastic behaviors can also create physical problems for energy networks. Taking into account physical constraints of energy networks makes the related optimization problems non-convex. Besides, multi-stage stochastic versions of the problems are necessary to account for stochastic production of renewable energy sources and dynamical constraints of storage systems, as they guarantee that decisions are not taken according to yet unobserved data. Such models are therefore difficult for two reasons: they are large-scale (as they are based on a scenario tree which is required to grow exponentially fast with the

[6] Icons made by Freepik and Smashicons from www.flaticon.com

time horizon to properly model the filtration [Sha06; HRS06; PP14]) and non-convex. Adequate approximations (linearizations) and/or convexification methods or local search methods can be used to deal with the non-convexity of such problems.

As combined optimal control of multi-energy networks is not further developed in this thesis (only electricity networks will be considered), we now focus on the challenges, as well promising modeling and solution approaches for this type of problems. A general framework for optimization of operations of multi-energy networks taking into account the physical laws of such networks (Optimal Power Flow) has been introduced in [GA05; GA07]. An inherent difficulty for such problems is that they physical laws of energy networks typically yield non-convex feasible regions. Combined with a stochastic framework, this yields difficult optimization problems. To solve such problems, non-linear programming approaches can be used. For instance, Chance-Constraints Programming models for optimization of gas networks are considered in [GHH17; Hei19], solved with descent methods. Chance-Constraints models for combined gas and electricity networks have also been considered [Ode+18]. Another category of approaches are convexification methods, which can provide optimality certificates, unlike non-linear programming approaches, and which are more precise than linearizations, taking into account non-linear effects such as pressure drops, thermal losses in pipelines (which are impacted by the mass flow). Convexification methods first appeared in the framework of electricity networks [LL11; Low14a; Low14b; MH+19], but recent works focus on such methods for other energy networks. With this respect, polynomial optimization (exploiting sparsity structure of networks), first used in the context of electricity networks [MH14; Mol+15; Jos16; Mol+16], has recently been employed to control other energy networks in the PhD dissertation [Hoh18]. These methods allow to consider Semi-Definite relaxations of the problems, which can be solved using interior points methods and are sufficiently tractable (using the sparsity pattern of such networks) to consider stochastic variants of the problems. For instance, a two-stage stochastic framework is considered to control a district heating network in [HWL19], solved using polynomial optimization.

1.6 Objectives and contents of the thesis

In this PhD thesis, we focus on the impact of the uncertainty of the production of renewable energy sources on two problems in energy management. First, we focus on the optimal control of single and multiple micro-grids in interaction to reduce the impact of uncertainty of renewable production (without taking into account network constraints). Second, we consider the optimal planning and control of an electricity network in a stochastic context (taking into account the non-linear Kirchhoff's circuit laws). This PhD dissertation introduces modeling, algorithmic and theoretical contributions to tackle these two issues. Let us give more details regarding the two parts of this PhD dissertation.

- Part I introduces two control frameworks which allow to reduce the impact of the uncertainty of renewable production on electricity networks, without taking into account the non-linear laws of physics of electricity networks. The mathematical models proposed lie in the field of stochastic optimal control, with a mean-field component. We also introduce two numerical methods based on the stochastic Pontryagin principle in order to solve non linear-quadratic problems. In Chapter 3, we focus on a single micro-grid and assume that it commits to a consumption profile on the network, say, for the next day, and controls an energy storage system to track this commitment in real-time, in order to reduce the volatility of its consumption on the network. The problem of optimal choice for the commitment profile (chosen before realization of randomness) and the control policy of the energy storage system in real-time is formulated as a McKean-Vlasov (MFV) stochastic control (also called Mean-Field Control (MFC)) problem, for which necessary and sufficient optimality conditions are derived, and numerical methods are designed to solve the problem. A decentralized control mechanism of a large population of Thermostatically Controlled Loads (TCLs) is introduced in 4 where many consumers control their individual storage systems in a cooperative way so as to provide ancillary services. An appropriate decomposition of the optimality system as well as a mean-field approximation are designed in order to tackle the curse of dimensionality issue and to obtain a decentralized control architecture ensuring privacy of the agents and minimal information sharing. Extending this framework to non-linear quadratic setting motivates a new theoretical and numerical method for solving stochastic control problems, which is a Newton method adapted to this setting, presented in Chapter 5. In particular, we show that the computation of the Newton step

reduces to solving Backward Stochastic Differential Equations, we design an appropriate line-search method and prove global convergence of the resulting Newton method with line-search in an appropriate space of stochastic processes. We illustrate the performance of this algorithm on a non-linear quadratic version of the problem presented in Chapter 4. In particular, we propose a fully implementable algorithm, based on regression techniques.

- Part II focuses on the optimal planning and operations of electricity network operations in a random environment. Chapter 6 gives an introduction to the Optimal Power Flow (OPF) problem, which models the optimization of electricity network operations taking into account the physical laws of such systems. It lists state-of-the-art results in the literature for this problem. In Chapter 7, we formulate a multi-stage stochastic extension of the Alternating Current Optimal Power Flow problem. This multi-stage setting is necessary to account for the uncertain production of renewable energy sources, temporal correlations in their production levels and the dynamical constraints of energy storage systems. We also consider conic relaxations [Low14a; Low14b; MH+19] of this non-convex problem (shown to be strongly NP-hard in [BV19]) which allow a good trade-off between tractability and accuracy. We give tractable and realistic a priori conditions ensuring a vanishing relaxation gap for the OPF problem in the multi-stage stochastic setting. We also provide an easily computable a posteriori bound on the relaxation gap of the multi-stage OPF problem. This yields theoretical insight on the conditions under which convex relaxations of the non-convex AC OPF problem are reliable. The proof of these results relies on primal reconstruction arguments inspired by [Hua+16], i.e., an algorithm which takes a solution of a convex relaxation as input and returns a solution of the non-convex problem. Using Shapley-Folkman-type results, we also obtain explicit bounds on the duality gap of adequate formulations of multi-stage stochastic non-convex problems in energy management. This shows that energy storage systems not only allow to reduce costs, but, in some cases, can also enhance the numerical tractability of such non-convex multi-stage stochastic problems.

1.7 Contributions of the chapters

We now give more details on the contributions of the individual chapters of this PhD dissertation.

1. **Chapter 3** introduces a control problem for a single micro-grid, where a consumer possibly equipped with photo-voltaic panels, a storage system and connected to the grid aims at reducing the day-ahead uncertainty of its consumption on the network, by committing in advance to a consumption profile. The consumer then controls its energy storage system in real time to minimize deviations of its consumption from its commitment and operational costs, see Figure 1.7[7]. We model this situation by a Mean-Field Control problem, with scalar interactions and in general filtrations. This allows the joint optimization of decisions taken at different times: the commitment profile is chosen initially and impacts the system on the whole time horizon, whereas the energy storage unit is controlled according to random data revealed along time. Necessary and sufficient optimality conditions are derived in the form of a Mean-Field Forward-Backward Stochastic Differential Equation (MF-FBSDE) see Theorems 3.2.3 and 3.2.10, and we study of solvability of the optimality system obtained. A numerical method to approximately solve problems which do not have the classical Linear-Quadratic structure is proposed, based on perturbation theory, see Theorem 3.3.7. Extensive numerical illustrations of the mechanism are provided.

2. In **Chapter 4**, motivated by the huge potential of Thermostatically Controlled Loads (TCLs) to provide ancillary services [Mat+12; Cam+18b], we propose a stochastic control problem to control cooperatively such devices to promote power balance. We develop a method to solve this stochastic control problem with a decentralized architecture, in order to respect privacy of individual users and to reduce both the telecommunications and the computational burden compared to the setting of an omniscient central planner, see Figure 1.8[8]. The optimality conditions are expressed in the form of a high-dimensional Forward-Backward Stochastic Differential Equation

[7] Icons made by Freepik and Smashicons from www.flaticon.com
[8] Icons made by Freepik and Smashicons from www.flaticon.com

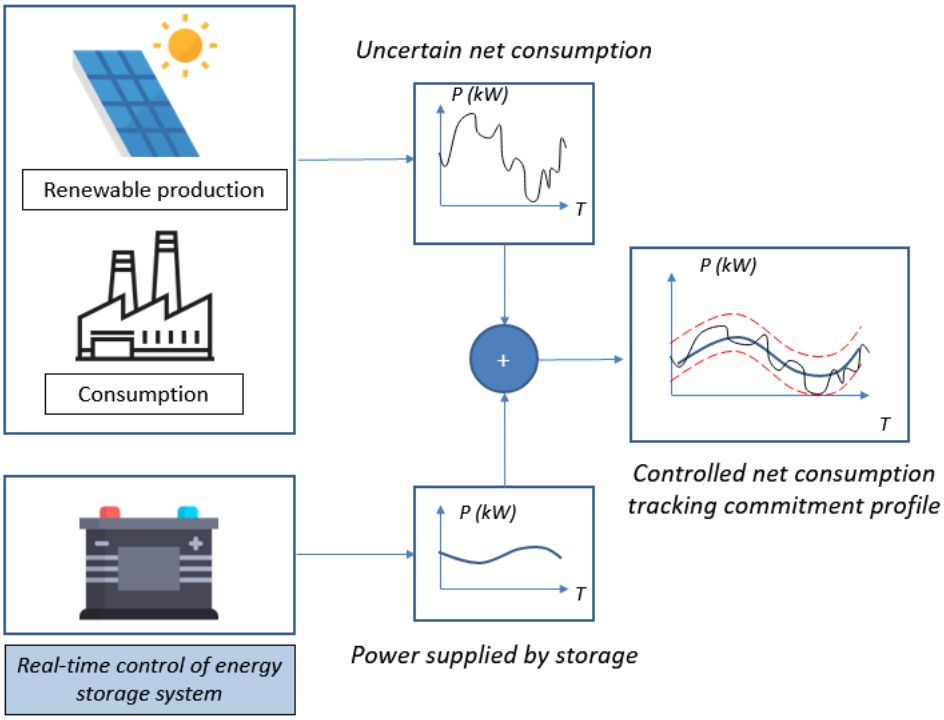


Figure 1.7: Day-ahead commitment with real-time tracking of the commitment profile

(FBSDE) (Theorem 4.2.1), which is decomposed into a coordination problem, which is a low-dimensional FBSDE modeling the optimal behaviors of the aggregate population of TCLs, and individual problems which are one-dimensional FBSDEs modeling the optimal control of individual devices. Existence and uniqueness results are derived for the solutions of these FBSDEs. We also show that these FBSDEs fully characterize the (unique) Nash equilibrium of a stochastic Stackelberg differential game. In this game, a coordinator (the leader) aims at controlling the aggregate behavior of the population, by sending appropriate signals, and agents (the followers) respond to this signal by optimizing their storage system locally. This allows a decentralized implementation of the optimal control from the point of view of the omniscient social planner. This is the reverse perspective of potential Mean-Field Games [FMHL19], which aim to find (convex) control problems with optimality system coinciding with the Nash system of a given Stochastic Differential game. A mean-field-type approximation of the coordination problem is proposed in which parameters do not depend on the real-time aggregated behaviors of the agents, so that the approximate problem can be solved observing only the common noise, see Theorem 4.4.5. This allows to circumvent telecommunication constraints and privacy issues. This approximation is based on the conditional Law of Large Numbers. Convergence results and error bounds are obtained for this approximation depending on the size of the population of TCLs, see Theorem 4.4.12. A numerical illustration is provided to show the interest of the control scheme and to exhibit the convergence of the approximation. An implementation which answers practical industrial challenges to deploy such a scheme is presented and discussed, see Algorithm 4.1.

3. In **chapitre 5**, we develop a new (theoretical and numerical) method based on stochastic Pontryagin principle to solve stochastic control problems. This method is nothing else than the Newton method extended to the framework of stochastic control in general filtrations, where the state dynamics is given by an ODE with stochastic coefficients. In particular, we show that the Newton method amounts to solve successively Linear-Quadratic approximations of the control problem, see Proposition 5.3.2, or equivalently, to perform successive linearizations from the optimality conditions obtained by the stochastic Pontryagin principle, see Proposition 5.3.3. A full methodology is proposed to compute theoretically the Newton step, which can be obtained by solving an affine-linear FBSDE with random coefficients see Theorem 5.2.9 and Definition 5.3.1. We then

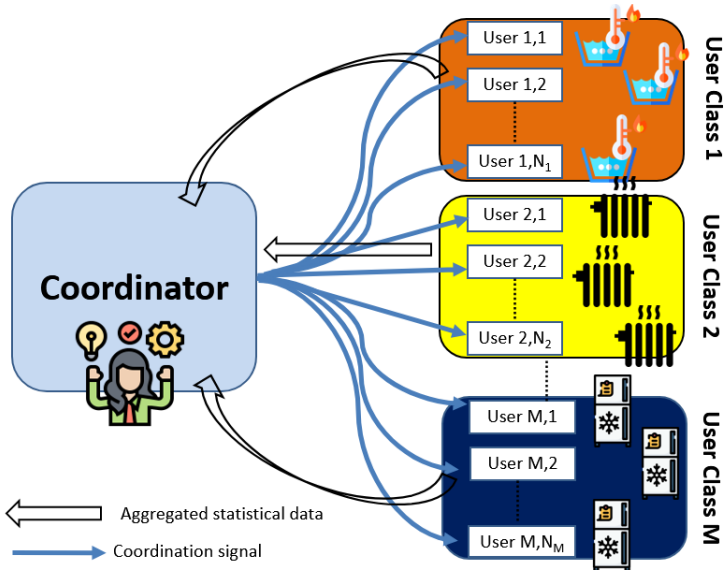


Figure 1.8: Decentralized control mechanism

show that solving such an FBSDE reduces to solving a Riccati Backward Stochastic Differential Equation (BSDE) and an affine-linear BSDE in Theorem 5.3.7, as expected in the framework of linear FBSDEs [Yon06] or Linear-Quadratic stochastic control problems [Bis76]. We then focus on convergence results for the Newton method. We show by a counter-example (Example 5.3.9) that sufficient regularity conditions ensuring convergence of the Newton method (like a Lipschitz-continuous second-order derivative [Kan48; NW06]) of the problem may not hold in our infinite-dimensional setting. To alleviate this issue, the problem is restricted to essentially bounded processes. We give another counter-example (Example 5.3.12) which shows that even for strongly convex problems with high regularity, global convergence of the method may fail. Then an adapted backtracking line-search method (different from the standard backtracking line-search presented in [NW06; BV04]) is developed to fit our infinite-dimensional setting and to obtain a proof of global convergence with asymptotic quadratic convergence of the Newton method with the line-search, see Theorem 5.3.15. An implementation with regression techniques to solve BSDEs arising in the computation of the Newton step is developed and applied to the control problem of a large number of batteries providing ancillary services to the grid.

4. **Chapter 6** presents a general introduction to an important problem in energy management, the Optimal Power Flow problem, which aims at controlling optimally an electricity network taking into account the non-linear Kirchhoff's circuit laws in the context of oscillating regimes (Alternating Current). The most common formulations of this problem as well as formulations used in the following chapter are given. Very little prior knowledge from the reader is assumed and this chapter aims to be as pedagogical as possible while giving pointers to recent state-of-the-art techniques, approximations, relaxations and extensions of the OPF problem available in the literature.

5. **Chapter 7** builds on the previous chapter. We formulate a multi-stage stochastic extension of the Alternating Current Optimal Power Flow problem, accounting for the random production of renewable energy sources and the dynamic constraints of storage system. We also give a conic relaxation of this non-convex problem. We then focus on developing estimates on duality gaps of the multi-stage stochastic Optimal Power Flow problem with storage, and more generally to other non-convex problems in energy management (like Unit Commitment). We first argue that a priori conditions guaranteeing no duality gap for the Alternating Current Optimal Power Flow (AC OPF) problem, see for instance [SL12; Gan+14; Hua+16], usually proved in a static and deterministic setting, can be extended to a multi-stage stochastic setting. In particular, we provide realistic and tractable a priori conditions guaranteeing a vanishing relaxation gap in the multi-stage stochastic framework, extending the results in [Hua+16]. We also give an easily computable a posteriori bound on the relaxation

gap of the multistage AC OPF problem. Using a different approach, we then derive a priori bounds on the duality gaps of multi-stage non-convex problems in energy management, where energy storage systems play a particular role, see Proposition 7.5.9 and Corollary 7.5.10. Such results are related to the Shapley-Folkman theorem, which gives a bound on the distance between the Minkowski sum of non-convex compact sets, and its convex envelop.

1.8 State-of-the-art

To compare our results with existing ones, we give an overview of methods in the literature for the two parts of this PhD dissertation.

1.8.1 Stochastic control of micro-grids (without network constraints modeling)

Optimization of a single micro-grid has been the focus of many recent works using very diverse approaches. Let us mention that the model presented in Chapter 3 lies in this stream of research, though taking into account a particularity: the joint optimization of a decision taken initially and impacting the system on the whole time horizon (consumption commitment) and of the control policy of a storage system in real-time.

- A popular approach for the control of energy storage systems in an uncertain setting is the Model Predictive Control. It amounts to treat stochastic optimization problems as deterministic ones, with the random parameters of the problem replaced by their forecasts. A detailed physical model of a battery and of Photo-Voltaic panels is considered in [SSM16] and a Model Predictive Control architecture is developed. However, the Model Predictive Control is a heuristic, and its performance can only be validated by simulation. This lack of optimality guarantee is investigated in [Pac+18], where Stochastic Dual Dynamic Programming and Model Predictive Control are used and compared to solve an optimization problem modeling the optimal control of a micro-grid. The flexibilities considered are a battery, a water heater, and the heating systems of a building.

- Considering a Markovian Decision Process setting and using Stochastic Dynamic Programming is also a popular approach when the dimension of the state variable is low, as is done in the case of optimal control of an electrical vehicle in a random environment [IMM14; Wu+16].

- Aging of a battery which help renewable energy sources integration is considered in [Hae14] and in [Car+19a]. The latter reference designs a specific temporal decomposition (dynamic programming) on two-different time scales to control a battery on a micro-grid while taking into account long-term aging.

- Deterministic and stochastic control problems in continuous time focusing on a single micro-grid are considered in [Hey+15; Hey+16], where the controller aims at optimizing simultaneously the operation of a diesel engine and a battery, in a deterministic [Hey+15] or stochastic [Hey+16] setting.

Decomposition methods have also been developed to handle several interacting micro-grids. A combined spatial and temporal decomposition of multistage stochastic problem is proposed in [Car+19b] and [Car+20], taking as application the optimal control of several individual micro-grids equipped with storage systems which can exchange energy through a power network. In particular, the method developed allows to consider problems for which the state dimension is much higher than what can be handled with standard methods like Stochastic Dynamic Programming or Stochastic Dual Dynamic Programming. The decomposition method relies on the formulation of sub-problems associated to the network nodes, each of them solved using dynamic programming. Note that in this application, a simple network flow model for the power network is considered. In particular, it does not take into account the oscillating nature of physical quantities on such networks, non-linear effects, such as thermal losses and reactive power. These physical characteristics of electricity networks are considered in the multistage stochastic Optimal Power Flow model of Chapter 7.

From the operational point of view, it is desirable to develop methods which do not suffer from the curse of dimensionality and allow to control many energy storage systems. Indeed, the large number of energy storage

systems, including batteries, batteries of Electric Vehicles, Thermostatically Controlled Loads (TCLs) (such as water heaters, heat pumps, air conditioners...) present on electricity networks can help provide the necessary ancillary services to integrated renewable energy sources. TCLs have been identified as particularly promising with this respect [Mat+12; Cam+18b]. However, optimal control problems of large numbers of such devices are challenging for several reasons:

1. The dimension of the state space corresponds to the number of controlled devices, which can be very large. Therefore, curse of dimensionality may rapidly occur when using methods such as Stochastic Dynamic Programming, Stochastic Dual Dynamic Programming,...

2. The controllers of individual energy storage systems should be situated locally. Indeed, a centralized control architecture would require a heavy telecommunication infrastructure, in order for the central planner to gather data of the individual agents and send control signals. Besides, the quality of service (i.e., the fact that temperature of the devices should lie in an admissible range) should be guaranteed but may not be ensured in case of loss of communication, if the controller is not local.

3. A particular focus should be made on the respect of privacy of individual data of the consumers, when implementing such control schemes.

Let us mention that the model proposed in Chapter 4 and the application illustrating the numerical method introduced in Chapter 5 answer these specific issues. Let us give an overview of methods used in the literature for the decentralized management of many local energy flexibilities.

A first stream of research includes cooperative control architectures which focus on practical implementation given the specific constraints we listed above, without a priori optimality guarantees. Local controllers which respond to a coordination signal designed using Partial Differential Equations-based models are considered in [TTS15; Tro+16]. A mean-field approximation of the large number of individual TCLs is considered in [BM16; Cam+18a] and local control loops are implemented so that the aggregated flexibility of the TCLs behaves like a virtual battery. For those reference, the (random) switching rate of individual devices plays the role of control variable, and no a priori optimality guarantee can be established.

Game-theoretic approaches are also considered as a convenient way to obtain a decentralized control architecture, with self-interested agents responding to price incentives. An hourly billing mechanism is considered in [Jac+18] to account for the fact that a high aggregated demand may have a negative impact on electricity networks. A game theoretic setting with heterogeneous players is considered in [JW18]. A cryptography technique is used in [Jac+19] to account for the privacy requirements for individual data when an aggregator only needs to estimate aggregated energy flexibility of several users.

Mean-field approximations are very appealing as they become more precise when the number of agents grows. Mean-Field Game models have recently drawn a lot of interest to model energy management problems [DP+19; ATM20; MMS19; KM13; KM16]. In [DP+19], individual TCLs respond to price signals determined by solving a Unit-Commitment problem to optimally allocate their flexible consumption and their participation to frequency reserves mechanisms. The problem is solved using a fixed-point method based on the Hamilton-Jacobi-Bellman and the Fokker-Plank equation, a coupled PDE system modeling the dynamics of the value function and of the probability measure of the states of the population. No a priori convergence guarantee for this method is given. In [ATM20], consumers use batteries to reduce their electricity bills, the spot price of electricity being determined by the aggregated demands of all consumers. The Nash equilibrium of this Mean-Field Game with common noise (representing geographical correlations of the weather, inducing correlations in renewable production) is found using Pontryagin principle, which yields explicit feedback formulas in the linear-quadratic case. The model of [ATM20], which considers only Brownian filtrations, is extended to a setting with jump processes in [MMS19]. A Mean-Field Game model without common noise is proposed in [KM13], where the mean temperature of a large number of TCLs is required to follow a specific profile. The model is Linear-Quadratic Gaussian, so that explicit feedback formulas are available for the solution of the Nash equilibrium. This model is extended to Markovian jump processes in [KM16], considering finer models of individual water heaters: the stratification phenomenon (i.e., non-uniformity of the temperature within a water heater) is taken into account. Note that Game-theoretic and Mean-Field Games models provide a practical way to obtain a decentralized control architecture, but yield a Nash equilibrium which may be sub-optimal from a

collective point of view. Bounds on the price of Anarchy can be derived [Jac+18]. In Linear-Quadratic frameworks and under some particular assumptions, Mean-Field Control problems (resp. Mean Field Games) have a unique optimum (resp. Nash equilibrium), which can be computed analytically. This gives as a by-product the price of anarchy. A Mean-Field type approximation combined with price (Lagrangian) decomposition techniques is proposed in the recent work [Seg+20] to solve stochastic control problems modeling aggregation problems arising in energy management.

1.8.2 State-of-the-art for optimal control of electricity networks with network constraints modeling

The problem of optimal control of an electricity network taking into account the non-linear Kirchhoff's circuit laws is called Optimal Power Flow. This problem is at the boundary between mathematical optimization, physics and electrical engineering. For this reason, we dedicate a specific chapter to properly introduce the OPF problem, see Chapter 6. We also give many pointers in the literature regarding theoretical results, mathematical formulations, approximations, convex relaxations, optimization methods, stochastic and dynamic extensions related to the OPF problem. We only summarize the main results in the literature regarding the OPF problem here.

Definition

The so-called Optimal Power Flow (OPF) problem is a mathematical optimization problem aiming at finding an operating point of a power network (taking into account the Kirchhoff's circuit laws, which yield the so-called load-flow equations) that minimizes a given objective function, such as generation costs, active power losses, subject to constraints on power injections and losses, voltage magnitudes and intensities in the lines. It has been introduced by Carpentier in 1962 ([Car62]) and has drawn a lot of attention since then thanks to its versatility.

Physically, the most accurate formulation of this problem is the so-called Alternating Current Optimal Power Flow (AC OPF). It is a non-convex quadratically constrained problem which has a natural formulation in complex variables. The complex variables allow a compact representation of oscillating signals, in the Alternating Current framework, which is the type of current used for most networks.

Single phase vs. multi-phase, radial vs. meshed networks

Note that there are three common types of networks: single phase, balanced three-phase and unbalanced three-phase networks. Single phase networks are composed of lines with two wires: a ground wire, and a wire where power can flow. The three-phase balanced and unbalanced networks are composed of four wires: a ground wire, and three wires carrying power and associated to a phase. In the case of balanced networks, the sizes of all phases are the same and they are differ from each other by 120 °, which is not the case for unbalanced networks. The AC OPF problem for three-phase balanced networks reduces to the AC OPF problem with single phase. Besides, networks can be either radial (acyclic) or meshed. Typically, transmission networks are meshed and operated with a single phase. Medium-voltage distribution networks are typically radial and balanced three-phase, hence the problem for such networks reduces to the case of single phase radial networks. Low-voltage distribution networks are typically operated radially and unbalanced three-phase. **We shall focus on the case of single phase networks**, which allows to consider balanced three phase networks. Some results for unbalanced multi-phase networks can be found in Chapter 6.

Methods to solve the AC-OPF problem for single phase networks in a static deterministic setting

The AC OPF problem has been recently shown to be strongly NP-hard [BV19]. To handle this difficulty, several methods can be considered, like Non-Linear Programming methods, linearizations of the problem or convex relaxations.

Non-Linear Programming methods include Newton Raphson [IS87], Sequential Quadratic Programming, steepest descent methods [For+10]. These optimization techniques and many more can be found in the very complete

surveys [FSR12a; FSR12b]. These methods are often sensitive to the initial point and may fail to converge to a global optimum, which can create substantial costs for network operators.

Linear approximations are often used in practice due to their tractability for many applications. However, when the share of renewable energy sources increases, they may neglect important phenomena, like variations of voltage magnitude [NT15]. Among the existing linear models, the most famous one is the Direct Current Optimal Power Flow (DC OPF) formulation, which assumes fixed voltage magnitudes, neglects power losses and reactive power flows. Methodologies to obtain error bounds between DC and AC OPF have been developed in [SJA09; DM16]. There exists many other linearizations, like the Linearized Distflow model, especially adapted for radial (i.e, acyclic) networks, or linearizations of the AC OPF model around a nominal setpoint.

Convex (conic) relaxations of the AC OPF problem can be efficiently solved using interior point methods and provide a certificate of optimality when they find an optimal solution which is feasible for the original non-convex problem. Among them, the Semi-Definite relaxation, Second-Order Cone relaxation and the chordal relaxation are the most famous ones. They are presented in [Low14a; Low14b; MH+19]. These relaxations exhibit low or no relaxation gap for many practical instances [LL11]. Conditions ensuring exactness of these relaxations have also been established in [LL11; SL12; FL13; Gan+14; Hua+16] but a gap remains between the theoretical conditions ensuring zero-duality gap and reality, as many instances of realistic problems have zero-duality gap without fitting assumptions of these works. The Semi-Definite (SD) relaxation typically has a lower relaxation gap than Second-Order Cone (SOC) relaxation in meshed network (i.e., with cycles) but at the expenses of higher computational burden, as it requires to work with matrices with number of entries growing quadratically with the number of buses of the network. For radial networks, both SD and SOC relaxations have the same relaxation gap, so that one should use the SOC relaxation due to its better numerical performance. Intermediate relaxations have been proposed to solve with good precision and reasonable computational burden large-scale instances of AC OPF with meshed networks, such as "partial" Semi-Definite Relaxations [BAD18], strong Second-Order relaxations [KDS16]. In case the Semi-Definite relaxation is not exact, penalization methods [MAL15] and sparse polynomial optimization methods [MH14] or combinations of both methods [Mol+15; Mol+16; Jos16] can be applied in order to solve large-scale instances for which the SD relaxation is inexact.

Extensions to dynamic and stochastic settings

A dynamic AC OPF model with energy storage system is formulated in [GKA13], using only the non-convex formulation and without considering uncertainty. The SOC relaxation is used in a dynamical setting in [GSGK18] in order to optimally size distribution networks. Conditions found in [LL11] under which the SD relaxation is exact are extended in [GT12] to the deterministic multi-period case with storage. A survey of OPF methods and tools with application to distribution networks with energy storage systems can be found in [SM16].

The most common stochastic models are the probabilistic OPF problems. A survey on such models can be found in [BCH14]. Some recent works propose a probabilistic version of the DC OPF model [Roa+13; Roa+16], while other focus on probabilistic versions of the linearized AC OPF model around a reference scenario [RMT17; RA17]. Other works consider the robust counterpart of the SD relaxation restricted to affine-linear decision-rules [Vra+13]. A Semi-Definite convex relaxation of the chance-constrained AC OPF problem is proposed in [Ven+17] using a scenario-based approach and assuming piece-wise linear decision rules, or assuming Gaussian uncertainty. A Second-Order Cone approximation of the chance constrained AC OPF, combined with a feasibility recovery method, is proposed in [HPC18], which allows good numerical performances. Another class of stochastic models are the multistage models, for which it is required that decision variables remain non-anticipative, i.e., do not depend on yet unobserved random data. References [Swa17; NCP14] consider the simplified case where decision are non-anticipative for the initial time steps only, as they consider a scenario tree with a comb structure (two-stage). Non-anticipativity is guaranteed in [JKK14] using affine-linear policies but with a DC OPF problem, hence inheriting the limitations of such approximate models. Non-anticipativity of the decisions is also guaranteed in [Sun+16] which considers an iterative heuristic procedure to optimize a decision policy.

Chapter 2

Introduction (en français)

2.1 Changement climatique et développement des énergies renouvelables

La majorité des pays industrialisés prennent de plus en plus conscience du changement climatique et soutiennent de manière active des solutions visant à réduire les émissions de gaz à effet de serre. Des efforts importants sont prévus afin de limiter la hausse de température à 2° C d'ici 2100, en accord avec la cible des accords de Paris de 2015. Les moyens de production d'électricité conventionnels comme les centrales à charbon, au fuel ou au gaz comptent parmi les principales sources de gaz à effet de serre. Certaines études montrent qu'ils représentent environ 25% des émissions de CO2 mondiale [Cha+14].

Afin de réduire l'empreinte carbone de la production électrique, les sources d'énergie renouvelable, tels les panneaux solaires et les éoliennes, constituent des solutions viables. Récemment, la Commission Européenne a établi un objectif de 32% des sources d'énergie renouvelable dans le mix énergétique d'ici 2030 [Com19]. Les activités de recherche sur ces moyens de production ont permis de réduire considérablement leurs coûts de production. Par exemple, le prix des cellules photovoltaïques a baissé de 99.6 % entre 1976 et 2018, passant de $79.3/W à $0.3/W [WB19]. Cependant, les sources d'énergie renouvelable précédemment mentionnées présentent des inconvénients majeurs qui nécessitent un traitement particulier lors de leur intégration dans le mix énergétique. Elles ne sont pas contrôlables, elles sont intermittentes [Suc+20] et leur niveau production est aléatoire, du fait d'erreurs possibles dans les prévisions météorologiques. Cela pose des problématiques complexes en termes de gestion de l'énergie. Pour les comprendre, nous présentons tout d'abord le contexte historique de la gestion de l'énergie.

2.2 Contexte historique en gestion de l'énergie

2.2.1 Spécificités historiques de la production, de la consommation et du transport de l'électricité

Aujourd'hui encore, la majorité de la production électrique est réalisée de manière centralisée par des centrales conventionnelles et pilotables, comme des centrales à charbon, au gaz, au fuel ou nucléaires.

L'électricité produite par ces centrales est envoyée aux consommateurs finaux par des câbles électriques du réseau de transport puis de réseaux de distributions. Le réseau de transport est un réseau opéré à haute tension (225 à 400 kV) avec une architecture maillée (i.e., avec des cycles), avec une seule phase. Ce réseau est opéré par Réseau de Transport d'Électricité (RTE) en France, qui est un monopole régulé. Les réseaux de distribution sont quant à eux des réseaux opérés à basse (de 220 à 430 V) ou moyenne tension (1 kV à 52 kV). Ils sont le plus souvent radiaux (i.e., sans cycles) triphasés et pilotés par Enedis en France, qui est un autre monopole régulé. L'architecture historique des réseaux électriques est représentée sur la Figure 2.1[1].

[1] Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

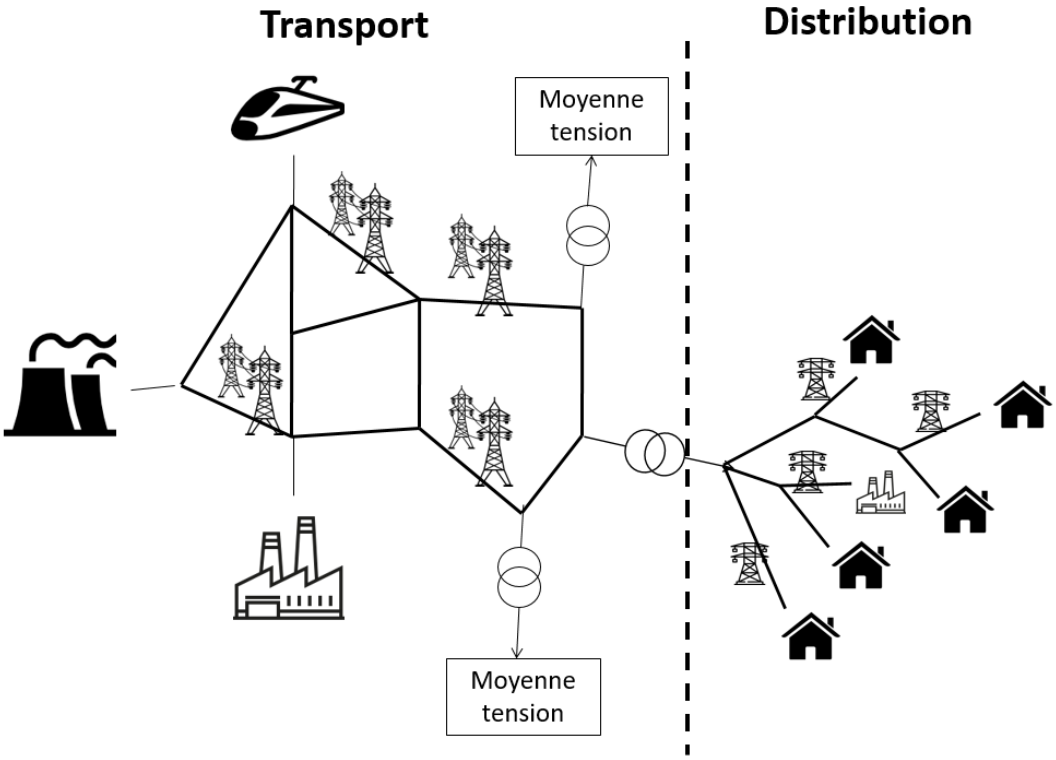


Figure 2.1: Structure historique des réseaux électriques

2.2.2 Équilibre offre-demande

Généralités

Une particularité de l'électricité est que, contrairement à d'autres biens, elle ne peut être stockée que marginalement, du fait des coûts élevés du stockage. En France, les Stations de Transfert d'Énergie par Pompage (STEP) constituent le principal levier de stockage de l'énergie actuellement. A cause de leurs coûts d'installation élevés et de leurs besoins, les STEPs ne peuvent absorber que 3% de la consommation électrique annuelle et permettent de déplacer environ 1.5% de la demande nationale annuelle.

Pour cette raison, l'équilibre offre-demande pour la puissance électrique doit être garanti à chaque instant. De nombreux acteurs en sont responsables: fournisseurs, producteurs, traders, agrégateurs. Tous ces acteurs doivent s'assurer de l'équilibre entre offre et demande au niveau de leur périmètre. Par exemple, EDF, en tant que producteur et fournisseur d'électricité, est responsable de l'équilibre entre sa production (incluant la production par ses centrales et ses achats sur les marchés) et la consommation de ses clients. EDF construit les plannings de production de ses centrales et achète ou vend de l'énergie sur les marchés afin de garantir l'équilibre entre la production prévue et la demande prévisionnelle. En d'autres termes, l'équilibre offre-demande est traditionnellement obtenu par pilotage de la production.

Prévision de la demande

La consommation quotidienne d'électricité dépend de plusieurs facteurs, comme le type de jour (semaine, week-end, férié), des évènements spécifiques (compétitions sportives) et les conditions météorologiques, comme la température ou l'ensoleillement. La température est l'un des principaux facteurs explicatifs de la consommation [PMV02], car la demande dépend fortement du chauffage et de la climatisation des consommateurs. Cependant, les bâtiments ont une inertie thermique et les consommateurs sont indifférents aux petites variations de température de leur environnement. Pour cette raison, la demande peut être prévue de manière assez précise en utilisant une

version lissée de la température (moyenne temporelle) [Jov+15]. Ainsi, les variations à court-terme des conditions météorologiques jouent un rôle moins important que leurs moyennes temporelles (sur quelques heures ou jours), et ces dernières peuvent être prévues avec davantage de précision que les conditions instantanées. Ceci explique que la demande peut être prévue de manière relativement précise à court-terme, dans un contexte de faible pénétration des énergies renouvelables.

Gestion de la production

Afin de réduire les coûts de production et les pénalités dues au déséquilibre offre-demande possible au niveau de leur périmètre, les producteurs doivent ajuster leur planning de production à la demande prévue. Ceci est fait à plusieurs échelle de temps.

- De 20 à 5 ans avant l'échéance, les décisions d'investissement dans de nouveaux moyens de production ainsi que dans les réseaux sont prises, et des partenariats stratégiques sont élaborés.
- De 5 ans à deux semaines avant l'échéance, les plannings de maintenance et de fonctionnement des centrales nucléaires et de barrages hydrauliques sont conçus. Les ressources (comme l'uranium dans les centrales nucléaires ou l'eau dans les barrages) sont allouées de manière optimale.
- De 2 semaines à un jour avant l'échéance, des prévisions plus précises de la demande sont disponibles, et les plannings des unités de production les plus flexibles sont ajustés.
- Un jour avant l'échéance, les prévisions de demande des clients et les plannings de production sont communiqués au régulateur (RTE en France).
- Le jour de l'échéance, des ajustements de dernière minute sont effectués, en recourant à des contrats bilatéraux ou en modifiant le point de fonctionnement d'unités de production flexibles.
- En temps réel, l'équilibre offre-demande est assuré par RTE à l'échelle du réseau de transport en activant des mécanismes de réserve.

De l'énergie est achetée ou vendue sur les marchés à toutes les échéances. Le niveau d'incertitude pris en compte ainsi que le réalisme physique de ces modèles est représenté en fonction de l'horizon de planification sur la Figure 2.2[2].

Mécanismes de réserve

Des évènements imprévus ou des erreurs de prévisions peuvent survenir et impacter l'offre ou la demande, ce qui peut perturber l'équilibre offre-demande. Des corrections en temps réel doivent donc être réalisées. En France, ces ajustements sont réalisés par RTE et sont basés sur 3 mécanismes de contrôles appelés réserves de fréquence.

1. La réserve primaire (activée automatiquement 1 seconde après un évènement spécifique et maintenue jusqu'à 30 secondes) a pour but de stopper la déviation de la fréquence du réseau de sa valeur nominale (50 Hz en Europe).
2. La réserve secondaire (activée automatiquement 30 secondes après un évènement et maintenue jusqu'à 30 minutes) permet de faire revenir la fréquence à sa valeur nominale.
3. La réserve tertiaire (activée manuellement à partir de 30 minutes après un évènement) consiste à modifier le point de fonctionnement de certaines unités de production afin de restaurer les capacités du système électrique à fournir des services de réserve primaire ou secondaire.

Les mécanismes de réserve sont à la main de RTE et sont principalement fournis par les unités de production, et dans une moindre mesure par de l'effacement de charge.

[2] Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

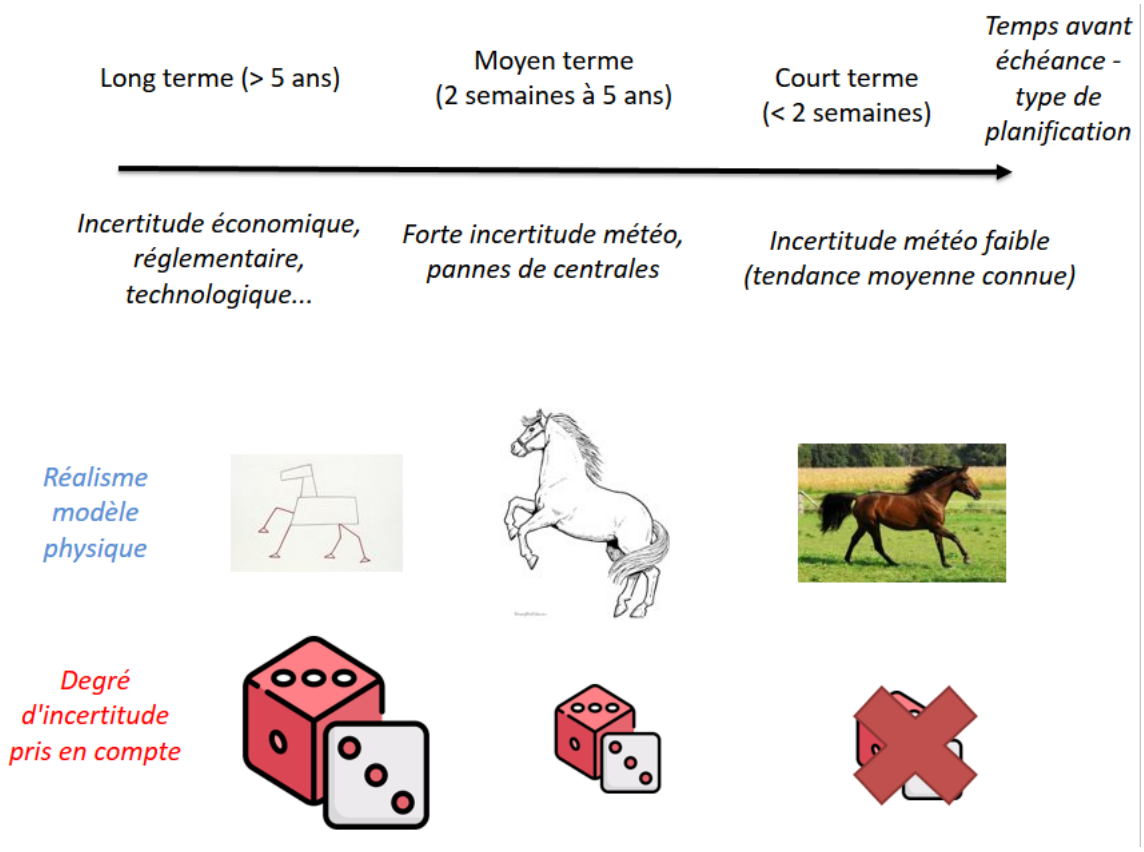


Figure 2.2: Structure historique des modèles de décision pour la gestion de l'énergie

2.2.3 Réseaux d'électricité

Les grandeurs physiques caractérisant l'état d'un réseau (amplitude de la tension aux nœuds ou de l'intensité dans les lignes du réseau) doivent être maintenues dans des plages opérationnelles acceptables. La puissance circule traditionnellement de grosses unités de production centralisées vers les consommateurs et est donc historiquement unidirectionnelle dans les réseaux de distribution (qui sont opérés radialement). Par conséquent, la tension aux nœuds du réseau est décroissante par rapport à la distance à la sous-station (point de connexion entre le réseau de distribution et le réseau de transport). Ainsi, un moyen simple de corriger la tension dans le réseau est de modifier la tension à la sous-station, qui est la racine de l'arbre formée par le réseau, par l'intermédiaire d'un transformateur. On peut effectuer un parallèle avec la régulation du trafic routier: lorsque toutes les routes sont en sens unique dans un réseau routier avec une structure d'arbre, il suffit de contrôler les arrivées de véhicules à la racine pour décongestionner les lignes du réseau. Cette métaphore est représentée Figure 2.3[3].

2.3 Changement de paradigme lié à l'essor des renouvelables

2.3.1 Équilibre offre-demande

Généralités

L'intermittence des sources d'énergie renouvelable comme les panneaux solaires ou les éoliennes peut provoquer des variations rapides de puissance dans le réseau [Suc+20]. Cela rend l'équilibre offre-demande plus complexe, car davantage de capacités doivent être allouées aux mécanismes de réserve, qui doivent en outre suivre des

[3] Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

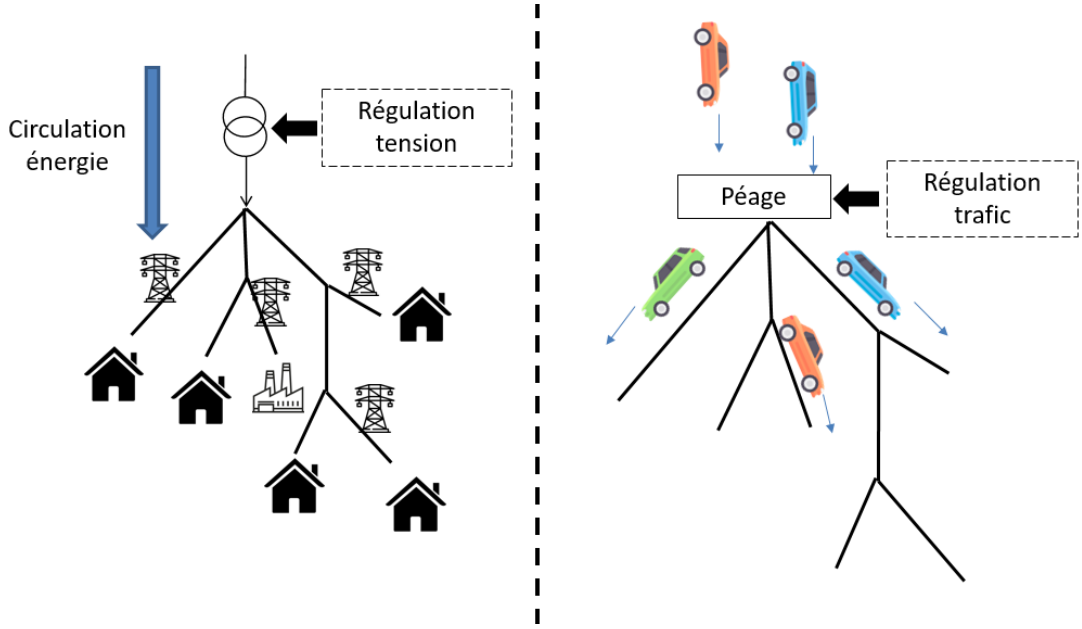


Figure 2.3: Régulation historique du réseau - Comparaison avec le trafic routier

variations de charges rapides. D'autre part, l'incertitude de production de ces moyens peut générer des erreurs dans les données des problèmes permettant de planifier la production, si l'on utilise des modèles déterministes.

Prévision de la demande et de la production renouvelable

La production d'électricité par des panneaux solaires ou des éoliennes dépend des conditions météorologiques instantanées, qui peuvent avoir de fortes corrélations temporelles ou spatiales [ADS99; ZDK16]. Par conséquent, plus la part des renouvelables est importante, plus la production dépend de la météo. Des modèles stochastiques pour la production solaire [Bad+18; SG16] ou éolienne [Pin+09; ADS99] peuvent apporter une plus-value non négligeable en permettant de diminuer les coûts de production et en étant utilisés comme données d'entrée de problèmes d'optimisation en gestion de l'énergie.

Gestion de la production

Les plannings de production doivent prendre en compte l'incertitude créée par les sources intermittentes: des modèles stochastiques doivent être employés, même à court-terme, voir la Figure 2.4[4], comme pour le problème de Unit Commitment [Håb19; Ack+18] par exemple.

Mécanismes d'ajustement

Dans un contexte de forte pénétration des renouvelables, l'intermittence de ces sources implique de fortes erreurs de prévision de la production. Par conséquent, les besoins en termes de réserves sont plus importants. Les centrales conventionnelles pourraient répondre à ces besoins supplémentaires, mais cette solution serait coûteuse, car les centrales ne fonctionneraient alors pas à pleine capacité. De plus, des contraintes opérationnelles (dynamiques par exemple) peuvent empêcher les variations de la production nécessaires pour suivre celles de la demande résiduelle (définie comme la demande moins la production renouvelable). D'autres mécanismes comme le Demand Response (pilotage de la demande) ou le pilotage conjoint de réseaux multi-énergies sont envisagées comme des alternatives prometteuses, mais nécessitent le développement de méthodes spécifiques d'optimisation et de contrôle.

[4] Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

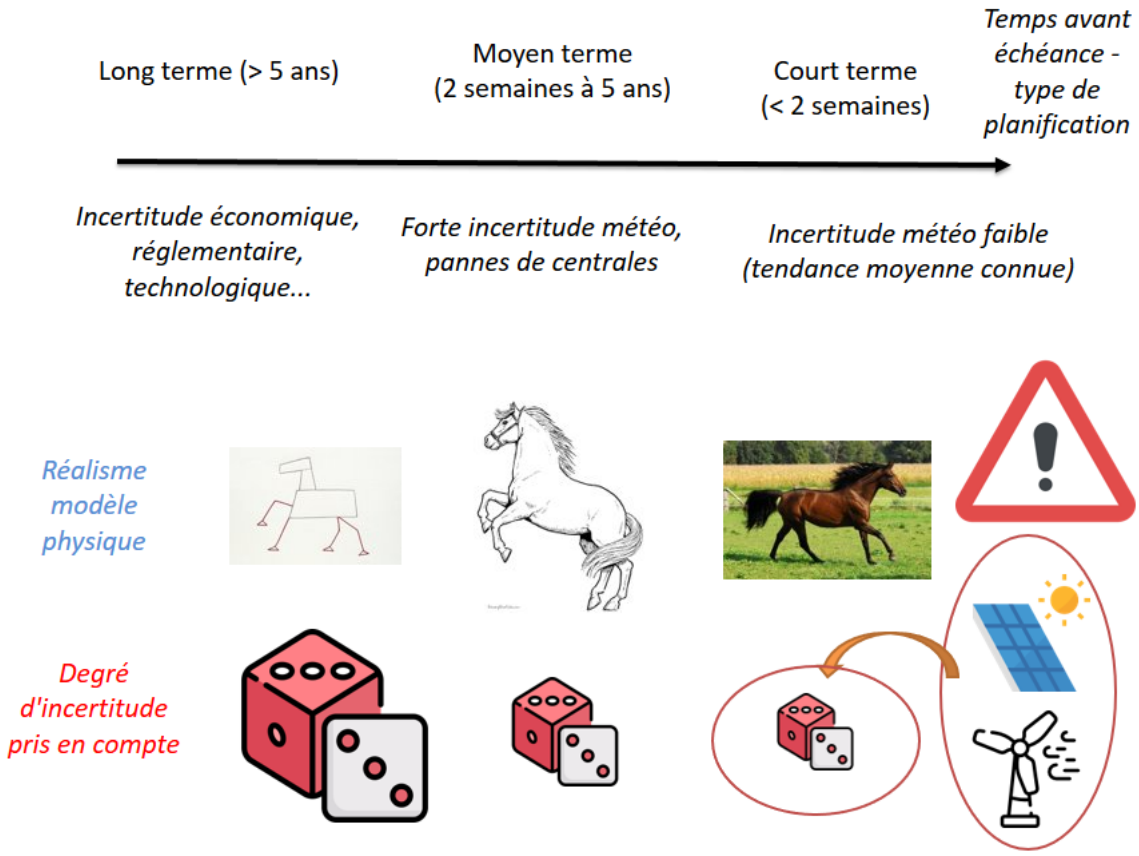


Figure 2.4: Structure des modèles d'aide à la décision pour la gestion de l'énergie avec une forte pénétration des renouvelables

2.3.2 Gestion des réseaux

Traditionnellement, la production d'électricité se faisait de façon centralisée, mais elle devient de plus en plus distribuée: de nombreux consommateurs sont maintenant équipés de panneaux solaires et de petites fermes solaires ou éoliennes sont connectées au réseau de distribution. Cette production distribuée peut contribuer à créer localement des violations de contraintes physiques sur le réseau. Par ailleurs, les sources renouvelables produisent de l'énergie sujette à de fortes erreurs de prédiction et ont donc un comportement aléatoire. Par conséquent, le point de fonctionnement du réseau est sujet à de l'incertitude également, ce qui peut être problématique pour la planification opérationnelle du réseau [ST14].

De plus, la production décentralisée pose deux problèmes potentiels: la non-monotonie des amplitudes de tension et des flux de puissance rétrograde (i.e., remontant dans le réseau) [NT15]. La production décentralisée peut dépasser la demande locale, ce qui peut provoquer une hausse du niveau de tension et des reflux de puissance du réseau de distribution vers le réseau de transport, qui n'a pas été conçu pour cela. La non-monotonie de la tension dans le réseau de distribution rend la régulation de tension plus complexe qu'auparavant, comme illustré dans la Figure 2.5[5]. Nous pouvons ainsi poursuivre la comparaison entre la gestion du réseau électrique et celle d'un réseau routier avec des routes à double sens.

[5]Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

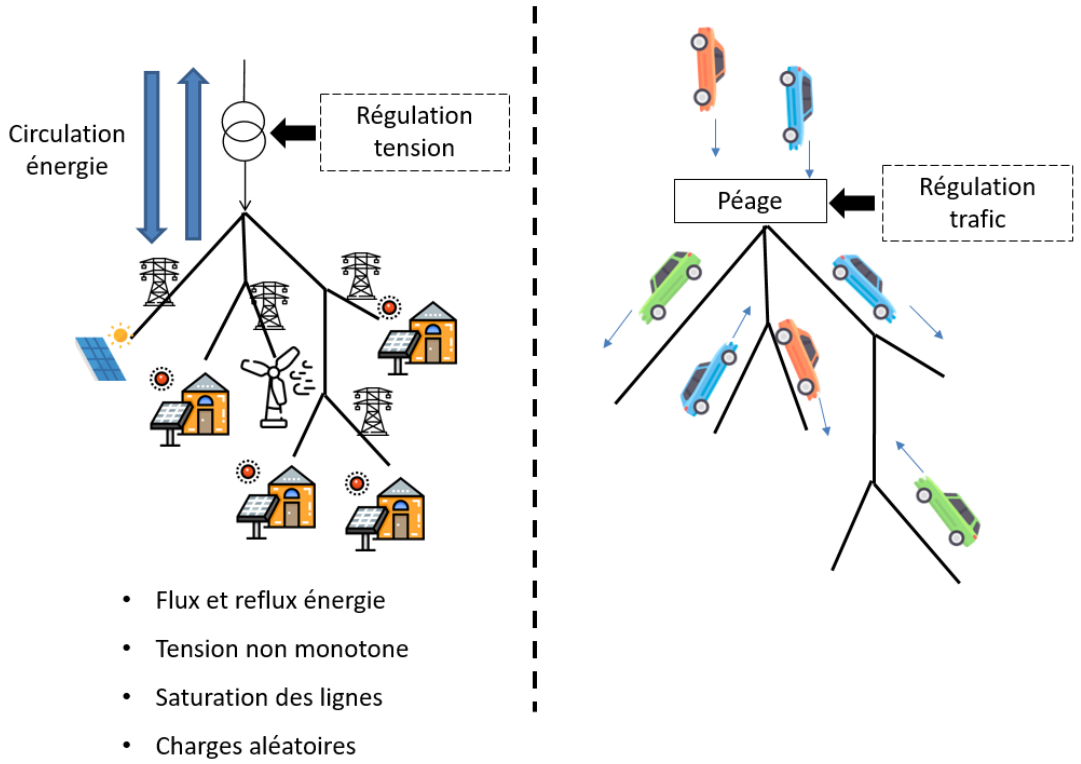


Figure 2.5: Pilotage de réseau avec des renouvelables - Comparaison avec la régulation de trafic routier

2.4 Nouvelles opportunités et nouveaux défis

2.4.1 Systèmes de stockage décentralisés

Des services système (réserve...) peuvent être fournis par des moyens de stockage situés en différents points du réseau. De tels projets existent déjà, comme les batteries à McHenry dans l'Illinois aux Etats-Unis, terminées en 2015, avec une capacité installée de 20 MW. Ce type de solution peut toutefois s'avérer coûteux. D'autres solutions peuvent être envisagées en utilisant des flexibilités énergétiques déjà existantes. C'est le cas du demand response (pilotage de la demande), qui a pour but de contrôler les flexibilités énergétiques présentes chez un grand nombre de consommateurs, ou la gestion conjointe de réseaux multi-énergies.

2.4.2 Demand response

Le pilotage de la demande peut être défini comme l'ensemble des techniques visant à modifier la consommation électrique des consommateurs finaux [AES08]. Cela peut être réalisé par le biais de signaux de prix ou d'effacement. Au niveau mondial, plusieurs marchés de l'électricité ont créé des programmes spécifiques permettant d'utiliser les charges pour fournir des services système (comme participer à des mécanismes de réserve, ou réduire un pic de consommation), qui se focalisent souvent sur les gros consommateurs industriels. La flexibilité agrégée obtenue par le contrôle simultané de plusieurs charges résidentielles pourrait s'avérer plus profitable [Cal11], mais la faible consommation de consommateurs résidentiels ne justifie pas leur participation directe aux marchés de capacité (réserve de fréquence). Cependant, de nouveaux acteurs émergent, avec parmi eux, les agrégateurs, dont le but est de se positionner en intermédiaires entre un grand nombre de consommateurs et les marchés de services système [CH10; Mat+12].

Des batteries virtuelles fournissant des services système peuvent être obtenues en contrôlant la charge d'une flotte de véhicules électriques [TRY16], dont la part de marché devrait monter significativement dans les prochaines années [Gna+18]. Les systèmes de stockage thermique (comme les réfrigérateurs, climatisations, pompes à

chaleur, chauffe-eaux, pompes de piscines...) peuvent également être pilotés pour participer aux mécanismes de réserves [Mat+12]. Ces appareils sont opérés de manières à maintenir la température d'un certain milieu dans une plage acceptable. L'inertie thermique de ces équipement en fait des candidats naturels pour fournir des services de flexibilité énergétique. Contrôler un grand nombre de tels appareils peut s'avérer moins coûteux que l'installation de batteries pour fournir des services système, avec des performances similaires [Cam+18b].

2.4.3 Multi-génération et pilotage conjoint de réseaux multi-énergies

Les réseaux de gaz [Koc+15] ou de chaud/froid peuvent être utilisés pour répondre aux besoins des réseaux électriques en termes de flexibilités énergétiques, grâce à des moyens de conversion énergétiques. Voir par exemple [Ord+17] pour un marché joint pour l'électricité et le gaz. La planification d'extension de réseaux est étudiée dans [UV+10] pour l'électricité et le gaz, dans [VBGS15; VB+17] pour des réseaux multi-énergies généraux. La gestion optimisée d'unité de multi-génération (cogénération par exemple) constitue également une piste prometteuse [CM09a; CM09b]. D'autres références et approches sont présentées dans la Section 2.5.

2.4.4 Moyens de contrôle locaux pour piloter les réseaux

Les gestionnaires de réseaux de distribution installent des moyens de contrôle locaux ainsi que de l'électronique de puissance afin de régler les violations de contraintes physiques des réseaux. Des composants permettent de maintenir la stabilité de la tension, comme les transformateurs, les compensateurs synchrones et les bancs de condensateurs [Liu+17]. Les systèmes de stockage d'énergie peuvent être employés pour éviter des phénomènes de congestion de lignes électriques ou des flux de puissance remontant dans le réseau de transport, et ainsi aider à l'insertion des renouvelables dans le mix énergétique [Das+18]. La comparaison entre pilotage de réseau électrique et gestion du trafic routier peut être développée en comparant les moyens de contrôle locaux pour les réseaux électriques à de la signalisation routière, voir la Figure 2.6[6].

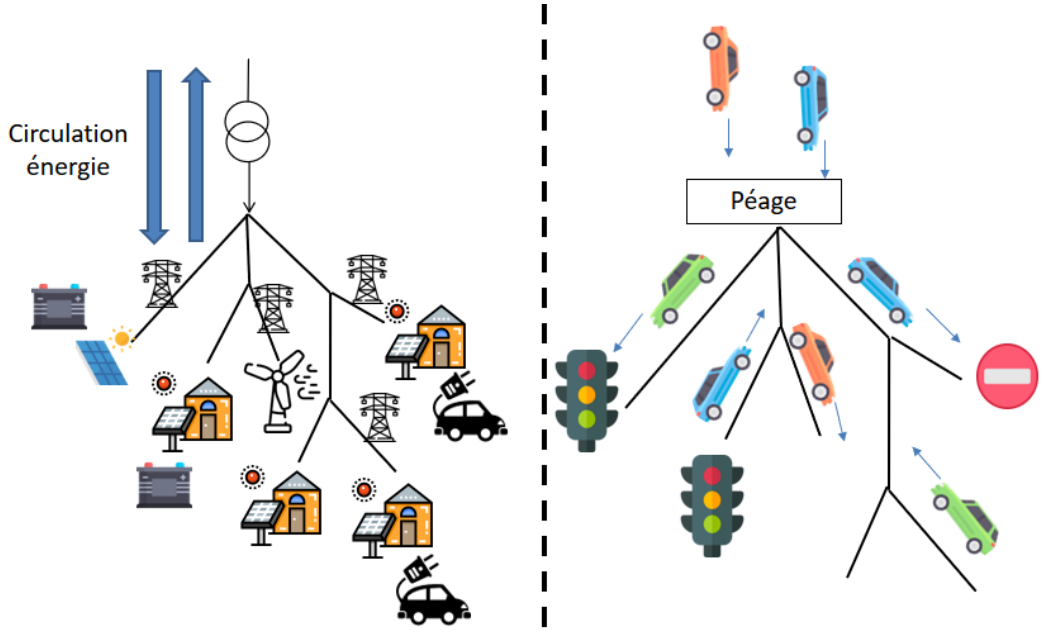


Figure 2.6: Pilotage intelligent de réseau électrique avec des renouvelables - Comparaison avec le trafic routier

[6]Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

2.5 Défis en gestion de l'énergie

L'insertion des renouvelables peut être améliorée par des adaptations de la structure des marchés, contribuant à inciter les acteurs à réduire l'impact de leurs moyens de production intermittents en utilisant par exemple des systèmes de stockage. Le pilotage optimisé de micro-grid consiste à contrôler des moyens de stockage dans un contexte de production et de consommation incertain. Pour le contrôle simultané d'un grand nombre de moyens de stockage fournissant des services système, un phénomène de fléau de la dimension peut rapidement apparaître, quand le nombre de variables d'état (stock) augmente. De plus, l'implémentation de ces moyens de contrôle exige de répondre à plusieurs problématiques. Premièrement, la confidentialité des données des consommateurs doit être respectée. Deuxièmement, les systèmes de contrôles de flexibilités locales doivent être locaux eux aussi, afin de garantir la qualité de service en cas d'interruption des télécommunications. Troisièmement, l'optimisation doit être réalisée en limitant les communications entre agents, du fait de contraintes de bande passante et de confidentialité des données.

Les sources d'énergie renouvelable et leur comportement aléatoire peuvent aussi créer des problèmes physiques locaux pour les réseaux électriques. Prendre en compte de manière fine les contraintes physiques des réseaux énergétiques est nécessaire pour cette raison mais cela rend les problèmes d'optimisation associés non-convexes. De plus, des formulations stochastiques multi-étapes sont nécessaires pour prendre en compte l'incertitude de la production renouvelable et les contraintes dynamiques des moyens de stockage, ce qui permet de garantir que les décisions n'anticipent pas des évènements pas encore survenus (non-anticipativité). De tels modèles sont difficiles à résoudre pour deux raisons: ce sont des problèmes de grande taille (car ils sont basés sur des arbres de scénarios qui doivent exploser en taille avec le nombre de pas de temps pour représenter de manière adéquate la filtration [Sha06; HRS06; PP14]) et non-convexes. Des méthodes d'approximation (linéarisations par exemple), de convexification ou de recherche locale peuvent être utilisées afin de gérer la non-convexité de ces problèmes.

Comme le contrôle joint de réseaux multi-énergies n'est pas développé davantage dans cette thèse (seuls les réseaux d'électricité seront considérés), nous présentons les défis et approches de modélisation et de résolution pour ce type de problèmes. Un cadre général pour l'optimisation de réseaux multi-énergies prenant en compte les contraintes physiques de ces réseaux (Optimal Power Flow) a été introduit dans les références [GA05; GA07]. Une difficulté de ces problèmes est que les lois physiques des réseaux d'énergie rendent les domaines admissibles non-convexes. Ajouté au comportement stochastique de la production renouvelable, cela rend ces problèmes particulièrement difficiles. Pour les résoudre, des méthodes de programmation non-linéaires peuvent être employés. Par exemple, des problèmes avec contraintes en probabilité pour le pilotage de réseaux de gaz sont considérés dans [GHH17; Hei19], et résolus à l'aide de méthodes de descente. Des problèmes avec contraintes en probabilité pour le pilotage conjoint de réseaux de gaz et d'électricité sont proposés dans [Ode+18]. Les méthodes de convexification constituent un autre type d'approches, qui peuvent fournir des garanties d'optimalité globale et sont plus précises que les approximations linéaires: elles prennent par exemple en compte des phénomènes non-linéaires comme les chutes de pressions et les pertes thermiques. Les premières approches de convexification pour les réseaux énergétiques sont apparues dans le contexte du pilotage de réseaux électriques [LL11; Low14a; Low14b; MH+19], mais des travaux récents se concentrent sur d'autres types de réseaux. L'optimisation polynomiale (exploitant la structure creuse des réseaux), utilisée dans un premier temps pour les réseaux électriques [MH14; Mol+15; Jos16; Mol+16], a récemment été employée pour modéliser des problèmes de contrôles de réseaux de chaleur dans la thèse de doctorat [Hoh18]. Cette méthode permet de considérer des relaxation Semi-Définies (donc convexes) des problèmes non-convexes, qui peuvent être résolues par des méthodes de point intérieur, qui sont suffisamment performantes pour considérer des problèmes stochastiques. Ainsi, un cadre bi-étapes stochastique est considéré pour contrôler un réseau de chaleur urbain dans [HWL19], et le problème est résolu à l'aide d'outils provenant de la théorie de l'optimisation polynomiale.

2.6 Objectifs et contenu de la thèse

Dans cette thèse de doctorat, nous nous concentrons sur l'impact de l'incertitude de la production renouvelable sur deux problèmes de la gestion de l'énergie. Premièrement, nous proposons des méthodes de pilotage d'un ou

plusieurs micro-grids en interaction afin de réduire l'impact de l'incertitude des énergies renouvelables, sans prendre en compte de manière fine les contraintes physiques du réseau. Dans un second temps, nous considérons la planification opérationnelle et le contrôle d'un réseau électrique dans un contexte incertain, en prenant en compte les lois de Kirchhoff, intrinsèquement non-linéaires. Cette thèse de doctorat introduit des contributions de modélisation, algorithmiques et théoriques permettant de répondre à ces deux problématiques.

- Dans la partie I de cette thèse, deux problèmes de contrôle sont introduits afin de réduire l'impact de l'incertitude de la production renouvelable sur les réseaux et les marchés de l'électricité, sans prendre en compte les contraintes non-linéaires des réseaux d'électricité. Ces deux problèmes sont des problèmes de contrôle optimal stochastique avec une composante de type champ moyen. Nous introduisons également deux méthodes numériques basées sur le principe d'optimalité de Pontriaguine afin de résoudre des problèmes sortant du cadre des problèmes linéaires-quadratiques, que l'on peut résoudre analytiquement. Dans le chapitre 3, nous considérons le pilotage d'un micro-grid associée à une consommation sur le réseau, qui est la différence entre la consommation et la production renouvelable à l'échelle de ce micro-grid. Celui-ci choisit un profil de consommation un jour à l'avance, et pilote en temps réel ses moyens de stockage afin de respecter son engagement. Cela permet de réduire la volatilité de la consommation résiduelle du micro-grid sur le réseau. Le problème d'optimisation jointe de la trajectoire d'engagement (choisie avant de connaître les erreurs de prévision de production et de consommation) et du contrôle des moyens de stockage en temps réel est formulé comme un problème de contrôle stochastique McKean-Vlasov (ou problème de contrôle à champ moyen). Pour ce type de problèmes, nous donnons des conditions nécessaires et suffisantes d'optimalité ainsi qu'une méthode numérique (approximative) de résolution de problèmes non linéaires-quadratiques. Un mécanisme de contrôle décentralisé d'un grand nombre de flexibilités thermostatiques est introduit dans le chapitre 4: un grand nombre de consommateurs contrôlent leurs systèmes de stockage thermiques individuels de manière coopérative de façon à fournir des services système. Une méthode de décomposition du système d'optimalité ainsi qu'une approximation de type champ moyen sont conçues afin de contourner le fléau de la dimension et d'obtenir une architecture de contrôle décentralisée garantissant la confidentialité des agents et un besoin réduit en télécommunications. Dans le chapitre 5, nous développons une méthode de Newton dans le cadre infini-dimensionnel des problèmes de contrôle stochastique. En particulier, nous montrons comment calculer le pas de Newton, théoriquement et numériquement en nous ramenant à des Equations Différentielles Stochastiques Rétrogrades, et nous proposons une méthode de recherche linéaire adaptée. Nous prouvons la convergence globale de la méthode de Newton combinée avec la procédure de recherche linéaire dans un espace vectoriel de processus adapté. Nous illustrons les performances de la méthode sur une version non-linéaire quadratique du problème présenté dans le chapitre 4. En particulier, nous proposons un algorithme complètement implémentable, basé sur des techniques de régression.

- Dans la partie II de cette thèse, nous nous intéressons à la planification et au pilotage optimal de réseaux d'électricité dans un contexte incertain. Nous introduisons le problème Optimal Power Flow (OPF) dans le chapitre 6, qui modélise le pilotage optimal de réseaux électriques, en prenant en compte les lois de la physique de manière fine. Ce chapitre donne des références et des résultats à la pointe de l'état de l'art pour le problème OPF. Dans le chapitre 7, nous formulons une version stochastique multi-étapes du problème Alternating Current Optimal Power Flow. Le cadre multi-étapes stochastique est nécessaire pour tenir compte des incertitudes de production renouvelable, de corrélations temporelles et des contraintes dynamiques pour les systèmes de stockage énergétique. Nous considérons des relaxations coniques [Low14a; Low14b; MH+19] de ce problème non-convexe (qui est fortement NP-difficile [BV19]) qui permettent un bon compromis entre temps de calcul et précision. Nous donnons des conditions réalistes et vérifiables garantissant l'absence de saut de dualité pour le problème OPF stochastique multi-étapes. Nous fournissons également une borne a posteriori sur le saut de dualité du problème. Ces résultats apportent un éclairage théorique sur des conditions sous lesquelles les relaxations convexes du problème OPF sont fiables. Les preuves des résultats sont basées sur des arguments de reconstruction primale (aussi appelés arguments d'arrondis) inspirés de [Hua+16], i.e., sur un algorithme prenant en entrée une solution faisable de la relaxation convexe et retournant une solution faisable du problème original non-convexe. En nous basant sur des résultats liés au théorème de Shapley-Folkman, nous obtenons par ailleurs des bornes sur le saut de dualité de problèmes stochastiques

multi-étapes et non-convexes en gestion de l'énergie. Cela montre que les systèmes de stockage peuvent améliorer les performances de résolution numérique de problèmes non-convexes multi-étapes, en plus de baisser les coûts de gestion.

2.7 Contributions des chapitres

Nous donnons davantage de détails sur les contributions associées à chacun des chapitres de cette thèse de doctorat.

1. Le **chapitre** 3 introduit un problème de contrôle d'un micro-grid, représentant un consommateur équipé de panneaux solaires et d'un système de stockage, et connecté au réseau. Ce consommateur a pour objectif de réduire l'incertitude de sa consommation résiduelle sur le réseau pour le jour suivant. Pour ce faire, il choisit un jour à l'avance un profil de consommation sur le réseau, puis il pilote en temps réel son système de stockage de manière à minimiser simultanément les coûts de gestion opérationnels et les déviations de sa consommation réelle de son engagement, voir la Figure 2.7[7]. Nous modélisons cette situation par un problème de contrôle à champ moyen, avec des interactions scalaires et des filtrations générales. Ceci permet une optimisation jointe de décisions avec des échéances temporelles différentes: l'engagement est pris initialement mais impacte le système sur tout l'horizon de temps, tandis que le système de stockage est piloté en réponse aux aléas survenus en temps réel. Nous établissons des conditions nécessaires et suffisantes d'optimalité pour ce problème sous la forme d'une équation différentielle stochastique progressive rétrograde à champ moyen, voir les Théorèmes 3.2.3 et 3.2.10. Nous étudions ensuite la solvabilité de ce système. Nous proposons une méthode numérique d'approximation des solutions de problèmes de contrôle à champ moyen non linéaires-quadratiques, basée sur la théorie des perturbations, voir le Théorème 3.3.7. Nous illustrons ces résultats numériquement.

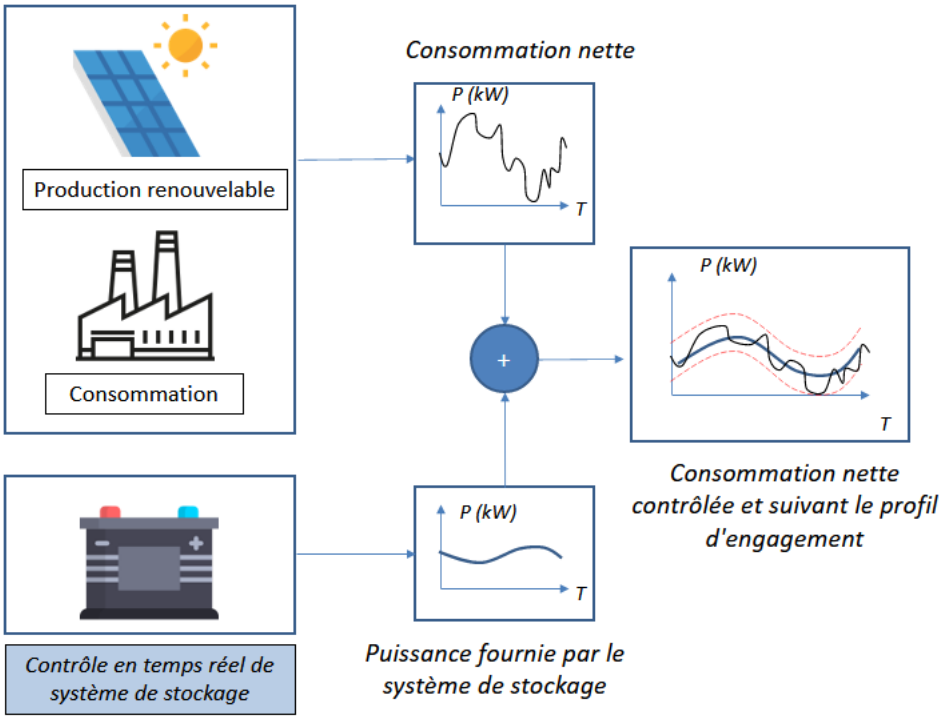


Figure 2.7: Engagement de consommation un jour à l'avance avec suivi en temps-réel du profil d'engagement

[7]Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

2. Le **chapitre** 4 est motivé par le potentiel important des unités de stockage thermique pour fournir des services système [Mat+12; Cam+18b]. Nous y proposons un problème de contrôle stochastique visant à contrôler de manière coopérative des charges thermostatiques pour assurer l'équilibre offre-demande. Nous développons une méthode pour résoudre ce problème de manière décentralisée, de façon à respecter la confidentialité des consommateurs, à réduire les télécommunications et les besoins en temps de calcul, comparé au cas d'un planificateur central, voir la Figure 2.8[8]. Les conditions d'optimalité de ce problème sont données sont la forme d'une équation différentielle stochastique progressive-rétrograde (Théorème 4.2.1). Ce système est décomposé en un problème de coordination, une équation différentielle stochastique progressive rétrograde de faible dimension modélisant le comportement optimal des agrégats de population de charges thermostatiques, et des problèmes individuels, qui sont des équations différentielles stochastiques progressives rétrogrades de dimension 1 modélisant les comportements optimaux individuels. Des résultats d'existence et d'unicité pour ces problèmes sont obtenus. Nous montrons également que ces équations caractérisent de façon unique l'équilibre de Nash d'un jeu différentiel stochastique de Stackelberg. Dans ce jeu, un coordinateur (le meneur) a pour but de contrôler les comportements agrégés de la population en envoyant des signaux de coordination, et chaque consommateur (suiveur) répond à ce signal en optimisant son système de stockage localement. Cela permet une implémentation décentralisée du contrôle optimal du point de vue d'un planificateur central. Il s'agit de la perspective inverse de jeux à champ moyen potentiel [FMHL19], qui ont pour but de trouver des problèmes de contrôle optimal (convexes) dont les conditions d'optimalité coïncident avec le système de Nash de jeux différentiels stochastiques. Nous proposons une approximation de type champ moyen du problème de coordination, dans laquelle les paramètres ne dépendent plus des comportements agrégés en temps réels des agents, garantissant que le problème peut être résolu en observant le bruit commun uniquement, voir le Théorème 4.4.5. Cela permet de répondre aux contraintes de télécommunication et de confidentialité. Cette approximation est basée sur la loi des grands nombres conditionnelle. Des résultats de convergence et des bornes d'erreur sont obtenus pour cette approximation en fonction de la taille de la population, voir le Théorème 4.4.12. Nous illustrons numériquement l'intérêt de cette architecture de contrôle et la convergence de l'approximation de type champ moyen. Nous proposons enfin une implémentation répondant à des problématiques industrielles réelles, voir l'Algorithme 4.1.

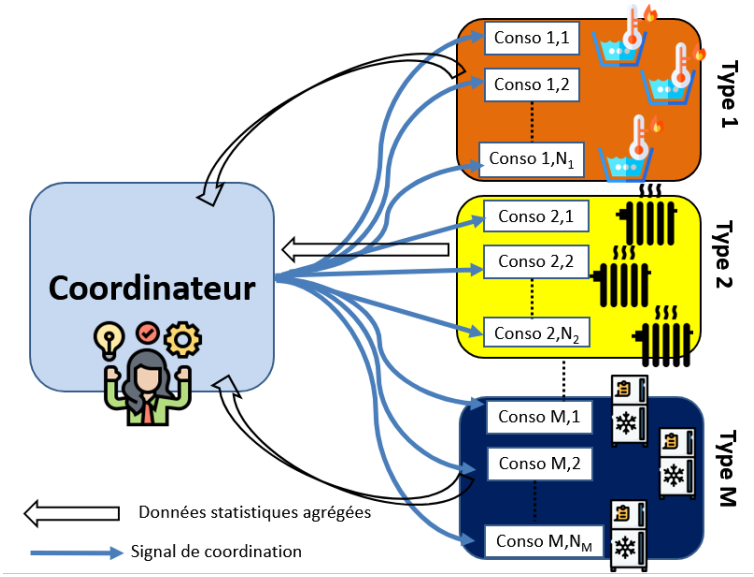


Figure 2.8: Mécanisme de contrôle décentralisé

3. Dans le **chapitre** 5, nous développons une nouvelle méthode (théorique et numérique) basée sur le principe de Pontriaguine pour résoudre des problèmes de contrôle stochastique. Il s'agit de la méthode de Newton étendue au cadre des problèmes de contrôle stochastique avec des filtrations générales, où la dynamique de

[8] Les icônes sont produits par Freepik et Smashicons et disponibles sur www.flaticon.com

l'état (contrôlé) du système est une équation différentielle ordinaire avec des coefficients stochastiques. En particulier, nous montrons que la méthode de Newton revient à résoudre de manière successive des approximations linéaires quadratiques du problème de contrôle stochastique (obtenues par développement limité), voir la Proposition 5.3.2. De façon équivalente, la méthode de Newton est obtenue par linéarisations successives du système d'optimalité établi en appliquant le principe de Pontriaguine, voir la Proposition 5.3.3. Nous proposons une méthodologie complète pour calculer théoriquement le pas de Newton, qui peut être calculé en résolvant une équation différentielle stochastique progressive rétrograde affine avec des coefficients aléatoires, voir le Théorème 5.2.9 et la Définition 5.3.1. Nous montrons ensuite que résoudre cette équation est équivalent à la résolution d'une Équation différentielle stochastique rétrogrades (EDSR) de type Riccati et d'une EDSR affine, voir le Théorème 5.3.7, comme attendu dans le cadre de telles équations linéaires [Yon06] ou de problèmes de contrôle stochastique linéaires-quadratiques [Bis76]. Nous étudions ensuite la convergence de la méthode de Newton pour les problèmes de contrôle stochastique. Nous montrons par un contre-exemple (Exemple 5.3.9) que des conditions de régularité garantissant la convergence locale de la méthode de Newton (comme une dérivée seconde Lipschitz-continue [Kan48; NW06]) ne sont pas nécessairement vérifiées dans l'espace de processus considéré initialement. Ceci nous conduit à nous restreindre à des processus essentiellement bornés. Nous donnons un autre contre-exemple (Exemple 5.3.12) qui montre que même pour des problèmes fortement convexes et très réguliers, la méthode de Newton ne converge pas nécessairement globalement. Nous développons une méthode de recherche linéaire adaptée (différente de celle présentée dans [NW06; BV04]), ce qui nous permet de prouver la convergence globale de la méthode de Newton avec cette recherche linéaire, voir le Théorème 5.3.15. Nous donnons une implémentation avec des techniques de régression pour résoudre les EDSR intervenant dans le calcul du pas de Newton, que nous appliquons à un problème de contrôle d'un grand nombre de batteries fournissant des services système.

4. Le **chapitre 6** introduit le problème Optimal Power Flow (OPF), qui modélise le contrôle optimal d'un réseau électrique en prenant en compte les lois de Kirchhoff, non-linéaires, dans le contexte de courant alternatif. Les formulations les plus courantes ainsi que celles utilisées dans le chapitre suivant y sont présentées. Nous supposons que le lecteur n'a pas nécessairement une connaissance approfondie du problème étudié et nous cherchons à avoir une approche la plus pédagogique possible, tout en donnant des pointeurs vers des techniques de résolution, des approximations, des relaxations et des extensions récentes du problème OPF.

5. Dans le **chapitre 7**, nous formulons une version stochastique multi-étapes du problème Alternating Current Optimal Power Flow, ce qui permet de prendre en compte à la fois l'incertitude des énergies renouvelables et les contraintes dynamiques des systèmes de stockage. Nous donnons également une relaxation conique de ce problème non-convexe. Nous développons ensuite des bornes sur le saut de dualité du problème OPF stochastique multi-étapes avec stockage. Nous donnons en particulier des conditions réalistes et facilement vérifiables a priori garantissant l'absence de saut de dualité pour le problème OPF dans un cadre stochastique multi-étapes, ce qui étend les résultats de [Hua+16]. Nous obtenons également une borne a posteriori facilement calculable sur le saut de relaxation de ce problème. En utilisant des outils différents, nous établissons des bornes a priori sur des problèmes stochastiques multi-étapes non-convexes en gestion de l'énergie, dans lesquels le stockage joue un rôle important, voir la Proposition 7.5.9 et le Corollaire 7.5.10. Ces résultats sont liés au Théorème de Shapley-Folkman, qui donne une borne sur la distance de Hausdorff entre la somme de Minkowski d'ensembles compacts non-convexes et son enveloppe convexe.

2.8 État de l'art

Nous donnons un panorama des méthodes existantes dans la littérature en lien avec les deux parties de cette thèse de doctorat, afin de comparer nos contributions à l'existant.

2.8.1 Contrôle de micro-grids en contexte incertain (sans prise en compte des contraintes physiques du réseau)

Différentes approches ont été récemment proposées pour contrôler un micro-grid. Le modèle proposé dans le chapitre 3 s'inscrit dans cette thématique de recherche, mais en prenant en compte une particularité: l'optimisation jointe d'une décision prise à l'instant initial et impactant le système sur tout l'horizon de temps (engagement de consommation) et du pilotage d'un système de stockage en temps réel.

- Une approche relativement populaire pour contrôler des systèmes de stockage énergétique en contexte incertain est l'approche Model Predictive Control. Celle-ci consiste à ignorer l'incertitude dans le modèle et à remplacer les paramètres incertains par leurs meilleures prévisions. Un modèle physique détaillé d'une batterie et de panneaux solaires est considéré dans [SSM16] et une architecture de type Model Predictive Control est développée. Cependant, l'approche Model Predictive Control est une heuristique dont la performance doit être validée par simulation. Ce manque de garantie d'optimalité est étudié dans [Pac+18], qui compare cette approche avec la Programmation Stochastique Dynamique Duale pour résoudre un problème d'optimisation de micro-grid. Dans ce modèle sont considérés une batterie, un ballon d'eau chaude et le système de chauffage d'un bâtiment.

- Une autre approche fréquemment utilisée est la modélisation par un Processus de Décision Markovien et une résolution du problème d'optimisation par Programmation Dynamique Stochastique, comme pour des problèmes de charge de véhicules électriques en contexte incertain [IMM14; Wu+16]. Cette approche fonctionne bien lorsque l'état du système peut être décrit par une variable de faible dimension.

- Le vieillissement est pris en compte dans des modèles de pilotage de batterie visant à aider à l'insertion de renouvelables dans [Hae14] et [Car+19a]. Cette dernière référence propose une méthode de décomposition temporelle (programmation dynamique) à deux échelles de temps pour contrôler la batterie à court-terme tout en prenant en compte son vieillissement sur le long terme.

- Une modélisation en temps continu est considérée pour le contrôle optimal d'un micro-grid équipé d'un moteur diesel et d'une batterie en contexte déterministe dans [Hey+15] et stochastique dans [Hey+16].

Des méthodes de décomposition ont par ailleurs été développées pour être capables de résoudre des problèmes de gestion de plusieurs micro-grids en interaction. Ainsi, une méthode de décomposition spatiale et temporelle est proposée dans [Car+19b; Car+20] pour résoudre le problème de micro-grids équipés de batterie et pouvant échanger de l'énergie par le biais d'un réseau. En particulier, la méthode développée permet de considérer des problèmes pour lequel la dimension de l'état du système est beaucoup plus élevée que ce qui peut être résolu par des méthodes standards comme la Programmation Dynamique Stochastique ou la Programmation Dynamique Stochastique Duale. Cette méthode de décomposition repose sur la formulation de sous-problèmes associés aux nœuds du réseaux, chacun pouvant être résolu par programmation dynamique. Un modèle simplifié et linéaire du réseau électrique y est considéré. En particulier, le modèle ne prend pas en compte les phénomènes liés au courant alternatif, ni les pertes thermiques et la puissance réactive. Ces phénomènes sont considérés de manière plus précise et réaliste dans le problème Optimal Power Flow stochastique multi-étapes étudié dans le chapitre 7. Par ailleurs, il est souhaitable de développer des méthodes ne souffrant pas du fléau de la dimension si l'on souhaite contrôler un grand nombre de systèmes de stockage. En effet, les réseaux électriques présentent déjà un grand nombre de systèmes de stockage énergétique connectés comme des batteries, des charges thermostatiques (tels les ballons d'eau chaude, des pompes à chaleurs, des systèmes de climatisation...) qui peuvent aider à fournir les services système requis pour l'essor des énergies renouvelables. Les systèmes de stockage thermique ont été identifiées comme des leviers prometteurs pour ce faire [Mat+12; Cam+18b]. Cependant, contrôler de façon optimale un grand nombre de ces appareils pose plusieurs défis:

1. La dimension de l'espace d'état correspond au nombre d'appareils contrôlés, qui peut être très grand. Par conséquent, un phénomène de fléau de la dimension peut rapidement apparaître avec des méthodes comme la programmation dynamique.

2. Les systèmes de contrôle de moyens de stockage doivent être locaux. En effet, une architecture centralisée nécessiterait des infrastructures de télécommunication lourdes pour que le planificateur central puisse rassembler les données des agents et envoyer des signaux de contrôle. De plus, la qualité de service (i.e., le maintien de la température associée à chacune des charges thermostatiques dans des plages admissibles) pourrait être dégradée en cas de perte de communication avec le planificateur central.

3. La confidentialité des données individuelles des consommateurs doit être respectée.

Le modèle proposé dans le chapitre 4 et le cas d'application illustrant la méthode numérique introduite au chapitre 5 répondent à ces problématiques. Nous donnons un aperçu des méthodes utilisées dans la littérature pour le contrôle décentralisé de nombreuse flexibilités énergétiques. Un première catégorie de travaux de recherche visant à répondre à ces problématiques est le développement d'architecture de contrôle pouvant être implémentées et basées sur la coopération des usagers. Des contrôleurs locaux répondant à un signal de coordination et basés sur des modèles d'Équations aux Dérivées Partielles sont considérés dans [TTS15; Tro+16]. Une approximation de type champ moyen d'un grand nombre de flexibilités thermostatiques est utilisée dans [BM16; Cam+18a] et des boucles de contrôle locales sont implémentées de manière à assurer que l'agrégation des flexibilités se comporte comme une batterie virtuelle. Pour ces références, le taux de commutation (aléatoire) des appareils joue le rôle de la variable de contrôle. De plus, ces méthodes n'apportent pas de garantie théorique d'optimalité et sont avant tout orientées vers l'implémentation.

Des approches basées sur la théorie des jeux constituent également un moyen pratique d'obtenir des architectures de contrôle décentralisées, avec des agents cherchant à satisfaire leurs propres intérêts et répondant à des signaux de prix. Un mécanisme de facturation horaire est proposé dans [Jac+18] pour inciter les consommateurs à consommer moins aux horaires de forte demande. Le cas des jeux avec des agents hétérogènes est considéré dans [JW18]. Une technique de cryptographie est employée dans [Jac+19] afin qu'un agrégateur puisse observer la flexibilité énergétique agrégée d'un ensemble de consommateurs sans avoir accès à leurs données individuelles.

Les approximations de type champ moyen sont par ailleurs particulièrement indiquées car elles sont d'autant plus précises que le nombre d'agents est important. Les modèles de jeux à champ moyen ont ainsi attiré l'intérêt de nombreux chercheurs pour modéliser des problèmes en gestion de l'énergie [DP+19; ATM20; MMS19; KM13; KM16]. Dans [DP+19], des systèmes de stockage thermique répondent à des signaux de prix déterminés par un problème de Unit-Commitment afin d'allouer la part flexible de leur consommation et leur participation à des services système. Le problème est résolu avec une méthode de point fixe visant à résoudre le système formé par les équations d'Hamilton-Jacobi-Bellman et de Fokker-Plank, qui est un système d'Équations aux Dérivées Partielles couplées modélisant la dynamique de la fonction valeur et de la mesure de probabilité des états de la population. Cette méthode ne présente pas de garantie de convergence. Dans [ATM20], des consommateurs utilisent des batteries afin de réduire leurs factures d'électricité, le prix de marché de l'électricité étant déterminé par la demande agrégée des consommateurs. L'équilibre de Nash de ce jeu à champ moyen avec bruit commun (représentant les corrélations géographiques dues à la météo, qui impactent la production renouvelable locale) est caractérisé en utilisant le principe de Pontriaguine, ce qui permet d'obtenir des formules sous forme de feedback (boucle de rétro-action) explicites dans le cas linéaire-quadratique. Le modèle de l'article [ATM20], qui ne considère que des filtrations Browniennes, est étendu au cas des processus à sauts dans [MMS19]. Un modèle de jeu à champ moyen sans bruit commun est proposé dans [KM13], dans lequel la température moyenne associée à un ensemble de flexibilités thermiques (comme des ballons d'eau chaude par exemple) doit suivre un profil spécifique, afin de fournir des services système. Le modèle considéré est linéaire-quadratique Gaussien, ce qui assure l'existence de formules quasi-explicites pour l'équilibre de Nash. Le modèle est étendu aux processus Markoviens à saut dans [KM16], avec des modèles plus fins des ballons d'eau chaude: le phénomène de stratification (i.e., la non-homogénéité de la température de l'eau contenue dans les ballons) y est pris en compte. Les modèles basés sur la théorie des jeux et des jeux à champ moyen fournissent un moyen pratique d'obtenir une architecture de contrôle décentralisée, mais ils ne fournissent qu'un équilibre de Nash qui peut être sous-optimal d'un point de vue collectif. Des bornes sur le prix de l'anarchie sont obtenues dans [Jac+18] pour de tels modèles de jeux appliqués à la gestion de l'énergie. Dans les modèles linéaires-quadratiques et sous certaines hypothèses, les modèles de contrôle à champ moyen (resp. de jeux à champ moyen) ont un unique optimum (resp. équilibre de Nash) pouvant être calculé analytiquement, ce qui permet d'obtenir le prix de l'anarchie, qui quantifie la perte d'optimalité entre le

cadre coopératif (contrôle à champ moyen) et compétitif (jeu à champ moyen). Une approximation de type champ moyen combinée avec une décomposition par les prix est également proposée dans le récent travail [Seg+20] afin de résoudre des problèmes de contrôle stochastique modélisant des problèmes d'agrégation de flexibilités en gestion de l'énergie.

2.8.2 État de l'art pour le contrôle optimal de réseaux électriques prenant en compte les contraintes physiques

Le problème de contrôle optimal d'un réseau électrique en prenant en compte les lois de Kirchhoff non-linéaires est appelé le problème Optimal Power Flow (OPF). Il est situé à la frontière entre l'optimisation mathématique, la physique et le génie électrique. Pour cette raison, nous avons fait le choix d'introduire en détail le problème OPF dans le chapitre 6. Nous y donnons de nombreux pointeurs vers de la littérature récente donnant des résultats théoriques, des formulations mathématiques, des approximations, des relaxations convexes et des méthodes d'optimisation associées au problème OPF. Dans la suite de ce paragraphe, nous ne faisons que résumer les principaux résultats, de manière plus succincte.

Définition

Le problème Optimal Power Flow (OPF) est un problème d'optimisation mathématique visant à trouver le point de fonctionnement d'un réseau électrique (en prenant en compte les lois de Kirchhoff) qui minimise une certaine fonction objectif, comme les coûts de production, les pertes thermiques dans les lignes, sous des contraintes physiques, comme des niveaux maximums et minimums d'amplitude de tension, d'intensité ou de puissance en certains points du réseau. Ce problème a été formulé pour la première fois par Carpentier en 1962 [Car62] et il a par la suite été beaucoup utilisé du fait de sa versatilité.

La formulation la plus précise de ce problème est le problème "Alternating Current Optimal Power Flow" (AC OPF). Il s'agit d'un problème d'optimisation non-convexes avec des contraintes quadratiques égalité, qui se formule naturellement avec des variables complexes. Celles-ci permettent en effet de représenter de manière compacte des signaux oscillants, que l'on retrouve dans les réseaux opérés en régime alternatif.

Réseaux à une ou plusieurs phases, réseaux radiaux ou maillés

Les réseaux les plus courants peuvent être à monophasés (à une phase), triphasés équilibrés ou triphasés déséquilibrés. Les réseaux monophasés sont constitués de lignes composées de deux câbles: un câble "terre" et un câble pouvant transporter de la puissance. Les réseaux triphasés équilibrés ou déséquilibrés comportent des lignes composées de quatre câbles: un câble terre et trois câbles transportant les charges électriques. Dans le cas des réseaux équilibrés, l'amplitude de la puissance circulant dans les trois câbles (appelés phases) est la même et les signaux ne diffèrent que par un déphasage de $120°$ les uns par rapport aux autres, contrairement aux réseaux déséquilibrés. Le problème AC OPF pour des réseaux triphasés équilibrés peut être réduit au cas monophasé.

De plus, les réseaux peuvent avoir une architecture radiale (acyclique) ou maillée (avec des cycles). Les réseaux de transport sont typiquement maillés et monophasés. Les réseaux de distribution de moyenne tension sont le plus souvent radiaux et triphasés équilibrés, tandis que les réseaux de distribution de basse tension sont radiaux et triphasés déséquilibrés. Nous nous concentrons sur le cas des réseaux monophasés, ce qui permet de considérer les réseaux triphasés équilibrés. Des résultats pour les réseaux triphasés déséquilibrés sont donnés dans le chapitre 6.

Méthodes de résolution du problème AC OPF pour des réseaux monophasés dans un cadre déterministe et statique

Il a été montré récemment que le problème AC OPF est fortement NP-difficile [BV19]. Pour le résoudre malgré sa difficulté, plusieurs méthodes peuvent être appliquées, comme des méthodes basées sur la programmation non-linéaire, des linéarisations ou des relaxations convexes.

Les méthodes issues de la programmation non-linéaire incluent la méthode de Newton Raphson [IS87], la programmation quadratique séquentielle, les méthodes de plus forte descente [For+10]. Ces techniques d'optimisation et beaucoup d'autres sont présentées dans [FSR12a; FSR12b]. Cependant, elles sont souvent sensibles aux conditions initiales et ne convergent pas forcément vers un optimum global, ce qui peut s'avérer coûteux pour les gestionnaires de réseaux.

Les approximations linéaires sont largement utilisées en pratique. Cependant, avec la part croissante des renouvelables dans le mix énergétique, elles peuvent négliger des phénomènes importants, comme des variations d'amplitude de tension [NT15]. Parmi les linéarisations existantes, la plus connue est sans aucun doute la formulation Direct Current Optimal Power Flow (DC OPF), qui suppose les amplitudes de tension fixées, et néglige les pertes et la puissance réactive. Des méthodologies pour obtenir des bornes d'erreurs entre les modèles DC et AC OPF ont été développées [SJA09; DM16]. Il existe de nombreuses autres linéarisations, comme le modèle Linearized Distflow, particulièrement adapté aux réseaux radiaux (i.e., acycliques), ou les linéarisations du problème AC OPF autour d'un point de fonctionnement nominal.

Les relaxations convexes (coniques) du problème AC OPF ont l'avantage de pouvoir être résolues efficacement par des méthodes de point intérieur et fournissent une garantie d'optimalité globale quand leur optimum est faisable pour le problème original non-convexe. Parmi elles, les relaxations Semi-Définie, conique quadratique et chordale sont les plus connues et sont présentées dans [Low14a; Low14b; MH+19]. Le saut de relaxation associé est faible ou nul pour de nombreuses instances réalistes [LL11]. Des conditions garantissant l'absence de saut de dualité ont été démontrées [LL11; SL12; FL13; Gan+14; Hua+16] mais il reste un écart entre les conditions théoriques garantissant l'absence de saut de dualité et la pratique. En effet, de nombreuses instances ont un saut de dualité nul sans pour autant satisfaire les conditions théoriques trouvées dans la littérature. La relaxation semi-définie a un saut de relaxation plus faible que la relaxation conique quadratique, mais au prix d'un temps de calcul plus élevé, car elle est formulée avec des matrices dont le nombre d'entrées croît avec le carré du nombre de nœuds du réseau. Pour des réseaux radiaux, ces deux relaxations ont le même saut de relaxation, si bien que la relaxation conique quadratique est toujours préférable du fait d'une complexité de résolution numérique moindre. Des relaxations intermédiaires pour résoudre des instances de grande taille du problème AC OPF avec une bonne précision et en temps raisonnable ont également été proposées, comme les relaxations semi-définies "partielles" [BAD18], des relaxations coniques quadratiques renforcées [KDS16]. Des méthodes de pénalisation [MAL15] ou provenant de l'optimisation polynomiale parcimonieuses [MH14] ou des combinaisons de ces deux méthodes [Mol+15; Mol+16; Jos16] peuvent être employées pour résoudre des instances de grande taille pour lesquelles la relaxation semi-définie a un saut de relaxation non nul.

Extensions au cas dynamique et/ou stochastique

Un modèle AC OPF dynamique avec système de stockage d'énergie est formulé dans [GKA13], basé sur la formulation non-convexe des contraintes et sans prise en compte des incertitudes. La relaxation conique quadratique est utilisée dans un cadre dynamique dans [GSGK18] pour dimensionner de manière optimale des réseaux de distribution. Des conditions trouvées dans [LL11] sous lesquelles la relaxation semi-définie est exacte sont étendues au cas dynamique avec stockage dans [GT12]. Une revue de la littérature pour les méthodes associées au problème OPF avec des application aux réseaux de distribution avec systèmes de stockage est effectuée dans [SM16].

Les modèles stochastiques les plus utilisés sont les modèles OPF avec contraintes en probabilité. Une revue de la littérature sur le sujet est effectuée dans [BCH14]. Des travaux récents proposent une version du problème DC OPF avec contraintes en probabilité [Roa+13; Roa+16], tandis que d'autres s'intéressent à des versions probabilistes du problème AC OPF linéarisé autour d'un point de fonctionnement de référence [RMT17; RA17]. D'autres travaux considèrent la formulation robuste associée à la relaxation semi-définie et restreinte à des règles de décision affines [Vra+13]. Une relaxation semi-définie du problème AC OPF avec contraintes en probabilité est proposée dans [Ven+17], se basant sur une approche par scénarios et avec des règles de décision affines, ou en faisant l'hypothèse de bruit gaussien. La référence [HPC18] présente une relaxation conique quadratique du problème AC OPF avec contraintes en probabilité, combinée avec des méthodes de récupération de solution faisable. Une autre classe de modèles stochastiques est celle des modèles multi-étapes, qui garantissent la non-anticipativité des décisions: celles-ci ne doivent dépendre que des quantités aléatoires observées à l'instant où elles sont prises, et

pas des aléas ultérieurs. Les références [Swa17; NCP14] considèrent le cas simplifié où les décisions sont non-anticipatives pour les premiers pas de temps uniquement, les arbres de scénarios ayant une structure de peigne. La non-anticipativité est garantie dans [JKK14] qui se restreint à des règles de décision affines avec un modèle DC OPF, et hérite donc des limites de ces modèles approximatifs. Elle est également assurée dans [Sun+16] qui considère une procédure heuristique itérative pour optimiser une politique de décision.

Part I

Optimal control of micro-grid in a stochastic environment

Chapter 3

Day-ahead Commitment

3.1 Introduction

General context in energy management. The energy sector is currently facing major changes because of the raising concern about climate change, the search for energy-efficiency and the need to reduce carbon footprint. In particular, the share of renewable energy (RE for short) production has increased in most industrialized countries over the last few years, and further effort has to be done to limit the temperature increase well below 2° C by 2100, as targeted by the 2015 Paris agreement. However, even if these renewable energies allow a huge reduction of carbon footprint during the energy production phase, they raise a major issue: the amount of energy produced is intermittent and uncertain, as a main difference with more conventional energy production units (coal/gas-fired units, or nuclear power plants).

Reducing uncertainty of net residual consumption. Since the electricity production has to meet consumption at all spatial and time scales, the load balancing operations become harder in this uncertain context, this leads to higher operating costs for the whole electricity system; furthermore, it sometimes lead to ecologically catastrophic solutions such as the use of coal units to compensate the deficit of clean energy production. See [Mor+14] for an overview on how to integrate renewables in electricity markets. Therefore, a major challenge is to smooth the electricity consumption by better predicting RE production and better managing the energy system. We address the latter in the context of a consumer equipped with its own RE production (e.g. PV panels), and formalize the problem as a stochastic control problem of McKean-Vlasov (MKV for short) type that we solve theoretically and numerically. More specifically, we study a decentralized mechanism aimed at reducing the variability of residual consumption on the electricity network; thus, operating the network could be done at lower costs and with a lower carbon footprint. This mechanism is a setting where a consumer has to commit in advance (say T=one day-ahead, to match the usual working of day-ahead markets) to a predefined load profile and then, he has to command optimally and dynamically his system according to his stochastic consumption/production. Both the optimal load profile and the optimal control are the outputs of the stochastic control problem described below. The above model is a simplified prototype of *smart grid* (as defined by the European Commission[1]): our so-called *consumer* is considered as an association of small consumers, with possibly individual RE production and individual storage facilities, that we aggregate and consider as a whole.

General setting and methodology. We take the point of view of a consumer supplied in energy by its own intermittent sources (PV panels for instance) and by the electrical public grid. We consider the situation where the non-flexible consumption and the intermittent production are exogenous and can not be predicted perfectly: a stochastic model should be used for both of them. See [Bad+18] about a recent methodology for deriving a probabilistic forecast for solar irradiance (and thus PV production). To smooth his residual consumption, the consumer can take advantage of storage facilities (for instance conventional batteries, electrical vehicle batteries, heating network,

[1] http://www.ieadsm.org/publication/functionalities-of-smart-grid-and-smart-meters-eutf/

flywheel etc) which we consider as a whole. At time t, his control is denoted by u_t, the level of storage is represented by X_t^u, its net consumption on the electrical public grid is $\mathrm{P}_t^{\mathrm{grid},u}$. The (deterministic) committed profile load is the curve $(\mathrm{P}_t^{\mathrm{grid,com.}} : 0 \leq t \leq T)$. Optimal control of a single micro-grid has already been considered in the literature, without the optimal committed load profile. A popular yet without theoretical optimality guarantee is Model Predictive Control [SSM16]. In discrete-time settings, Stochastic Dynamic Programming [IMM14; Wu+16] and Stochastic Dual Dynamic Programming [Pac+18] are popular approaches to get theoretical optimality guarantees. Long-term aging of the battery equipping a micro-grid is taken into account by two-time scales time decomposition in [Car+19a]. Continuous time optimal control problems are considered in [Hey+15] in a deterministic setting, and in [Hey+16] in a stochastic environment. By jointly optimizing with the profile $\mathrm{P}^{\mathrm{grid,com.}}$, we change the nature of the stochastic control problem, compared to these works. We shall consider general filtrations with processes possibly exhibiting jumps, to account for sudden variations of solar irradiance or consumption for instance.

In short, in a simplified setting, the optimization criterion takes the form of the following cost functional

$$\mathbb{E}\left[\int_0^T \left\{\mathrm{C}_t \mathrm{P}_t^{\mathrm{grid},u} + \frac{\mu}{2}u_t^2 + \frac{\nu}{2}\left(X_t^u - \frac{1}{2}\right)^2 + l_1\left(\mathrm{P}_t^{\mathrm{grid},u} - \mathrm{P}_t^{\mathrm{grid,com.}}\right)\right\}\mathrm{d}t + \frac{\gamma}{2}\left(X_T^u - \frac{1}{2}\right)^2\right],$$

minimized over admissible controls $(u_t)_t$. The first term in the above cost functional is the cost of buying electricity to the electrical public grid, at a price C_t which can be random. The second term in the cost functional accounts for a penalization of the use of the storage (e.g. aging cost in the case of a battery). The third and fifth terms are penalization of the deviation from the desired state of charge of the storage, which we define as $\frac{1}{2}$ by convention. The fourth term is a penalization (through a convex loss function l_1) of the deviation of the power supplied by the electrical public grid $\mathrm{P}^{\mathrm{grid},u}$ from the commitment profile $\mathrm{P}^{\mathrm{grid,com.}}$. If the later were exogenously given, it would take the form of standard stochastic control problem. In our model, it is endogenous and we set

$$\mathrm{P}_t^{\mathrm{grid,com.}} = \mathbb{E}\left[\mathrm{P}_t^{\mathrm{grid},u}\right]. \tag{3.1.1}$$

This choice is inspired by the quadratic case for l_1: indeed, solving the optimal stochastic control for a given $\mathrm{P}^{\mathrm{grid,com.}}$, and then minimizing the resulting cost functional over $\mathrm{P}^{\mathrm{grid,com.}}$ would lead to (3.1.1), as the reader can easily check. Doing so, we obtain a stochastic control problem of MKV type, see later.

Going back to the applications, once identified the optimal control $(u_t : 0 \leq t \leq T)$, the consumer can commit to the profile $\mathrm{P}^{\mathrm{grid,com.}}$ as in (3.1.1) and then execute the optimal command, so that the variability of its residual consumption on the electrical public grid is minimized in a consistent way. On the side of the electricity supplier on the electrical public grid, since the consumption is smoothed, the operating costs are lower and the use of "brown" generation units can be likely avoided. We shall highlight that presumably, good loss functions l_1 should penalize more the consumption exceedance than the consumption deficit: indeed, exceedance possibly requires the use of extra production units with high carbon footprint, this is clearly to discard as often as possible. A typical example of loss function would be:

$$l_1(x) = \alpha x^2 + \alpha_+ \max(x, 0)^2; \tag{3.1.2}$$

see Figure 3.1 for an example with $\alpha_+ = 1, \alpha = 1$. This choice is somehow related to generalized risk measures accounting for both left and right tails of the distribution, such as expectiles [Bel+14]. Another point to stress is the need to account for jumps in the production/consumption dynamics – i.e. the consumption might have discontinuities as appliances/devices are switched-on/off, the power production by a solar panel might suddenly drop to zero if a cloud hides the sun. To summarize, in order to fit application needs, we shall consider non quadratic loss functions and a probabilistic setting of general filtration (allowing jumps).

MKV stochastic control problems: background results. We embed the previous example in a more general setting:

$$\left.\begin{aligned} &\mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t, \omega, u_t, X_t^u, \mathbb{E}\left[g(t, \omega, u_t, X_t^u)\right]\right)\mathrm{d}t + \psi\left(\omega, X_T^u, \mathbb{E}\left[k(\omega, X_T^u)\right]\right)\right] \\ &\text{s.t. } X_t^u = x_0 + \int_0^t \phi(s, \omega, u_s, X_s^u)\mathrm{d}s. \end{aligned}\right\} \longrightarrow \min_u . \tag{3.1.3}$$

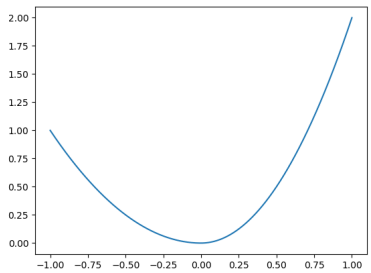


Figure 3.1: Loss function l_1 penalizing more the consumption exceedance

The functions l, g, ψ, k, ϕ depend on time, control, state variable and on the ambient randomness ω, precise assumptions are given later. Note that the control only appears in the drift of the state variable: we could also have considered a more general model $X_t^u = x + \int_0^t \phi(s, \omega, u_s, X_s^u)\mathrm{d}s + Z_t$ where Z is càdlàg semi-martingale (independent of u), but actually, this extended model is equivalent to the current one by setting $\widetilde{X_t^u} = X_t^u - Z_t$ as a new state variable and by adjusting the (already random) coefficients. Besides, note that the above dynamics for X^u is compatible with usual battery dynamics [Hus+07], like for example models of the form

$$\frac{\mathrm{d}\ \text{State of charge}}{\mathrm{d}t} = \text{constant} \cdot \text{Battery power}. \tag{3.1.4}$$

The problem (3.1.3) is of McKean-Vlasov (MKV) type since the distribution of (u, X^u) enters into the functional cost. But since this is through generalized moments via the functions g and k, the interactions are so-called scalar, which avoids to use the notion of derivatives with respect to probability measures, while maintaining some interesting flexibility. For a full account on control of Stochastic Differential Equations (SDE for short) of MKV type and the link with Mean Field Games, see the recent books [CD18] and in particular Chapter 6 of Volume I. However, in the above reference, only the distribution of SDE enters in the coefficients, not that of the control as in our setting. We refer to this more general setting as *extended* MKV stochastic optimal control.

Studies in such an extended framework are quite unusual in the literature. In [PW16], the general discrete case is studied. In [Yon13] and very recently in [BP18], both the probability distributions of the state and control variables appear in the dynamic of the state and the cost function, but only through their first and second order moments (Linear-Quadratic problems, LQ for short). In [PW18], the cost functional and the dynamic depend both on the joint probability distribution of the state and control variables, but the authors consider closed-loop controls, which allows them to consider the probability distribution of the state variable only: in our setting, *we do not make any Markovian assumptions* for the characterization of the optimal control. During the preparation of this work (started in 2016), we have been aware of the recent preprint [ABVC19] which deals also with the extended MKV stochastic optimal control, with fully non-linear interaction, Markovian dynamics, in the case of a Brownian filtration.

Our contributions. As a difference with the previous references, we do not restrict ourselves to the LQ setting, we deal with extended MKV stochastic optimal control, without Markovian assumptions, and we do not assume that the underlying filtration is Brownian (allowing jump processes). Besides, apart "expected" results about existence/uniqueness, we provide some numerical approximations by using some perturbations analysis around the LQ case. We shall insist that MKV stochastic control is a very recent field and numerical methods are still in their infancy; see [Ang+19] for a scheme based on tree methods for solving some MKV Forward-Backward SDE (FBSDE for short) that characterize optimal stochastic controls. Our perturbation approach is different from theirs. As a consequence, we design an effective numerical scheme to address the problem raised by the optimal management of storage facilities able to reduce the variability of residual electricity consumption on the electrical public grid, in the context of uncertain production/consumption of an aggregated consumer. This presumably opens the door to a wider use of these approaches in real smart grid applications.

Now let us go into the details of mathematical/computational arguments. For characterizing the optimal control, we follow a quite standard methodology (see e.g.[CD15]), although details are quite different. This is made in three steps: necessary first order conditions, which become sufficient under additional convexity assumptions, existence

of solutions to the first order equations. The derivation of the first order conditions follows the stochastic Pontryagin principle, see for instance [Ben88; Pen90; CD15]. This is achieved for general running and terminal cost functions. In particular, to account for jumps in the production/consumption dynamics, our mathematical analysis is performed in the context of general filtration. It gives rise to an optimality system (see Theorems 3.2.2 and 3.2.3), composed of a forward degenerate SDE and of a backward SDE (the adjoint equation), with possibly discontinuous martingale term, and an optimality condition linking the values and probability laws of the state and control variables with the adjoint variable.

In Section 3.2.4, we establish that this system of equations has a unique solution under some regularity conditions, an invertibility assumption and for small time horizon T (see Theorem 3.2.4). The condition on T is quite explicit from the proof, which makes the verification on practical examples easy. Here the proof has to be specific and restricted to small time because of non-Brownian filtration and of non-Markovian dynamics: indeed, we can not invoke neither a drift-monotony condition, as in [PT99], nor a non-degeneracy condition as in [DG06]. In Section 3.2.5, we discuss how the unique solution to the first order condition may or may not be the optimal solution; we provide a counter-example (Proposition 3.2.6), which is interesting for its own, we believe that this kind of situation is already known but we could not find an appropriate reference.

Then we show in Section 3.2.7 that the necessary optimality conditions established in Theorem 3.2.3 become sufficient if we assume some convexity conditions on the Hamiltonian and the terminal cost. We shall highlight that the usual Hamiltonian [CD15] (when the distribution of the control is not optimized) can not match with our framework; alternatively, we define a version in expectation (Lemma 3.2.9). The final optimality result is stated in Theorem 3.2.10.

In Section 3.3, we exemplify our study to the toy model presented in introduction, motivated by practical applications to smart grid management. To get a tractable and effective solution, we perform a perturbation approach around the LQ case. We establish error bounds and as an approximation, we select the expansion with the second order error terms. Numerical experiments illustrate the performance and accuracy of the method, as well the behavior on the optimally controlled system.

Long and technical proofs are postponed to Section 3.4 in order to smooth the reading.

Notations. We list the most common notations used in all this work.

▷ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers and positive integers. The notation $|x|$ stands for the Euclidean norm of a vector x, without further reference to its dimension. For a given matrix $A \in \mathbb{R}^p \otimes \mathbb{R}^d$, $A^\top$ refers to its transpose. Its norm is that induced by the Euclidean norm, i.e. $|A| := \sup_{x\in\mathbb{R}^d, |x|=1} |Ax|$. Recall that $|A^\top| = |A|$. For $p \in \mathbb{N}^*$, Id_p stands for the identity matrix of size $p \times p$.

▷ *Functions, derivatives.* When a function (or a process) ψ depends on time, we write indifferently $\psi_t(z)$ or $\psi(t, z)$ for the value of ψ at time t, where z represents all other arguments of ψ.
For a smooth function $g : \mathbb{R}^q \mapsto \mathbb{R}^p$, g_x represents the Jacobian matrix of g with respect to x, i.e. the matrix $(\partial_{x_j} g_i)_{i,j} \in \mathbb{R}^p \otimes \mathbb{R}^q$. However, a subscript x_t refers to the value of a process x at time t (and not to a partial derivative with respect to t). We also introduce $\nabla_x f := f_x^\top$.

▷ *Probability.* To model the random uncertainty on the time interval $[0, T]$ ($T > 0$ fixed), we consider a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$, we assume that the filtration $\{\mathcal{F}_t\}_{0\le t\le T}$ is right-continuous, augmented with the $\mathbb{P}$-null sets. For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$. Denote by $\mathcal{P}$ the σ-field of predictable sets of $[0, T] \times \Omega$.
All the quantities impacted by the control u are upper-indexed by u, like Z^u for instance.
As usually, càdlàg processes stand for processes that are right continuous with left-hand limits. All the martingales are considered with their càdlàg modifications.

▷ *Spaces.* Let $k \in \mathbb{N}^*$. We define $\mathbb{L}^2([0, T], \mathbb{R}^k)$ (resp. $\mathbb{L}^\infty([0, T], \mathbb{R}^k)$) as the Banach space of deterministic functions f on $[0, T]$ with values in $\mathbb{R}^k$ such that $\int_0^T |f_t|^2 \mathrm{d}t < +\infty$ (resp. $\sup_{t\in[0,T]} |f(t)| < +\infty$). Since the arrival space $\mathbb{R}^k$ will be unimportant, we will skip the reference to it in the notation and write the related norms as

$$\|f\|_{\mathbb{L}^2_T} := \left(\int_0^T |f(t)|^2 \mathrm{d}t \right)^{\frac{1}{2}}, \qquad \|f\|_{\mathbb{L}^\infty_T} := \sup_{t\in[0,T]} |f(t)|.$$

Let $p \geq q \geq 1$. The Banach space of $\mathbb{R}^k$-valued random variables X such that $\mathbb{E}\left[|X|^p\right] < +\infty$ is denoted by $\mathbb{L}^p(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^p_\Omega$; the associated norm is

$$\|X\|_{\mathbb{L}^p_\Omega} := \mathbb{E}\left[|X|^p\right]^{\frac{1}{p}}.$$

The Banach space $\mathbb{H}^{p,q}([0,T] \times \Omega, \mathbb{R}^k)$ (resp. $\mathbb{H}^{p,q}_{\mathcal{P}}([0,T] \times \Omega, \mathbb{R}^k)$) is the set of all $\mathbb{F}$-progressively measurable (resp. $\mathbb{F}$-predictable) processes $\psi : [0,T] \times \Omega \to \mathbb{R}^k$ such that $\int_0^T \mathbb{E}\left[|\psi_t|^q\right]^{p/q} \mathrm{d}t < +\infty$. Here again we will omit the reference to $\mathbb{R}^k$, which will be clear from the context. The associated norm is

$$\|\psi\|_{\mathbb{H}^{p,q}} := \left(\int_0^T \mathbb{E}\left[|\psi_t|^q\right]^{p/q} \mathrm{d}t\right)^{\frac{1}{p}}.$$

The Banach space $\mathbb{H}^{\infty,q}\left([0,T] \times \Omega, \mathbb{R}^k\right)$ stands for the elements of $\mathbb{H}^{p,q}\left([0,T] \times \Omega, \mathbb{R}^k\right)$ satisfying $\sup_{t\in[0,T]} \mathbb{E}\left[|\psi_t|^q\right] < +\infty$, and the related norm is

$$\|\psi\|_{\mathbb{H}^{\infty,q}([0,T]\times\Omega,\mathbb{R}^k)} := \sup_{t\in[0,T]} \mathbb{E}\left[|\psi_t|^q\right]^{\frac{1}{q}}.$$

We shall most often consider $p = q = 2$.

3.2 Stochastic control and MKV-FBSDEs

The aim is to analyze the control problem, about minimizing (3.1.3). We first discuss the smart grid setting and the class of admissible controls u; second we derive the first-order condition (Pontryagin principle) which writes as a MKV-FBSDE; third we derive sufficient conditions for the existence and uniqueness to the above; fourth in the absence of convexity conditions we provide a counter-example to optimality; last, with suitable convexity assumptions we establish that the MKV-FBSDE solution characterizes the optimal control.

3.2.1 Stochastic model and smart grid framework

As explained in introduction, (3.1.3) may describe the optimal energy management of an aggregated consumer, with storage facilities (e.g. battery), with his own RE production (e.g. building equipped with solar panel), with a connection to the electrical public grid. The management horizon T is typically short, e.g. 24 hours for reasons explained in introduction.

The control is made through a $\mathbb{R}^d$-valued vector process $u = (u_t : 0 \leq t \leq T)$, $d \in \mathbb{N}^*$. We consider u as a $\mathcal{F}_t$-predictable process in $\mathbb{H}^{2,2}_{\mathcal{P}}$: the intuition behind it is that decisions occurring at time t have to be made in accordance with the information available up to this time. This is coherent with the smart grid application. In particular, there has to be a slight delay between sudden events and the decisions taken by the controller, whence the predictability assumption.

The dynamics of the system are represented by a $\mathbb{R}^p$-valued state variable, denoted by X, which satisfies the following ODE

$$X^u_t = x_0 + \int_0^t \phi(s, \omega, u_s, X^u_s)\mathrm{d}s. \tag{3.2.1}$$

The state variable can include various information in the smart grid application, like for example the state of charge of the battery (see (3.1.4)), the PV production, the building electricity consumption, etc. Moreover, the possible dependence in time of $\phi(\cdot)$ is a degree of freedom suitable to account for energy losses over time or aging of the battery, both impacting the state of charge of the battery.

The cost functional is described by $\mathcal{J}(u)$, given in (3.1.3). In the smart grid application, Markovian-type costs would take the form, for instance, $l(t, \omega, u, x, \bar{g}) = \tilde{l}(t, Z_t(\omega), u, x, \bar{g})$ where Z would represent a multidimensional stochastic factor modeling the evolution of the exogenous uncontrolled variables (weather, consumption...), but we

also allow non Markovian models. In the sequel, we omit ω when we write terms inside $\mathcal{J}(u)$ and X^u, since it is now clear that we deal with random coefficients. All in all, the optimal control problem we study is

$$\left.\begin{aligned} \mathcal{J}(u) &:= \mathbb{E}\left[\int_0^T l\left(t, u_t, X_t^u, \mathbb{E}\left[g(t, u_t, X_t^u)\right]\right) \mathrm{d}t + \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right] \\ \text{s.t. } X_t^u &= x_0 + \int_0^t \phi(s, u_s, X_s^u)\mathrm{d}s. \end{aligned}\right\} \longrightarrow \min_{u \in \mathbb{H}_{\mathcal{P}}^{2,2}}. \tag{3.2.2}$$

Last, we summarize the coefficients from the toy example described p.48.

Example 3.2.1 (Smart grid toy example)**.** *Let $\mathrm{P}^{\mathtt{load}}$ be the difference between the instantaneous consumer local consumption and his RE production: we assume this is a process in $\mathbb{H}^{2,2}([0,T]\times\Omega,\mathbb{R})$. The control $u \in \mathbb{H}_{\mathcal{P}}^{2,2}([0,T]\times\Omega,\mathbb{R})$ corresponds to the power supplied by the battery, while the state X^u corresponds to the normalized state of charge of the battery which dynamics is linear with respect to the control u, see [Hey+15]:*

$$X_t^u = x_0 - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s \mathrm{d}s.$$

If $\mathrm{P}^{\mathtt{grid},u}$ is the power supplied by the electrical public grid, the power balance imposes that

$$\mathrm{P}_{t-}^{\mathtt{load}} = \mathrm{P}_{t-}^{\mathtt{grid},u} + u_t.$$

Then set $d = p = 1$ and

$$\begin{aligned} l(t,\omega,u,x,\bar{g}) &:= \mathrm{C}_{t-}(\omega)\,(\mathrm{P}_{t-}^{\mathtt{load}}(\omega) - u) + \frac{\mu_t}{2}u^2 + \frac{\nu_t}{2}(x - \frac{1}{2})^2 + l_1(\mathrm{P}_{t-}^{\mathtt{load}}(\omega) - u - \bar{g}), \\ g(t,\omega,u,x) &:= \mathrm{P}_{t-}^{\mathtt{load}}(\omega) - u, \\ \psi(\omega,x,\bar{k}) &:= \frac{\gamma}{2}(x - \frac{1}{2})^2, \\ k(\omega,x) &:= 0, \\ \phi(t,u,x) &:= -\frac{u}{\mathcal{E}_{\max}}. \end{aligned} \tag{3.2.3}$$

The time-dependent coefficients μ_t and ν_t give the flexibility to include hourly effect in the management. We recall that the convex loss function l_1 may take the form (3.1.2). Considering the left-hand limit $t-$ in the above definitions is a technicality to fulfill the following assumptions.

3.2.2 Standing assumptions

From now on, we assume the following hypotheses hold. When we refer to a *constant*, we mean a *finite deterministic constant*.

(H.x) $x_0 \in \mathbb{L}_\Omega^2$ and is $\mathcal{F}_0$-measurable.

(H.l) $l : (t,\omega,u,x,\bar{g}) \in [0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^q \mapsto l(t,\omega,u,x,\bar{g}) \in \mathbb{R}$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)$-measurable. Furthermore, $l(\cdot,\cdot,0,0,0) \in \mathbb{H}^{1,1}$, l is continuously differentiable in $(u,x,\bar{g})$ with the growth condition

$$|\nabla_u l(t,\omega,u,x,\bar{g})| + |\nabla_x l(t,\omega,u,x,\bar{g})| + |\nabla_{\bar{g}} l(t,\omega,u,x,\bar{g})| \le C\,(|u| + |x| + |\bar{g}|) + C_l^{(0)}(t,\omega)$$

for any $(t,u,x,\bar{g}) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^q$ a.s., for some constant C and some random process $C_l^{(0)}$ in $\mathbb{H}^{2,2}$.

(H.g) $g : (t,\omega,u,x) \in [0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^p \mapsto g(t,\omega,u,x) \in \mathbb{R}^q$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $g(\cdot,\cdot,0,0) \in \mathbb{H}^{2,1}$, g is continuously differentiable in (u,x) and there exist constants $C_{g,u}$ and $C_{g,x}$ such that

$$|\nabla_x g(t,\omega,u,x)| \le C_{g,x} \qquad \text{and} \qquad |\nabla_u g(t,\omega,u,x)| \le C_{g,u}$$

for any $(t,u,x) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^p$ a.s. .

(H.ψ) $\psi : (\omega, x, \bar{k}) \in \Omega \times \mathbb{R}^p \times \mathbb{R}^r \mapsto \psi(\omega, x, \bar{k}) \in \mathbb{R}$ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^r)$-measurable. Furthermore, $\psi(\cdot, 0, 0) \in \mathbb{L}^1_\Omega$, ψ is continuously differentiable in $(x, \bar{k})$ and the growth condition

$$|\nabla_x \psi(\omega, x, \bar{k})| + |\nabla_{\bar{k}} \psi(\omega, x, \bar{k})| \leq C\,(|x| + |\bar{k}|) + C^{(0)}_\psi(\omega)$$

holds for any $(x, \bar{k}) \in \mathbb{R}^p \times \mathbb{R}^r$ a.s., for some constant C and some random variable $C^{(0)}_\psi$ in $\mathbb{L}^2_\Omega$.

(H.k) $k : (\omega, x) \in \Omega \times \mathbb{R}^p \mapsto k(\omega, x) \in \mathbb{R}^r$ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $k(\cdot, 0) \in \mathbb{L}^1_\Omega$, k is continuously differentiable in x and there exists a constant $C_{k,x}$ such that

$$|\nabla_x k(\omega, x)| \leq C_{k,x}$$

holds for any $x \in \mathbb{R}^p$ a.s..

(H.ϕ) $\phi : (t, \omega, u, x) \in [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}^p \mapsto \phi(t, \omega, u, x) \in \mathbb{R}^p$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $\phi(\cdot, \cdot, 0, 0) \in \mathbb{H}^{2,2}$, ϕ is continuously differentiable in (u, x) and there exist constants $C_{\phi,u}$ and $C_{\phi,x}$ such that

$$|\nabla_u \phi(t, \omega, u, x)| \leq C_{\phi,u} \qquad \text{and} \qquad |\nabla_x \phi(t, \omega, u, x)| \leq C_{\phi,x}$$

hold for any $(t, u, x) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^p$ a.s..

It is easy to check these conditions in Example 3.2.1.

As a consequence of **(H.ϕ)**, the dynamics of X^u in (3.2.1) writes as a ODE with Lipschitz-continuous stochastic coefficient: the uniqueness and existence stem from the Cauchy existence theorem for ODE, applied ω by ω. In addition, we easily show

$$|X^u_t| \leq |x_0| + \int_0^t \left(|\phi(s, 0, 0)| + C_{\phi,u}|u_s| + C_{\phi,x}|X^u_s|\right) \mathrm{d}s \leq C_T \left(|x_0| + \int_0^t \left(|\phi(s, 0, 0)| + C_{\phi,u}|u_s|\right) \mathrm{d}s\right)$$

where the second inequality comes from Gronwall's lemma. Then one directly shows that, since u and $\phi(\cdot, 0, 0)$ are in $\mathbb{H}^{2,2}$, X^u is in $\mathbb{H}^{\infty,2} \subset \mathbb{H}^{2,2}$. Then, a careful inspection of the assumptions **(H.l)**-**(H.g)**-**(H.ψ)**-**(H.k)** shows that it implies that the cost $\mathcal{J}(u)$ is finite.

3.2.3 Necessary condition for optimality

For admissible controls u and v, we now provide a representation of the derivative

$$\dot{\mathcal{J}}(u, v) = \partial_\varepsilon \mathcal{J}(u + \varepsilon v)|_{\varepsilon=0},$$

using an adjoint process Y^u.

Theorem 3.2.2 (Gâteaux derivatives)**.** *Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and set $\bar{g}^u_t := \mathbb{E}\left[g(t, u_t, X^u_t)\right]$. Let $\tilde{L}^u$ be the unique solution of*

$$\tilde{L}^u_0 = \mathrm{Id}_p, \qquad \frac{\mathrm{d}\tilde{L}^u_t}{\mathrm{d}t} = \tilde{L}^u_t \nabla_x \phi(t, u_t, X^u_t).$$

Then $\tilde{L}^u$ is invertible and its inverse satisfies (see Lemma 3.4.1)

$$(\tilde{L}^u_0)^{-1} = \mathrm{Id}_p, \qquad \frac{\mathrm{d}(\tilde{L}^u_t)^{-1}}{\mathrm{d}t} = -\nabla_x \phi(t, u_t, X^u_t)(\tilde{L}^u_t)^{-1}.$$

Define also $L^u := ((\tilde{L}^u)^{-1})^\top$. The following $\mathbb{R}^p$-valued process Y^u is well defined as a càdlàg process in $\mathbb{H}^{\infty,2}$:

$$\begin{aligned} Y^u_t = {} & \mathbb{E}_t\left[(\tilde{L}^u_t)^{-1}\tilde{L}^u_T\left(\nabla_x \psi\left(X^u_T, \mathbb{E}\left[k(X^u_T)\right]\right) + \nabla_x k(X^u_T)\mathbb{E}\left[\nabla_{\bar{k}}\psi\left(X^u_T, \mathbb{E}\left[k(X^u_T)\right]\right)\right]\right)\right] \\ & + \mathbb{E}_t\left[\int_t^T (\tilde{L}^u_t)^{-1}\tilde{L}^u_s\left(\nabla_x l(s, u_s, X^u_s, \bar{g}^u_s) + \nabla_x g(s, u_s, X^u_s)\mathbb{E}\left[\nabla_{\bar{g}} l(s, u_s, X^u_s, \bar{g}^u_s)\right]\right) \mathrm{d}s\right]. \end{aligned} \tag{3.2.4}$$

In particular, there exists a $\mathbb{R}^p$-valued càdlàg martingale M^u in $\mathbb{H}^{\infty,2}$, vanishing at time 0, such that (Y^u, M^u) is the unique solution in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ of the following BSDE in (Y, M):

$$\begin{aligned} -\mathrm{d}Y_t &= \Big(\nabla_x \phi(t, u_t, X_t^u) Y_t + \nabla_x l(t, u_t, X_t^u, \bar{g}_t^u) + \nabla_x g(t, u_t, X_t^u) \mathbb{E}\left[\nabla_{\bar{g}} l(t, u_t, X_t^u, \bar{g}_t^u)\right]\Big) \mathrm{d}t - \mathrm{d}M_t, \\ Y_T &= \nabla_x \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right) + \nabla_x k(X_T^u) \mathbb{E}\left[\nabla_{\bar{k}} \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right]. \end{aligned} \tag{3.2.5}$$

Besides, for any $u, v \in \mathbb{H}^{2,2}_{\mathcal{P}}$, the directional derivative $\dot{\mathcal{J}}(u, v)$ exists and is given by

$$\dot{\mathcal{J}}(u, v) = \mathbb{E}\left[\int_0^T \left\{l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u)\right\} v_t \mathrm{d}t\right].$$

The proof is postponed to Subsection 3.4.1. At the optimal control u (whenever it exists), the above derivative $\dot{\mathcal{J}}(u, v)$ must be 0, in any direction $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$. Take for instance v given by:

$$\forall \in [0, T], \qquad v_t := l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u),$$

which ensures that $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ under our assumptions. This justifies the following statement.

Theorem 3.2.3 (Necessary condition for optimality)**.** *Under the notations and assumptions of Theorem 3.2.2, if a control $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ is optimal, then there exists a unique couple $\left(X^u, Y^u\right) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ fulfilling (3.2.1) and (3.2.4) such that*

$$l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u) = 0 \tag{3.2.6}$$

holds $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.e.

3.2.4 Solvability of the MKV Forward-Backward SDE

Our aim is now to provide sufficient conditions to ensure existence of solution to the system of forward-backward equations (3.2.1)-(3.2.4)-(3.2.6), which we call MKV-FBSDE. For this, we strengthen previous assumptions.

(H.l.2) (H.l) holds and there exist constants $C_{l_{*,\star}}$ where $*$ stands for x and $\bar{g}$, and $\star$ stands for u, x or $\bar{g}$ such that:

$$\begin{aligned} |\nabla_x l(t, \omega, u_1, x_1, \bar{g}_1) - \nabla_x l(t, \omega, u_2, x_2, \bar{g}_2)| &\leq C_{l_{x,u}} |u_1 - u_2| + C_{l_{x,x}} |x_1 - x_2| + C_{l_{x,\bar{g}}} |\bar{g}_1 - \bar{g}_2|, \\ |\nabla_{\bar{g}} l(t, \omega, u_1, x_1, \bar{g}_1) - \nabla_{\bar{g}} l(t, \omega, u_2, x_2, \bar{g}_2)| &\leq C_{l_{\bar{g},u}} |u_1 - u_2| + C_{l_{\bar{g},x}} |x_1 - x_2| + C_{l_{\bar{g},\bar{g}}} |\bar{g}_1 - \bar{g}_2| \end{aligned}$$

holds for any $(u_1, u_2, x_1, x_2, \bar{g}_1, \bar{g}_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$, $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e..

(H.g.2) (H.g) holds and g is affine-linear in x, of the form $g(t, u, x) = a_t^{(g)} x + b^{(g)}(t, u)$.

(H.ψ.2) (H.ψ) holds and there exist constants $C_{\psi_{*,\star}}$ where $*$ and $\star$ stand for x or $\bar{k}$ such that:

$$\begin{aligned} |\nabla_x \psi(x_1, \bar{k}_1) - \nabla_x \psi(x_2, \bar{k}_2)| &\leq C_{\psi_{x,x}} |x_1 - x_2| + C_{\psi_{x,\bar{k}}} |\bar{k}_1 - \bar{k}_2|, \\ |\nabla_{\bar{k}} \psi(x_1, \bar{k}_1) - \nabla_{\bar{k}} \psi(x_2, \bar{k}_2)| &\leq C_{\psi_{\bar{k},x}} |x_1 - x_2| + C_{\psi_{\bar{k},\bar{k}}} |\bar{k}_1 - \bar{k}_2| \end{aligned}$$

holds for any $(x_1, x_2, \bar{k}_1, \bar{k}_2) \in \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^r \times \mathbb{R}^r$, $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e..

(H.k.2) (H.k) holds and k is affine-linear in x, of the form $k(x) = a^{(k)} x + b^{(k)}$.

(H.ϕ.2) (H.ϕ) holds and the dynamic of X^u is affine-linear in x, given by $\phi(t, u, x) = a_t^{(\phi)} x + b^{(\phi)}(t, u)$.

Observe again that this set of conditions is consistent with Example 3.2.1. We now aim at establishing the solvability of the system composed of (3.2.1), (3.2.5) and (3.2.6). We are going to show that this system has a unique solution for a small enough time horizon T, hence the existence and uniqueness of a solution to the optimal control problem, under the sufficient conditions of Theorem 3.2.10.

Theorem 3.2.4. *Assume* (H.l.2)-(H.g.2)-(H.ψ.2)-(H.k.2)-(H.ϕ.2) *hold. Assume furthermore that there exists a* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)\otimes\mathcal{B}(\mathbb{R}^q)$*-measurable function* $h:(t,\omega,x,y,\bar{g},\bar{\lambda})\mapsto h(t,\omega,x,y,\bar{g},\bar{\lambda})\in\mathbb{R}^d$ *such that*

$$\begin{aligned}&l_u(t,u_t,X_t^u,\bar{g}_t^u)+\mathbb{E}\left[l_{\bar{g}}(t,u_t,X_t^u,\bar{g}_t^u)\right]g_u(t,u_t,X_t^u)+(Y_{t-}^u)^\top\phi_u(t,u_t,X_t^u)=0,\quad d\mathbb{P}\otimes\mathrm{d}t-a.e.\\ &\iff u_t=h\left(t,X_t^u,Y_{t-}^u,\bar{g}_t^u,\mathbb{E}\left[\nabla_{\bar{g}}l(t,\omega,u_t,X_t^u,\bar{g}_t^u)\right]\right),\quad d\mathbb{P}\otimes\mathrm{d}t-a.e..\end{aligned}\tag{3.2.7}$$

If h *is Lipschitz continuous in* $(x,y,\bar{g},\bar{\lambda})$*, with Lipschitz constants denoted by* $C_{h,x},C_{h,y},C_{h,\bar{g}},C_{h,\bar{\lambda}}$*, and if* $\left(h(t,\omega,0,0,0,0)\right)_{(t,\omega)\in\mathcal{P}}\in\mathbb{H}^{2,2}_{\mathcal{P}}$*,*

$$\Theta:\begin{cases}\mathbb{H}^{2,2}_{\mathcal{P}} & \to\mathbb{H}^{2,2}_{\mathcal{P}}\\ u & \mapsto\tilde{u}\end{cases},$$

where

$$\Theta(u)_t:=\tilde{u}_t=h\left(t,\omega,X_t^u,Y_{t-}^u,\bar{g}_t^u,\mathbb{E}\left[\nabla_{\bar{g}}l(t,u_t,X_t^u,\bar{g}_t^u)\right]\right),\quad \mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e.,$$

is well defined and Lipschitz continuous. If moreover,

$$C_{h,\bar{g}}C_{g,u}+C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u}+C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)<1,\tag{3.2.8}$$

then for T *small enough,* Θ *is a contraction and therefore has a unique fixed point* $u^\star$*. In that case, there exists a unique* $u\in\mathbb{H}^{2,2}_{\mathcal{P}}$ *satisfying* (3.2.1)-(3.2.5)-(3.2.6) *and* $u=u^\star$*.*

The proof is available in Subsection 3.4.2. Regarding the proof of a fixed point when the time interval $[0,T]$ is arbitrary large, observe that, as a difference with [PT99] and [DG06] for instance, in our setting we cannot rely on a monotony condition of the drifts, nor a non-degeneracy condition. This is why we shall restrict to small time condition.

Remark 3.2.5. *If one can exhibit a* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)\otimes\mathcal{B}(\mathbb{R}^q)$*-measurable function* h *such that for all* $(\tilde{u},x,y,\bar{g},\bar{\lambda})\in\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^q$*:*

$$\begin{aligned}&d\mathbb{P}\otimes\mathrm{d}t-a.e.,\ l_u(t,\omega,\tilde{u},x,\bar{g})+\bar{\lambda}^\top g_u(t,\omega,\tilde{u},x)+y^\top\phi_u(t,\omega,\tilde{u},x)=0\\ &\iff d\mathbb{P}\otimes\mathrm{d}t-a.e.,\ \tilde{u}=h\left(t,\omega,x,y,\bar{g},\bar{\lambda}\right),\end{aligned}$$

then (3.2.7) *is satisfied with the same function* h*.*

3.2.5 Existence and uniqueness of critical point do not necessarily imply existence of a minimum

If there exists a unique solution to the first order optimality condition (unique critical point), and under other assumptions like continuity, growth properties, it is tempting to conclude that this point is a minimum. However, this is not necessarily the case in infinite dimension. This section aims at clarifying this fact by providing an example[2] where continuity, coercivity and unique critical point are ensured, but without existence of minimum. Therefore, extra conditions are necessary to get the existence of a minimum, see later the discussion in Section 3.2.6.

Proposition 3.2.6. *Set*

$$F:\begin{cases}\mathbb{L}^2_1:=\mathbb{L}^2([0,1],\mathbb{R}) & \mapsto\mathbb{R}\\ u & \mapsto(\|u\|^2_{\mathbb{L}^2_1}-1)^2+\int_0^1 t|u_t|^2\mathrm{d}t.\end{cases}$$

Then F *satisfies the following properties:*

[2] Such examples might exist in the literature, but we have not been able to find a reference for this.

1. ***Continuity:*** *F is continuous*
2. ***Coercivity:*** *$F(u)$ tends to $+\infty$ when $\|u\|_{\mathbb{L}^2_1}$ tends to $+\infty$*
3. ***Existence and uniqueness of critical point:*** *F is Gateaux-differentiable and has a unique critical point.*

However, F does not have a minimum.

The proof is postponed to Subsection 3.4.3. The function F defined in this example cannot be quasi-convex (and a fortiori F cannot be convex), since it would then have a minimum, as stated in the next section.

3.2.6 Existence of an optimal control

We now give sufficient conditions for the existence of an optimal control, i.e. existence of a minimizer of $\mathcal{J}$. In such a favorable case, and if the necessary optimality conditions (3.2.1)-(3.2.4)-(3.2.6) have a unique solution u^*, then u^* is the unique minimum of $\mathcal{J}$. We start with a general result.

Theorem 3.2.7. *Let E be a reflexive Banach space, let $F : E \to \mathbb{R}$ be a lower semi-continuous, quasi-convex function which satisfies the coercivity condition $\lim_{\|u\|_E \to +\infty} F(u) = +\infty$. Then F has a minimum on E.*

Proof. We adapt the arguments of [Bre10a, Corollary 3.23, pp. 71], where the operator considered is assumed to be continuous and convex. However, the hypothesis can be relaxed to lower semi-continuity and quasi-convexity of the function F, since we only need closedness and convexity of the sub-level sets $\Gamma_\alpha^{(F)} := \{u \in E | F(u) \leq \alpha\}$ for all $\alpha \in \mathbb{R}$. □

Let us add a few comments. In the finite dimensional case, any lower semi-continuous and coercive function has a minimum (since any closed and bounded set is compact). In the infinite dimensional case, the example in Subsection 3.2.5 illustrates that this may be not the case without the quasi-convexity assumption. Besides, note that without the coercivity condition, the existence of minimum may not hold, even in finite dimension (take $E = \mathbb{R}$ and $F(x) = \exp(x)$). Moreover, without the lower semi-continuity of F, the result may fail as well (take $F : (-\infty, 0] \mapsto \mathbb{R}$ defined by $F(x) = |x|\mathbf{1}_{x<0} + \mathbf{1}_{x=0}$, which is coercive and convex).

Apply the previous result with $E = \mathbb{H}^{2,2}$ and $F = \mathcal{J}$: E is an Hilbert space, thus a reflexive Banach space. The functional $\mathcal{J}$ is continuous, hence lower semi-continuous. Therefore, we have proved the following.

Corollary 3.2.8. *Assume that $\mathcal{J}$ defined in (3.2.2) is quasi-convex and that $\lim_{\|u\|_{\mathbb{H}^{2,2}} \to +\infty} \mathcal{J}(u) = +\infty$. Then the optimal control problem has a solution $u^* \in \mathbb{H}^{2,2}_{\mathcal{P}}$.*

3.2.7 Sufficient condition for optimality

Let us now give conditions under which the necessary optimality conditions are sufficient. Additionally to **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**, we assume the following conditions.

(Conv)

1. The mapping $\mathcal{T} : \begin{cases} \mathbb{L}^2_\Omega & \to \mathbb{R} \\ X & \mapsto \mathbb{E}\left[\psi(X, \mathbb{E}\left[k(X)\right])\right] \end{cases}$ is convex.
2. The mapping $\mathcal{I} : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} & \to \mathbb{R} \\ (\tilde{u}, X) & \mapsto \int_0^T \mathbb{E}\left[l\left(t, \tilde{u}_t, X_t, \mathbb{E}\left[g\left(t, \tilde{u}_t, X_t\right)\right]\right)\right] \mathrm{d}t \end{cases}$ is convex.
3. The mapping: $\phi : \begin{cases} [0,T] \times \mathbb{R}^d \times \mathbb{R}^p & \to \mathbb{R}^p \\ (t, u, X) & \mapsto \phi(t, u, X) \end{cases}$ is affine-linear in (u, X).

Lemma 3.2.9. *Under (Conv), $\mathcal{J}$ is convex. If furthermore, $\mathcal{I}$ is strictly convex in $\tilde{u}$, or $\mathcal{I}$ is strictly convex in X and ϕ_u has full column rank (which implies $p \geq d$) for almost every t in $[0,T]$, then $\mathcal{J}$ is strictly convex.*

Proof. Under the assumption on ϕ, $u \mapsto X^u$ is affine-linear. This yields the first result using the fact that a composition of an affine-linear function by a convex function is convex. If $\mathcal{I}$ is strictly convex in u then so is $\mathcal{J}$. If ϕ_u has full column rank, $u \mapsto X^u$ is an affine-linear injection and if besides $\mathcal{I}$ is strictly convex in X, we get that $\mathcal{J}$ is strictly convex. □

Let us emphasize the difference with usual stochastic maximum principle (when distributions do not enter in the cost functions). In that case, i.e. without the dependence w.r.t. $\mathbb{E}[g(t, \tilde{u}_t, X_t)]$ of the running cost and w.r.t. $\mathbb{E}[k(X)]$ of the terminal cost, the sufficient optimality condition is the affine-linearity in (u, X) of ϕ, the *point-wise* convexity in (u, x) of

$$(t, u, x) \mapsto l(t, u, x),$$

for any t and the *point-wise* convexity of ψ in x.
In the current MKV setting, it would be tempting to require:

$$\xi : (t, u, x) \mapsto l(t, u, x, \mathbb{E}[g(t, u, x)])$$

to be convex in $(u, X) \in \mathbb{L}^2_\Omega \times \mathbb{L}^2_\Omega$ for any t and

$$X \mapsto \psi(X, \mathbb{E}[k(X)])$$

to be convex in X in $\mathbb{L}^2_\Omega$. However, even for the simple linear-quadratic case with $d = p = q = 1$, i.e.

$$l(t, u, x, \bar{g}) = (1 + \kappa)u^2 - \kappa\bar{g}^2, \qquad g(t, u, x) = u, \qquad \phi(t, u, x) = u, \qquad \psi = 0,$$

with parameter $\kappa > 0$, this fails to be true. Indeed, denoting $\zeta(u) = \xi(t, u, x)$, we get:

$$\zeta\left(\frac{u_1 + u_2}{2}\right) - \frac{1}{2}(\zeta(u_1) + \zeta(u_2)) = \frac{1}{4}\left(\kappa(\mathbb{E}[u_1 - u_2])^2 - (1 + \kappa)(u_1 - u_2)^2\right).$$

Now if u_1 is a Bernoulli random variable with parameter $\frac{1}{2}$, and $u_2 = -u_1$, then on the set $\{\omega : u_1(\omega) = u_2(\omega) = 0\}$ of positive probability, the above equals $\frac{\kappa}{4} > 0$, which violates the convexity condition for these ω. On the contrary, $\mathbb{E}\left[\zeta(\frac{u_1+u_2}{2}) - \frac{\zeta(u_1)+\zeta(u_2)}{2}\right] \leq 0$ for $\kappa \geq 0$, and it is easy to see that $\mathbb{E}[\zeta(u)]$ is convex in u, for such κ. This discussion clarifies better why the correct convexity condition for the integrated Hamiltonian $\mathcal{I}$ or the point-wise one $\mathcal{H}$ is in expectation and not ω-wise, as stated in (**Conv**).

We now summarize all the results for having existence and uniqueness of an optimal stochastic control. This is one of the main results of this section.

Theorem 3.2.10. *Assume* (**H.x**)-(**H.g**)-(**H.l**)-(**H.k**)-(**H.ψ**)-(**H.ϕ**)-(**Conv**) *hold.*

1. *If $\mathcal{J}$ defined in* (3.2.2) *satisfies the coercivity condition:*

$$\lim_{\|u\|_{\mathbb{H}^{2,2}} \to +\infty} \mathcal{J}(u) = +\infty,$$

   then there exists an optimal control $u^\star \in \mathbb{H}^{2,2}_{\mathcal{P}}$, i.e. a minimum of $\mathcal{J}$ on $\mathbb{H}^{2,2}_{\mathcal{P}}$.

2. *$u^\star$ is an optimal control for the problem* (3.2.2) *if and only if there exists $(X^\star, Y^\star) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ such that $(u^\star, X^\star, Y^\star)$ fulfills* (3.2.1)-(3.2.5)-(3.2.6).

3. *If $\mathcal{J}$ is strictly convex, then it admits at most one minimizer.*

Proof. 1. This is a direct consequence of Theorem 3.2.8 and Lemma 3.2.9.
2. If $(u^\star, X^{u^\star}, Y^{u^\star})$ satisfies (3.2.1)-(3.2.5)-(3.2.6), then $\dot{\mathcal{J}}(u^\star, v) = 0$ for any $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ according to Theorem 3.2.2. Besides, under our assumptions, $\mathcal{J}$ is convex and therefore, for all $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and $t \in (0, 1]$,

$$\mathcal{J}(v) - \mathcal{J}(u^\star) \geq \frac{\mathcal{J}(u^\star + t(v - u^\star)) - \mathcal{J}(u^\star)}{t}.$$

By taking the limit when $t \to 0$, we obtain $\mathcal{J}(v) - \mathcal{J}(u^\star) \geq \dot{\mathcal{J}}(u^\star, v - u^\star) = 0$, hence the optimality of $u^\star$. The direct implication $\Rightarrow$ has been established in Theorem 3.2.3. □

3.3 Effective computation and approximation of battery control

3.3.1 Model/Context

For simplicity, we assume one-dimensional processes ($p = q = r = 1$), but the results can be easily extended to any dimension, since the arguments are based on the solution of Linear-Quadratic FBSDE, which are well known (see [Yon06]). Let us consider the following toy problem:

$$\min_{u\in\mathbb{H}^{2,2}_{\mathcal{P}}} \mathbb{E}\left[\int_0^T \left\{\mathtt{C}_{t-}\mathtt{P}^{\mathtt{grid},u}_{t-} + \frac{\mu}{2}u_t^2 + \frac{\nu}{2}\left(X^u_t - \frac{1}{2}\right)^2 + l\left(\mathtt{P}^{\mathtt{grid},u}_{t-} - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid},u}_{t-}\right]\right)\right\}\mathrm{d}t + \frac{\gamma}{2}\left(X^u_T - \frac{1}{2}\right)^2\right]$$
$$\text{s.t.} \begin{cases} X^u_t = x - \frac{1}{\mathcal{E}_{\max}}\int_0^t u_s\mathrm{d}s, \\ \mathtt{P}^{\mathtt{grid},u}_{t-} = \mathtt{P}^{\mathtt{load}}_{t-} - u_t. \end{cases}$$

This model is the same as the one presented in the introduction and has the same interpretation. We consider the following hypothesis:

(Toy)

1. The parameters μ, ν, γ are deterministic and satisfy $\mu > 0$, $\nu \geq 0$, $\gamma \geq 0$.
2. The mapping l is deterministic, convex, continuously differentiable with the growth condition $|l'(x)| \leq C_{l,x}(1+|x|)$ for all x, for some constant $C_{l,x} > 0$.
3. $\mathtt{P}^{\mathtt{load}} \in \mathbb{H}^{2,2}$, $C \in \mathbb{H}^{2,2}$ are $\mathbb{F}$-adapted and càdlàg.

Under assumptions **(Toy)**, **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**-**(Conv)** hold. Besides, one can show the strict convexity of $\mathcal{J}$. Then, it remains to apply Theorem 3.2.10 to conclude the following.

Proposition 3.3.1. *Under assumptions* **(Toy)**, *there exists a unique optimal control* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$. *Besides, there exist unique processes* $X^u \in \mathbb{H}^{\infty,2}$ *and* $Y^u \in \mathbb{H}^{\infty,2}$ *such that* (u, X^u, Y^u) *satisfies the following McKean-Vlasov Forward Backward SDE:*

$$\begin{cases} X^u_t = x - \frac{1}{\mathcal{E}_{\max}}\int_0^t u_s\mathrm{d}s, \\ Y^u_t = \mathbb{E}_t\left[\int_t^T \nu(X^u_s - \frac{1}{2})\mathrm{d}s + \gamma\left(X^u_T - \frac{1}{2}\right)\right], \\ \mu u_t - \mathtt{C}_{t-} - l'\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right) + \mathbb{E}\left[l'\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)\right] = \frac{Y^u_{t-}}{\mathcal{E}_{\max}}. \end{cases} \tag{3.3.1}$$

Although we can derive specific results for the control problem under assumption **(Toy)** (see Propositions 3.3.1 and 3.3.4), solving explicitly the system (3.3.1) remains difficult for general convex l. To get approximation results, we consider a specific form of l.

(ToyBis) The mapping l is given by $l(x) := \frac{\lambda}{2}x^2 + \frac{\varepsilon(\lambda+\mu)}{2}(x_+)^2$ with $\lambda \geq 0$, $|\varepsilon| < 1$.

From the application point of view, we remind that we want to penalize more consumption excess (compared to the commitment) than consumption deficit. The asymmetry parameter ε should thus be taken non-negative. Under assumptions **(Toy)** and **(ToyBis)**, the last equation in (3.3.1) writes:

$$(\lambda+\mu)u_t - \lambda\mathbb{E}[u_t] - \mathtt{C}_{t-} - \lambda(\mathtt{P}^{\mathtt{load}}_{t-} - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-}\right]) - \varepsilon(\lambda+\mu)\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)_+$$
$$+ \varepsilon(\lambda+\mu)\mathbb{E}\left[\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)_+\right] = \frac{Y^u_{t-}}{\mathcal{E}_{\max}}.$$

We now provide a first order expansion of the solution of this problem with respect to the parameter $\varepsilon \to 0$.

3.3.2 Computation of first order expansion

Preliminary result

The computation of a first order expansion of the solution of the MKV FBSDE (3.3.1) will rely extensively on the following result.

Proposition 3.3.2. *Let a, b, c, e, f, g be deterministic real parameters with $a > 0$, $g > 0$, $b \geq 0$ and $e \geq 0$. Let $(h_t)_t$ be a stochastic process in $\mathbb{H}^{2,2}_{\mathcal{P}}$ and $x_0 \in \mathbb{L}^2(\Omega)$ be $\mathcal{F}_0$-measurable. Define:*

$$\theta_t := \begin{cases} \frac{1}{2}\left(1 + e\sqrt{\frac{ag}{b}}\right)\exp(\sqrt{abg}(T-t)) + \frac{1}{2}\left(1 - e\sqrt{\frac{ag}{b}}\right)\exp(\sqrt{abg}(t-T)) & \text{if } b > 0, \\ eag(T-t) + 1 & if b = 0, \end{cases} \tag{3.3.2}$$

$$p_t := -\frac{\mathrm{d}\theta_t}{\mathrm{d}t}\frac{1}{ag\theta_t}, \tag{3.3.3}$$

$$\pi_t = \frac{1}{\theta_t}\left(f - \int_t^T (ap_s\mathbb{E}_t[h_s] - c)\theta_s \mathrm{d}s\right). \tag{3.3.4}$$

Define x, y and v by:

$$\begin{cases} x_t = x_0\frac{\theta_t}{\theta_0} - \int_0^t (ag\pi_s + ah_s)\frac{\theta_t}{\theta_s}\mathrm{d}s, \\ y_t = p_t x_t + \pi_t, \\ v_t = gp_t x_t + g\pi_{t-} + h_t. \end{cases} \tag{3.3.5}$$

Then (x, y, v) is a solution in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{2,2}_{\mathcal{P}}$ of the Forward-Backward system:

$$\begin{cases} x_t = x_0 - \int_0^t av_s \mathrm{d}s, \\ y_t = \mathbb{E}_t\left[\int_t^T (bx_s + c)\mathrm{d}s + ex_T + f\right], \\ v_t = gy_{t-} + h_t. \end{cases} \tag{3.3.6}$$

Besides, for T small enough, this solution to (3.3.6) is the unique one in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{2,2}_{\mathcal{P}}$.

The proof is postponed to Subsection 3.4.4.

Remark 3.3.3. *Uniqueness of the solution of the FBSDE (3.3.6) could be proved for arbitrary time horizon T, using the fact that (3.3.6) characterizes the solution of a (linear-quadratic) stochastic control which has a unique solution (as the associated cost function is continuous, convex and coercive [Bre10a, Corollary 3.23, pp. 71]).*

Average processes

We introduce the following notations for the average (in the sense of expectation) of the solutions of (3.3.1):

$$\bar{u} := \mathbb{E}[u], \qquad \bar{X} := \mathbb{E}[X^u], \qquad \bar{Y} := \mathbb{E}[Y^u], \qquad \bar{\mathsf{C}} := \mathbb{E}[\mathsf{C}].$$

By taking the expectation in (3.3.1), we immediately get the following simple but remarkable result: the average processes do not depend on l.

Proposition 3.3.4. *Assume (**Toy**), $(\bar{u}, \bar{X}, \bar{Y})$ solves*

$$\begin{cases} \bar{X}_t = \mathbb{E}[x] - \frac{1}{\mathcal{E}_{\max}}\int_0^t \bar{u}_s \mathrm{d}s, \\ \bar{Y}_t = \int_t^T \nu(\bar{X}_s - \frac{1}{2})\mathrm{d}s + \gamma\left(\bar{X}_T - \frac{1}{2}\right), \\ \bar{u}_t = \frac{\bar{Y}_{t-}}{\mu\mathcal{E}_{\max}} + \frac{\bar{\mathsf{C}}_{t-}}{\mu}. \end{cases} \tag{3.3.7}$$

In particular, $(\bar{u}, \bar{X}, \bar{Y})$ does not depend on l.

Note that the FBSDE (3.3.7) is explicitly solvable, as a particular case of Equation (3.3.6) with $x_0 := x$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = -\frac{\nu}{2}$, $e = \gamma$, $f = -\frac{\gamma}{2}$, $g = \frac{1}{\mu\mathcal{E}_{\max}}$ and $h_t = \frac{\bar{\mathsf{C}}_{t-}}{\mu}$.

Notations

From now on, assume that (**Toy**) and (**ToyBis**) hold. From Proposition 3.3.4, $(\bar{u}, \bar{X}, \bar{Y})$ does not depend on ε. We denote the processes u, X^u and Y^u by $u^{(\varepsilon)}$, $X^{(\varepsilon)}$ and $Y^{(\varepsilon)}$ respectively to insist on the dependency w.r.t. the parameter ε. $(u^{(\varepsilon)}, X^{(\varepsilon)}, Y^{(\varepsilon)})$ satisfies (3.3.1) with $l'(x) = \lambda x + \varepsilon(\lambda + \mu)x_+$.

For the ease of the proofs, let us define the recentered processes

$$u^{\Delta,(\varepsilon)} := u^{(\varepsilon)} - \bar{u}, \qquad X^{\Delta,(\varepsilon)} := X^{(\varepsilon)} - \bar{X}, \qquad Y^{\Delta,(\varepsilon)} := Y^{(\varepsilon)} - \bar{Y},$$
$$\mathrm{P}^{\mathrm{load},\Delta} := \mathrm{P}^{\mathrm{load}} - \mathbb{E}\left[\mathrm{P}^{\mathrm{load}}\right], \qquad \mathrm{C}^{\Delta} := \mathrm{C} - \mathbb{E}\left[\mathrm{C}\right].$$

Then, $(u^{\Delta,(\varepsilon)}, X^{\Delta,(\varepsilon)}, Y^{\Delta,(\varepsilon)})$ satisfies:

$$\begin{cases} X_t^{\Delta,(\varepsilon)} = x - \mathbb{E}\left[x\right] - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s^{\Delta,(\varepsilon)} \mathrm{d}s, \\ Y_t^{\Delta,(\varepsilon)} = \mathbb{E}_t\left[\int_t^T \nu X_s^{\Delta,(\varepsilon)} \mathrm{d}s + \gamma X_T^{\Delta,(\varepsilon)}\right], \\ \mu u_t^{\Delta,(\varepsilon)} - \mathrm{C}_{t-}^{\Delta} - \lambda\left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right) - \varepsilon(\lambda+\mu)\left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+ + \varepsilon(\lambda+\mu)\mathbb{E}\left[\left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+\right] = \frac{Y_{t-}^{\Delta,(\varepsilon)}}{\mathcal{E}_{\max}}. \end{cases} \tag{3.3.8}$$

We now seek a first order expansion of the solution of (3.3.1) w.r.t. ε, as $\varepsilon \to 0$, and equivalently, as the average processes do not depend on ε (see Proposition 3.3.4), we will perform it for the recentered processes, by showing

$$u^{\Delta,(\varepsilon)} = u^{\Delta,(0)} + \varepsilon \dot{u} + o(\varepsilon), \qquad X^{\Delta,(\varepsilon)} = X^{\Delta,(0)} + \varepsilon \dot{X} + o(\varepsilon), \qquad Y^{\Delta,(\varepsilon)} = Y^{\Delta,(0)} + \varepsilon \dot{Y} + o(\varepsilon),$$

where $\dot{u}$, $\dot{X}$ and $\dot{Y}$ are suitable processes in $\mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ (independent of ε) and the convergence $o(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$ holds in $\mathbb{H}^{2,2}$-norm.

Proposition 3.3.5. *Assume* (**Toy**) *and* (**ToyBis**). *Then* $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ *satisfies:*

$$\begin{cases} X_t^{\Delta,(0)} = x - \mathbb{E}\left[x\right] - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s^{\Delta,(0)} \mathrm{d}s, \\ Y_t^{\Delta,(0)} = \mathbb{E}_t\left[\int_t^T \nu X_s^{\Delta,(0)} \mathrm{d}s + \gamma X_T^{\Delta,(0)}\right], \\ u_t^{\Delta,(0)} = \frac{Y_{t-}^{\Delta,(0)}}{(\lambda+\mu)\mathcal{E}_{\max}} + \frac{\mathrm{C}_{t-}^{\Delta} + \lambda \mathrm{P}_{t-}^{\mathrm{load},\Delta}}{\lambda+\mu}. \end{cases} \tag{3.3.9}$$

Observe that the FBSDE (3.3.9) is known in a closed form, as a particular case of Equation (3.3.6) with $x_0 := x - \mathbb{E}\left[x\right]$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = 0$, $e = \gamma$, $f = 0$, $g = \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$ and $h_t = \frac{\mathrm{C}_{t-}^{\Delta} + \lambda \mathrm{P}_{t-}^{\mathrm{load},\Delta}}{\lambda+\mu}$.

Proposition 3.3.6. *Assume* (**Toy**) *and* (**ToyBis**). *Define the finite differences*

$$\dot{u}^{(\varepsilon)} := \frac{u^{\Delta,(\varepsilon)} - u^{\Delta,(0)}}{\varepsilon}, \qquad \dot{X}^{(\varepsilon)} := \frac{X^{\Delta,(\varepsilon)} - X^{\Delta,(0)}}{\varepsilon}, \qquad \dot{Y}^{(\varepsilon)} := \frac{Y^{\Delta,(\varepsilon)} - Y^{\Delta,(0)}}{\varepsilon},$$

which solve

$$\begin{cases} \dot{X}_t^{(\varepsilon)} = -\frac{1}{\mathcal{E}_{\max}} \int_0^t \dot{u}_s^{(\varepsilon)} \mathrm{d}s, \\ \dot{Y}_t^{(\varepsilon)} = \mathbb{E}_t\left[\int_t^T \nu \dot{X}_s^{(\varepsilon)} \mathrm{d}s + \gamma \dot{X}_T^{(\varepsilon)}\right], \\ \dot{u}_t^{(\varepsilon)} = \frac{\dot{Y}_{t-}^{(\varepsilon)}}{(\lambda+\mu)\mathcal{E}_{\max}} + \left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+ - \mathbb{E}\left[\left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+\right]. \end{cases} \tag{3.3.10}$$

Besides, for small enough time horizon T, $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ is uniformly bounded in $\mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ as $\varepsilon \to 0$. Define $(\dot{u}, \dot{X}, \dot{Y})$ as a solution (unique when T is small enough) to

$$\begin{cases} \dot{X}_t = -\frac{1}{\mathcal{E}_{\max}} \int_0^t \dot{u}_s \mathrm{d}s, \\ \dot{Y}_t = \mathbb{E}_t\left[\int_t^T \nu \dot{X}_s + \gamma \dot{X}_T \mathrm{d}s\right], \\ \dot{u}_t = \frac{\dot{Y}_{t-}}{(\lambda+\mu)\mathcal{E}_{\max}} + \left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(0)}\right)_+ - \mathbb{E}\left[\left(\mathrm{P}_{t-}^{\mathrm{load},\Delta} - u_t^{\Delta,(0)}\right)_+\right]. \end{cases} \tag{3.3.11}$$

Then, for small enough time horizon T, the finite differences $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ are close (at order 1 in ε) to $(\dot{u}, \dot{X}, \dot{Y})$:

$$\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}_{\mathcal{P}}} + \|\dot{X}^{(\varepsilon)} - \dot{X}\|_{\mathbb{H}^{2,2}} + \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} = O(\varepsilon).$$

The proof is postponed to Subsection 3.4.5. Note again that the FBSDE (3.3.11) is explicitly solvable, as a particular case of Equation (3.3.6) with $x_0 := 0$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = 0$, $e = \gamma$, $f = 0$, $g = \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$ and $h_t = \left(\mathrm{P}^{\mathrm{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+ - \mathbb{E}\left[\left(\mathrm{P}^{\mathrm{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+\right]$.

Collecting all the previous results, we get the following theorem, which fully characterizes the first order expansion of the solution to the control problem.

Theorem 3.3.7. *Assume* **(Toy)** *and* **(ToyBis)** *hold. For small enough time horizon T, the unique solution $(u^{(\varepsilon)}, X^{(\varepsilon)}, Y^{(\varepsilon)})$ in $\mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ of (3.3.1) can be expanded at first order w.r.t. ε (with error of second order as $\varepsilon \to 0$):*

$$u^{(\varepsilon)} = \bar{u} + u^{\Delta,(0)} + \varepsilon \dot{u} + O(\varepsilon^2), \qquad X^{(\varepsilon)} = \bar{X} + X^{\Delta,(0)} + \varepsilon \dot{X} + O(\varepsilon^2), \qquad Y^{(\varepsilon)} = \bar{Y} + Y^{\Delta,(0)} + \varepsilon \dot{Y} + O(\varepsilon^2),$$

where errors $O(\varepsilon^2)$ are measured in $\mathbb{H}^{2,2}$-norm, with $(\bar{u}, \bar{X}, \bar{Y})$ solution of (3.3.7), $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ solution of (3.3.9) and $(\dot{u}, \dot{X}, \dot{Y})$ solution of (3.3.11).

We shall emphasize that all terms in these expansions are solutions of FBSDEs of the form (3.3.6) for different input parameters (see Table 3.1) and thus they are explicitly solvable.

For other problems with more regularity (notice that $x \mapsto (x_+)^2$ is not twice continuously differentiable), the previous approach could actually be extended to a second order expansion or even higher order, but it would lead to more and more nested FBSDEs: on the mathematical side, there is no hard obstacle to derive these equations under appropriate regularity conditions. The concerns would be rather on the computational side since it would require larger and larger computational time.

3.3.3 Effective simulation of first order expansion of optimal control

Models for random uncertainties

We assume the electricity price C is constant ($\bar{\mathrm{C}} = \mathrm{C}$ and $\mathrm{C}^{\Delta} = 0$), and we suppose $\mathrm{P}^{\mathrm{load}}$ is given by $\mathrm{P}^{\mathrm{load}} = \mathrm{P}^{\mathrm{cons}} - \mathrm{P}^{\mathrm{sun}}$, where $\mathrm{P}^{\mathrm{cons}}$ and $\mathrm{P}^{\mathrm{sun}}$ are two independent scalar SDEs[3], representing respectively the consumption and the photovoltaic power production. For the consumption $\mathrm{P}^{\mathrm{cons}}$, we use the jump process:

$$\mathrm{d}\mathrm{P}^{\mathrm{cons}}_t = -\rho^{\mathrm{cons}}(\mathrm{P}^{\mathrm{cons}}_t - \mathrm{p}^{\mathrm{cons,ref}}_t)\mathrm{d}t + h^{\mathrm{cons}}\mathrm{d}N^{\mathrm{cons}}_t, \tag{3.3.12}$$

where N^{cons} is a compensated Poisson Process with intensity λ^{cons}. Regarding the PV production, we follow [Bad+18] by setting $\mathrm{P}^{\mathrm{sun}} = \mathrm{P}^{\mathrm{sun,max}}\mathrm{X}^{\mathrm{sun}}$ where $\mathrm{P}^{\mathrm{sun,max}} : [0,T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) and $\mathrm{X}^{\mathrm{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathrm{X}^{\mathrm{sun}}_t = -\rho^{\mathrm{sun}}(\mathrm{X}^{\mathrm{sun}}_t - \mathrm{x}^{\mathrm{sun,ref}}_t)\mathrm{d}t + \sigma^{\mathrm{sun}}(\mathrm{X}^{\mathrm{sun}}_t)^{\alpha}(1 - \mathrm{X}^{\mathrm{sun}}_t)^{\beta}\mathrm{d}W_t, \tag{3.3.13}$$

with $\alpha, \beta \geq 1/2$. As proved in [Bad+18], there is a strong solution to the above SDE and the solution $\mathrm{X}^{\mathrm{sun}}$ takes values in $[0,1]$.

Since the drifts are affine-linear, the conditional expectation of the solution is known in closed forms (this property is intensively used in [BSS05]):

$$\mathbb{E}_t\left[\mathrm{P}^{\mathrm{sun}}_s\right] = \left(\frac{\mathrm{P}^{\mathrm{sun}}_t}{\mathrm{P}^{\mathrm{sun,max}}_t}\exp(-\rho^{\mathrm{sun}}(s-t)) + \int_t^s \rho^{\mathrm{sun}}\mathrm{x}^{\mathrm{sun,ref}}_\tau \exp(-\rho^{\mathrm{sun}}(s-\tau))\mathrm{d}\tau\right)\mathrm{P}^{\mathrm{sun,max}}_s, \tag{3.3.14}$$

$$\mathbb{E}_t\left[\mathrm{P}^{\mathrm{cons}}_s\right] = \mathrm{P}^{\mathrm{cons}}_t \exp(-\rho^{\mathrm{cons}}(s-t)) + \int_t^s \rho^{\mathrm{cons}}\mathrm{p}^{\mathrm{cons,ref}}_\tau \exp(-\rho^{\mathrm{cons}}(s-\tau))\mathrm{d}\tau, \tag{3.3.15}$$

for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\mathrm{P}^{\mathrm{load}}_s\right]$ as required when deriving the optimal control.

Algorithm 3.1 Sample of a path of (x, y, v), solution of (3.3.6)

1: **Inputs:** $x_0 \in \mathbb{L}^2_\Omega, a > 0, b \geq 0, c \in \mathbb{R}, e \geq 0, f \in \mathbb{R}, g > 0, h \in \mathbb{H}^{2,2}_{\mathcal{P}}, N_T > 0$
2: Sample x_0 and set $X(0) \leftarrow x_0$. Set $\tau = \frac{T}{N_T}$.
3: **for** $n = 0, ..., N_T - 1$ **do**
4: Compute the conditional expectations $(\mathbb{E}_{n\tau}[h_s])_{n\tau \leq s \leq T}$
5: Compute $\pi(n\tau)$ by numerical integration, as given in (3.3.4)
6: Compute $p(n\tau)$ as in (3.3.3)
7: $v(n\tau) \leftarrow gp(n\tau)X(n\tau) + g\pi(n\tau) + h(n\tau)$
8: $x((n+1)\tau) \leftarrow x(n\tau) - av(n\tau)\tau$
9: **end for**
10: **return** (x, y, v)

FBSDE Parameters

Proposition 3.3.2 is repeatedly used to solve the affine-linear FBSDEs $(\bar{u}, \bar{X}, \bar{Y})$, $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ and $(\dot{u}, \dot{X}, \dot{Y})$ arising in the first order expansion of the optimal control w.r.t. ε (see Theorem 3.3.7). In Algorithm 3.1 we give the pseudo-code of the scheme used to compute solutions of the FBSDE of the form (3.3.6).

In Table 3.1, we give the correspondence between the input parameters $(a, b, c, d, e, f, g, h_t)$ for the generic FBSDE of Proposition 3.3.2 and the parameters defining the 3 FBSDEs. Merged columns indicate common values of parameters. As the data involved in the system defining $(\bar{u}, \bar{X}, \bar{Y})$ is deterministic, one only needs to perform

	$(\bar{u}, \bar{X}, \bar{Y})$	$(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$	$(\dot{u}, \dot{X}, \dot{Y})$
a	$\frac{1}{\mathcal{E}_{\max}}$		
b	ν		
c	$\frac{-\nu}{2}$	0	
e	γ		
f	$\frac{-\gamma}{2}$	0	
g	$\frac{1}{\mu\mathcal{E}_{\max}}$	$\frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$	
h_t	$\frac{\bar{\mathrm{C}}_{t-}}{\mu}$	$\frac{\mathrm{C}^{\Delta}_{t-} + \lambda \mathrm{P}^{\mathrm{load},\Delta}_{t-}}{\lambda+\mu}$	$\left(\mathrm{P}^{\mathrm{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+ - \mathbb{E}\left[\left(\mathrm{P}^{\mathrm{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+\right]$

Table 3.1: Table of parameters needed to compute the expansion terms

numerical integrations to compute π and therefore $(\bar{u}, \bar{X}, \bar{Y})$. For $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ and $(\dot{u}, \dot{X}, \dot{Y})$, it becomes a bit more involved. Let us provide some details on the implementation.

- For the computation of $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$, the conditional expectations $(\mathbb{E}_{n\tau}[h_s])_{n\tau \leq s \leq T}$ are given by affine-linear combinations of $\mathrm{P}^{\mathrm{cons}}_{n\tau}$ and $\mathrm{P}^{\mathrm{sun}}_{n\tau}$ with deterministic coefficients, depending on s and n, by assumption on our models for $\mathrm{P}^{\mathrm{cons}}$ and $\mathrm{P}^{\mathrm{sun}}$ (see (3.3.14)-(3.3.15)). Therefore, $\pi(n\tau)$ is also given by an affine-linear combination of $\mathrm{P}^{\mathrm{cons}}_{n\tau}$ and $\mathrm{P}^{\mathrm{sun}}_{n\tau}$ with deterministic coefficients. This allows to speed up Steps 4 and 5 in Algorithm 3.1.

- For the computation of $(\dot{u}, \dot{X}, \dot{Y})$, the conditional expectations $\left(\mathbb{E}_{n\tau}\left[\left(\mathrm{P}^{\mathrm{load},\Delta}_s - u^{\Delta,(0)}_s\right)_+\right]\right)_{n\tau \leq s \leq T}$ at Step 4 is estimated by Monte-Carlo methods. The procedure for doing so is given in Algorithm 3.2. This Step 4 has a complexity of order $O((N_T - n)M_0)$, which is the most costly Step in the loop of Algorithm 3.1; hence sampling $(\dot{u}, \dot{X}, \dot{Y})$ has a computational cost of order $O(N_T^2 M_0)$.

[3] we consider Brownian SDEs for simplicity, but note that the current setting allows more general processes.

Algorithm 3.2 Evaluation of $\left(\mathbb{E}_{n\tau}\left[\left(\mathrm{P}_s^{\mathrm{load},\Delta} - u_s^{\Delta,(0)}\right)_+\right]\right)_{s=n\tau,\dots,N_T\tau}$

Inputs: $n < N_T$, $X_{n\tau}^{\Delta,(0)}$, $\mathrm{P}_{n\tau}^{\mathrm{sun}}$, $\mathrm{P}_{n\tau}^{\mathrm{load}}$, $M_0 > 0$
Initialization: $(R[n], R[n+1], \dots, R[N_T]) \leftarrow (0, 0, \dots, 0)$.
Compute $u^{\Delta,(0)}(n\tau)$ using a similar procedure as in Algorithm 3.1.
$R[n] \leftarrow \left(\mathrm{P}_{n\tau}^{\mathrm{load},\Delta} - u_{n\tau}^{\Delta,(0)}\right)_+$.
for $m = 1, \dots, M_0$ **do**
  for $k = n+1, \dots, N_T$ **do**
    Sample $(\mathrm{P}_{k\tau}^{\mathrm{cons}}, \mathrm{P}_{k\tau}^{\mathrm{sun}})$ conditionally to $(\mathrm{P}_{(k-1)\tau}^{\mathrm{cons}}, \mathrm{P}_{(k-1)\tau}^{\mathrm{sun}})$ using (3.3.12)-(3.3.13), independently from all other random variables simulated so far.
    Compute $u_{k\tau}^{\Delta,(0)}$ with Steps 5 to 8 of Algorithm 3.1 with the data of the FBSDE (3.3.9). Compute $X_{k\tau}^{\Delta,(0)}$.
    $R[k] \leftarrow R[k] + \frac{1}{M_0}\left(\mathrm{P}_{k\tau}^{\mathrm{load},\Delta} - u_{k\tau}^{\Delta,(0)}\right)_+$
  end for
end for
return $(R[n], R[n+1], \dots, R[N_T])$

Numerical values of parameters

We report the values chosen for the next experiments.

Parameters for smart grid. We consider the following values for the time horizon, the size of the storage system and the initial value of its normalized state of charge.

Parameter	T	$\mathcal{E}_{\max}$	x_0
Value	24 h	200 kWh	0.5

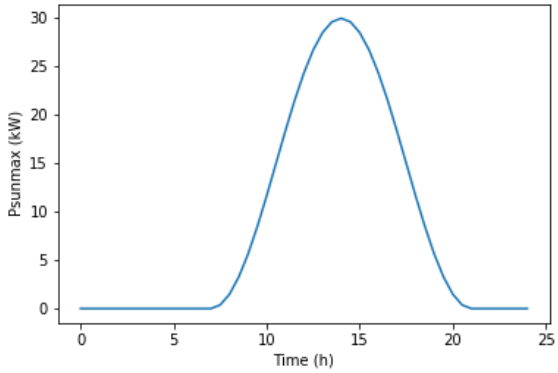
Parameters for uncertain consumption/production. The following table gives the values of the parameters used in the modeling of the underlying exogenous stochastic processes impacting the system.

$\mathrm{P}^{\mathrm{sun}}$	ρ^{sun}	$0.75h^{-1}$
	$\mathrm{x}_t^{\mathrm{sun,ref}}$	0.5
	σ^{sun}	0.8
	α	0.8
	β	0.7
	$\mathrm{P}^{\mathrm{sun,max}}$	see Figure 3.2a
$\mathrm{P}^{\mathrm{cons}}$	ρ^{cons}	$0.9h^{-1}$
	$\mathrm{p}^{\mathrm{cons,ref}}$	see Figure 3.2b
	h^{cons}	5 kW
	λ^{cons}	0.5 h^{-1}

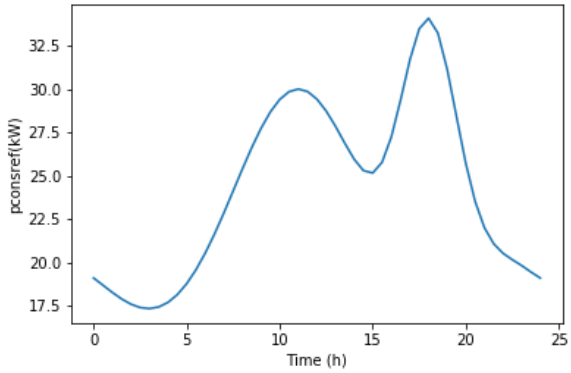
In Figure 3.2, we plot the time-evolution of the deterministic functions $\mathrm{P}^{\mathrm{sun,max}}$ and $\mathrm{p}^{\mathrm{cons,ref}}$, 10 independent samples of processes $\mathrm{P}^{\mathrm{sun}}$ and $\mathrm{P}^{\mathrm{cons}}$, and the time-evolution of quantiles (computed with $M_1 = 100000$ i.i.d. simulations).

Parameters of input data and optimization problem. The values of the parameters of the optimization problem are chosen such that:

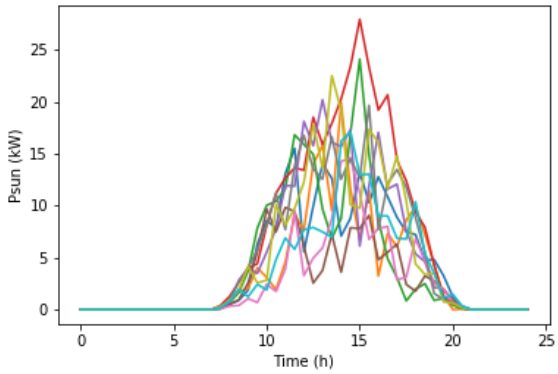
1. the state of charge of the battery remains close to a reference level, which we set to 0.5,
2. we observe a clear reduction of the random fluctuation of $\mathrm{P}^{\mathrm{grid}}$ on the time interval.



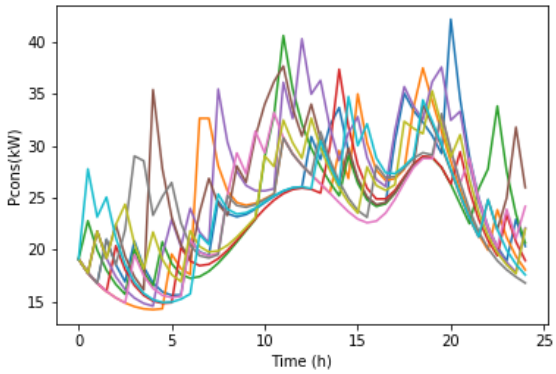
(a) Time evolution of $\mathrm{P}^{\mathrm{sun,max}}$, accounting for clear sky model



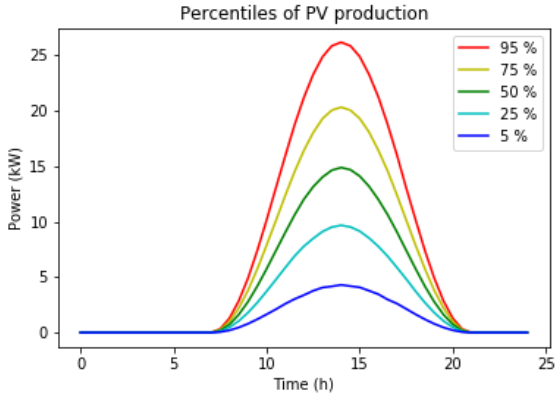
(b) Time evolution of $\mathrm{p}^{\mathrm{cons,ref}}$, accounting for intraday peaks



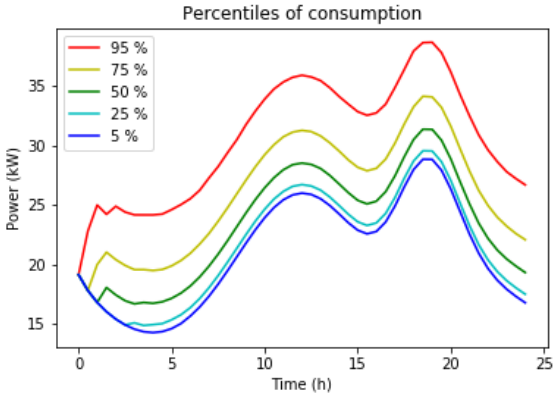
(c) Example trajectories of $\mathrm{P}^{\mathrm{sun}}$



(d) Example trajectories of $\mathrm{P}^{\mathrm{cons}}$



(e) Time evolution of quantiles of $\mathrm{P}^{\mathrm{sun}}$



(f) Time evolution of quantiles of $\mathrm{P}^{\mathrm{cons}}$

Figure 3.2: Graphical statistics of the evolution of $\mathrm{P}^{\mathrm{sun}}$ and $\mathrm{P}^{\mathrm{cons}}$

The following table gives the values of the parameters of the cost functional of the control problem.

Parameter	ε	λ	μ	ν	γ	C
Value	0.2	0.49	0.01	0.1	500	0.27
Unit	-	euros.kW^{-2}.h^{-1}	euros.kW^{-2}.h^{-1}	euros.h^{-1}	euros	euros.kW.h^{-1}

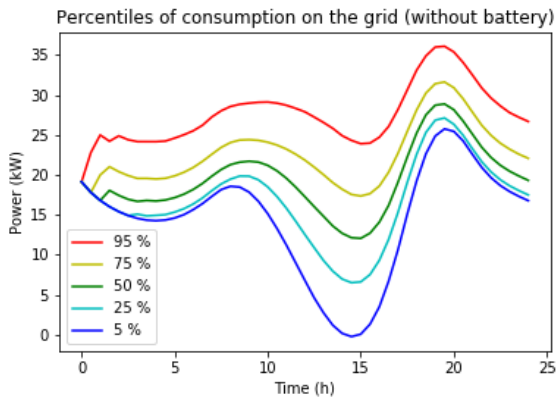
Time discretization. The average processes $(\bar{u}, \bar{X})$ are computed explicitly (up to numerical integration), while the

recentered processes $(u^{\Delta,(0)}, X^{\Delta,(0)})$ and first order correction processes $(\dot{u}, \dot{X})$ are computed using discretization schemes (detailed in Algorithms 3.1 and 3.2) with time-step equal to $0.5h$.

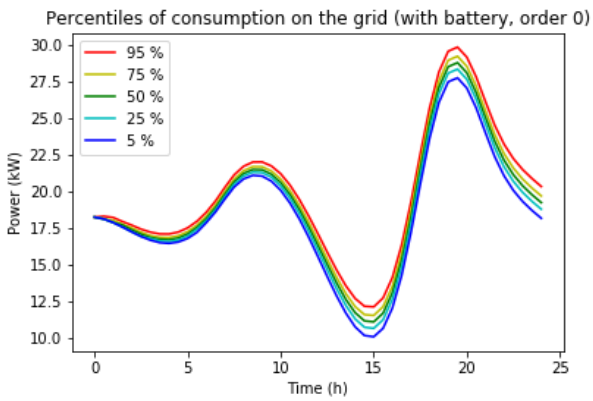
Monte-Carlo simulations. To compute the first order correction, we need Monte-Carlo estimations, as explained in Algorithm 3.2. We choose $M_0 = 4000$. For assessing the statistical performances of the optimal control associated to a symmetric loss function ($\varepsilon = 0$), we consider $M_1 = 100000$ macro-runs. Among those M_1 trajectories, we only consider the first $M_2 = 4000$ trajectories for the computation of the first order corrections associated to $\varepsilon = 0.2$.

Results from the experiments

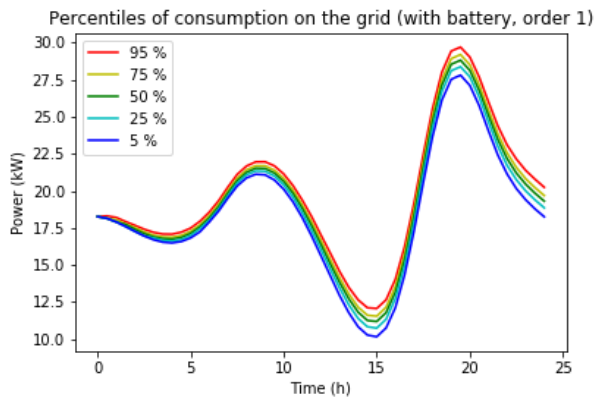
Computational time. The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. We have computed the optimal control associated to a symmetric penalization of deviations of $\mathrm{P}^{\mathrm{grid}}$ from its average ($\varepsilon = 0$) and for $M_1 = 100000$ i.i.d. simulations, which takes about 3 seconds. The computation of the first order correction when $\varepsilon = 0.2$ for $M_2 = 4000$ i.i.d. simulations takes about 80 minutes.



(a) Without flexibility



(b) With controlled flexibility ($\varepsilon = 0$)



(c) With controlled flexibility ($\varepsilon = 0.2$)

Figure 3.3: Quantiles of $\mathrm{P}^{\mathrm{grid}}$ as a function of time

Reduction of fluctuations. We plot the time-evolution of quantiles (see Figure 3.3) of the power supplied by the network in 3 cases: using no flexibility, with optimal control of the battery with $\varepsilon = 0$, and with the approximated optimal control associated to $\varepsilon = 0.2$ respectively. The comparison of the first graph with the two others shows that the quantiles are much closer to each other in the case of storage use, meaning that the variability of the power supplied by the grid has been much reduced, as expected. However, the difference between the optimal control with symmetric and asymmetric loss functions is not much visible on these plots.

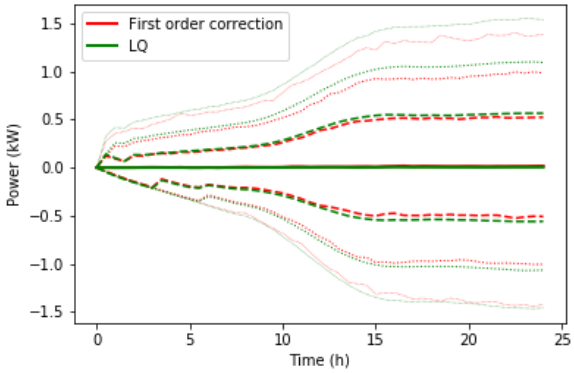
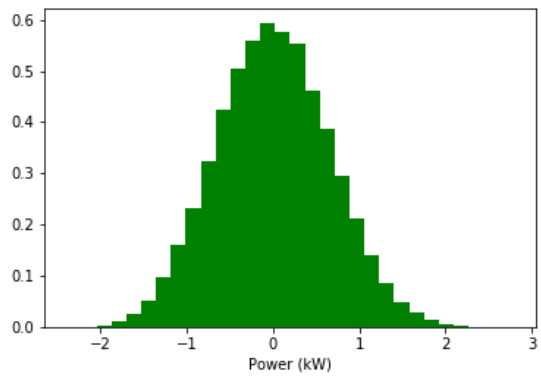
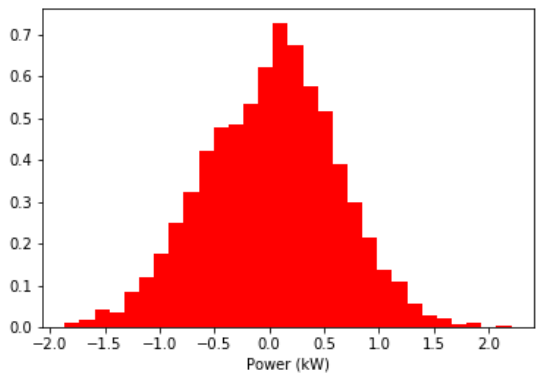


Figure 3.4: Time evolution of quantiles of deviations $\mathrm{P}^{\mathrm{grid}} - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}\right]$

Impact of first order correction. Overall, the effect of the first order correction $\dot{u}$ (which has theoretically an average value of 0), is to lower the probability of large upper deviations of $\mathrm{P}^{\mathrm{grid}}$ from its expectation. This is quite visible if we plot the time-evolution of quantiles of the deviations $\mathrm{P}^{\mathrm{grid}}(t) - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}(t)\right]$ for the case $\varepsilon = 0$, in green in Figure 3.4, refered as "LQ" and $\varepsilon = 0.2$, in red, refered as "First Order Correction". In Figure 3.4, we have represented from top to bottom, the quantiles of $\mathrm{P}^{\mathrm{grid}}(t) - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}(t)\right]$ associated to levels 99%, 95%, 80%, 50%, 20%, 5% and 1%. We observe that the empirical estimations of the lower quantiles are left unchanged, while the upper quantiles 99% and 95% have been notably decreased, which was the effect sought by the choice of this loss function. To have a even more clear visualization of the change of distribution of the deviations $\mathrm{P}^{\mathrm{grid}} - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}\right]$, we have represented the empirical histograms of $\mathrm{P}^{\mathrm{grid}}(T) - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}(T)\right]$ for both cases $\varepsilon = 0$ in Figure 3.5a ($M_1 = 100000$ i.i.d. simulations) and $\varepsilon = 0.2$ in Figure 3.5b ($M_2 = 4000$ i.i.d. simulations). Observe that the impact of the first order term is to break the symmetry of the distribution around 0, and to reduce the probability of the highest values of $\mathrm{P}^{\mathrm{grid}}(T) - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}(T)\right]$. These results suggest that we have reached our goal of reducing the probability of high upper deviations of $\mathrm{P}^{\mathrm{grid}}$ from its average.



(a) Symmetric penalization ($\varepsilon = 0$)



(b) Asymmetric penalization ($\varepsilon = 0.2$)

Figure 3.5: Empirical histograms of deviations $\mathrm{P}^{\mathrm{grid}}(T) - \mathbb{E}\left[\mathrm{P}^{\mathrm{grid}}(T)\right]$

Distribution of state of charge of the battery. As the first order correction term has only minor impact on the distribution of the state of charge of the storage system, we only consider the case with $\varepsilon = 0$ in this paragraph. Figure 3.6 shows the time-evolution of the quantiles 95%, 50%, 5% of the state of charge of the battery with $\varepsilon = 0$ (computed using $M_1 = 100000$ i.i.d. simulations) for several initial conditions on the state of charge of the battery, namely $x_0 = 0.75$, $x_0 = 0.5$ and $x_0 = 0.25$. What we observe is that independently on the initial condition chosen, the state of

charge remains between 0.15 and 0.75 with high probability. Besides, the terminal distribution of the state of charge is quite independent from the initial condition: the terminal values of the quantiles (levels 95%, 50%, 5%) of the state of charge are almost the same, for all initial conditions $x_0 = 0.75$, $x_0 = 0.5$ and $x_0 = 0.25$. This is presumably due to the term in the cost functional which penalizes the deviations of state of charge from a medium value (here $1/2$).

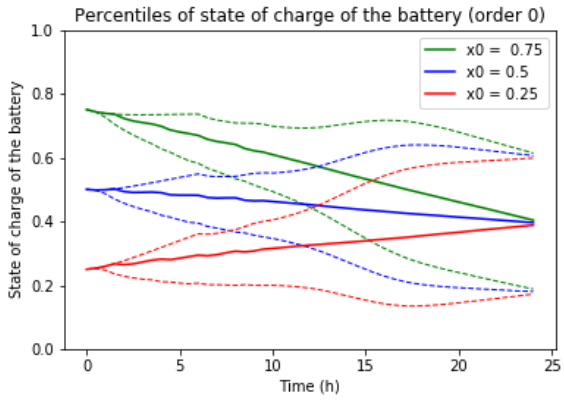


Figure 3.6: Time evolution of the empirical quantiles 95%, 50%, 5% of the state of charge of the storage system

Simulation-based bound on approximation error of first order expansion. Following the proof of Proposition 3.3.6, with our choice of parameters, we obtain an upper bound on the error in the approximation of the optimal control $u^{(\varepsilon)}$:

$$\begin{aligned}\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} &= \varepsilon\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} \\ &\leq \frac{4\varepsilon^2}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)}\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}.\end{aligned}$$

We would like to obtain a bound on the relative error committed $\frac{\|u^{(\varepsilon)}-\bar{u}-u^{\Delta,(0)}-\varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\|u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}}$. To do this, we have by triangular inequality:

$$\begin{aligned}\frac{\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\|u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}} &\leq \frac{\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\left|\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - \|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}\right|} \\ &\leq \frac{4\varepsilon^2}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)}\frac{\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{\left|\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - \|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}\right|} \\ &\leq \frac{4\varepsilon^2\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - 4\varepsilon^2\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}.\end{aligned}$$

In the last inequality, we used the fact that $\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$ is asymptotically small compared to $\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$ when ε goes to 0, as well as the previous bound on $\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$. Hence we obtain an upper bound which depends only on quantities which can be estimated by simulations in the algorithm. This is very convenient to assess the relative accuracy of our approximation. The left-hand-side in the last inequality is estimated using the $M_2 = 4000$ simulations of the first order expansion and we find a value of 0.03. In other words, the relative error is smaller than 3% when taking the first order expansion of the control instead of its true value. Note that we do not take into account errors due to the time discretization or due to residual noise in the Monte-Carlo estimations.

3.4 Proofs

3.4.1 Proof of Theorem 3.2.2

a) Observe first that, in view of **(H.ϕ)** and Lemma 3.4.1, $\tilde{L}^u$ and $(\tilde{L}^u)^{-1}$ are uniformly bounded by a constant $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e (take $A : (t,\omega) \mapsto \nabla_x\phi(t,\omega,u_t,X_t^u)$). Therefore, and owing to **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**, the random

variable inside the conditional expectation defining Y^u in (3.2.4) is bounded by

$$\begin{aligned}\Gamma_T :=& C_T\left(C_\psi^{(0)} + |X_T^u| + \mathbb{E}\left[|k(0)|\right] + \mathbb{E}\left[|X_T^u|\right]\right)\\ &+ C_T\int_0^T\left(C_l^{(0)}(s) + |X_s^u| + |u_s| + \mathbb{E}\left[|g(s,0,0)|\right] + \mathbb{E}\left[|u_s|\right] + \mathbb{E}\left[|X_s^u|\right] + \mathbb{E}\left[C_l^{(0)}(s)\right]\right)\mathrm{d}s\end{aligned}$$

for some constant C_T depending on the bounds in (**H.x**)-(**H.g**)-(**H.l**)-(**H.k**)-(**H.ψ**)-(**H.ϕ**). Hence by the Cauchy Schwartz inequality, for some other constant C_T:

$$\begin{aligned}\mathbb{E}\left[|\Gamma_T|^2\right] \leq& C_T\left(\mathbb{E}\left[\left(C_\psi^{(0)}\right)^2\right] + \mathbb{E}\left[|X_T^u|^2\right] + \mathbb{E}\left[|k(0)|\right]^2\right)\\ &+ C_T\int_0^T\left(\mathbb{E}\left[\left(C_l^{(0)}(s)\right)^2\right] + \mathbb{E}\left[|g(s,0,0)|\right]^2 + \mathbb{E}\left[|u_s|^2\right] + \mathbb{E}\left[|X_s^u|^2\right]\right)\mathrm{d}s.\end{aligned}$$

Note that this bound is finite and independent from t (since $C_\psi^{(0)} \in \mathbb{L}^2_\Omega$, $X^u \in \mathbb{H}^{\infty,2} \subset \mathbb{H}^{2,2}$, $k(0) \in \mathbb{L}^1$, $C_l^{(0)} \in \mathbb{H}^{2,2}$, $g(\cdot,0,0) \in \mathbb{H}^{2,1}$ and $u \in \mathbb{H}^{2,2}$). Consequently $Y^u \in \mathbb{H}^{\infty,2}$.

b) Now observe that, by definition of Y^u,

$$N_t^u := \tilde{L}_t^u Y_t^u + \int_0^t \tilde{L}_s^u\Big(\nabla_x l(s,u_s,X_s^u,\bar{g}_s^u) + \nabla_x g(s,u_s,X_s^u)\mathbb{E}\left[\nabla_{\bar{g}} l(s,u_s,X_s^u,\bar{g}_s^u)\right]\Big)\mathrm{d}s = \mathbb{E}_t\left[N_T^u\right], \tag{3.4.1}$$

with N_T^u square integrable (using the same arguments as before) and therefore, N^u is a càdlàg martingale in $\mathbb{H}^{\infty,2}$. As a by-product, we obtain that Y^u is a semi-martingale, which dynamics has now to be identified.

c) To justify that Y^u defined in (3.2.4) solves the BSDE (3.2.5) for some M^u, left-multiply both sides of (3.4.1) by $(\tilde{L}_t^u)^{-1}$, then apply the integration by parts formula in [Pro03, Corollary 2, p. 68] to $(\tilde{L}^u)^{-1}N^u$ and use the fact that $(\tilde{L}^u)^{-1}$ is continuous with finite variations. After reorganizing terms and using that N^u has countable jumps, we retrieve (3.2.5) with $M_t^u := \int_{0^+}^t (\tilde{L}_s^u)^{-1}\mathrm{d}N_s^u$ (which is also a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [Pro03, Theorem 20 p.63, Corollary 3 p.73, Theorem 29 p.75]).

d) We now claim that the solution (Y,M) to (3.2.5) is unique in $\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}$, and thus given by (Y^u,M^u). In fact, the uniqueness is a classical result for linear BSDE, see for instance [EPQ97, Theorem 5.1 with $p=2$] in our context of general filtration.

e) Let us now prove the last claim about $\dot{\mathcal{J}}(u,v)$. We first study the differentiability properties of $X^{u+\varepsilon v}$ with respect to ε. In the following computations we use different constants which we denote generically by C (although their values may change from line to line), they do not depend on u,v,ε, they only depend on $T>0$ and on the bounds from the assumptions (**H.x**)-(**H.g**)-(**H.l**)-(**H.k**)-(**H.ψ**)-(**H.ϕ**). At this point of the proof, it is more convenient to work with Jacobian matrices than with gradients (as in the statement). Only at the very end are we going to make the link with Y^u and go back to the gradient notation.

Set $\theta_t^u := (t,u_t,X_t^u)$ and let $\dot{X}_t^{u,v}$ be the solution to the following linear equation

$$\dot{X}_t^{u,v} := \int_0^t\left[\phi_u(\theta_s^u)v_s + \phi_x(\theta_s^u)\dot{X}_s^{u,v}\right]\mathrm{d}s = \int_0^t (L_t^u)^{-1}L_s^u\phi_u(\theta_s^u)v_s\mathrm{d}s, \tag{3.4.2}$$

since it can be noticed that L^u is the unique solution of

$$L_0^u = \mathrm{Id}_p, \qquad \frac{\mathrm{d}L_t^u}{\mathrm{d}t} = -L_t^u\phi_x(\theta_t^u),$$

using Lemma 3.4.1 and $\nabla_x\phi = (\phi_x)^\top$. Note, whenever useful, that $\int_0^T |v_s|^2\mathrm{d}s < +\infty$ a.s. since $v \in \mathbb{H}^{2,2}$.

For $\varepsilon \neq 0$, set $\Delta X_t^{u,v,\varepsilon} := \frac{X_t^{u+\varepsilon v}-X_t^u}{\varepsilon}$ and $RX_t^{u,v,\varepsilon} := \Delta X_t^{u,v,\varepsilon} - \dot{X}_t^{u,v}$: we claim that a.s. $RX_t^{u,v,\varepsilon} \to 0$ as $\varepsilon \to 0$, i.e. $\dot{X}_t^{u,v}$ is the derivative of $X_t^{u+\varepsilon v}$ at $\varepsilon = 0$. To justify this, we proceed in a few steps. First the Taylor formula equality gives, for smooth φ,

$$\varphi(z,x') - \varphi(z,x) = \Big(\int_0^1 \varphi_x(z,x+\lambda(x'-x))\mathrm{d}\lambda\Big)(x'-x) := \overline{\varphi}_x(z,[x,x'])(x'-x);$$

applying that to ϕ_x and ϕ_u, we obtain

$$\begin{aligned}\Delta X_t^{u,v,\varepsilon} &= \int_0^t \left[\overline{\phi}_x(s, u_s + \varepsilon v_s, [X_s^u, X_s^{u+\varepsilon v}])\Delta X_s^{u,v,\varepsilon} + \overline{\phi}_u(s, [u_s, u_s + \varepsilon v_s], X_s^u)v_s\right]\mathrm{d}s \\ &= (\overline{L_t^{u,v,\varepsilon}})^{-1} \int_0^t \overline{L_s^{u,v,\varepsilon}}\, \overline{\phi}_u(s, [u_s, u_s + \varepsilon v_s], X_s^u)v_s\mathrm{d}s \end{aligned} \tag{3.4.3}$$

where $\overline{L^{u,v,\varepsilon}}$ is the unique solution of

$$\overline{L_0^{u,v,\varepsilon}} = \mathrm{Id}_p, \qquad \frac{\mathrm{d}\overline{L_t^{u,v,\varepsilon}}}{\mathrm{d}t} = -\overline{L_t^{u,v,\varepsilon}}\left(\overline{\phi}_x(t, u_t + \varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}])\right).$$

By hypothesis on ϕ_x, L^u and $(L^u)^{-1}$ are $dt \times \mathrm{d}\mathbb{P}$-a.e. uniformly bounded by a constant. Besides, $(t, \omega) \mapsto \overline{\phi}_x(t, u_t + \varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}])$ satisfies the hypothesis of Lemma 3.4.1 with a constant C independent from ε. Therefore $\overline{L^{u,v,\varepsilon}}$ is invertible and $L^{u,v,\varepsilon}$ and $(L^{u,v,\varepsilon})^{-1}$ are $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. uniformly bounded by a constant independent from ε. From (3.4.2)-(3.4.3) and using the fact that ϕ_u and $\overline{\phi}_u(\cdot, [u, u+\varepsilon v], X^u)$ are $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. bounded by a constant independent from ε (by hypothesis on ϕ_u), we derive

$$|\Delta X_t^{u,v,\varepsilon}| + |\dot{X}_t^{u,v}| \leq C \int_0^t |v_s|\mathrm{d}s, \tag{3.4.4}$$

where C is a constant independent from ε. With similar arguments, we can represent the residual error $RX^{u,v,\varepsilon}$ as follows:

$$RX_t^{u,v,\varepsilon} = \int_0^t \left(\eta_s^{u,v,\varepsilon} + \phi_x(s, u_s, X_s^u)RX_s^{u,v,\varepsilon}\right)\mathrm{d}s = \int_0^t (L_t^u)^{-1} L_s^u \eta_s^{u,v,\varepsilon}\mathrm{d}s,$$

with L^u as before and

$$\begin{aligned}\eta_t^{u,v,\varepsilon} &:= \alpha_t^{u,v,\varepsilon}\Delta X_t^{u,v,\varepsilon} + \beta_t^{u,v,\varepsilon}v_t, \\ \alpha_t^{u,v,\varepsilon} &:= \overline{\phi}_x(t, u_t + \varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}]) - \phi_x(t, u_t, X_t^u), \\ \beta_t^{u,v,\varepsilon} &:= \overline{\phi}_u(t, [u_t, u_t + \varepsilon v_t], X_t^u) - \phi_u(t, u_t, X_t^u).\end{aligned}$$

By boundedness of L^u and $(L^u)^{-1}$, and by (3.4.4), we have:

$$\sup_{t\in[0,T]} |RX_t^{u,v,\varepsilon}| \leq C\left(\int_0^T |\alpha_t^{u,v,\varepsilon}|\mathrm{d}t\right)\left(\int_0^T |v_t|\mathrm{d}t\right) + C\int_0^T |\beta_t^{u,v,\varepsilon}||v_t|\mathrm{d}t,$$

for some constant $C > 0$.

By continuity of the state with respect to the control and continuous differentiability of ϕ with respect to (u, x), $\alpha^{u,v,\varepsilon}$ and $\beta^{u,v,\varepsilon}$ converge point-wise to 0 when ε goes to 0 and are uniformly bounded by assumption on ϕ. Therefore $\left(\int_0^T |\alpha_t^{u,v,\varepsilon}|\mathrm{d}t\right)$ converges to 0 as ε goes to 0, by Lebesgue's domination theorem. Besides, $|\beta^{u,v,\varepsilon}||v|$ converges point-wise to 0 since $\{s \in [0, T] : |v_s| < +\infty\}$ has a full Lebesgue measure (recall that a.s. $\int_0^T |v_s|^2\mathrm{d}s < +\infty$) and is dominated by $|v|$, which is a.s. integrable. Applying again Lebesgue's domination theorem yields the convergence of $\int_0^T |\beta_t^{u,v,\varepsilon}||v_t|\mathrm{d}t$ to 0 when ε goes to 0.

Putting everything together, one gets that a.s. $\sup_{t\in[0,T]} |RX_t^{u,v,\varepsilon}| \to 0$ as $\varepsilon \to 0$, i.e. $\partial_\varepsilon X_t^{u+\varepsilon v}\big|_{\varepsilon=0} = \dot{X}_t^{u,v}$ a.s.

f) We now switch to the differentiability of $\mathcal{J}(u + \varepsilon v)$ w.r.t. ε at $\varepsilon = 0$. Similar arguments as before yield

$$|\partial_\varepsilon X_t^{u+\varepsilon v}| \leq C \int_0^t |v_s|\mathrm{d}s, \tag{3.4.5}$$

and for $\varepsilon \in [-1, 1]$,

$$|X_t^{u+\varepsilon v}| \leq |X_t^u| + |\varepsilon\Delta X_t^{u,v,\varepsilon}| \leq |X_t^u| + |\Delta X_t^{u,v,\varepsilon}| \leq |X_t^u| + C\int_0^t |v_s|\mathrm{d}s \tag{3.4.6}$$

$\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. for a constant C independent from $\varepsilon \in [-1, 1]$.

The above, combined with $X^u \in \mathbb{H}^{\infty,2}$, $u \in \mathbb{H}^{2,2}$, the smoothness of the functions l, g, ψ, k with the bounds on their derivatives (assumptions **(H.l)**-**(H.g)**-**(H.ψ)**-**(H.k)**) allow to apply the Lebesgue differentiation theorem and to obtain

$$\dot{\mathcal{J}}(u,v) = \mathbb{E}\Bigg[\int_0^T \Big(l_u(\theta_t^u, \bar{g}_t^u)v_t + l_x(\theta_t^u, \bar{g}_t^u)\dot{X}_t^{u,v} + l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\mathbb{E}\left[g_u(\theta_t^u)v_t + g_x(\theta_t^u)\dot{X}_t^{u,v}\right]\Big)\mathrm{d}t$$
$$+ \psi_x\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\dot{X}_T^{u,v} + \psi_{\bar{k}}\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\mathbb{E}\left[k_x(X_T^u)\dot{X}_T^{u,v}\right]\Bigg].$$

Using Fubini's theorem and reorganizing terms, we get

$$\dot{\mathcal{J}}(u,v) = \mathbb{E}\Bigg[\int_0^T \Big(\Big\{l_u(\theta_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\right]g_u(\theta_t^u)\Big\}v_t$$
$$+ \Big\{l_x(\theta_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\right]g_x(\theta_t^u)\Big\}\dot{X}_t^{u,v}\Big)\mathrm{d}t$$
$$+ \Big\{\psi_x\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right) + \mathbb{E}\left[\psi_{\bar{k}}\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right]k_x(X_T^u)\Big\}\dot{X}_T^{u,v}\Bigg].$$

Apply now the Itô lemma to $Y_t^u \cdot \dot{X}_t^{u,v}$ between $t = 0$ and $t = T$, with

$$Y_T^u \cdot \dot{X}_T^{u,v} = \Big\{\psi_x\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right) + \mathbb{E}\left[\psi_{\bar{k}}\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right]k_x(X_T^u)\Big\}\dot{X}_T^{u,v}, \qquad Y_0^u \cdot \dot{X}_0^{u,v} = 0,$$

note that $\dot{X}^{u,v}$ has finite variations, combine with (3.2.5) and (3.4.2), and take the expectation: it gives

$$\dot{\mathcal{J}}(u,v) = \mathbb{E}\Bigg[\int_0^T \Big\{l_u(\theta_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\right]g_u(\theta_t^u) + (Y_t^u)^\top \phi_u(\theta_t^u)\Big\}v_t\mathrm{d}t\Bigg].$$

The formula is also valid for Y_{t-}^u since the jumps of Y^u are countable. Theorem 3.2.2 is proved. □

3.4.2 Proof of Theorem 3.2.4

In the proof, $T \leq 1$ and C denotes a generic (deterministic) constant which only depends on the bounds in the hypothesis (and not on T). For $u^{(1)}$ and $u^{(2)}$ in $\mathbb{H}^{2,2}$, if a process or variable F^u depends on u we write $F^{(1)} := F^{u^{(1)}}$ and $F^{(2)} := F^{u^{(2)}}$. Besides, for any function, operator or process F which depends on u, X^u,..., we write $\Delta F := F^{(2)} - F^{(1)}$.

The proof is decomposed into several steps.

1. First, notice that by our assumptions on ϕ, L^u (resp. $\tilde{L}^u = ((L^u)^{-1})^\top$) (defined in Theorem 3.2.2) is independent from u, therefore, we simply write L (resp. $\tilde{L}$) instead. Using Lemma 3.4.1, L and $\tilde{L}$ are bounded by constants.

2. Consider the application

$$\Theta^{(X)} : \begin{cases} \mathbb{H}_{\mathcal{P}}^{2,2} & \to \mathbb{H}^{\infty,2} \\ u & \mapsto X^u \end{cases}.$$

It is well-defined, since we have already seen that $X^u \in \mathbb{H}^{\infty,2}$ whenever $u \in \mathbb{H}_{\mathcal{P}}^{2,2}$. We want to show that $\Theta^{(X)}$ is Lipschitz continuous and its Lipschitz constant is such that

$$C_{\Theta^{(X)},u}(T) = O\left(\sqrt{T}\right) \quad (T \to 0).$$

Using assumption **(H.ϕ.2)** and computations as in (3.4.2):

$$\Delta X_t = \int_0^t \Delta\phi_s \mathrm{d}s = \int_0^t \left(a_s^{(\phi)}\Delta X_s + b^{(\phi)}(s, u_s^{(2)}) - b^{(\phi)}(s, u_s^{(1)})\right)\mathrm{d}s$$

$$= \int_0^t (L_t)^{-1} L_s (b^{(\phi)}(s, u_s^{(2)}) - b^{(\phi)}(s, u_s^{(1)}))\mathrm{d}s.$$

Therefore by assumption on ϕ, we get

$$|\Delta X_t|^2 \leq C\left(\int_0^T |\Delta u_s|\mathrm{d}s\right)^2 \leq CT\int_0^T |\Delta u_s|^2\mathrm{d}s,$$

whence, $\|\Delta X\|^2_{\mathbb{H}^{\infty,2}} \leq CT\|\Delta u\|^2_{\mathbb{H}^{2,2}}$ and the Lipschitz continuity of $\Theta^{(X)}$ as announced.

3. Now consider

$$\Theta^{(Y)} : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} & \to \mathbb{H}^{\infty,2} \\ u & \mapsto Y^u \end{cases},$$

with Y^u as in (3.2.4). Theorem 3.2.2 guarantees that $\Theta^{(Y)}$ is well defined. Let us prove that it is Lipschitz continuous and its Lipschitz constant is such that

$$C_{\Theta^{(Y)},u}(T) = \mathcal{O}\left(\sqrt{T}\right) \quad (T \to 0).$$

Using the hypothesis on ϕ, g and k and the notation $\theta_s^u = (s, u_s, X_s^u)$, we get $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e.$

$$\begin{aligned}
\Delta Y_t =& \mathbb{E}_t\left[\tilde{L}_t^{-1}\tilde{L}_T\left(\nabla_x\psi(X_T^{(2)}, \mathbb{E}\left[k(X_T^{(2)})\right]) - \nabla_x\psi(X_T^{(1)}, \mathbb{E}\left[k(X_T^{(1)})\right])\right)\right] \\
&+ \mathbb{E}_t\left[\tilde{L}_t^{-1}\tilde{L}_T(a^{(k)})^\top\mathbb{E}\left[\nabla_{\bar{k}}\psi(X_T^{(2)}, \mathbb{E}\left[k(X_T^{(2)})\right]) - \nabla_{\bar{k}}\psi(X_T^{(1)}, \mathbb{E}\left[k(X_T^{(1)})\right])\right]\right] \\
&+ \mathbb{E}_t\left[\int_t^T \tilde{L}_t^{-1}\tilde{L}_s\left(\nabla_x l(\theta_s^{(2)}, \mathbb{E}\left[g(\theta_s^{(2)})\right]) - \nabla_x l(\theta_s^{(1)}, \mathbb{E}\left[g(\theta_s^{(1)})\right])\right)\mathrm{d}s\right] \\
&+ \mathbb{E}_t\left[\int_t^T \tilde{L}_t^{-1}\tilde{L}_s(a_s^{(g)})^\top\mathbb{E}\left[\nabla_{\bar{g}} l(\theta_s^{(2)}, \mathbb{E}\left[g(\theta_s^{(2)})\right]) - \nabla_{\bar{g}} l(\theta_s^{(1)}, \mathbb{E}\left[g(\theta_s^{(1)})\right])\right]\mathrm{d}s\right].
\end{aligned}$$

Now, owing to assumptions, the Cauchy-Schwartz inequality, the previous estimate on $\tilde{L}$, on its inverse and $\|\Delta X\|_{\mathbb{H}^{\infty,2}}$, the inequality $\|\cdot\|_{\mathbb{H}^{2,2}} \leq \sqrt{T}\|\cdot\|_{\mathbb{H}^{\infty,2}}$, one gets:

$$\|\Delta Y\|^2_{\mathbb{H}^{\infty,2}} \leq C\left(\mathbb{E}\left[|\Delta X_T|^2\right] + T\left\{\|\Delta u\|^2_{\mathbb{H}^{2,2}} + \|\Delta X\|^2_{\mathbb{H}^{2,2}}\right\}\right) \leq CT\|\Delta u\|^2_{\mathbb{H}^{2,2}}.$$

This yields the Lipschitz continuity of $\Theta^{(Y)}$ with $C_{\Theta^{(Y)},u}(T) = \mathcal{O}\left(\sqrt{T}\right) \quad (T \to 0)$.

4. Our goal is to prove that:

$$\Theta : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} & \to \mathbb{H}^{2,2}_{\mathcal{P}} \\ u & \mapsto \tilde{u} \end{cases},$$

with

$$\begin{aligned}
&\tilde{u}_t = h\left(t, X_t^u, Y_{t-}^u, \bar{g}_t^u, \mathbb{E}\left[\nabla_{\bar{g}} l(t, u_t, X_t^u, \bar{g}_t^u)\right]\right), \\
&X^u = \Theta^{(X)}(u), \qquad Y^u = \Theta^{(Y)}(u), \qquad \bar{g}_t^u = \mathbb{E}\left[g(t, u_t, X_t^u)\right]
\end{aligned}$$

is well defined, Lipschitz continuous and its Lipschitz constant satisfies:

$$C_{\Theta,u}(T) = C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right) + \mathcal{O}(T) \quad (T \to 0).$$

By construction, $\tilde{u}$ is predictable. Besides, for $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$,

$$\mathbb{E}\left[\int_0^T |\tilde{u}_s|^2\mathrm{d}s\right] \leq \mathbb{E}\left[\int_0^T \left(|h(s,0,0,0,0)| + (C_{h,\bar{g}} + C_{h,\bar{\lambda}}C_{l_{\bar{g}},\bar{g}})\mathbb{E}\left[|g(s,0,0)|\right] + C_{h,x}|X_s^u|\right.\right.$$

$$+ C_{h,y}|Y^u_s| + \left(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\right)\mathbb{E}\left[|u_s|\right]$$
$$+ \left(C_{h,\bar{g}}C_{g,x} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},x} + C_{l_{\bar{g}},\bar{g}}C_{g,x}\right)\right)\mathbb{E}\left[|X^u_s|\right]\Big)^2 \mathrm{d}s\Bigg].$$

Using Minkowski's inequality, this shows that the right-hand side is finite since $g(\cdot,0,0) \in \mathbb{H}^{2,1}$, $h(\cdot,0,0,0,0)$, X^u, Y^u and u are in $\mathbb{H}^{2,2}$, whence the well-posedness of Θ. Similar computations give

$$\|\Delta\tilde{u}\|_{\mathbb{H}^{2,2}} \leq \Bigg(\mathbb{E}\Bigg[\int_0^T \Big(C_{h,x}|\Delta X_s| + \left(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\right)\mathbb{E}\left[|\Delta u_s|\right]$$
$$+ C_{h,y}|\Delta Y_s| + \left(C_{h,\bar{g}}C_{g,x} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},x} + C_{l_{\bar{g}},\bar{g}}C_{g,x}\right)\right)\mathbb{E}\left[|\Delta X_s|\right]\Big)^2 \mathrm{d}s\Bigg]\Bigg)^{1/2}.$$

Again, from Minkowski's inequality it follows that

$$\|\Delta\tilde{u}\|_{\mathbb{H}^{2,2}} \leq \left(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\right)\|\Delta u\|_{\mathbb{H}^{2,2}} + C\Big(\|\Delta X\|_{\mathbb{H}^{2,2}} + \|\Delta Y\|_{\mathbb{H}^{2,2}}\Big).$$

Using $\|\cdot\|_{\mathbb{H}^{2,2}} \leq \sqrt{T}\|\cdot\|_{\mathbb{H}^{\infty,2}}$ and our estimates on $\|\Delta X\|_{\mathbb{H}^{\infty,2}}$ and $\|\Delta Y\|_{\mathbb{H}^{\infty,2}}$, we obtain that Θ is Lipschitz continuous and its Lipschitz constant satisfies:

$$C_{\Theta,u}(T) \leq C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right) + \mathcal{O}(T) \quad (T \to 0).$$

5. Under assumption (3.2.8), for T small enough, Θ is a contraction in the complete space $\mathbb{H}^{2,2}_{\mathcal{P}}$ and has therefore a unique fixed point u in $\mathbb{H}^{2,2}_{\mathcal{P}}$.
6. To conclude, notice (3.2.1) - (3.2.4) - (3.2.6) are satisfied by (u, X^u, Y^u) with $X^u = \Theta^{(X)}(u)$ and $Y^u = \Theta^{(Y)}(u)$ if and only if u is a fixed point of Θ. □

3.4.3 Proof of Proposition 3.2.6

▷ The continuity and coercivity of F are obvious. Similar computations as in the proof of Theorem 3.2.2 show that F is Gateaux-differentiable and that the Gateaux-derivative of F at u in direction v is given by:

$$\dot{F}(u,v) = 4(\|u\|^2_{\mathbb{L}^2_1} - 1)\int_0^1 u_t v_t \mathrm{d}t + 2\int_0^1 t u_t v_t \mathrm{d}t := \int_0^1 \mathcal{F}(u)_t v_t \mathrm{d}t,$$

where $\mathcal{F} : \mathbb{L}^2_1 \mapsto \mathbb{L}^2_1$ is defined by $\mathcal{F}(u) : t \mapsto \left(4(\|u\|^2_{\mathbb{L}^2_1} - 1) + 2t\right)u_t$.

▷ Let us identify the critical points $u^\star \in \mathbb{L}^2_1$: for such element, we must have $\dot{F}(u^\star, \mathcal{L}(u^\star)) = \int_0^1 |\mathcal{F}(u^\star)_t|^2 \mathrm{d}t = 0$, which implies $(4(\|u^\star\|^2_{\mathbb{L}^2_1} - 1) + 2t)u^\star_t = 0$ a.e. on $[0,1]$. Clearly, it leads to $u^\star_t = 0$ a.e. on $[0,1]$ and therefore, 0 is the unique critical point of F.

▷ Let us show that $\inf_{u\in\mathbb{L}^2_1} F(u) = 0$. Since F takes values in $\mathbb{R}^+$, it is enough to exhibit a sequence $u^{(n)} \in \mathbb{L}^2_1$ s.t. $F(u^{(n)}) \to 0$ as $n \to +\infty$. Define, $\forall n \in \mathbb{N}$,

$$u^{(n)} : t \mapsto \sqrt{n+1}\mathbb{1}_{[0,\frac{1}{n+1}]}(t).$$

Then,

$$\int_0^1 |u^{(n)}_t|^2 \mathrm{d}t = 1, \qquad \int_0^1 t|u^{(n)}_t|^2 \mathrm{d}t = \int_0^{1/(n+1)} (n+1)t\mathrm{d}t = \frac{1}{2(n+1)},$$

therefore $F(u^{(n)}) = \frac{1}{2(n+1)} \to 0$, as it was sought.

▷ Last, we prove that the minimum is not achieved. Assume the contrary with the existence of $u^\star \in \mathbb{L}^2_1$ s.t. $F(u^\star) = 0$. We must have $\|u^\star\|_{\mathbb{L}^2_1} = 1$ and $t|u^\star_t|^2 = 0$ a.e. on $[0,1]$: the second condition requires $u^\star = 0$ which is incompatible with the first condition. We are done, F does not have a minimum. □

3.4.4 Proof of Proposition 3.3.2

Usual results about the solution to affine-linear FBSDEs hold for Brownian filtration, see [Yon06] for instance. Here, we consider more general filtrations, but the arguments are quite similar. For the sake of completeness, we give the proof.

The function θ is the unique solution of the following affine-linear second order ODE

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} - bag\theta_t = 0 \text{ for } t \in [0,T], \\ \theta_T = 1, \\ \frac{\mathrm{d}\theta_t}{\mathrm{d}t}|_{t=T} = -eag, \end{cases}$$

and does not vanish on $[0,T]$ according to the sign conditions on the coefficients. We can therefore define p as in (3.3.3). Besides, p and θ are continuous and bounded on $[0,T]$. By standard arguments, one can check that p is the unique solution of the following Riccati ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} - agp_t^2 + b = 0, \\ p_T = e. \end{cases} \tag{3.4.7}$$

Define the following BSDE:

$$\begin{cases} -\mathrm{d}\tilde{\pi}_t = -(agp_t\tilde{\pi}_t + ap_th_t - c)\mathrm{d}t - \mathrm{d}M_t, \\ \tilde{\pi}_T = f, \end{cases}$$

which has a unique solution in $\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}$ (see [EPQ97, Theorem 5.1 with $p = 2$]) in our context of general filtrations. By the integration by parts formula applied to $\tilde{\pi}_t \exp\left(\int_t^T agp_s\mathrm{d}s\right)$ (see [Pro03, Corollary 2, p.68]), we get that $\tilde{\pi}$ is also given by:

$$\begin{aligned} \tilde{\pi}_t &= \mathbb{E}_t\left[f\exp\left(-\int_t^T agp_\tau\mathrm{d}\tau\right) - \int_t^T (ap_sh_s - c)\exp\left(-\int_t^s agp_\tau\mathrm{d}\tau\right)\mathrm{d}s\right] \\ &= \mathbb{E}_t\left[f\exp\left(\int_t^T \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right) - \int_t^T (ap_sh_s - c)\exp\left(\int_t^s \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right)\mathrm{d}s\right] \\ &= \mathbb{E}_t\left[f\frac{\theta_T}{\theta_t} - \int_t^T (ap_sh_s - c)\frac{\theta_s}{\theta_t}\mathrm{d}s\right] \\ &= \frac{1}{\theta_t}\left(f - \int_t^T (ap_s\mathbb{E}_t[h_s] - c)\theta_s\mathrm{d}s\right) = \pi_t, \end{aligned}$$

where we used the definitions of p and π.

We deduce that the process π also has the following representation:

$$\pi_t = \mathbb{E}_t\left[f - \int_t^T (agp_s\pi_s + ap_sh_s - c)\mathrm{d}s\right]. \tag{3.4.8}$$

Since θ and p are bounded on $[0,T]$, we easily prove that $\pi \in \mathbb{H}^{\infty,2}$. From that and our assumptions on the data of the problem, it is clear that (x,y,v) as defined in (3.3.5) belong to $\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{2,2}_{\mathcal{P}}$. In particular, v is predictable since x is continuous by construction.

We now prove that (x,y,v) defined by (3.3.5) solves (3.3.6). By definition of y and v in (3.3.5), we can check that:

$$v_t = gy_{t-} + h_t.$$

Define $\tilde{x}$ the unique solution of the following affine-linear ODE:

$$\tilde{x}_t := x_0 - \int_0^t (agp_s\tilde{x}_s + ag\pi_s + ah_s)\mathrm{d}s. \tag{3.4.9}$$

It is also given by:

$$\begin{aligned}
\tilde{x}_t &= x_0 \exp\left(-\int_0^t agp_\tau \mathrm{d}\tau\right) - \int_0^t \left\{(ag\pi_s + ah_s)\exp\left(-\int_s^t agp_\tau \mathrm{d}\tau\right)\right\}\mathrm{d}s \\
&= x_0 \exp\left(\int_0^t \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right) - \int_0^t \left\{(ag\pi_s + ah_s)\exp\left(\int_s^t \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right)\right\}\mathrm{d}s \\
&= x_0\frac{\theta_t}{\theta_0} - \int_0^t (ag\pi_s + ah_s)\frac{\theta_t}{\theta_s}\mathrm{d}s \\
&= x_t;
\end{aligned}$$

hence, x is the unique solution of (3.4.9). Since π has countably many jumps and changing the Lebesgue integral is left unchanged by changing the integrand at countably many points, we then get by definition of v:

$$x_t = x_0 - \int_0^t av_s \mathrm{d}s.$$

It remains to show that the second equation in (3.3.6) is verified. Using that p is solution of (3.4.7), π verifies (3.4.8) and x is solution of (3.4.9), we get:

$$\begin{aligned}
y_t &= p_t x_t + \pi_t \\
&= \mathbb{E}_t\left[p_T x_T - \int_t^T \left(\frac{\mathrm{d}p_s}{\mathrm{d}s}x_s + \frac{\mathrm{d}x_s}{\mathrm{d}s}p_s\right)\mathrm{d}s + \pi_T - \int_t^T (agp_s\pi_s + ap_s h_s - c)\mathrm{d}s\right] \\
&= \mathbb{E}_t\left[ex_T + f + \int_t^T \left\{-(agp_s^2 - b)x_s + (agp_s x_s + ag\pi_s + ah_s)p_s - (agp_s\pi_s + ap_s h_s - c)\right\}\mathrm{d}s\right] \\
&= \mathbb{E}_t\left[ex_T + f + \int_t^T (bx_s + c)\,\mathrm{d}s\right].
\end{aligned}$$

This shows that (v, x, y) is a solution of (3.3.6) in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{2,2}_{\mathcal{P}}$.
The uniqueness of the solution for small time T follows from a fixed point argument as in the proof of Theorem 3.2.4. We do not repeat the arguments here. □

3.4.5 Proof of Proposition 3.3.6

▷ By definition of $\dot{u}^{(\varepsilon)}$, $\dot{X}^{(\varepsilon)}$ and $\dot{Y}^{(\varepsilon)}$, they are clearly solutions of (3.3.10). Let us turn to uniform boundedness. The first two equations yield:

$$\|\dot{X}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \leq \frac{\sqrt{T}}{\mathcal{E}_{\max}}\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}, \qquad \|\dot{Y}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \leq \frac{\nu T + \gamma\sqrt{T}}{\mathcal{E}_{\max}}\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}.$$

This can be easily proved following the arguments given in the proof of Theorem 3.2.7, details are left to the reader. From the last equation of (3.3.10) and the 1-Lipschitz continuity of $x \to x_+$, we get:

$$\begin{aligned}
\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} &\leq \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}\|\dot{Y}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} + 2\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(\varepsilon)})_+\|_{\mathbb{H}^{2,2}} \\
&\leq \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2}\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} + 2\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}} + 2\varepsilon\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}.
\end{aligned}$$

When T and ε are small, such that,

$$\alpha(T) + 2\varepsilon := \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2} + 2\varepsilon < 1,$$

we obtain

$$\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \leq \frac{2\|(\mathrm{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{1 - \alpha(T) - 2\varepsilon},$$

whence the uniform boundedness of $\dot{u}^{(\varepsilon)}$ as $\varepsilon \to 0$, provided $\alpha(T) < 1$.
▷ Now we prove the convergence of $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ to $(\dot{u}, \dot{X}, \dot{Y})$ in $\mathbb{H}^{2,2}$-norms as $\varepsilon \to 0$. Similarly as before, we have:

$$\|\dot{X}^{(\varepsilon)} - \dot{X}\|_{\mathbb{H}^{2,2}} \leq \frac{\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}}, \qquad \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} \leq \frac{\nu T + \gamma\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}}.$$

Besides, the last equations in (3.3.10) and (3.3.11) as well as the 1-Lipschitz continuity of $x \to x_+$ give:

$$\begin{aligned}
\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} &\leq \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}} \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} + 2\|u^{\Delta,(\varepsilon)} - u^{\Delta,(0)}\|_{\mathbb{H}^{2,2}} \\
&\leq \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} + 2\varepsilon\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \\
&= \alpha(T)\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} + 2\varepsilon\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}.
\end{aligned}$$

For T small enough s.t. $\alpha(T) < 1$ and and for $\varepsilon < \frac{1-\alpha(T)}{2}$, we thus obtain:

$$\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} \leq \frac{2\varepsilon}{1-\alpha(T)} \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \leq \frac{4\varepsilon}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)} \|(\mathbf{P}^{\mathrm{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}.$$

This completes the proof. □

3.4.6 Boundedness of solutions to linear ODE with bounded stochastic coefficient

This following result is used in the proof of Theorems 3.2.2 and 3.2.4.

Lemma 3.4.1. *Let $A : [0,T] \times \Omega \mapsto \mathbb{R}^p \times \mathbb{R}^p$ be a random matrix-valued process. Suppose there exists a constant C such that $|A(t,\omega)| \leq C$, $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. .*
Let R and L be the unique (continuous) solutions of the following linear ODEs:

$$\begin{cases} \frac{\mathrm{d}L_t}{\mathrm{d}t} = L_t A_t, \\ L_0 = \mathrm{Id}_p, \end{cases} \qquad \text{and} \qquad \begin{cases} \frac{\mathrm{d}R_t}{\mathrm{d}t} = -A_t R_t, \\ R_0 = \mathrm{Id}_p. \end{cases}$$

Then, L and R are invertible with $L^{-1} = R$. Besides, $|L_t|$ and $|R_t|$ are uniformly bounded on $[0,T]$ by $\exp(CT)$.

Proof. A direct computation shows that

$$\mathrm{d}t \times \mathrm{d}\mathbb{P}\text{-a.e.}, \qquad \frac{\mathrm{d}(L_t R_t)}{\mathrm{d}t} = 0,$$

thus $\forall t \in [0,T]$, $L_t R_t = L_0 R_0 = \mathrm{Id}_p$. Therefore R and L are invertible with $R = L^{-1}$. Let us now turn to the uniform boundedness. Let $v \in \mathbb{R}^p$, we have

$$\frac{\mathrm{d}|L_t^\top v|^2}{\mathrm{d}t} = v^\top L_t (A_t + A_t^\top) L_t^\top v \leq |A_t + A_t^\top|\,|L_t^\top v|^2 \leq 2C\,|L_t^\top v|^2, \quad \mathrm{d}t \times \mathrm{d}\mathbb{P}\text{-a.e.}.$$

Therefore, by integration, $|L_t^\top v|^2 \leq |v|^2 \exp(2CT)$ for $t \in [0,T]$, which yields $\sup_{0\leq t\leq T} |L_t^\top|^2 \leq \exp(2CT)$. This proves $\exp(CT) \geq \sup_{0\leq t\leq T} |L_t^\top| = \sup_{0\leq t\leq T} |L_t|$, whence the announced bound for L. For bounding $|R|$ start from $\frac{\mathrm{d}|R_t v|^2}{\mathrm{d}t}$ and proceed similarly. □

3.5 Conclusion

In this work, we have identified the optimal control of storage facilities of a smart grid under uncertain consumption/production, in order to reduce the stochastic fluctuations of the residual consumption on the electrical public grid. It has been possible thanks to the resolution of a new extended McKean-Vlasov stochastic control problem, using

Pontryagin principle and Forward Backward Stochastic Differential Equations. For situations where the costs are close to quadratic functions, we have derived quasi-explicit formulas for the control, using perturbation arguments.

In further works, we will consider subsequent issues like more realistic dynamics of the battery flow accounting with aging/boundary effect, sizing of the smart grid and of the storage/production capacities, risk aggregation of optimized smart grids with dependent solar productions, impact of model mis-specification on the optimal solution (risk model).

Chapter 4

Decentralized control of heterogeneous energy storage systems in a stochastic environment

4.1 Introduction

Context. To meet the goal of low carbon footprint for mitigating the climate change, the energy sector is seeking solutions for better energy-efficiency. Among them, the use of renewable energy (like solar or wind power) is appealing but on the other hand, the intermittency of their production raises challenges to satisfy the power balance between production and consumption. In this work, we focus on demand-side flexibilities, more specifically, Thermostatically Control Loads (TCLs) like for instance, fridges, air conditioners, hot water tanks, swimming pool heaters... These devices aim at maintaining a set-point temperature, but the realized temperature X has an inertia and tolerates a range of admissible values, which gives flexibility in controlling the appliances [BM16]. Leveraging this flexibility to provide services to the grid has an enormous potential [Mat+12; Cam+18b].

Statement of the problem and objectives. In this chapter, we consider the problem of optimally controlling a large population of N TCLs owned by individual consumers (also called agents), in a stochastic environment modeled by a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. From the application perspective, it corresponds to a centralized control architecture, where an omniscient planner solves a high-dimensional control problem in order to both minimize operational costs and promote energy balance.
The model will incorporate a common weather noise for all agents. This weather noise models the exogenous conditions that impact both the solar production through the irradiance [Bad+18] and the household consumption through non-constant lighting, heating, cooling, see [PMV02]. Conditionally to the common weather noise, the net consumption (household consumption without flexible appliance minus solar production) of different agents will be assumed independent. This allows us to account for spatial correlations of meteorological conditions [ADS99; ZDK16] between different agent locations, for instance.
To allow for general and realistic situations, the agents will differ w.r.t. their flexibilities and their consumption/production (difference of size and habits of the households, of renewable energy equipment, of appliances, etc). Additionally, we will assume that the agents are split in M classes in which the agents' flexibilities share the same physical characteristics (similar appliance), see Section 4.2 for the precise modeling. The net consumption of the i-th agent in the k-th class is denoted by $(\mathrm{P}^{\mathtt{load},(k,i)})_{k\in[M],i\in[N_k]}$ using the notation $[n] := \{1,\dots,n\}$ for any integer $n \geq 1$. The power consumed by its flexible appliance is denoted by $u^{(k,i,N)}$; hence, its total consumption is $\mathrm{P}^{\mathtt{load},(k,i)} + u^{(k,i,N)}$. Given a average (per agent) power $\mathrm{P}^{\mathtt{prod}}$ available on the public grid, the average power imbalance is $\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u^{(l,j,N)} + \mathrm{P}^{\mathtt{load},(l,j)}) - \mathrm{P}^{\mathtt{prod}}$. Although not mathematically essential, $\mathrm{P}^{\mathtt{prod}}$ is assumed deterministic for simplicity, which makes sense since it comes from power demand forecast made by the planner (and usually supplied by conventional generation units).

In addition, all the above quantities depend on time.
The optimization criterion will consist in minimizing the average power imbalance over a finite interval $[0, T]$ ($T > 0$ fixed), while maintaining each flexible power $u^{(k,i,N)}$ around a nominal value $u^{\mathtt{ref},(k,i)}$ and the temperature $X^{(k,i,N)}$ around a set-point temperature $x^{\mathtt{ref},(k,i)}$. All in all, the controls are $(u^{(k,i,N)})_{k\in[M],i\in[N_k]}$, adapted to the ambient filtration $\mathbb{F}$, and the cost functional takes the form

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k,i,N)}-u_t^{\mathtt{ref},(k,i)}\right)^2+\frac{\nu_t^{(k)}}{2}\left(X_t^{(k,i,N)}-x_t^{\mathtt{ref},(k,i)}\right)^2\right)\mathrm{d}t+\frac{\rho^{(k)}}{2}\left(X_T^{(k,i,N)}-x_T^{\mathtt{f},(k,i)}\right)^2\right\}\right]$$
$$+\,\mathbb{E}\left[\int_0^T\mathcal{L}_t\left(\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u_t^{(l,j,N)}+\mathrm{P}_t^{\mathtt{load},(l,j)})-\mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right]. \tag{4.1.1}$$

The fact that the parameters $\mu^{(k)}$, $\nu^{(k)}$, $\rho^{(k)}$ may depend on the class of device allows to model heterogeneity among the devices. For instance, a fridge may not have the same temperature dead-band tolerance as a heat pump, which justifies to consider different values of the parameter $\nu^{(k)}$ for these two classes of devices. Actually, we are looking for a solution method where agents can keep their individual data private. This privacy preservation is nowadays a major concern in grid management, see [AA19] for a recent overview and references therein. We are also interested in deriving a practical implementation of the control where minimal communication between actors is required.

Methodology and main contributions. Mathematically speaking, our problem fits the setting of stochastic control problem in high dimension (the number N of agents). The dynamics for each state variable $X^{(k,i,N)}$ (modeling the temperature inertia) is linear w.r.t. the control (appliance power): for each agent, it typically writes under the form[1] [2]

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \underbrace{u_t/C}_{\text{appliance power consumption}} - \underbrace{(X_t - X_t^{out})/RC}_{\text{thermal losses}} + \underbrace{\Lambda_t}_{\text{exogenous perturbations}}. \tag{4.1.2}$$

The filtration can be quite general and allows, for instance, for jump processes in the net consumption. The stochastic (Pontryagin) maximum principle enables us to characterize the optimal controls as solution to a coupled system of Forward-Backward Stochastic Differential Equations (FBSDE), see Theorem 4.2.1. However, the FBSDE is a high-dimensional coupled equation, and hence curse of dimensionality occurs.

To overcome this, we design a decoupled system, with a so-called coordination problem and individual problems associated to each agent. The coordination problem is a FBSDE which can be interpreted as the optimality conditions of a (convex) optimal control problem that a coordinator has to solve to compute a coordination signal. Each individual problem is an FBSDE which can be interpreted as the optimality conditions of the control problem of a selfish agent controlling (with locally available information) its individual storage system to minimize operational costs, while responding to the coordination signal. In other words, we show that the optimal solution of the control problem of the central planner corresponds to the (unique) Nash equilibrium of a stochastic Stackelberg differential game, which allows for possible decentralized control schemes. This is in some way the inverse perspective of potential stochastic differential games [FMHL19], in which one seeks a stochastic control problem which optimality conditions coincide with the Nash system of a given stochastic differential game. In order to avoid the need for real-time communication from agents to the coordinator, we design a mean-field-type approximation of the decoupled system: the approximation of the coordination problem mainly depends on the population statistics and not anymore on the data of agents. Under the assumption that the size of each class of agents is large enough and under the assumption that agents are conditionally independent in some sense, we prove that this new decoupled system yields a control which preserves privacy and converges to the omniscient control in the limit of infinite population (see Theorem 4.4.12). Error bounds on performance loss are also derived. Besides, there again, we can interpret the approximations of the coordination and individual problems as the system of FBSDE characterizing the (unique)

[1] the coefficient C is the calorific capacity of the system C, R is the thermal resistance of the system, X^{out} is the temperature of the environment, and Λ models random perturbations like opening the fridge, using the hot water tank for a shower etc

[2] the affine-linear dynamic is in agreement with the "leaky battery model" presented in [Hao+14; TTS16] and with the first-order dynamical model used to model the temperature evolution of a TCL in [DP+19].

equilibrium of a stochastic Stackelberg differential game with a leader, the coordinator, and many (non-symmetric) followers, the agents. To get the convergence results and error bounds, we leverage the conditional Law of Large Numbers, stability results for FBSDEs and probabilistic properties related to immersion of filtration (to deal with the common noise). The new control boils down to solving a system of $M + 1 \ll N$ weakly coupled FBSDEs, and thus one suffers much less from the curse of dimensionality than what may be expected. We illustrate these results numerically on an example involving a large population of two types of devices (water heaters and heat pumps) with realistic characteristics. In this chapter, we also discuss a practical online and decentralized implementation of the privacy-preserving control, see Algorithm 4.1. It allows real-time computations of a coordination signal by a coordinator. Broadcasting this signal allows each agent to compute its optimal response in real-time. In particular, the coordination signal sent at each time t to all agents is a function of time measurable with respect to the information available to the coordinator at time t. Besides, the problems solved online by the agents are easy to solve and only require information available locally. This allows to preserve the privacy of individual consumers and maintain quality of service if communication loss occurs.

Literature background. Stochastic control of large population of micro-grids has recently drawn significant interest, but standard methods like Stochastic Dynamic Programming, Stochastic Dual Dynamic Programming (in a discrete-time setting) suffer rapidly from the curse of dimensionality, when considering more than a few tens of micro-grids, even when considering spatial decomposition techniques [Car+19b; Car+20]. To tackle this issue, mean-field approximations are particularly promising, as they become more accurate when the number of agents grows. There is a recent and abundant literature about stochastic control with large population, commonly known as mean-field games (MFG)/McKean-Vlasov (MKV) stochastic control problems, see [CDL13; CD18; BFY+13] among recent contributions. Mean-Field Games models for control of large populations of micro-grids without common noise have been proposed in [DP+19; KM13; KM16]. In [DP+19], self-interested consumers allocate their flexible consumption and choose a level of participation to electricity reserves mechanisms according to price signals derived from a Unit-Commitment problem solved by a coordinator. In [KM13], water heaters are controlled so that their average profile tracks a specific profile sent by a coordinator. This model is enriched in [KM16] to consider Markovian jumps dynamics for individual water heaters and non-uniformity of the temperature within water tanks.

In the literature of MFG and MKV control, agents are usually assumed to be symmetric. Let us mention however [HMC+06; KM13; KM16; ATM20], [BFY+13, Chapter 8, pp. 67-72] which consider an heterogeneous population by introducing user classes in the setting of Mean-Field Games.
Our model is defined with a common noise, which seemingly connects our contribution to the recent developments of the theory of MFG/MKV with a common noise. In [ATM20], similarly as in our work, a setting with common and individual noises is considered, heterogeneity among agents is introduced, and the structure of equations obtained is similar. However, this work considers applications related to price-arbitrage and peak-shaving, and directly studies the Mean-Field approximation. By contrast, our work focuses on tracking power imbalance from an application point of view, and we are also interested in approximation and convergence results. The model of [ATM20] is extended to the case with jumps in [MMS19]. These works mostly consider the case of an infinite population, whereas we consider a finite number of agents.

Note that Mean-Field Games assume a competitive setting and seek for Nash equilibria, which may be far from optimal from a collective point of view. This performance loss can easily be assessed in Linear-Quadratic frameworks, as both MKV control and MFG admit explicit feedback formulas for the control, which allows to compute the Price of Anarchy. In this chapter, by contrast to the MFG frameworks for the control of micro-grids [DP+19; KM13; KM16; ATM20; MMS19], we assume a cooperative setting, so that our problem is closely related to the field of MKV stochastic control problems. We show that the optimal solution of our control problem is also the (unique) Nash equilibrium of a Stochastic Stackelberg Differential Game with a leader (the coordinator) and many heterogeneous followers (the agents), whose decisions are impacted by the decisions of the leader. Hence, we follow the inverse perspective of potential stochastic differential games [FMHL19], in which one seeks a stochastic control problem whose optimality conditions coincide with the Nash system of a given stochastic game. See [ŞC14] or [CD18, section 7.1, pp. 541-610, Volume II] for an introduction to Mean-Field Games with major and minor players.

We also mention other works presenting cooperative control architectures of TCLs, without a priori optimality guarantees. [Hao+13; Hao+14] propose a control architecture based on priority queue to decide which device to

control. A control architecture based on PDE models is proposed in [TTS15] to track a reference profile where each device builds an ensemble model for the whole population and uses this model to compute an appropriate random switching rate. Another control architecture is proposed in [Tro+16] which ensures that the aggregate consumption of a large population of TCLs depends linearly on the frequency of the network (which is a good indicator for power balance) and its rate of change. Our approach differs from these works since we use tools coming from stochastic optimal control theory, which provides a priori optimality guarantee (up to model errors).

Organization of the chapter. The probabilistic model as well as first-order necessary and sufficient optimality conditions (Theorem 4.2.1) are given in Section 4.2. A decomposition method for the optimality system is obtained in Section 4.3, with definitions of the coordination problem (Proposition 4.3.1) and individual problems (Proposition 4.3.3). Then approximations of these problems are given in Section 4.4. In particular, the solution of the approximation of the coordination problem is shown to be progressively measurable with respect to the common noise filtration, see Theorem 4.4.5, which is a desirable property in a decentralized control scheme. Indeed it allows a third party (called coordinator) to solve the approximate coordination problem without having to observe the aggregated individual parameters, circumventing privacy and telecommunication issues. Error bounds between the privacy-preserving control and the omniscient control are then presented, see Theorem 4.4.12. Section 4.5 collects a few numerical illustrations in the case of agents equipped with heat pumps or water heaters. Practical interest of the approach is demonstrated, as well as the convergence of the mean-field approximation in the limit of large populations. Then, a decentralized online implementation with minimal information sharing of the approximate solution method for the control problem is presented in Section 4.6. Some of the proofs are postponed to Section 4.7.

Most commonly used notations. We list the most common notations used in this chapter.
▷ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers and positive integers. The notation $|x|$ stands for the Euclidean norm of a vector x. For $k \in \mathbb{N}$, the notation $[k]$ stands for the integer set $\{1, ..., k\}$.
▷ *Function derivatives.* For a smooth function $g : \mathbb{R}^p \mapsto \mathbb{R}$, g'_x represents the partial derivative of g with respect to x. However, the notation x_t refers to the value of a process x at time t (and not to the partial derivative of x with respect to t).

▷ *Probability.* The randomness on the interval $[0, T]$ is modeled on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with a right-continuous filtration $\mathbb{F} := \{\mathcal{F}_t\}_{0 \le t \le T}$ augmented with the $\mathbb{P}$-null sets. We consider another filtration $\mathbb{G} \subset \mathbb{F}$ (meaning $\mathcal{G}_t \subset \mathcal{F}_t$ for all t in $[0, T]$), assumed immersed in $\mathbb{F}$ (see [CD18, Definition 1.2, p.5, Volume II]: all $\mathbb{G}$ square integrable martingale are $\mathbb{F}$-martingales): this filtration will model the structure of information for the common weather noise. The immersion property implies the independence of $\mathcal{G}_T$ and $\mathcal{F}_t$ conditionally on $\mathcal{G}_t$, for any $t \in [0, T]$. This assumption is equivalent to the fact that for any $t \in [0, T]$, for any random variable $X \in \mathbb{L}^1(\mathcal{F}_t), \mathbb{E}[X|\mathcal{G}_T] = \mathbb{E}[X|\mathcal{G}_t]$ (see [CD18, Proposition 1.3, p. 6, Volume II]).

The set of square integrable variables is denoted by $\mathbb{L}^2$. The notation $\mathbb{L}^2_T$ stands for the set of $\mathcal{F}_T$-measurable square integrable variables.
▷ *Stochastic processes.* For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$.
All the martingales are considered with their càdlàg modifications.
The space $\mathcal{S}$ (resp. $\mathcal{H}$) stands for the $\mathbb{F}$-adapted càdlàg (resp. $\mathbb{F}$-progressively measurable) processes $(\Psi_t : t \in [0, T])$ valued in an Euclidean space $\mathcal{E}$ such that $\sqrt{\mathbb{E}\left[\sup_{t\in[0,T]} |\Psi_t|^2\right]} =: \|\Psi\|_{\mathcal{S}}$ (resp. $\sqrt{\mathbb{E}\left[\int_0^T |\Psi_t|^2 \mathrm{d}t\right]} =: \|\Psi\|_{\mathcal{H}}$). Since the space $\mathcal{E}$ will be clear from the context (typically $\mathbb{R}$, $\mathbb{R}^M$ or $\mathbb{R}^N$), we will skip the reference to it in the notation. The space $\mathcal{H}_{\mathbb{G}}$ (resp. $\mathcal{S}_{\mathbb{G}}$) is the subspace of processes in $\mathcal{H}$ (resp. $\mathcal{S}$) which are $\mathbb{G}$-progressively measurable.

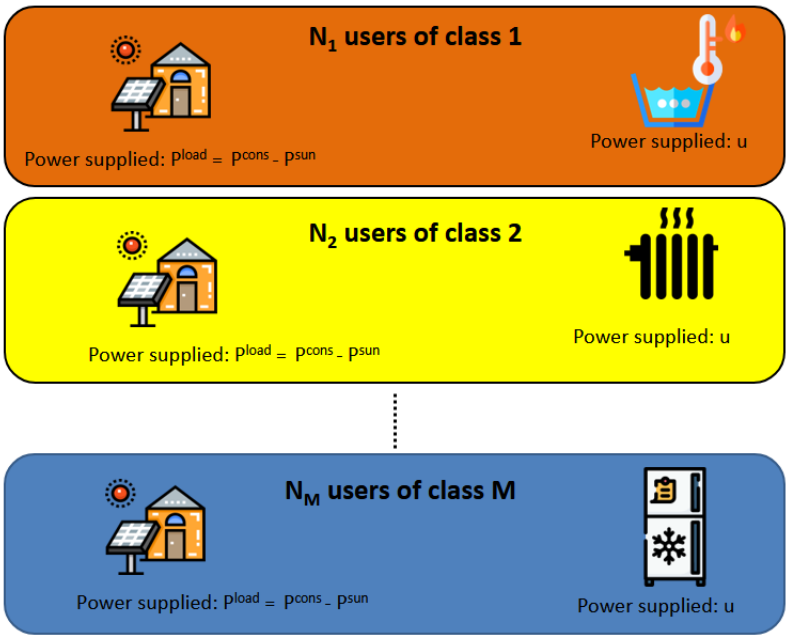
Remark 4.1.1. *By comparison with Chapter 3 where we sought the control as a predictable process in order to account for latency in the application, here, by simplicity, we will consider controls which are progressively measurable.*

4.2 Model, assumptions and first properties

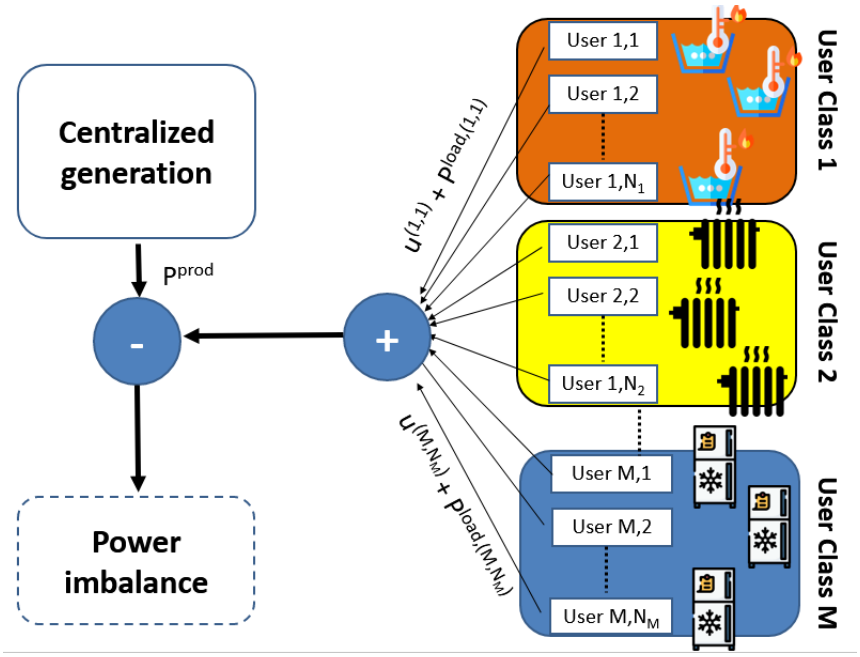
4.2.1 Assumptions

We follow the model presented in introduction with $N = \sum_{k=1}^{M} N_k$ agents split into M classes of N_k agents each (see Fig. 4.1a[3]). The state variable (temperature) for the i-th agent of the k-th class is $X^{(k,i,N)}$ and satisfies to the dynamics

[3] Icons made by Freepik and Smashicons from www.flaticon.com



(a) Description of users in the classes



(b) Power balance

Figure 4.1: Heterogeneity of agents and power imbalance

$$X_t^{(k,i,N)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i,N)} + \beta_s^{(k)} X_s^{(k,i,N)} + \gamma_s^{(k,i)}\right) \mathrm{d}s, \tag{4.2.1}$$

where $u^{(k,i,N)}$ is the control for the flexible appliance of agent (k,i). The above dynamics is consistent with the example in (4.1.2), in particular the coefficients $\alpha^{(k)}$ and $\beta^{(k)}$ are the same within the class k (similar device). This allows to incorporate heterogeneity for the devices, which do not have the same performances nor thermal behaviors. On the mathematical side, we assume from now on that

(H-X) $\alpha^{(k)}, \beta^{(k)}$ are measurable deterministic functions, uniformly bounded on $[0,T]$. Each process $\gamma^{(k,i)}$ is in $\mathcal{H}$. Each $x_0^{(k,i)}$ is deterministic.

Given the control $u^{(N)} := (u^{(k,i,N)})_{k,i} \in \mathcal{H}$, the functional to minimize is

$$\begin{aligned}\mathcal{J}(u) =& \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k,i,N)} - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k,i,N)} - x_t^{\mathtt{ref},(k,i)}\right)^2\right)\mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k,i,N)} - x_T^{\mathtt{f},(k,i)}\right)^2\right\}\right] \\ &+ \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u_t^{(l,j,N)} + \mathrm{P}_t^{\mathtt{load},(l,j)}) - \mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right],\end{aligned} \tag{4.2.2}$$

corresponding to an omniscient planner aiming to control TCLs to track the power imbalance signal, represented in Figure 4.1b, while keeping each individual flexibility around a possibly stochastic nominal state (associated to $u^{\mathtt{ref},(k,i)}$, $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$).

(H-J) $\mathrm{P}^{\mathtt{prod}}$ is a measurable deterministic function, square integrable on $[0,T]$.
All the processes $u^{\mathtt{ref},(k,i)}, x^{\mathtt{ref},(k,i)}, \mathrm{P}^{\mathtt{load},(k,i)}$ are in $\mathcal{H}$ and $x_T^{\mathtt{f},(k,i)}$ is in $\mathbb{L}_T^2$.
The coefficients $\mu^{(k)}, \nu^{(k)}$ are deterministic measurable functions, the $\rho^{(k)}$ are deterministic. They are all bounded. In addition, for some $\varepsilon > 0$, we have $\mu_t^{(k)} \geq \varepsilon$ and $\nu_t^{(k)}$ for any t and k. The function $\mathcal{L} : (t,x) \in [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ is deterministic measurable. We assume that for any $t \in [0,T]$, $x \mapsto \mathcal{L}_t(x) = \mathcal{L}(t,x)$ is convex, twice continuously differentiable, with uniformly bounded second order derivative.

4.2.2 Differentiability, convexity, characterization of optimality

We start with a somehow standard result. We show that, under our assumptions, $\mathcal{J}$ is strongly convex and admits a unique minimizer which can be obtained by solving a Forward-Backward Stochastic Differential Equation, obtained using the Stochastic Pontryagin Principle. The proof is postponed to Subsection 4.7.1.

Theorem 4.2.1. *The function $\mathcal{J} : \mathcal{H} \mapsto \mathbb{R}$ is strongly convex. It admits a unique minimizer denoted $u^{(N)} := (u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$. Define $X^{(N)} = (X^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ by (4.2.1) and $Y^{(N)} = (Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ by:*

$$Y_t^{(k,i,N)} = \mathbb{E}_t\left[\rho^{(k)}(X_T^{(k,i,N)} - x_T^{\mathtt{f},(k,i)}) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k,i,N)} + \nu_s^{(k)}(X_s^{(k,i,N)} - x_s^{\mathtt{ref},(k,i)})\right)\mathrm{d}s\right]. \tag{4.2.3}$$

Then $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ is the unique solution in $\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ of the coupled FBSDE with unknowns $(u^{(k,i)}, X^{(k,i)}, Y^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$:

$$\begin{aligned}&\forall t \in [0,T], \forall k \in [M], \forall i \in [N_k], \\ &\begin{cases} X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}\right)\mathrm{d}s, \\ Y_t^{(k,i)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k,i)} + \nu_s^{(k)}\left(X_s^{(k,i)} - x_s^{\mathtt{ref},(k,i)}\right)\right)\mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right) + \mathcal{L}_x'\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}) - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k,i)} = 0. \end{cases}\end{aligned} \tag{4.2.4}$$

Our assumptions guarantee that the optimal control problem

$$\min_{u\in\mathcal{H}} \quad \mathcal{J}(u)$$

$$s.t. \quad (4.2.1)$$

has a unique solution $u^{(N)}$, which can be equivalently computed by solving (4.2.4). This system is a high-dimensional coupled FBSDE, the dimension being the number N of agents. Hence, it suffers from the curse of dimensionality, as the number of agents may be very large. To tackle this issue, we propose a decomposition method of the FBSDE (4.2.4) in the next section.

4.3 Decomposition of the problem and equivalent representation as a stochastic differential game

Notations

- *Empirical and statistical means:* We introduce the notation $\bar{\mathrm{P}}^{\mathtt{load},(N)} := \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathrm{P}^{\mathtt{load},(l,j)}$ for the empirical mean process of net consumption of agents.
- *Empirical means over a class:* We introduce the following notations for the empirical means: $\bar{\gamma}^{(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k}\gamma^{(k,j)}$, $\bar{u}^{\mathtt{ref},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k}u^{\mathtt{ref},(k,j)}$, $\bar{x}^{\mathtt{ref},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k}x^{\mathtt{ref},(k,j)}$, $\bar{x}_T^{\mathtt{f},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k}x_T^{\mathtt{f},(k,j)}$.

In this section, we show that the control problem is equivalent to two types of control problems arising in a nested structure. We call the first problem the *coordination problem*, as it allows to compute a coordination signal. Once the coordination signal has been computed, the control problem can be decomposed into N sub-problems, the *individual problems*, each of them associated to an individual consumer/agent. The parameters of the individual problem of each agent only involve the individual data of the corresponding agent (consumption and preferences), the shared information G and the coordination signal. We also show that the coordination problem (resp. each individual problem) can be interpreted as the optimality conditions of a control problem of the coordinator (resp. of each agent). This shows that the optimal solution of the control problem corresponds to the (unique) Nash equilibrium of a stochastic differential game, allowing for a decentralized implementation.

4.3.1 The coordination problem

Proposition 4.3.1. *Consider the empirical mean processes for all $k \in [M]$:*

$$(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)}) := \left(\frac{1}{N_k}\sum_{j=1}^{N_k}u^{(k,j,N)}, \frac{1}{N_k}\sum_{j=1}^{N_k}X^{(k,j,N)}, \frac{1}{N_k}\sum_{j=1}^{N_k}Y^{(k,j,N)}\right) \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}.$$

The empirical mean process $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]}$ is the unique solution of the following FBSDE with unknowns $(u^{(k)}, X^{(k)}, Y^{(k)})_{k\in[M]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$, which we call the coordination problem:

$$\forall t \in [0,T], \forall k \in [M],$$
$$\begin{cases} X_t^{(k)} = \frac{1}{N_k}\sum_{j=1}^{N_k}x_0^{(k,j)} + \int_0^t\left(\alpha_s^{(k)}u_s^{(k)} + \beta_s^{(k)}X_s^{(k)} + \bar{\gamma}_s^{(k,N)}\right)\mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k,N)}\right) + \int_t^T\left(\beta_s^{(k)}Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k,N)}\right)\right)\mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k,N)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^{M}\pi^{(l)}u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)}Y_t^{(k)} = 0. \end{cases} \tag{4.3.1}$$

The proof of the above Proposition is postponed to Subsection 4.7.2. This proof shows that the FBSDE (4.3.1) is the optimality system of the stochastic control problem (4.7.3), which can be interpreted as the control problem of a coordinator aiming at controlling the aggregate behaviors of the agents within different classes.

Definition 4.3.2. *Let $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ be the unique solution of the coordination problem (4.3.1). We define the coordination signal $\bar{v}^{(N)} \in \mathcal{H}$ by:*

$$\forall t \in [0,T], \quad \bar{v}_t^{(N)} := \mathcal{L}_x'\left(t, \sum_{l=1}^{M}\pi^{(l)}\bar{u}_t^{(l,N)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right). \tag{4.3.2}$$

4.3.2 The individual problems

Proposition 4.3.3. *Let $\bar{v}^{(N)} \in \mathcal{H}$ be the coordination signal defined in (4.3.2). For any $k \in [M], i \in [N_k]$, consider the FBSDE with unknown $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$, called individual problem of agent i of class k:*

$$\forall t \in [0,T],$$
$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + \bar{v}_t^{(N)} + \alpha_t^{(k)} Y_t = 0. \end{cases} \tag{4.3.3}$$

Then (4.3.3) has a unique solution $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$. Besides $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M], i\in[N_k]}$ is the unique solution of the FBSDE (4.2.4) and $(u^{(k,i,N)})_{k\in[M], i\in[N_k]}$ is the unique solution of control problem $\min_{u^{(N)}\in\mathcal{H}} \mathcal{J}(u^{(N)})$.

Proof. The FBSDE (4.3.3) fully characterizes the solutions of the following stochastic control problem:

$$\min_{u\in\mathcal{H}} \quad \mathbb{E}\left[\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t - x_t^{\mathtt{ref},(k,i)}\right)^2 + \bar{v}_t^{(N)} u_t\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T - x_T^{\mathtt{f},(k,i)}\right)^2\right], \tag{4.3.4}$$
$$s.t. \quad X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s.$$

This control problem (4.3.4) has a unique solution, by similar arguments as in the proof of Proposition 4.3.1, hence so does the FBSDE (4.3.3). The fact that $(u^{(k,i,N)})_{k\in[M], i\in[N_k]}$ is the unique minimizer of $\mathcal{J}$ is a consequence of Proposition 4.3.1 and of the uniqueness of the solutions of (4.2.4), (4.3.1),(4.3.3). □

The stochastic control problem (4.3.4) can be interpreted as the control problem of agent i of class k interacting with an aggregator which sends him a coordination signal. It can be interpreted as the agent aiming at minimizing operational costs and a cost of contribution to global power imbalance, where the coordination signal $\bar{v}^{(N)}$ plays the role of a price signal cast by the coordinator. In particular, replacing $\bar{v}^{(N)}$ by its expression, the above Proposition shows that the optimal solution $(u^{(N)})$ of the control problem corresponds to the (unique) Nash equilibrium of a Stochastic differential game.

The following Proposition gives another way to compute the solution of the individual problems and allows to focus on the fluctuations of the controls and states of individual players around the empirical means over classes.

Proposition 4.3.4. *Let $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the individual problem of agent $i \in [N_k]$ of class $k \in [M]$. Let $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the coordination problem. Then $(u^{(k,i,N)} - \bar{u}^{(k,N)}, X^{(k,i,N)} - \bar{X}^{(k,N)}, Y^{(k,i,N)} - \bar{Y}^{(k,N)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ is the unique solution of the following FBSDE with unknowns $(\Delta u, \Delta X, \Delta Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\forall t \in [0,T],$$
$$\begin{cases} \Delta X_t = x_0^{(k,i)} - \bar{x}_0^{(k,N)} + \int_0^t \left(\alpha_s^{(k)} \Delta u_s + \beta_s^{(k)} \Delta X_s + \gamma_s^{(k,i)} - \bar{\gamma}_s^{(k,N)}\right) \mathrm{d}s, \\ \Delta Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(\Delta X_T - x_T^{\mathtt{f},(k,i)} + \bar{x}_T^{\mathtt{f},(k,N)}\right) + \int_t^T \left(\beta_s^{(k)} \Delta Y_s + \nu_s^{(k)}\left(\Delta X_s - x_s^{\mathtt{ref},(k,i)} + \bar{x}_s^{\mathtt{ref},(k,N)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(\Delta u_t - u_t^{\mathtt{ref},(k,i)} + \bar{u}_t^{\mathtt{ref},(k,N)}\right) + \alpha_t^{(k)} \Delta Y_t = 0. \end{cases} \tag{4.3.5}$$

We have developed a decomposition method of the N-dimensional FBSDE (4.2.4), which shows that it is equivalent to solve one M-dimensional FBSDE, with $M << N$ typically, the coordination problem, and N one-dimensional FBSDE. However the parameters of the FBSDE depend on aggregate data of individual agents. In practical applications, the coordination problem shall be solved by a third party, called coordinator, which may not have access in real time to the aggregate data of individual agents (namely $\bar{\mathrm{P}}^{\mathtt{load},(k,N)}$, $\bar{\gamma}^{(k,N)}$, $\bar{u}^{\mathtt{ref},(k,N)}$, $\bar{x}^{\mathtt{ref},(k,N)}$, $\bar{x}_T^{\mathtt{f},(k,N)}$). The reason for that is two-fold: a privacy concern, as agents may not wish to share their private data (namely $\mathrm{P}^{\mathtt{load},(k,i)}$, $\gamma^{(k,i)}$, $u^{\mathtt{ref},(k,i)}$, $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$), and a technical reason, as heavy telecommunication infrastructures would be required to allow individual agents to share their private data with the coordinator.

4.4 Approximate decentralized control architecture preserving privacy

In this section, we present an approximation of the coordination problem which parameters only depend on the statistical behaviors of the agents (namely the conditional means of their associated parameters), rather than the actual realization of their individual data. In particular, this is a desirable feature in decentralized implementations where the coordination problem is solved by a third party not observing in real-time the aggregate behaviors of agents, for privacy reasons or in order to reduce real-time telecommunication requirements. Solving the approximate coordination problem allows to compute an approximation of the coordination signal (4.3.2), which may be used to decouple individual problems. This is done at the cost of a small performance loss, which can be estimated and which vanishes asymptotically when the number of consumers/agents goes to infinity.

Notations

- *Empirical conditional means:* We introduce the notation $\bar{\mathrm{P}}_t^{\mathtt{load}} := \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{E}\left[\mathrm{P}_t^{\mathtt{load},(l,j)}|\mathcal{G}_t\right]$ for the conditional average of the empirical average consumption.
- *Empirical conditional means over a class:* We introduce the following notations for the conditional means of the empirical average over classes of individual parameters: for $t \in [0,T]$, $\bar{\gamma}_t^{(k)} := \frac{1}{N_k}\sum_{j=1}^{N_k}\mathbb{E}\left[\gamma_t^{(k,j)}|\mathcal{G}_t\right]$, $\bar{u}_t^{\mathtt{ref},(k)} := \frac{1}{N_k}\sum_{j=1}^{N_k}\mathbb{E}\left[u_t^{\mathtt{ref},(k,j)}|\mathcal{G}_t\right]$, $\bar{x}_t^{\mathtt{ref},(k)} := \frac{1}{N_k}\sum_{j=1}^{N_k}\mathbb{E}\left[x_t^{\mathtt{ref},(k,j)}|\mathcal{G}_t\right]$, $\bar{x}_T^{\mathtt{f},(k)} := \frac{1}{N_k}\sum_{j=1}^{N_k}\mathbb{E}\left[x_T^{\mathtt{f},(k,j)}|\mathcal{G}_T\right]$.

Lemma 4.4.1. *The processes $(\bar{\gamma}^{(k)}, \bar{u}^{\mathtt{ref},(k)}, \bar{x}^{\mathtt{ref},(k)}, \bar{x}_T^{\mathtt{f},(k)})_{k\in[M]}$, $\bar{\mathrm{P}}^{\mathtt{load}}$ are $\mathbb{G}$-measurable and can be assumed progressively measurable without loss of generality, and therefore in $\mathcal{H}_{\mathbb{G}}$.*

Proof. Let Z denote any of the processes $(\bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k\in[M]}$, $\bar{\mathrm{P}}^{\mathtt{load},(N)}$. Then $Z \in \mathcal{H}$ is $\mathbb{F}$-progressively measurable and since $\mathbb{G}$ is immersed in $\mathbb{F}$:

$$\forall t \in [0,T], \quad \mathbb{E}\left[Z_t|\mathcal{G}_t\right] = \mathbb{E}\left[Z_t|\mathcal{G}_T\right].$$

As in [CD18, Volume II, pp. 265-266], one can redefine $t \mapsto \mathbb{E}\left[Z_t|\mathcal{G}_T\right]$ as a $\mathcal{B}([0,T]) \otimes \mathbb{G}$-measurable process, up to a $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-null set. By [KS98, Proposition 1.12, p. 5], there exists a $\mathbb{G}$-progressively measurable modification of the above process. □

Additional assumptions We assume that the coordinator does not observe the parameters resulting from the aggregation of individual data $(\bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k\in[M]}$, $\bar{\mathrm{P}}^{\mathtt{load},(N)}$ but can use statistical estimators of these quantities, measurable with respect to the common noise information $\mathbb{G}$. Similarly, the aggregator does not observe the empirical mean of the initial states over classes $(\frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)})_{k\in[M]}$, but can use (statistical) approximations of these parameters, denoted $(\bar{x}_0^{(k)})_{k\in[M]}$. This is realistic and desirable from the point of view of the application, as it allows to consider a non-intrusive and non omniscient coordinator.

(H-lim) Assume **(H-X)**, **(H-J)** and in addition, the following assumptions.
We denote $\pi^{(k)} := \frac{N_k}{N}$, for all $k \in [M]$ and assume $\pi^{(k)} \geq \eta$ for all $k \in [M]$ for some $\eta > 0$. In particular, $M \leq \frac{1}{\eta}$.
The meteorological conditions are represented by a stochastic process $\mathrm{X}^{\mathtt{sun}} \in \mathcal{H}$, assumed $\mathbb{G} = \{\mathcal{G}_t\}_{0\leq t\leq T}$-progressively measurable, where $\mathbb{G}$ satisfies the usual conditions and is immersed in $\mathbb{F}$.
We assume $\left|\bar{x}_0^{(k)} - \frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)}\right| \leq \frac{C}{\sqrt{N_k}} = \frac{C}{\sqrt{\pi_k N}}$, for some constant C independent from N.
$(u^{\mathtt{ref},(k,i)})_{k,i}$ (resp. $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$) are independent conditionally to $\mathcal{G}_T$.
$\mathrm{P}^{\mathtt{load},(k,i)}$, $\gamma^{(k,i)}$, $u^{\mathtt{ref},(k,i)}$ and $x^{\mathtt{ref},(k,i)}$ are uniformly bounded in $\mathcal{H}$ by a constant independent from N.
Similarly, $x_T^{\mathtt{f},(k,i)}$ is bounded in $\mathbb{L}^2$ by a constant independent from N.
The processes $(\mathrm{P}^{\mathtt{load},(k,i)}, \gamma^{(k,i)}, u^{\mathtt{ref},(k,i)}, x^{\mathtt{ref},(k,i)}, x_T^{\mathtt{f},(k,i)})_{k\in[M],i\in[N_k]}$ are independent conditionally to $\mathcal{G}_T$.

Example 4.4.2. *Let us illustrate the assumption of independence of $(\mathrm{P}^{\mathtt{load},(k,i)})_{k,i}$, $(u^{\mathtt{ref},(k,i)})_{k,i}$, $(x^{\mathtt{ref},(k,i)})_{k,i}$ and $(\gamma_{k,i})_{k,i}$ conditionally to $\mathbb{G}$. For this discussion, assume $\mathrm{P}^{\mathtt{load},(k,i)} = \mathrm{P}^{\mathtt{cons},(k,i)} - \mathrm{P}^{\mathtt{sun},(k,i)}$.*

One can consider that the power consumption $\mathrm{P}^{\mathrm{cons}}$ by a household can be explained by the temperature (as it impacts the heating and cooling), the solar irradiance (as it impacts the lighting) and other random individual factors. Our assumption states that the random individual factors are independent for distinct households. It can be interpreted as the consumers having independent behaviors, once the weather is known. Similarly, the exogenous solicitations γ of individual energy storage systems (thermal losses for instance), the target power and state of charge profiles u^{ref} and x^{ref} depend on meteorological conditions and individual factors. Our assumption states that these individual factors are statistically independent.

Last, if the average solar irradiance $\mathrm{P}^{\mathrm{sun}}$ is known in a region, there remains some random local fluctuations of solar irradiance, due to clouds passing by, for instance. Our assumption includes the case where these random local fluctuations (specific to each agent) are independent conditionally to the average solar irradiance on the region.

4.4.1 Convergence of the coordination problem for large populations

When the number of agents becomes large enough, the input parameters of the *coordination problem* converge, and their limits can be easily computed, based on the conditional law of large numbers, without having to observe individual data.

Proposition 4.4.3. *We have the following convergence properties:*

$$\|\bar{\mathrm{P}}^{\mathrm{load},(N)} - \bar{\mathrm{P}}^{\mathrm{load}}\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N}},$$
$$\forall k \in [M], \quad \|\bar{u}^{\mathrm{ref},(k,N)} - \bar{u}^{\mathrm{ref},(k)}\|_{\mathcal{H}} + \|\bar{x}^{\mathrm{ref},(k,N)} - \bar{x}^{\mathrm{ref},(k)}\|_{\mathcal{H}} + \|\bar{x}_T^{\mathrm{f},(k,N)} - \bar{x}_T^{\mathrm{f},(k)}\|_{\mathcal{H}} + \|\bar{\gamma}^{(k,N)} - \bar{\gamma}^{(k)}\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N_k}}.$$

for some constant $C > 0$ independent from N.

The proof of the above Proposition is postponed to Subsection 4.7.3.

Replacing the coefficients of the FBSDE (4.3.1) by their $\mathbb{G}$-measurable approximations leads us to consider the following FBSDE with unknown $(u^{(k)}, X^{(k)}, Y^{(k)})_{k\in[M]} \in (\mathcal{H} \times \mathcal{S} \times \mathcal{S})^M$:

$$\forall k \in [M],$$
$$\begin{cases} X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathrm{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathrm{ref},(k)}\right) + \mathcal{L}'_x\left(t, \sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right) + \alpha_t^{(k)} Y_t^{(k)} = 0. \end{cases} \tag{4.4.1}$$

This FBSDE is called the *limiting coordination problem* and has a unique solution denoted by $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{1\leq k\leq M} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$. Indeed, it is the optimality system of the stochastic control problem:

$$\begin{aligned} \min_{(u^{(k)})\in\mathcal{H}^M} \quad & \mathbb{E}\left[\sum_{k=1}^M \pi^{(k)} \left\{\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)} - \bar{u}_t^{\mathrm{ref},(k)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k)} - \bar{x}_t^{\mathrm{ref},(k)}\right)^2\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right)^2\right\}\right] \\ & + \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right) \mathrm{d}t\right], \\ s.t. \quad & X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \quad \forall k \in [M]. \end{aligned} \tag{4.4.2}$$

This control problem can be interpreted as the problem of the coordinator (which plays the role of major player in the stochastic differential game) aiming to control the aggregate behavior of the population by sending appropriate signals, in the asymptotic regime of large populations. By our assumptions and using [Bre10b, Corollary 3.23, pp.71], the above control can be shown to have a unique solution, which is fully characterized by the FBSDE (4.4.1), by convexity of the stochastic control problem.

Definition 4.4.4. *Let* $\left(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)}\right) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ *be the unique solution of the limiting coordination problem* (4.4.1)*. We define the limiting coordination signal* $\bar{v}^{(\infty)} \in \mathcal{H}$ *by:*

$$\forall t \in [0,T], \quad \bar{v}_t^{(\infty)} := \mathcal{L}_x'\left(t, \sum_{l=1}^{M} \pi^{(l)} \bar{u}_t^{(l,\infty)} + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right). \tag{4.4.3}$$

The interest of considering the *limiting coordination problem* instead of the *coordination problem* is that it can be solved in the filtration $\mathbb{G}$ instead of the general filtration $\mathbb{F}$, allowing to consider a non-intrusive coordinator. This fact is made clear by the following theorem.

Theorem 4.4.5. *The unique solution* $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{k\in[M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ *of the FBSDE* (4.4.1) *is* $\mathbb{G}$*-progressively measurable.*

The proof of the above theorem is postponed to Subsection 4.7.4.

We then show the convergence of the solution of the *coordination problem* to the solution of *limiting coordination problem* when N goes to infinity at speed $1/\sqrt{N}$, as expected. To do it, we prove stability of the solution of FBSDE with similar structure as the one of (4.3.1) and (4.4.1) with respect to parameters of the FBSDE.

Proposition 4.4.6. *Let* $\theta = (x, v, w, u^{\mathrm{ref}}, x^{\mathrm{ref}}, x_T^{\mathrm{f}})$ *with* $x = (x^{(k)})_{k\in[M]} \in \mathbb{R}^M$, $v \in \mathcal{H}$, $w = (w^{(k)})_{k\in[M]} \in \mathcal{H}(\mathbb{R}^M)$, $u^{\mathrm{ref}} = (u^{\mathrm{ref},(k)})_{k\in[M]} \in \mathcal{H}$, $x^{\mathrm{ref}} = (x^{\mathrm{ref},(k)})_{k\in[M]} \in \mathcal{H}$, $x_T^{\mathrm{f}} = (x_T^{\mathrm{f},(k)})_{k\in[M]} \in \mathbb{L}_T^2$.

Consider the FBSDE with unknowns $(u^{(k)}, X^{(k)}, Y^{(k)})_{k\in[M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ *parameterized by :*

$$\begin{cases} X_t^{(k)} = x^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + w_s^{(k)}\right) \mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)} - x_T^{\mathrm{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - x_s^{\mathrm{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - u_t^{\mathrm{ref},(k)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)} u_t^{(l)} + v_t\right) + \alpha_t^{(k)} Y_t^{(k)} = 0. \end{cases}$$

This FBSDE has a unique solution $(\bar{u}^{(k),\theta}, \bar{X}^{(k),\theta}, \bar{Y}^{(k),\theta})_{k\in[M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ *for any* $\theta = (x, v, w, u^{\mathrm{ref}}, x^{\mathrm{ref}}, x_T^{\mathrm{f}}) \in \mathbb{R}^M \times \mathcal{H} \times \mathcal{H}(\mathbb{R}^M) \times \mathcal{H}(\mathbb{R}^M) \times \mathbb{L}_T^2(\mathbb{R}^M)$*. Besides, for any* $\theta^1 = (x^1, v^1, w^1, u^{\mathrm{ref},1}, x^{\mathrm{ref},1}, x_T^{\mathrm{f},1}) \in \mathbb{R}^M \times \mathcal{H} \times \mathcal{H}(\mathbb{R}^M) \times \mathcal{H}(\mathbb{R}^M) \times \mathbb{L}_T^2(\mathbb{R}^M)$ *and* $\theta^2 = (x^2, v^2, w^2, u^{\mathrm{ref},2}, x^{\mathrm{ref},2}, x_T^{\mathrm{f},2}) \in \mathbb{R}^M \times \mathcal{H} \times \mathcal{H}(\mathbb{R}^M) \times \mathcal{H}(\mathbb{R}^M) \times \mathbb{L}_T^2(\mathbb{R}^M)$*, we have, for* T *small enough:*

$$\|(\bar{u}^{\theta^1} - \bar{u}^{\theta^2}, \bar{X}^{\theta^1} - \bar{X}^{\theta^2}, \bar{Y}^{\theta^1} - \bar{Y}^{\theta^2})\|_{\mathcal{H}} \leq C_T \|\theta^1 - \theta^2\|.$$

The proof of the above Proposition is postponed to Subsection 4.7.5.

Remark 4.4.7. *This stability result could be extended to arbitrary time horizon* T*, for instance by an adaptation of the continuation method, presented for instance in [CD18, p. 560, Volume I], or maybe as a consequence of [Ma+15, Theorem 8.1, p. 2203].*

Corollary 4.4.8. *Under our assumptions, we have:*

$$\sum_{k=1}^{M} \|(\bar{u}^{(k,\infty)} - \bar{u}^{(k,N)}, \bar{X}^{(k,\infty)} - \bar{X}^{(k,N)}, \bar{Y}^{(k,\infty)} - \bar{Y}^{(k,N)})\|_{\mathcal{H}} \leq \frac{C_T}{\sqrt{N}}.$$

In particular, we have the following estimation of the error between the coordination signal and the limiting coordination signal:

$$\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}} \leq \frac{C_T}{\sqrt{N}}.$$

The proof of the above Corollary is postponed to Subsection 4.7.6.

4.4.2 Convergence of the individual problems for large populations

Proposition 4.4.9. *For $k \in [M], i \in [N_k]$, consider the following FBSDE with unknown $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$, called the limiting individual problem of agent i of class k:*

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s \\ Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right] \\ \mu_t^{(k)}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + \bar{v}_t^{(\infty)} + \alpha_t^{(k)} Y_t = 0 \end{cases} \tag{4.4.4}$$

This FBSDE has a unique solution $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$.

Proof. We use similar arguments as in the proof of Theorem 4.2.1. The FBSDE (4.4.4) is the optimality system associated to the stochastic control problem:

$$\begin{aligned} \min_{u \in \mathcal{H}} \quad & \mathbb{E}\left[\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t - x_t^{\mathtt{ref},(k,i)}\right)^2 + \bar{v}_t^{(\infty)} u_t\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T - x_T^{\mathtt{f},(k,i)}\right)^2\right], \\ s.t. \quad & X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s. \end{aligned} \tag{4.4.5}$$

By [Bre10b, Corollary 3.23, pp.71], this stochastic control problem has a unique solution by our assumptions (ensuring strong convexity, continuity and coercivity of the mapping minimized in problem (4.4.5)). □

The stochastic control problem (4.4.5) can be interpreted as the individual optimization problem of agent i of class k, responding to the limiting coordination signal sent by the coordinator. Similarly as in Theorem 4.4.5, it could be proved that the solution of the individual problem of a given agent (k, i) is progressively measurable with respect to the completed filtration generated by the processes $x^{\mathtt{ref},(k,i)}$, $u^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$, $\gamma^{(k,i)}$, $\bar{v}^{(\infty)}$. This is of practical interest as it ensures that individual decisions can be taken using only locally available information.

The following Proposition gives another way to compute the solution of the limiting individual problems and allows to focus on the fluctuations of the controls and states of individual players around the empirical means over classes.

Proposition 4.4.10. *Let $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the limiting individual problem (4.4.4) of agent $i \in [N_k]$ of class $k \in [M]$. Let $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{k \in [M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the limiting coordination problem (4.4.1). Then $(u^{(k,i,\infty)} - \bar{u}^{(k,\infty)}, X^{(k,i,\infty)} - \bar{X}^{(k,\infty)}, Y^{(k,i,\infty)} - \bar{Y}^{(k,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ is the unique solution of the following FBSDE with unknowns $(\Delta u, \Delta X, \Delta Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\forall t \in [0, T],$$
$$\begin{cases} \Delta X_t = x_0^{(k,i)} - \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} \Delta u_s + \beta_s^{(k)} \Delta X_s + \gamma_s^{(k,i)} - \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \\ \Delta Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(\Delta X_T - x_T^{\mathtt{f},(k,i)} + \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} \Delta Y_s + \nu_s^{(k)}\left(\Delta X_s - x_s^{\mathtt{ref},(k,i)} + \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(\Delta u_t - u_t^{\mathtt{ref},(k,i)} + \bar{u}_t^{\mathtt{ref},(k)}\right) + \alpha_t^{(k)} \Delta Y_t = 0. \end{cases} \tag{4.4.6}$$

To show the convergence of the solutions of the *individual problems* to the solutions of the *limiting individual problems*, we have to use the following stability result for a class of FBSDE with respect to its input parameters.

Proposition 4.4.11. *Let $v \in \mathcal{H}$. For $k \in [M], i \in [N_k]$, consider the FBSDE with unknowns $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + v_t + \alpha_t^{(k)} Y_t = 0. \end{cases}$$

This FBSDE has a unique solution $(u^{(k,i),v}, X^{(k,i),v}, Y^{(k,i),v}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ for any $v \in \mathcal{H}$. Besides, for T small enough, for any v, v' in $\mathcal{H}$, for all $k \in [M], i \in [N_k]$:

$$\|(u^{(k,i),v} - u^{(k,i),v'}, X^{(k,i),v} - X^{(k,i),v'}, Y^{(k,i),v} - Y^{(k,i),v'})\|_{\mathcal{H}} \leq C_T \|v - v'\|_{\mathcal{H}}.$$

The proof of the above Proposition is similar to the proof of Proposition 4.4.6.

Theorem 4.4.12. *The solutions of the limiting individual problems are close to the solutions of the individual problems. In other words, for $k \in [M], i \in [N_k]$, for T small enough, for some constant C independent from N,*

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N}},$$

and we have the following estimation on the sub-optimality of $u^{(\infty)} := (u^{(k,i,\infty)})_{k\in[M],i\in[N_k]}$ compared to the optimal solution $u^{(N)} := (u^{(k,i,N)})_{k\in[M],i\in[N_k]}$, for another constant C independent from N:

$$0 \leq \mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) = \mathcal{J}(u^{(\infty)}) - \min_{v\in\mathcal{H}} \mathcal{J}(v) \leq \frac{C}{N}.$$

The proof of the above Theorem is postponed to Subsection 4.7.7.

The above result shows that the (unique) Nash equilibrium a Stochastic Stackelberg Differential Game with followers (the agents) and a leader (the aggregator) corresponds to a quasi-optimal solution of the centralized control problem of many TCLs. In particular, no real-time communication from the agents to the aggregator is required, and the problems of the agents and the coordinator can be solved using locally available data only.

4.5 Numerical experiments

4.5.1 The model

Models for exogenous random uncertainties

In the following, $M = 2$, $W = (W^{(k,i)})_{k\in[2],i\in[N_k]}$, $\tilde{W}$ and $N^{\mathrm{cons}} = (N^{\mathrm{cons},(k,i)}_{k,i})_{k\in[2],i\in[N_k]}$ denote respectively a N-dimensional Brownian motion, a one-dimensional Brownian motion and a N independent compensated Poisson processes with intensity λ^{cons}. These processes are assumed independent.

We assume $\mathrm{P}^{\mathrm{load},(k,i)} = \mathrm{P}^{\mathrm{cons},(k,i)} - \mathrm{P}^{\mathrm{sun}}$ for all $k \in [2], i \in N_k$. For the consumption process $\mathrm{P}^{\mathrm{cons},(k,i)}$, we assume the following dynamic:

$$\mathrm{d}\mathrm{P}^{\mathrm{cons},(k,i)}_t = -\rho^{\mathrm{cons}}(\mathrm{P}^{\mathrm{cons},(k,i)}_t - \mathrm{p}^{\mathrm{cons,ref}}_t)\mathrm{d}t + \sigma^{\mathrm{cons}}_t \mathrm{d}W^{(k,i)}_t + h^{\mathrm{cons}}\mathrm{d}N^{\mathrm{cons},(k,i)}_t. \tag{4.5.1}$$

In practice, only conditional independence of the processes $(\mathrm{P}^{\mathrm{cons},(k,i)})$ is necessary for our results to hold, but for simplicity of our presentation, we assume that they are (unconditionally) independent identically distributed.

Regarding the PV production, we follow [Bad+18] by setting $\mathrm{P}^{\mathrm{sun}} = \mathrm{P}^{\mathrm{sun,max}}\mathrm{X}^{\mathrm{sun}}$ where $\mathrm{P}^{\mathrm{sun,max}} : [0,T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) and $\mathrm{X}^{\mathrm{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathrm{X}^{\mathrm{sun}}_t = -\rho^{\mathrm{sun}}(\mathrm{X}^{\mathrm{sun}}_t - \mathrm{x}^{\mathrm{sun,ref}}_t)\mathrm{d}t + \sigma^{\mathrm{sun}}(\mathrm{X}^{\mathrm{sun}}_t)^{k_1}(1 - \mathrm{X}^{\mathrm{sun}}_t)^{k_2}\mathrm{d}\tilde{W}_t, \tag{4.5.2}$$

with $k_1, k_2 \geq 1/2$. As proved in [Bad+18], there is a strong solution to the above SDE and the solution $\mathrm{X}^{\mathrm{sun}}$ takes values in $[0,1]$.

Since the drifts are affine-linear, the conditional expectations of the solutions are known in closed form (this property is intensively used in [BSS05]):

$$\mathbb{E}_t\left[\mathrm{P}^{\mathrm{sun}}_s\right] = \left(\frac{\mathrm{P}^{\mathrm{sun}}_t}{\mathrm{P}^{\mathrm{sun,max}}_t}\exp(-\rho^{\mathrm{sun}}(s-t)) + \int_t^s \rho^{\mathrm{sun}}\mathrm{x}^{\mathrm{sun,ref}}_\tau \exp(-\rho^{\mathrm{sun}}(s-\tau))\mathrm{d}\tau\right)\mathrm{P}^{\mathrm{sun,max}}_s, \tag{4.5.3}$$

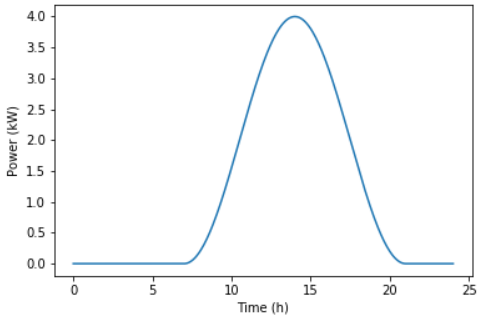
$$\mathbb{E}_t\left[\mathrm{P}^{\mathrm{cons},(k,i)}_s\right] = \mathrm{P}^{\mathrm{cons},(k,i)}_t \exp(-\rho^{\mathrm{cons}}(s-t)) + \left(\int_t^s \rho^{\mathrm{cons}}\mathrm{p}^{\mathrm{cons,ref}}_\tau \exp(-\rho^{\mathrm{cons}}(s-\tau))\mathrm{d}\tau\right), \tag{4.5.4}$$

for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\bar{\mathrm{P}}^{\mathrm{load}}_s\right]$ and $\mathbb{E}_t\left[\bar{\mathrm{P}}^{\mathrm{load},(N)}_s\right]$ as required when deriving the optimal control. Moreover, throughout our application, we assume $\mathrm{P}^{\mathrm{prod}}_t = \mathbb{E}\left[\bar{\mathrm{P}}^{\mathrm{load},(N)}_t\right]$, which can be easily computed using the previous remark.

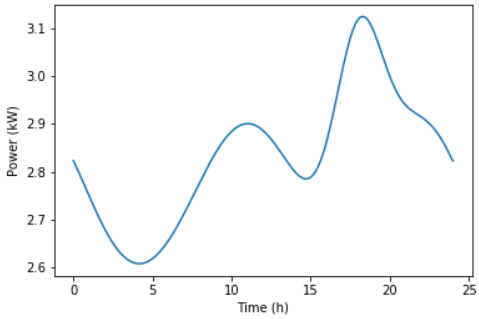
The values of the parameters used are given in the following table, while $\mathrm{P}^{\mathrm{sun,max}}$ and $\mathrm{p}^{\mathrm{cons,ref}}$ are plotted in Figures 4.2a and 4.2b. Empirical quantile plot (obtained by simulation of 40000 i.i.d. trajectories) as well as one example trajectory of $\mathrm{P}^{\mathrm{sun}}$ and $\mathrm{P}^{\mathrm{cons}}$ are given in Figures 4.2c and 4.2d. The simulations are performed using Euler scheme on a time horizon $T = 24$ h with step length $1/16$ h.

Table 4.1: Parameter values for the simulation of input processes

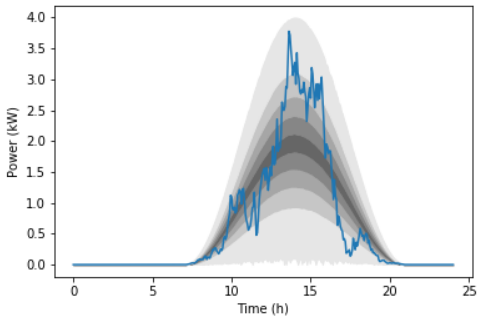
ρ^{cons}	σ^{cons}	λ^{cons}	h^{cons}	ρ^{sun}	$\mathbf{x}^{\text{sun,ref}}$	σ^{sun}	k_1	k_2
0.9 h^{-1}	$0.1(\mathrm{p}^{\text{cons,ref}} - 2.5) + 0.1$	0.25 h^{-1}	0.25 kW	0.75 h^{-1}	0.5	0.8	0.8	0.7



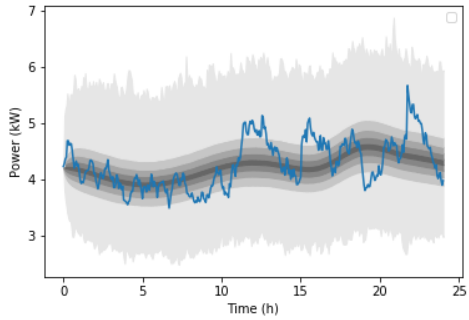
(a) Time evolution of $\mathrm{P}^{\text{sun,max}}$



(b) Time evolution of $\mathrm{p}^{\text{cons,ref}}$



(c) Quantile plot and one example trajectory of P^{sun}



(d) Quantile plot and one example trajectory of P^{cons}

Figure 4.2: Parameters and statistical evolution of P^{sun} and P^{cons}

Models for Thermostatically Controlled Loads

We consider two classes (i.e., $M = 2$) of Thermostatically Controlled Loads: heat pumps and water heaters. We control the power consumption of these devices around a nominal value defined such that thermal losses are exactly compensated when their target temperature $x_T^{\text{f},(k,i)} = x^{\text{ref},(k,i)}$, assumed constant, is reached. We assume first order models for the temperature, as in [Tro+16]. The dynamics of the temperature associated to individual devices is given by:

$$\frac{\mathrm{d}X_t^{(k,i)}}{\mathrm{d}t} = \frac{COP^{(k)}}{C^{(k)}} u_t^{(k,i)} - \frac{1}{R^{(k)}C^{(k)}}(X_t^{(k,i)} - x^{\text{ref},(k,i)}),$$

where $COP^{(k)}$ denotes the coefficient of performance of devices of type k, $C^{(k)}$ denotes its thermal capacitance, $R^{(k)}$ its thermal resistance, $X^{out,(k,i)}$ the environment temperature of the device. This gives $\alpha^{(k)} = \frac{COP^{(k)}}{C^{(k)}}$, $\beta^{(k)} = -\frac{1}{R^{(k)}C^{(k)}}$ and $\gamma^{(k,i)} = \frac{x^{\text{ref},(k,i)}}{R^{(k)}C^{(k)}}$. Realistic parameter values for various types of TCLs (AC, refrigerators, water heaters, heat pumps) can be found in [Mat+12]. We use these values to set the parameters of our models of devices.

The temperature deadband $\delta^{(k)}$ of a TCLs denotes admissible deviation of the temperature $X^{(k,i)}$ of a device from its reference $x^{\text{ref},(k,i)}$. We assume $(x_0^{(k,i)})_{k\in[2],i\in[N_k]} = (x^{\text{ref},(k,i)})_{k\in[2],i\in[N_k]} = (x_T^{\text{f},(k,i)})_{k\in[2],i\in[N_k]}$ are independent and $x^{\text{ref},(k,i)} = x_0^{(k,i)} = x_T^{\text{f},(k,i)}$ is drawn from a uniform distribution on the interval $[\bar{x}^{\text{ref},(k)} - \delta^{(k)}, \bar{x}^{\text{ref},(k)} + \delta^{(k)}]$.

Table 4.2: Parameter values for the Thermostatically Control Loads

Class index k	type of device	$R^{(k)}$	$C^{(k)}$	$COP^{(k)}$	$\bar{x}^{\mathtt{ref},(k)}$	Deadband $\delta^{(k)}$
1	Water heater	120 (℃/kW)	0.4 (kWh/ ℃)	1	48 (℃)	4 (℃)
2	Heat pump	2 (℃/kW)	2 (kWh/℃)	3.5	20 (℃)	1 (℃)

Cost parameters

In the case where $\mathcal{L}$ is not quadratic, the coordination problem is a low-dimensional control problem which does not have explicit solutions in general, so that a numerical method is required to solve it. We assume in this chapter that $\mathcal{L}$ is simply the quadratic function $\mathcal{L} : (t, x) \mapsto \frac{\lambda}{2}x^2$ for some deterministic constant $\lambda \geq 0$. In this case, all the control problems are Linear-quadratic, the associated FBSDEs are affine-linear, and therefore have quasi-explicit solutions. Individual cost parameters are scaled according to the square of the temperature deadband of the type of device considered, in order to guarantee proper scaling of the temperature range sizes.

Table 4.3: Parameter values for the cost functional

$\mu^{(k)}(\delta^{(k)})^2$	$\nu^{(k)}(\delta^{(k)})^2$	$\rho^{(k)}(\delta^{(k)})^2$	λ
4 (℃)^2 kW^{-2} h^{-1}	5 h^{-1}	0	30 kW^{-2} h^{-1}

4.5.2 Solution method of affine-linear FBSDEs

The solution method for linear FBSDEs is standard, see for instance [Bis76; Yon99; Yon06].

The limiting coordination problem

The optimality system of the limiting coordination problem is a linear FBSDE which writes, up to rescaling of the optimality condition by $\boldsymbol{\pi}$:

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\alpha} u_t + \boldsymbol{\beta} X_t + \bar{\gamma}, \\ Y_t = \mathbb{E}_t\Big[\boldsymbol{\rho}(X_T - \bar{x}^{\mathtt{ref}}) + \int_t^T \Big(\boldsymbol{\nu}(X_s - \bar{x}^{\mathtt{ref}}) + \boldsymbol{\beta} Y_s\Big)\,\mathrm{d}s\Big], \\ \boldsymbol{\mu\pi} u_t + \lambda\pi(\pi^\top u_t + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}) + \boldsymbol{\alpha\pi} Y_t = 0, \end{cases} \tag{4.5.5}$$

where:

$$\boldsymbol{\alpha} := diag(\alpha^{(k)})_{k=1,2}, \quad \boldsymbol{\beta} := diag(\beta^{(k)})_{k=1,2}, \quad \boldsymbol{\mu} := diag(\mu^{(k)})_{k=1,2}, \quad \boldsymbol{\nu} := diag(\nu^{(k)})_{k=1,2}, \quad \boldsymbol{\rho} := diag(\rho^{(k)})_{k=1,2},$$
$$\pi := \begin{pmatrix} \pi^{(1)} \\ \pi^{(2)} \end{pmatrix}, \quad \boldsymbol{\pi} := diag(\pi), \quad \bar{\gamma} = (\bar{\gamma}^{(k)})_{k=1,2}, \quad \bar{u} = (\bar{u}^{(k)})_{k=1,2}, \quad \bar{X} = (\bar{X}^{(k)})_{k=1,2}, \quad \bar{Y} = (\bar{Y}^{(k)})_{k=1,2}.$$

The last equation allows to eliminate u, which writes:

$$u_t = -(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\Big(\boldsymbol{\alpha\pi} Y_t + \lambda(\bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}})\pi\Big).$$

Therefore, we obtain the following FBSDE with unknowns (X, Y):

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\beta} X_t - \boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha\pi} Y_t + \bar{\gamma} + \lambda\boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathtt{prod}} - \bar{\mathrm{P}}_t^{\mathtt{load}})\pi, \\ Y_t = \mathbb{E}_t\Big[\boldsymbol{\rho}(X_T - \bar{x}^{\mathtt{ref}}) + \int_t^T \Big(\boldsymbol{\nu}(X_s - \bar{x}^{\mathtt{ref}}) + \boldsymbol{\beta} Y_s\Big)\,\mathrm{d}s\Big]. \end{cases} \tag{4.5.6}$$

Introduce the Matrix-valued Riccati ODE with unknown ϕ:

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + \phi_t\boldsymbol{\beta} + \boldsymbol{\beta}\phi_t + \boldsymbol{\nu\pi} - \phi_t\boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}\phi_t = 0, \\ \phi_T = \boldsymbol{\rho\pi}. \end{cases}$$

According to [Bis76, Theorem 6.1], this Riccati ODE has a unique bounded solution $\bar{\boldsymbol{\phi}}$, since $\boldsymbol{\rho\pi}$ and $\boldsymbol{\mu\pi} + \lambda\pi\pi^\top$ are positive semi-definite matrices (the second matrix is even positive definite).

Introduce the linear BSDE in $(\psi, M) \in \mathcal{S} \times \mathcal{S}$ with M martingale vanishing at $t = 0$:

$$\begin{cases} -\mathrm{d}\psi_t = \left\{\left(\boldsymbol{\beta} - \bar{\boldsymbol{\phi}}_t\boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}\right)\psi_t - \boldsymbol{\nu\pi}\bar{x}^{\mathtt{ref}} + \bar{\boldsymbol{\phi}}_t\bar{\gamma} + \lambda\bar{\boldsymbol{\phi}}_t\boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathtt{prod}} - \bar{\mathrm{P}}_t^{\mathtt{load}})\pi\right\}\mathrm{d}t - \mathrm{d}M_t, \\ \psi_T = -\boldsymbol{\rho\pi}\bar{x}^{\mathtt{ref}}. \end{cases}$$

This linear BSDE admits a unique solution $(\bar{\psi}, \bar{M}) \in \mathcal{S} \times \mathcal{S}$ with $\bar{M}$ martingale vanishing at $t = 0$, according to [EPQ97, Theorem 5.1]. Let $\bar{X}$ be the unique solution of the affine-linear ODE:

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\beta}X_t - \boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_tX_t + \bar{\psi}_t) + \bar{\gamma} + \lambda\boldsymbol{\alpha}(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathtt{prod}} - \bar{\mathrm{P}}_t^{\mathtt{load}})\pi. \end{cases}$$

Then, using Integration by Parts Formula in [Pro03, Corollary 2, p. 68], one can show that $\bar{X}$ and $\bar{Y} = \boldsymbol{\pi}^{-1}\left(\bar{\boldsymbol{\phi}}\bar{X} + \bar{\psi}\right)$ are solutions of the FBSDE (4.5.6). Therefore, (4.5.5) has a unique solution $(\bar{u}, \bar{X}, \bar{Y})$ with the following feedback expression for $\bar{u}$:

$$\bar{u}_t = -(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_t\bar{X}_t + \bar{\psi}_t) + \lambda(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\pi(\mathrm{P}_t^{\mathtt{prod}} - \bar{\mathrm{P}}_t^{\mathtt{load}}).$$

Besides, the limiting coordination signal is given by the feedback expression:

$$\begin{aligned} \bar{v}_t^{(\infty)} &= \lambda\left(\pi^\top\bar{u}_t + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}\right) \\ &= \lambda\left(-\pi^\top(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_t\bar{X}_t + \bar{\psi}_t) + \lambda\pi^\top(\boldsymbol{\mu\pi} + \lambda\pi\pi^\top)^{-1}\pi(\mathrm{P}_t^{\mathtt{prod}} - \bar{\mathrm{P}}_t^{\mathtt{load}}) + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}\right). \end{aligned}$$

The coordination problem

The coordination problem has the same structure as the limiting coordination problem with coefficients $\bar{x}^{\mathtt{ref}}$, $\bar{x}_0$, $\bar{\gamma}$ and $\bar{\mathrm{P}}^{\mathtt{load}}$ formally replaced respectively by $\bar{x}^{\mathtt{ref},(N)}$, $\bar{x}_0^{(N)}$, $\bar{\gamma}^{(N)}$ and $\bar{\mathrm{P}}^{\mathtt{load},(N)}$. Solving the coordination problem can be done similarly as for the limiting coordination problem, by formally replacing parameters $\bar{x}^{\mathtt{ref}}$, $\bar{x}_0$, $\bar{\gamma}$ and $\bar{\mathrm{P}}^{\mathtt{load}}$ by $\bar{x}^{\mathtt{ref},(N)}$, $\bar{x}_0^{(N)}$, $\bar{\gamma}^{(N)}$ and $\bar{\mathrm{P}}^{\mathtt{load},(N)}$.

The individual and limiting individual problems

Once the solution of the coordination problem is computed, solving the individual and limiting individual problems amounts to solve (4.3.5) and (4.4.6). By eliminating the (recentered) control using the last equation in both these FBSDEs, one can show that the recentered individual problems and recentered limiting individual problems are both equivalent to one-dimensional FBSDEs with the following structure:

$$\begin{cases} X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + f + \int_t^T (CX_s + AY_s + b_s)\mathrm{d}s\right], \end{cases} \tag{4.5.7}$$

with A, B, C, Γ deterministic, $B < 0$, $C, \Gamma \geq 0$, a, b in $\mathcal{H}$, f in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega)$. The structure of the online limiting individual problems is similar.

Lemma 4.5.1 (1-dimension Riccati ODE with constant coefficients)**.** *Consider the following Riccati ODE:*

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + a\phi_t + b\phi_t^2 + c = 0, \\ \phi_T = \gamma, \end{cases} \tag{4.5.8}$$

with a, b, c, γ deterministic, $b < 0$, $c, \gamma \geq 0$. Then this equation admits a unique bounded solution on $[0, T]$, denoted by ϕ. Define θ as the unique solution of the second-order linear ODE:

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} + a\frac{\mathrm{d}\theta_t}{\mathrm{d}t} + bc\theta_t = 0, \\ \frac{\mathrm{d}\theta_T}{\mathrm{d}t} = \gamma b, \\ \theta_T = 1. \end{cases} \tag{4.5.9}$$

Then θ is positive and the unique solution ϕ of (4.5.8) on $[0, T]$ is given by:

$$\forall t \in [0, T], \quad \phi_t = \left(\frac{\mathrm{d}\theta_t}{\mathrm{d}t}\right)\frac{1}{b\theta_t}$$

Proof. As $x \mapsto -ax - bx^2 - c$ is locally Lipschitz-continuous, (4.5.8) has a unique solution ϕ on some maximal interval (t_0, t_1) with $t_0 < T < t_1$ and $t_0 \in \bar{\mathbb{R}}$ and $t_1 \in \bar{\mathbb{R}}$ are unique. Consider the following Riccati ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} + ap_t + bp_t^2 = 0, \\ p_T = 0. \end{cases} \tag{4.5.10}$$

The null function is the unique solution of 4.5.10 on $(-\infty, +\infty)$. Consider as well the following linear ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} + ap_t + c = 0, \\ p_T = \gamma. \end{cases}$$

It admits a unique solution $\overline{\phi}$ on $(-\infty, +\infty)$. Besides, by comparison theorem for Ordinary differential equations, we have :

$$\forall t \in (t_0, t_1), \quad 0 \leq \phi_t \leq \overline{\phi}_t$$

which shows that ϕ can not explode in finite time, and hence $t_0 = -\infty$ and $t_1 = +\infty$. Hence ϕ is well-defined and bounded on $[0, T]$. Now, let us define θ as the unique solution of the following ODE:

$$\begin{cases} \frac{\mathrm{d}\theta_t}{\mathrm{d}t} = b\phi_t\theta_t, \\ \theta_T = 1. \end{cases} \tag{4.5.11}$$

Then we immediately get that θ is well-defined, and positive on $\mathbb{R}$. Besides, θ is C^2 and:

$$\begin{aligned} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} &= b\frac{\mathrm{d}\phi_t}{\mathrm{d}t}\theta_t + b\phi_t\frac{\mathrm{d}\theta_t}{\mathrm{d}t} \\ &= b(-a\phi_t - b\phi_t^2 - c)\theta_t + b^2\phi_t^2\theta_t \\ &= -ab\phi_t\theta_t - bc\theta_t \\ &= -a\frac{\mathrm{d}\theta_t}{\mathrm{d}t} - bc\theta_t, \end{aligned}$$

where we used successively that ϕ solves (4.5.8) and that θ solves (4.5.11). In particular, this shows that θ is also the unique solution of (4.5.9). This completes the proof. □

Theorem 4.5.2 (Verification theorem for affine-linear FBSDE with constant coefficients)**.** *Let A, B, C, Γ be deterministic constants, $B < 0$, $C, \Gamma \geq 0$, a, b in $\mathcal{H}$, f in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega)$. Let ϕ be the unique solution of the following Riccati ODE:*

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + 2A\phi_t + B\phi_t^2 + C = 0, \\ \phi_T = \Gamma, \end{cases} \tag{4.5.12}$$

and let $(\psi, M) \in \mathcal{S} \times \mathcal{M}_0^2$ be the unique solution of the following BSDE:

$$\begin{cases} -\mathrm{d}\psi_t = \big((B\phi_t + A)\psi_t + \phi_t a_t + b_t\big)\mathrm{d}t - \mathrm{d}M_t, \\ \psi_T = f, \end{cases} \tag{4.5.13}$$

where $\mathcal{M}_0^2$ denotes the space of martingales in $\mathcal{S}$ vanishing at $t = 0$. Denoting θ the unique (non-negative) solution of:

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} + 2A\frac{\mathrm{d}\theta_t}{\mathrm{d}t} + BC\theta_t = 0, \\ \frac{\mathrm{d}\theta_T}{\mathrm{d}t} = \Gamma B, \\ \theta_T = 1, \end{cases} \tag{4.5.14}$$

we have the explicit formula for ψ:

$$\psi_t = \mathbb{E}_t[f]\left(\frac{\theta_T}{\theta_t}\right)\exp(A(T-t)) + \mathbb{E}_t\left[\int_t^T (a_s\phi_s + b_s)\left(\frac{\theta_s}{\theta_t}\right)\exp(A(s-t))\mathrm{d}s\right].$$

If θ is the unique solution of (4.5.14), define X by:

$$X_t = x_0\frac{\theta_t}{\theta_0}\exp(At) + \int_0^t (B\psi_s + a_s)\frac{\theta_t}{\theta_s}\exp(A(t-s))\mathrm{d}s. \tag{4.5.15}$$

Define also $Y := \phi X + \psi$. Then $(X, Y) \in \mathcal{S} \times \mathcal{S}$ is a solution of the following FBSDE:

$$\begin{cases} X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + f + \int_t^T (CX_s + AY_s + b_s)\mathrm{d}s\right]. \end{cases} \tag{4.5.16}$$

Proof. By Lemma 4.5.1, the Riccati ordinary differential equation (4.5.12) has a unique solution ϕ. The uniqueness of the solution $(\psi, M) \in \mathcal{S} \times \mathcal{M}_0$ of (4.5.13) arises from an application of [EPQ97, Theorem 5.1, p. 54]. To obtain the explicit expression of ψ, we use the Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to the product $\tilde{\psi}$ defined by $\tilde{\psi}_t := \psi_t \exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)$ between 0 and T, using the fact that the second term is bounded, continuous with finite variations. This shows:

$$\begin{cases} -\mathrm{d}\tilde{\psi}_t = (\phi_t a_t + b_t)\exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)\mathrm{d}t - \exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)\mathrm{d}M_t, \\ \tilde{\psi}_T = f. \end{cases}$$

In particular, using the boundedness of ϕ, the last term in the above BSDE is a true martingale, so that:

$$\tilde{\psi}_t = \mathbb{E}_t\left[f + \int_t^T (\phi_s a_s + b_s)\exp\left(-\int_s^T (B\phi_r + A)\mathrm{d}r\right)\mathrm{d}s\right].$$

We obtain the explicit expression of ψ by using $\psi_t = \tilde{\psi}_t \exp\left(\int_t^T (B\phi_s + A)\mathrm{d}s\right)$ and using the fact that $\phi_t = \frac{\mathrm{d}\theta_t}{\mathrm{d}t}\frac{1}{B\theta_t}$, which yields $\exp\left(\int_t^s (B\phi_r + A)\mathrm{d}r\right) = \frac{\theta_s}{\theta_t}\exp(A(s-t))$, since θ is positive. Let $\hat{X}$ be given by:

$$\hat{X}_t = x_0 + \int_0^t ((A + B\phi_s)\hat{X}_s + B\psi_s + a_s)\mathrm{d}s. \tag{4.5.17}$$

We want to show that $\hat{X} = X$ given in (4.5.15). Define $\tilde{X}$ by $\tilde{X}_t = \hat{X}_t \exp\left(-\int_0^t (\phi_s B + A)\mathrm{d}s\right)$. Then, by integration by part, we obtain $\tilde{X}_t = x_0 + \int_0^t (B\psi_s + a_s)\exp\left(-\int_0^s (\phi_r B + A)\mathrm{d}r\right)\mathrm{d}s$. We can finally show that $\hat{X} = X$ using $\hat{X}_t = \tilde{X}_t \exp\left(\int_0^t (\phi_s B + A)\mathrm{d}s\right)$ and $\exp\left(\int_s^t (B\phi_r + A)\mathrm{d}r\right) = \frac{\theta_t}{\theta_s}\exp(A(t-s))$ for any t and s in $[0, T]$, using the form $\phi = \frac{\dot{\theta}}{B\theta}$ given by Lemma 4.5.1 (where $\dot{\theta}$ is the derivative of $t \mapsto \theta_t$). Then using the definition of $Y := \phi X + \psi$ and by (4.5.17), we get that X satisfies:

$$X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s.$$

One can verify using $Y = \phi X + \psi$ and an integration by parts that $(Y, M) \in \mathcal{S} \times \mathcal{M}_0^2$ is solution of the following BSDE:

$$\begin{cases} -\mathrm{d}Y_t = (CX_t + AY_t + b_t)\mathrm{d}t - \mathrm{d}M_t, \\ Y_T = \Gamma X_T + f, \end{cases}$$

which yields the result. □

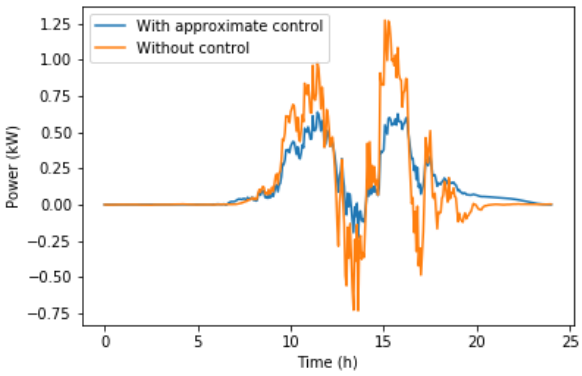
In the simulations, we rely heavily on Theorem 4.5.2 to solve the one-dimensional FBSDEs.

4.5.3 Numerical simulations and results

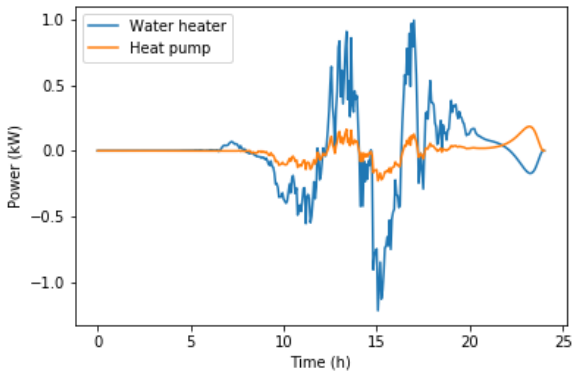
The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. We simulate one realization of the stochastic process $\mathrm{P}^{\mathrm{sun}}$ and N i.i.d scenarios of $\mathrm{P}^{\mathrm{cons}}$ (N being the number of agents) on $[0, T]$ with $T = 24$ hours using Euler schemes with step length $1/16$ h. The solutions of ordinary differential equation and linear backward stochastic differential equations are computed using Euler scheme. For the solution of linear BSDE, we rely heavily on the assumption of affine-linear processes given in Section 4.5.1.

Results with identical population sizes

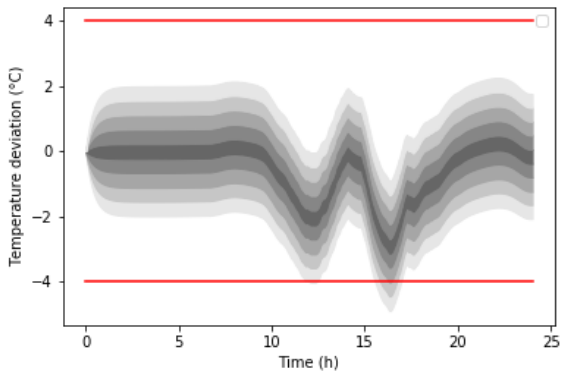
We consider $N = 40000$ users, with $N_1 = N_2 = 20000$ users in each class (which yields the relative population sizes $\pi_1 = \pi_2 = 0.5$). The results of the simulation for one weather scenario are given in Figure 4.3. In particular, the first graph 4.3a shows that the power imbalance using the approximate control (obtained by solving the limiting coordination and limiting individual problems) is closer to 0 than without control. This is done without violating temperature bounds for the populations of water heaters and heat pumps (at least not often and with low probability), see Figures 4.5c and 4.5d. This shows the interest of our approach: power imbalance may be reduced by distributed TCLs while guaranteeing good quality of service, i.e., while maintaining the temperatures of the devices in their admissible ranges.



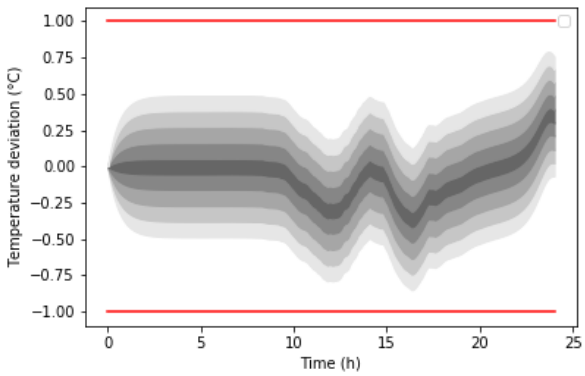
(a) Time evolution of power imbalance



(b) Time evolution of $\bar{u}^{(k)}$ for both types of devices



(c) Quantile plot of the temperature deviations of individual water heaters $\left(X^{(1,i,\infty)} - x^{\mathrm{ref},(1,i)}\right)_{1\le i\le N_1}$

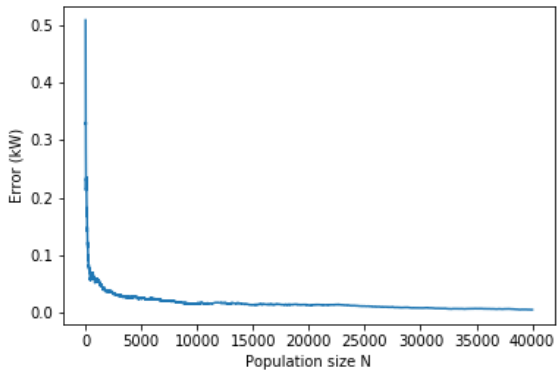


(d) Quantile plot of the temperature deviations of individual heat pumps $\left(X^{(2,i,\infty)} - x^{\mathrm{ref},(2,i)}\right)_{1\le i\le N_2}$

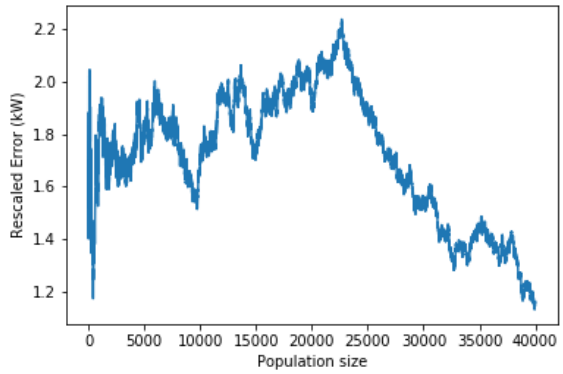
Figure 4.3: Evolution of the system (1 scenario of solar irradiance, quantiles computed within each population class)

Numerical illustration of the convergence of the coordination signal to the limiting coordination signal

We plot the error between the real and limiting coordination signals $\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of the population size N and conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$ in Figure 4.4a and the rescaled error $\sqrt{N}\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of the population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$) 4.4b, to empirically illustrate the convergence of $\bar{v}^{(N)}$ to $\bar{v}^{(\infty)}$ at speed $O\left(\frac{1}{\sqrt{N}}\right)$ when the population size N goes to infinity. .



(a) Error $\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$)



(b) Rescaled error $\sqrt{N}\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$)

Figure 4.4: Convergence of the coordination signal to the limiting coordination signal in the limit of large populations

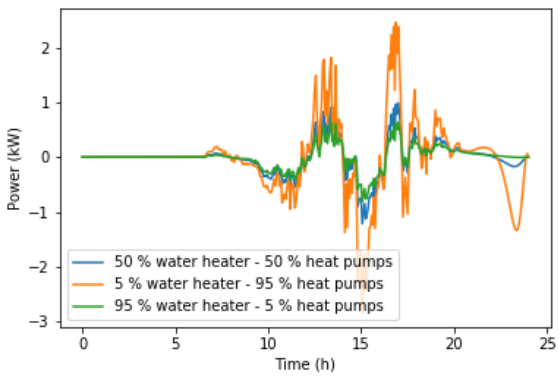
Impact of relative population sizes

We consider the relative sizes of the population given in Table 4.4 without modifying other parameters of the problem.

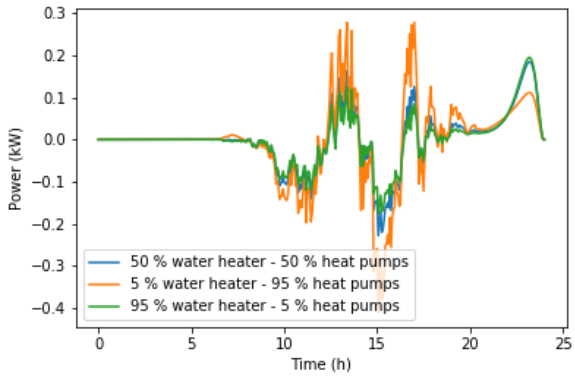
Table 4.4: Relative sizes of populations of water heaters and heat pumps

Scenario	Proportion of water heaters π_1	Proportion of heat pumps π_2
Case 1	50 %	50 %
Case 2	95 %	5 %
Case 3	5 %	95 %

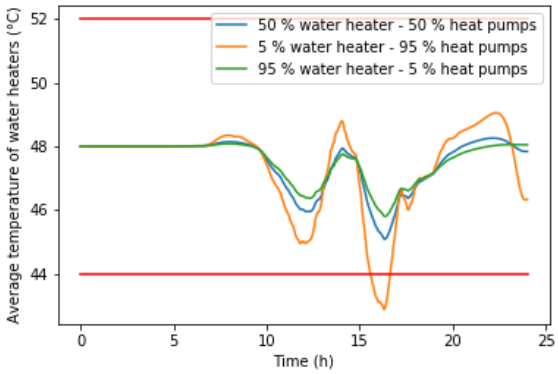
The evolution of the controlled power imbalance in the three cases is represented in Figure 4.5e. We observe a similar magnitude of this signal in the three cases considered. However, whenever the relative size of water heaters population is small (5%), the magnitudes of average controls of both types of population increase (see Figures 4.5a and 4.5b). Besides, the temperature of both types of TCLs vary more, and in the case of water heaters, the temperature may even go outside of the range defined by the deadband temperature (see Figures 4.5c and 4.5d). This may be explained intuitively. Water heaters have more capabilities to provide power without violating their operational constraints than heat pump do. As a result, water heater provide more power than heat pumps (see Figures 4.5a and 4.5b). Hence, when the relative population size of water heaters is reduced, the overall system has a smaller capability of providing and absorbing power. As a result, individual devices of both types are more solicited. These experiments show that there is a trade-off between ensuring individual constraints (power levels, temperature) and global power balance. Appropriate tuning of the parameters of the cost function is therefore required.



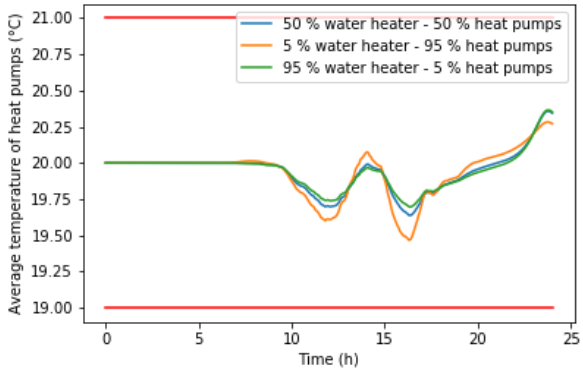
(a) Mean control of water heaters $\bar{u}^{(1)}$



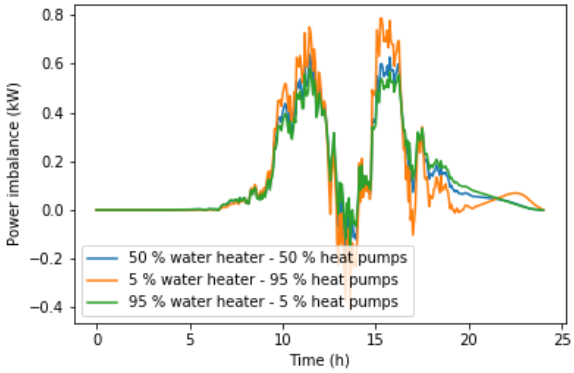
(b) Mean control of heat pumps $\bar{u}^{(2)}$



(c) Mean temperature of water heaters $\bar{X}^{(1)}$



(d) Mean temperature of heat pumps $\bar{X}^{(2)}$



(e) Power imbalance

Figure 4.5: Impact of relative sizes of populations of TCLs for 1 scenario of $\mathrm{P}^{\mathrm{sun}}$ (means computed over the population)

4.6 Online decentralized control scheme with minimal telecommunication

We go back to the setting of Section 4.4, with time dependent coefficients and with a non-quadratic loss function $\mathcal{L}$. We have developed a decomposition and a mean-field approximation which allow to solve approximately the stochastic control with limited telecommunications. In the control architecture developed, a coordinator solves the so-called limiting coordination problem (4.4.1) which allows him to compute the limiting coordination signal $\bar{v}^{(\infty)} \in \mathcal{H}$.

This signal is sent to all agents and used as input parameter of the limiting individual problems 4.4.4. In particular, the (conditional) distribution of the limiting coordination signal $\bar{v}^{(\infty)}$ is required a priori by the agents to solve their individual problems. This raises two issues: sending the distribution of the limiting coordination signal $\bar{v}^{(\infty)}$ is costly and conceptually complex from a telecommunication point of view, and this information needs to be stored locally by agents, which requires heavy memory needs. To tackle these issues, we show that the affine-linear structure of the limiting individual problems allows to compute the current value of the control using a simpler coordination signal. This results in the same control as one would obtain by solving the limiting coordination and individual problems, without additional error. We describe an online control scheme where, at each instant, a coordinator sends the current best estimation (conditional expectation) of the limiting coordination signal on the remaining time horizon. Replacing the distribution of the limiting coordination signal by its current best estimation in the limiting individual problems allows to compute the current value of the control variables using a diagonal scheme. This allows a decentralized architecture with one-way real-time communication of a simple and agent-independent coordination signal from the coordinator to individual consumers, see Figure 4.6[4]. Let us give more details.

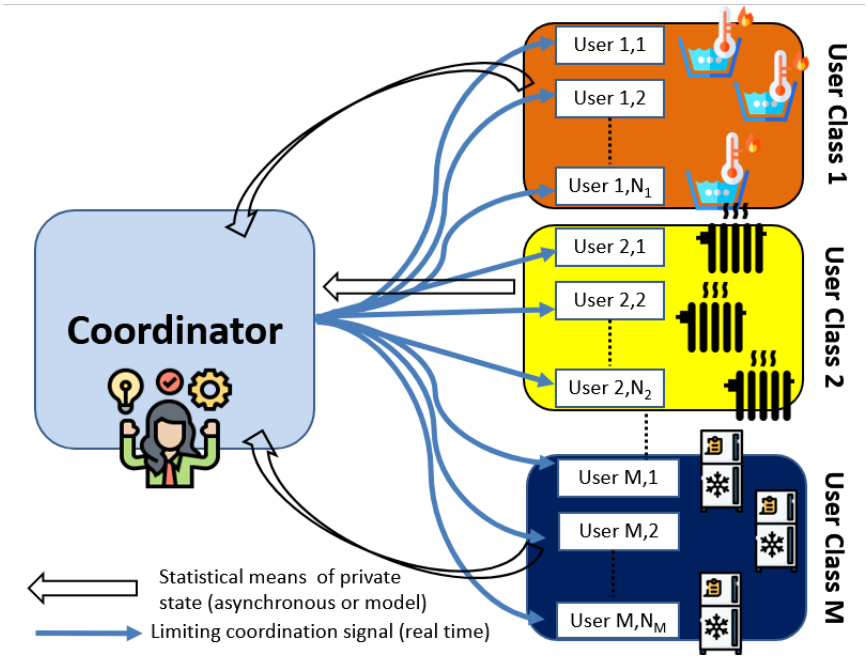


Figure 4.6: Coordination mechanism

Let us adopt the point of view of agent $i \in [N_k]$ of class $k \in [M]$. Consider its limiting individual problem:

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(X_T - x_T^{\mathrm{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - x_s^{\mathrm{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t - u_t^{\mathrm{ref},(k)}\right) + \bar{v}_t^{(\infty)} + \alpha_t^{(k)} Y_t = 0, \end{cases} \tag{4.6.1}$$

where the limiting coordination signal $\bar{v}^{(\infty)} \in \mathcal{H}$ is defined in (4.4.3) and computed by the coordinator by solving the limiting coordination problem (4.4.1). We recall that the solution of the individual problem of agent i of class k is denoted by $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)})$.

Introduce for $t \in [0, T]$ the online coordination signal at time t, denoted by $\bar{v}^{(\infty,t)}$ and defined by:

$$\bar{v}_\tau^{(\infty,t)} := \mathbb{E}_t\left[\bar{v}_\tau^{(\infty)}\right] = \mathbb{E}_t\left[\mathcal{L}_x'\left(\tau, \sum_{l=1}^M \pi^{(l)} \bar{u}_\tau^{(l,\infty)} + \bar{\mathrm{P}}_\tau^{\mathrm{load}} - \mathrm{P}_\tau^{\mathrm{prod}}\right)\right]. \tag{4.6.2}$$

[4]Icons made by Freepik and Smashicons from www.flaticon.com

As $\bar{v}^{(\infty)}$ is $\mathbb{G}$-progressively measurable and as $\mathbb{G}$ is immersed in $\mathbb{F}$, we also have:

$$\bar{v}_\tau^{(\infty,t)} = \mathbb{E}\left[\mathbb{E}\left[\bar{v}_\tau^{(\infty)}|\mathcal{G}_T\right]|\mathcal{F}_t\right] = \mathbb{E}\left[\bar{v}_\tau^{(\infty)}|\mathcal{G}_t\right].$$

In particular, $\bar{v}^{(\infty,t)}$ is a $\mathcal{G}_t$-measurable function of time which can be fully computed by the coordinator (which observes $\mathcal{G}_t$ only at time t). Let us consider the point of view of the agent i of class k at a fixed time $t \in [0,T]$, which aims at computing the current value of its control $u_t^{(k,i)}$. Informally, using the linearity of the conditional expectation and of (4.6.1), replacing $\bar{v}^{(\infty)}$ by its conditional expectation $\bar{v}^{(\infty,t)}$, we can heuristically justify that $(\mathbb{E}_t\left[u^{(k,i,\infty)}\right], \mathbb{E}_t\left[X^{(k,i,\infty)}\right], \mathbb{E}_t\left[Y^{(k,i,\infty)}\right])$ is solution of the following linear FBSDE with unknowns $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:

$$\begin{cases} X_\tau = x_0^{(k,i)} + \int_0^\tau \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \mathbb{E}_t\left[\gamma_s^{(k,i)}\right]\right) \mathrm{d}s \\ Y_\tau = \rho^{(k)}\left(X_T - \mathbb{E}_t\left[x_T^{\mathtt{f},(k)}\right]\right) + \int_\tau^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - \mathbb{E}_t\left[x_s^{\mathtt{ref},(k)}\right]\right)\right) \mathrm{d}s \\ \mu_\tau^{(k)}\left(u_\tau - \mathbb{E}_t\left[u_\tau^{\mathtt{ref},(k)}\right]\right) + \bar{v}_\tau^{(\infty,t)} + \alpha_\tau^{(k)} Y_\tau = 0. \end{cases} \tag{4.6.3}$$

Using the "diagonal" identity

$$(u_t^{(k,i,\infty)}, X_t^{(k,i,\infty)}, Y_t^{(k,i,\infty)}) = (\mathbb{E}_t\left[u_t^{(k,i,\infty)}\right], \mathbb{E}_t\left[X_t^{(k,i,\infty)}\right], \mathbb{E}_t\left[Y_t^{(k,i,\infty)}\right]), \tag{4.6.4}$$

we can recover the current value of the control $u^{(k,i)}$ of the agent at time t. This procedure can be repeated for each time t: the coordinator solves the coordination problem, evaluates the online coordination signal $\bar{v}^{(\infty,t)}$ on the remaining time horizon and sends it to all agents. The agents can then compute the current value of their control by solving a one-dimensional affine-linear FBSDE without additional approximation error. Such a problem is easy to solve, see Theorem 4.5.2.

The above discussion justifies the following decentralized control scheme 4.1.

Algorithm 4.1 Decentralized control scheme

1: **Inputs:** Time grid $(\tau_0, ..., \tau_{N_T})$.
2: **for** $j = 0, ..., N_T$ **do**
3: Wait for $t = \tau_j$.
4: Aggregator observes the common noise $\mathcal{G}_{\tau_j}$, computes the online coordination signal at time τ_j $(\bar{v}_\tau^{(\infty,\tau_j)})_{\tau_j \le \tau \le T}$ given in (4.6.2) and sends it to all agents. This coordination signal is a $\mathcal{G}_{\tau_j}$ measurable function of time.
5: Each agent solves its limiting online individual problem (4.6.3) to get its optimal control $u_{\tau_j}^{(k,i,\infty)} = u_{\tau_j}^{(k,i,\infty,\tau_j)}$ by the diagonal identity (4.6.4).
6: Each agent implements control $u_{\tau_j}^{(k,i,\infty)}$ for its storage system at time τ_j.
7: **end for**

4.6.1 Discussion on the decentralized control scheme

We can make the following remarks on this scheme.

1. **Fast computation of the coordination signal is possible with reasonable computational resource.** The limiting coordination problem (4.4.1) is a M-dimensional FBSDE, equivalent to a M-dimensional control problem. The control problem is relatively easy to solve in a linear-quadratic setting, using similar arguments as in Section 4.5.2, and it can be solved using numerical methods in other cases. Computation of the parameters of this problem is easy as well, under some assumption like affine-linear stochastic processes.

2. **Fast computation of the individual controls are possible by agents equipped with limited computational resources.** The online limiting individual problems at time t (4.6.3) are linear one-dimensional FBSDEs, hence particularly easy to solve, see Section 4.5.2. Considering deterministic coefficients or stochastic processes with affine-linear drift can make computations even easier.

3. **The parameters of the problems solved by the agents and the coordinator are locally available.** Indeed, the parameters of the limiting coordination problem at time t can be computed by an aggregator only observing the common information $\mathbb{G}$. The parameters of the online limiting individual problems of agent (k,i) at time t are all available locally for agent (k,i). This includes the shared information $\mathcal{G}_t$, the online limiting coordination signal received, and individual parameters of the energy storage system of agent (k,i). In particular, the computation of conditional expectations of the parameters of individual energy storage systems is simplified by our assumption of conditional independence of these parameters.

4. **Limited telecommunication is required.** A single online coordination signal at time t $(\bar{v}_\tau^{(\infty,t)})\tau \in [0,T]$ is sent to agents by the coordinator, so that no specific routing is needed. This signal is a one dimensional $\mathcal{G}_t$-measurable function of time, hence it encoding is easy, for instance by discretization/interpolation or by regression against a function basis (like Fourier). No real-time communication from agents to the coordinator is required.

4.6.2 On the privacy of individual users habits

In order for an aggregator to come up with good stochastic models of the empirical averages of the class parameters $(\bar{\mathrm{P}}^{\mathtt{load},(k,N)}, \bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k\in[M]}$, one may imagine that the aggregator is given some historical realization of these processes. As these processes are aggregates of individual data of consumers, which may be subject to some privacy requirements, one can use the Secure Multiparty Computation (SMC) technique in [Yao86] in order to deal with privacy concerns. Indeed, this technique would allow the coordinator to compute the empirical averages of the class parameters, while guaranteeing that the values of the parameters of individual agents remain unknown to him. This method has already been used in the context of energy management in [Jac+19].

4.7 Proofs

4.7.1 Proof of Theorem 4.2.1

The existence and uniqueness of an optimal control are proved using standard arguments of functional analysis. We give main arguments and leave full details to the reader. The convexity directly stems from the linearity of the dynamic and the quadratic/convex functions in the definition of $\mathcal{J}$. In addition, the strong convexity of $\mathcal{J}$ comes from the uniform lower bound on μ. It directly yields the coercivity of $\mathcal{J}$. As $\mathcal{J}$ is additionally continuous, we get the existence of a minimizer of $\mathcal{J}$ from [Bre10b, Corollary 3.23, pp.71]. This minimizer is unique from the strict convexity of $\mathcal{J}$. We denote it by $(u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$.

The characterization of optimality is proved applying the stochastic Pontryagin principle. However, our setting of optimal control of ODE with non-Markovian coefficients in general filtrations differs from standard references: see [Yon99] for the case of Brownian filtrations, see [CD18, pp. 543-552, Volume I] when incorporating McKean-Vlasov terms. The closest reference to our setting is presumably [Cad02] in the (Markovian) SDE case with jumps, under different integrability conditions. This motivates us to give a proof of our result in our specific setting.

Let $u = (u^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$. By our integrability assumptions, there is existence and uniqueness of $X = (X^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ solution of the ODE:

$$\forall k \in [M], \forall i \in [N_k], \quad X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}\right) \mathrm{d}s. \tag{4.7.1}$$

Now, let us consider the Backward Stochastic Differential Equation in $(Y,\tilde{M}) := (Y^{(k,i)}, \tilde{M}^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}\times\mathcal{S}$ with $\tilde{M}$ a square integrable martingale vanishing at $t=0$:

$$\forall k \in [M], \forall i \in [N_k], \begin{cases} -\mathrm{d}Y_t^{(k,i)} = \left(\beta_t^{(k)} Y_t^{(k,i)} + \nu_t^{(k)} (X_t^{(k,i)} - x_t^{\mathtt{ref},(k,i)})\right) \mathrm{d}t - \mathrm{d}\tilde{M}_t^{(k,i)}, \\ Y_T^{(k,i)} = \rho^{(k)} \left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right). \end{cases}$$

It is an affine-linear BSDE in a general filtration, and our boundedness and integrability assumptions on its coefficients ensure existence and uniqueness of its solution, see [EPQ97, Theorem 5.1, p. 54].

Now, the arguments for proving the Gateaux-differentiability of $\mathcal{J}$ are standard and follow the ones in [CD18, pp. 543-548, Volume I]: we show Gateaux-differentiability of the state variable and of the cost functional successively. Define the application $\phi_X : u := (u^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H} \mapsto X^u := (X^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ by (4.7.1). Then ϕ_X is Gateaux differentiable and its Gateaux derivative at $u := (u^{(k,i)})_{k\in[M],i\in[N_k]}$ in direction $v := (v^{(k,i)})_{k\in[M],i\in[N_k]}$ is given by $\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}X^{u+\varepsilon v}\right)_{|\varepsilon=0} = \dot{X}^v := (\dot{X}^{v,(k,i)})_{k\in[M],i\in[N_k]}$ with:

$$\dot{X}_t^{v,(k,i)} = \int_0^t \left(\alpha_s^{(k)} v_s^{(k,i)} + \beta_s^{(k)} \dot{X}_s^{v,(k,i)}\right) \mathrm{d}s.$$

This can be proved following arguments of the proof of [CD18, Lemma 6.10, pp.544-545, Volume I]. Now, following arguments in [CD18, pp. 546-548, Volume I], we get that $\mathcal{J}$ is Gateaux-differentiable and its Gateaux derivative at u in direction v is given by $\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(u+\varepsilon v)\right)_{|\varepsilon=0} = \dot{\mathcal{J}}(u,v)$ where:

$$\begin{aligned}\dot{\mathcal{J}}(u,v) =& \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\int_0^T \left(\mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right) v_t^{(k,i)} + \nu_t^{(k)}\left(X_t^{(k,i)} - x_t^{\mathtt{ref},(k,i)}\right)\dot{X}_t^{v,(k,i)}\right)\mathrm{d}t\right] \\ &+ \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right)\dot{X}_T^{v,(k,i)} + \int_0^T \mathcal{L}_x'\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}\right) - \mathrm{P}_t^{\mathtt{prod}}\right) v_t^{(k,i)}\mathrm{d}t\right)\right].\end{aligned}$$

Then, applying Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to $Y \cdot \dot{X}^v := \sum_{k=1}^{M}\sum_{i=1}^{N_k} Y^{(k,i)}\dot{X}^{v,(k,i)}$ between $t = 0$ and $t = T$ yields, using $\dot{X}_0^v = 0$ and $Y_T^{(k,i)} = \rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right)$, we finally obtain the following expression for the Gateaux derivative of $\mathcal{J}$ at u in direction v:

$$\dot{\mathcal{J}}(u,v) = \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\int_0^T \left\{\mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right) + \mathcal{L}_x'\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}\right) - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k,i)}\right\} v_t^{(k,i)}\mathrm{d}t\right]. \tag{4.7.2}$$

By convexity and differentiability of $\mathcal{J}$ and by uniqueness of its minimizer, $(u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$ is also the unique critical point of $\mathcal{J}$. Combining this with the expression of the Gateaux derivative of $\mathcal{J}$, we get that the term inside the brackets in (4.7.2) is 0 for all t. Therefore, $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ is the unique solution of the FBSDE (4.2.4). □

4.7.2 Proof of Proposition 4.3.1

Using Theorem 4.2.1 and the definition of the empirical mean processes, one can directly show that the empirical mean processes are solution of (4.3.1). This FBSDE fully characterizes the solution in $(u^{(k)})_{k\in[M]} \in \mathcal{H}^M$ of the following stochastic control problem:

$$\begin{aligned}\min_{(u^{(k)})_{k\in[M]}\in\mathcal{H}^M} \quad & \mathbb{E}\left[\sum_{k=1}^{M}\pi^{(k)}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k,N)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k)} - \bar{x}_t^{\mathtt{ref},(k,N)}\right)^2\right)\mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k,N)}\right)^2\right\}\right] \\ & + \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\sum_{l=1}^{M}\pi^{(l)}u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right], \\ s.t. \quad & X_t^{(k)} = \frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)} + \int_0^t\left(\alpha_s^{(k)}u_s^{(k)} + \beta_s^{(k)}X_s^{(k)} + \bar{\gamma}_s^{(k,N)}\right)\mathrm{d}s, \quad \forall k\in[M].\end{aligned} \tag{4.7.3}$$

This results from applying Pontryagin's principle, up to scaling of the k-th adjoint variable by $\frac{1}{\pi^{(k)}}$, and by similar arguments as the ones used in the proof of Theorem 4.2.1. The uniqueness of the solution of the control problem, and hence of the solution of the FBSDE (4.3.1) from similar arguments as the ones used in the proof of Theorem 4.2.1. □

4.7.3 Proof of Proposition 4.4.3

By independence of $(\mathrm{P}_t^{\mathrm{load},(l,j)})_{1\le l\le M, 1\le j\le N_l}$ conditionally to $\mathcal{G}_T$:

$$\mathbb{V}\mathrm{ar}\left[\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathrm{P}_t^{\mathrm{load},(l,j)}|\mathcal{G}_T\right] = \frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{V}\mathrm{ar}\left[\mathrm{P}_t^{\mathrm{load},(l,j)}|\mathcal{G}_T\right],$$
$$\mathbb{E}\left[\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathrm{P}_t^{\mathrm{load},(l,j)}|\mathcal{G}_T\right] = \bar{\mathrm{P}}_t^{\mathrm{load}}.$$

This yields:

$$\mathbb{E}\left[(\bar{\mathrm{P}}_t^{\mathrm{load},(N)} - \bar{\mathrm{P}}_t^{\mathrm{load}})^2\right] = \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\bar{\mathrm{P}}_t^{\mathrm{load},(N)}|\mathcal{G}_T\right]\right] = \frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\mathrm{P}_t^{\mathrm{load},(l,j)}|\mathcal{G}_T\right]\right] \le \frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{E}\left[\left(\mathrm{P}_t^{\mathrm{load},(l,j)}\right)^2\right].$$

This yields, integrating over time and using the fact that all $\mathrm{P}^{\mathrm{load},(l,j)}$ are bounded in $\mathcal{H}$ by a constant independent from N:

$$||\bar{\mathrm{P}}^{\mathrm{load},(N)} - \bar{\mathrm{P}}^{\mathrm{load}}||_{\mathcal{H}}^2 \le \frac{C}{N}$$

Similarly, we obtain the convergence in $\mathcal{H}$ of $\bar{\gamma}^{(k,N)}$ (resp. $\bar{u}^{\mathrm{ref},(k,N)}$, resp. $\bar{x}^{\mathrm{ref},(k,N)}$) to $\bar{\gamma}^{(k)}$ (resp. $\bar{u}^{\mathrm{ref},(k)}$, resp. $\bar{x}^{\mathrm{ref},(k)}$) at speed $\frac{1}{\sqrt{N_k}}$, and the convergence in $\mathbb{L}^2$ of $\bar{x}_T^{\mathrm{f},(k)}$ to $\bar{x}_T^{\mathrm{f},(k)}$ at speed $\frac{1}{\sqrt{N_k}}$. □

4.7.4 Proof of Theorem 4.4.5

Consider the following FBSDE with $\mathbb{G}$-progressively measurable coefficients:

$$\forall k\in[M],$$
$$\begin{cases} X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)}u_s^{(k)} + \beta_s^{(k)}X_s^{(k)} + \bar{\gamma}_s^{(k)}\right)\mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)}Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathrm{ref},(k)}\right)\right)\mathrm{d}s|\mathcal{G}_t\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathrm{ref},(k)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)}u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right) + \alpha_t^{(k)}Y_t^{(k)} = 0. \end{cases} \tag{4.7.4}$$

The above FBSDE is the optimality system associated to the following stochastic control problem considered in in $(\Omega, \mathcal{G}_T, \mathbb{G}, \mathbb{P})$:

$$\min_{(u^{(k)})_{1\le k\le M}\in\mathcal{H}} \quad \mathbb{E}\left[\sum_{k=1}^M \pi^{(k)}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)} - \bar{u}_t^{\mathrm{ref},(k)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k)} - \bar{x}_t^{\mathrm{ref},(k)}\right)^2\right)\mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right)^2\right\}\right]$$
$$+ \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\sum_{l=1}^M \pi^{(l)}u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right)\mathrm{d}t\right],$$
$$s.t. \quad X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)}u_s^{(k)} + \beta_s^{(k)}X_s^{(k)} + \bar{\gamma}_s^{(k)}\right)\mathrm{d}s, \quad \forall k\in[M].$$

Our assumptions and [Bre10b, Corollary 3.23, pp.71] show that the above problem has a unique solution $\tilde{u}\in\mathcal{H}_{\mathbb{G}}$ and therefore, the FBSDE (4.7.4) has a unique solution $(\tilde{u}, \tilde{X}, \tilde{Y})\in\mathcal{H}_{\mathbb{G}}\times\mathcal{S}_{\mathbb{G}}\times\mathcal{S}_{\mathbb{G}}$. Now consider the $\mathbb{G}$-martingales:

$$\tilde{M}_t^{(k)} := \mathbb{E}\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right) + \int_0^T \left(\beta_s^{(k)}\tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathrm{ref},(k)}\right)\right)\mathrm{d}s|\mathcal{G}_t\right].$$

Noting that $\mathbb{G}$ is immersed in $\mathbb{F}$ (see [CD18, Definition 1.2, p. 5, Volume II]), for all $k\in[M]$, $\tilde{M}^{(k)}$ is a $\mathbb{G}$-square integrable martingale and therefore, by definition, it is a $\mathbb{F}$-square integrable martingale, so that:

$$\tilde{M}_t^{(k)} := \mathbb{E}_t\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathrm{f},(k)}\right) + \int_0^T \left(\beta_s^{(k)}\tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathrm{ref},(k)}\right)\right)\mathrm{d}s\right].$$

Therefore, we have for all $k \in [M]$, by the previous two expressions of $\tilde{M}^{(k)}$ and using the fact that $\tilde{X}^{(k)}$, $\tilde{Y}^{(k)}$ and $\bar{x}^{\mathtt{ref},(k)}$ are $\mathbb{G}$ and $\mathbb{F}$ progressively measurable (as $\mathbb{G}$ is assumed immersed in $\mathbb{F}$):

$$\begin{aligned}
\tilde{Y}_t^{(k)} &= \mathbb{E}\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)}\tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s | \mathcal{G}_t\right] \\
&= \tilde{M}_t^{(k)} - \int_0^t \left(\beta_s^{(k)}\tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s \\
&= \mathbb{E}_t\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)}\tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right].
\end{aligned}$$

Besides, $\tilde{u}$ and $\tilde{X}$ are $\mathbb{F}$-progressively measurable. Therefore, $(\tilde{u}, \tilde{X}, \tilde{Y})$ is also solution of the optimality system of the control problem considered with the filtration $\mathbb{F}$. By uniqueness of such a solution, we deduce that $(\bar{u}, \bar{X}, \bar{Y})$ coincides with $(\tilde{u}, \tilde{X}, \tilde{Y})$ and therefore $(\bar{u}, \bar{X}, \bar{Y})$ is $\mathbb{G}$-progressively measurable. □

4.7.5 Proof of Proposition 4.4.6

The uniqueness of solution of the above FBSDE arises from similar arguments as in the proof of Theorem 4.2.1.

Consider the function $\chi : [0,T] \times \mathbb{R}^M \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M \mapsto \mathbb{R}$:

$$\chi : (t, u, v, y, u^{\mathtt{ref}}) \mapsto \sum_{k=1}^M \left(\frac{\pi^{(k)}\mu^{(k)}}{2}(u_k - u^{\mathtt{ref},(k)})^2 + \pi^{(k)}\alpha_t^{(k)} y_k u_k\right) + \mathcal{L}_t\left(\sum_{l=1}^M \pi^{(l)} u_i + v\right).$$

It is straightforward to observe that χ is twice continuously differentiable in $(u, v, y, u^{\mathtt{ref}})$.

For any $(t, v, y, u^{\mathtt{ref}}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M$, $u \mapsto \chi(t, u, v, y, u^{\mathtt{ref}})$ is $\min_k\{\pi_k\mu^{(k)}\}$-strongly convex, and as such, this function admits a unique minimizer and its Hessian is positive semi-definite. Besides, its Hessian is invertible with inverse bounded by $\frac{1}{\min_k\{\pi_k\mu^{(k)}\}}$.

By the implicit function theorem, this directly implies, for any $(t, v, y, u^{\mathtt{ref}}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M$, the equation in $u \in \mathbb{R}^M$

$$\nabla_u \chi(t, u, v, y, u^{\mathtt{ref}}) = 0$$

has a unique solution $u = \tilde{u}(t, v, y, u^{\mathtt{ref}})$ with $\tilde{u}$ continuously differentiable in $(v, y, u^{\mathtt{ref}})$. Besides, we have:

$$\begin{aligned}
\nabla_v \tilde{u}(t, v, y, u^{\mathtt{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right)^{-1}\left(\nabla^2_{u,v}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right), \\
\nabla_y \tilde{u}(t, v, y, u^{\mathtt{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right)^{-1}\left(\nabla^2_{u,y}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right), \\
\nabla_{u^{\mathtt{ref}}} \tilde{u}(t, v, y, u^{\mathtt{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right)^{-1}\left(\nabla^2_{u,u^{\mathtt{ref}}}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right).
\end{aligned}$$

Using the bound $\|\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathtt{ref}}), v, y, u^{\mathtt{ref}})\right)^{-1}\| \leq \frac{1}{\min_k\{\pi_k\mu^{(k)}\}}$ and as the second order derivative of $\mathcal{L}$ with respect to x is uniformly bounded, we obtain:

$$\|\nabla_v \tilde{u}(t, v, y, u^{\mathtt{ref}})\| + \|\nabla_y \tilde{u}(t, v, y, u^{\mathtt{ref}})\| + \|\nabla_{u^{\mathtt{ref}}} \tilde{u}(t, v, y, u^{\mathtt{ref}})\| \leq \frac{C}{\min_k\{\pi_k\mu_t^{(k)}\}}.$$

Then there exists a constant C which grows like $\frac{1}{\min_{1\leq k\leq M, t\in[0,T]}\{\pi_k\mu_t^{(k)}\}}$ such that, for any $y^1, y^2 \in \mathbb{R}^M$, $v^1, v^2 \in \mathbb{R}$ and $u^{\mathtt{ref},1}, u^{\mathtt{ref},2} \in \mathbb{R}^M$, we have:

$$\|\tilde{u}(t, v^1, y^1, u^{\mathtt{ref},1}) - \tilde{u}(t, v^2, y^2, u^{\mathtt{ref},2})\|_{\mathbb{R}^M} \leq C\left(|v^1 - v^2| + \|y^1 - y^2\|_{\mathbb{R}^M} + \|u^{\mathtt{ref},1} - u^{\mathtt{ref},2}\|_{\mathbb{R}^M}\right).$$

This implies, for $\theta^1 = (x^1, v^1, w^1, u^{\mathtt{ref},1}, x^{\mathtt{ref},1}, x_T^{\mathtt{f},1})$ and $\theta^2 = (x^2, v^2, w^2, u^{\mathtt{ref},2}, x^{\mathtt{ref},2}, x_T^{\mathtt{f},2})$ in $\mathbb{R}^M \times \mathcal{H} \times \mathcal{H}(\mathbb{R}^M) \times \mathcal{H}(\mathbb{R}^M) \times \mathbb{L}^2_T(\mathbb{R}^M)$:

$$\|u^{\theta^1} - u^{\theta^2}\|_{\mathcal{H}} \leq C\left(\|v^1 - v^2\|_{\mathcal{H}} + \|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{H}} + \|u^{\mathtt{ref},1} - u^{\mathtt{ref},2}\|_{\mathcal{H}}\right).$$

In the following, C_T denotes a constant depending on the input parameters of the problem and depending continuously on T, with finite limit when $T \to 0$.

Applying Gronwall's lemma to the state equations, we obtain:

$$\|X^{\theta^1} - X^{\theta^2}\|_{\mathcal{S}} \leq C_T \left(\|x^1 - x^2\|_{\mathbb{R}^M} + \|u^{\theta^1} - u^{\theta^2}\|_{\mathcal{H}} + \|w^1 - w^2\|_{\mathcal{H}} \right).$$

Applying Gronwall's lemma and Cauchy-Schwartz inequality to the adjoint equations yields:

$$\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} \leq C_T \left(\|X^{\theta^1} - X^{\theta^2}\|_{\mathcal{S}} + \|x^{\mathtt{ref},1} - x^{\mathtt{ref},2}\|_{\mathcal{H}} + \|x_T^{\mathtt{f},1} - x_T^{\mathtt{f},2}\|_{\mathbb{L}^2} \right).$$

Combining the previous inequalities, we get:

$$\begin{aligned} \|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} &\leq C_T \left(\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{H}} + \|\theta^1 - \theta^2\| \right) \\ &\leq C_T \left(\sqrt{T} \|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} + \|\theta^1 - \theta^2\| \right). \end{aligned}$$

Then, using the fact that C_T is bounded for T small, we obtain, for any T small enough, so that $C_T \sqrt{T} < 1$:

$$\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} \leq C_T \|\theta^1 - \theta^2\|.$$

Combining with the above estimations, this finally yields:

$$\|(\bar{u}^{\theta^1} - \bar{u}^{\theta^2}, \bar{X}^{\theta^1} - \bar{X}^{\theta^2}, \bar{Y}^{\theta^1} - \bar{Y}^{\theta^2})\|_{\mathcal{H}} \leq C_T \|\theta^1 - \theta^2\|.$$

□

4.7.6 Proof of Corollary 4.4.8

Apply Proposition 4.4.6 to $x^1 = \left(\bar{x}_0^{(k)}\right)_{1 \leq k \leq M}$, $x^2 = \left(\frac{1}{N_k} \sum_{j=1}^{N_k} x_0^{(k,j)}\right)_{1 \leq k \leq M}$, $v^1 = \bar{\mathrm{P}}^{\mathtt{load}} - \mathrm{P}^{\mathtt{prod}}$, $v^2 = \bar{\mathrm{P}}^{\mathtt{load},(N)} - \mathrm{P}^{\mathtt{prod}}$, $w^1 = \left(\bar{\gamma}^{(k)}\right)_{1 \leq k \leq M}$, $w^2 = \left(\bar{\gamma}^{(k,N)}\right)_{1 \leq k \leq M}$, $u^{\mathtt{ref},1} = (\bar{u}^{\mathtt{ref},(k)})_{1 \leq k \leq M}$, $u^{\mathtt{ref},2} = (\bar{u}^{\mathtt{ref},(k,N)})_{1 \leq k \leq M}$, $x^{\mathtt{ref},1} = (\bar{x}^{\mathtt{ref},(k)})_{1 \leq k \leq M}$, , $x^{\mathtt{ref},2} = (\bar{x}^{\mathtt{ref},(k,N)})_{1 \leq k \leq M}$, , $x_T^{\mathtt{f},1} = (\bar{x}_T^{\mathtt{f},(k)})_{1 \leq k \leq M}$, $x_T^{\mathtt{f},1} = (\bar{x}_T^{\mathtt{f},(k)})_{1 \leq k \leq M}$. Then use Proposition 4.4.3 and the assumption $\pi^{(k)} \geq \eta > 0$ for all $1 \leq k \leq N$ to conclude. □

4.7.7 Proof of Theorem 4.4.12

By applying the previous proposition to $v = \bar{v}^{(N)} = \frac{1}{N} \sum_{l=1}^{M} \sum_{j=1}^{N_l} \left(u^{(l,j,N)} + \mathrm{P}^{\mathtt{load},(l,j)}\right) - \mathrm{P}^{\mathtt{prod}}$ and $v' = \bar{v}^{(\infty)} = \sum_{l=1}^{M} \pi^{(l)} \left(\bar{u}^{(l,\infty)} + \bar{\mathrm{P}}^{\mathtt{load},(l)}\right) - \mathrm{P}^{\mathtt{prod}}$, we obtain using Corollary 4.4.8 and the definition of $\bar{v}^{(\infty)}$ and $\bar{v}^{(N)}$:

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \leq C_T \|\bar{v}^{(\infty)} - \bar{v}^{(N)}\|_{\mathcal{H}} \leq \frac{C_T}{\sqrt{N}}.$$

For $k \in [M]$, $i \in [N_k]$, we introduce the notation $u^{\Delta,(k,i)} := u^{(k,i,\infty)} - u^{(k,i,N)}$, for $\sigma \in [0,1]$, $u^{(\sigma),(k,i)} := u^{(k,i,N)} + \sigma(u^{(k,i,\infty)} - u^{(k,i,N)})$, $\bar{u}^{(\sigma)} := \frac{1}{N} \sum_{k=1}^{M} \sum_{i=1}^{N_k} u^{(\sigma),(k,i)}$.

We have by Taylor formula, by the form of the Gateaux derivative of $\mathcal{J}$ given in (4.7.2) and since $\dot{\mathcal{J}}(u^{(N)}, u^{(\infty)} - u^{(N)}) = 0$ by optimality of $u^{(N)}$:

$$\begin{aligned} \mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) &= \int_0^1 \dot{\mathcal{J}}(u^{(N)} + \sigma(u^{(\infty)} - u^{(N)}), u^{(\infty)} - u^{(N)}) \mathrm{d}\sigma \\ &= \int_0^1 \left(\dot{\mathcal{J}}(u^{(N)} + \sigma(u^{(\infty)} - u^{(N)}), u^{(\infty)} - u^{(N)}) - \dot{\mathcal{J}}(u^{(N)}, u^{(\infty)} - u^{(N)}) \right) \mathrm{d}\sigma \\ &= \int_0^1 \mathbb{E}\left[\frac{1}{N} \sum_{k=1}^{M} \sum_{i=1}^{N_k} \int_0^T \left\{ \mu_t^{(k)} \sigma\left(u_t^{(k,i,\infty)} - u_t^{(k,i,N)}\right) + \alpha_t^{(k)} \sigma(Y_t^{(k,i,\infty)} - Y^{(k,i,N)}) \right\} (u_t^{(k,i,\infty)} - u_t^{(k,i,N)}) \mathrm{d}t \right] \mathrm{d}\sigma \\ &\quad + \int_0^1 \mathbb{E}\left[\int_0^T \left\{ \mathcal{L}_x'\left(t, \bar{u}_t^{(\sigma)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right) - \mathcal{L}_x'\left(t, \bar{u}_t^{(0)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right) \right\} \left(\frac{1}{N} \sum_{k=1}^{M} \sum_{i=1}^{N_k} u_t^{\Delta,(k,i)} \right) \mathrm{d}t \right] \mathrm{d}\sigma, \end{aligned}$$

where we used the affine-linearity of the state and adjoint variables with respect to the control variable.

In what follows, C denotes a constant independent from N, which depends on data of the problem and may change from one line to another.

We then use Cauchy-Schwartz inequality, Taylor formula applied to $\sigma \mapsto \mathcal{L}'_x\left(t, \bar{u}_t^{(\sigma)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right)$ as well as $\bar{u}^{(\sigma)} = \bar{u}^{(0)} + \sigma\left(\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u^{\Delta,(k,i)}\right)$ to obtain the following upper bound:

$$\begin{aligned}\mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) \leq \frac{C}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\|Y^{(k,i,N)} - Y^{(k,i,\infty)}\|_{\mathcal{H}} + \|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}\right)\|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}} \\ + \int_0^1 \mathbb{E}\left[\int_0^T\int_0^\sigma \mathcal{L}''_{xx}\left(t, \bar{u}_t^{(r)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right)\left(\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u_t^{\Delta,(k,i)}\right)^2 \mathrm{d}r\mathrm{d}t\right]\mathrm{d}\sigma.\end{aligned}$$

Using the boundedness of the second-order derivative of $(t,x) \mapsto \mathcal{L}''_{xx}(t,x)$ uniformly in $t \in [0,T]$ and x, we get:

$$\mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) \leq \frac{C}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\|Y^{(k,i,N)} - Y^{(k,i,\infty)}\|_{\mathcal{H}} + \|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}\right)\|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}.$$

Combine this inequality with the previous bound:

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N}},$$

and the fact that $u^{(N)}$ minimizes $\mathcal{J}$ to get:

$$0 \leq \mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) \leq \frac{C}{N}.$$

4.8 Conclusion

We have formulated a control problem to model a cooperative setting where Thermostatically Controlled Loads distributed among a large population of agents are used to balance power production and consumption in a context of strong uncertainty created by renewable energy sources. Necessary and sufficient optimality conditions are given, in the form of a high-dimensional FBSDE. The curse of dimensionality one may expect can be dealt with by an appropriate decomposition method, which shows that the high-dimensional FBSDE is equivalent to lower-dimension FBSDEs: a coordination problem and individual problems. In particular, we show the optimal solution of the (centralized) stochastic control problem can be obtained by computing the (unique) Nash equilibrium of an associated Stochastic Stackelberg Differential Game, with a coordinator (leader) solving a control problem and sending coordination signal to the agents (followers), solving their own individual problems. This allows a decentralized implementation. Under a conditional independence-type assumption and in the limit of large population, we show a mean-field type approximation of the problem of the coordinator, which does not require the aggregator to observe the behaviors of the agents, in the framework of the associated Stochastic differential game. This is desirable for both preserving privacy of consumers and reducing the need for real-time telecommunication between agents and the coordinator. Numerical results show the performance of the approach and the quality of the mean-field approximation. The experiments demonstrate the need to carefully tune the cost parameters of the problem in order to maximize the contribution of the TCLs to power balancing while ensuring that individual constraints of the devices are not violated. A decentralized and online implementation of the control mechanism with minimal one-way communication from the aggregator to the agents is also proposed, allowing for the coordinator and the agents to solve in real-time their respective problems.

Chapter 5

Newton method for stochastic control problems, with applications to the management of energy storage systems

5.1 Introduction

In this chapter, we introduce a new method to solve stochastic control problems, which is a generalization of the Newton method to the particular (infinite-dimensional) setting of stochastic control problems. We consider problems in general filtrations with a linear dynamic. For the sake of simplicity, we restrict ourselves to the one-dimensional setting (meaning the state and control variables are real-valued stochastic processes). The general form of problems we consider is:

$$\left.\begin{aligned} \mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t,\omega,u_{t,\omega},X^u_{t,\omega}\right)\mathrm{d}t + \Psi(\omega,X^u_{T,\omega})\right] \\ \text{s.t. } X^u_{t,\omega} = x_0 + \int_0^t (\alpha_{s,\omega}u_{s,\omega} + \beta_{s,\omega}X^u_{s,\omega})\mathrm{d}s + M_{t,\omega}. \end{aligned}\right\} \longrightarrow \min_u . \tag{5.1.1}$$

To properly introduce the Newton method for stochastic control problems, we give an overview of state-of-the-art numerical methods for this class of problems, then a brief introduction to the Newton method for the optimization of functions taking values in $\mathbb{R}^d$ with $d \in \mathbb{N}$.

State of the art of numerical methods for stochastic control problems Standard approaches to solve stochastic control problems are based either on Bellman dynamic programming principle, either on Pontryagin's optimality principle.

The dynamic programming principle gives rise to a non-linear Partial Differential Equation (PDE) called Bellman's equation, which is satisfied by the value function under reasonable conditions [Pha09; Kry08]. Finite-difference methods to solve this type of PDE have been studied in [GŠ09] and [BZ03] for instance, which allow to reduce the problem to a high dimensional non-linear system of equations. Among other methods based on Bellman's principle, the Howard's policy improvement algorithm is an iterative algorithm where a linearized version of the Bellman's equation is solved at each step. This method has been introduced by Howard in [How60], in the context of Markovian decision processes. A global convergence rate for Howard policy improvement algorithm (and a variant) for stochastic control problems have been recently established in [KŠS20b]. Deep-learning methods have also been applied to solve the non-linear Bellman PDE arising in this context [HJW18].

Another approach to solve stochastic optimal control is the Pontryagin principle, which gives rise to a Forward-Backward Stochastic Differential Equation, see [Zha17]. Methods to solve this type of equations include fixed-point methods such as Picard iterations or the method of Markovian iterations for coupled FBSDEs in [BZ+08]. They converge for small time horizon under the assumption of Lipschitz coefficients, but convergence can be proved

for arbitrary time horizon under some monotony assumption [PT99], using norms with appropriate exponential weights. Among methods related to fixed-point iteration, the Method of Successive Approximations is an iterative method based on Pontryagin's principle, which was proposed in [CL82] for deterministic control problems. However, convergence is not guaranteed in general. This method is refined in the yet to be published work [KŠS20a], for stochastic control problems, using an a modification of the Hamiltonian, called augmented Hamiltonian, which allows to show global convergence, and even establish a convergence rate for particular structures of problems. Solvability of FBSDEs for arbitrary time horizon under monotony conditions can be proved using the continuation method in [HP95; Yon97; PW99]. This method is well developed theoretically to prove the existence of solutions of a FBSDE, but rarely used to design algorithms to solve the problem numerically. Another method to solve FBSDEs is the Four step scheme introduced in [MPY94] in the case of a non-degenerate diffusion term σ, which allows to compute a so-called decoupling field as solution of a quasi-linear PDE. This decoupling field allows to express the adjoint variable as a feedback of the state variable. Some Deep-learning based algorithms have been recently proposed to solve FBSDEs in [Ji+20] and [HL20].

The case of linear FBSDEs and of linear quadratic stochastic optimal control problems has been extensively studied see for instance [Bis76; Yon99; Yon06]. Our result builds on these works, as our algorithm is based on successive linearizations of non-linear FBSDEs obtained by a Taylor expansion.

Preliminary on the Newton method in $\mathbb{R}^d$. Consider a twice-differentiable convex function $f : x \in \mathbb{R}^d \mapsto \mathbb{R}$. We wish to solve the minimization problem $\min_{x\in\mathbb{R}^d} f(x)$ If f is strongly convex and its second-order derivative is Lipschitz-continuous [Kan48; NW06; BV04] or if f is strictly convex and self-concordant [NN94; BV04] (meaning that f is three times differentiable and $\frac{d}{d\alpha}\nabla^2 f(x+\alpha y)\big|_{\alpha=0} \leq 2\sqrt{y^T\nabla^2 f(x)\, y}\,\nabla^2 f(x)$, for all x, y in $\mathbb{R}^d$), then the Newton method gives a sequence of points which converges locally quadratically to the global minimizer of f. This means that if the initial point is sufficiently close to the optimum, the convergence of the sequence to this point is very fast. The pseudo-code for the Newton method is given in Algorithm 5.1.

Algorithm 5.1 Newton's method

1: $\varepsilon > 0$, $k = 0$, $x^{(0)}$ fixed
2: **while** $|\nabla f(x^{(k)})| > \varepsilon$ **do**
3: Compute Newton direction Δ_x solution of the linear equation $\nabla^2 f(x^{(k)})(\Delta_x) = -\nabla f(x^{(k)})$
4: Compute new iterate $x^{(k+1)} = x^{(k)} + \Delta_x$.
5: $k \leftarrow k+1$
6: **end while**
7: **return** $x^{(k)}$

To obtain global convergence of the Newton method (i.e., convergence to the optimum no matter the choice of the initial point), a common procedure is to use line search methods, allowing to choose a step size σ and to define new iterates by the formula $x^{(k+1)} = x^{(k)} + \sigma\Delta_x$ instead of the standard Newton method which considers the case $\sigma = 1$. Among them, given $\beta \in (0,1)$, the backtracking line search procedure allows to find the largest value among $\{1, \beta, \beta^2, \beta^3, ...\}$ which satisfies a sufficient decrease condition. Its pseudo-code is given in Algorithm 5.2. The combination of Newton method with backtracking line search gives a globally convergent method [BV04] in $\mathbb{R}^d$.

Algorithm 5.2 Backtracking line search procedure

1: **Inputs:** Current point $x \in \mathbb{R}$, Current search direction Δ_x, $\beta \in (0,1)$, $\gamma \in (0,1)$.
2: $\sigma = 1$.
3: **while** $f(x+\sigma\Delta_x) > f(x) + \gamma\sigma\nabla f(x)\cdot\Delta_x$ **do**
4: $\sigma \leftarrow \beta\sigma$.
5: **end while**
6: **return** $x + \sigma\Delta_x$.

Our contributions. To solve numerically the problem proposed in (5.1.1), we extend the Newton method to the infinite-dimensional setting of convex stochastic control problems in general filtrations, where the dynamics is an affine-linear Ordinary Differential Equation (ODE). This iterative method generates a sequence of points which are solution of successive Linear-Quadratic approximation of the stochastic control problem around a current estimate of the solution, see Proposition 5.3.2. Equivalently, it can be interpreted as successive linearizations of the Forward Backward Stochastic Differential Equation (FBSDE) arising from the stochastic Pontryagin's principle around a current estimate of the solution, see Proposition 5.3.3.

In section 5.3.2, a full methodology is proposed to solve affine-linear FBSDEs with random coefficients in general filtrations, which arise when computing the Newton step. The methodology is quite standard, though the framework is a bit unusual as we do not assume Brownian filtrations. In particular, we show that solving the solution of linear FBSDEs or the computation of the Newton step require solving a Riccati BSDE and an affine-linear BSDE, see Theorem 5.3.7 and Corollary 5.3.8.

The convergence of Newton's method typically requires sufficient regularity of the second-order derivative of the cost functional, see [Kan48; NW06] for the case of a Lipschitz second order derivative and [NN94] for the self-concordant case. Such regularity is not guaranteed in our case: a counter-example (Example 5.3.9) is given to show that even under strong assumptions (namely, the regularity of the running and terminal costs), the second-order derivative of the cost functional $\mathcal{J}$ may fail to be sufficiently regular in the infinite-dimensional space considered. To tackle this issue, we show that an appropriate restriction of the problem to essentially bounded processes allows to obtain the desired regularity of the second-order derivative of the cost function to minimize, see Theorem 5.3.11. Local quadratic convergence can thus be expected in this framework [Kan48]-[NW06, Theorem 3.5, p. 44]. However, as in the case of Newton's method in $\mathbb{R}^n$, global convergence may fail in our infinite-dimensional setting, even when the function to minimize is strongly convex with Lipschitz-continuous bounded second-order derivative. We give a counter-example (Example 5.3.12) to show that such pathological behaviors may occur in our setting. To obtain global convergence, a new line-search method tailored to our infinite-dimensional framework is proposed (see Algorithm 5.4) and global convergence results are derived for the Newton method combined with this line-search method (see Algorithm 5.5) under convexity assumptions, see Theorem 5.3.15.

We then apply our results to solve an energy management problem, which consists in a set of many weakly-interacting symmetric batteries controlled to minimize the total operational costs and power imbalance. A Markovian framework is assumed and regression techniques are used to compute efficiently all the conditional expectations required for the computation of the Newton direction. This allows to obtain a fully implementable version of the Newton method with Backtracking line search (see Algorithm 5.8).

Numerical results show the performance of the Newton method which of the proposed Backtracking line-search procedure. On the other hand, we show numerically that the natural extension of the standard Backtracking line search is not adapted to our infinite-dimensional setting: the algorithm takes ridiculously small steps and the gradient norm does not decrease after a few iterations, see Figures 5.6b and 5.6d. The numerical results are consistent with what we expect, with the asymmetric loss function allowing to penalize more heavily positive than negative power imbalance. We then discuss the choice of some hyper-parameters of the regression methods used to solve the BSDEs.

Organization of the chapter. Section 5.2 introduces the general framework of stochastic control problems studied. Classical results under suitable assumptions are derived: well-posedness, existence and uniqueness of a minimizer, Gateaux and Fréchet differentiability, as well as necessary and sufficient optimality conditions 5.2.6. We then prove the second-order differentiability of the problem (Proposition 5.2.8) and show that the second-order differential is valued in the space of isomorphisms of the ambient process space (Corollary 5.3.8). Section 5.3 defines the Newton step and its two equivalent interpretations (Propositions 5.3.2 and5.3.3). We show that the computation of the Newton step amounts to solve an affine-linear FBSDE, for which we show existence and uniqueness of the solutions, and we prove the computation of the Newton step reduces to solving a Backward Riccati Stochatic Differential equation and an affine-linear BSDE (Theorem 5.3.7 and Corollary 5.3.8). We then show Lipschitz-continuity of the second-order derivative of the cost when considering bounded processes. An adapted line-search, called Gradient Backtracking Line Search, as well as Newton's method with line search are given. We prove global convergence for this method as well as quadratic convergence after finitely many iterations (Theorem 5.3.15). Section

5.4 provides a full implementation and the numerical results of the Newton method on the stochastic optimal control problem of a large number of batteries tracking power imbalance. Some proofs are postponed in Section 5.5 to ease the reading.

Notations. We list the most common notations used in all this work.

▷ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers, positive integers. For $n \in \mathbb{N}^*$, $[n]$ denotes the set of integers $\{1, ..., n\}$, and for $m, p \in \mathbb{N}$ with $m \leq p$, $[m : p]$ denotes the set $\{m, ..., p\}$.The notation $|x|$ stands for the Euclidean norm of a vector x, without further reference to its dimension. For a given matrix $A \in \mathbb{R}^p \otimes \mathbb{R}^d$, $A^\top \in \mathbb{R}^d \otimes \mathbb{R}^p$ refers to its transpose. Its norm is that induced by the Euclidean norms in $\mathbb{R}^p$ and $\mathbb{R}^d$, i.e. $|A| := \sup_{x \in \mathbb{R}^d, |x|=1} |Ax|$. Recall that $|A^\top| = |A|$. For $p \in \mathbb{N}^*$, Id_p stands for the identity matrix of size $p \times p$.

▷ *Functions, derivatives.* When a function (or a process) ψ depends on time, we write indifferently $\psi_t(z)$ or $\psi(t, z)$ for the value of ψ at time t, where z represents all other arguments of ψ.

For a smooth function $g : \mathbb{R}^q \mapsto \mathbb{R}^p$, g_x represents the partial derivative of g with respect to x. However, a subscript x_t refers to the value of a process x at time t (and not to a partial derivative with respect to t).

▷ *Probability.* To model the random uncertainty on the time interval $[0, T]$ ($T > 0$ fixed), we consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that the filtration $\mathbb{F} := \{\mathcal{F}_t\}_{0 \leq t \leq T}$ is right-continuous, augmented with the $\mathbb{P}$-null sets. For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$. Denote by $\mathcal{P}$ the σ-field of predictable sets of $[0, T] \times \Omega$.

All the quantities impacted by the control u are upper-indexed by u, like Z^u for instance.

As usually, càdlàg processes stand for processes that are right continuous with left-hand limits. All the martingales are considered with their càdlàg modifications.

▷ *Spaces.* Let $k \in \mathbb{N}^*$. We define $\mathbb{L}^2([0, T], \mathbb{R}^k)$ (resp. $\mathbb{L}^\infty([0, T], \mathbb{R}^k)$) as the Banach space of square integrable (resp. bounded) deterministic functions f on $[0, T]$ with values in $\mathbb{R}^k$. Since the arrival space $\mathbb{R}^k$ will be unimportant, we will skip the reference to it in the notation and write the related norms as

$$\|f\|_{\mathbb{L}^2_T} := \left(\int_0^T |f(t)|^2 \mathrm{d}t \right)^{\frac{1}{2}}, \qquad \|f\|_{\mathbb{L}^\infty_T} := \sup_{t \in [0,T]} |f(t)|.$$

The Banach space of $\mathbb{R}^k$-valued square integrable random variables X is denoted by $\mathbb{L}^2(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^2_\Omega$. We also define the Banach space of $\mathbb{R}^k$-valued essentially bounded random variables X, denoted by $\mathbb{L}^\infty(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^\infty_\Omega$. The associated norms are

$$\|X\|_{\mathbb{L}^2_\Omega} := \mathbb{E}\left[|X|^2\right]^{\frac{1}{2}}; \qquad \|X\|_{\mathbb{L}^\infty_\Omega} := \mathrm{essup}|X| = \inf_{M} \{M \mid \mathbb{P}(|X| \leq M) = 1\}.$$

The Banach space $\mathbb{H}^{2,2}([0, T] \times \Omega, \mathbb{R}^k)$ (resp. $\mathbb{H}^{2,2}_{\mathcal{P}}([0, T] \times \Omega, \mathbb{R}^k)$) is the set of all $\mathbb{F}$-adapted (resp. $\mathbb{F}$-predictable) processes $\psi : [0, T] \times \Omega \to \mathbb{R}^k$ such that $\mathbb{E}\left[\left(\int_0^T |\psi_t|^2 \mathrm{d}t\right)\right] < +\infty$. The Banach space $\mathbb{H}^{\infty,2}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ stands for the elements of $\mathbb{H}^{2,2}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ satisfying $\mathbb{E}\left[\sup_{t \in [0,T]} |\psi_t|^2\right] < +\infty$. The Banach space $\mathbb{H}^{\infty,\infty}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ (resp. $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}\left([0, T] \times \Omega, \mathbb{R}^k\right)$) stands for the space of essentially bounded processes in $\mathbb{H}^{2,2}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ (resp. $\mathbb{H}^{2,2}_{\mathcal{P}}\left([0, T] \times \Omega, \mathbb{R}^k\right)$). Here again we will omit the reference to $\mathbb{R}^k$ and $[0, T] \times \Omega$, which will be clear from the context. The associated norms are:

$$\|\psi\|_{\mathbb{H}^{2,2}} := \mathbb{E}\left[\left(\int_0^T |\psi_t|^2 \mathrm{d}t\right)\right]^{\frac{1}{2}}; \qquad \|\psi\|_{\mathbb{H}^{\infty,2}} := \mathbb{E}\left[\sup_{t \in [0,T]} |\psi_t|^2\right]^{\frac{1}{2}}; \qquad \|\psi\|_{\mathbb{H}^{\infty,\infty}} := \mathrm{essup} \sup_{t \in [0,T]} |\psi_t|.$$

The space of martingales in $\mathbb{H}^{\infty,2}$ is denoted $\mathcal{M}^2$ and the space of martingales vanishing at $t = 0$ is denoted $\mathcal{M}^2_0$.

5.2 Control problem: setting, assumptions and preliminary results

5.2.1 Setting and assumptions

We consider a stochastic control problem where the state dynamic is given by an ordinary differential equation, which is relevant for applications such as control of energy storage/conversion systems. Problems with a more general state dynamics, given by a stochastic differential equation with uncontrolled diffusion or jump terms can also be embedded in this framework, see Remark 5.2.1. In the field of energy management, this can be used to model water dams for instance, where the level of stored water depends on decisions (pumping, ...) and exogenous random processes, like water inflows arising from the rain or the ice melting in the mountains. More generally it allows to model a problem of control of an energy storage system subject to an exogenous random environment, which makes sense in a context of high renewable penetration.

We assume that the coefficients of the control problem are random, without any Markovian assumption and we do not suppose that the filtration is Brownian. We do not consider control nor state constraints. For clarity of the presentation, the results are established in the one dimensional-case, i.e., both the state and control variables are real-valued processes. However, they could be established in a higher dimension setting.

$$\left.\begin{array}{l}\mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t,\omega,u_{t,\omega},X^u_{t,\omega}\right)\mathrm{d}t + \Psi(\omega,X^u_{T,\omega})\right] \\ \text{s.t. } X^u_{t,\omega} = x_0 + \int_0^t \alpha_{s,\omega}u_{s,\omega}\mathrm{d}s.\end{array}\right\} \longrightarrow \min_{u\in\mathbb{H}^{2,2}_{\mathcal{P}}}. \tag{5.2.1}$$

We consider the following regularity assumptions on the problem:

(Reg-1) The function $l : (t,\omega,u,x) \in [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R} \mapsto l(t,\omega,u,x) \in \mathbb{R}$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$-measurable. The function $\Psi : (\omega,x) \in \Omega\times\mathbb{R} \mapsto \psi(\omega,x) \in \mathbb{R}$ is $\mathcal{F}_T\otimes\mathcal{B}(\mathbb{R})$-measurable. Besides, l and Ψ satisfy the growth conditions:

$$\begin{aligned}&|l(t,\omega,u,x)| \le C^{(l)}_{t,\omega} + C\left(|u|^2+|x|^2\right),\\ &|\Psi(\omega,x)| \le C^{(\Psi)}_{\omega} + C|x|^2,\end{aligned}$$

with $C^{(l)} \in \mathbb{H}^{1,1}$, $C^{(\Psi)} \in \mathbb{L}^\infty_T$ and $C > 0$ a deterministic constant. We assume besides $x_0 \in \mathbb{L}^2_\Omega$.

(Reg-2) Assumption **(Reg-1)** holds and the function l is C^1 with respect to (u,x) and Ψ is C^1 with respect to x with derivatives satisfying:

$$\begin{aligned}&|l'_v(t,\omega,u,x)| \le C^{(l')}_{t,\omega} + C'\left(|u|+|x|\right), && v\in\{u,x\},\\ &|\Psi'_x(\omega,x)| \le C^{(\Psi')}_{\omega} + C'|x|,\end{aligned}$$

with $C^{(l')} \in \mathbb{H}^{2,2}$, $C^{(\Psi')} \in \mathbb{L}^2_\Omega$ and $C' > 0$ a deterministic constant.

(Reg-3) Assumptions **(Reg-1)**-**(Reg-2)** hold and the functions l and Ψ are C^1 with respect to (u,x) with Lipschitz continuous derivatives.

(Reg-4) Assumptions **(Reg-1)**-**(Reg-2)**-**(Reg-3)** hold and the functions l and Ψ are C^2 with respect to (u,x) with bounded second derivatives (uniformly in (t,ω,u,x)).

(Reg-5) Assumptions **(Reg-1)**-**(Reg-2)**-**(Reg-3)**-**(Reg-4)** hold and the mappings l and Ψ have lipschitz-continuous second derivatives. Besides, the bounds $C^{(l)}$ and $C^{(l')}$ introduced earlier are in $\mathbb{H}^{\infty,\infty}$, and the constants $C^{(\Psi)}$ and $C^{(\Psi')}$ are in $\mathbb{L}^\infty_\Omega$. We assume besides $x_0 \in \mathbb{L}^\infty_\Omega$.

We introduce the assumption of linearity of the dynamic:

(Lin-Dyn) The dynamic is affine-linear given by $\Phi : (t,\omega,u,x) \mapsto \alpha_{t,\omega}u_{t,\omega}$, with $\alpha \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.

We consider the following convexity assumptions on the problem:

(Conv-1) The mapping l is convex in (u,x) and Ψ is convex in x.

(Conv-2) Assumption **(Conv-1)** holds and the mapping l is μ-strongly convex in u, with $\mu > 0$. In particular, under Assumption **(Reg-4)**, l''_{uu} is uniformly bounded from below by μ.

Remark 5.2.1. *The assumption* **(Lin-Dyn)** *is not restrictive and one could consider general affine-linear dynamics of the form:*

$$X^u_{t,\omega} = x_0 + \int_0^t (\alpha_{s,\omega} u_{s,\omega} + \beta_{s,\omega} X^u_{s,\omega}) \mathrm{d}s + M_{t,\omega},$$

with M an uncontrolled $(\mathcal{F}_t)$-adapted càdlàg uncontrolled process. Without loss of generality, we can assume $M = 0$, as we can reformulate the obtained problem in terms of $\tilde{X}^u = X^u - M$, up to minor modifications of l and Ψ. In the case $M = 0$, we can directly show for general $\beta \in \mathbb{H}^{\infty,\infty}$:

$$X^u_{t,\omega} = \exp\left(\int_0^t \beta_{s,\omega} \mathrm{d}s\right)\left(x_0 + \int_0^t \alpha_{s,\omega} u_{s,\omega} \exp\left(-\int_0^s \beta_{r,\omega} \mathrm{d}r\right) u_{s,\omega} \mathrm{d}s\right)$$

and thus the problem is equivalent to:

$$\left.\begin{array}{l} \mathcal{J}(u) := \mathbb{E}\left[\int_0^T \tilde{l}\left(t, \omega, u_{t,\omega}, \tilde{X}^u_{t,\omega}\right) \mathrm{d}t + \tilde{\Psi}(\tilde{X}^u_{T,\omega})\right] \\ \text{s.t. } \tilde{X}^u_{t,\omega} = x_0 + \int_0^t \tilde{\alpha}_{s,\omega} u_{s,\omega} \mathrm{d}s. \end{array}\right\} \longrightarrow \min_{u \in \mathbb{H}^{2,2}_{\mathcal{P}}}.$$

with:

$$\begin{cases} \tilde{l}(t, \omega, u, x) := l\left(t, \omega, u, \exp\left(\int_0^t \beta_{s,\omega} \mathrm{d}s\right) x\right), \\ \tilde{\alpha}_{t,\omega} := \alpha_{t,\omega} \exp\left(-\int_0^t \beta_{s,\omega} \mathrm{d}s\right), \\ \tilde{\Psi}(\omega, x) = \Psi\left(\omega, \exp\left(\int_0^T \beta_{s,\omega} \mathrm{d}s\right) x\right). \end{cases}$$

5.2.2 Well-posedness, existence and uniqueness of an optimal control

Proposition 5.2.2. *Under Assumption* **(Reg-1)** *and* **(Lin-Dyn)**, *for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, one can define $X^u \in \mathbb{H}^{\infty,2}$ by:*

$$X^u_t = x_0 + \int_0^t \alpha_s u_s \mathrm{d}s. \tag{5.2.2}$$

Besides, we have $\|X^u\|_{\mathbb{H}^{\infty,2}} \leq \sqrt{T}\|\alpha\|_{\mathbb{H}^{\infty,\infty}}\|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega}$, $\|X^u - X^v\|_{\mathbb{H}^{\infty,2}} \leq \sqrt{T}\|\alpha\|_{\mathbb{H}^{\infty,\infty}}\|u - v\|_{\mathbb{H}^{2,2}}$ and $\mathcal{J}(u) < +\infty$.

Proposition 5.2.3. *Under Assumption* **(Reg-2)**-**(Lin-Dyn)**-**(Conv-2)**, *$\mathcal{J}$ is continuous and strongly convex, coercive (i.e., $\lim_{\|u\|_{\mathbb{H}^{2,2}}} \mathcal{J}(u) = +\infty$) and hence $\mathcal{J}$ has a unique minimizer in $\mathbb{H}^{2,2}_{\mathcal{P}}$.*

Proof. The continuity of $u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ holds thanks to **(Lin-Dyn)**, by Lebesgue's continuity theorem. The continuity of $\mathcal{J}$ stems from this fact, Lebesgue's continuity theorem and **(Reg-2)**. Under assumption **(Conv-2)**, $\mathcal{J}$ is μ-strongly convex and coercive. Besides, $\mathbb{H}^{2,2}_{\mathcal{P}}$ is reflexive, as it is a Hilbert space. Therefore, by [Bre10a, Corollary 3.23, pp.71], $\mathcal{J}$ has a unique minimizer $u^* \in \mathbb{H}^{2,2}_{\mathcal{P}}$. □

5.2.3 First-order necessary and sufficient optimality conditions

We first prove first-order differentiability properties of the state variable and the cost function with respect to the control variable under suitable assumptions.

Lemma 5.2.4. *The application $\Phi_X : u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ is Fréchet-differentiable. Besides, for any $(u, v) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2$, the derivative of Φ_X at point u in direction v is independent from u and given by:*

$$\dot{X}^v := \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Phi_X(u + \varepsilon v)\right)_{|\varepsilon=0} = \int_0^t \alpha_s v_s \mathrm{d}s = X^v - x_0. \tag{5.2.3}$$

Besides, we have the following estimate:

$$\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}.$$

Proof. The application $u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ is Gateaux-differentiable at $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ in direction $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ with derivative $\dot{X}^v := X^v - x_0 \in \mathbb{H}^{\infty,2}$. In particular, Φ_X is continuously differentiable and therefore Fréchet-differentiable. The bound $\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}$ arises from Cauchy-Schwarz inequality and the assumption $\alpha \in \mathbb{H}^{\infty,\infty}$. □

Proposition 5.2.5. *Suppose Assumptions* **(Reg-3)** *and* **(Lin-Dyn)** *hold. Then for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, consider $X^u \in \mathbb{H}^{\infty,2}$ given in (5.2.2) and define Y^u by:*

$$Y^u_t = \mathbb{E}_t\left[\Psi'_x(X^u_T) + \int_t^T l'_x(s, u_s, X^u_s)\mathrm{d}s\right]. \tag{5.2.4}$$

Then, Y^u is well-defined and in $\mathbb{H}^{\infty,2}$. Besides, $\mathcal{J}$ is Fréchet-differentiable and admits a gradient at u denoted $\nabla\mathcal{J}(u) \in \mathbb{H}^{2,2}_{\mathcal{P}}$ given by:

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad (\nabla\mathcal{J}(u))_t = l'_u(t, u_t, X^u_t) + \alpha_t Y^u_{t-}. \tag{5.2.5}$$

Besides, we have the following estimates for a deterministic constant C independent of u and v:

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \qquad \|Y^u\|_{\mathbb{H}^{\infty,2}} + \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{2,2}} \leq C(1 + \|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega}),$$
$$\forall (u,v) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2, \qquad \|Y^u - Y^v\|_{\mathbb{H}^{\infty,2}} + \|\nabla\mathcal{J}(u) - \mathcal{J}(v)\|_{\mathbb{H}^{2,2}} \leq C\|u - v\|_{\mathbb{H}^{2,2}}.$$

Proof. The regularity assumptions combined with the estimation on $\|X^u\|_{\mathbb{H}^{\infty,2}}$ directly show that $Y^u \in \mathbb{H}^{\infty,2}$ with:

$$\|Y^u\|_{\mathbb{H}^{\infty,2}} \leq C(1 + \|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega})$$
$$\|Y^u - Y^v\|_{\mathbb{H}^{\infty,2}} \leq C\|u - v\|_{\mathbb{H}^{2,2}}.$$

Admitting first the expression (5.2.5) for $\nabla\mathcal{J}(u)$, we can deduce from the regularity assumptions the bounds claimed.

Let us now prove that $\mathcal{J}$ is Fréchet-differentiable as well as the expression of $\nabla\mathcal{J}(u)$. By Lebesgue's differentiation theorem and Lemma 5.2.4, the application $\mathcal{J}$ is Gateaux-differentiable at u in direction v with derivative given by:

$$\dot{\mathcal{J}}(u,v) := \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(u + \varepsilon v)\right)_{|\varepsilon=0} = \mathbb{E}\left[\Psi'_x(X^u_T)\dot{X}^v_T + \int_0^T \left(l'_x(s, u_s, X^u_s)\dot{X}^v_s + l'_u(s, u_s, X^u_s)v_s\right)\mathrm{d}s\right]. \tag{5.2.6}$$

Define the martingale $M^u \in \mathbb{H}^{\infty,2}$ by $M^u_t := \mathbb{E}_t\left[\psi'_x(X^u_T) + \int_0^T l'_x(s, u_s, X^u_s)\mathrm{d}s\right]$. Then $Y^u_t = M^u_t - \int_0^t l'_x(s, u_s, X^u_s)\mathrm{d}s$ so that (Y^u, M^u) satisfies the following BSDE in $(Y, M) \in \mathbb{H}^{2,2} \times \mathcal{M}^2_0$:

$$\begin{cases} -\mathrm{d}Y_t = l'_x(t, u_t, X^u_t)\mathrm{d}t - \mathrm{d}M_t, \\ Y_T = \Psi'_x(X^u_T). \end{cases}$$

Then, applying Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to the product $Y^u \cdot \dot{X}^v$ between 0 and T yields, using $\dot{X}^v_0 = 0$, $Y^u_T = \Psi'_x(X^u_T)$, the fact that $\dot{X}^v$ is continuous with finite variations and the fact that $\int_{0^+}^t \dot{X}^v_s \mathrm{d}M^u_s$ is a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [Pro03, Theorem 20 p.63, Corollary 3 p.73, Theorem 29 p.75]:

$$\begin{aligned}
\dot{\mathcal{J}}(u,v) &= \mathbb{E}\left[Y^u_T\dot{X}^v_T + \int_0^T \left(l'_x(s, u_s, X^u_s)\dot{X}^v_s + l'_u(s, u_s, X^u_s)v_s\right)\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^T \left(l'_x(s, u_s, X^u_s)\dot{X}^v_s + l'_u(s, u_s, X^u_s)v_s + Y^u_s\alpha_s v_s - \dot{X}^v_s l'_x(s, u_s, X^u_s)\right)\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^T \left(\alpha_s Y^u_s + l'_u(s, u_s, X^u_s)\right)v_s\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^T \left(\alpha_s Y^u_{s-} + l'_u(s, u_s, X^u_s)\right)v_s\mathrm{d}s\right].
\end{aligned}$$

In the last inequality, we used the fact that Y^u has countably many jumps, and that the Lebesgue integral is left unchanged by changing the integrand on a countable set of points. This yields the expression of $\nabla\mathcal{J}(u)$. In particular, the previous estimates imply that $\mathcal{J}$ has a Lipschitz continuous gradient and is therefore Fréchet-differentiable.

□

Theorem 5.2.6 (First order necessary and sufficient optimality conditions). *1. Suppose Assumptions* **(Reg-3)** *and* **(Lin-Dyn)** *hold. Assume $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ is a minimizer of $\mathcal{J}$. Define $X^u \in \mathbb{H}^{\infty,2}$ by (5.2.2) and $Y^u \in \mathbb{H}^{\infty,2}$ by (5.2.4). Then, necessarily,*

$$l'_u(t, u_t, X^u_t) + \alpha_t Y^u_{t-} = 0, \quad \mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e. \tag{5.2.7}$$

2. Under Assumptions **(Reg-3)**, **(Conv-1)** *and* **(Lin-Dyn)**, *if $(u, X^u, Y^u) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ satisfies (5.2.7) with X^u given by (5.2.2) and Y^u given by (5.2.4), then u is a solution of (5.2.1), i.e., a minimizer of $\mathcal{J}$.*

Proof. 1. Under **(Reg-3)**, $\mathcal{J}$ is Gateaux-differentiable and an optimal control is necessary a critical point of $\mathcal{J}$, hence $\nabla\mathcal{J}(u) = 0$, which yields (5.2.7).

2. Under **(Conv-1)** and **(Lin-Dyn)**, $\mathcal{J}$ is convex and $\mathcal{J}$ is Gateaux-differentiable under **(Reg-3)**, so that (5.2.7) is a sufficient optimality condition.

□

5.2.4 Second-order differentiability

We now turn to second-order differentiability of the cost functional, necessary for the Newton method. We then prove a key result showing the invertibility of the second order derive of $\mathcal{J}$, and the form of the inverse. This shows the existence and provides a characterization of the Newton step.

Lemma 5.2.7. *Suppose Assumptions* **(Reg-4)** *and* **(Lin-Dyn)** *hold. Then the mapping*

$$\Phi_Y : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{H}^{\infty,2} \\ u \mapsto Y^u \end{cases}$$

is Gateaux-differentiable. Furthermore, for all u, v in $\mathbb{H}^{2,2}_{\mathcal{P}}$, $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$ is defined by the following affine-linear BSDE with Lipschitz coefficients:

$$\dot{Y}^{u,v}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\dot{X}^v_T + \int_t^T \left(l''_{xu}(s, u_s, X^u_s)v_s + l''_{xx}(s, u_s, X^u_s)\dot{X}^v_s\right)\mathrm{d}s\right]. \tag{5.2.8}$$

Besides, we have the following estimate:

$$\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}.$$

Proof. By our assumptions, $\dot{Y}^{u,v}$ is well defined and by Lebesgue's differentiation theorem and Lemma 5.2.4, it is straightforward that Φ_Y is Gateau-differentiable and that: $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$. Applying Lebesgue's theorem and using our estimation on $\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}}$, one gets $\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}$. □

Proposition 5.2.8 (Second-order differentiability). *Suppose Assumptions* **(Reg-4)** *and* **(Lin-Dyn)** *hold. Then the mapping $\mathcal{J}$ is twice Gateaux differentiable and its second-order derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ is given by:*

$$\forall v \in \mathbb{H}^{2,2}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad \left(\nabla^2\mathcal{J}(u)(v)\right)_t = l''_{uu}(t, u_t, X^u_t)v_t + l''_{ux}(t, u_t, X^u_t)\dot{X}^v_t + \alpha_t \dot{Y}^{u,v}_{t-}. \tag{5.2.9}$$

Besides, we have the following estimate:

$$\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{2,2}} \leq C\|v\|_{\mathbb{H}^{2,2}},$$

where C is a constant independent of u. In other words, for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, $\nabla^2\mathcal{J}(u)$ is a continuous endomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$.

Proof. Applying Lebesgue's differentiation theorem to $\nabla\mathcal{J}(u)$ given by (5.2.5) yields (5.2.9), using Lemmas 5.2.4 and 5.2.7. The continuity of $\nabla^2\mathcal{J}(u)$ for all $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ results from the previous estimates, our assumptions and Lebesgue's differentiation theorem. □

The computation of the Newton step Δu amounts to solve the equation $\nabla^2\mathcal{J}(u)(\Delta u) = -\nabla\mathcal{J}(u)$, as expected by an infinite-dimensional generalization of the Newton method in $\mathbb{R}^d$. The following theorem is the key result which guarantees that this infinite dimensional equation has a unique solution, as we show invertibility of the second order derivative of the cost function at any admissible point. This theorem also makes the connection between the computation of the Newton step and the solution of an auxiliary Linear-Quadratic stochastic control problem, or equivalently, with the solution of an affine-linear FBSDE with random coefficients.

Theorem 5.2.9. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let* $(u, w) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{2,2}_{\mathcal{P}}$ *and define* $X^u \in \mathbb{H}^{\infty,2}$ *by* (5.2.2)*. We introduce the following auxiliary (linear-quadratic) stochastic control problem:*

$$\begin{aligned} &\min_{v\in\mathbb{H}^{2,2}_{\mathcal{P}}} \tilde{\mathcal{J}}^{quad,u,w}(v) \\ &\text{s.t. } \tilde{X}_t = \int_0^t \alpha_s v_s \mathrm{d}s. \end{aligned} \tag{5.2.10}$$

where $\tilde{\mathcal{J}}^{quad,u,w}(v)$ *is defined by:*

$$\mathbb{E}\left[\int_0^T \left\{\frac{1}{2} l''_{uu}\left(t, u_t, X^u_t\right) v_t^2 + \frac{1}{2} l''_{xx}\left(t, u_t, X^u_t\right) \tilde{X}_t^2 + l''_{ux}\left(t, u_t, X^u_t\right) \tilde{X}_t v_t - w_t v_t\right\} \mathrm{d}t + \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{X}_T^2\right].$$

Then $\tilde{\mathcal{J}}^{quad,u,w}$ *has a unique minimizer* $\tilde{u}^{u,w} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *defined by*

$$l''_{uu}(t, u_t, X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t, u_t, X^u_t)\tilde{X}^{u,w}_t + \alpha_t \tilde{Y}^{u,w}_{t-} = w_t,$$

where $(\tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ *are given by:*

$$\begin{cases} \tilde{X}^{u,w}_t = \int_0^t \alpha_s \tilde{u}^{u,w}_s \mathrm{d}s, \\ \tilde{Y}^{u,w}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\tilde{X}^{u,w}_T + \int_t^T \left(l''_{xu}(s, u_s, X^u_s)\tilde{u}^{u,w}_s + l''_{xx}(s, u_s, X^u_s)\tilde{X}^{u,w}_s\right)\mathrm{d}s\right]. \end{cases}$$

In particular, $(\tilde{u}^{u,w}, \tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ *is the unique solution of an affine-linear FBSDE with random coefficients (which depend on the stochastic process* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$*). Besides, for any* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$*,* $\nabla^2\mathcal{J}(u) \in \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ *is invertible and for any* $w \in \mathbb{H}^{2,2}_{\mathcal{P}}$*,* $\left(\nabla^2\mathcal{J}(u)\right)^{-1}(w) = \tilde{u}^{u,w}$*.*

Proof. Introduce the auxiliary running cost function $\tilde{l}^{u,w} : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ defined by $\tilde{l}^{u,w}(t, \omega, \tilde{u}, \tilde{x}) = \frac{1}{2} l''_{uu}\left(t, u_t, X^u_t\right)\tilde{u}_t^2 + \frac{1}{2} l''_{xx}\left(t, u_t, X^u_t\right)\tilde{x}_t^2 + l''_{ux}\left(t, u_t, X^u_t\right)\tilde{u}_t\tilde{x}_t - w_t\tilde{u}_t$, where we dropped the reference to ω for simplicity. Introduce as well the auxiliary terminal cost function $\tilde{\Psi}^{u,w} : \Omega \times \mathbb{R} \mapsto \mathbb{R}$ defined by $\tilde{\Psi}^{u,w}(\omega, \tilde{x}) = \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{x}^2$. Then Assumptions **(Reg-4)**, **(Lin-Dyn)** and **(Conv-2)** are verified for the auxiliary minimization problem of $\tilde{\mathcal{J}}^{quad,u,w}$, with l and Ψ respectively replaced by $\tilde{l}^{u,w}$ and $\tilde{\Psi}^{u,w}$. Applying Proposition 5.2.3 to the auxiliary problem shows the existence and uniqueness of a minimizer, denoted $\tilde{u}^{u,w}$. Applying Theorem 5.2.6 to the auxiliary problem, we have existence and uniqueness of $(\tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ such that $(\tilde{u}^{u,w}, \tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ is the (unique) solution of the FBSDE:

$$\begin{cases} \tilde{X}^{u,w}_t = \int_0^t \alpha_s \tilde{u}^{u,w}_s \mathrm{d}s, \\ \tilde{Y}^{u,w}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\tilde{X}^{u,w}_T + \int_t^T \left(l''_{xu}(s, u_s, X^u_s)\tilde{u}^{u,w}_s + l''_{xx}(s, u_s, X^u_s)\tilde{X}^{u,w}_s\right)\mathrm{d}s\right], \\ l''_{uu}(t, u_t, X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t, u_t, X^u_t)\tilde{X}^{u,w}_t + \alpha_t \tilde{Y}^{u,w}_{t-} = w_t. \end{cases}$$

In particular, one has $\tilde{X}^{u,w} = \dot{X}^{\tilde{u}^{u,w}}$ and $\tilde{Y}^{u,w} = \dot{Y}^{\tilde{u}^{u,w}}$. The last equation therefore writes $l''_{uu}(t, u_t, X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t, u_t, X^u_t)\dot{X}^{\tilde{u}^{u,w}}_t + \alpha_t \dot{Y}^{\tilde{u}^{u,w}}_{t-} = w_t$. Recognizing $\nabla^2\mathcal{J}(u)(\tilde{u}^{u,w})$ in the left-hand side (see (5.2.9)), we get in particular the existence and uniqueness of $\tilde{u}^{u,w} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ solution of the equation $\nabla^2\mathcal{J}(u)(\tilde{u}^{u,w}) = w$. This holds for all $w \in \mathbb{H}^{2,2}_{\mathcal{P}}$. We thus obtain the invertibility of $\nabla^2\mathcal{J}(u)$ and the expression $(\nabla^2\mathcal{J}(u))^{-1}(w) = \tilde{u}^{u,w}$, for all $u, w \in \mathbb{H}^{2,2}_{\mathcal{P}}$. □

5.3 The Newton method for stochastic control problems

5.3.1 Definition and interpretation of the Newton step

Definition 5.3.1. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *and define* $X^u \in \mathbb{H}^{\infty,2}$ *by* (5.2.2)*. The Newton step* Δ_u *of* $\mathcal{J}$ *at the point* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *is defined by* $\Delta_u = -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u)) \in \mathbb{H}^{2,2}_{\mathcal{P}}$*.*

The following Proposition shows that computation of the Newton step at point u, $-(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u))$ amounts to solve a Linear-Quadratic approximation of the original problem around the current point u, based on a point-wise second order expansion of the cost and a first order expansion of the dynamic.

Proposition 5.3.2. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define $X^u \in \mathbb{H}^{\infty,2}$ by (5.2.2). Denote by $\theta^u_t := (t, u_t, X^u_t)$. The Newton step $\Delta_u = -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u)) \in \mathbb{H}^{2,2}_{\mathcal{P}}$ of $\mathcal{J}$ at the point u is the unique minimizer in $\mathbb{H}^{2,2}_{\mathcal{P}}$ of the Linear-Quadratic approximation $\mathcal{J}^{LQ,u}$ of $\mathcal{J}$ around u, defined by:*

$$\forall v \in \mathbb{H}^{2,2}_{\mathcal{P}}, \quad \mathcal{J}^{LQ,u}(v) := \mathbb{E}\left[\int_0^T \left\{\frac{1}{2}l''_{uu}(\theta^u_t)v_t^2 + \frac{1}{2}l''_{xx}(\theta^u_t)(\dot{X}^v_t)^2 + l''_{ux}(\theta^u_t)\dot{X}^v_t v_t + l'_u(\theta^u_t)v_t + l'_x(\theta^u_t)\dot{X}^v_t + l(\theta^u_t)\right\}\mathrm{d}t\right] + \mathbb{E}\left[\frac{1}{2}\Psi''_{xx}(X^u_T)(\dot{X}^v_T)^2 + \Psi'_x(X^u_T)\dot{X}^v_T + \Psi(X^u_T)\right],$$

where $\dot{X}^v_t = \int_0^t \alpha_s v_s \mathrm{d}s$.

Proof. Introduce the auxiliary running cost function $l^{LQ,u} : [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$ defined by $l^{LQ,u}(t,\tilde{u},\tilde{x}) = \frac{1}{2}l''_{uu}\left(\theta^u_t\right)\tilde{u}_t^2 + \frac{1}{2}l''_{xx}\left(\theta^u_t\right)\tilde{x}_t^2 + l''_{ux}\left(\theta^u_t\right)\tilde{u}\tilde{x} + l'_u\left(\theta^u_t\right)\tilde{u} + l'_x\left(\theta^u_t\right)\tilde{x} + l\left(\theta^u_t\right)$, where we dropped the reference to ω for simplicity. Introduce as well the auxiliary terminal cost function $\Psi^{LQ,u} : \Omega\times\mathbb{R} \mapsto \mathbb{R}$ defined by $\Psi^{LQ,u}(\omega,\tilde{x}) = \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{x}^2 + \Psi'_x(X^u_T)\tilde{x} + \Psi(X^u_T)$. Then Assumptions **(Reg-4)**, **(Lin-Dyn)** and **(Conv-2)** are verified for the auxiliary minimization problem of $\mathcal{J}^{LQ,u}$, with l and Ψ respectively replaced by $l^{LQ,u}$ and $\Psi^{LQ,u}$. Applying Proposition 5.2.3 to the auxiliary problem shows the existence and uniqueness of a minimizer of $\mathcal{J}^{LQ,u}$, denoted by $\hat{u}$. We can then apply Theorem 5.2.6 to the auxiliary problem and get existence and uniqueness of $(\hat{X},\hat{Y}) \in \mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}$ such that $(\hat{u},\hat{X},\hat{Y}) \in \mathbb{H}^{2,2}_{\mathcal{P}}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}$ is the (unique) triple satisfying:

$$\begin{cases} \hat{X}_t = \int_0^t \alpha_s\hat{u}_s\mathrm{d}s, \\ \hat{Y}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\hat{X}_T + \Psi'_x(X^u_T) + \int_t^T \left(l''_{xu}(s,u_s,X^u_s)\hat{u}_s + l''_{xx}(s,u_s,X^u_s)\hat{X}_s + l'_x(s,u_s,X^u_s)\right)\mathrm{d}s\right], \\ l''_{uu}(t,u_t,X^u_t)\hat{u}_t + l''_{ux}(t,u_t,X^u_t)\hat{X}_t + l'_u(t,u_t,X^u_t) + \alpha_t\hat{Y}_{t-} = 0. \end{cases}$$

In particular, we have $\hat{X} = \dot{X}^{\hat{u}}$ by (5.2.3), $\hat{Y} = Y^u + \dot{Y}^{u,\hat{u}}$ by (5.2.4) and(5.2.8). Besides, the last equation is equivalent to $\nabla\mathcal{J}(u) + \nabla^2\mathcal{J}(u)(\hat{u}) = 0$ by (5.2.5) and (5.2.9). In particular, the minimizer $\hat{u}$ of $\mathcal{J}^{LQ,u}$ is nothing else than the Newton step of $\mathcal{J}$ at point u. □

Without surprises, solving such a linear-quadratic stochastic control problem is equivalent to solving an affine-linear FBSDE. This gives a second interpretation of the Newton step as solution of the linearized first order optimality conditions of the control problem, summed up in the following Proposition.

Proposition 5.3.3. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define $X^u \in \mathbb{H}^{\infty,2}$ by (5.2.2). Let $\Delta_u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ be the Newton step of $\mathcal{J}$ at the point $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$. Define $(X^u, Y^u, \dot{X}^{\Delta_u}, \dot{Y}^{\Delta_u}) \in \mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}$ by:*

$$\begin{cases} X^u_t = x_0 + \int_0^t \alpha_s u_s\mathrm{d}s, \\ Y^u_t = \mathbb{E}_t\left[\Psi'_x(X^u_T) + \int_t^T l'_x(s,u_s,X^u_s)\mathrm{d}s\right], \\ \dot{X}^{\Delta_u}_t = \int_0^t \alpha_s(\Delta_u)_s\mathrm{d}s, \\ \dot{Y}^{\Delta_u}_t = \mathbb{E}_t\left[\Psi''_x(X^u_T)\dot{X}^{\Delta_u}_T + \int_t^T \left(l''_{xx}(s,u_s,X^u_s)\dot{X}^{\Delta_u}_s + l''_{xu}(s,u_s,X^u_s)(\Delta_u)_s\right)\mathrm{d}s\right]. \end{cases}$$

Then $X^u + \dot{X}^{\Delta_u}$ is the first-order approximation of $v \mapsto X^{u+v}$ evaluated at $v = \Delta_u$, $Y^u + \dot{Y}^{\Delta_u}$ is the first-order approximation of $v \mapsto Y^{u+v}$ evaluated at $v = \Delta_u$ and Δ_u is the zero of the linearized gradient of $\mathcal{J}$ around u, i.e., $\nabla\mathcal{J}(u) + \nabla^2\mathcal{J}(u)(\Delta_u) = 0$.

Proof. This is a direct consequence of Lemmas 5.2.4 and 5.2.7, as well as the definition of the Newton step 5.3.1. □

Remark 5.3.4. *In Brownian filtrations, one can derive similar results with more general dynamics of the state variable. Consider the following dynamic (controlled diffusion):*

$$X_t = x_0 + \int_0^t (\alpha_s u_s + \beta_s X_s + \gamma_s)\mathrm{d}s + \int_0^t (\bar{\alpha}_s u_s + \bar{\beta}_s X_s + \bar{\gamma}_s)\mathrm{d}W_s.$$

We make the same assumptions on the cost functional. One can apply Pontryagin principle to show necessary and sufficient optimality condition of order 1. The computation of the second-order derivative can be adapted to this particular case as well. Then, one can show that the Newton step can be defined in this setting, and that its computation amounts to solve an affine-linear FBSDE with stochastic coefficients. In this case, as the filtration is Brownian, the affine-linear FBSDE is computable using directly the results in [Yon06].

5.3.2 Solution of affine-linear FBSDEs and computation of the Newton step

Computing the Newton step is equivalent to compute the inverse of $(\nabla^2\mathcal{J}(u))^{-1}(w)$ with $w = -\nabla\mathcal{J}(u)$. By Theorem 5.2.9, computing this quantity is equivalent to solving the following affine-linear FBSDE:

$$\begin{cases} \tilde{X}_t^{u,w} = \displaystyle\int_0^t \left\{ -\frac{\alpha_s l''_{ux}(s,u_s,X_s^u)}{l''_{uu}(s,u_s,X_s^u)}\tilde{X}_s^{u,w} - \frac{\alpha_s^2}{l''_{uu}(s,u_s,X_s^u)}\tilde{Y}_s^{u,w} + \frac{\alpha_s w_s}{l''_{uu}(s,u_s,X_s^u)} \right\}\mathrm{d}s, \\ \tilde{Y}_t^{u,w} = \mathbb{E}_t\left[\Psi''_{xx}(X_T^u)\tilde{X}_T^{u,w} + \displaystyle\int_t^T \left(\frac{l''_{uu}(s,u_s,X_s^u)l''_{xx}(s,u_s,X_s^u) - (l''_{ux}(s,u_s,X_s^u))^2}{l''_{uu}(s,u_s,X_s^u)}\tilde{X}_s^{u,w} - \frac{l''_{xu}(s,u_s,X_s^u)\alpha_s}{l''_{uu}(s,u_s,X_s^u)}\tilde{Y}_s^{u,w} + \frac{l''_{xu}(s,u_s,X_s^u)w_s}{l''_{uu}(s,u_s,X_s^u)} \right)\mathrm{d}s \right], \end{cases} \tag{5.3.1}$$

and then $(\nabla^2\mathcal{J}(u))^{-1}(w)$ is given by:

$$\left((\nabla^2\mathcal{J}(u))^{-1}(w)\right)_t = -\frac{1}{l''_{uu}(t,u_t,X_t^u)}\left(l''_{ux}(t,u_t,X_t^u)\tilde{X}_t^{u,w} + \alpha_t\tilde{Y}_{t-}^{u,w} - w_t\right). \tag{5.3.2}$$

Note that $(\tilde{X}^{u,w}, \tilde{Y}^{u,w})$ are solution of an affine-linear FBSDE with stochastic coefficients which has the following structure:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right] \end{cases} \tag{5.3.3}$$

with $C \geq 0$ (convexity of l with respect to (u,x) and strong convexity of l with respect to u), $\Gamma \geq 0$, $B \leq 0$. A, B and C are in $\mathbb{H}^{\infty,\infty}$, a and b are in $\mathbb{H}^{2,2}$, $\Gamma \in \mathbb{L}^\infty(\mathcal{F}_T)$, $\eta \in \mathbb{L}^2(\mathcal{F}_T)$ and $x \in \mathbb{L}^2(\mathcal{F}_0)$.

Affine-linear FBSDEs have been studied in the literature, and a solution method is based on a so-called decoupling field, assumed affine-linear, by the structure of the equation. The assumption is that the solution verifies $Y = PX + \Pi$ for some P and Π to determine. Standard results in the literature [Yon06] show that P solves a matrix Riccati BSDE and Π solves an affine-linear BSDE. However, we are outside of the scope of [Yon06], which assumes a Brownian filtration. The following Lemma gives some results on solutions of Riccati BSDEs in general filtrations.

Lemma 5.3.5 (One-dimensional Riccati-BSDE under general filtrations)**.** *Let A, B, C be processes in $\mathbb{H}^{\infty,\infty}$ and $\Gamma \in \mathbb{L}^\infty(\mathcal{F}_T)$. Suppose additionally that $\Gamma \geq 0$ $\mathrm{d}\mathbb{P}$-a.e., $B_t \leq 0$ $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e., $C_t \geq 0$ $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e. Then the following Riccati BSDE with unknown P and stochastic coefficients:*

$$P_t = \mathbb{E}_t\left[\Gamma + \int_t^T \left(2A_s P_s + B_s P_s^2 + C_s\right)\mathrm{d}s\right] \tag{5.3.4}$$

has a unique solution in $\mathbb{H}^{\infty,\infty}([0,T])$ and we have the estimation $0 \leq P_t \leq \overline{P}_t$ $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e. with:

$$\overline{P}_t = \mathbb{E}_t\left[\Gamma\exp\left(2\|A\|_{\mathbb{H}^{\infty,\infty}}(T-t)\right) + \int_t^T C_s\exp\left(2\|A\|_{\mathbb{H}^{\infty,\infty}}(s-t)\right)\mathrm{d}s\right]. \tag{5.3.5}$$

A general result of Bismut for existence and uniqueness of the solution of Riccati BSDE can be found in [Bis76, Theorem 6.1]. In section 5.5.1, we provide our own proof in the one-dimensional case. It is based on the comparison principle for BSDEs, and allows to prove that the solution of the Riccati BSDE coincides with the solution of a BSDE with a truncated drift, globally Lipschitz continuous. As a limitation, the comparison principle applies only for one-dimensional BSDEs. Therefore, our proof cannot be expected to be generalized to higher dimension, except if an analogous comparison principle for BSDEs with square symmetric matrix unknown is developed, using the order defined by the cone of positive semi-definite matrices.

We now give a result on the second ingredient allowing to solve coupled linear FBSDEs.

Lemma 5.3.6. *[1-dimensional affine-linear BSDE in general filtrations] Let A, B, C and P be as before. Suppose $a, b \in \mathbb{H}^{2,2}$ and $\eta \in \mathbb{L}^2(\mathcal{F}_T)$. Define $\Pi \in \mathbb{H}^{\infty,2}$ by:*

$$\Pi_t = \mathbb{E}_t\left[\eta \exp\left(\int_t^T (P_s B_s + A_s)\mathrm{d}s\right) + \int_t^T (a_s P_s + b_s)\exp\left(\int_t^s (P_r B_r + A_r)\mathrm{d}r\right)\mathrm{d}s\right]. \tag{5.3.6}$$

Then Π is the unique solution in $\mathbb{H}^{\infty,2}$ of the BSDE:

$$\Pi_t = \mathbb{E}_t\left[\eta + \int_t^T ((P_s B_s + A_s)\Pi_s + a_s P_s + b_s)\,\mathrm{d}s\right]. \tag{5.3.7}$$

Additionally, we have the estimation:

$$\|\Pi\|_{\mathbb{H}^{\infty,2}} \leq \left(\|\eta\|_{\mathbb{L}^2} + \sqrt{T}\|a\|_{\mathbb{H}^{2,2}}\|P\|_{\mathbb{H}^{\infty,\infty}} + \sqrt{T}\|b\|_{\mathbb{H}^{2,2}}\right) e^{\|PB+A\|_{\mathbb{H}^{\infty,\infty}}T}.$$

The proof is given in Section 5.5.2. We are now in position of deriving a verification Theorem for the solution of affine-linear FBSDEs, based on the two previous Lemmas.

Theorem 5.3.7. *[Scalar affine-linear FBSDEs with exogenous noise under general filtrations] Let A, B, C be processes in $\mathbb{H}^{\infty,\infty}$ and $\Gamma \in \mathbb{L}^\infty(\mathcal{F}_T)$. Let $a, b \in \mathbb{H}^{2,2}$, $x \in \mathbb{L}^2(\mathcal{F}_0)$ and $\eta \in \mathbb{L}^2(\mathcal{F}_T)$. Suppose additionally that $\Gamma \geq 0$ $\mathrm{d}\mathbb{P}$-a.e., $B_t \leq 0$ $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e., $C_t \geq 0$ $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e. Then:*

1. *The following FBSDE has a unique solution $(X, Y) \in (\mathbb{H}^{\infty,2})^2$:*

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right]. \end{cases}$$

2. *Define $P \in \mathbb{H}^{\infty,\infty}$ by (5.3.4) and $\Pi \in \mathbb{H}^{\infty,2}$ by (5.3.6). Define as well X by:*

$$X_t = x + \int_0^t ((A_s + B_s P_s)X_s + B_s \Pi_s + a_s)\,\mathrm{d}s. \tag{5.3.8}$$

and

$$Y = PX + \Pi. \tag{5.3.9}$$

The processes X and Y are well-defined and in $\mathbb{H}^{\infty,2}$ with the estimates:

$$\|X\|_{\mathbb{H}^{\infty,2}} \leq \left(\|x\|_{\mathbb{L}^2_\Omega} + T\|B\|_{\mathbb{H}^{\infty,\infty}}\|\Pi\|_{\mathbb{H}^{\infty,2}} + \sqrt{T}\|a\|_{\mathbb{H}^{2,2}}\right) e^{\|A+PB\|_{\mathbb{H}^{\infty,\infty}}T},$$

and

$$\|Y\|_{\mathbb{H}^{\infty,2}} \leq \|P\|_{\mathbb{H}^{\infty,\infty}}\|X\|_{\mathbb{H}^{\infty,2}} + \|\Pi\|_{\mathbb{H}^{\infty,2}}.$$

Besides, $(X, Y) \in (\mathbb{H}^{\infty,2})^2$ is the unique solution of the FBSDE:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right]. \end{cases}$$

The proof of this Theorem is postponed to Section 5.5.3. Applying the above result to (5.3.1) and using (5.3.2), we get the following Corollary.

Corollary 5.3.8 (Explicit computation of the inverse of $\nabla^2\mathcal{J}(u)$)**.** *Suppose assumptions* **(Reg-4)**, **(Conv-2)** *and* **(Lin-Dyn)** *hold. Let $u, w \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define X^u and Y^u in $\mathbb{H}^{\infty,2}$ as in (5.2.2) and (5.2.4). Define as well:*

$$\begin{cases} A^u_t = -\dfrac{\alpha_t l''_{ux}(t,u_t,X^u_t)}{l''_{uu}(t,u_t,X^u_t)}, \\ B^u_t = -\dfrac{\alpha_t^2}{l''_{uu}(t,u_t,X^u_t)}, \\ C^u_t = \dfrac{l''_{uu}(t,u_t,X^u_t)l''_{xx}(t,u_t,X^u_t) - (l''_{ux}(t,u_t,X^u_t))^2}{l''_{uu}(t,u_t,X^u_t)}, \\ \Gamma^u = \Psi''_{xx}(X^u_T), \\ a^{u,w}_t = \dfrac{\alpha_t w_t}{l''_{uu}(t,u_t,X^u_t)}, \\ b^{u,w}_t = \dfrac{l''_{ux}(t,u_t,X^u_T)w_t}{l''_{uu}(t,u_t,X^u_t)}. \end{cases} \tag{5.3.10}$$

Then the following Riccati BSDE with unknown P:

$$P_t = \mathbb{E}_t\left[\Gamma^u + \int_t^T \left(2A^u_s P_s + B^u_s P_s^2 + C^u_s\right)\mathrm{d}s\right] \tag{5.3.11}$$

has a unique solution in $\mathbb{H}^{\infty,\infty}$, denoted P^u and the following affine-linear BSDE with unknown Π:

$$\Pi_t = \mathbb{E}_t\left[\int_t^T ((P^u_s B^u_s + A^u_s)\Pi_s + a^{u,w}_s P^u_s + b^{u,w}_s)\,\mathrm{d}s\right]. \tag{5.3.12}$$

has a unique solution in $\mathbb{H}^{\infty,2}$, which is denoted $\Pi^{u,w}$. Define $\tilde{X}^{u,w} \in \mathbb{H}^{\infty,2}$ as the unique solution of the following ordinary differential equation:

$$\tilde{X}_t = \int_0^t \left((A^u_s + B^u_s P^u_s)\tilde{X}_s + B^u_s \Pi^{u,w}_s + a^{u,w}_s\right)\mathrm{d}s. \tag{5.3.13}$$

Then:

$$\left((\nabla^2\mathcal{J}(u))^{-1}(w)\right)_t = -\frac{1}{l''_{uu}(t,u_t,X^u_t)}\left(\{l''_{ux}(t,u_t,X^u_t) + \alpha_t P^u_{t-}\}\,\tilde{X}^{u,w}_t + \alpha_t \Pi^{u,w}_{t-} - w_t\right). \tag{5.3.14}$$

Besides, we have:

$$\begin{aligned} &\|P^u\|_{\mathbb{H}^{\infty,\infty}} \le C, \\ &\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}} \le C\|w\|_{\mathbb{H}^{2,2}}, \\ &\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}} \le C\|w\|_{\mathbb{H}^{2,2}}, \end{aligned}$$

for some constant C independent of u and w. In particular, $\nabla^2\mathcal{J}(u)$ is a bi-continuous isomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$ for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$.

The proof of this Corollary is postponed to Section 5.5.4.

5.3.3 Global convergence of Newton's method with an adapted line-search method

For linear-quadratic problems with random coefficient, the Newton direction is equal to the optimal solution of the problem minus the current point, so that the Newton method converges in one iteration.

In the finite-dimensional case, the local quadratic convergence of Newton's method typically requires the function to minimize to have a Lipschitz continuous second-order derivative, see [Kan48]-[NW06, Theorem 3.5, p. 44], or

to be self-concordant [NN94]. As self-concordance is not a notion well-defined in our setting, we focus on the first assumption.

We provide next a -example of $\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{R}$ to show that, even under assumption **(Reg-5)**, $\mathcal{J}$ may have a second-order derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2} \mapsto \mathcal{L}(\mathbb{H}^{2,2})$ which is not Lipschitz-continuous. This shows that the (local) convergence of the Newton method should be established in a different space.

Example 5.3.9. *Let us assume $T = 1$ and consider $\mathcal{J}$ given by:*

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \quad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 l(u_t)\mathrm{d}t\right], \tag{5.3.15}$$

$$s.t. \quad X^u_t = 0, \forall t \in [0,1]. \tag{5.3.16}$$

where l is deterministic twice continuously differentiable with derivative given by the oscillating function represented in Figure 5.1, which is Lipschitz-continuous, bounded and non-negative. In particular, Assumptions **(Reg-5)**, **(Lin-**

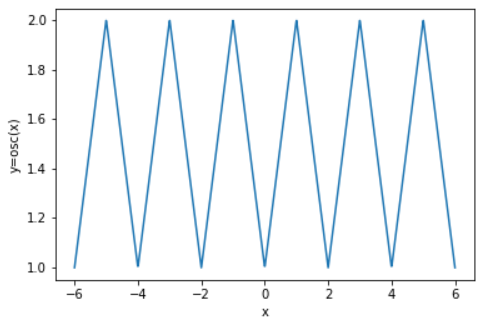


Figure 5.1: Oscillating function l''

Dyn) *and* **(Conv-2)**, *therefore, $\mathcal{J}$ is 1-strongly-convex, twice continuously differentiable, with second order-derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ given by:*

$$\forall (u,v) \in \mathbb{H}^{2,2}_{\mathcal{P}}, \forall t \in [0,1], \quad (\nabla^2\mathcal{J}(u)(v))_t = l''(u_t)v_t.$$

This second-order derivative is bounded, and for all $u \in \mathbb{H}^{2,2}$, $\nabla^2\mathcal{J}(u) : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{H}^{2,2}_{\mathcal{P}}$ is a bi-continuous endomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$, i.e., $\nabla^2\mathcal{J}(u)$ is a continuous invertible endomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$, and its inverse is bounded as well. We have:

$$\forall (u,w) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2, \forall t \in [0,1], \quad ((\nabla^2\mathcal{J}(u))^{-1}(w))_t = \frac{w_t}{l''(u_t)}.$$

In particular, Assumptions **(Reg-4)**, **(Lin-Dyn)** *and* **(Conv-2)** *are verified. However, for $n \in \mathbb{N}_*$, let us define $u^{(n)} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ the constant process with value $U^{(n)}$ given by a Bernoulli random variable with parameter $p^{(n)} = 1/n$, i.e., $U^{(n)} = 1$ with probability $1/n$ and 0 else. Let $v = 0 \in \mathbb{H}^{2,2}_{\mathcal{P}}$. Then we have, using $l''(1) = 1$ and $l''(0) = 0$, $\mathbb{E}\left[(U^{(n)})^2\right] = \mathbb{E}\left[(U^{(n)})^4\right] = \frac{1}{n}$:*

$$\begin{aligned}
\frac{\|\nabla^2\mathcal{J}(u^{(n)}) - \nabla^2\mathcal{J}(v)\|_{\mathcal{L}(\mathbb{H}^{2,2})}}{\|u^{(n)} - v\|_{\mathbb{H}^{2,2}}} &\geq \frac{\|\nabla^2\mathcal{J}(u^{(n)})(u^{(n)}) - \nabla^2\mathcal{J}(v)(u^{(n)})\|_{\mathbb{H}^{2,2}}}{\|u^{(n)}\|_{\mathbb{H}^{2,2}}\|u^{(n)} - v\|_{\mathbb{H}^{2,2}}} \\
&= \frac{\mathbb{E}\left[(l''(U^{(n)}) - l''(0))^2(U^{(n)})^2\right]^{1/2}}{\mathbb{E}\left[(U^{(n)})^2\right]} \\
&= \frac{\mathbb{E}\left[(U^{(n)})^4\right]^{1/2}}{\mathbb{E}\left[(U^{(n)})^2\right]} \\
&= \sqrt{n} \underset{n \longrightarrow +\infty}{\longrightarrow} +\infty.
\end{aligned}$$

In particular $\nabla^2\mathcal{J}$ is not Lipschitz-continuous in $\mathbb{H}^{2,2}_{\mathcal{P}}$ endowed with $\|.\|_{\mathbb{H}^{2,2}}$.

Under additional uniform boundedness and regularity assumptions, one can actually show that $\mathcal{J}$ actually defines an operator from the space of uniformly bounded process in $\mathbb{H}^{\infty,\infty}$ to reals, that $\nabla^2\mathcal{J}$ send $\mathbb{H}^{\infty,\infty}$ on $\mathcal{L}(\mathbb{H}^{\infty,\infty})$ and is Lipschitz-continuous.

Lemma 5.3.10. *Let the conditions of Theorem 5.3.7 hold. Define $P \in \mathbb{H}^{\infty,\infty}$ as in (5.3.4), $\Pi \in \mathbb{H}^{\infty,2}$ by (5.3.6), $X \in \mathbb{H}^{\infty,2}$ as in (5.3.8), $Y = PX + \Pi \in \mathbb{H}^{\infty,2}$. Suppose additionally that $a, b \in \mathbb{H}^{\infty,\infty}$ and $x, \eta \in \mathbb{L}^\infty_\Omega$. Then Π, X and Y are in $\mathbb{H}^{\infty,\infty}$ and we have the estimates:*

$$\begin{aligned}
&\|\Pi\|_{\mathbb{H}^{\infty,\infty}} \leq \left(\|\eta\|_{\mathbb{L}^\infty_\Omega} + T\|a\|_{\mathbb{H}^{\infty,\infty}}\|P\|_{\mathbb{H}^{\infty,\infty}} + T\|b\|_{\mathbb{H}^{\infty,\infty}}\right) e^{\|PB+A\|_{\mathbb{H}^{\infty,\infty}}T},\\
&\|X\|_{\mathbb{H}^{\infty,\infty}} \leq \left(\|x\|_{\mathbb{L}^\infty_\Omega} + T\|B\|_{\mathbb{H}^{\infty,\infty}}\|\Pi\|_{\mathbb{H}^{\infty,\infty}} + T\|a\|_{\mathbb{H}^{\infty,\infty}}\right) e^{\|A+PB\|_{\mathbb{H}^{\infty,\infty}}T},\\
&\|Y\|_{\mathbb{H}^{\infty,\infty}} \leq \|P\|_{\mathbb{H}^{\infty,\infty}}\|X\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Proof. The fact that $\Pi \in \mathbb{H}^{\infty,\infty}$ and the estimate on Π are immediate using formula (5.3.6). The fact that $X \in \mathbb{H}^{\infty,\infty}$ and the estimate on X are a consequence of this latter fact, from definition (5.3.8) and from Gronwall's lemma. The fact that $Y \in \mathbb{H}^{\infty,\infty}$ and the estimate on Y directly come from the fact that P, X and Π are in $\mathbb{H}^{\infty,\infty}$. □

Theorem 5.3.11 (Stability of $\mathbb{H}^{\infty,\infty}$)**.** *Suppose assumptions* **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. Then, for all $(u,v) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $X^u, Y^u, \nabla\mathcal{J}(u), \dot{X}^v, \dot{Y}^{u,v}, \nabla^2\mathcal{J}(u)(v)$ and $(\nabla^2\mathcal{J}(u))^{-1}(v)$ are all in $\mathbb{H}^{\infty,\infty}$ and besides:*

$$\begin{aligned}
&\forall (u,v,w) \in (\mathbb{H}^{\infty,\infty}_{\mathcal{P}})^3,\\
&\|P^u\|_{\mathbb{H}^{\infty,\infty}} \leq C,\\
&\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}},\\
&\|X^u\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla\mathcal{J}(u) - \nabla\mathcal{J}(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u-v\|_{\mathbb{H}^{\infty,\infty}},\\
&\|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}},\\
&\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla^2\mathcal{J}(u)(w) - \nabla^2\mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u-v\|_{\mathbb{H}^{\infty,\infty}}\|w\|_{\mathbb{H}^{\infty,\infty}},\\
&\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}},
\end{aligned}$$

where the (generic) constant $C > 0$ is independent of u, v, w. Note that this implies that $\nabla^2\mathcal{J}$ defines a Lipschitz-continuous operator from $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$ to the space of continuous endomorphisms of $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, and that $\nabla^2\mathcal{J}(u)$ and $(\nabla^2\mathcal{J}(u))^{-1}$ are bounded linear operators, uniformly in u. This also implies that for any $u \in \mathbb{H}^{\infty,\infty}$, the Newton direction $-\left(\nabla^2\mathcal{J}(u)\right)^{-1}(\nabla\mathcal{J}(u))$ is also in $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.

Proof. Note that for any $(X,Y) \in (\mathbb{H}^{\infty,\infty})^2$, $XY \in \mathbb{H}^{\infty,\infty}$ and:

$$\|XY\|_{\mathbb{H}^{\infty,\infty}} \leq \|X\|_{\mathbb{H}^{\infty,\infty}}\|Y\|_{\mathbb{H}^{\infty,\infty}}.$$

Throughout the proof, C denotes a generic deterministic constant depending only on T and the bounds on the data of the problem and their derivatives in $\mathbb{H}^{\infty,\infty}$. We have immediately for $(u,v) \in (\mathbb{H}^{\infty,\infty})^2$:

$$\begin{aligned}
&\|X^u\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u-v\|_{\mathbb{H}^{\infty,\infty}},\\
&\|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

We also have:

$$\begin{aligned}
&\|Y^u\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}} + \|X^u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u-v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}}).
\end{aligned}$$

Combining the first two upper bounds with the estimates on X^u yields the inequality on Y^u. Combining the third estimate with the bound on $\dot{X}^v$ yields:

$$\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}}.$$

Using the Lipschitz-continuity of the second-order derivatives of l and Ψ, one gets:

$$\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\left(\|u-v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}}\right)\left(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^w\|_{\mathbb{H}^{\infty,\infty}}\right).$$

We can use all our previous estimates to get the claimed bound on $\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}}$.

We have:

$$\begin{aligned}
&\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}} + \|X^u\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\nabla\mathcal{J}(u) - \nabla\mathcal{J}(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u-v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}}).
\end{aligned}$$

Combining this with the estimates on X^u and Y^u yields the estimates on $\nabla\mathcal{J}(u)$. Using the expression of $\nabla^2\mathcal{J}(u)(v)$ derived earlier, we easily get the estimate:

$$\begin{aligned}
&\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\nabla^2\mathcal{J}(u)(w) - \nabla^2\mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u-v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}})(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^w\|_{\mathbb{H}^{\infty,\infty}}) + C\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

We get the claimed estimates on $\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\nabla^2\mathcal{J}(u)(w) - \mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}}$ using the previous bounds.

We have the bounds on parameters appearing in(5.3.10):

$$\begin{aligned}
&\|A^u\|_{\mathbb{H}^{\infty,\infty}} + \|B^u\|_{\mathbb{H}^{\infty,\infty}} + \|C^u\|_{\mathbb{H}^{\infty,\infty}} + \|\Gamma^u\|_{\mathbb{L}^\infty_\Omega} \leq C,\\
&\|a^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + \|b^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Therefore, we are in the framework of application of Lemma 5.3.10 for $\Pi = \Pi^{u,w}$, $P = P^u$ and $X = \tilde{X}^{u,w}$, which are defined in (5.3.11), (5.3.12) and (5.3.13) (with $\eta = 0$ and $x = 0$). This yields:

$$\begin{aligned}
&\|P^u\|_{\mathbb{H}^{\infty,\infty}} \leq C,\\
&\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|P^u\|_{\mathbb{H}^{\infty,\infty}})\|w\|_{\mathbb{H}^{\infty,\infty}},\\
&\|\tilde{X}^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}}).
\end{aligned}$$

We have by (5.3.14):

$$\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + (1 + \|P^u\|_{\mathbb{H}^{\infty,\infty}})\|\tilde{X}^{u,w}\|_{\mathbb{H}^{\infty,\infty}}),$$

which gives the estimate on $\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}}$. □

It is well known that, for strongly convex functions, with Lipschitz-continuous second order derivative, local convergence can be shown for the Newton method, see for instance [Kan48] - [NW06, Theorem 3.5, p. 44]. However, even under such demanding assumptions, it is well-known in a finite-dimensional setting that Newton method is not guaranteed to converge globally. We provide next a counter-example to the global convergence of Newton method in our infinite-dimensional framework.

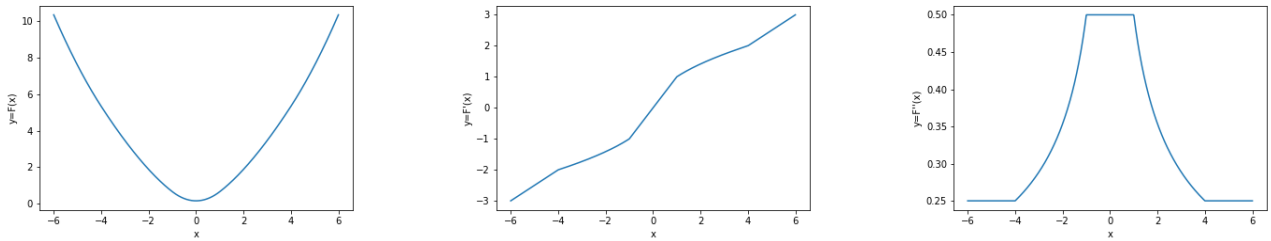
Example 5.3.12 (Counter-example to global convergence of Newton method)**.** *We consider $\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{R}$ given by:*

$$\forall u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}, \quad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 F(u_t)\mathrm{d}t\right], \tag{5.3.17}$$

$$s.t. \quad X^u_t = 0, \forall t \in [0,1], \tag{5.3.18}$$

where $F : x \mapsto \mathbb{R}$ is defined by:

$$F(x) := \begin{cases} \frac{x^2}{4} + \frac{4}{3}, & \text{if } |x| > 4,\\ \frac{2|x|^{\frac{3}{2}}}{3}, & \text{if } 1 \leq |x| \leq 4,\\ \frac{x^2}{2} + \frac{1}{6}, & \text{if } |x| < 1. \end{cases}$$



(a) Graph of function F (b) Graph of the first derivative F' (c) Graph of the second derivative F''

Figure 5.2: Graphs of F, its first and second derivatives

Then F is strongly convex, twice continuously differentiable with bounded second order derivative, see Figure 5.2.

Assumptions **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. Let $u^{(0)} \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$ be any stochastic process such that $1 < |u^{(0)}_{t,\omega}| < 4$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. Due to the particular structure of our problem, Newton's method reduces to the Newton method applied to $F : \mathbb{R} \mapsto \mathbb{R}$ applied ω by ω and t by t. For $1 < |x_0| < 4$, Newton's method (in $\mathbb{R}$) applied to F with initial guess x_0 produces the sequence $(x_k)_{k\in\mathbb{N}}$ with general term $x_k = (-1)^k x_0$. Indeed, for $1 \leq |x| \leq 4$, we have $F'(x) = sign(x)\sqrt{|x|}$ and $F''(x) = \frac{1}{2\sqrt{|x|}}$, so that the Newton iteration is given by:*

$$x_{k+1} = x_k - \frac{F'(x_k)}{F''(x_k)} = -x_k.$$

Therefore, Newton's method (in $\mathbb{H}^{\infty,\infty}$) applied to $\mathcal{J}$ with initial guess $u^{(0)}$ produces the sequence $(u^{(k)})_{k\in\mathbb{N}}$ with general term $u^{(k)} = (-1)^k u^{(0)}$, which does not converge.

The previous counter-example motivates globalization procedures. We consider Backtracking-line search methods, which are iterative procedures which allow to select appropriate step lengths such that the Goldstein conditions, presented in [NW06, p. 36], hold.

The standard Backtracking line-search method is given in Algorithm 5.3. It is directly adapted from Backtracking line-search method in $\mathbb{R}^n$ [NW06, Algorithm 3.1, p.37] or [BV04, Algorithm 9.2, p. 464]. Under the assumption of a strongly convex function with bounded and Lipschitz second order derivative, it can be shown in the finite dimensional case that the Newton method with the Standard Backtracking line search converges globally, see [BV04, Section 9.5.3, pp. 488-491].

Algorithm 5.3 Standard Backtracking line search (compact generic version)

1: **Inputs:** Current point $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, Current search direction $\Delta u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
2: $\sigma = 1$.
3: **while** $\mathcal{J}(u + \sigma\Delta u) > \mathcal{J}(u) + \gamma\sigma\langle\nabla\mathcal{J}(u), \Delta u\rangle_{\mathbb{H}^{2,2}}$ **do**
4: $\quad \sigma \leftarrow \beta\sigma$.
5: **end while**
6: **return** $u + \sigma\Delta u$.

However, since we do not work in a finite-dimensional setting, the global convergence of the method is not guaranteed to our knowledge. This is due to the fact that $\mathcal{J}$ is a criteria in expectation, whereas we are working with the norm in $\mathbb{H}^{\infty,\infty}$. We shall see in our numerical applications that the Standard Backtracking line-search may prevent the Newton method to converge, see Figures 5.6b and 5.6d.

To alleviate this issue, we propose a new Backtracking line-search rule, based on the sup-essential norm of the gradient, described in Algorithm 5.4. Hence, we are only working with the norm in $\mathbb{H}^{\infty,\infty}$, which makes more sense as we are working with variables in $\mathbb{H}^{\infty,\infty}$.

We show that the Newton method combined with the Gradient Backtracking line search algorithm 5.4 ensures global convergence of the method, and that after a finite number of iterations, the algorithm takes full Newton steps, which ensures quadratic convergence.

Algorithm 5.4 Gradient Backtracking line search (compact generic version)

Inputs: Current point $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, Current search direction $\Delta u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
$\sigma = 1$.
while $\|\nabla \mathcal{J}(u + \sigma \Delta u)\|_{\mathbb{H}^{\infty,\infty}} > (1 - \gamma\sigma)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}$ **do**
  $\sigma \leftarrow \beta\sigma$.
end while
return $u + \sigma \Delta u$.

Lemma 5.3.13. *Suppose assumptions* **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. For any* $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, *the Gradient Backtracking line search terminates in finitely many iterations for the Newton step* $\Delta u := -(\nabla^2 \mathcal{J}(u))^{-1}(\nabla \mathcal{J}(u)) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.

Besides, if the algorithm returns $\sigma = 1$, *then the new point* $u + \Delta u$ *satisfies:*

$$\|\nabla \mathcal{J}(u + \Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \min(1 - \gamma, C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}})\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}},$$

where $C = \frac{L_{\nabla^2 \mathcal{J}} C^2_{(\nabla^2 \mathcal{J})^{-1}}}{2}$ *with* $L_{\nabla^2 \mathcal{J}}$ *the Lipschitz constant of* $\nabla^2 \mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty})$, *and*

$$C_{(\nabla^2 \mathcal{J})^{-1}} := \sup_{u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}} \|(\nabla^2 \mathcal{J}(u))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})},$$

which is finite by Theorem 5.3.11). Conversely, if $C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq (1 - \gamma)$ *then the algorithm returns* $\sigma = 1$. *On the other hand, if the algorithm returns* $\sigma < 1$, *then the new iterate* $u + \sigma \Delta u$ *satisfies:*

$$\|\nabla \mathcal{J}(u + \sigma \Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} - \frac{\beta\gamma(1 - \gamma)}{C}.$$

Proof. By Taylor-Lagrange formula, using $\Delta u = -(\nabla^2 \mathcal{J}(u))^{-1}(\nabla \mathcal{J}(u))$, we have

$$\begin{aligned}
\nabla \mathcal{J}(u + \sigma \Delta u) &= \nabla \mathcal{J}(u) + \int_0^1 \nabla^2 \mathcal{J}(u + s\sigma \Delta u)(\sigma \Delta u)\mathrm{d}s, \\
&= (1 - \sigma)\nabla \mathcal{J}(u) + \sigma \int_0^1 \left(\nabla^2 \mathcal{J}(u + s\sigma \Delta u) - \nabla^2 \mathcal{J}(u)\right)(\Delta u)\mathrm{d}s,
\end{aligned}$$

which yields by Lipschitz continuity of $\nabla^2 \mathcal{J}$ in $\mathcal{L}(\mathbb{H}^{\infty,\infty})$:

$$\begin{aligned}
\|\nabla \mathcal{J}(u + \sigma \Delta u)\|_{\mathbb{H}^{\infty,\infty}} &\leq (1 - \sigma)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} + \sigma \int_0^1 \|\nabla^2 \mathcal{J}(u) - \nabla^2 \mathcal{J}(u + s\sigma \Delta u)\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}\mathrm{d}s\|\Delta u\|_{\mathbb{H}^{\infty,\infty}} \\
&\leq (1 - \sigma)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} + \frac{L_{\nabla^2 \mathcal{J}}}{2}\sigma^2\|\Delta u\|^2_{\mathbb{H}^{\infty,\infty}} \\
&\leq \left(1 - \sigma + \frac{L_{\nabla^2 \mathcal{J}} C^2_{(\nabla^2 \mathcal{J})^{-1}}}{2}\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\sigma^2\right)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Notice that, since $0 < \gamma < 1$, for $\sigma > 0$ small enough, we have:

$$1 - \sigma + C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\sigma^2 \leq 1 - \gamma\sigma. \tag{5.3.19}$$

In particular, the Gradient Backtracking line search terminates after finitely many iterations. Besides, if $\frac{C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}}{1-\gamma} \leq 1$ then the algorithm stops and returns $\sigma = 1$.

Suppose that the algorithm stops with $\sigma = 1$, then:

$$\|\nabla \mathcal{J}(u + \Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \min(1 - \gamma, C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}})\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}},$$

by the termination criterion and the previous estimate.

Suppose that the algorithm returns $\sigma < 1$. Then by the termination criteria of the algorithm, (5.3.19) implies in particular that:

$$-(\sigma/\beta) + C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}(\sigma/\beta)^2 > -\gamma(\sigma/\beta).$$

This yields:

$$\frac{\beta(1-\gamma)}{C\|\nabla\mathcal{J}(u).\|_{\mathbb{H}^{\infty,\infty}}} < \sigma < 1.$$

This yields, using the termination criterion of the algorithm, the fact that σ, γ and β are in $(0,1)$ and the previous inequality:

$$\begin{aligned}\|\nabla\mathcal{J}(u+\sigma\Delta u)\|_{\mathbb{H}^{\infty,\infty}} &\leq (1-\gamma\sigma)\,\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\\ &\leq \left(1-\frac{\beta\gamma(1-\gamma)}{C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}}\right)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\\ &\leq \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} - \frac{\beta\gamma(1-\gamma)}{C}.\end{aligned}$$

□

Remark 5.3.14. *Note that the properties of the backtracking line-search algorithm can be enforced for values of parameters which are independent on the problem, and in particular, independent from a priori unknown constants of the problem (bounds on derivatives, constant of strong convexity...). The properties of the problem (regularity, convexity) only impact the number of iterations of the algorithm.*

Algorithm 5.5 Newton's method with Backtracking line search (compact generic version)

1: **Inputs:** $u^{(0)} \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\gamma \in (0,1)$, $\beta \in (0,1)$, $\varepsilon > 0$. $k = 0$
2: **while** $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \varepsilon$ **do**
3: Compute Newton direction $-(\nabla^2\mathcal{J}(u^{(k)}))^{-1}(\nabla\mathcal{J}(u^{(k)})) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$
4: Compute new iterate using Backtracking line-search rule $u^{(k+1)} = u^{(k)} - \sigma(\nabla^2\mathcal{J}(u^{(k)}))^{-1}(\nabla\mathcal{J}(u^{(k)})) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$
5: $k \leftarrow k+1$
6: **end while**
7: **return** $u^{(k)}$

Theorem 5.3.15. *Suppose assumptions* **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. Then $\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{R}$ is twice continuously differentiable, with Lipschitz continuous first and second derivatives. Besides $\nabla^2\mathcal{J}$ is an invertible, bi-continuous endomorphism of $\mathbb{H}^{\infty,\infty}$. Let $u^{(0)} \in \mathbb{H}^{\infty,\infty}$. Let γ and β be parameters in $(0,1)$. Define $C = \frac{L_{\nabla^2\mathcal{J}}C^2_{(\nabla^2\mathcal{J})^{-1}}}{2}$ with $L_{\nabla^2\mathcal{J}}$ the Lipschitz constant of $\nabla^2\mathcal{J} : \mathbb{H}^{\infty,\infty}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty}_{\mathcal{P}})$ and*

$$C_{(\nabla^2\mathcal{J})^{-1}} := \sup_{u\in\mathbb{H}^{\infty,\infty}_{\mathcal{P}}} \|(\nabla^2\mathcal{J}(u))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}.$$

Define as well:

$$\begin{aligned}\eta &= \frac{\beta\gamma(1-\gamma)}{C},\\ k_1 &= \inf\left\{k\in\mathbb{N} \quad | \quad \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{1-\gamma}{C}\right\}.\end{aligned}$$

Then k_1 is finite. Besides, after k_1 iterations, the step length is always 1 and Newton method with Gradient Backtracking line search converges quadratically, i.e.,

$$\forall k \geq k_1, \quad C\|\nabla\mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq \left(C\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)^2,$$

$$\forall k \geq k_1, \quad \|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{(1-\gamma)^{2^{k-k_1}}}{C}.$$

Besides, $u^{(k)}$ converges to $u^* \in \mathbb{H}^{\infty,\infty}$ which is the minimizer of $\mathcal{J}$, and the asymptotic convergence is quadratic, i.e.,

$$\forall k \geq k_1, \quad C\|u^{(k+1)} - u^*\|_{\mathbb{H}^{\infty,\infty}} \leq C^2 \|u^{(k)} - u^*\|^2_{\mathbb{H}^{\infty,\infty}},$$
$$\forall k \geq k_1 + 1, \quad \|u^{(k)} - u^*\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{\gamma C}(1-\gamma)^{2^{k-k_1}}.$$

Proof. By the previous lemma, the sequence $\left(\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ is monotone decreasing.

We first prove that $k_1 < +\infty$. Let $k \leq k_1$. Then $\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \frac{1-\gamma}{C}$. If at iteration k, the gradient backtracking line search returns a unit step length $\sigma = 1$, then by Lemma 5.3.13:

$$\begin{aligned}\|\nabla \mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} &\leq (1-\gamma)\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} - \eta,\end{aligned}$$

where we used the assumption $\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \frac{1-\gamma}{C}$ and $\beta \in (0,1)$.

Else, at iteration k, the gradient backtracking line search returns a step length $\sigma < 1$ and still by Lemma 5.3.13,

$$\|\nabla \mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} - \eta.$$

This yields:

$$\forall k \leq k_1, \quad \|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla \mathcal{J}(u^{(0)})\|_{\mathbb{H}^{\infty,\infty}} - k\eta.$$

Since $\eta > 0$, this yields existence and finiteness of k_1, which is bounded from above by $\frac{\|\nabla \mathcal{J}(u^{(0)})\|_{\mathbb{H}^{\infty,\infty}}}{\eta} + 1$.

Besides, for all $k \geq k_1$, we have $\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{1-\gamma}{C}$ since the sequence $\left(\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ is monotone decreasing. In that case, Lemma 5.3.13 shows that the algorithm takes unit step length and:

$$C\|\nabla \mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq C^2 \|\nabla \mathcal{J}(u^{(k+1)})\|^2_{\mathbb{H}^{\infty,\infty}}.$$

This combined with $C\|\nabla \mathcal{J}(u^{(k_1)})\| \leq 1-\gamma$ yields:

$$\forall k \geq k_1, \quad C\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq (1-\gamma)^{2^{k-k_1}}.$$

From that and since the algorithm takes unit step length after iteration k_1, we deduce:

$$\begin{aligned}\forall k \geq k_1, \quad \|u^{(k+1)} - u^{(k)}\|_{\mathbb{H}^{\infty,\infty}} &= \|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)}))\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}\frac{(1-\gamma)^{2^{k-k_1}}}{C}.\end{aligned}$$

In particular, this yields the absolute convergence of the series $S_n = \sum_{k=0}^n v_n$ of general term $v_0 = u^{(0)}$ and $v_k = u^{(k)} - u^{(k-1)}$ for $k > 1$. Hence $(S_n)_{n\in\mathbb{N}}$ converges in $\mathbb{H}^{\infty,\infty}$ and so does $(u^{(k)})_{k\in\mathbb{N}}$. Denote $u^* \in \mathbb{H}^{\infty,\infty}$ the limit point of $(u^{(k)})_{k\in\mathbb{N}}$. By continuity of $\nabla \mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{H}^{\infty,\infty}$, u^* is a critical point of $\mathcal{J}$ and hence, by strong convexity of $\mathcal{J}$, u^* is the unique minimizer of $\mathcal{J}$. Besides, as the algorithm takes unit step length $\sigma = 1$ after k_1 iterations,

$$\begin{aligned}\forall k \geq k_1, \|u^{(k+1)} - u^*\|_{\mathbb{H}^{\infty,\infty}} &= \|u^{(k)} - (\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)})) - u^*\|_{\mathbb{H}^{\infty,\infty}} \\ &= \|u^{(k)} - u^* - (\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)}) - \nabla \mathcal{J}(u^*))\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}\|\nabla \mathcal{J}(u^{(k)}) - \nabla \mathcal{J}(u^*) - \nabla^2 \mathcal{J}(u^{(k)})(u^{(k)} - u^*)\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \frac{\|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})} L_{\nabla^2 \mathcal{J}}}{2}\|u^{(k)} - u^*\|^2_{\mathbb{H}^{\infty,\infty}},\end{aligned}$$

by Taylor-Lagrange's formula and Lipschitz continuity of $\nabla^2\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty})$.

Besides, for all $k > k_1$,

$$\begin{aligned}
\|u^{(k)} - u^*\|_{\mathbb{H}^{\infty,\infty}} &\leq \sum_{j=k}^{+\infty} \|u^{(j+1)} - u^{(j)}\|_{\mathbb{H}^{\infty,\infty}} \\
&\leq \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} \left(\sum_{j=k}^{+\infty} (1-\gamma)^{2^{j-k_1}} \right) \\
&= \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} \left(\sum_{j=0}^{+\infty} (1-\gamma)^{2^{j+k-k_1}} \right) \\
&\leq \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} (1-\gamma)^{2^{k-k_1}} \left(1 + \sum_{j=0}^{+\infty} (1-\gamma)^{2^j} \right) \\
&\leq \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} (1-\gamma)^{2^{k-k_1}} \left(\sum_{j=0}^{+\infty} (1-\gamma)^{j} \right) \\
&\leq \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C\gamma} (1-\gamma)^{2^{k-k_1}}
\end{aligned}$$

where we used $k > k_1$ and the fact that $2^{j+k-k_1} \geq 2^j + 2^{k-k_1}$ for all $j \geq 1$. □

5.4 Application: energy storage system control for power balancing

5.4.1 Problem setting

We consider N identical batteries with energy capacity $\mathcal{E}_{\max}$ operated in order to balance production and consumption on an electricity network. Other types of energy storage systems could be considered. For instance, one could replace batteries in this application by a large population of Thermostatically Controlled Loads (TCLs), which include water heaters, Air Conditioners, Heat pumps,... provided a first order affine-linear model of their temperature dynamic is used. The global consumption on the network is given by a deterministic function $N\mathrm{P}^{\mathrm{cons}}$, where $\mathrm{P}^{\mathrm{cons}}$ is the total consumption divided by the number of batteries. The assumption of a deterministic consumption profile can be justified by the fact that it is the aggregation of a large number of small independent consumption profiles, which allows to use the law of large numbers. We assume additionally a total solar power production $N\mathrm{P}^{\mathrm{sun}}$, i.e. $\mathrm{P}^{\mathrm{sun}}$ is the total solar production divided by the number of batteries on the network (we do not account for wind power, although this could easily be included in the model). We follow [Bad+18] by setting $\mathrm{P}^{\mathrm{sun}} = \mathrm{P}^{\mathrm{sun,max}}\mathrm{X}^{\mathrm{sun}}$ where $\mathrm{P}^{\mathrm{sun,max}} : [0,T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) represented in 5.3a and $\mathrm{X}^{\mathrm{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathrm{X}_t^{\mathrm{sun}} = -\rho^{\mathrm{sun}}(\mathrm{X}_t^{\mathrm{sun}} - \mathrm{x}_t^{\mathrm{sun,ref}})\mathrm{d}t + \sigma^{\mathrm{sun}}(\mathrm{X}_t^{\mathrm{sun}})^{k_1}(1-\mathrm{X}_t^{\mathrm{sun}})^{k_2}\mathrm{d}\tilde{W}_t, \tag{5.4.1}$$

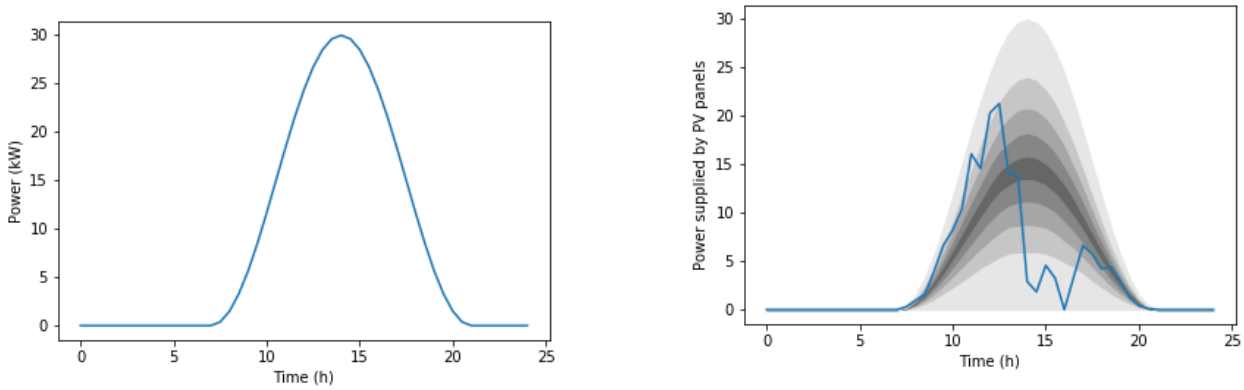
with $k_1, k_2 \geq 1/2$. As proved in [Bad+18], there is a strong solution to the above SDE and the solution $\mathrm{X}^{\mathrm{sun}}$ takes values in $[0,1]$. Since the drifts are affine-linear, the conditional expectation of the solution of (5.4.1) is known in closed forms (this property is intensively used in [BSS05]):

$$\mathbb{E}_t\left[\mathrm{P}_s^{\mathrm{sun}}\right] = \left(\frac{\mathrm{P}_t^{\mathrm{sun}}}{\mathrm{P}_t^{\mathrm{sun,max}}} \exp(-\rho^{\mathrm{sun}}(s-t)) + \int_t^s \rho^{\mathrm{sun}}\mathrm{x}_\tau^{\mathrm{sun,ref}} \exp(-\rho^{\mathrm{sun}}(s-\tau))\mathrm{d}\tau \right)\mathrm{P}_s^{\mathrm{sun,max}}, \tag{5.4.2}$$

for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\mathrm{P}_s^{\mathrm{sun}}\right]$ and $\mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]$ as required when deriving the optimal control. The value of the parameters used are given in the following table. Empirical quantile plot (obtained by simulation of 10000 i.i.d. trajectories) as well as one example trajectory of $\mathrm{P}^{\mathrm{sun}}$ are given in Figure 5.3b.

Table 5.1: Parameter values for the simulation of PV power production

ρ^{sun}	$\mathrm{x}^{\mathrm{sun,ref}}$	σ^{sun}	k_1	k_2
0.75 h^{-1}	0.5	0.8	0.8	0.7



(a) Time evolution of $\mathrm{P}^{\mathrm{sun,max}}$

(b) Empirical quantiles of $\mathrm{P}^{\mathrm{sun}}$, obtained with $M = 10000$ samples, and one realization of $\mathrm{P}^{\mathrm{sun}}$

Figure 5.3: Graphical statistics of the evolution of $\mathrm{P}^{\mathrm{sun}}$

Our goal is to minimize global cost for the control of N batteries, which are composed of operational costs for battery management and a penalization for power balance imbalance, represented in Figure 5.4[1].

We assume a production profile per battery $\mathrm{P}_t^{\mathrm{prod}} = \mathrm{P}_t^{\mathrm{cons}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$, which can be easily computed using the model on $\mathrm{P}^{\mathrm{sun}}$.

Denoting by $u^{(n)}$ the power supplied by battery $n \in [N]$ and by $X^{(n)}$ its normalized state of charge, we wish to solve the following stochastic control problem:

$$\min_{u \in \mathbb{H}_{\mathcal{P}}^{2,2}} \mathbb{E}\left[\int_0^T \left\{\frac{1}{N}\sum_{n=1}^N \frac{\mu}{2}(u_t^{(n)})^2 + \frac{1}{N}\sum_{n=1}^N \frac{\nu}{2}\left(X_t^{(n)} - \frac{1}{2}\right)^2 + \mathcal{L}\left(\frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\} \mathrm{d}t + \frac{1}{N}\sum_{n=1}^N \frac{\rho}{2}\left(X_T^{(n)} - \frac{1}{2}\right)^2\right], \tag{5.4.3}$$

$$s.t. \quad X_t^{(n)} = x_0^{(n)} - \int_0^t \frac{u_s^{(n)}}{\mathcal{E}_{\max}} \mathrm{d}s. \tag{5.4.4}$$

The first two terms as well as the last term in the cost functional represent the sum of the operational costs for individual batteries. We penalize quadratically power supplied or absorbed by the batteries (first term) and penalize deviations of the normalized states of charge of the batteries from the reference value $1/2$ (second and last term). The third term represents a penalization term for the power imbalance $\frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathrm{P}_t^{\mathrm{cons}} + \mathrm{P}_t^{\mathrm{prod}} = \frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$, using $\mathrm{P}_t^{\mathrm{prod}} = \mathrm{P}_t^{\mathrm{cons}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$. The state variable $X^{(n)}$ represents the normalized state of charge of battery n, i.e., the energy stored divided by the maximal capacity $\mathcal{E}_{\max} = 150$ kWh. In particular, the total installed storage capacity corresponds to 5 hours of the PV panels production at full capacity, which is 30 kW, which corresponds to about 300 squared meters of photo-voltaic panels, with the current technology. Equivalently, assuming a availability rate of 12% for solar (accounting for seasonality, intermittency and unavailability at night), about 40 hours of the average solar production, where the availability rate is defined as the average power production divided by the maximal power capacity. We consider simple ideal batteries with charging and discharging efficiencies equal to 1. We do not enforce the state constraints $X^{(n)} \in [0,1]$. We will consider a non-quadratic loss function $\mathcal{L}$

[1] Icons made by Freepik and Smashicons from www.flaticon.com

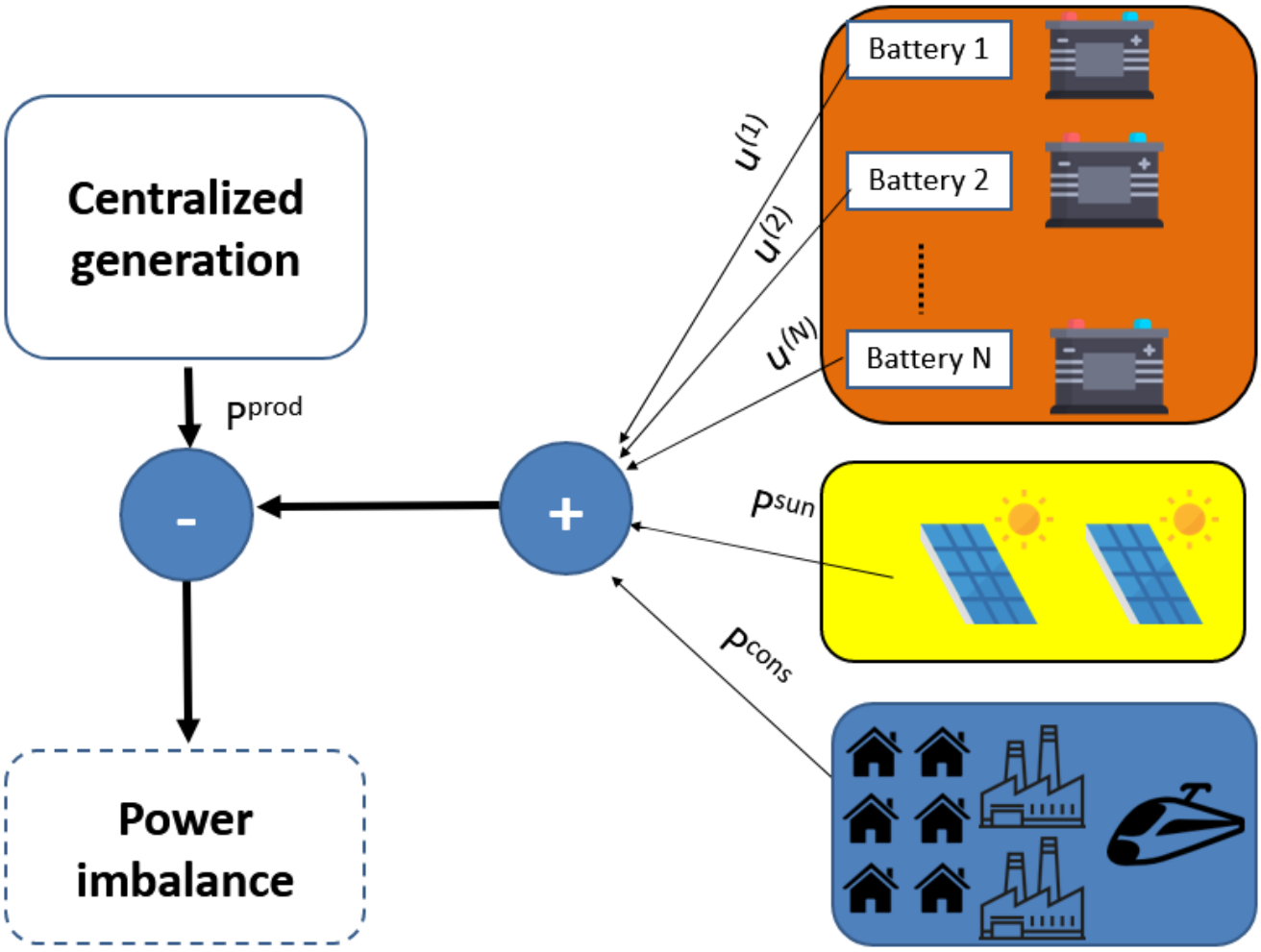


Figure 5.4: Power imbalance on the network

given by:

$$\mathcal{L}(x) = \begin{cases} \frac{1}{2}\lambda x^2 - \frac{\delta}{6\varepsilon}x^3 & \text{if } -\varepsilon \leq x \leq \varepsilon, \\ \frac{1}{2}(\lambda - \delta)x^2 + \frac{\delta\varepsilon}{2}x - \frac{\delta}{6}\varepsilon^2 & \text{if } x \geq \varepsilon, \\ \frac{1}{2}(\lambda + \delta)x^2 + \frac{\delta\varepsilon}{2}x + \frac{\delta}{6}\varepsilon^2 & \text{if } x \leq -\varepsilon. \end{cases}$$

In this case, $\mathcal{L}$ is C^2 with $\frac{\delta}{\varepsilon}$-Lipschitz-continuous second order derivatives and is represented in Figure 5.5 for $\lambda = 2, \delta = 1, \varepsilon = 0.1$.

The function $\mathcal{L}$ penalizes more energy production deficit (as compared to its expected value). Indeed, such situation possibly requires the use of extra production units with high carbon footprint, which is clearly to discard as often as possible. We will use the following parameter values for the cost functional.

Table 5.2: Parameter values for the cost functional

μ	ν	ρ	λ	δ	ε
1 kW^{-2} h^{-1}	5 h^{-1}	10000	10 kW^{-2} h^{-1}	5 kW^{-2} h^{-1}	1 kW

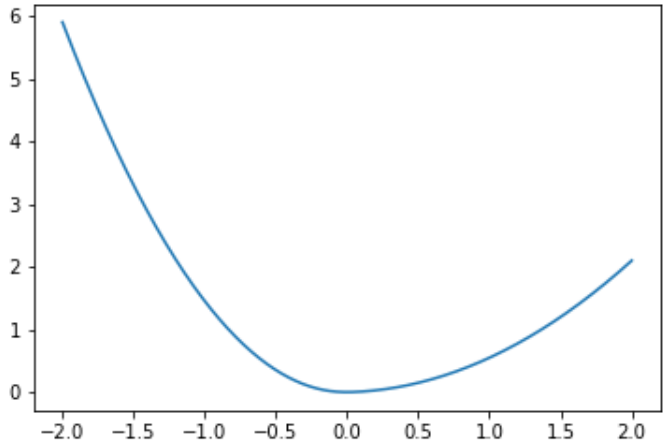


Figure 5.5: Loss function $\mathcal{L}$ with $\lambda = 2, \delta = 1, \varepsilon = 0.1$

5.4.2 Solving the stochastic control problem

Applying a similar methodology as in Theorem 5.2.6, one can show that it is equivalent to solve the stochastic control problem (5.4.3) and the following high-dimension coupled FBSDE:

$$\forall n \in [N], \quad \begin{cases} X_t^{(n)} = x_0^{(n)} - \int_0^t \frac{u_s^{(n)}}{\mathcal{E}_{\max}} \mathrm{d}s, \\ Y_t^{(n)} = \mathbb{E}_t \left[\rho(X_T^{(n)} - 1/2) + \int_t^T \nu(X_s^{(n)} - 1/2) \mathrm{d}s \right], \\ \mu u_t^{(n)} + \mathcal{L}' \left(\frac{1}{N} \sum_{j=1}^N u_t^{(j)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}} \right] \right) - \frac{Y_{t-}^{(n)}}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{5.4.5}$$

There exists a unique solution $(u^{(n)}, X^{(n)}, Y^{(n)})_{n\in[N]} \in (\mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2})^N$ of the above FBSDE, and $(u^{(n)})_{n\in[N]}$ is the unique solution of the stochastic control problem (5.4.3). The above FBSDE (5.4.5) is a high-dimensional fully coupled FBSDE. To solve it, introduce

$$(\bar{u}, \bar{X}, \bar{Y}) := \left(\frac{1}{N} \sum_{j=1}^N u^{(j)}, \frac{1}{N} \sum_{j=1}^N X^{(j)}, \frac{1}{N} \sum_{j=1}^N Y^{(j)} \right).$$

Introduce as well $\left(u^{(n),\Delta}, X^{(n),\Delta}, Y^{(n),\Delta}\right)_{n\in[N]} := \left(u^{(n)} - \bar{u}, X^{(n)} - \bar{X}, Y^{(n)} - \bar{Y}\right)$. Then for all $n \in [N]$, $\left(u^{(n),\Delta}, X^{(n),\Delta}, Y^{(n),\Delta}\right)$ is a solution of the linear FBSDE:

$$\begin{cases} X_t^{(n),\Delta} = x_0^{(n)} - \bar{x}_0 - \int_0^t \frac{u_s^{(n),\Delta}}{\mathcal{E}_{\max}} \mathrm{d}s, \\ Y_t^{(n),\Delta} = \mathbb{E}_t \left[\rho X_T^{(n),\Delta} + \int_t^T \nu X_s^{(n),\Delta} \mathrm{d}s \right], \\ \mu u_t^{(n),\Delta} - \frac{Y_{t-}^{(n),\Delta}}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{5.4.6}$$

To solve efficiently this linear FBSDE, we use Theorem 4.5.2. Let us turn to the computation of $(\bar{u}, \bar{X}, \bar{Y})$. Notice that $(\bar{u}, \bar{X}, \bar{Y})$ is solution of the following FBSDE, where we denoted $\bar{x}_0 = \frac{1}{N} \sum_{j=1}^N x_0^{(j)}$:

$$\begin{cases} \bar{X}_t = \bar{x}_0 - \int_0^t \frac{\bar{u}_s}{\mathcal{E}_{\max}} \mathrm{d}s, \\ \bar{Y}_t = \mathbb{E}_t \left[\rho(\bar{X}_T - 1/2) + \int_t^T \nu(\bar{X}_s - 1/2) \mathrm{d}s \right], \\ \mu \bar{u}_t + \mathcal{L}' \left(\bar{u}_t + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}} \right] \right) - \frac{\bar{Y}_t}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{5.4.7}$$

We will assume $\bar{x}_0 = 0.5$ in our case study. This FBSDE fully characterizes the solution of the following stochastic control problem, called coordination problem:

$$\min_{\bar{u}\in\mathbb{H}^{2,2}} \quad \bar{\mathcal{J}}(\bar{u}) := \mathbb{E}\left[\int_0^T \left\{\frac{\mu}{2}\bar{u}_t^2 + \frac{\nu}{2}\left(\bar{X}_t - \frac{1}{2}\right)^2 + \mathcal{L}(\bar{u}_t + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right])\right\}\mathrm{d}t + \frac{\rho}{2}\left(\bar{X}_T - \frac{1}{2}\right)^2\right]$$
$$s.t. \quad \bar{X}_t = \bar{x}_0 - \int_0^t \frac{\bar{u}_s}{\mathcal{E}_{\max}}\mathrm{d}s.$$

Proposition 5.2.3 shows that there exists a unique solution of the coordination problem. By applying Theorem 5.2.6, we deduce the existence and uniqueness of a solution $(\bar{u}, \bar{X}, \bar{Y}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$. To solve the non-linear FBSDE (5.4.7), we use Newton's method globalized with (Gradient) Backtracking Line Search, noting that the random parameters of the problem are uniformly bounded.

Let $\bar{u}^{(k)}$ be the (candidate) control variable at iteration k. We define the associated state variable $\bar{X}^{(k)}$ at iteration k:

$$\bar{X}_t^{(k)} = \bar{x}_0 - \int_0^t \frac{\bar{u}_s^{(k)}}{\mathcal{E}_{\max}}\mathrm{d}s, \tag{5.4.8}$$

the adjoint variable $\bar{Y}^{(k)}$ at iteration k:

$$\bar{Y}_t^{(k)} = \mathbb{E}_t\left[\rho(\bar{X}_T^{(k)} - 1/2) + \int_t^T \nu(\bar{X}_s^{(k)} - 1/2)\mathrm{d}s\right]. \tag{5.4.9}$$

Applying Proposition 5.2.5 to $\bar{\mathcal{J}}$, the gradient of the cost at $\bar{u}^{(k)}$ is given by:

$$(\nabla\bar{\mathcal{J}}(\bar{u}^{(k)}))_t = \mu\bar{u}_t^{(k)} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right) - \frac{\bar{Y}_{t-}^{(k)}}{\mathcal{E}_{\max}}. \tag{5.4.10}$$

The Newton direction $\dot{u}^{(k)} = -(\nabla^2\bar{\mathcal{J}}(\bar{u}^{(k)}))^{-1}(\nabla\bar{\mathcal{J}}(\bar{u}^{(k)}))$ at the point $\bar{u}^{(k)}$ is given by:

$$\dot{u}_t^{(k)} = \frac{\dot{Y}_{t-}^{(k)} + \bar{Y}_{t-}^{(k)} - \left\{\mu\bar{u}_t^{(k)} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}},$$

where $(\dot{X}^{(k)}, \dot{Y}^{(k)})$ satisfy:

$$\begin{cases} \dot{X}_t^{(k)} = -\int_0^t \frac{\dot{u}_s^{(k)}}{\mathcal{E}_{\max}}\mathrm{d}s, \\ \dot{Y}_t^{(k)} = \mathbb{E}_t\left[\rho\dot{X}_T^{(k)} + \int_t^T \nu\dot{X}_s^{(k)}\mathrm{d}s\right]. \end{cases}$$

This comes from the application of Theorem 5.2.9 which gives the expression of the inverse of the second order derivative at $\bar{u}^{(k)}$, applied to $-\nabla\bar{\mathcal{J}}(\bar{u}^{(k)})$. Eliminating $\dot{u}^{(k)}$, we obtain the following affine-linear FBSDE for $(\dot{X}^{(k)}, \dot{Y}^{(k)})$:

$$\begin{cases} \dot{X}_t^{(k)} = \displaystyle\int_0^t \frac{-\dot{Y}_s^{(k)} - \bar{Y}_s^{(k)} + \left\{\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}\mathrm{d}s, \\ \dot{Y}_t^{(k)} = \mathbb{E}_t\left[\rho\dot{X}_T^{(k)} + \int_t^T \nu\dot{X}_s^{(k)}\mathrm{d}s\right]. \end{cases}$$

To solve this FBSDE arising at iteration k, we apply Theorem 5.3.7 in our particular framework, with the following

values for the parameters:

$$\begin{cases} A_t = 0, \\ B_t = \dfrac{-1}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}, \\ C_t = \nu, \\ \Gamma = \rho, \\ a_t = \dfrac{-\bar{Y}_t^{(k)} + \mu\bar{u}_t^{(k)}\mathcal{E}_{\max} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}, \\ b_t = 0, \\ \eta = 0, \\ x = 0. \end{cases}$$

Introduce the Riccati BSDE with stochastic coefficients:

$$P_t^{(k)} = \mathbb{E}_t\left[\rho + \int_t^T \left(\nu - \frac{1}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}(P_s^{(k)})^2\right)\mathrm{d}s\right], \tag{5.4.11}$$

and the linear BSDE:

$$\Pi_t^{(k)} = \mathbb{E}_t\left[\int_t^T \left(\frac{-P_s^{(k)}}{\mathcal{E}_{\max}^2\left(\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)}\Pi_s^{(k)} + \frac{\left(\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)\mathcal{E}_{\max} - \bar{Y}_s^{(k)}}{\mathcal{E}_{\max}^2\left(\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)}P_s^{(k)}\right)\mathrm{d}s\right]. \tag{5.4.12}$$

Then, $\dot{u}^{(k)}$ is given by the following feedback expression:

$$\dot{u}_t^{(k)} = \frac{P_{t-}^{(k)}\dot{X}_t^{(k)} + \Pi_{t-}^{(k)} + \bar{Y}_{t-}^{(k)} - \left\{\mu\bar{u}_t^{(k)} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}. \tag{5.4.13}$$

The process $\dot{X}^{(k)}$ satisfies:

$$\dot{X}_t^{(k)} = \int_0^t \frac{-P_s^{(k)}\dot{X}_s^{(k)} - \Pi_s^{(k)} - \bar{Y}_s^{(k)} + \left\{\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}\mathrm{d}s,$$

and $\dot{Y}^{(k)} = P^{(k)}\dot{X}^{(k)} + \Pi^{(k)}$, according to Theorem 5.3.7.

To be able to practically implement the Newton method with (Gradient) Backtracking line search, the conditional expectations in the equations or expressions of $\bar{Y}^{(k)}$, $P^{(k)}$, $\Pi^{(k)}$ in (5.4.9), (5.4.11) and (5.4.12) need to be estimated. We focus on these aspects in the next section.

5.4.3 Practical implementation

The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. The process $\mathrm{X}^{\mathrm{sun}}$ is simulated using an Euler scheme with time step $h = \frac{T}{N_T} = 0.5$ h, with $T = 24$ h and $N_T = 48$. The number of Monte-Carlo simulations is $M = 10000$.

Linear Least-Squares Regression

To provide an implementation to compute the conditional expectation in the expression of $\bar{Y}^{(k)}$, $P^{(k)}$, $\Pi^{(k)}$ in (5.4.9), (5.4.11) and (5.4.12), we take advantage on the Markovian framework and use Linear-Least Square regression [GT16a]. We use this method to obtain closed-loop feedback expressions of the solutions of each BSDE, with respect to the Markovian underlying extended state process $(\mathrm{X}^{\mathrm{sun}}, \bar{X}^{(k)})$.

Notations To simplify the notations, we write $(u^{(k)}, X^{(k)}, Y^{(k)}) = (u^{(k)}_\tau, X^{(k)}_\tau, Y^{(k)}_\tau)_{\tau\in[N_T]}$ the discretized process associated $(\bar{u}^{(k)}, \bar{X}^{(k)}, \bar{Y}^{(k)})$ in the following on the time grid $(\tau h)_{\tau\in[N_T]}$(not be be confused with $(u^{(n)}, X^{(n)}, Y^{(n)})$ which is the optimal control, state and adjoint variable of battery n). We also use the notation $\mathbf{x}^{\text{sun},h}_{\tau,m}$ for values of the mth simulated (discretized) approximation of $\mathbf{x}^{\text{sun}}$ at time τh.

Definition 5.4.1 (Linear Least-Squares Regression (LLSR) [GT16a]). *For $l \geq 1$ and for probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l), \nu)$, let S be a $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable $\mathbb{R}$-valued function such that $S(\omega,\cdot) \in \mathbb{L}^2(\mathcal{B}(\mathbb{R}^l), \nu)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. Let $\mathcal{K} := span(\phi_f)_{f=1,\dots,N_f}$ be the vector space spanned by N_f deterministic functions $(\phi_f)_{f=1,\dots,N_f}$. The Least-Squares approximation of S in the space $\mathcal{K}$ with respect to ν is the $\mathrm{d}\mathbb{P}\otimes\mathrm{d}\nu$-a.e. unique $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable function $S^\star$ given by:*

$$S^\star(\omega,\cdot) = \arg\inf_{\phi\in\mathcal{K}} \int |\phi(x) - S(\omega,x)|^2 \nu(\mathrm{d}x).$$

*We say that $S^\star$ solves **OLS**$(S, \mathcal{K}, \nu)$.*

In particular, if $\nu_M = \frac{1}{M}\sum_{m=1}^M \delta_{\chi^{(m)}}$ is a discrete probability measure on $(\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l))$ where $\chi^{(1)}, \chi^{(2)}, \dots, \chi^{(M)} : \Omega \to \mathbb{R}^l$ are i.i.d. random variables with distribution ν. For an $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable $\mathbb{R}$-valued function S such that $|S(\omega, \chi^{(m)}(\omega))| < \infty$ for any m and $\mathbb{P}$-a.e. $\omega \in \Omega$, the Least-Squares approximation of S in the space $\mathcal{K}$ with respect to ν_M is the $\mathbb{P}-a.e.$ unique $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable function $S^\star$ given by:

$$S^\star(\omega,\cdot) = \arg\inf_{\phi\in\mathcal{K}} \frac{1}{M}\sum_{m=1}^M |\phi(\chi^{(m)}(\omega)) - S(\omega, \chi^{(m)}(\omega))|^2.$$

Informally, relying on the Markovian framework, we wish to use LLSR to obtain approximations of the solutions of the BSDE at time step τ in the form of closed-loop feedback with respect to the current value of the extended state variable $(\mathbf{x}^{\text{sun},h}_\tau, X^{(k)}_\tau)$, i.e., we wish to determine $\Phi^{(k)}_{Y,\tau}$, $\Phi^{(k)}_{P,\tau}$, $\Phi^{(k)}_{\Pi,\tau}$ to obtain estimates of the form:

$$Y^{(k)}_\tau \simeq \Phi_{Y,\tau}(\mathbf{x}^{\text{sun},h}_\tau, X^{(k)}_\tau); \quad P^{(k)}_t \simeq \Phi_{P,\tau}(\mathbf{x}^{\text{sun},h}_\tau, X^{(k)}_t); \quad \Pi^{(k)}_\tau \simeq \Phi_{\Pi,\tau}(\mathbf{x}^{\text{sun},h}_\tau, X^{(k)}_\tau).$$

We introduce the notation $\nu^{(k)}_{\tau,[M]} = \frac{1}{M}\sum_{m=1}^M \delta_{\mathbf{x}^{\text{sun},h}_{[\tau:N_T],m}, u^{(k)}_{[\tau:N_T],m}, X^{(k)}_{[\tau:N_T],m}}$ for the empirical measure. We define $\mathcal{K}_\tau$ as the 3-dimensional vector space of functions spanned by $(\phi_{\tau,1}, \phi_{\tau,2}, \phi_{\tau,3})$ taking as arguments $(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) \in \mathbb{R}^{3(N_T-\tau+1)}$ and returning respectively 1, z_τ and x_τ. Hence ϕ_1 spans the vector space of constant functions, while ϕ_2 and ϕ_3 span the vector space of linear functions depending only on the two state variables at time τ.

We could consider more features in the function space $\mathcal{K}_\tau$ to allow more accurate functional representation. This is left for further investigation.

Ensuring the respect of a priori bounds for the solution of Riccati BSDE

In addition to this, we truncate $P^{(k)}$ at each iteration, using a priori upper and lower bounds. This helps to stabilize the LLSR algorithm [GT16b]. Let us introduce P^{UB} and P^{LB} the unique solutions of the Riccati ordinary differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}P^{UB}_t = \frac{(P^{UB}_t)^2}{(\lambda+\mu+\delta)\mathcal{E}^2_{\max}} - \nu, \qquad P^{UB}_T = \rho$$
$$\frac{\mathrm{d}}{\mathrm{d}t}P^{LB}_t = \frac{(P^{LB}_t)^2}{(\lambda+\mu-\delta)\mathcal{E}^2_{\max}} - \nu, \qquad P^{LB}_T = \rho.$$

Note that both equations can be easily solved analytically and numerically using similar arguments as when computing P^Δ. Using the (uniform) bounds:

$$\forall x \in \mathbb{R}, \quad \lambda - \delta \leq \mathcal{L}''(x) \leq \lambda + \delta,$$

it can be show by comparison principle for BSDEs in general filtrations [ØZ12, Theorem 3.4, p. 710] that

$$P^{LB}_t \leq P^{(k)}_t \leq P^{UB}_t, \tag{5.4.14}$$

for all $t \in [0, T]$, $\mathrm{d}\mathbb{P}$-a.e. We use this to truncate the numerical approximation of $P^{(k)}$.

Other parameters and implementation details

We choose the parameter values $\gamma = \beta = 0.1$. The parameter γ is linked to the acceptance rate of the (possibly reduced) Newton step, and 0.1 is a good trade-off between the need of sufficient reduction and a high acceptance rate. The choice $\beta = 0.1$ ensures that when sufficient reduction is not achieved, the step is sufficiently reduced to provide an acceptable step length with high probability.

The algorithms considered

We implement and compare Newton's method combined with two backtracking line-search: the standard backtracking line-search Algorithm 5.3 and the Gradient Backtracking line-search 5.4 designed in this chapter. To estimate the mean value of a random variable given an i.i.d. samples of this random variable, we use the empirical mean, which is an unbiased estimator. To estimate $\|X\|_{\mathbb{H}^{\infty,\infty}}$ given i.i.d. sample trajectories $(X_{\tau,m})_{\tau\in[N_T],m\in[M]}$, we use the estimator $\sup_{\tau\in[N_T],m\in[M]} X_{\tau,m}$, which is lower biased (neglecting the impact of time discretization on the bias). More accurate estimators based on extreme-value theory could be used, see for instance [ANR17]. Practical implementations of Algorithms 5.3, 5.4 and 5.5 are respectively given by Algorithms 5.6, 5.7 and 5.8. We use the initial guess $u^{(0)} = 0$ for Newton method. However, using the easily computable solution of the linear quadratic problem obtained by replacing the non-quadratic loss $\mathcal{L}$ by the quadratic loss function $\mathcal{L}^{\text{quad}} : x \mapsto \frac{\lambda x^2}{2}$ could allow to find a better initial guess (warm start). Though we do not show the results of such a procedure, numerical experiments show that this amounts to reduce by 1 the number of Newton iterations required to obtain a given accuracy.

On the stopping criteria of the Newton method

Ideally, the stopping criteria of the Newton method with Gradient Backtracking line-search should be $\|\nabla\bar{\mathcal{J}}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \varepsilon$. However, the norm of the gradient as estimated is erroneous, due to discretization and regression errors, and shold be estimated on a test set, distinct from the training set used in the algorithm. Hence, finding a relevant stopping criteria is a difficult task and left for further investigation. In practice, we shall replace the "while" loop by a "for" loop with a fixed number of iterations, and monitor the estimated norm of the gradient along iterations.

Algorithm 5.6 Standard Backtracking line search with Linear Least Square Regression

Inputs: Current control: $(u^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, current state $(X^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\dot{u}^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ Newton direction, $(\hat{Y}^{(k)}_{\tau})_{\tau\in[N_T]}$ regression functions for adjoint variable $Y^{(k)}$, $(\beta,\gamma)\in(0,1)^2$, M trajectories of solar irradiance $(\mathbf{X}^{\mathrm{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.
$\sigma = 1$.
repeat
  $u^{(k+1)} = u^{(k)} + \sigma\dot{u}^{(k)}$.
  Compute $(X^{(k+1)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (5.4.8).
  $\sigma \leftarrow \beta\sigma$.
  {*Computation of discretized gradient* (5.4.10).}
  $\nabla\mathcal{J}^{(k)}_{\tau,m} = \mu u^{(k)}_{\tau,m} + \mathcal{L}'\left(u^{(k)}_{\tau,m} + \mathbf{X}^{\mathrm{sun},h}_{\tau,m}\mathrm{P}^{\mathrm{sun,max}}_{\tau h} - \mathbb{E}\left[\mathrm{P}^{\mathrm{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k)}_{\tau}(\mathbf{X}^{\mathrm{sun},h}_{\tau,m}, u^{(k)}_{\tau,m}, X^{(k)}_{\tau,m})$.
  {*Computation of cost function.*}
  $\mathcal{J}^{(k+1)}_{m} = \sum_{\tau=1}^{N_T}\left(\frac{\mu}{2}(u^{(k+1)}_{\tau,m})^2 + \frac{\nu}{2}\left(X^{(k+1)}_{\tau,m} - \frac{1}{2}\right)^2 + \mathcal{L}\left(u^{(k+1)}_{\tau,m} + \mathbf{X}^{\mathrm{sun},h}_{\tau,m}\mathrm{P}^{\mathrm{sun,max}}_{\tau h} - \mathbb{E}\left[\mathrm{P}^{\mathrm{sun}}_{\tau h}\right]\right)\right)h + \frac{\rho}{2}\left(X^{(k+1)}_{N_T,m} - \frac{1}{2}\right)^2$.
  $\mathcal{J}^{(k)}_{m} = \sum_{\tau=1}^{N_T}\left(\frac{\mu}{2}(u^{(k)}_{\tau,m})^2 + \frac{\nu}{2}\left(X^{(k)}_{\tau,m} - \frac{1}{2}\right)^2 + \mathcal{L}\left(u^{(k)}_{\tau,m} + \mathbf{X}^{\mathrm{sun},h}_{\tau,m}\mathrm{P}^{\mathrm{sun,max}}_{\tau h} - \mathbb{E}\left[\mathrm{P}^{\mathrm{sun}}_{\tau h}\right]\right)\right)h + \frac{\rho}{2}\left(X^{(k)}_{N_T,m} - \frac{1}{2}\right)^2$.
until $\frac{1}{M}\sum_{m=1}^{M}\mathcal{J}^{(k+1)}_{m} \leq \frac{1}{M}\sum_{m=1}^{M}\left(\mathcal{J}^{(k)}_{m} + \gamma\sigma\sum_{\tau=1}^{N_T}\nabla\mathcal{J}^{(k)}_{\tau,m}\dot{u}^{(k)}_{\tau,m}h\right)$ {*Sufficient decrease of cost*}
$\sigma \leftarrow \sigma/\beta$ {*Correction of σ which has been reduced one too many times.*}
for $\tau = N_T, .., 1$ **do**
  Define the empirical measure $\nu^{(k+1)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathbf{X}^{\mathrm{sun},h}_{[\tau:N_T],m}, u^{(k+1)}_{[\tau:N_T],m}, X^{(k+1)}_{[\tau:N_T],m}}$.
  {*Regression of adjoint variable. See* (5.4.9) *and Definition* 5.4.1.}
  Compute $\hat{Y}^{(k+1)}_{\tau}$ solution of **OLS**$(S_{Y_\tau}, \mathcal{K}_\tau, \nu^{(k+1)}_{\tau,[M]})$ with $S_{Y_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho(x_{N_T} - 1/2) + \sum_{j=\tau+1}^{N_T}\nu(x_j - 1/2)h$.
end for
return $u^{(k+1)}$, $X^{(k+1)}$, $\hat{Y}^{(k+1)}$.

Algorithm 5.7 Gradient Backtracking line search with Linear Least Square Regression

Inputs: $(u^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(X^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\dot{u}^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\hat{Y}^{(k)}_{\tau})_{\tau\in[N_T]}$, $(\beta,\gamma)\in(0,1)^2$, M trajectories of solar irradiance $(\mathbf{X}^{\mathrm{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.
$\sigma = 1$.
repeat
  $u^{(k+1)} = u^{(k)} + \sigma\dot{u}^{(k)}$.
  Compute $(X^{(k+1)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (5.4.8).
  for $\tau = N_T, .., 1$ **do**
    Define the empirical measure $\nu^{(k+1)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathbf{X}^{\mathrm{sun},h}_{[\tau:N_T],m}, u^{(k+1)}_{[\tau:N_T],m}, X^{(k+1)}_{[\tau:N_T],m}}$.
    {*Adjoint variable regression. See* (5.4.9) *and Definition* 5.4.1.}
    Compute $\hat{Y}^{(k+1)}_{\tau}$ solution of **OLS**$(S_{Y_\tau}, \mathcal{K}_\tau, \nu^{(k+1)}_{\tau,[M]})$ with $S_{Y_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho(x_{N_T} - 1/2) + \sum_{j=\tau+1}^{N_T}\nu(x_j - 1/2)h$.
  end for
  $\sigma \leftarrow \beta\sigma$
  {*Computation of discretized gradient* (5.4.10).}
  $\nabla\mathcal{J}^{(k+1)}_{\tau,m} = \mu u^{(k+1)}_{\tau,m} + \mathcal{L}'\left(u^{(k+1)}_{\tau,m} + \mathbf{X}^{\mathrm{sun},h}_{\tau,m}\mathrm{P}^{\mathrm{sun,max}}_{\tau h} - \mathbb{E}\left[\mathrm{P}^{\mathrm{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k+1)}_{\tau}(\mathbf{X}^{\mathrm{sun},h}_{\tau,m}, u^{(k+1)}_{\tau,m}, X^{(k+1)}_{\tau,m})$.
  $\nabla\mathcal{J}^{(k)}_{\tau,m} = \mu u^{(k)}_{\tau,m} + \mathcal{L}'\left(u^{(k)}_{\tau,m} + \mathbf{X}^{\mathrm{sun},h}_{\tau,m}\mathrm{P}^{\mathrm{sun,max}}_{\tau h} - \mathbb{E}\left[\mathrm{P}^{\mathrm{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k)}_{\tau}(\mathbf{X}^{\mathrm{sun},h}_{\tau,m}, u^{(k)}_{\tau,m}, X^{(k)}_{\tau,m})$.
until $\max_{\tau\in[N_T],m\in[M]}\left|\nabla\mathcal{J}^{(k+1)}_{\tau,m}\right| \leq (1-\gamma\sigma)\left|\max_{\tau\in[N_T],m\in[M]}\left|\nabla\mathcal{J}^{(k)}_{\tau,m}\right|\right.$ {*Sufficient decrease condition*}
$\sigma \leftarrow \sigma/\beta$ {*Correction of σ which has been reduced one too many times.*}
return $u^{(k+1)}$, $X^{(k+1)}$, $\hat{Y}^{(k+1)}$.

Algorithm 5.8 Newton method with Least-Square Regression and backtracking line search

Initialization: M trajectories of solar irradiance $(\mathtt{X}^{\mathrm{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.
Set $k=0$.
Set $(u^{(0)}_{\tau,m})_{\tau\in[N_T],m\in[M]}=0$.
Compute $(X^{(0)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (5.4.8).
for $\tau=N_T,..,1$ **do**
  {*Regression function of adjoint state variable, see* (5.4.9) *and Definition* 5.4.1.}
  Define the empirical measure $\nu^{(0)}_{\tau,[M]}:=\frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathrm{sun},h}_{[\tau:N_T],m},u^{(0)}_{[\tau:N_T],m},X^{(0)}_{[\tau:N_T],m}}$.
  $\hat{Y}^{(0)}_{\tau}$ solution of **OLS**$(S_{Y_\tau},\mathcal{K}_\tau,\nu^{(0)}_{\tau,[M]})$ with $S_{Y_\tau}(z_{[\tau:N_T]},v_{[\tau:N_T]},x_{[\tau:N_T]})=\rho(x_{N_T}-1/2)+\sum_{j=\tau+1}^{N_T}\nu(x_j-1/2)h$.
end for
while Stopping criteria not met **do**
  for $\tau=N_T,...,1$ **do**
    {*Regression functions for* $P^{(k)}$ *and* $\Pi^{(k)}$ *solutions of* (5.4.11) *and* (5.4.12). *See Definition* 5.4.1. *Truncation of the estimator of* $P^{(k)}$ *to verify the a priori bounds* (5.4.14).}
    Define the empirical measure $\nu^{(k)}_{\tau,[M]}:=\frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathrm{sun},h}_{[\tau:N_T],m},u^{(k)}_{[\tau:N_T],m},X^{(k)}_{[\tau:N_T],m}}$.
    $\hat{P}^{(k)}_{\tau}$ as the projection on the convex set $[P^{LB}_\tau,P^{UB}_\tau]$ of the solution of **OLS**$(S_{P^{(k)}_\tau},\mathcal{K}_\tau,\nu^{(k)}_{\tau,[M]})$ with
    $$S_{P^{(k)}_\tau}(z_{[\tau:N_T]},v_{[\tau:N_T]},x_{[\tau:N_T]})=\rho+\sum_{j=\tau+1}^{N_T}\left(-\frac{1}{\mathcal{E}^2_{\max}\left\{\mu+l''\left(z_j,v_j\right)\right\}}\left(\hat{P}^{(k)}_j(z_j,x_j)\right)^2+\nu\right)h.$$
    $\hat{\Pi}^{(k)}_{\tau}$ solution of **OLS**$(S_{\Pi^{(k)}_\tau},\mathcal{K}_\tau,\nu^{(k)}_{\tau,[M]})$ with $S_{\Pi^{(k)}_\tau}(z_{[\tau:N_T]},v_{[\tau:N_T]},x_{[\tau:N_T]})$ given by
    $$\sum_{j=\tau+1}^{N_T}\left(-\frac{\hat{P}^{(k)}_j(z_j,x_j)}{\mathcal{E}^2_{\max}\left\{\mu+l''_j(z_j,v_j)\right\}}\hat{\Pi}^{(k)}_j(z_j,x_j)+\frac{\left(\mu v_j+l'_j(z_j,v_j)\right)\mathcal{E}_{\max}-\hat{Y}^{(k)}_j(z_j,x_j)}{\mathcal{E}^2_{\max}\left\{\mu+l''_j(z_j,v_j)\right\}}\hat{P}^{(k)}_j(z_j,x_j)\right)h,$$
    where we used the notations $l''_j(z,v):=\mathcal{L}''\left(v+z\mathtt{P}^{\mathrm{sun,max}}_{jh}-\mathbb{E}\left[\mathtt{P}^{\mathrm{sun}}_{jh}\right]\right)$ and $l'_j(z,v):=\mathcal{L}'\left(v+z\mathtt{P}^{\mathrm{sun,max}}_{jh}-\mathbb{E}\left[\mathtt{P}^{\mathrm{sun}}_{jh}\right]\right)$
  end for
  {*Computation of Newton step by feedback expression* (5.4.13).}
  for $m=1\in[M]$ **do**
    $\dot{X}^{(k)}_{1,m}=0$.
    for $\tau=1,..N_T$ **do**
      Denote $\hat{P}^{(k)}_{\tau,m}=\hat{P}^{(k)}_{\tau}(\mathtt{X}^{\mathrm{sun},h}_{\tau,m},X^{(k)}_{\tau,m})$, $\hat{\Pi}^{(k)}_{\tau,m}=\hat{\Pi}^{(k)}_{\tau}(\mathtt{X}^{\mathrm{sun},h}_{\tau,m},X^{(k)}_{\tau,m})$ and $\hat{Y}^{(k)}_{\tau,m}=\hat{Y}^{(k)}_{\tau}(\mathtt{X}^{\mathrm{sun},h}_{\tau,m},X^{(k)}_{\tau,m})$.
      $$\dot{u}^{(k)}_{\tau,m}=\frac{\hat{P}^{(k)}_{\tau,m}\dot{X}^{(k)}_{\tau,m}+\hat{\Pi}^{(k)}_{\tau,m}+\hat{Y}^{(k)}_{\tau,m}}{\mathcal{E}_{\max}\left\{\mu+\mathcal{L}''\left(u^{(k)}_{\tau,m}+\mathtt{X}^{\mathrm{sun},h}_{\tau,m}\mathtt{P}^{\mathrm{sun,max}}_{\tau h}-\mathbb{E}\left[\mathtt{P}^{\mathrm{sun}}_{\tau h}\right]\right)\right\}}-\frac{\mu u^{(k)}_{\tau,m}+\mathcal{L}'\left(u^{(k)}_{\tau,m}+\mathtt{X}^{\mathrm{sun},h}_{\tau,m}\mathtt{P}^{\mathrm{sun,max}}_{\tau h}-\mathbb{E}\left[\mathtt{P}^{\mathrm{sun}}_{\tau h}\right]\right)}{\mu+\mathcal{L}''\left(u^{(k)}_{\tau,m}+\mathtt{X}^{\mathrm{sun},h}_{\tau,m}\mathtt{P}^{\mathrm{sun,max}}_{\tau h}-\mathbb{E}\left[\mathtt{P}^{\mathrm{sun}}_{\tau h}\right]\right)}$$
      $\dot{X}^{(k)}_{\tau+1,m}=\dot{X}^{(k)}_{\tau,m}-\frac{1}{\mathcal{E}_{\max}}\dot{u}^{(k)}_{\tau,m}h$.
    end for
  end for
  Backtracking line search to get $u^{(k+1)}$, $X^{(k+1)}$ and $\hat{Y}^{(k+1)}$.
  $k\leftarrow k+1$.
end while
return $(u^\star,X^\star,\hat{Y}^\star)$:=$(u^{(k)},X^{(k)},\hat{Y}^{(k)})$.

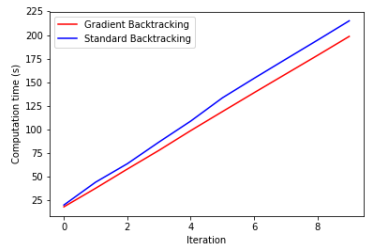
5.4.4 Analysis of the numerical performance

Figure 5.6b shows that initially both backtracking line-search methods return full Newton steps, which suggests that our initial guess $u^{(0)}$ is located in the quadratic convergence area for the Newton method. However, from iteration 3, the standard Backtracking line-search takes ridiculously small step lengths, as $\sigma = \beta^{13} = 10^{-13}$. Hence, the method fails to converge, as is suggest by Figures 5.6c and 5.6d which show that the $\mathbb{H}^{\infty,\infty}$ and $\mathbb{H}^{2,2}$ norm of the gradient is stationary from iteration 3. This shows that the Standard Backtracking line-search is not adapted to our setting.

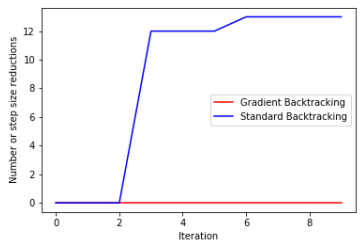
On the other hand, Figures 5.6c and 5.6d suggest that the Newton method with Gradient Backtracking line-search converges, as the norm of the gradient (as considered in $\mathbb{H}^{2,2}$ and $\mathbb{H}^{\infty,\infty}$) decreases along iterations. Hence, this shows that the Gradient Backtracking line-search procedure is better suited for our application than the (naive) standard backtracking line-search method.

Moreover, we would expect theoretically that $\left(\|\nabla\bar{\mathcal{J}}(\bar{u}^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ decreases quadratically fast. However, this is not the case, see Figure 5.6c: after the third iteration, the convergence is not quadratic anymore, although the algorithm takes full steps $\sigma = 1$ at all iterations, see Figure 5.6b. We believe this comes from the regression steps, which introduce some residual errors in the computations of $\bar{Y}^{(k)}$, $P^{(k)}$ and $\Pi^{(k)}$.

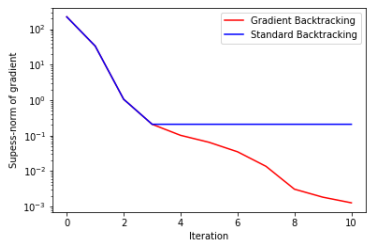
The cost decreases quickly for the first iterations for the Newton method combined with both Backtracking line-search procedures, see Figure 5.6e. From iteration 3, the Newton method with standard backtracking line-search does not make any progress (the step size is ridiculously small), while it is no longer decreasing for the Gradient Backtracking line-search, see Figure 5.6f. This is not surprising as our result states the convergence of the norm of the gradient (in $\mathbb{H}^{\infty,\infty}$) to 0, and not that the cost is decreasing along iterations. Besides, the number of samples considered $M = 10000$ may explain this non-monotonic behavior of the cost along iterations of the Newton method gradient backtracking line-search.



(a) Running computation time (in seconds)

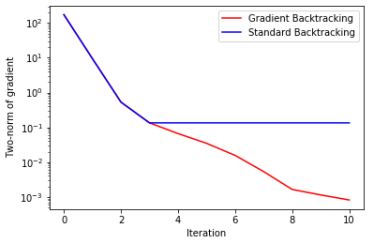


(b) Number of step size reductions by the backtracking line-search

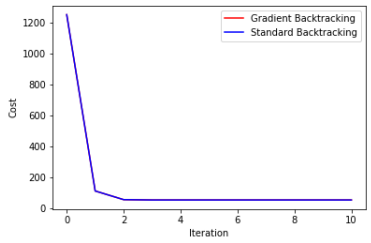


(c) Numerical approximation of $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}$ along iterations

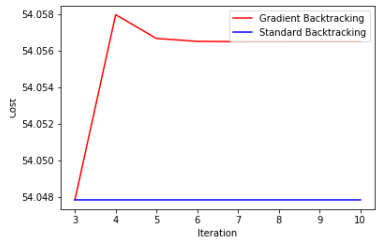
//



(d) Numerical approximation of $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{2,2}}$ along iterations



(e) Cost along iterations



(f) Cost along iterations from iteration 3

Figure 5.6: Comparison of performances of Newton method with the two Backtracking line search methods

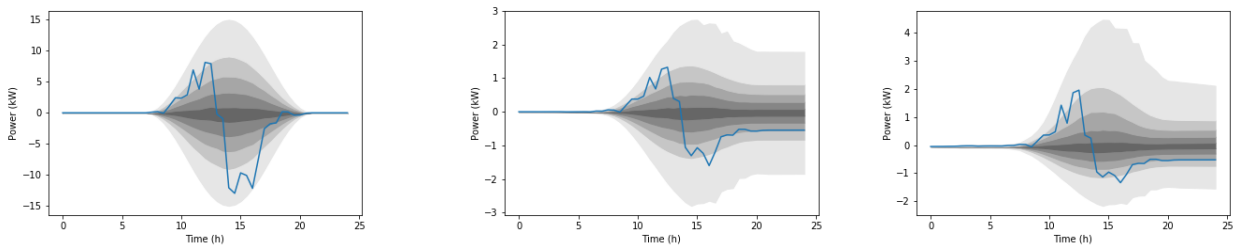
5.4.5 Over-fitting, under-fitting, and automatic tuning of regression parameters by cross-validation

Over-fitting or under-fitting may occur in the regression steps and could be dealt with by introducing validation steps (by splitting the sample into a training set and a validation set) allowing automatic parameter tuning at each

step. As examples, one could consider introducing regularization or use other functional spaces. Fine-tuning of the regression parameters by cross-validation would allow lower generalization errors, but at the expenses of a higher computational cost. Moreover, there are many heterogeneous regression steps, each of them requiring a particular treatment: good regression parameters have no reason to be the same for $Y^{(k)}$, $P^{(k)}$ and $\Pi^{(k)}$ and may change according to the time step considered and along the iterations. To give one example, one could consider regressing against the state variable $\bar{X}^{(k)}$ instead of $(\bar{X}^{(k)}, \mathtt{X}^{\mathrm{sun}})$ for the regression steps of processes at time steps after the sun set, because the solar irradiance does not play any role in the problem as it is canceled out by $\mathtt{P}^{\mathrm{sun,max}}$ which is null after sunset. We do not focus on these over-fitting or under-fitting issues in this chapter, and simply consider a general regression procedure with a functional space spanned by affine-linear functions of the extended state $(\bar{X}^{(k)}, \mathtt{X}^{\mathrm{sun}})$. Incorporating cross-validation steps for automatic parameter tuning (functional space, regularization) in our algorithm is an interesting perspective of our work.

5.4.6 Analysis of the results from the application point of view

For completeness, we give some brief comments on the numerical results, from the point of view of the application. Figure 5.7a represents the evolution of the power imbalance without the control mechanism. Figure 5.7b represents the evolution of the power imbalance with a quadratic (symmetric) loss function $\mathcal{L}^{\mathrm{quad}} : x \mapsto \lambda x^2$, which gives a linear quadratic structure to the coordination sub-problem, which makes it particularly easy to solve. Figure 5.7c represents the evolution of the power imbalance with the asymmetric loss function $\mathcal{L}$. By comparing the uncontrolled case with the two controlled case, we can see that the power imbalance range has been significantly reduced (noticing the change of scale of the graphs). This shows efficiency of the proposed control mechanism to reduce the power imbalance. The asymmetric loss function $\mathcal{L}$ tends to penalize more heavily negative imbalance than positive imbalance, which creates an asymmetry in the probability distribution of the power imbalance, see Figure 5.7c, to be compared with the symmetry of the probability distribution of the power imbalance in the case of a symmetric loss function $\mathcal{L}^{\mathrm{quad}}$, see Figure 5.7b. For all plots, we add the realization of the power imbalance for one scenario of solar irradiance (the same scenario as the one plotted in Figure 5.3b). The power imbalance is null at night in the uncontrolled case, as there is no solar production. There is no power imbalance in the controlled case before sunrise in the controlled case. However, for some scenarios, there is a non-zero power imbalance in the controlled case, which arises from the fact that the state of charge of the batteries at sunset might be far away from its target terminal value $1/2$. Hence, the batteries are used after sunset in this case in order to take into account the target terminal value of the state of charge.



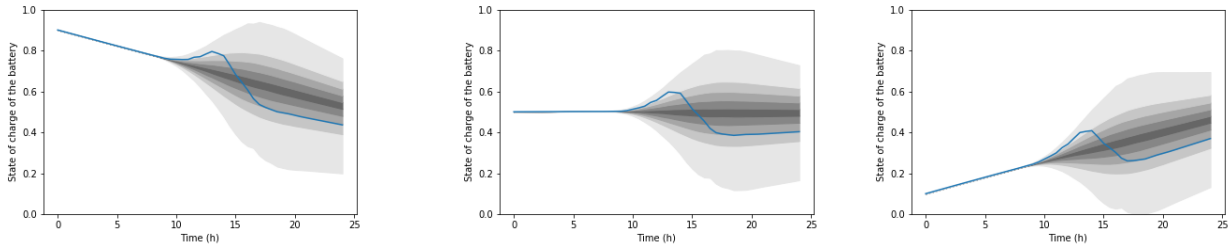
(a) Uncontrolled power imbalance $\mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$

(b) Controlled power imbalance $\bar{u} + \mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$ - symmetric loss $\mathcal{L}^{\mathrm{quad}}$

(c) Controlled power imbalance $\bar{u} + \mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$ - asymmetric loss $\mathcal{L}$

Figure 5.7: Power imbalance

One may wonder if the control mechanism proposed respects the constraint on the states of charge of the batteries, which must lie in $[0, 1]$, even for an initial state of charge close to 0 or 1. The quantiles of the state of charge of one of the batteries participating to the control mechanism is plotted in Figure 5.8, depending on the initial value of the state of charge. One can in particular see that, even for initial value of the state of charge close to 0 or 1, the state of charge of the battery remains between these two values with high probability. For all plots, we add the realization of the state of charge of the battery considered for one scenario of solar irradiance (the same scenario as the one plotted in Figure 5.3b).



(a) $x_0 = 0.9$ (b) $x_0 = 0.5$ (c) $x_0 = 0.1$

Figure 5.8: Percentiles of the state of charge of a battery participating to the control mechanism - Sensitivity to initial condition

5.5 Proofs

5.5.1 Proof of Lemma 5.3.5

Proof. For $M > 0$ we introduce:

$$\begin{cases} \overline{f}(t,p) := 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t, \\ f^{(M)}(t,p) := C_M\left(2A_t p + B_t p^2 + C_t\right), \\ \underline{f}(t,p) := -2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| - \|B\|_{\mathbb{H}^{\infty,\infty}}p^2, \end{cases}$$

where the clipping operator C_M is defined by $C_M : x \in \mathbb{R} \mapsto \min(\max(x,-M),M) = \max(\min(x,M),-M) \in [-M,M]$. Notice that for any $p \in \mathbb{R}$ and any $M > 0$, by our assumptions, we have $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. on $[0,T]$:

$$\begin{aligned} \underline{f}(t,p) &= \min(\underline{f}(t,p), M) && (M > 0, \underline{f}(t,p) \le 0) \\ &\le \min(2A_t p + B_t p^2 + C_t, M) && \text{(Monotony of min)} \\ &\le C_M\left(2A_t p + B_t p^2 + C_t\right) && (\forall x \in \mathbb{R}, \min(x,M) \le C_M(x)) \\ &= f^{(M)}(t,p) \\ &\le \max\left(-M, 2A_t p + B_t p^2 + C_t\right) && (\forall x \in \mathbb{R}, C_M(x) \le \max(-M,x)) \\ &\le \max\left(-M, 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t\right) && \text{(Monotony of max}, B \le 0) \\ &= \overline{f}(t,p) && (-M \le 0 \le C_t \le 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t). \end{aligned}$$

Consider the BSDEs:

$$\underline{P}_t := \int_t^T \underline{f}(s,\underline{P}_s)\mathrm{d}s, \tag{5.5.1}$$

$$P_t^{(M)} := \mathbb{E}_t\left[\Gamma + \int_t^T f^{(M)}(s,P_s^{(M)})\mathrm{d}s\right], \tag{5.5.2}$$

$$\overline{P}_t := \mathbb{E}_t\left[\Gamma + \int_t^T \overline{f}(s,\overline{P}_s)\mathrm{d}s\right]. \tag{5.5.3}$$

First, the BSDE (5.5.1) is actually an ODE with locally Lipschitz-continuous driver, and therefore, by Cauchy-Lipschitz theorem has a unique solution on some maximal interval $(\tau, T]$. The null function is clearly the unique solution of (5.5.1) and $\tau = -\infty$. This yields the well-posedness of $\underline{P}$ and the explicit expression $\underline{P}_t = 0, \forall t \in [0,T]$.

Second, we notice that (5.5.2) and (5.5.3) are BSDE with Lipschitz drivers, so that they are well-defined on $[0,T]$, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. Notice that $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e., for any $p \in \mathbb{R}$,

$$0 \le \Gamma \quad ; \quad \underline{f}(t,p) \le f^{(M)}(t,p) \le \overline{f}(t,p).$$

By comparison theorem for BSDEs (see [ØZ12, Theorem 3.4, p. 710].), we get for all $M > 0$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. on $[0, T]$:

$$0 = \underline{P}_t \leq P_t^{(M)} \leq \overline{P}_t.$$

Besides, $\overline{P}$ satisfies for some martingale $M \in \mathcal{M}_0$:

$$\begin{cases} -\mathrm{d}\overline{P}_t = \left(2\|A\|_{\mathbb{H}^{\infty,\infty}}\overline{P}_t + C_t\right)\mathrm{d}t - \mathrm{d}M_t, \\ \overline{P}_T = \Gamma. \end{cases} \tag{5.5.4}$$

Using the Integration by Parts formula in [Pro03, Corollary 2, p. 68] to $t \mapsto \overline{P}_t \exp\left(-2\|A\|_{\mathbb{H}^{\infty,\infty}}(T-t)\right)$ yields the explicit expression (5.3.5) and the estimation on $\|P\|_{\mathbb{H}^{\infty,\infty}}$. Now notice that for $M > M_0 := 2\|A\|_{\mathbb{H}^{\infty,\infty}}\|\overline{P}\|_{\mathbb{H}^{\infty,\infty}} + \|B\|_{\mathbb{H}^{\infty,\infty}}\|\overline{P}\|^2_{\mathbb{H}^{\infty,\infty}} + \|C\|_{\mathbb{H}^{\infty,\infty}}$, we have:

$$\forall p \in [0, \|\overline{P}\|_{\mathbb{H}^{\infty,\infty}}], \mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad f^{(M)}(t,p) = 2A_t p + B_t p^2 + C_t. \tag{5.5.5}$$

Since $0 \leq P_t^{(M)} \leq \overline{P}_t$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e., we get, for $M > M_0$:

$$P_t^{(M)} = \mathbb{E}_t\left[\Gamma + \int_t^T \left(2A_s P_s^{(M)} + B_s\left(P_s^{(M)}\right)^2 + C_s\right)\mathrm{d}s\right].$$

This shows existence of solution of the Riccati BSDE (5.3.4).

Let us now turn to uniqueness. Consider two solutions P and Q of the Riccati BSDE (5.3.4). By application of the comparison principle for BSDEs, we obtain with similar arguments as before:

$$\mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad 0 \leq P_t \leq \overline{P}_t, \qquad 0 \leq Q_t \leq \overline{P}_t,$$

and therefore, for $M > M_0$, (5.5.5) shows that P and Q are both solutions of the BSDE (5.5.2), which has Lipschitz driver, hence a unique solution, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. In particular, $P = Q$, which yields uniqueness of solutions of (5.3.4). □

5.5.2 Proof of Lemma 5.3.6

Proof. Define Π by (5.3.6) and define as well:

$$R_t := \exp\left(\int_0^t (P_s B_s + A_s)\mathrm{d}s\right)\Pi_t + \int_0^t (a_s P_s + b_s)\exp\left(\int_0^s (P_r B_r + A_r)\mathrm{d}r\right)\mathrm{d}s = \mathbb{E}_t\left[R_T\right],$$
$$S_t := \exp\left(-\int_0^t (P_s B_s + A_s)\mathrm{d}s\right).$$

Then $R \in \mathbb{H}^{\infty,2}$ and R is an $(\mathcal{F}_t)$-adapted càdlàg martingale. Then, apply integration by parts formula [Pro03, Corollary 2, p. 68] to the product SR, using the fact that S is continuous with finite variations. After reorganizing terms and using that R has countable jumps, we find:

$$\begin{cases} -\mathrm{d}\Pi_t = \left((P_t B_t + A_t)\Pi_t + a_t P_t + b_t\right)\mathrm{d}t + S_t \mathrm{d}R_t, \\ \Pi_T = \eta, \end{cases}$$

and $\int_{0^+}^t S_s \mathrm{d}R_s$ is a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [Pro03, Theorem 20 p.63, Corollary 3 p.73, Theorem 29 p.75]. We then find that Π solves (5.3.7) and it is the unique solution of this BSDE, as the BSDE has a Lipschitz driver, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. □

5.5.3 Proof of Theorem 5.3.7

Proof. 1. FBSDEs have been studied in the general case in [Zha17] and [MY99]. In the affine-linear case, the result is a consequence of [Yon06] and the Martingale Representation Theorem if the filtration is Brownian. However, for more general filtrations, the result is outside the scope of [Yon06] and we provide a proof for this case, restricting ourselves to one-dimensional control and state processes.

Consider the following auxiliary linear-quadratic stochastic control problem:

$$\left.\begin{aligned} \mathcal{J}^{quad}(u) &:= \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}u_t^2 + \tfrac{1}{2}C_t X_t^2 + b_t X_t\right)\mathrm{d}t + \tfrac{1}{2}\Gamma X_T^2 + \eta X_T\right] \\ \text{s.t. } X_t &= x + \int_0^t (A_s X_s + \sqrt{-B_s}u_s + a_s)\mathrm{d}s. \end{aligned}\right\} \longrightarrow \min_{u\in\mathbb{H}^{2,2}_{\mathcal{P}}}. \tag{5.5.6}$$

Our assumptions show that $\mathcal{J}^{quad}$ satisfies the hypothesis of Proposition 5.2.3 and therefore, it has a unique minimizer $u^* \in \mathbb{H}^{2,2}$.

The function $\mathcal{J}^{quad}$ also satisfies the assumptions of first order sufficient optimality conditions (see second point of Theorem 5.2.6), so that, if we define $(X^*, Y^*) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ by:

$$\begin{cases} X_t^* = x + \int_0^t (A_s X_s^* + \sqrt{-B_s}u_s^* + a_s)\mathrm{d}s, \\ Y_t^* = \mathbb{E}_t\left[\Gamma X_T^* + \eta + \int_t^T (C_s X_s^* + A_s Y_s^* + b_s)\,\mathrm{d}s\right], \end{cases}$$

we have

$$u_t^* + \sqrt{-B_t}Y_{t-}^* = 0.$$

By eliminating u^* using the last equation and using the fact that the Lebesgue integral is left unchanged by changing the value of the integrand on a countable set, this shows that (X^*, Y^*) satisfies the FBSDE:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right], \end{cases}$$

Let us turn to uniqueness. Consider two solutions (X^1, Y^1) and (X^2, Y^2) of the above FBSDE. Then (u^1, X^1, Y^1) and (u^2, X^2, Y^2) with $u_t^i = -\sqrt{-B}Y_{t-}^i$ for $i = 1, 2$ are both solutions of:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + \sqrt{-B_s}u_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right], \\ u_t + \sqrt{-B_t}Y_{t-} = 0. \end{cases}$$

Hence u^1 and u^2 are both solutions of the first order conditions characterizing minimizers of $\mathcal{J}^{quad}$ by Theorem 5.2.6 and by Proposition (5.2.3), $u^1 = u^2$. This shows $X^1 = X^2$, then $Y^1 = Y^2$, hence the existence and uniqueness of a solution of the FBSDE.

2. By our previous results, P and Π are well defined in $\mathbb{H}^{\infty,\infty}$ and $\mathbb{H}^{\infty,2}$ respectively. Then X given in (5.3.8) solves an affine-linear ODE and the assumption on the coefficients show that it is well-defined (non-explosion) and given by:

$$\forall t \in [0,T], \quad X_t = x\exp\left(\int_0^t (A_s + B_s P_s)\mathrm{d}s\right) + \int_0^t (B_s \Pi_s + a_s)\exp\left(\int_s^t (A_r + B_r P_r)\mathrm{d}r\right)\mathrm{d}s.$$

The estimates on X and $Y = PX + \Pi$ in the spaces $\mathbb{H}^{\infty,2}$ come directly from that and Lemma 5.3.6. Let us now prove that (X, Y) is a solution of the affine-linear FBSDE, which will conclude the proof, by uniqueness of the solution of such FBSDE, by the previous point. Using $Y = PX + \Pi$, it is easy to show that X satisfies:

$$X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s.$$

It remains to show that Y satisfies the BSDE:

$$Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right].$$

To do that, use the fact that $Y = PX + \Pi$ by definition so that $Y_T = P_T X_T + \Pi_T = \Gamma X_T + \eta$ so that the terminal condition is verified. Introduce $M^{(P)}$ in $\mathcal{M}_0^2 \cap \mathbb{H}^{\infty,\infty}$ and $M^{(\Pi)}$ in $\mathcal{M}_0^2$ such that:

$$\begin{cases} -\mathrm{d}P_t = (2A_t P_T + B_t P_t^2 + C_T)\mathrm{d}t - \mathrm{d}M_t^{(P)}, \\ P_T = \Gamma, \end{cases}$$

and:

$$\begin{cases} -\mathrm{d}\Pi_t = ((P_t B_t + A_t)\Pi_t + a_t P_t + b_t)\,\mathrm{d}t - \mathrm{d}M_t^{(\Pi)}, \\ P_T = \eta. \end{cases}$$

Then, use the integration by parts formula in Protter [Pro03, Corollary 2, p. 68], combined with the fact that X is continuous with finite variations. We get:

$$\begin{aligned} -\mathrm{d}Y_t &= -(\mathrm{d}P_t)X_t - P_t(\mathrm{d}X_t) - \mathrm{d}\Pi_t \\ &= (2A_t P_T + B_t P_t^2 + C_T)X_t\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - P_t(A_t X_t + B_t Y_t + a_t)\mathrm{d}t + ((P_t B_t + A_t)\Pi_t + a_t P_t + b_t)\,\mathrm{d}t - \mathrm{d}M_t^{(\Pi)} \\ &= (A_t(P_t X_t + \Pi_t) + B_t P_t(P_t X_t + \Pi_t - Y_t) + C_t X_t + b_t)\,\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - \mathrm{d}M_t^{(\Pi)} \\ &= (A_t Y_t + C_t X_t + b_t)\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - \mathrm{d}M_t^{(\Pi)}. \end{aligned}$$

Using the fact that the last two terms are the increments of true martingales in $\mathcal{M}_0^2$ (as $X \in \mathbb{H}^{\infty,2}$ and $M^{(P)} \in \mathcal{M}_0^2 \cap \mathbb{H}^{\infty,\infty}$), this concludes the proof.

□

5.5.4 Proof of Corollary 5.3.8

Proof. It just remains to prove the estimates:

$$\begin{aligned} &\|P^u\|_{\mathbb{H}^{\infty,\infty}} \leq C, \\ &\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}} \leq C\|w\|_{\mathbb{H}^{2,2}}, \\ &\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}} \leq C\|w\|_{\mathbb{H}^{2,2}}, \end{aligned}$$

for some constant C independent of u and w. Using the assumption of bounded second-order derivative and the fact that l is strongly convex in u (implying that l''_{uu} is uniformly bounded from below by a non-negative constant), we get for a constant C independent from u and w:

$$\begin{aligned} &\|A^u\|_{\mathbb{H}^{\infty,\infty}} + \|B^u\|_{\mathbb{H}^{\infty,\infty}} + \|C^u\|_{\mathbb{H}^{\infty,\infty}} + \|\Gamma^u\|_{\mathbb{L}^\infty_\Omega} \leq C, \\ &\|a^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + \|b^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}}. \end{aligned}$$

We get the bounds on $\|P^u\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}}$ by using the estimates on $\|P\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\Pi\|_{\mathbb{H}^{\infty,2}}$ obtained in Lemmas 4.5.1 and 5.3.6, with $\eta = 0$ and $x = 0$. The bound on $\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}}$ is then obtained using the strong convexity of l with respect to u and using the expression (5.3.2). □

5.6 Conclusion

In this chapter, we extend the Newton method to the framework of stochastic control problems, which amounts to consider successive linearizations of the optimality system found by using the stochastic Pontryagin principle. We

show that the computation of the Newton step amounts to solve a linear FBSDE with random coefficients (with some sign conditions), which in turn reduces to solving a Riccati BSDE and a linear BSDE. Then, an appropriate restriction of the space of processes is considered to obtained desirable regularity for the control problem, allowing to prove convergence results for the Newton method. To obtain a global convergence, an appropriate line-search which fits our infinite-dimensional setting is proposed. Global convergence of the Newton method combined with this adapted line-search is then proved theoretically. The Newton method is implemented on a problem of joint control of many identical batteries in order to maintain power balance on a given network. In particular, regression techniques are used in order to compute the solutions of the linear and non-linear BSDEs arising when computing the Newton step. So far, we have considered low dimensional problems: the regression steps are performed in $\mathbb{R}^2$ and the control and state variables are one-dimensional. In higher dimension, we expect a curse of dimensionality when solving the BSDEs using regression method. Other methods like Deep-learning could help solve the issue. However, the Newton method is iterative and training a network at each iteration seems computationally expensive. We also expect the Newton method to be applicable to other settings, like controlled diffusions for instance, which would change the form of the Riccati BSDEs arising when solving the successive linearizations of the FBSDE characterizing the optimal control. Other interesting perspectives to our work include designing appropriate stopping criteria in the Newton method implemented using regression techniques, or incorporate automatic tuning procedures for the hyper-parameters in regression steps in the algorithm.

Part II

The Alternating Current Optimal Power Flow problem and its extension to the multi-stage stochastic setting

Chapter 6

Introduction to the Optimal Power Flow problem

6.1 General presentation of the OPF problem

The so-called Optimal Power Flow (OPF) problem is a mathematical optimization problem aiming at finding an operating point of a power network that minimizes a given objective function, such as generation costs, active power losses, subject to constraints on power injections and losses, voltage magnitudes and intensities in the lines. It has been introduced by Carpentier in 1962 [Car62] and has drawn a lot of attention since then thanks to its versatility.

Physically, the most accurate formulation of this problem is the so-called Alternating Current Optimal Power Flow (AC OPF). There are three common types of networks: single phase, balanced three-phase and unbalanced three-phase networks. Single phase networks are composed of lines with two wires: a ground wire, and a wire where power can flow. The balanced and unbalanced networks are composed of four wires: a ground wire, and three wires carrying power and associated to a phase. In the case of balanced networks, the sizes of all phases are the same and they are differ from each other by $120\ °$, which is not the case for unbalanced networks. The AC OPF problem for three-phase balanced networks reduces to the AC OPF problem with single phase. We focus on the case of single phase networks. It can be formulated as an optimization problem with complex variables, in order to take into account the oscillating behavior of the physical quantities in the network, operated in alternating current.

6.2 Surveys

Surveys on the OPF, related problems and perspectives can be found in [Cap+11; Cap16], while an overview of optimization methods to tackle the OPF problem is available in [FSR12a; FSR12b]. See [BCH14] for a survey on Chance-Constrained AC OPF, [MH+19] if one is interested in approximations and convex relaxations of AC OPF while [Zoh+20] focuses more specifically on conic relaxations of the AC OPF problem. A series of video tutorials on convex relaxations of AC OPF can be found in [CR18].

6.3 The Bus Injection Model (BIM): formulation and equivalence

There exist two main variants of the AC OPF problem for networks with a single phase: the Bus Injection Model (BIM), and the Branch Flow Model (BFM), both presented in [Low14a]. The Bus Injection Model focuses on nodal quantities, such as complex voltages and apparent power injections. The variables of the BIM models are given in Table 6.1. The notation $\mathbf{i}$ denotes the imaginary number such that $\mathbf{i}^2 = -1$.

Some useful physical quantities can be expressed in terms of complex voltages only, and these expressions are given in Table 6.2. We use the notation $z \leq_{\mathbb{C}} z'$ which means $\mathtt{Re}(z) \leq \mathtt{Re}(z')$ and $\mathtt{Im}(z) \leq \mathtt{Im}(z')$. Each line (i, j) is

Table 6.1: Decision variables for the Bus Injection Model

Variable	Voltage at bus i	Active/reactive power injection at bus i	Power injection at bus i
Notation	$V_i \in \mathbb{C}$	$p_i \in \mathbb{R}$ and $q_i \in \mathbb{R}$	$s_i = p_i + \mathbf{i}q_i \in \mathbb{C}$

characterized by a complex admittance denoted by $y_{i,j}$ (or equivalently an impedance $z_{i,j} = 1/y_{i,j}$) which gives its physical characteristics, i.e., the connection between current intensity in a line and voltage difference between its extremities. The *sending-end power* in line (i, j) at bus i is defined as the complex power leaving bus i to go through line (i, j). It is equal to the opposite of the sending-end power in (i, j) at bus j plus the power losses in line (i, j).

Table 6.2: Expression of some physical quantities in the Bus Injection Model

Physical quantities	Expression in terms of complex voltage
Intensity in line (i, j)	$y_{i,j}\left(V_i - V_j\right)$
Sending-in power in line (i, j) at bus i	$y^*_{i,j}\left(\lvert V_i\rvert^2 - V_i V^*_j\right)$
Power losses in line (i, j)	$y^*_{i,j}\left(\lvert V_i\rvert^2 + \lvert V_j\rvert^2 - 2\mathrm{Re}(V_i V^*_j)\right)$

The usual constraints of the optimal power flow problem, as well as their formulation in terms of complex voltages are given in Table 6.3.

Table 6.3: Constraints for the Bus Injection Model

Constraint	Expression
Voltage magnitude at bus i	$\underline{V}_i \leq \lvert V_i\rvert \leq \overline{V}_i$
Intensity magnitude in line i, j	$\lvert y_{i,j}\rvert\lvert V_i - V_j\rvert \leq \overline{I}_{i,j}$
Sending-in power magnitude in line i, j at bus i	$\lvert y_{i,j}\rvert\lvert V_i(V_i - V_j)^*\rvert \leq \overline{S}_{i,j}$
Power injection limits at bus i	$\underline{s}_i \leq_{\mathbb{C}} s_i \leq_{\mathbb{C}} \overline{s}_i$
Complex power conservation at bus i	$y^*_{i,i}\lvert V_i\rvert^2 + \sum_{(i,j)\in\mathcal{L}} y^*_{i,j}\left(\lvert V_i\rvert^2 - V_i V^*_j\right) = s_i$

The cost for the problem usually depends on the power injections at the buses of the network as well as on power losses in the network. There exists an equivalent reformulation of the BIM formulation [LL11, Theorem 1], where the complex voltage variables are replaced by the lifting variables $W_{i,j} := V_i V^*_j$, for all buses i, j. Equivalently, denoting $V = (V_i)_i \in \mathbb{C}^n$ the vector of complex voltages at all n buses, the lifting variable $W \in \mathbb{C}^{n\times n}$ is defined by $W = (W_{i,j})_{i,j} = VV^T$. To obtain an equivalent reformulation of the problem in terms of the lifting variable W, all constraints and terms in the cost should be reformulated in terms of variable W instead of V, and one should enforce additional constraints: a positive semi-definite condition and a rank constraint. This follows from the following fact:

$$(\exists V \in \mathbb{C}^n, W = VV^\top) \Leftrightarrow (W \succcurlyeq 0, \quad rank(W) = 1)\,.$$

The notation $W \succcurlyeq 0$ means that W is hermitian positive semi-definite, i.e., $W = W^*$, where W^* denotes the transposition of the conjugate of W, and $u^\top W u \in \mathbb{R}_+ \forall u \in \mathbb{C}^n$. The usual constraints for this reformulation of the problem are given in Table 6.4. The set of all lines of the network is denoted $\mathcal{L}$. With the exception of the rank condition, all constraints are convex, either linear, convex quadratic or semi-definite constraints.

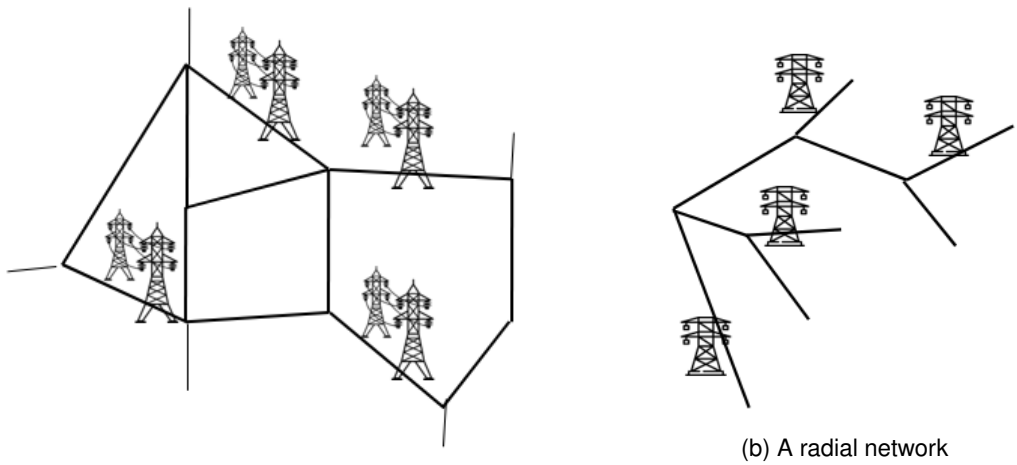
For radial networks, i.e., acyclic networks, it can be shown that an equivalent formulation of the BIM formulation with lifting variable W can be obtained by replacing the constraints $W \succcurlyeq 0$ and $rank(W) = 1$ by the following weaker constraints $W_{i,i}W_{j,j} = \lvert W_{i,j}\rvert^2$ and $W_{i,j} = W^*_{j,i}$, for all lines $(i, j) \in \mathcal{L}$, see [SL12, Corollary 1] (z^* denoting the complex conjugate of $z \in \mathbb{C}$).

Table 6.4: Constraints for the Bus Injection Model reformulated with lifting variables

Constraint	Expression
Voltage magnitude at bus i	$\underline{V}_i^2 \leq W_{i,i} \leq \overline{V}_i^2$
Intensity magnitude in line i, j	$\lvert y_{i,j}\rvert^2 \left(W_{i,i} + W_{j,j} - 2\mathrm{Re}(W_{i,j})\right) \leq \overline{I}_{i,j}^2$
Sending-in power magnitude in line i, j	$\lvert y_{i,j}\rvert\lvert W_{i,i} - W_{i,j}\rvert \leq \overline{S}_{i,j}$
Power injection limits at bus i	$\underline{s}_i \leq_{\mathbb{C}} s_i \leq_{\mathbb{C}} \overline{s}_i$
Complex power conservation at bus i	$y_{i,i}^* W_{i,i} + \sum_{(i,j)\in\mathcal{L}} y_{i,j}^* \left(W_{i,i} - W_{i,j}\right) = s_i$
Positivity constraint	$W \succcurlyeq 0$
Rank constraint	$rank(W) = 1$

6.4 The Branch Flow Model (BFM)

The Branch Flow Model is based on a description of intensity, apparent power flows and losses in the lines of the network. It is adapted for radial networks (i.e., without cycles), as opposed to the BIM, which is adapted for both meshed and radial networks, see Figure 6.1[1].



(a) A meshed network

(b) A radial network

Figure 6.1: Meshed vs. radial networks

For the Branch Flow Model, we assume the lines of the network are oriented, from the leaves to the root of the tree, which corresponds generally to a substation. With that convention, for each bus i except the root node, there is exactly one edge with starting point i. This edge is directed towards the unique ancestor of i in the tree. Besides, we assume edges are indexed according to their depths in the tree, see Figure 6.2 for an example.

The variables for the BFM formulation of the AC OPF problem are given in the Table 6.5.

With this model, the power losses in $\overrightarrow{(i,j)}$ can be expressed by $z_{i,j}\mathcal{I}_{\overrightarrow{(i,j)}}$, with $z_{i,j} = \frac{1}{y_{i,j}}$ the complex impedance of line (i, j). The expression of the constraints of the OPF in the BFM formulation are given in the following table. The set of directed edges (directed towards the root of the tree) is denoted by $\mathcal{E}$.

[1] Icons made by Freepik and Smashicons from www.flaticon.com

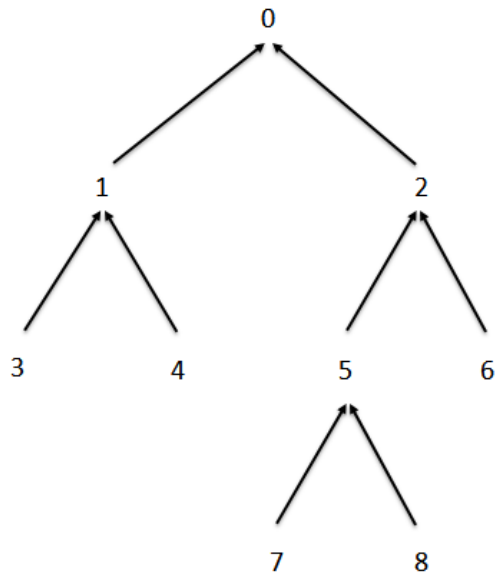


Figure 6.2: Meshed vs. radial networks

Table 6.5: Decision variables for the Branch Flow Model

Variable	Notation
Voltage squared magnitude at bus i	$v_i \in \mathbb{R}_+$
Intensity squared magnitude in edge $\overrightarrow{(i,j)}$	$\mathcal{I}_{\overrightarrow{(i,j)}} \in \mathbb{R}_+$
Sending-in power in edge $\overrightarrow{(i,j)}$ at i	$S_{\overrightarrow{(i,j)}} \in \mathbb{C}$
Active power injection at bus i	$p_i \in \mathbb{R}$
Reactive power injection at bus i	$q_i \in \mathbb{R}$
Power injection at bus i	$s_i = p_i + \mathbf{i}q_i \in \mathbb{C}$

Table 6.6: Constraints for the Branch Flow Model

Constraint	Expression
Voltage magnitude at bus i	$\underline{V}_i^2 \leq v_i \leq \overline{V}_i^2$
Intensity magnitude in edge $\overrightarrow{(i,j)}$	$0 \leq \mathcal{I}_{\overrightarrow{(i,j)}} \leq \overline{I}^2_{\overrightarrow{(i,j)}}$
Sending-in power magnitude in line i,j	$\left\lvert S_{\overrightarrow{(i,j)}} \right\rvert \leq \overline{S}_{\overrightarrow{(i,j)}}$
Power injection limits at bus i	$\underline{s}_i \leq_{\mathbb{C}} s_i \leq_{\mathbb{C}} \overline{s}_i$
Complex power conservation at edge $\overrightarrow{(i,j)}$	$S_{\overrightarrow{(i,j)}} = \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S_{\overrightarrow{(k,i)}} - z_{k,i}\mathcal{I}_{\overrightarrow{(k,i)}}) + s_i$
Complex power conservation at root bus 0	$0 = \sum_{\overrightarrow{(k,0)} \in \mathcal{E}} (S_{\overrightarrow{(k,0)}} - z_{k,0}\mathcal{I}_{\overrightarrow{(k,0)}}) + s_0$
Voltage propagation	$v_i - v_j = 2\mathrm{Re}(z^*_{i,j} S_{\overrightarrow{(i,j)}}) - \lvert z_{i,j}\rvert^2 \mathcal{I}_{\overrightarrow{(i,j)}}$
Link between intensity and sending-end power in $\overrightarrow{(i,j)}$	$v_i \mathcal{I}_{\overrightarrow{(i,j)}} = \lvert S_{\overrightarrow{(i,j)}}\rvert^2$

6.5 Equivalence between BIM and BFM formulations

Actually, it has been proved that the BIM and BFM formulations are equivalent for radial connected networks without lines with zero impedance: there exists a bijection between their feasible sets which preserves the cost, see [Bos+14, Theorem 6] and [Din+19]. For meshed network, the BFM formulation does not account for all the physical constraints of the problem and is equivalent to a relaxation of the BIM formulation. In the case of a radial connected network, one can choose either the BIM or BFM formulation, and one formulation or the other may be more adapted for the specific use. For instance, some algorithms exploit the recursive structure of the equations in the BFM formulation to efficiently compute load flows (in single phase, balanced or unbalanced three phase networks), like Forward-Backward sweep methods [BS11; EH08]. These methods consist in applying repeatedly a Forward sweep, starting from the leaves and going to the root, where power flows (and sometimes intensities in the lines) are updated followed by a backward sweep, where voltage magnitudes are updated (and sometimes intensities in the lines). Some exactness conditions of conic relaxations which exploit the recursive structure of the equations in the BFM formulation are given in [FL13; Gan+14; Hua+16]. Other exactness conditions are more easily proved using the BIM formulation [LL11; SL12; Bos+15].

6.6 Difficulty of AC OPF problem

The power flow is governed by the non-convex Kirchhoff's circuit Laws, which are quadratic equality constraints. Therefore, the AC OPF problem is a non-convex optimization problem, and is in general NP-hard even in the deterministic static case for a tree network as shown in [LGVH15, Theorem 1, p.3]. The problem has also been recently shown to be strongly NP-hard in a general setting in [BV19]. We can distinguish three approaches to deal with the computational issue of NP-hardness of the AC OPF problem. The first one is to approximate the power flow equations (in order to make it linear for instance). The second is to look for a local optimum of the problem, using heuristic methods or methods from non-linear programming. The third one is to consider convex relaxations of the AC OPF problem.

6.7 Approximation/Linearization of the AC OPF problem

6.7.1 Single-phase networks

A popular linear approximation of the AC OPF problem is the so-called Direct Current Optimal Power Flow (DC OPF), which is obtained by linearizing the power flow equations in the BIM formulation. It is an approximation neglecting power losses and reactive power flows in the network, while assuming fixed voltage magnitude and small phase angle differences between adjacent buses (which allows to linearize the trigonometric functions in the polar representation of complex voltage). This approximation is successfully used in several works in many contexts (deterministic, security constrained, stochastic, dynamic), and is widely used in the industry today. It is an appropriate approximation for market-related applications which require the computation of reliable Locational Marginal Prices, see [OCS04]. However, due to the change of paradigm in the energy sector, with the rising share of distributed generation and more demand-side management, more reverse flows in distribution networks are to be expected, which may result in low voltage operations for distribution networks, [NT15]. Therefore, assumptions justifying the use of the DC models may fail to hold in this new context. Besides, as the DC approximation assumes fixed voltage magnitude and neglects losses, it is inappropriate for application such as voltage regulation and loss minimization... Methodologies to obtain error bounds between DC and AC OPF have been developed in [SJA09; DM16].

There exist other linearizations of the AC OPF problem, for instance, some authors linearize the power flow equations around a given set-point (sensitivity analysis) to account for small deviations of the power injections from a nominal value. See [SJA09] for an overview of variants of the DC OPF formulation, as well as [MH+19, Chapter 5] and the references therein. The latter reference also presents other convex non-linear approximations of the AC OPF problem, which range from second-order cone to convex quadratic approximations.

We present more in details another linear model, which is a linearization of the BFM formulation of the AC OPF problem, presented in [Low14a, Section VI.B], called *Simplified Distflow equations* in [BW89] as it will be used in Chapter 7. This model neglects losses and yields overestimation of voltage magnitudes. The variables for this model are given in the following table.

Table 6.7: Decision variables for the Linear DistFlow model

Variable	Notation
Linearized voltage squared magnitude at bus i	$v_i^{\text{Lin}} \in \mathbb{R}_+$
Linearized sending-in power in edge $\overrightarrow{(i,j)}$ at i	$S^{\text{Lin}}_{\overrightarrow{(i,j)}} \in \mathbb{C}$
Active power injection at bus i	$p_i \in \mathbb{R}$
Reactive power injection at bus i	$q_i \in \mathbb{R}$
Power injection at bus i	$s_i = p_i + \mathbf{i}q_i \in \mathbb{C}$

In this model, the power losses are neglected, and there is no variable related to the square magnitude of intensity in the lines. The expression of the constraints of the OPF in the BFM formulation are given in the following table. We recall that the set of directed edges (directed towards the root of the tree) is denoted by $\mathcal{E}$.

Table 6.8: Constraints for the Linear DistFlow model

Constraint	Expression
Voltage magnitude at bus i	$\underline{V}_i^2 \leq v_i^{\text{Lin}} \leq \overline{V}_i^2$
Sending-in power magnitude in line i, j	$\left\lvert S^{\text{Lin}}_{\overrightarrow{(i,j)}} \right\rvert \leq \overline{S}_{\overrightarrow{(i,j)}}$
Power injection limits at bus i	$\underline{s}_i \leq_{\mathbb{C}} s_i \leq_{\mathbb{C}} \overline{s}_i$
Complex power conservation at edge $\overrightarrow{(i,j)}$	$S^{\text{Lin}}_{\overrightarrow{(i,j)}} = \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,i)}} + s_i$
Complex power conservation at root bus 0	$0 = \sum_{\overrightarrow{(k,0)} \in \mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,0)}} + s_0$
Voltage propagation	$v_i^{\text{Lin}} - v_j^{\text{Lin}} = 2\text{Re}(z_{i,j}^* S^{\text{Lin}}_{\overrightarrow{(i,j)}})$

6.7.2 Unbalanced multi-phase networks

An extension of the Linearized DistFlow model to unbalanced multiphase networks has also been recently established in [Arn+16], neglecting losses and assuming fixed ratio between voltage phasors of $\exp\left(\pm\mathbf{i}\frac{2\pi}{3}\right)$. This model is extended in [San+16] which proposes an iterative algorithm where at each iteration, a linearized OPF model is solved with losses and ratio between voltage phasors fixed, then these parameters are updated using the solution of the linearized OPF problem. This is of particular importance as most low voltage distribution networks are typically unbalanced.

6.8 Non-linear programming

There are several methods used to look for local minimizers of the OPF problem. Among them, the methods which perform probably the best are based on the Newton-Raphson method [IS87]. Alternative approaches include Sequential Quadratic Programming, steepest descent methods [For+10]. These optimization techniques and many more can be found in the very complete surveys [FSR12a; FSR12b]. However, these approaches do not offer any theoretical guarantees regarding the optimality of the solution they return and may exhibit a strong sensitivity to the initial point.

6.9 Convex relaxations of the OPF problem

A very extensive overview on relaxations and approximations of AC Power flow equations can be found in the book [MH+19]. A particular focus on conic relaxations, related literature and results is given in [Zoh+20]. A series of video tutorials can be found in [CR18]. We give a more in-depth overview of the literature in this field, as it is closely related to the results developed in Chapter 7.

6.9.1 Semi-definite, chordal and second-order cone relaxations

We focus here on results on the conic relaxations of AC OPF problems for monophase radial networks. These models allow to treat the case with balanced three-phase radial networks by a simple transformation the problem. The most famous conic relaxation of the AC OPF Problem are the Second-Order Cone (SOC), first introduced in [Jab06], the Semi-Definite (SD) in [Bai+08] and the Chordal relaxations. All three are presented in details in [Low14a; BAD18]. For the BFM formulation, the SD and Chordal relaxation do not exist. The SOC relaxation for the BFM formulation is obtained by relaxing the non-linear equality constraints:

$$v_i I_{\overrightarrow{(i,j)}} = |S_{\overrightarrow{(i,j)}}|^2$$

and replacing them by the rotated Second-order cone constraints (which can be expressed as standard second-order cone constraints by an appropriate change of variable):

$$v_i I_{\overrightarrow{(i,j)}} \geq |S_{\overrightarrow{(i,j)}}|^2.$$

For the BIM formulation, we only give the construction for the SD and SOC relaxations, and refer to previously mentioned references for the Chordal relaxation. The SD relaxation of the BIM formulation is obtained by considering its reformulation with lifting variables W and removing the rank constraint $rank(W) = 1$. Similarly, the SOC relaxation of the BIM formulation is obtained by considering its reformulation with lifting variables W, by removing the rank constraint $rank(W) = 1$ and by replacing the positivity constraint $W \succcurlyeq 0$ by the weaker constraint (positivity of 2 by 2 principal sub-matrices associated to the lines of the network):

$$\forall (i,j) \in \mathcal{L}, \quad W_{i,i} W_{j,j} \geq |W_{i,j}|^2,$$

which are rotated second-order cone constraints. The Chordal and SD relaxations are equivalent in terms of precision but the Chordal relaxation exhibits better numerical performance and is therefore always preferable to the SD relaxation (see [Low14a]). The SOC relaxation achieves better computational cost than the SD relaxation at the expenses of an a priori higher relaxation gap: depending on the application, one may prefer one relaxation or the other. However, for radial networks, the SOC, chordal and SD relaxations are equivalent, and thus, the SOC is preferable for such topologies. Similarly as the original non-convex BIM and BFM formulations are equivalent for radial connected networks (without line with zero impedance), the SOC relaxations for the BIM and BFM formulations are equivalent, i.e., there exists a bijection between their feasibility sets, under the same assumptions [Bos+14, Theorem 6]. Conic relaxations are often used in the literature due to their enhanced numerical performance, compared to the non-convex formulation, see for instance [Swa17; Vra+13] and they are exact (no relaxation gap) for many practical instances of OPF problem, as pointed out in [LL11]. For this reason, it comes at no surprise that many authors have studied exactness conditions for these convex relaxations.

6.9.2 Exactness conditions of convex relaxations of OPF problem

A posteriori conditions on the solution of the convex dual problem of the Bus Injection Model (BIM) formulation of the OPF problem (a Semi-Definite optimization problem) under which the SD relaxation is exact are given in [LL11]. These conditions are extended in [GT12] to the deterministic multi-period case with storage. A priori exactness conditions of the SD and SOC relaxations are first given in [SL11; SL12], under an assumption called over-generation, which amounts to relaxing the power balance equality constraint and replacing it by an inequality constraint. Using the BFM formulation, these results are extended in [FL13], using a related condition, called load

over-satisfaction. Weaker conditions are obtained in [ZT11; ZT12; Bos+15]: not all power injections are assumed unbounded from below. Another a priori exactness condition is the absence of upper bounds on voltage magnitude, see [Gan+12]. Let us mention [Low14a; Low14b] which give a general overview of standard convex relaxation of AC OPF and some conditions under which they are exact.

The assumption of infinite lower bounds on power injections or infinite upper bounds on voltage magnitude at buses of the network are typically not verified for OPF instances. For this reason, other authors have obtained more realistic a priori exactness conditions, see [Gan+14; Hua+16]. It is typically required that the upper bounds on apparent power injections in the network to be small enough, which ensures the absence of big reverse active and reactive power flows. For pedagogical reasons, we introduce more in details the main result of [Hua+16], as it will be generalized in Chapter 7. For all bus i, define $\mathcal{E}_i$ as the set of directed edges belonging to the sub-tree starting from i, as shown in Figure 6.3. In this example, we have $\mathcal{E}_2 = \left\{\overrightarrow{(5,2)}, \overrightarrow{(6,2)}, \overrightarrow{(7,5)}, \overrightarrow{(8,5)}\right\}$, $\mathcal{E}_5 = \left\{\overrightarrow{(7,5)}, \overrightarrow{(8,5)}\right\}$, $\mathcal{E}_1 = \left\{\overrightarrow{(3,1)}, \overrightarrow{(4,1)}\right\}$.

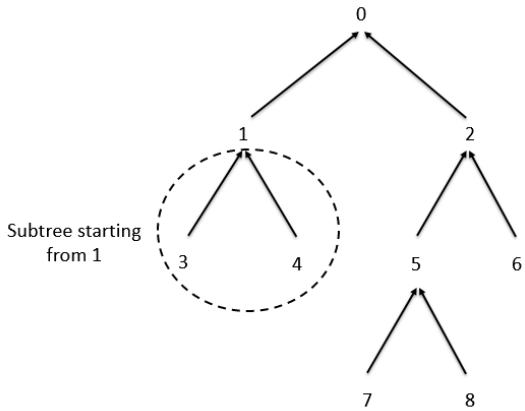


Figure 6.3: Example of sub-tree $\mathcal{E}_1$

Consider the following modified BFM model, denoted (P), where the cost functional only depends on active power injections at the buses.

$$\min_{s,S,\mathcal{I},v} \quad f(\mathrm{Re}(s)) \tag{6.9.1}$$

$$s.t. \quad \underline{V}_i^2 \leq v_i \leq \overline{V}_i^2, \qquad \forall i \in \mathcal{B}, \tag{6.9.2}$$

$$0 \leq \mathcal{I}_{\overrightarrow{(i,j)}} \leq \overline{I}^2_{\overrightarrow{(i,j)}}, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.3}$$

$$\left|S_{\overrightarrow{(i,j)}}\right| \leq \overline{S}_{\overrightarrow{(i,j)}}, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.4}$$

$$\underline{s}_i \leq_{\mathbb{C}} s_i \leq_{\mathbb{C}} \overline{s}_i, \qquad \forall i \in \mathcal{B} \setminus \{0\}, \tag{6.9.5}$$

$$S_{\overrightarrow{(i,j)}} = \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S_{\overrightarrow{(k,i)}} - z_{k,i} I_{\overrightarrow{(k,i)}}) + s_i, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.6}$$

$$0 = \sum_{\overrightarrow{(k,0)} \in \mathcal{E}} (S_{\overrightarrow{(k,0)}} - z_{k,0} I_{\overrightarrow{(k,0)}}) + s_0, \tag{6.9.7}$$

$$v_i - v_j = 2\mathrm{Re}(z^*_{i,j} S_{\overrightarrow{(i,j)}}) - |z_{i,j}|^2 I_{\overrightarrow{(i,j)}}, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.8}$$

$$v_i \mathcal{I}_{\overrightarrow{(i,j)}} = |S_{\overrightarrow{(i,j)}}|^2, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}. \tag{6.9.9}$$

The SOC relaxation of this model, denoted (P_{SOC}), is obtained by replacing constraint (6.9.9) by the rotated second-order cone:

$$v_i \mathcal{I}_{\overrightarrow{(i,j)}} \geq |S_{\overrightarrow{(i,j)}}|^2, \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.10}$$

Reference [Hua+16] considers the following condition.

(C1) Assume that for any $x := (s, S, \mathcal{I}, v)$ feasible for the problem (P), the following holds:

$$v_i^{\text{Lin}} \leq \overline{V}_i^2, \qquad \forall i \in \mathcal{B}, \tag{6.9.11}$$

$$S^{\text{Lin}}_{\overrightarrow{(i,j)}} = \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,i)}} + s_i, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.12}$$

$$0 = \sum_{\overrightarrow{(k,0)} \in \mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,0)}} + s_0 \tag{6.9.13}$$

$$v_i^{\text{Lin}} - v_j^{\text{Lin}} = 2\text{Re}(z^*_{i,j} S^{\text{Lin}}_{\overrightarrow{(i,j)}}), \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \tag{6.9.14}$$

$$\text{Re}(z^*_{k,l} S^{\text{Lin}}_{\overrightarrow{(i,j)}}) \leq 0, \qquad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \forall \overrightarrow{(k,l)} \in \mathcal{E}_i. \tag{6.9.15}$$

The linear inequalities (6.9.11)-(6.9.12)-(6.9.13)-(6.9.14) come from the *simplified Distflow equations* introduced earlier and mean feasibility of the linearized power flow equations. The inequality (6.9.15) has a simple interpretation: active (resp. reactive) backward flows in the network are sufficiently compensated by forward (resp. active) forward power flows for any feasible point. The following result is proved in [Hua+16].

Theorem 6.9.1. *[Hua+16, Proposition 3] Consider the above BIM model, under the following assumptions.*

1. *The cost function f is monotone non-decreasing in $\text{Re}(s_0)$.*
2. *The network is radial, connected and passive (i.e., line resistances and reactances are non-negative).*
3. *Condition (**C1**) holds.*

Then the SOC relaxation is exact: $val(P) = val(P_{SOC})$.

Many of the previously mentioned works focus on exactness conditions for formulations in a deterministic and static setting, with the exception of [GT12], which extends the results of [LL11] to the dynamic case with energy storage systems. This will motivate the subsequent Chapter 7.

6.9.3 Advanced convex relaxations of the OPF problem

Some authors have exhibited some limitations of the SDP and SOC relaxations (namely a non-vanishing relaxation gap or a high computational burden for large-scale networks in the case of the SD relaxation) and proposed corrective approaches. Both [KDS15; KDS16] focus on tightening the SOC relaxation in a meshed network, whereas [BAD18] develops a SD relaxation which is a trade-off between the precision of the standard SD relaxation and the low computational cost of the SOC relaxation. When the standard SD relaxation fails to be exact, some recent works focus on Moment Relaxations and polynomial optimization [Mol+15; Jos16] in order to ensure feasibility of the power flow. This approach is based on a hierarchy of Semi-Definite optimization problems which can be made more accurate at the expense of a higher computational cost. There exist asymptotic results showing the convergence of the moment relaxations towards the value of the problem, when the order of the hierarchy goes to infinity [KL19]. It is however more difficult to obtain a priori conditions under which moment relaxations of fixed order are exact. Some methods allow to retrieve feasible solutions of the non-convex problem using convex relaxations, like penalization methods. Reference [Mol+15] combines the moment relaxation approach with a penalization of reactive power injections at some nodes. References [MAL15; Mol+16] consider the SD relaxation and use a penalization of losses in problematic lines. Such problematic lines are detected by a graph theoretical algorithm in [MAL15] and by studying power mismatch between the current solution of the penalized relaxation and a close rank one solution in [Mol+16]. Those works propose iterative methods to update the cost functional in order to encourage the solution to be feasible for the original non-convex problem. This enlarges the usability of convex optimization methods to non-convex problems, at the expense of potential sub-optimality. Reference [VCM20] investigates the best AC feasibility recovery methods for instances for which the convex relaxation of the AC OPF problem is inexact. It also develops

two metrics to quantify AC infeasibility of an optimal solution of the convex relaxation of the AC OPF problem. Let us mention that even in the presence of a relaxation gap (preventing the possibility to retrieve a feasible solution of the original problem given an optimal solution of the relaxed problem), solving a convex relaxation still yields useful information, such as a lower bound on the optimal cost, which can be used to evaluate a feasible point of the non-convex problem, or provide a certificate of infeasibility of the original problem if the relaxed problem is infeasible. Last but not least, solving a convex relaxation is generally less challenging than solving the original non-convex problem.

6.9.4 Unbalanced multiphase networks

For completeness, let us mention [DZG13; GL14] which consider semi-definite relaxations of AC OPF problem in the case of unbalanced multiphase radial networks, which represent most of the low voltage distribution networks. In particular, [GL14] presents two SD relaxations, which can be interpreted as the natural extensions of the standard BIM and BFM conic relaxations in the multiphase unbalanced setting. It then proves equivalence between these two relaxations. References [Liu+17; LLW18] propose to map the structure of the unbalanced network into a chordal graph, which allows to propose a SD chordal relaxation for this particular setting.

6.10 Optimal Power Flow in dynamic setting

A dynamic AC OPF model with energy storage system is formulated in [GKA13], using only the non-convex formulation and without considering uncertainty. The SOC relaxation is used in a dynamical setting in [GSGK18] in order to optimally size distribution networks. Conditions found in [LL11] under which the SD relaxation is exact are extended in [GT12] to the deterministic multi-period case with storage. A survey of OPF methods and tools with application to distribution networks with energy storage systems can be found in [SM16].

6.11 Optimal Power Flow in stochastic environment

With the increasing share of intermittent energy sources and demanding loads, like electrical vehicles, forecast errors for power demands are expected to become higher and jeopardize network operations security. Introducing uncertainty in the OPF model allows to account for these forecasting errors.

6.11.1 Chance-Constraints OPF

The most common stochastic models are the probabilistic OPF problems. A survey on the field of Chance-Constrained OPF can be found in [BCH14]. Some recent research proposes a probabilistic version of the DC OPF model [Roa+13; Roa+16], which are approximations neglecting losses and voltage magnitude variations. A similar but more physically accurate methodology is to consider probabilistic versions of the linearized AC OPF model around a reference scenario [RMT17; RA17]. Other works consider the robust counterpart of the SD relaxation restricted to affine-linear decision-rules [Vra+13]. A Semi-Definite convex relaxation of the chance-constrained AC OPF problem is proposed in [Ven+17] using a scenario based-approach and assuming piece-wise linear decision rules or Gaussian uncertainty. A Second-Order Cone approximation of the chance-constrained AC OPF is proposed in [HPC18] which allows good numerical performances, combined with a feasibility recovery method. These works approximate the impact of uncertainty by a linearization step combined either with the assumption of a known distribution (usually Gaussian) of the noise or with a scenario-based approach, which allows a deterministic analytical reformulations of the chance constraints. All these aforementioned works related to the probabilistic OPF do not consider flexibility brought by storage and do not take into account the dynamical aspects associated to such storage systems in a completely satisfactory way. Other authors consider distributionally robust OPF [ZSM16], accounting for mis-specification of the probability distribution of the random data impacting the problem.

6.11.2 Multistage stochastic OPF

To the best of our knowledge, there are very few works focusing on the combined difficulties of a stochastic and dynamic model of the AC OPF. Three elements seem essential to us in order to obtain relevant results: the AC OPF or a conic relaxation is necessary in order to obtain sufficient accuracy of the physics of the network, a dynamical model is necessary when considering small storage systems studied in the short run, a stochastic model is necessary with high renewable integration. In particular, a consistent stochastic multi-period model should guarantee that decisions are non-anticipative, i.e., are not taken with respect to yet unobserved stochastic quantities.

A dynamic stochastic AC OPF model solved through an SOC relaxation is considered in [NCP14]. However, the scenario tree considered is a comb, as opposed to a branching tree. In particular, the approximation of the underlying process is not consistent, as the structure of information (the filtration) is not discretized, as recommended by [HRS06; Sha06; PP12; PP14; PP15]. As a consequence, decisions after the first stage are anticipative, thus underestimating the real cost due to over-fitting. The work [LS17] considers a two-stage adaptative robust multi-period version of the OPF problem: the first stage corresponds to the first time period, while in the second stage, decisions in a multi-period setting are taken according to a worst case scenario. The non-convexity of the problem is dealt with using the SOC relaxation of the problem. Though the model considered allows good physical accuracy due to the use of SOC relaxations, and good robustness properties, solution of this problem remain anticipative. This yields an underestimation of the cost of the worst case scenario, which is counterbalanced by the adaptative robust model chosen. Hence, the cost may be under-estimated or over-estimated depending on the scenarios considered. In the PhD thesis [Swa17], the author uses the SOC relaxation of the BFM formulation in a two-stage setting in order to optimally control energy flexibilities on a distribution network, with a very detailed model. This amounts to consider the simplified case where decisions are non-anticipative only for the initial time steps. In [JKK14], the author circumvents the difficulties of multistage stochastic programming by considering affinely adjustable robust counterpart (AARC) of the multi-period DC OPF problem, hence in particular, it inherits the limitations of DC OPF model. This methodology has been extended to SOC relaxations of AC OPF in the case of 2-stage stochastic optimization in [BQQ15]. However, the uncertainty only impacts the decision rules on active power generation: the model does not account for the impact of uncertainty on other variables, like voltage or reactive power injections. An iterative scheme adjusting parameters of the optimization problem is proposed in [Sun+16]. Each iteration consists in three steps. First, a dynamic AC OPF problem is solved in order to optimize a control policy. Second, the probability of violation of the constraints is estimated by simulation. Last, the bounds on some decision variables are adjusted if this estimated violation probability is too high, otherwise, the algorithm stops. This method is interesting but offers no theoretical guarantee of optimality.

Chapter 7

Bounding the duality gap of the multi-stage stochastic Alternating Current Optimal Power Flow problem with storage

We propose a generic multistage stochastic model for the Alternating Current Optimal Power Flow (AC OPF) problem, to account for the randomness of the electricity production by decentralized renewable energy sources and the dynamic constraints of storage systems. The multi-stage stochastic framework allows to account for non-anticipativity of decisions and requires, in practice, the formulation of a scenario tree, whose size typically has to grow exponentially with the number of time stages. This induces a large scale optimization problem, which, combined with the non-convex nature of the AC OPF, makes it a priori extremely challenging to solve to global optimality. We derive easy to check and realistic a priori conditions guaranteeing a vanishing relaxation gap for the multi-stage AC OPF problem, which can thus be solved using convex optimization algorithms. We also give an easily computable a posteriori upper bound on the relaxation gap. In particular, we show that a null or low relaxation gap may be expected for applications with light reverse power flows (low installed decentralized production capacity for instance) or if sufficient storage capacities with low cost are available. We then show that bounds on duality gaps of multi-stage stochastic problems arising in energy management can be obtained, where storage devices can be used in a limited way. Such results are based on the node formulation of multi-stage problems and on Shapley-Folkman-type results.

Nomenclature

Complex numbers

$\mathbf{i}$	Imaginary number with $\mathbf{i}^2 = -1$.
$\mathtt{Re}(z)$	For $z = a + \mathbf{i}b \in \mathbb{C}$, with $a, b \in \mathbb{R}$, $\mathtt{Re}(z) = a$ denotes its real part.
$\mathtt{Im}(z)$	For $z = a + \mathbf{i}b \in \mathbb{C}$, with $a, b \in \mathbb{R}$, $\mathtt{Im}(z) = b$ denotes its imaginary part.
$z \leq_{\mathbb{C}} z'$	$\mathtt{Re}(z) \leq \mathtt{Re}(z')$ and $\mathtt{Im}(z) \leq \mathtt{Im}(z')$.
z^*	For $z \in \mathbb{C}$, z^* denotes its complex conjugate.

Sets and indices

$\mathcal{B}$	Set of buses, indexed by i, with 0 reference bus and root of the tree.
$\mathcal{E}$	Set of (directed) lines, directed towards the substation 0.
$\mathcal{E}_i$	Set of (directed) lines of the sub-tree starting from i (i: root of $\mathcal{E}_i$).
$\mathcal{T}$	Set of time steps, indexed by t, $\mathcal{T} = 0, 1, 2, ...T$.
Ω	Discrete probability space, indexed by ω.

Decision Variables

$s = (s_{i,t,\omega})_{i\in\mathcal{B},t\in\mathcal{T},\omega\in\Omega}$	Complex power injections (generator convention).
$s_0 = (s_{0,t,\omega})_{t\in\mathcal{T},\omega\in\Omega}$	Complex power injections (generator convention) at the slack bus 0.
$p^{inj} = (p^{inj}_{i,t,\omega})_{i\in\mathcal{B}\setminus\{0\},t\in\mathcal{T},\omega\in\Omega}$	Active power injections by batteries.
$p^{abs} = (p^{abs}_{i,t,\omega})_{i\in\mathcal{B}\setminus\{0\},t\in\mathcal{T},\omega\in\Omega}$	Active power absorbed by batteries.
$q = (q_{i,t,\omega})_{i\in\mathcal{B}\setminus\{0\},t\in\mathcal{T},\omega\in\Omega}$	Flexible reactive power injections.
$q^{\text{sol}} = (q^{\text{sol}}_{i,t,\omega})_{i\in\mathcal{B}\setminus\{0\},t\in\mathcal{T},\omega\in\Omega}$	Flexible reactive power injections by solar panels.
$v = (v_{i,t,\omega})_{i\in\mathcal{B},t\in\mathcal{T},\omega\in\Omega}$	Squared voltage magnitude.
$\mathcal{I} = \left(\mathcal{I}_{\overrightarrow{(i,j)},t,\omega}\right)_{\overrightarrow{(i,j)}\in\mathcal{E},t\in\mathcal{T},\omega\in\Omega}$	Squared magnitude of intensity.
$S = \left(S_{\overrightarrow{(i,j)},t,\omega}\right)_{\overrightarrow{(i,j)}\in\mathcal{E},t\in\mathcal{T},\omega\in\Omega}$	Complex sending-end complex power in the lines.
$X = (X_{i,t,\omega})_{i\in\mathcal{B},t\in\mathcal{T},\omega\in\Omega}$	States of charge of storage systems.
y	Collection of all decision variables of the problem considered.
$s_0^{\text{Lin}} = (s^{\text{Lin}}_{0,t,\omega})_{t\in\mathcal{T},\omega\in\Omega}$	Linearized power injections at bus 0.
$v^{\text{Lin}} = (v^{\text{Lin}}_{i,t,\omega})_{i\in\mathcal{B},t\in\mathcal{T},\omega\in\Omega}$	Squared linearized voltage magnitude.
$S^{\text{Lin}} = \left(S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}\right)_{\overrightarrow{(i,j)}\in\mathcal{E},t\in\mathcal{T},\omega\in\Omega}$	Sending-end complex power.

Input Parameters

$z_{i,j} = r_{i,j} + \mathbf{i}x_{i,j}$	Impedance of $\overrightarrow{(i,j)} \in \mathcal{E}$.
$\overline{v}_i/\underline{v}_i$	Maximum/Minimum squared voltage magnitude at bus i.
$\overline{S}_{\overrightarrow{(i,j)}}$	Maximal magnitude of sending-in power in $\overrightarrow{(i,j)} \in \mathcal{E}$
$\overline{\mathcal{I}}_{\overrightarrow{(i,j)}}$	Maximal squared magnitude of intensity in $\overrightarrow{(i,j)} \in \mathcal{E}$
$\overline{p}_i^{inj}/\overline{p}_i^{abs}$	Maximum active power injected/absorbed by battery at bus i.
$\overline{X}_i/\underline{X}_i$	Maximum/minimum state of charge of battery at bus i.
$x_{0,i}$	Initial value of level of battery at bus i.
$\rho_i^{abs}/\rho_i^{inj}$	Charging/discharging efficiency parameter of battery at i.
$s^d_{i,t,\omega}$	Exogenous residual complex power demand.
$p^{\text{sol}}_{i,t,\omega}$	Exogenous active power production by solar panels.
I^{sol}	Exogenous solar irradiance.
$\overline{x}^{sol,norm}$	Normalized envelop of solar production.
$\overline{p}^{sol,tot}$	Total installed solar capacity on the network.
$\overline{p}_i^{sol}$	Installed solar capacity at bus i.
$s^{cons}_{i,t,\omega}$	Exogenous complex consumption.
S_i	Peak demand at bus i / Size of bus i.
$s_t^{cons,ref}$	Normalized consumption at time t.
τ_t	Time corresponding to time index $t \in \mathcal{T} \cap \{T+1\}$.
Δ_t	Length of time step $t \in \mathcal{T}$, given by $\tau_{t+1} - \tau_t$.

7.1 Introduction

Distribution networks are currently facing a major change due to the increasing share of decentralized Renewable Energy Sources (RES). They can create local physical violations and induce uncertainty on the network operating point: their production levels are random, as a result of the weather. To alleviate these issues, many Distribution Network Operators (DNOs) are required to become able to respond locally to unforeseen events. This can be done by the means of energy flexibilities located at the nodes of the network, such as demand-response, energy storage systems, energy conversion systems, power electronics... The dynamical and sometimes uncertain nature of these

subsystems as well as the uncertainty brought by RES should be taken into account in the operation planning tools used by DNOs. For instance, electrical vehicles providing ancillary services have random and dynamical availability.

One of these relevant tools for network operation planning is the so-called Optimal Power Flow (OPF) problem. It is a mathematical optimization problem which aims at finding an operating point of a power network that minimizes a given objective function, such as generation costs, active power losses, subject to constraints on power injections and losses, voltage magnitudes and intensities in the lines. Recent surveys can be found in [Cap+11; Cap16]. We are interested in the Alternating Current Optimal Power Flow (AC OPF) problem which is an accurate physical model of the operating point of a network. It is non-convex and traditionally used in a static deterministic framework. However, as argued before, a stochastic dynamic framework with storage systems and RES is more reasonable for the future, which is the focus of this chapter.

7.1.1 Optimal Power Flow Problem

There exist two main declinations of the AC OPF problem: the Bus Injection Model (BIM), and the Branch Flow Model (BFM), both presented in [Low14a]. Both formulations are equivalent for radial connected networks without lines with zero impedance [Din+19]. We choose the BFM formulation in this paper, since the proofs and arguments are easier to present in this context. The AC OPF problem is a non-convex optimization problem, and is in general NP-hard even in the deterministic static case for a tree network as shown in [LGVH15], and it has been recently shown to be strongly NP-hard in a general setting in [BV19]. Several optimization techniques from non-linear programming and heuristic methods have been proposed to solve the problem, see [FSR12a; FSR12b]. However, descent and heuristic methods to solve such non-convex problems may fail to converge to a global optimum and can be sensitive to the initial starting point. On the other hand, some recent works have shown that many real-world instances can be solved to global optimality using convex relaxations of the original non-convex AC OPF problem [LL11].

Let us mention another popular formulation of the OPF problem, the so-called Direct Current Optimal Power Flow (DC OPF). This model is an approximation, as opposed to a relaxation of the AC OPF problem. This approximation neglects non-linear effects, such as power losses through the network, reactive power flows and voltage variations. In particular, decentralized production can induce low voltage operations for distribution networks [NT15], which cannot be modeled in the DC formulation. Methodologies to obtain error bounds between DC and AC OPF have been developed in [SJA09; DM16]. Another popular approximation is the Linearized DistFlow model which is nothing else than a linear approximation of the BFM formulation [BW89] which neglects losses and provides overestimation of voltage magnitudes (see Lemma 7.3.7). We only focus on the AC OPF formulation.

7.1.2 Existing work on convex relaxations of AC OPF

A very extensive overview on relaxations and approximations of AC Power flow equations can be found in the book [MH+19], while a particular focus on conic relaxations, related literature and results is given in [Zoh+20]. A series of video tutorials on this topic can be found in [CR18]. We focus here on results on the conic relaxations of AC OPF problems for single phase radial networks. These models allow to treat the case of balanced three-phase radial networks, which reduces to the former case. The most famous conic relaxations of the AC OPF Problem are the Second-Order Cone (SOC) relaxation, first introduced in [Jab06] and the Semi-Definite (SD) relaxation, introduced in [Bai+08]. Both are presented in details in [Low14a; BAD18]. For radial networks, the SOC and SD relaxations are equivalent, but the SOC relaxation exhibits better numerical performance, and is preferable for such topologies. Conic relaxations are often used in the literature due to their enhanced numerical performance and for the certificate of optimality they may provide, compared to the non-convex formulation, see for instance [Swa17; Vra+13]. They are exact (no relaxation gap) for many practical instances of OPF problem, as pointed out in [LL11]. For this reason, it comes at no surprise that many authors have studied exactness conditions for these convex relaxations. A posteriori conditions on the solution of the dual problem are given in [LL11], which are then extended to the deterministic multi-period case with storage in [GT12]. It has also been proved that the relaxation is exact for radial connected network under an assumption like over-generation [SL12] and load over-satisfaction [FL13; ZT11; ZT12; Bos+15]. Another a priori exactness condition is the absence of upper bounds on voltage magnitude [Gan+12]. As these conditions are

not verified by practical instances, other authors [Gan+14; Hua+16] have obtained more realistic a priori exactness conditions. It is typically required that the upper bounds on apparent power injections in the network be small enough, which ensures the absence of big reverse active and reactive power flows.

7.1.3 Our contributions

In this chapter, we develop a generic model for the multi-stage stochastic version of the AC Optimal Power Flow problem with storage systems and intermittent RES. Multistage stochastic models allow to ensure the non-anticipativity of decision variables and are a pre-requisite in order to avoid the modeling issue of decisions taken according to realization of yet unknown data. To ensure this property, these models require the formulation of a scenario tree, the size of which, has to grow exponentially with the number of time stages, see for instance [Sha06; PP14]. The difficulty of solving a multistage stochastic AC OPF problem is therefore two-fold: it is non-convex and large-scale. To alleviate the non-convexity issue, we consider the SOC relaxation of the problem.

Our main contribution is to show that approaches guaranteeing the absence of relaxation gap for the AC OPF problem, originally developed in a static deterministic setting, can be extended to the multi-stage stochastic setting. Inspired by the approach of [Hua+16] in the deterministic case, we propose to restrict the feasible set of the problem by adding a finite number of linear constraints, which ensures a conservative behavior with respect to some physical limits (namely feasibility of a linearized power flow), and impose compensations for active or reactive reverse power flows in the network. We show that the restricted problem has the same optimal value as the original one under realistic assumptions which can easily be checked a priori, see Proposition 7.3.1. The restricted problem has no relaxation gap, and hence its optimal value is easily computable, see Theorem 7.3.10. This allows us to easily compute a feasible solution of the original problem and an a posteriori upper bound on its relaxation gap, see Theorem 7.3.11. Besides, the result provides realistic and tractable a priori conditions guaranteeing that the relaxation gap of the original problem is zero, see Theorem 7.3.12. The interest of the result is numerically illustrated on two realistic distribution networks with respectively 56 and 47 buses found in [Far+12] and [Far+11], equipped with distributed storage and solar panels.

Using a different approach, we then derive explicit bounds on duality gaps of appropriate formulations of dynamic or multi-stage stochastic non-convex problems in energy management can be obtained. Such estimates are connected to the Shapley-Folkman theorem, which provides bounds on the distance between a Minkowski sum of non-convex sets and its convex envelop.

7.1.4 Related work

By comparison with other works guaranteeing zero relaxation gap for the static deterministic AC OPF problem [LL11; SL12; FL13; ZT12; Bos+15; Gan+14; Hua+16], we consider a multistage stochastic setting. The conditions given in this chapter extend to the stochastic case the realistic zero relaxation gap conditions given in [Hua+16]. Moreover, we use this approach to provide a posteriori bounds on the relaxation gap. We also allow a more general cost functional, which may depend on power flows and losses in the network and voltage magnitude.

Several works consider a deterministic dynamic AC OPF model, like [GKA13], which uses the non-convex formulation or [GSGK18], which considers the SOC relaxation.

Some recent research proposes probabilistic (indifferently called Chance-Constraints) versions of the OPF problem. A probabilistic DC OPF problem is tackled in [Roa+13; Roa+16], whereas a probabilistic AC OPF model linearized around a reference scenario is suggested in [RMT17; RA17]. Other works consider the robust counterpart of the SD relaxation restricted to affine-linear decision-rules [Vra+13]. A Semi-Definite convex relaxation of the chance-constrained AC OPF problem is proposed in [Ven+17] using a scenario based-approach and assuming piece-wise linear decision rules, or assuming Gaussian uncertainty. A Second-Order Cone approximation of the chance-constrained AC OPF is proposed in [HPC18] which allows good numerical performances, combined with a feasibility recovery method. None of these references account for storage systems, nor do they consider dynamical aspects of the problem.

Several other works deal with dynamic stochastic models. In particular, an important requirement for such problems is to ensure that decision variables remain non-anticipative, i.e., do not depend on yet unobserved random

data. References [Swa17; NCP14] consider the simplified case where decision are non-anticipative for the initial time steps only, as they consider a scenario tree with a comb structure (two-stage). Non-anticipativity is guaranteed in [JKK14] using affine-linear policies but with a DC OPF problem, hence inheriting the limitations of such approximate models. Non-anticipativity of the decisions is also guaranteed in [Sun+16] which considers an iterative heuristic procedure to optimize a decision policy. By comparison, the approach by scenario trees used here accounts for the non-anticipativity and benefits from theoretical convergence guarantees: scenario tree methods provide an approximation of the value of the original problem with continuous distribution of the random data, and this approximation converges to this value when the number of scenarios goes to infinity [Sha06; PP14].

7.1.5 Outline of the chapter

In Section 7.2, we introduce the multi-stage stochastic AC-OPF problem and its convex relaxation. In particular, we give a brief recall on the formulation of non-anticipativity constraints in the case of the scenario formulation of multi-stage stochastic problems in 7.2.2. Then, in Section 7.3, we present a modified version of the original non-convex problem, obtained by restricting the feasible set. We show that the value of the restricted and original problems are equal under realistic a priori verifiable assumptions, see section 7.3.3 and Proposition 7.3.1. We then prove that the second order cone relaxation of the restricted problem is exact under some mild assumptions, see Theorem 7.3.10. This provides a convenient way to compute an a posteriori bound on the relaxation gap of the original problem, see Theorem 7.3.11. This also gives a tractable and realistic a priori condition ensuring a vanishing relaxation gap of the original problem, see Theorem 7.3.12. We illustrate numerically these results on two realistic distribution networks with respectively 56 and 47 buses in Section 7.4. Then, in Section 7.5, we focus on obtaining explicit a priori bounds on duality gaps of multi-stage stochastic problems in energy management, using Shapley-Folkman type of results, recalled in Subsection 7.5.1. In particular, we show how the node formulation of multi-stage stochastic problems (see Subsection 7.5.2 for a recall on this type of formulation of multistage stochastic problems) is appropriate to obtain explicit bounds on the duality gap of the problem, when storage devices can be used in a limited way, see Proposition 7.5.9 and Corollary 7.5.10.

7.2 The multistage stochastic AC OPF model

7.2.1 Formulation of the problem

Throughout the chapter, we will make the assumption that the network is radial, connected and passive, i.e., $\mathtt{Re}(z_{\overrightarrow{(i,j)}}) \geq 0$ and $\mathtt{Im}(z_{\overrightarrow{(i,j)}}) \geq 0$ for all lines $\overrightarrow{(i,j)} \in \mathcal{E}$ of the network. We formulate a multi-stage stochastic AC OPF problem with a battery storage system at each bus of the network (except for the reference bus 0) using the Branch Flow Model. We next describe the multistage stochastic AC OPF problem in Branch Flow Model formulation. We impose the following constraints. First, we consider the constraints on voltage squared magnitudes v:

$$v_{0,t,\omega} = 1, \quad t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.1}$$

$$\underline{v}_i \leq v_{i,t,\omega} \leq \overline{v}_i, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.2}$$

We also incorporate bounds on intensity squared magnitude $\mathcal{I}$:

$$0 \leq \mathcal{I}_{\overrightarrow{(i,j)},t,\omega} \leq \overline{\mathcal{I}}_{\overrightarrow{(i,j)}}, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.3}$$

We consider bounds on sending-end power flow S magnitudes in the lines of the network, which is a convex quadratic constraint:

$$\left| S_{\overrightarrow{(i,j)},t,\omega} \right| \leq \overline{S}_{\overrightarrow{(i,j)}}, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.4}$$

We consider bound constraints on active power injected p^{inj}, absorbed p^{abs} by batteries and flexible reactive power injected q:

$$0 \leq p^{inj}_{i,t,\omega} \leq \overline{p}^{inj}_i, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.5}$$

$$0 \leq p^{abs}_{i,t,\omega} \leq \overline{p}^{abs}_i, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.6}$$

$$\underline{q}_i \leq q_{i,t,\omega} \leq \overline{q}_i, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.7}$$

We introduce the constraints on the states of charge of the batteries X, which represent respectively their dynamics, their initial values and their physical bounds:

$$X_{i,t+1,\omega} = X_{i,t,\omega} + \rho^{abs}_i p^{abs}_{i,t,\omega}\Delta_t - \rho^{inj}_i p^{inj}_{i,t,\omega}\Delta_t, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T} \setminus \{T\}, \omega \in \Omega, \tag{7.2.8}$$

$$X_{i,0,\omega} = x_{0,i}, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T} \setminus \{T\}, \omega \in \Omega, \tag{7.2.9}$$

$$\underline{X}_i \leq X_{i,t,\omega} \leq \overline{X}_i, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.10}$$

Then we consider the expression of the complex power injections s at the buses of the network (except at the slack bus):

$$s_{i,t,\omega} = p^{inj}_{i,t,\omega} - p^{abs}_{i,t,\omega} + \mathbf{i}q_{i,t,\omega} - s^d_{i,t,\omega}, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.11}$$

This allows to formulate the power balance equations at the non-slack buses and the slack bus 0:

$$S_{\overrightarrow{(i,j)},t,\omega} = \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S_{\overrightarrow{(k,i)},t,\omega} - z_{k,i}\mathcal{I}_{\overrightarrow{(k,i)},t,\omega}) + s_{i,t,\omega}, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.12}$$

$$0 = \sum_{\overrightarrow{(k,0)} \in \mathcal{E}} (S_{\overrightarrow{(k,0)},t,\omega} - z_{k,0}\mathcal{I}_{\overrightarrow{(k,0)},t,\omega}) + s_{0,t,\omega}, \quad t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.13}$$

We consider also the voltage propagation constraint and the constraint making the link between voltage v, intensity $\mathcal{I}$ and power flow S in the network:

$$v_{i,t,\omega} - v_{j,t,\omega} = 2\mathrm{Re}(z^*_{i,j}S_{\overrightarrow{(i,j)},t,\omega}) - |z_{i,j}|^2\mathcal{I}_{\overrightarrow{(i,j)},t,\omega}, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.14}$$

$$v_{i,t,\omega}\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} = |S_{\overrightarrow{(i,j)},t,\omega}|^2, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.2.15}$$

Constraint (7.2.15) is a non-convex quadratic equality constraint. Last, we consider the non-anticipativity constraint, which encodes the fact that decision variables y should not depend on yet unknown realization of random data of the problem:

$$y \text{ is non anticipative.} \tag{7.2.16}$$

When the random processes (and their probability distributions) impacting the system are discretized and represented as a scenario tree, the non-anticipativity constraints can be represented as a set of linear equality constraints [SDR14]. More details shall be given later on their formulation. We consider a general (possibly random, progressively-measurable) convex cost function C depending on all decision variables of the problem. We can now formulate the multi-stage stochastic AC-OPF problem:

$$\min_{y=(s_0,s,p^{inj},p^{abs},q,X,S,\mathcal{I},v)} \quad \mathbb{E}\left[C(s_0, p^{inj}, p^{abs}, q, S, \mathcal{I}, X)\right] \tag{7.2.17}$$

$$s.t. \quad (7.2.1) - (7.2.16). \tag{7.2.18}$$

We denote this optimization problem by (P). This problem is non-convex because of constraint (7.2.15).

7.2.2 On the formulation of the non-anticipativity constraints

To formulate the non-anticipativity constraints, we define a scenario tree, which allows to formulate a problem with a finite number of scenarios while accounting for the filtration structure. Each scenario $\omega \in \Omega$ is associated with a trajectory $\xi_\omega = (\xi_{t,\omega})_{t\in\mathcal{T}}$ of the exogenous random process impacting the system and can be visualized as a path from the root to the leaves of the scenario tree. To give a concrete example, we consider the symmetric random walk on $\mathbb{Z}$ with steps -1 or $+1$, with $T = 2$. We give the corresponding scenario tree in Figure 7.1. Nodes of the

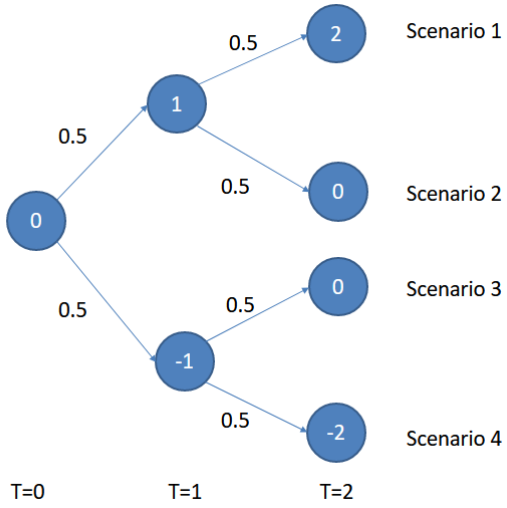


Figure 7.1: The random walk on $\mathbb{Z}$

tree are labeled with a time and an associated value of the process at this time. Edges in this tree are labeled with a transition probability. For instance, for scenario ω_1 has probability $1/8$, and the values of the process for this scenario are $\xi_{0,\omega_1} = 0$; $\xi_{1,\omega_2} = 1$; $\xi_{2,\omega_1} = 2$.

Scenarios ω and ω' are said to be indistinguishable up to time t if $\xi_{\tau,\omega} = \xi_{\tau,\omega'}$ for any $\tau \leq t$. Indistinguishability up to time t defines an equivalence relationship for scenarios. We use the notation $\omega \sim_t \omega'$ for the fact that ω and ω' are indistinguishable scenarios up to time t. For instance, for the example of the symmetric random walk, ω_1 and ω_2 are indistinguishable up to time 1, and so are ω_3 and ω_4. For any time step $t \in \mathcal{T}$ and any scenario ω, $rep_t(\omega)$ denotes a fixed representative of the equivalence class of ω with respect to $\sim_t$, which means that $\omega \sim_t \omega' \Leftrightarrow rep_t(\omega) = rep_t(\omega')$. The choice of such a representative is not unique in general. In the example of the symmetric random walk on $\mathbb{Z}$, a possible choice for the representatives is given in Table 7.1.

Table 7.1: Representative scenarios

time step t	$rep_t(\omega_1)$	$rep_t(\omega_2)$	$rep_t(\omega_3)$	$rep_t(\omega_2)$
$t = 0$	1	1	1	1
$t = 1$	1	1	3	3
$t = 2$	1	2	3	4

If $x = (x_{t,\omega})_{t\in\mathcal{T},\omega\in\Omega}$ denotes the decision variable of a multi-stage stochastic problem, and $x_{t,\omega}$ denotes the decision taken at time t for scenario ω, non-anticipativity (i.e., the fact that for all time steps t and all scenarios ω, $x_{t,\omega}$ should depend only on the values $(\xi_{\tau,\omega})_{\tau\leq t}$) can be expressed by the following constraint:

$$\forall t \in \mathcal{T}, \omega \in \Omega, \quad x_{t,\omega} = x_{t,rep_t(\omega)}. \tag{7.2.19}$$

According to standard works on multi-stage stochastic programming where measurability constraints are expressed using linear equality constraints, good scenario trees should be branching. In other words, in order to correctly approximate both the (typically continuous) distribution of the random input data and the information structure, equivalence classes with respect to $\sim_t$ for $t < T$ should typically not be singletons, see for instance [Sha06; HRS06; PP12; PP14; PP15].

Remark 7.2.1. *Getting back to problem (P), non-anticipativity constraints can be forgotten for variables which are measurable with respect to other non-anticipative variables. For instance, if we enforce non-anticipativity of p^{inj} and p^{abs}, X is also non-anticipative, as X_t is measurable with respect to $(p^{inj}_\tau, p^{abs}_\tau)_{\tau\leq t}$. Therefore, we do not need to*

add explicitly a non-anticipativity constraint on X. Besides, non-anticipativity constraints do not need to be explicitly incorporated for variables s_0, v, $\mathcal{I}$ and S, by measurable selection arguments.

7.2.3 Second-Order Cone relaxation of the problem

As we already observed, constraint (7.2.15) is non-convex. Relaxing it into an inequality constraint:

$$v_{i,t,\omega}\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} \geq |S_{\overrightarrow{(i,j)},t,\omega}|^2, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.2.20}$$

yields a convex problem, denoted (P_{SOC}). This problem is called the Second-Order Cone Relaxation of the problem, and it is given by:

$$\min_{y=(s_0,s,p^{inj},p^{abs},q,X,S,\mathcal{I},v)} \quad \mathbb{E}\Big[C(s_0,p^{inj},p^{abs},q,S,\mathcal{I},X)\Big] \tag{7.2.21}$$

$$s.t. \quad (7.2.1) - (7.2.14), (7.2.16), (7.2.20). \tag{7.2.22}$$

Indeed, (7.2.20) has the structure of a rotated second-order cone constraint $x_1x_2 \geq x_3^2 + x_4^2$. It can be equivalently reformulated as the SOC constraint:

$$\Big(2|z_{i,j}||S_{\overrightarrow{(i,j)},t,\omega}|\Big)^2 + (|z_{i,j}|^2\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} - v_{i,t,\omega})^2 \leq (|z_{i,j}|^2\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} + v_{i,t,\omega})^2, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega.$$

The (non-zero) factors $|z_{i,j}|$ are introduced in order to ensure physical homogeneity. As (P_{SOC}) is a relaxation of (P), we immediately deduce that the value of the relaxation is a lower bound of the value of the original problem, i.e., $val(P_{SOC}) \leq val(P)$. Besides, (P_{SOC}) is a Second-Order Cone problem, and as such, can be solved efficiently, using for instance interior point methods.

7.3 Restriction of the feasible set ensuring a vanishing relaxation gap

We propose to show a condition ensuring a vanishing relaxation gap for the multi-stage stochastic AC OPF problem.

7.3.1 Presentation of the problem with restricted feasible set

We present a variant of Problems (P) and (P_{SOC}), obtained by adding a finite number of linear inequalities, which, in many cases, do not modify fundamentally the feasible set. Consider additional variables $v^{\text{Lin}} = (v^{\text{Lin}}_{i,t,\omega})_{i\in\mathcal{B},t\in\mathcal{T},\omega\in\Omega}$, which play the role of the square voltage magnitude variable, $S^{\text{Lin}} = (S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega})_{\overrightarrow{(i,j)}\in\mathcal{E},t\in\mathcal{T},\omega\in\Omega}$, which play the role of the sending-end power flow variable, and $s_0^{\text{Lin}} = (s^{\text{Lin}}_{0,t,\omega})_{t\in\mathcal{T},\omega\in\Omega}$, which play the role of the power injections at the slack bus 0. We first consider the constraints of the Linearized DistFlow model [BW89], which neglects thermal losses, without considering lower bounds on the linear model of squared voltage magnitudes:

$$v^{\text{Lin}}_{0,t,\omega} = 1, \quad ,t \in \mathcal{T}, \omega \in \Omega, \tag{7.3.1}$$

$$v^{\text{Lin}}_{i,t,\omega} \leq \overline{v}_i, \quad i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.3.2}$$

$$S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega} = \sum_{\overrightarrow{(k,i)}\in\mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,i)},t,\omega} + s_{i,t,\omega}, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \tag{7.3.3}$$

$$0 = \sum_{\overrightarrow{(k,0)}\in\mathcal{E}} S^{\text{Lin}}_{\overrightarrow{(k,0)},t,\omega} + s^{\text{Lin}}_{0,t,\omega}, \quad t \in \mathcal{T}, \omega \in \Omega, \tag{7.3.4}$$

$$v^{\text{Lin}}_{i,t,\omega} - v^{\text{Lin}}_{j,t,\omega} = 2\text{Re}(z^*_{i,j}S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}), \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.3.5}$$

For all buses i, define $\mathcal{E}_i$ as the set of directed edges belonging to the sub-tree starting from i, as shown in Figure 7.2. In this example, we have $\mathcal{E}_2 = \Big\{\overrightarrow{(5,2)}, \overrightarrow{(6,2)}, \overrightarrow{(7,5)}, \overrightarrow{(8,5)}\Big\}$, $\mathcal{E}_5 = \Big\{\overrightarrow{(7,5)}, \overrightarrow{(8,5)}\Big\}$, $\mathcal{E}_1 = \Big\{\overrightarrow{(3,1)}, \overrightarrow{(4,1)}\Big\}$.

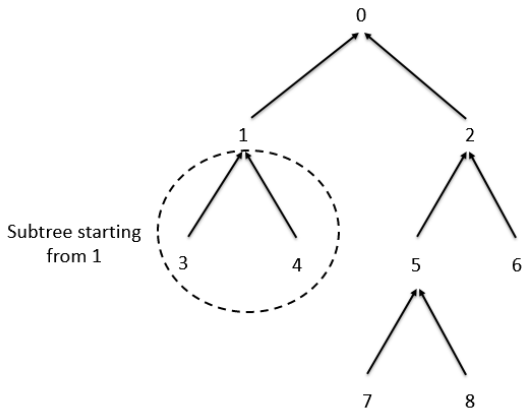


Figure 7.2: Example of sub-tree $\mathcal{E}_1$

Consider the additional constraint:

$$\mathrm{Re}(z^*_{k,l} S^{\mathrm{Lin}}_{\overrightarrow{(i,j)},t,\omega}) \leq 0, \quad \overrightarrow{(i,j)} \in \mathcal{E}, \overrightarrow{(k,l)} \in \mathcal{E}_i, t \in \mathcal{T}, \omega \in \Omega. \tag{7.3.6}$$

This constraint imposes compensation of active (resp. reactive) reverse power flows in the lines of the network by forward reactive (resp. active) power flow along the same lines.

We can now introduce the problem (P'):

$$\min_{y=(s_0,s,p^{inj},p^{abs},q,X,S,\mathcal{I},v,S^{\mathrm{Lin}},v^{\mathrm{Lin}},s_0^{\mathrm{Lin}})} \quad \mathbb{E}\left[C(s_0,p^{inj},p^{abs},q,S,\mathcal{I},X)\right] \tag{7.3.7}$$

$$s.t. \quad (7.2.1) - (7.2.16), (7.3.1) - (7.3.6). \tag{7.3.8}$$

This problem is similar to (P) but with additional constraints, which amounts to restricting the feasible set to the region of feasible linearized flow and of light reverse power flows. In particular, the value of (P') is an upper bound on the value of (P), i.e., $val(P) \leq val(P')$.

7.3.2 Second-order cone relaxation of the problem with restricted feasible set

Similarly as before, we can introduce the second-order cone relaxation of the new problem (P') by replacing the non-convex quadratic equality constraints (7.2.15) by the rotated second-order cone constraints (7.2.20). This convex relaxation is denoted (P'_{SOC}), and given by:

$$\min_{y=(s_0,s,p^{inj},p^{abs},q,X,S,\mathcal{I},v,S^{\mathrm{Lin}},v^{\mathrm{Lin}},s_0^{\mathrm{Lin}})} \quad \mathbb{E}\left[C(s_0,p^{inj},p^{abs},q,S,\mathcal{I},X)\right] \tag{7.3.9}$$

$$s.t. \quad (7.2.1) - (7.2.14), (7.2.20), (7.2.16), (7.3.1) - (7.3.6). \tag{7.3.10}$$

As (P'_{SOC}) is a relaxation of (P'),(P'_{SOC}) yields a lower bound on the optimal value of (P'), i.e., $val(P'_{SOC}) \leq val(P')$. As the feasible set of (P_{SOC}) is included in the feasible set of (P'_{SOC}), $val(P_{SOC}) \leq val(P'_{SOC})$. Besides, just as (P_{SOC}), (P'_{SOC}) is a Second-Order Cone problem, and as such, can be solved efficiently using for instance interior point methods.

7.3.3 Conditions ensuring equality of the feasible sets of the original and restricted problems

The definition of the restricted problem (P') and its relaxation (P'_{SOC}) might seem a bit artificial. We show that, for some realistic instances, the feasible sets of (P') and (P) coincide, which shows that the additional constraints introduced in (P') are valid on the feasible set of (P).

Denote $\bar{s}_{i,t,\omega} = \bar{p}_i^{inj} + \mathbf{i}\bar{q}_i - s^d_{i,t,\omega}$ an upper bound on total power injections at bus i at time t for scenario Ω, obtained for instance using Constraints (7.2.5)-(7.2.6)-(7.2.7)-(7.2.11). Define $\bar{v}^{\text{Lin}}$ and $\bar{S}^{\text{Lin}}$ by:

$$\begin{cases} \bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega} = \sum_{\overrightarrow{(k,i)}\in\mathcal{E}} \bar{S}^{\text{Lin}}_{\overrightarrow{(k,i)},t,\omega} + \bar{s}_{i,t,\omega}, & \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \\ \bar{v}^{\text{Lin}}_{0,t,\omega} = 1, & t \in \mathcal{T}, \omega \in \Omega, \\ \bar{v}^{\text{Lin}}_{i,t,\omega} - \bar{v}^{\text{Lin}}_{j,t,\omega} = 2\text{Re}(z^*_{i,j}\bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}), & \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \end{cases} \tag{7.3.11}$$

The quantities $\bar{v}^{\text{Lin}}$ and $\bar{S}^{\text{Lin}}$ are easily computable and depend only on bounds $\bar{s}$ on s and impedances of the network lines.

Proposition 7.3.1. *Define $(\bar{v}^{Lin}, \bar{S}^{Lin})$ as the unique solution of the system (7.3.11). Assume the network is radial, connected and passive (i.e., for all lines $\overrightarrow{(i,j)} \in \mathcal{E}$, we have $z_{i,j} \geq_{\mathbb{C}} 0$) and moreover suppose that it holds that:*

$$\begin{cases} \bar{v}^{Lin}_{i,t,\omega} \leq \bar{v}_i, & i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega, \\ \text{Re}(z^*_{k,l}\bar{S}^{Lin}_{\overrightarrow{(i,j)},t,\omega}) \leq 0, & \overrightarrow{(i,j)} \in \mathcal{E}, \overrightarrow{(k,l)} \in \mathcal{E}_i, t \in \mathcal{T}, \omega \in \Omega. \end{cases} \tag{7.3.12}$$

For any feasible point y of (P) (resp. (P_{SOC})), define $(s_0^{Lin}, v^{Lin}, S^{Lin})$ as the unique solution of the linear system defined by (7.3.1)-(7.3.3)-(7.3.4)-(7.3.5). Then $y' := (y, s_0^{Lin}, v^{Lin}, S^{Lin})$ is feasible with respect to (P') (resp. (P'_{SOC})). In particular, $val(P) = val(P')$ and $val(P_{SOC}) = val(P'_{SOC})$.

Proof. Consider a feasible point $y := (s, v, S, \mathcal{I})$ of (P) (resp. (P_{SOC})). Then, we have $s_{i,t,\omega} \leq_{\mathbb{C}} \bar{s}_{i,t,\omega}$ for all $i \in \mathcal{B} \setminus \{0\}$, $t \in \mathcal{T}$, $\omega \in \Omega$. Define $(s_0^{\text{Lin}}, v^{\text{Lin}}, S^{\text{Lin}})$ by (7.3.1)-(7.3.3)-(7.3.4)-(7.3.5). In particular, we have:

$$S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega} \leq \bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega} \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega.$$

Using the above, the assumption of a passive network and (7.3.12), we have for all $\overrightarrow{(i,j)} \in \mathcal{E}$, $\overrightarrow{(k,l)} \in \mathcal{E}_i$, $t \in \mathcal{T}$, $\omega \in \Omega$:

$$\text{Re}(z^*_{k,l}S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}) \leq \text{Re}(z^*_{k,l}\bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}) \leq 0,$$

which shows that (7.3.6) holds. We also get for all $\overrightarrow{(i,j)} \in \mathcal{E}$, $t \in \mathcal{T}$, $\omega \in \Omega$:

$$v^{\text{Lin}}_{i,t,\omega} - v^{\text{Lin}}_{j,t,\omega} = 2\text{Re}(z^*_{i,j}S^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}) \leq 2\text{Re}(z^*_{i,j}\bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}) = \bar{v}^{\text{Lin}}_{i,t,\omega} - \bar{v}^{\text{Lin}}_{j,t,\omega} \ ,$$

which implies for all $i \in \mathcal{B}$, using (7.2.1), (7.3.1) and the orientations of the edges towards the slack bus 0:

$$v^{\text{Lin}}_{i,t,\omega} \leq \bar{v}^{\text{Lin}}_{i,t,\omega} \leq \bar{v}_i.$$

This shows that $y' := (y, s_0^{\text{Lin}}, v^{\text{Lin}}, S^{\text{Lin}})$ is feasible for (P') (resp (P'_{SOC})). □

The above Proposition provides a simple condition which can be verified a priori guaranteeing that the optimal values of (P) and (P') coincide. This condition can be interpreted as the feasibility of the linearized model, and light reverse power flows in the network.

Remark 7.3.2. *The above Proposition implies condition C1 in [Hua+16], which is an abstract assumption on the feasible set of the problem (recalled in Chapter 6, condition (**C1**)), but it is easier to check. We simply give sign conditions on the unique solution of a linear system, which input parameters are bounds on power injections and impedances of network lines.*

The following Proposition gives an interpretable sufficient condition under which (7.3.12) holds, namely the absence of reverse power flows in the network.

Proposition 7.3.3. *Define $(\bar{v}^{Lin}, \bar{S}^{Lin})$ as the unique solution of the system (7.3.11). Assume the network is radial, connected and passive, that $\bar{v}_i \geq 1$ for all buses $i \in \mathcal{B}$ and the following condition holds:*

$$\bar{S}^{Lin}_{\overrightarrow{(i,j)},t,\omega} \leq_{\mathbb{C}} 0, \quad \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.3.13}$$

Then (7.3.12) holds, and therefore $val(P) = val(P')$ and $val(P_{SOC}) = val(P'_{SOC})$.

Proof. Under (7.3.13) and the assumption of a passive network, we have for all $\overrightarrow{(i,j)} \in \mathcal{E}$, $t \in \mathcal{T}$, $\omega \in \Omega$:

$$\bar{v}^{\text{Lin}}_{i,t,\omega} - \bar{v}^{\text{Lin}}_{j,t,\omega} = 2\text{Re}(z^*_{i,j}\bar{S}^{\text{Lin}}_{\overrightarrow{(i,j)},t,\omega}) \leq 0,$$

which implies for all $i \in \mathcal{B}$, using (7.3.1) and the orientations of the edges towards the slack bus 0:

$$\bar{v}^{\text{Lin}}_{i,t,\omega} \leq \bar{v}^{\text{Lin}}_{0,t,\omega} = 1 \leq \bar{v}_i.$$

One can then easily show that (7.3.12) holds. □

Remark 7.3.4. *The assumption of a passive network holds in practice for most real-world networks. The assumption on the upper bounds on voltage magnitude also holds since one typically aims to maintain the voltage magnitudes of the buses around the value 1 p.u., up to some tolerance threshold. The assumption of absence of reverse power flows holds for some realistic instances, for instance when demand exceeds local generation.*

For applications where the demand exceeds local power generation and injections by the battery, the following corollary shows that the values of (P) (resp. (P_{SOC})) and (P') (resp. (P'_{SOC})) coincide.

Corollary 7.3.5. *Define $(\bar{v}^{Lin}, \bar{S}^{Lin})$ as the unique solution of the system (7.3.11). Assume the network is radial, connected and passive, that $\bar{v}_i \geq 1$ for all buses $i \in \mathcal{B}$ and the following condition holds:*

$$s_{i,t,\omega} \leq_{\mathbb{C}} 0, \quad i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega. \tag{7.3.14}$$

Then (7.3.12) holds, and therefore $val(P) = val(P')$ and $val(P_{SOC}) = val(P'_{SOC})$.

Proof. Under (7.3.14) and using (7.3.3), we can show that (7.3.13) holds. We conclude using Proposition 7.3.3. □

7.3.4 Vanishing relaxation gap for the problem with restricted feasible set

We now prove that the problem with restricted feasible set has no relaxation gap, i.e., $val(P'_{SOC}) = val(P')$. The interest of this result is the following: the value of the non-convex problem (P') can be computed efficiently and is an upper bound on the value of the original problem (P) (as the feasible set of (P') is included in the feasible set of (P)). It also provides a practical way to compute an upper bound on the relaxation gap of (P), given by the inequality:

$$\begin{aligned} val(P) - val(P_{SOC}) &\leq val(P') - val(P_{SOC}) \\ &= val(P'_{SOC}) - val(P_{SOC}). \end{aligned}$$

As (P_{SOC}) and (P'_{SOC}) are both Second-Order Cone problems, the bound above can be efficiently computed.

The idea of the proof of a vanishing relaxation gap for (P') relies on an appropriate relabeling of the buses, then an iterative scheme inspired by [Hua+16]. By comparison to the latter reference, we consider a multi-stage stochastic setting and allow more general cost functions. We make the following assumption, which can be ensured by appropriately re-indexing the buses:

(H.Lab) The buses are labeled in non-decreasing order according to their depths in the tree, see Figure 7.2.

Then, the iterative scheme we consider takes as input a feasible point $y^{(0)}$ of (P'_{SOC}), and at every iteration, construct a new feasible point of (P'_{SOC}) using a Forward-Backward Sweep method, see Algorithm 7.1. We shall see that the repeated applications of the Forward-Backward Sweep method 7.1 generates a convergent sequence of feasible points $(y^{(k)})_{k\in\mathbb{N}}$ of (P'_{SOC}), and that the limit satisfies the constraints of the non-convex problem (P'). The proof of convergence relies on the boundedness of the sequence generated and the monotone convergence theorem. The limit of this sequence is also feasible for (P'_{SOC}) and is a fixed point of the Forward Backward Sweep method, which is designed in such a way that a fixed point satisfies the non-convex constraint (7.2.15), see the key step at line 7 of Algorithm 7.1. Hence it is feasible for (P').

Remark 7.3.6. *There exists other Forward-Backward Sweep methods with similar structure as Algorithm 7.1, which aim at solving efficiently load-flow problems on single phase, balanced or unbalanced three-phase radial network, using the radial structure of the network [BS11; EH08].*

Algorithm 7.1 Forward-Backward sweep method

Inputs: $(s_0, S, \mathcal{I}, v)$.
for $\omega \in \Omega$, $t \in \mathcal{T}$ **do**
  {Forward pass}
  for $i = n, n-1, ..., 1$ **do**
    Let j be the unique node in $\mathcal{B}$ such that $\overrightarrow{(i,j)} \in \mathcal{E}$ with the new labels.
    $S'_{\overrightarrow{(i,j)},t,\omega} \leftarrow s_{i,t,\omega} + \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S'_{\overrightarrow{(k,i)},t,\omega} - z_{k,i} \mathcal{I}'_{\overrightarrow{(k,i)},t,\omega})$.
    $\mathcal{I}'_{\overrightarrow{(i,j)},t,\omega} \leftarrow \frac{|S'_{\overrightarrow{(i,j)},t,\omega}|^2}{v_{i,t,\omega}}$.
  end for
  {Root node}
  $s'_{0,t,\omega} \leftarrow -\sum_{\overrightarrow{(k,0)} \in \mathcal{E}} (S'_{\overrightarrow{(k,0)},t,\omega} - z_{k,0} \mathcal{I}'_{b,0,t,\omega})$.
  $v'_{0,t,\omega} \leftarrow 1$.
  {Backward pass}
  for $i = 1, 2, ..., n$ **do**
    Let j be the unique node in $\mathcal{B}$ such that $\overrightarrow{(i,j)} \in \mathcal{E}$ with the new labels.
    $v'_{i,t,\omega} \leftarrow v'_{j,t,\omega} + 2\mathrm{Re}(z^*_{i,j} S'_{\overrightarrow{(i,j)},t,\omega}) - |z_{i,j}|^2 \mathcal{I}'_{\overrightarrow{(i,j)},t,\omega}$.
  end for
end for
Outputs: $(s'_0, S', \mathcal{I}', v')$.

The following Lemma shows some a priori bounds on some variables of any feasible point of (P'_{SOC}), which will be useful to apply the monotone convergence theorem.

Lemma 7.3.7. *Assume that $y := (s_0, s, p^{inj}, p^{abs}, X, S, \mathcal{I}, v, v^{Lin}, S^{Lin}, s_0^{Lin})$ is a feasible solution of (P'_{SOC}). Then we have the following:*

$$S_{\overrightarrow{(i,j)},t,\omega} \leq_{\mathbb{C}} S^{Lin}_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega,$$
$$s_{0,t,\omega} \geq_{\mathbb{C}} s^{Lin}_{0,t,\omega}, \quad \forall t \in \mathcal{T}, \omega \in \Omega,$$
$$v_{i,t,\omega} \leq v^{Lin}_{i,t,\omega}, \quad \forall i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega.$$

Proof. The claimed inequalities are established t by t and ω by ω. Throughout the proof, we consider fixed values of $t \in \mathcal{T}$ and $\omega \in \Omega$, and we drop these indices for simplicity of the notations. The inequalities on S and S^{Lin} arise from (7.2.3) which implies that $\mathcal{I}$ is non-negative component-wise, from passivity of the network and from constraints (7.2.12) and (7.3.3). The inequalities on s_0 and s_0^{Lin} can then be deduced by the inequality between S and S^{Lin} and constraints (7.2.13) and (7.3.4). Comparing (7.2.14) and (7.3.5), using the passivity of the network and the inequalities between S and S^{Lin}, one gets for all $\overrightarrow{(i,j)}$ in $\mathcal{E}$:

$$v_i - v_j \leq v_i^{\mathrm{Lin}} - v_j^{\mathrm{Lin}}.$$

We can then show the inequalities on v and v^{Lin} using the fact that $v_0^{\mathrm{Lin}} = 1 = v_0$, by (7.2.1) and (7.3.1), and using the fact that edges are directed towards the root indexed by 0. □

The following Lemma shows some monotonic properties of the sequence generated by repeated application of the Forward-Backward sweep method 7.1.

Lemma 7.3.8. *Algorithm 7.1 is well-defined. Consider a feasible solution $y := (s_0, s, p^{inj}, p^{abs}, X, S, \mathcal{I}, v, v^{Lin}, S^{Lin}, s_0^{Lin})$ of (P'_{SOC}). Apply Algorithm 7.1 once to $(s_0, S, \mathcal{I}, v)$ and denote by $(s'_0, S', \mathcal{I}', v')$ its output. Then we have:*

$$S_{\overrightarrow{(i,j)},t,\omega} \leq_{\mathbb{C}} S'_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega,$$
$$\left|S_{\overrightarrow{(i,j)},t,\omega}\right| \geq \left|S'_{\overrightarrow{(i,j)},t,\omega}\right|, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega,$$

$$\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} \geq \mathcal{I}'_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega,$$
$$s_{0,t,\omega} \geq_{\mathbb{C}} s'_{0,t,\omega}, \quad \forall t \in \mathcal{T}, \omega \in \Omega,$$
$$v_{i,t,\omega} \leq v'_{i,t,\omega}, \quad \forall i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega.$$

Moreover, $y' := (s_0, s, p^{inj}, p^{abs}, X, S, \mathcal{I}, v, v^{Lin}, S^{Lin}, s_0^{Lin})$ *is feasible for* (P'_{SOC}).

Proof. The claimed inequalities are established t by t and ω by ω. Throughout the proof, we consider fixed values of $t \in \mathcal{T}$ and $\omega \in \Omega$, and we drop these indices for simplicity of the notations. The definition of S for leaves of the tree in the forward pass is well-defined as the sum in the LHS is empty in this case by our labels. The labels chosen ensure that the forward pass always explores leaves before their ancestors, which ensures that the forward pass is well-defined. Therefore, the whole algorithm is well-posed. Consider the forward pass, with $i = n$, n being the index of the last bus after setting the new labels (see Assumption (**H.Lab**)). In particular, the bus $i = n$ is a leaf. Denoting j its unique ancestor, we have $S_{\overrightarrow{(n,j)}} = s_n = S'_{\overrightarrow{(n,j)}}$ by construction. We then obtain, using the fact that y is feasible for (P'_{SOC}) and thus satisfies (7.2.20):

$$\mathcal{I}'_{\overrightarrow{(n,j)}} = \frac{|S'_{\overrightarrow{(n,j)}}|^2}{v_n} = \frac{|S_{\overrightarrow{(n,j)}}|^2}{v_n} \leq \mathcal{I}_{\overrightarrow{(n,j)}}.$$

Let us now assume $i < n$, and we assume the inequalities for S, S', $\mathcal{I}$ and $\mathcal{I}'$ have been proved for all $k = i+1, ..., n$. If i is a leaf, we can prove the inequalities similarly as for $i = n$. Consider the case where $i < n$ is not a leaf. Let j be its unique ancestor. Then, by passivity of the network and since the inequalities $S_{\overrightarrow{(k,i)}} \leq_{\mathbb{C}} S'_{\overrightarrow{(k,i)}}$ and $\mathcal{I}_{\overrightarrow{(k,i)}} \geq \mathcal{I}'_{\overrightarrow{(k,i)}}$ have been established for all k such that $\overrightarrow{(k,i)} \in \mathcal{E}$ (by our choice of bus labels):

$$\begin{aligned}
S'_{\overrightarrow{(i,j)}} &= s_i + \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S'_{\overrightarrow{(k,i)}} - z_{k,i} I'_{\overrightarrow{(k,i)}}) \\
&\geq_{\mathbb{C}} s_i + \sum_{\overrightarrow{(k,i)} \in \mathcal{E}} (S_{\overrightarrow{(k,i)}} - z_{k,i} I_{\overrightarrow{(k,i)}}) \\
&= S_{\overrightarrow{(i,j)}}.
\end{aligned}$$

Besides, denoting $P := \mathtt{Re}(S)$, $Q := \mathtt{Im}(S)$, $P^{\text{Lin}} := \mathtt{Re}(S^{\text{Lin}})$ and $Q^{\text{Lin}} := \mathtt{Im}(S^{\text{Lin}})$:

$$\begin{aligned}
|S'_{\overrightarrow{(i,j)}}|^2 - |S_{\overrightarrow{(i,j)}}|^2 &= (P'_{\overrightarrow{(i,j)}} - P_{\overrightarrow{(i,j)}})(P'_{\overrightarrow{(i,j)}} + P_{\overrightarrow{(i,j)}}) + (Q'_{\overrightarrow{(i,j)}} - Q_{\overrightarrow{(i,j)}})(Q'_{\overrightarrow{(i,j)}} + Q_{\overrightarrow{(i,j)}}) \\
&\leq 2(P'_{\overrightarrow{(i,j)}} - P_{\overrightarrow{(i,j)}}) P^{\text{Lin}}_{\overrightarrow{(i,j)}} + 2(Q'_{\overrightarrow{(i,j)}} - Q_{\overrightarrow{(i,j)}}) Q^{\text{Lin}}_{\overrightarrow{(i,j)}} \\
&= -2\left(\sum_{(k,l) \in \mathcal{E}_i} r_{k,l}(\mathcal{I}'_{\overrightarrow{(k,l)}} - \mathcal{I}_{\overrightarrow{(k,l)}})\right) P^{\text{Lin}}_{\overrightarrow{(i,j)}} - 2\left(\sum_{\overrightarrow{(k,l)} \in \mathcal{E}_i} x_{k,l}(\mathcal{I}'_{\overrightarrow{(k,l)}} - \mathcal{I}_{\overrightarrow{(k,l)}})\right) Q^{\text{Lin}}_{\overrightarrow{(i,j)}} \\
&= -2 \sum_{\overrightarrow{(k,l)} \in \mathcal{E}_i} \mathtt{Re}(z^*_{k,l} S^{\text{Lin}}_{\overrightarrow{(k,l)}})(\mathcal{I}'_{\overrightarrow{(k,l)}} - \mathcal{I}_{\overrightarrow{(k,l)}}) \\
&\leq 0.
\end{aligned}$$

In the inequality in the third line, we used Lemma 7.3.7, then we use the definition of S' and the fact that S satisfies (7.2.12) to obtain the following equality. The last inequality is obtained using $\mathcal{I}'_{\overrightarrow{(k,l)}} \geq \mathcal{I}_{\overrightarrow{(k,l)}}$ for all $\overrightarrow{(k,l)} \in \mathcal{E}_i$ and the fact that S^{Lin} satisfies (7.3.6). We then obtain, using the inequalities $|S'_{\overrightarrow{(i,j)}}|^2 \leq |S_{\overrightarrow{(i,j)}}|^2$ and the fact that y satisfies (7.2.20):

$$\mathcal{I}'_{\overrightarrow{(i,j)}} = \frac{|S'_{\overrightarrow{(i,j)}}|^2}{v_i} \leq \frac{|S_{\overrightarrow{(i,j)}}|^2}{v_i} \leq \mathcal{I}_{\overrightarrow{(i,j)}}.$$

This ends the proof of the inequalities:

$$S_{\overrightarrow{(i,j)}} \leq_{\mathbb{C}} S'_{\overrightarrow{(i,j)}}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E},$$

$$\left|S_{\overrightarrow{(i,j)}}\right| \geq \left|S'_{\overrightarrow{(i,j)}}\right|, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E},$$
$$\mathcal{I}_{\overrightarrow{(i,j)}} \geq \mathcal{I}'_{\overrightarrow{(i,j)}}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}.$$

The inequality $s_0 \geq_{\mathbb{C}} s'_0$ can then be deduced from the first and third above inequality and the assumption of passivity of the network. Let us now show:

$$v_i \leq v'_i, \quad \forall i \in \mathcal{B}.$$

We have $v_0 = 1 = v'_0$ by construction and by (7.2.1). Notice then that for all $\overrightarrow{(i,j)} \in \mathcal{E}$, by construction of v' and by (7.2.14):

$$\begin{aligned} v'_i - v_i &= v'_j - v_j + 2\mathrm{Re}\left(z^*_{i,j}\left\{S'_{\overrightarrow{((i,j))}} - S'_{\overrightarrow{((i,j))}}\right\}\right) - |z_{i,j}|^2\left(\mathcal{I}'_{\overrightarrow{((i,j))}} - \mathcal{I}_{\overrightarrow{((i,j))}}\right) \\ &\geq v'_j - v_j, \end{aligned}$$

where we used the earlier inequalities on S, S', $\mathcal{I}$ and $\mathcal{I}'$ and the assumption of passivity of the network. By propagating this in the network from 0 to n, we get the desired inequality. By Lemma 7.3.7, we have:

$$\begin{aligned} &S'_{\overrightarrow{(i,j)},t,\omega} \leq_{\mathbb{C}} S^{\mathrm{Lin}}_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \\ &s'_{0,t,\omega} \geq_{\mathbb{C}} s^{\mathrm{Lin}}_{0,t,\omega}, \quad \forall t \in \mathcal{T}, \omega \in \Omega, \\ &v'_{i,t,\omega} \leq v^{\mathrm{Lin}}_{i,t,\omega}, \quad \forall i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega. \end{aligned}$$

We have also:

$$\forall \overrightarrow{(i,j)} \in \mathcal{E}, \quad \mathcal{I}'_{\overrightarrow{(i,j)}} = \frac{|S'_{\overrightarrow{(i,j)}}|^2}{v_i} \geq \frac{|S'_{\overrightarrow{(i,j)}}|^2}{v'_i}.$$

This shows that $(S', v', \mathcal{I}')$ satisfies (7.2.20). Define $y' := (s_0, s, p^{inj}, p^{abs}, X, S, \mathcal{I}, v, v^{\mathrm{Lin}}, S^{\mathrm{Lin}}, s_0^{\mathrm{Lin}})$. Non-anticipativity of y' arises from the fact that for all $t \in \mathcal{T}$ and $\omega \in \Omega$, $y'_{t,\omega}$ is measurable with respect to $y_{t,\omega}$. Hence non-anticipativity of y yields the same property for y'. By construction and using the inequalities derived above, one can show that if y is feasible for (P'_{SOC}), then so is y' is feasible. □

Corollary 7.3.9. *Let $y := (s_0, s, p^{inj}, p^{abs}, X, S, \mathcal{I}, v, v^{Lin}, S^{Lin}, s_0^{Lin})$ be a feasible solution of (P'_{SOC}). Then, there exists $y' := (s'_0, s, p^{inj}, p^{abs}, X, S', \mathcal{I}', v', v^{Lin}, S^{Lin}, s_0^{Lin})$ which is feasible for (P') and such that:*

$$\begin{aligned} &S_{\overrightarrow{(i,j)},t,\omega} \leq_{\mathbb{C}} S'_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \\ &\left|S_{\overrightarrow{(i,j)},t,\omega}\right| \geq \left|S'_{\overrightarrow{(i,j)},t,\omega}\right|, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \\ &\mathcal{I}_{\overrightarrow{(i,j)},t,\omega} \geq \mathcal{I}'_{\overrightarrow{(i,j)},t,\omega}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, t \in \mathcal{T}, \omega \in \Omega, \\ &s_{0,t,\omega} \geq_{\mathbb{C}} s'_{0,t,\omega}, \quad \forall t \in \mathcal{T}, \omega \in \Omega, \\ &v_{i,t,\omega} \leq v'_{i,t,\omega}, \quad \forall i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega. \end{aligned}$$

Proof. The claimed inequalities are established t by t and ω by ω. Throughout the proof, we consider fixed values of $t \in \mathcal{T}$ and $\omega \in \Omega$, and we drop these indices for simplicity of the notations. Apply recursively Algorithm 7.1 to $x^{(0)} := (s_0, S, \mathcal{I}, v)$. This defines a sequence $(x^{(k)})_{k\in\mathbb{N}} := (s_0^{(k)}, S^{(k)}, \mathcal{I}^{(k)}, v^{(k)})_k$. By Lemmas 7.3.7 and 7.3.8, we have the following inequalities for all $k \in \mathbb{N}$:

$$\begin{aligned} &S^{(k)}_{\overrightarrow{(i,j)}} \leq_{\mathbb{C}} S^{(k+1)}_{\overrightarrow{(i,j)}} \leq_{\mathbb{C}} S^{\mathrm{Lin}}_{\overrightarrow{(i,j)}}, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \\ &\mathcal{I}^{(k)}_{\overrightarrow{(i,j)}} \geq \mathcal{I}^{(k+1)}_{\overrightarrow{(i,j)}} \geq 0, \quad \forall \overrightarrow{(i,j)} \in \mathcal{E}, \end{aligned}$$

$$s_0^{(k)} \geq_{\mathbb{C}} s_0^{(k+1)} \geq_{\mathbb{C}} s_0^{\text{Lin}}, \quad \forall$$
$$v_i^{(k)} \leq v_i^{(k+1)} \leq v_i^{\text{Lin}}, \quad \forall i \in \mathcal{B}.$$

By the monotone convergence theorem, $(x^{(k)})$ converges to a point $(s_0', S', \mathcal{I}', v')$. Define:

$$y' := (s_0', s, p^{inj}, p^{abs}, X, S', \mathcal{I}', v', v^{\text{Lin}}, S^{\text{Lin}}, s_0^{\text{Lin}}),$$

which is feasible for (P'_{SOC}) as limit of feasible points of (P'_{SOC}). Besides, it satisfies (7.2.15) as $(s_0', S', \mathcal{I}', v')$ is a fixed point of Algorithm 7.1. This implies that y is a feasible point of (P'). The inequalities claimed arise from the monotone behavior of the sequence $x^{(k)}$. □

Theorem 7.3.10. *Assume the following:*

1. *The network is radial and connected.*
2. *The network is passive, i.e.,* $\mathtt{Re}(z_{\overrightarrow{(i,j)}}) \geq 0$ *and* $\mathtt{Im}(z_{\overrightarrow{(i,j)}}) \geq 0$ *for all lines* $\overrightarrow{(i,j)} \in \mathcal{E}$ *of the network.*
3. *The cost function* C *is convex, component-wise monotone non-increasing in* $\mathtt{Re}(S), \mathtt{Im}(S), v$*, component-wise monotone non-decreasing in* $\mathcal{I}$*,* $\mathtt{Re}(s_0)$*,* $\mathtt{Im}(s_0)$ *and* $|S|$*.*

Then (P') *has no relaxation gap, i.e., its optimal value coincides with the optimal value of* (P'_{SOC})*.*

Proof. If (P'_{SOC}) is infeasible, then so is (P') and the result holds. If (P'_{SOC}) is feasible and bounded from below, consider its optimal solution y^*. Corollary 7.3.9 and the monotonicity assumptions on the cost then show that there exists $\tilde{y}^*$ which is feasible for (P') and with lower cost than y^*. This yields the result. If (P'_{SOC}) is feasible and unbounded from below, given a sequence of feasible points of (P'_{SOC}), whose costs goes to $-\infty$, we build a sequence of feasible point of (P') with lower costs, using Corollary 7.3.9. This shows that (P') is also feasible and unbounded from below. □

The assumptions of Theorem 7.3.10 are quite realistic. In practice, the cost functional is often independent from $\mathtt{Re}(S), \mathtt{Im}(S), v, \mathtt{Im}(s_0)$ and is monotone non-decreasing in the active power injections at the substation $\mathtt{Re}(s_0)$ and in thermal losses, proportional to $\mathcal{I}$. Besides, no assumption is made on the behavior of the cost regarding power injections at buses other than the slack bus. This allows to apply the result to a wide range of applications.

Theorem 7.3.11 (A posteriori bound on the relaxation gap)**.** *Under the assumptions of Theorem 7.3.10, the relaxation gap of* (P)*, given by* $val(P) - val(P_{SOC})$ *is bounded from above by* $val(P'_{SOC}) - val(P_{SOC})$*, which can be computed by solving the two conic quadratic problems* (P_{SOC}) *and* (P'_{SOC})*.*

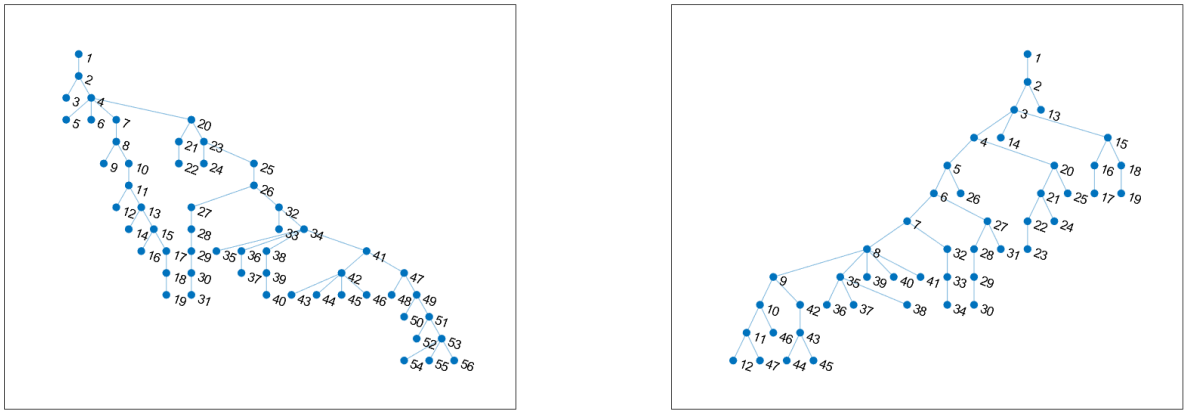
Theorem 7.3.12. *[A priori condition for a vanishing relaxation gap] Under the assumptions of Proposition 7.3.1 and Theorem 7.3.10, the problem* (P) *has no relaxation gap, i.e.,* $val(P) = val(P_{SOC})$*.*

The assumptions of Proposition 7.3.1 and Theorem 7.3.10 are overlapping and verified by realistic instances of the OPF problem, for example instances with a convex cost which penalize active power injections at the substation and power losses, and with light reverse power flows.

Remark 7.3.13. *The result can be generalized to other types of storage systems, like for instance a virtual battery arising from the aggregation of Thermostatically Controlled Loads [Hao+13; Hao+14]. Other constraints, static or dynamic (i.e., linking variables of two distinct time steps), can be incorporated to power injections at all buses except the slack bus* 0*.*

7.4 Numerical illustration

This numerical study is implemented using Matlab R2018b combined with YALMIP R20200116, interfaced with conic solver Gurobi 9.0.0 with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory.



(a) 56 buses network [Far+12]

(b) 47 buses network [Far+11]

Figure 7.3: Visualization of the networks before relabelling of the buses (according to their depths)

7.4.1 General framework

We consider an optimization window of $31h$, divided into $T + 1 = 9$ sub-intervals Each time step t corresponds to a time interval in the model $[\tau_t, \tau_{t+1}]$. The correspondence is given in Table 7.2.

Table 7.2: Time steps and their corresponding time window

t	0	1	2	3	4	5	6	7	8
$[\tau_t, \tau_{t+1}]$	$[0, 7]$	$[7, 10]$	$[10, 12]$	$[12, 14]$	$[14, 16]$	$[16, 18]$	$[18, 21]$	$[21, 24]$	$[24, 31]$

7.4.2 Network model

Network topology and characteristics

We consider two distribution networks on the Southern California Edison system with respectively 56 buses [Far+12] and 47 buses [Far+11]. A visualization of both networks is given in Figure 7.3.

We leave network topology and line resistances and reactances unchanged for both networks, compared to [Far+12; Far+11], but we modify the installed solar capacities and do not assume shunt capacitors at any bus. Additionally, we shall considered storage capacities at many buses. The peak loads are also assumed unchanged, with the exception of the load at bus 1 for the 47 buses network, which is set to 0 MVA here.

We attribute to each bus i a size parameter S_i, given by the peak load of the bus i, taken from the column Load data, Peak MVA. Non-specified buses are attributed the size $S_i = 0$. The values of the parameters of both networks are given in Tables 7.4 and 7.5. In particular, the total peak load is given by $\sum_i S_i = 3.835$ MVA for the 56 buses network and $\sum_i S_i = 11.3$ MVA for the 47 buses network, which is therefore more loaded than the 56 buses network.

For the 56 buses network, we assume the maximal intensity magnitude in each line is 300 A, which yields the bound $\overline{\mathcal{I}}_{\overrightarrow{(i,j)}} = 90000$ A^2. We accept 5% deviations of voltage magnitude from the reference value, i.e., $\underline{v}_{\overrightarrow{(i,j)}} = (0.95)^2$ p.u. and $\overline{v}_{\overrightarrow{(i,j)}} = (1.05)^2$ p.u. We assume also $\overline{S}_{\overrightarrow{(i,j)}} = 5$ MVA.

For the 47 buses network, we assume the maximal intensity magnitude in each line is 1000 A, which yields the bound $\overline{\mathcal{I}}_{\overrightarrow{(i,j)}} = 10^6$ A^2. We accept 10% deviations of voltage magnitude from the reference value, i.e., $\underline{v}_{\overrightarrow{(i,j)}} = (0.9)^2$ p.u. and $\overline{v}_{\overrightarrow{(i,j)}} = (1.1)^2$ p.u. We assume also $\overline{S}_{\overrightarrow{(i,j)}} = 20$ MVA. The bound constraints for the 47 buses network are therefore less tight than for the 56 buses network.

Network Data																	
Line Data				Line Data				Line Data				Load Data		Load Data		Load Data	
From Bus.	To Bus.	R (Ω)	X (Ω)	From Bus.	To Bus.	R (Ω)	X (Ω)	From Bus.	To Bus.	R (Ω)	X (Ω)	Bus No.	Peak MVA	Bus No.	Peak MVA	Bus No.	Peak MVA
1	2	0.160	0.388	20	21	0.251	0.096	39	40	2.349	0.964	3	0.057	29	0.044	52	0.315
2	3	0.824	0.315	21	22	1.818	0.695	34	41	0.115	0.278	5	0.121	31	0.053	54	0.061
2	4	0.144	0.349	20	23	0.225	0.542	41	42	0.159	0.384	6	0.049	32	0.223	55	0.055
4	5	1.026	0.421	23	24	0.127	0.028	42	43	0.934	0.383	7	0.053	33	0.123	56	0.130
4	6	0.741	0.466	23	25	0.284	0.687	42	44	0.506	0.163	8	0.047	34	0.067		
4	7	0.528	0.468	25	26	0.171	0.414	42	45	0.095	0.195	9	0.068	35	0.094	V_{base} = 12kV	
7	8	0.358	0.314	26	27	0.414	0.386	42	46	1.915	0.769	10	0.048	36	0.097	S_{base} = 1MVA	
8	9	2.032	0.798	27	28	0.210	0.196	41	47	0.157	0.379	11	0.067	37	0.281	Z_{base} = 144Ω	
8	10	0.502	0.441	28	29	0.395	0.369	47	48	1.641	0.670	12	0.094	38	0.117		
10	11	0.372	0.327	29	30	0.248	0.232	47	49	0.081	0.196	14	0.057	39	0.131		
11	12	1.431	0.999	30	31	0.279	0.260	49	50	1.727	0.709	16	0.053	40	0.030		
11	13	0.429	0.377	26	32	0.205	0.495	49	51	0.112	0.270	17	0.057	41	0.046		
13	14	0.671	0.257	32	33	0.263	0.073	51	52	0.674	0.275	18	0.112	42	0.054		
13	15	0.457	0.401	32	34	0.071	0.171	51	53	0.070	0.170	19	0.087	43	0.083		
15	16	1.008	0.385	34	35	0.625	0.273	53	54	2.041	0.780	22	0.063	44	0.057		
15	17	0.153	0.134	34	36	0.510	0.209	53	55	0.813	0.334	24	0.135	46	0.134		
17	18	0.971	0.722	36	37	2.018	0.829	53	56	0.141	0.340	25	0.100	47	0.045		
18	19	1.885	0.721	34	38	1.062	0.406					27	0.048	48	0.196		
4	20	0.138	0.334	38	39	0.610	0.238					28	0.038	50	0.045		

Figure 7.4: Table of the data of the 56 buses network [Far+12]: line resistances, reactances, peak loads S_i

Network Data															
Line Data				Line Data				Line Data				Load Data		Load Data	
From Bus	To Bus	R (Ω)	X (Ω)	From Bus	To Bus	R (Ω)	X (Ω)	From Bus	To Bus	R (Ω)	X (Ω)	Bus No	Peak MVA	Bus No	Peak MVA
1	2	0.259	0.808	8	41	0.107	0.031	21	22	0.198	0.046	1	0	34	0.2
2	13	0	0	8	35	0.076	0.015	22	23	0	0	11	0.67	36	0.27
2	3	0.031	0.092	8	9	0.031	0.031	27	31	0.046	0.015	12	0.45	38	0.45
3	4	0.046	0.092	9	10	0.015	0.015	27	28	0.107	0.031	14	0.89	39	1.34
3	14	0.092	0.031	9	42	0.153	0.046	28	29	0.107	0.031	16	0.07	40	0.13
3	15	0.214	0.046	10	11	0.107	0.076	29	30	0.061	0.015	18	0.67	41	0.67
4	20	0.336	0.061	10	46	0.229	0.122	32	33	0.046	0.015	21	0.45	42	0.13
4	5	0.107	0.183	11	47	0.031	0.015	33	34	0.031	0.010	22	2.23	44	0.45
5	26	0.061	0.015	11	12	0.076	0.046	35	36	0.076	0.015	25	0.45	45	0.2
5	6	0.015	0.031	15	18	0.046	0.015	35	37	0.076	0.046	26	0.2	46	0.45
6	27	0.168	0.061	15	16	0.107	0.015	35	38	0.107	0.015	28	0.13		
6	7	0.031	0.046	16	17	0	0	42	43	0.061	0.015	29	0.13	Base Voltage (kV) = 12.35	
7	32	0.076	0.015	18	19	0	0	43	44	0.061	0.015	30	0.2	Base kVA = 1000	
7	8	0.015	0.015	20	21	0.122	0.092	43	45	0.061	0.015	31	0.07	Substation Voltage = 12.35	
8	40	0.046	0.015	20	25	0.214	0.046					32	0.13		
8	39	0.244	0.046	21	24	0	0					33	0.27		

Figure 7.5: Table of the data of the 47 buses network [Far+11]: line resistances, reactances, peak loads S_i

We assume that buses i are equipped with batteries and solar panels with capacity proportional to their sizes S_i. We also consider time dependent exogenous demand at the buses, proportional to their sizes S_i as well.

Batteries

We assume the total installed capacity of batteries is $\overline{X}^{tot} = 1$ MWh for the 56 buses network and $\overline{X}^{tot} = 5$ MWh for the 47 buses network. Each bus $i \in \mathcal{B}$ is equipped with a battery with maximal energy capacity proportional to the size S_i of the bus $\overline{X}_i = \frac{S_i}{\sum_{i\in\mathcal{B}} S_i}\overline{X}^{tot}$. In other words, the storage capacity is split between buses according to their peak loads. We assume p_i^{inj} (resp. p_i^{abs}) represents the power supplied (absorbed) by the battery at bus i, at time step $t \in \mathcal{T}$, for scenario $\omega \in \Omega$. We assume each battery can be charged/discharged in 2 hours, i.e., $\overline{p}_i^{inj} = \overline{p}_i^{abs} = \frac{\overline{X}_i}{2hours}$. We assume a charging efficiency $\rho_i^{abs} = 0.95$ while the discharging efficiency is $\rho_i^{inj} = 1/\rho_i^{abs}$. Additionally, we impose the periodicity constraints:

$$\forall i \in \mathcal{B}, \omega \in \Omega, \quad X_{i,T+1,\omega} = X_{i,1,\omega}, \tag{7.4.1}$$

which ensure that the batteries will have the same states of charge on consecutive days. This allows us to take into account the daily repetition of the problem considered. In the light of Remark 7.3.13, this additional constraint does not jeopardize the earlier results of this chapter, as it is a dynamic constraint not impacting voltage squared magnitudes, intensity squared magnitudes, sending-end power flows nor power injections at the slack bus 0. Besides, we assume the batteries cannot provide nor absorb reactive power.

Photovoltaic panels

The total installed solar capacity is denoted $\overline{p}^{sol,tot}$, and we shall consider several values for this parameter in our application. Each bus $i \in \mathcal{B}$ is equipped with photovoltaic panels with maximal capacity proportional to the size S_i of the bus, i.e., $\overline{p}_i^{\text{sol}} = \frac{S_i}{\sum_{i\in\mathcal{B}} S_i}\overline{p}^{sol,tot}$. In other words, the solar capacity is split between buses according to their peak loads. The solar panels are equipped with power electronics allowing to absorb reactive power. The reactive power supplied by the solar panels $q_{i,t,\omega}^{\text{sol}}$ is a decision variable and we have the bound constraints:

$$\forall i \in \mathcal{B}, t \in \mathcal{T}, \omega \in \Omega, \quad -0.3\overline{p}_i^{\text{sol}} \leq q_{i,t,\omega}^{\text{sol}} \leq 0.$$

Exogenous residual demand

We consider an exogenous demand profile at node i and for time t given by the difference between a deterministic consumption profile and a solar power production profile

$$s_{i,t,\omega}^{d} = s_{i,t}^{cons} - p_{i,t,\omega}^{\text{sol}}.$$

The consumption profiles are assumed deterministic given by $s_{i,t}^{cons} = s_t^{cons,ref}\frac{1+\mathbf{i}0.2}{\sqrt{1+(0.2)^2}}S_i$, i.e., the consumption at each node i is proportional to a reference deterministic evolution $s^{cons,ref}$ (independent from i) represented in Figure 7.6 and the size parameter S_i.

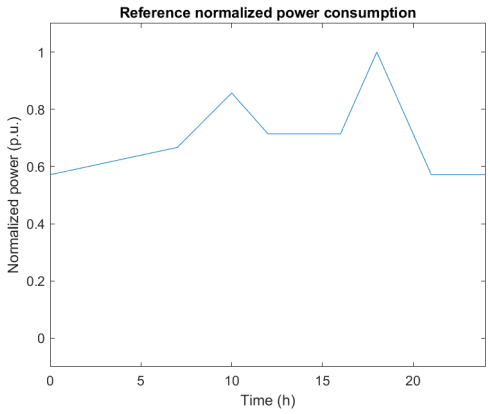


Figure 7.6: Normalized reference consumption profile $s^{cons,ref}$

The active solar production $p_{i,t,\omega}^{\text{sol}}$ is generated according to the model described in the following Section 7.4.3.

Vanishing relaxation gap for the problem with low installed solar capacity

The power injections s at all buses of the network (except the substation 0) are bounded according to:

$$\forall i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \quad p_{i,t,\omega}^{\text{sol}} - \mathbf{i}0.3\overline{p}_i^{\text{sol}} - s_t^{cons,ref}\frac{1+\mathbf{i}0.2}{\sqrt{1+(0.2)^2}}S_i - \overline{p}_i^{abs} \leq_{\mathbb{C}} s_{i,t,\omega} \leq_{\mathbb{C}} p_{i,t,\omega}^{\text{sol}} - s_t^{cons,ref}\frac{1+\mathbf{i}0.2}{\sqrt{1+(0.2)^2}}S_i + \overline{p}_i^{inj}$$

As the solar power production cannot exceed the solar capacity, the upper bounds on power injections satisfy, no matter the choice of scenario tree:

$$\forall i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \quad p^{\text{sol}}_{i,t,\omega} - s^{cons,ref}_t \frac{1 + \mathbf{i}0.2}{\sqrt{1+(0.2)^2}} S_i + \overline{p}^{inj}_i \leq_{\mathbb{C}} \overline{p}^{\text{sol}}_i + \overline{p}^{inj}_i - \frac{s^{cons,ref}_t}{\sqrt{1+(0.2)^2}} S_i - \mathbf{i}\, \frac{0.2 s^{cons,ref}_t}{\sqrt{1+(0.2)^2}} S_i$$
$$= \left(\frac{\overline{p}^{sol,tot}}{\sum_{i\in\mathcal{B}} S_i} + \frac{\overline{X}^{tot}}{(\sum_{i\in\mathcal{B}} S_i)2h} - \frac{s^{cons,ref}_t}{\sqrt{1+(0.2)^2}} - \mathbf{i}\frac{0.2 s^{cons,ref}_t}{\sqrt{1+(0.2)^2}} \right) S_i$$

For the 56 buses network, with the values $\sum_i S_i = 3.835$ MVA, $\overline{X}^{tot}/2h = 0.5$ MW and $\overline{p}^{sol,tot} \leq 1.6$ MW, using $s^{cons,ref}_t \geq 0.55$ p.u. for all $t \in \mathcal{T}$, we obtain:

$$\forall i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \quad s_{i,t,\omega} \leq_{\mathbb{C}} 0.$$

For the 47 buses network, with the values $\sum_i S_i = 11.3$ MVA, $\overline{X}^{tot}/2h = 2.5$ MW and $\overline{p}^{sol,tot} \leq 3.5$ MW, using $s^{cons,ref}_t \geq 0.55$ p.u. for all $t \in \mathcal{T}$, we obtain:

$$\forall i \in \mathcal{B} \setminus \{0\}, t \in \mathcal{T}, \omega \in \Omega, \quad s_{i,t,\omega} \leq_{\mathbb{C}} 0.$$

In particular we are in the framework of Corollary 7.3.5. Then, for any cost functional satisfying the assumptions of Theorem 7.3.10, **the problems have vanishing relaxation gaps, no matter the choice of scenario tree and scenarios** for low installed solar capacities ($\overline{p}^{sol,tot} \leq 1.6$ MW for the 56 buses network and $\overline{p}^{sol,tot} \leq 3.5$ MW for the 47 buses network). For larger values of total installed solar capacity $\overline{p}^{sol,tot}$, the assumptions of Proposition 7.3.1 may not hold.

7.4.3 Generation of i.i.d. scenarios for solar power production

Recall that τ_t is the time associated with time step $t \in \mathcal{T}$. We assume the solar power is given by:

$$p^{\text{sol}}_{i,t,\omega} = \overline{p}^{\text{sol}}_i I^{\text{sol}}_{\tau_t,\omega} \overline{x}^{sol,norm}_{\tau_t}, \tag{7.4.2}$$

where the clear-sky index I^{sol} is a random process taking values between 0 and 1, which models the clearness of the sky. Its value is 0 when the sky is completely cloudy (i.e., even at day, there would be no light) and 1 for a completely clear sky. The deterministic time-dependent envelop $\overline{x}^{sol,norm}$ models the time evolution of the normalized solar power we would observe if the sky were clear all day, i.e., the ratio of the power produced by a solar panel at a given time with a clear sky and the capacity of the solar panel. We suppose it is given by $\overline{x}^{sol,norm}_\tau = 0$ for $\tau < T_{day} = 7$ and $\tau > T_{night} = 21$ and by

$$\overline{x}^{sol,norm}_\tau = 0.5 - 0.5\cos\left(\frac{2\pi(\tau - T_{night})}{T_{night} - T_{day}}\right)$$

for $T_{day} \leq \tau \leq T_{night}$. We plot the evolution of the normalized envelop $\overline{x}^{sol,norm}$ over one day in Figure 7.7.

To build a scenario tree, we use the stochastic model in [Bad+18] for the clear-sky index I^{sol}, based on a Fisher-Wright-type Stochastic Differential Equation (SDE), given by:

$$I^{\text{sol}}_\tau = I^{\text{sol}}_0 - \int_0^\tau a(I^{\text{sol}}_s - I^{\text{ref}})\mathrm{d}s + \int_0^\tau \sigma (I^{\text{sol}}_s)^\alpha (1 - I^{\text{sol}}_s)^\beta \mathrm{d}B_s, \tag{7.4.3}$$

with B a Brownian motion. The parameter $a \geq 0$ is a mean-reversion speed parameter, and $I^{\text{ref}} \in [0,1]$ is a reference value for the clear-sky index. Under the assumption $a \geq 0$, $\alpha, \beta \geq 0.5$, this SDE has a unique strong solution with values in $[0,1]$ almost surely, see [Bad+18].

The branching structure of the scenario tree is characterized by pre-specified vector $(C_t)_{t\in\mathcal{T}}$ where C_t corresponds to the number of children nodes of a node at stage t. In particular, the total number of scenarios is given by $N = \prod_{t=1}^T C_t$. Given a structure of a scenario tree, we use the quantile-based Algorithm 7.2 with parameters given in Table 7.3 to instantiate the values of solar irradiance at the nodes of the scenario tree. This algorithm

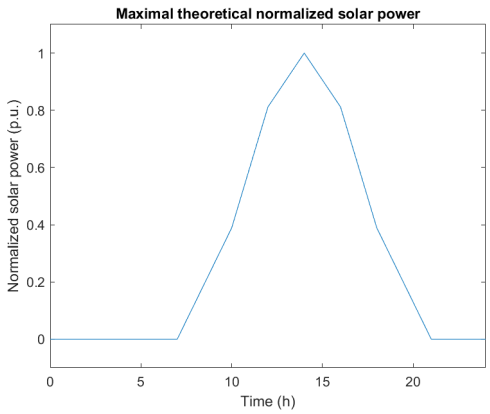


Figure 7.7: Normalized solar power envelop $\overline{x}^{sol,norm}$

estimates evenly spaced conditional quantiles using i.i.d. sample trajectories of the irradiance starting from nodes of the former time step. These conditional quantiles are then used to define nodes of the scenario tree associated to the next time step. Each scenario is assigned probability $1/N$, which is consistent with the fact that we consider evenly spaced quantiles for the values of successors of each node.

Table 7.3: Parameters of SDE of solar irradiance I^{sol} and of Algorithm 7.2

I^{ref}	a	σ	α	β	I_0^{sol}	M	τ^{Euler}
0.75	0.75 h^{-1}	0.7	0.8	0.7	0.5	10000	$0.1h$

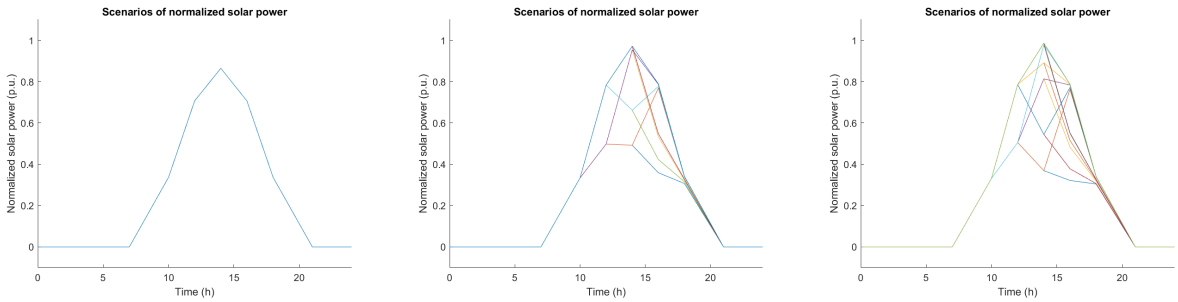
Algorithm 7.2 Generation of scenario tree of solar irradiance I^{sol}

Given: M, $(C_t)_{t=1,\dots,T}$, $\{\tau_1,\dots,\tau_T\}$, τ^{Euler}.
val: table of values of the nodes of scenario tree
$val[n_1] \leftarrow I_{init}^{\text{sol}}$ with n_1 node at time $t = 1$
for $t = 2,\dots,T$ **do**
  for n_{t-1} node at stage $t-1$ **do**
    Simulate M i.i.d. trajectories of I^{sol} in (7.4.3) on $[\tau_{t-1},\tau_t]$ conditionally to $I_{\tau_{t-1}}^{\text{sol}} = val[n_{t-1}]$ using Euler scheme with step τ^{Euler}.
    for $i = 1\dots C_{t-1}$ **do**
      Create i^{th} child of node n_{t-1}, denoted n_t.
      Set $val[n_t]$ to quantile $\frac{100(2i-1)}{2C_{t-1}}\%$ of simulated values of $I_{\tau_t}^{\text{sol}}$.
    end for
  end for
end for
Return val.

Other existing methods to generate scenario trees could be considered to obtain a scenario tree with branching structure from a set of scenario (i.e., a scenario tree with a comb structure). For instance, some authors consider moment-matching approaches [HW01]. Others use a distance obtained by adding the Wasserstein distance and a distance for filtrations [HRS06; HR09]. It has also been proposed to combine the moment-matching and metric minimization approaches [ROW14]. Recently, another distance for stochastic processes, adapted to the framework of multi-stage stochastic problems, has been developed and used to build scenario trees [PP12; PP15].

7.4.4 Choice of scenario tree structure

We consider various scenario trees which approximate the stochastic process I^{sol}, where branching (i.e., $C_t > 1$) occurs at the time steps where $\overline{x}_t^{\text{sol, norm}}$ is big. In other words, for the time steps t when there is no sun ($\overline{x}_t^{\text{sol, norm}} = 0$), the scenario tree will not be branching at t. This is heuristically justified by the fact that instantaneous volatility of the solar power at time t is directly proportional to $\overline{x}_{\tau_t}^{sol,norm}$. The time evolution of the scenarios of normalized solar power $(I_{t,\omega}^{\text{sol}}\overline{x}_{\tau_t}^{sol,norm})_{t=1,\dots,T,\omega=1,\dots,N}$ is given in Figure 7.8.



(a) 1 scenario (b) 8 scenarios (c) 12 scenarios

Figure 7.8: Scenarios of daily normalized solar power production considered $\overline{x}^{sol,norm} I^{sol}$

7.4.5 Numerical results

We assume that the cost functional is given by:

$$\frac{1}{N}\sum_{\omega=1}^{N}\sum_{t=0}^{T}\Delta_t\left(c_{0,+}(p_{0,t,\omega})_+ - c_{0,-}(p_{0,t,\omega})_- + c_{loss}\left(\sum_{\overrightarrow{(i,j)}\in\mathcal{E}} r_{i,j}I_{\overrightarrow{(i,j)},t,\omega}\right)\right), \tag{7.4.4}$$

where $c_{0,+}, c_{0,-}, c_{loss}$ are respectively equal to $1, 0.5, 2$. Since $c_{0,+} > c_{0,-}$, the cost functional is piece-wise linear and convex. Therefore an appropriate reformulation of the problem allows to consider a linear cost, by introducing specific decision variables for the positive $p_{0,t,\omega}^{+} \geq 0$ and negative $p_{0,t,\omega}^{-} \geq 0$ part of active power injections at the slack bus 0. The complementarity relation $p_{0,t,\omega}^{+}p_{0,t,\omega}^{-} = 0$ can then be shown to hold for all $t \in \mathcal{T}$ and $\omega \in \Omega$ for any optimal point: otherwise, a point with strictly lower cost could be exhibited. We do not incorporate storage costs, which requires a specific model accounting for the investment costs and aging of the energy storage systems [Hae14; Swa17].

When the total installed solar capacity is $\overline{p}^{sol,tot}$ is low enough, one can invoke Corollary 7.3.5 and Theorem 7.3.12 to prove theoretically that the relaxation gap is zero, no matter the choice of scenario tree. We consider the values $\overline{p}^{sol,tot} = 3$ MW in the case of the 56 buses network and $\overline{p}^{sol,tot} = 7.5$ MW in the case of the 47 buses network. One cannot invoke Theorem 7.3.12 guaranteeing a vanishing relaxation gap. Instead, we use Theorem 7.3.11 to compute an a posteriori bound on the relative relaxation gap ε, obtained by the following formula:

$$\varepsilon = \frac{2\left(val(P'_{SOC}) - val(P_{SOC})\right)}{|val(P_{SOC})| + \left|val(P'_{SOC})\right|}. \tag{7.4.5}$$

Table 7.4 represents the bound on the relative relaxation gap and the computation time as a function of the number of scenarios for both networks.

For the two networks and the three scenario trees considered, the a posteriori bound on relative relaxation gap of (P) is zero (up to a numerical tolerance) and the computation times remain reasonable.

We now exhibit typical cases where the a posteriori bound on the relaxation gap can be expected to be good (i.e., close to the true value of the relaxation gap) or not. To this end, we consider a deterministic case with different

Table 7.4: A posteriori bound on relative relaxation gap ε and computation time depending on the number of scenarios N

Network considered	56 buses	56 buses	56 buses	47 buses	47 buses	47 buses
Installed solar capacity $\overline{p}^{sol,tot}$	3 MW	3 MW	3 MW	7.5 MW	7.5 MW	7.5 MW
number of scenarios N	1	8	12	1	8	12
A posteriori bound ε	4.4×10^{-9}	7.5×10^{-8}	1.8×10^{-6}	0	5.1×10^{-7}	0
Yalmip time (P_{SOC})	1.56 s	14.3 s	21.5 s	1.03 s	9.98 s	28.8 s
Gurobi time (P_{SOC})	0.21 s	3.06 s	132.4 s	0.200 s	1.53 s	46.0 s
Yalmip time (P'_{SOC})	1.47 s	17.9 s	23.5 s	1.08 s	10.4 s	36.9 s
Gurobi time (P'_{SOC})	0.27 s	17.8 s	70.4 s	0.207 s	2.10 s	65.2 s

values for the storage cost and the total installed solar capacity $\overline{p}^{sol,tot}$. We consider the cost functional given by:

$$\sum_{t=0}^{T} \Delta_t \left(c_{0,+}(p_{0,t})_+ - c_{0,-}(p_{0,t})_- + c_{loss} \left(\sum_{\overrightarrow{(i,j)} \in \mathcal{E}} r_{i,j} I_{\overrightarrow{(i,j)},t} \right) + \sum_{i \in \mathcal{B}} c_{bat}(p_{i,t}^{inj} + p_{i,t}^{abs}) \right). \tag{7.4.6}$$

We give in Tables 7.5 and 7.6 the values of upper bound on the relative relaxation gap ε defined in (7.4.5) for the 56 and 47 buses networks respectively, as a function of the installed solar capacity and storage cost.

Table 7.5: A posteriori bound ε depending on storage cost and installed solar capacity (56 buses network)

$\overline{p}^{sol,tot}$	1.5 MW	3 MW	3.5 MW	4 MW	4.5 MW
$c_{bat} = 0$	0	0	7.7×10^{-6}	4.0×10^{-4}	$+\infty$
$c_{bat} = 1$	3.7×10^{-8}	0	7.2×10^{-4}	6.2×10^{-3}	$+\infty$
$c_{bat} = 2$	0	3.4×10^{-7}	7.2×10^{-4}	3.1×10^{-2}	$+\infty$

Table 7.6: A posteriori bound ε depending on storage cost and installed solar capacity (47 buses network)

$\overline{p}^{sol,tot}$	5 MW	7.5 MW	10 MW	11 MW	12 MW
$c_{bat} = 0$	2.0×10^{-5}	0	1.5×10^{-6}	3.1×10^{-5}	$+\infty$
$c_{bat} = 1$	2.1×10^{-7}	9.1×10^{-7}	2.7×10^{-3}	6.2×10^{-3}	$+\infty$
$c_{bat} = 2$	2.2×10^{-8}	1.4×10^{-8}	1.6×10^{-2}	3.7×10^{-2}	$+\infty$

One can show that for all the instances considered in Tables 7.5 and 7.6, the optimal solution of (P_{SOC}) found is feasible for (P), showing that the relaxation gap is zero in all cases investigated: $val(P) = val(P_{SOC})$. Therefore, the a posteriori bound on the relaxation gap is good when ε is close to 0. This is the case if the installed solar capacity is low (in agreement with Theorem 7.3.12) or intermediate (i.e., roughly equal to the peak load). For intermediate values of the installed solar capacity, the a posteriori bound is of better quality for lower storage cost. This makes sense since the batteries can be leveraged to absorb power to avoid reverse power flow at low cost (guaranteeing 7.3.6). For higher values of the installed solar capacity, the a posteriori bound ε becomes even infinite, although the relaxation gap of (P) is still zero. In that case, the storage capacity is not sufficient to absorb power, and hence constraint 7.3.6 cannot be satisfied, even using batteries. The numerical results show that the a posteriori bound on the relaxation gap is good in the case of low decentralized production capacity or if reverse power flows can be avoided at low cost, which is the case, in particular, if active or reactive power can be absorbed locally and at low cost.

7.5 Explicit bounds on duality gap of multi-stage stochastic problems in energy management

For many instances of the AC-OPF problem, convex relaxations of the problem are often very good, as shown in the literature [LL11]. In this section, we use metric inequalities to establish bounds on the duality gaps of multi-stage stochastic formulations of the AC-OPF problem, and to more general non-convex problems with a partially separable structure arising in energy management (like Unit Commitment for instance), where batteries could be leveraged to reduce the costs. In particular, we show that, for problems where batteries are used in a limited (conservative) way on the time horizon, an upper bound on the duality gap can be established and this bound reduces on average with the number of time steps considered. Hence batteries can help reduce cost of such problems and also reduce their duality gap, which enhances their tractability. This type of result is based on Shapley-Folkman Theorem and related results, which are presented in the following paragraph.

7.5.1 Shapley-Folkman Theorem and related results

For a set $S \in \mathbb{R}^n$, we denote by $conv(S)$ its convex hull, i.e., the smallest convex set containing S. For two sets A and B in $\mathbb{R}^n$, the Minkowski sum $A + B$ is defined by $A + B = \{a + b \mid a \in A, b \in B\}$.

Theorem 7.5.1 (Shapley-Folkman)**.** *Let $S_1, ..., S_m$ be compact sets in $\mathbb{R}^n$, and let $x \in conv\left(\sum_{i=1}^m S_i\right) = \sum_{i=1}^m conv(S_i)$. Then there exists an index set $I \subset \{1, ..., m\}$ with $card(I) \leq n$ such that*

$$x \in \sum_{i \in I} conv(S_i) + \sum_{i \notin I} S_i.$$

A proof of this result can be found in [Sch14, Theorem 3.1.2].

Definition 7.5.2. *For $S \in \mathbb{R}^n$ we define its radius by*

$$rad(S) := \inf_{x \in \mathbb{R}^n} \sup_{a \in S} \|x - a\|,$$

i.e., it is the radius of the smallest ball enclosing S. We define its inner radius by

$$inrad(S) = \sup_{x \in conv(S)} \inf_{\lambda \geq 0} \{\lambda | x \in conv(S \cap B(x, \lambda))\}.$$

The inner radius of S is the smallest radius of a ball which can contain points of S spanning any point of $conv(S)$. We have $0 \leq inrad(S) \leq rad(S)$ with $inrad(S) = 0$ if and only if S is convex. On the other hand, $inrad(S) = rad(S)$ if and only S is given by a set of affinely independent points.

For A, B subsets of $\mathbb{R}^n$, define the Hausdorff distance between A and B by:

$$\delta(A, B) := \max\left\{\sup_{x \in A} \inf_{y \in B} |x - y|, \sup_{y \in B} \inf_{X \in A} |x - y|\right\}.$$

A straightforward consequence of the above theorem is the following result.

Theorem 7.5.3. *Let $S_1, ..., S_m$ be compact sets in $\mathbb{R}^n$, then*

$$\delta\left(\sum_{i=1}^m S_i, conv\left(\sum_{i=1}^m S_i\right)\right) \leq \sum_{k=1}^{\min(m,n)} R_{[k]},$$

where $\max_{1 \leq i \leq m}(rad(S_i)) =: R_{[1]} \geq R_{[2]} \geq ... \geq R_{[m]} := min_{1 \leq i \leq m}(rad(S_i))$ are the order statistics of $(rad(S_i))_{1 \leq i \leq m}$.

In particular, the Shapley-Folkman Theorem shows that the Minkowski average of a uniformly bounded non-convex set becomes convex in the limit, when the number of sets considered becomes much larger than the ambient dimension in which the sets live. Indeed, we by the above result:

$$\lim_{m \to +\infty} \delta\left(\frac{1}{m}\sum_{i=1}^m S_i, conv\left(\frac{1}{m}\sum_{i=1}^m S_i\right)\right) = 0.$$

The following result is called Starr's refinement and has been first introduced in [Sta69].

Theorem 7.5.4 (Shapley-Folkman-Starr [Sta69]). *Let $S_1, ..., S_m$ be compact sets in $\mathbb{R}^n$, endowed with the euclidean norm $|.|_2$, and let $x \in Conv\left(\sum_{i=1}^m S_i\right)$. Then, there are points $(a_i)_{1\leq i\leq m} \in \prod_{i=1}^m S_i$, such that*

$$|x - \sum_{i=1}^{m} a_i|_2^2 \leq \sum_{k=1}^{\min(m,n)} r_{[k]},$$

where $\max_{1\leq i\leq m}(inrad(S_i)) = r_{[1]} \geq r_{[2]} \geq ... \geq min_{1\leq i\leq m}(inrad(S_i)) = r_{[m]}$ are the order statistics of $(inrad(S_i))_{1\leq i\leq m}$.

The Shapley-Folkman theorem finds application in optimization and allows to estimate the duality gap of optimization problems with a separable structure. The first results were obtained by Aubin and Ekeland in [AE76]. We first give some definitions regarding convex envelops of functions, found in [UB16]. Let $g : S \subset \mathbb{R}^n \mapsto \mathbb{R}$. We denote by g^{**} its convex envelop, i.e., the largest closed convex function smaller than g, i.e. $g^{**}(x) \leq g(x)$ for all $x \in S$, see [Roc70, Theorem 17.2]. When g is lower semi-continuous (i.e. its sub-level sets are closed) and S is compact and convex, g^{**} is closed, proper and convex [Roc70]. We can now define the following measure of non-convexity of a function f.

Definition 7.5.5 ([UB16]). *The non-convexity of $g : S \subset \mathbb{R}^n \mapsto \mathbb{R}$ is defined by:*

$$\rho(g) := \sup_{x\in S} g(x) - g^{**}(x),$$

where, for convenience, we assume that g and its convex envelop g^{**} are $+\infty$ outside of their domains (extended-valued functions) and we use the convention $\infty - \infty = 0$.

Obviously, $\rho(g) \geq 0$ with equality if and only if g is convex. The non-convexity $\rho(g)$ is finite if and only if $g : S \mapsto \mathbb{R}$ is bounded and lower semi-continuous and S is compact and convex.

Let $g_i : S_i \subset \mathbb{R}^{n_i} \mapsto \mathbb{R}$ with $i = 1, ..., m$ and assume $\rho(g_1) \geq \rho(g_2) \geq ... \geq \rho(g_m)$. Consider the optimization problem $\mathcal{P}$ with separable structure:

$$\begin{aligned} \min_{x=(x_i)_{1\leq i\leq k}\in\mathbb{R}^n} & \sum_{i=1}^{m} g_i(x_i) \\ s.t. \quad & \forall i \in \{1, ..., m\}, \quad x_i \in S_i, \\ & Ax \leq b, \\ & Gx = h. \end{aligned}$$

where $n = \sum_{i=1}^m n_i$, $A \in \mathbb{R}^{k_1\times n}$, $G \in \mathbb{R}^{k_2\times n}$, $b \in \mathbb{R}^k$ with $k = k_1 + k_2$. Without loss of generality, we assume that scalar constraints involving only one block variable x_i are implicitly represented in S_i by an appropriate modification. In particular, all scalar constraints in $Ax \leq b$ involve at least two different block variables x_i and x_j with $i \neq j$. We denote by $val(\mathcal{P})$ the value of $\mathcal{P}$. The constraints $Ax \leq b$ are called coupling or complicating constraints. Define the convexified problem $\mathcal{P}^{**}$:

$$\begin{aligned} \min_{x=(x_i)_{1\leq i\leq k}\in\mathbb{R}^n} & \sum_{i=1}^{m} g_i^{**}(x_i) \\ s.t. \quad & \forall i \in \{1, ..., m\}, \quad x_i \in S_i, \\ & Ax \leq b, \\ & Gx = h. \end{aligned}$$

We denote by $val(\mathcal{P}^{**})$ the value of $\mathcal{P}^{**}$.

We can now formulate the result of Aubin and Ekeland [AE76] which relies on Shapley-Folkman Theorem and the fact that the epigraph of the function defined in $\mathcal{P}$ can be formulated as the Minkowski sum of m sets in dimension $k + 1$.

Theorem 7.5.6 ([AE76]). *We have the following estimation on the duality gap of $\mathcal{P}$:*

$$0 \leq val(\mathcal{P}) - val(\mathcal{P}^{**}) \leq \sum_{i=1}^{\min(m,k+1)} \rho(g_i).$$

This result has been refined in [UB16]. Let us define $\tilde{k}_1$ as the maximal number of scalar inequality constraints in $Ax \leq b$ which can be simultaneously verified, and let $\tilde{k} := \tilde{k}_1 + k_2$. In particular, $\tilde{k}_1 \leq k_1$ and $\tilde{k} \leq k$.

Theorem 7.5.7 ([UB16]). *There exists $I \subset \{1, ..., m\}$ with $card(I) \leq \tilde{k}$ and an optimal solution $\hat{x} = (\hat{x}_i)_{1 \leq i \leq m}$ of $\mathcal{P}^{**}$ such that:*

$$\forall i \notin I, \quad g_i^{**}(\hat{x}_i) = g_i(\hat{x}_i).$$

In particular, we have the following estimation on the duality gap of $\mathcal{P}$:

$$0 \leq val(\mathcal{P}) - val(\mathcal{P}^{**}) \leq \sum_{i=1}^{\min(m,\tilde{k})} \rho(g_i).$$

We wish to use the latter result to obtain estimates on the duality gap of multi-stage stochastic problems arising in energy management, where energy storage systems allow to modify operations. In particular, dynamic and non-anticipativity constraints play the role of complicating constraints. As we wish to reduce this number in order to use Theorem 7.5.7, the node formulation of multi-stage stochastic programming on scenario trees is particularly adapted. Indeed, as we shall see, this formulation allows us to eliminate the non-anticipativity constraints, compared to the scenario formulation presented in Section 7.2.2.

7.5.2 On node formulation of multi-stage stochastic programming problems

The node formulation for multi-stage stochastic programming is equivalent to the scenario formulation presented in Section 7.2.2. Variables are no longer associated with a time and a scenario, but with a node of the scenario tree. This formulation can be obtained from the scenario formulation by eliminating duplicated variables, using the non-anticipativity linear equality constraints, see (7.2.19). Nodes of the scenario tree are indexed by (t, n_t), with a time index t and a node index n_t taking value in a time-dependent set of integers between 1 and N_t. An example with the symmetric random walk on $\mathbb{Z}$ is given in Figure 7.9. For this particular example $N_0 = 1$, $N_1 = 2$ and $N_2 = 4$.

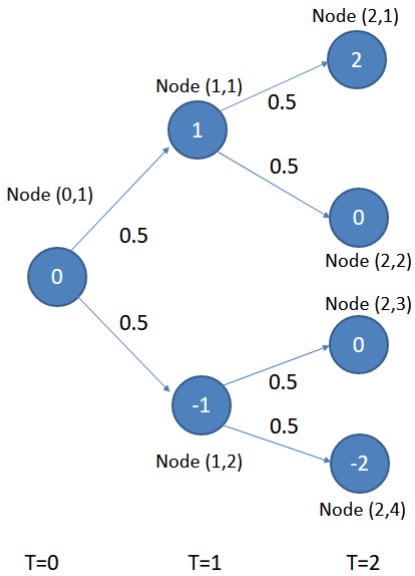


Figure 7.9: The random walk on $\mathbb{Z}$

Besides, in order to formulate the multi-stage stochastic problem, we need to use the fact that each node (t, i) with $t > 0$ in the scenario tree has a (unique) predecessor (t', i') at time $t' < t$. We denote this by: $(t', i') = pred_{t',t}(t, i)$. Each node (t, n) has an associated probability which is the product of transition probabilities on the arcs of the unique path from the root to node (t, n). For our particular example in Figure 7.9, $pred_{0,2}(2, i) = (0, 1)$ for all $i \in \{1, 2, 3, 4\}$,

$pred_{1,2}(2,i) = (1,1)$ for $i \in \{1,2\}$ and $pred_{1,2}(2,i) = (1,2)$ for $i \in \{3,4\}$. Then a dynamic constraint involving decision variables $x_{t,\omega}$ and $x_{t',\omega}$ for given time steps $t, t' \in \mathcal{T}$ with $t < t'$ and a scenario $\omega \in \Omega$, generically denoted:

$$\phi(x_{t,\omega}, x_{t',\omega}) \leq 0$$

writes in our new formulation:

$$\phi(x_{t,pred_{t,t'}(n)}, x_{t',n}) \leq 0,$$

assuming the node (t', n) corresponds to the node at time t' associated with scenario ω.

7.5.3 Motivation of the problem

Consider a scenario tree, the nodes of which are indexed by $(t, n_t)_{t\in\mathcal{T}, 1\leq n_t\leq N_t}$, with associated probabilities $(\pi_{t,n_t})_{0\leq t\leq T, 1\leq n_t\leq N_t}$. For $t \in \mathcal{T}$, $v_{t,t} : (u_t^+, u_t^-) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \bar{\mathbb{R}}$ is a (non-convex) function, and $g_{n_T} : \mathbb{R}^d \times \mathbb{R}^d \mapsto \bar{\mathbb{R}}$ is assumed convex for all $1 \leq n_T \leq N_T$. These functions are assumed extended-valued.

We consider the following multi-stage stochastic problem, denoted by $(\mathcal{P})$:

$$\min_{(u^+,u^-,E^+,E^-)} \sum_{t=0}^{T} \sum_{n_t=1}^{N_t} \pi_{t,n_t} v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) + \sum_{n_T=1}^{N_T} \pi_{T,n_T} g_{n_T}(E^+_{T,n_T}, E^-_{T,n_T}) \tag{7.5.1}$$

$$s.t. \quad \forall t \in \mathcal{T}, \forall n_t \in \{1,...,N_t\}, \forall i \in \{1,...,d\}, \quad E^+_{t,n_t,i} = \sum_{\tau=0}^{t} u^+_{pred_{\tau,t}(t,n_t),i} \Delta t, \tag{7.5.2}$$

$$\forall t \in \mathcal{T}, \forall n_t \in \{1,...,N_t\}, \forall i \in \{1,...,d\}, \quad E^-_{t,n_t,i} = \sum_{\tau=0}^{t} u^-_{pred_{\tau,t}(t,n_t),i} \Delta t, \tag{7.5.3}$$

$$\forall t \in \mathcal{T}, \forall n_t \in \{1,...,N_t\}, \forall i \in \{1,...,d\}, \quad -E_{0,i} \leq E^+_{t,n_t,i} - E^-_{t,n_t,i} \leq \overline{E}_i - E_{0,i}. \tag{7.5.4}$$

Let us give an interpretation of this problem. We consider a setting with $d \geq 1$ batteries. The decision variable $u^+_{t,n_t,i}$ (resp. $u^-_{t,n_t,i}$) denotes the power absorbed (resp. supplied) by battery i at time t for node (t, n_t) of the scenario tree. The quantity $v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t})$ represents the optimal value of an energy management problem (like AC OPF or unit-commitment for instance) at node (t, n_t) of the scenario tree where the power absorbed and supplied by the batteries are fixed at level $(u^+_{t,n_t}, u^-_{t,n_t})$. The decision variables E^+ and E^- represent the total charge and discharge processes associated with the batteries. Their dynamics are respectively given in constraints (7.5.2) and (7.5.3). Constraints on the state of charge of the batteries are given in (7.5.4), while the value of the terminal state is represented by the functions $(g_{n_T})_{1\leq n_T\leq N_T}$. We distinguish between power supplied and absorbed by batteries in order to account for different efficiency loss parameters associated with charge and discharge of the batteries, see constraint (7.2.8) for instance.

Without loss of generality, for all $t \in \mathcal{T}, n_t \in \{1,...,N_t\}$, the bound constraints

$$\forall i \in \{1,...,d\}, \quad 0 \leq u^+_{t,n_t,i} \leq \overline{u}^+_{t,i}, \tag{7.5.5}$$

$$\forall i \in \{1,...,d\}, \quad 0 \leq u^-_{t,n_t,i} \leq \overline{u}^-_{t,i}, \tag{7.5.6}$$

are incorporated in v_{t,n_t}, which is an extended-valued function.

Remark 7.5.8. *More generally, any **static** constraints on $(u^+_{t,n_t}, u^-_{t,n_t})$ for $(t, n_t) \in \mathcal{T} \times \{1,...,N_T\}$ can be incorporated in the extended-valued function v_{t,n_t} and any constraint on $(E^+_{T,n_T}, E^-_{T,n_T})$ can be incorporated in g_{n_T}, for $n_T \in \{1,...,N_T\}$, without loss of generality.*

Problem $(\mathcal{P})$ has $4d\sum_{t=0}^{T} N_t$ scalar constraints in (7.5.4). Let us eliminate variables $E^+_{t,n_t,i}$ and $E^-_{t,n_t,i}$ for all $t \in \mathcal{T} \setminus \{T\}, n_t \in \{1,...,N_t\}$ and $i \in \{1,...,d\}$ using constraints (7.5.2)-(7.5.3) and let us replace g_{n_T} by $h_{n_T} = g_{n_T} + \chi$ with $\chi : \mathbb{R}^d \times \mathbb{R}^d$ and $\chi(E^+, E^-) = +\infty$ if $-E_{0,i} \leq E^+_i - E^-_i \leq \overline{E}_i - E_{0,i}$ is violated for some $i \in \{1,...,d\}$. We can reformulate the problem $(\mathcal{P})$ as:

$$\min_{(u^+,u^-,E^+,E^-)} \sum_{t=0}^{T} \sum_{n_t=1}^{N_t} \pi_{t,n_t} v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) + \sum_{n_T=1}^{N_T} \pi_{T,n_T} h_{n_T}(E^+_{T,n_T}, E^-_{T,n_T}) \tag{7.5.7}$$

$$s.t. \quad \forall n \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad E^+_{T,n,i} = \sum_{\tau=0}^{t} u^+_{pred_{\tau,T}(T,n),i}\Delta t, \tag{7.5.8}$$

$$\forall n \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad E^-_{T,n,i} = \sum_{\tau=0}^{t} u^-_{pred_{\tau,T}(T,n),i}\Delta t, \tag{7.5.9}$$

$$\forall t \in \mathcal{T} \setminus \{T\}, \forall n_t \in \{1,...,N_t\}, \forall i \in \{1,...,d\}, \quad -E_{0,i} \leq \sum_{\tau=0}^{t} (u^+_{pred_{\tau,t}(t,n_t),i} - u^-_{pred_{\tau,t}(t,n_t),i})\Delta t, \tag{7.5.10}$$

$$\forall t \in \mathcal{T} \setminus \{T\}, \forall n_t \in \{1,...,N_t\}, \forall i \in \{1,...,d\}, \quad \sum_{\tau=0}^{t} (u^+_{pred_{\tau,t}(t,n_t),i} - u^-_{pred_{\tau,t}(t,n_t),i})\Delta t \leq \overline{E}_i - E_{0,i}. \tag{7.5.11}$$

For this reformulation, the maximal number of simultaneously active complicating constraints is $N_T d + d\sum_{t=1}^{T} N_t$. In particular, applying Theorem 7.5.7 does not allow us to obtain a better bound on the duality gap than the naive one $\sum_{t=0}^{T}\sum_{n_t=1}^{N_t} \rho(v_{t,n_t})$. For this reason, we propose to consider a modification of the problem, allowing us to substantially reduce the number of scalar constraints while maintaining reasonable interest of the model.

We consider the constraints:

$$\forall n_T \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad 0 \leq E^+_{T,n_T,i} \leq \overline{E}_i - E_{0,i}, \tag{7.5.12}$$

$$\forall n_T \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad 0 \leq E^-_{T,n_T,i} \leq E_{0,i}. \tag{7.5.13}$$

These constraints restrict the total charge and discharge of the batteries on the whole time horizon. It can be shown that inequalities (7.5.2)-(7.5.3)-(7.5.12)-(7.5.13) are sufficient for (7.5.2)-(7.5.3)-(7.5.4) to hold.

For $1 \leq n_T \leq N_T$, define $f_{n_T} = g_{n_T} + \psi$ where $\psi : (E^+, E^-) \in \mathbb{R}^d \times \mathbb{R}^d$ is the indicator function of the set $A = \left(\prod_{1\leq i\leq d}[0, \overline{E}_i - E_{0,i}]\right) \times \left(\prod_{1\leq i\leq d}[0, E_{0,i}]\right)$, which is 0 on A and $+\infty$ outside. Note in particular that f_{n_T} is convex. Consider the problem $(\mathcal{Q})$:

$$\min_{(u^+,u^-,E^+,E^-)} \sum_{t=0}^{T}\sum_{n_t=1}^{N_t} \pi_{t,n_t} v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) + \sum_{n_T=1}^{N_T} \pi_{T,n_T} f_{n_T}(E^+_{n_T}, E^-_{n_T}) \tag{7.5.14}$$

$$s.t. \quad \forall n \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad E^+_{n,i} = \sum_{\tau=0}^{T} u^+_{pred_{\tau,t}(t,n),i}\Delta t, \tag{7.5.15}$$

$$\forall n \in \{1,...,N_T\}, \forall i \in \{1,...,d\}, \quad E^-_{n,i} = \sum_{\tau=0}^{T} u^-_{pred_{\tau,t}(t,n),i}\Delta t. \tag{7.5.16}$$

Then, using the definition of $(h_n)_{1\leq n\leq N_T}$ and the fact that (7.5.12) and (7.5.13) are stronger than (7.5.4), we obtain $val(\mathcal{Q}) \geq val(\mathcal{P})$. Let us now justify the model $(\mathcal{Q})$. For many applications in energy management, it may be desirable to avoid batteries to be overused to avoid excessive aging [Hae14; Car+19a]. We also expect the optimal values of $(\mathcal{Q})$ and $(\mathcal{P})$ to be close for high storage costs. The problem $(\mathcal{Q})$ may thus be seen as a more conservative yet reasonable alternative to $(\mathcal{P})$ in some cases. The interest of problem $(\mathcal{Q})$ is that it has only $2dN_T$ scalar complicating constraints, which can be much smaller than the number of terms in the cost $\sum_{t=0}^{T} N_t + N_T$ yields potential room for using the result [UB16, Theorem 2].

7.5.4 Bounds of the duality gap

Denote v^{**}_{t,n_t} as the convex envelop of the non-convex function v_{t,n_t} for all $(t, n_t) \in \mathcal{T} \times \{1,...,N_t\}$. Define the convexified problem $(\mathcal{Q}^{**})$ by:

$$\min_{(u^+,u^-,E^+,E^-)} \sum_{t=0}^{T}\sum_{n_t=1}^{N_t} \pi_{t,n_t} v^{**}_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) + \sum_{n_T=1}^{N_T} \pi_{T,n_T} h_{n_T}(E^+_{n_T}, E^-_{n_T})$$
$$s.t. \quad (7.5.15) - (7.5.16).$$

We have the following result using Theorem 7.5.7.

Proposition 7.5.9. *There exists an optimal solution (u^+, u^-, E^+, E^-) of Q^{**} and an index set of nodes of the scenario tree $I \subset \{(t, n_t) | t \in \mathcal{T}, n_t \in \{1, ..., N_t\}\}$ with $|I| \leq M := 2dN_T$ such that:*

$$\forall t \in \mathcal{T}, \forall n_t \in \{1, ..., N_t\}, \quad ((t, n_t) \notin I) \Rightarrow \left(v^{**}_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) = v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t})\right).$$

Proof. The number of active constraints in (7.5.15) and (7.5.16) is fixed equal to $2dN_T$, therefore, by the proof of [UB16, Theorem 2], we get the existence of an optimal solution $x = (u^+, u^-, E^+, E^-)$ of (Q^{**}) and of $I \in \mathcal{T} \times \Omega$ with $|I| \leq 2dN_T$ satisfying:

$$\begin{cases} (7.5.2) - (7.5.3) \\ \forall (t, n_t) \notin I, v^{**}_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}) = v_{t,n_t}(u^+_{t,n_t}, u^-_{t,n_t}). \end{cases}$$

Note that we use in particular that h_n are convex for all $n \in \{1, ..., N_T\}$, which proves that the only terms which may not coincide with their convex envelops are the ones related to the non-convex functions. □

By convexity of h_n for all $1 \leq n \leq N_T$, we get the following corollary.

Corollary 7.5.10. *Denote by $\rho_{[i]}$ the i-th order statistics of $(\rho(\pi_{t,n_t} v_{t,n_t}))_{t \in \mathcal{T}, 1 \leq n_t \leq N_t}$ defined by*

$$\max_{(t,n_t)_{t \in \mathcal{T}, 1 \leq n_t \leq N_t}} \rho(\pi_{t,n_t} v_{t,n_t}) = \rho_{[1]} \geq \rho_{[2]} \geq ... \geq \rho_{[\sum_{t=0}^T N_t]} = \min_{(t,n_t)_{t \in \mathcal{T}, 1 \leq n_t \leq N_t}} \rho(\pi_{t,n_t} v_{t,n_t}).$$

We have:

$$0 \leq val(Q) - val(Q^{**}) \leq \sum_{i=1}^{\min(\sum_{t=0}^T n_t, 2dN_T)} \rho_{[i]}.$$

The above corollary is particularly informative if $2dN_T << \sum_{t=0}^T N_t$. Let us give examples of such situations.

Example 7.5.11 (Deterministic dynamic problem)**.** *Assume a deterministic setting, i.e., $N_t = 1$ for all $t \in \mathcal{T}$ with a single battery, i.e., $d = 1$ and consider the problem:*

$$\begin{aligned} \min_{(u^+,u^-,E^+,E^-)} & \frac{1}{T+1} \sum_{t=0}^T v(u_t^+, u_t^-) + g(E^+, E^-) \\ s.t. \quad & E^+ = \frac{1}{T+1} \sum_{t=0}^T u_t^+, \\ & E^- = \frac{1}{T+1} \sum_{t=0}^T u_t^-. \end{aligned}$$

Then we obtain the a priori bound $\frac{1}{T+1} \min(2, T+1)\rho(v) \underset{T \longrightarrow +\infty}{\longrightarrow} 0$ on the duality gap of the problem to be compared with the naive a priori bound $\rho(v)$.

Example 7.5.12 (Two-stage stochastic dynamic problem)**.** *Assume a scenario tree with a comb structure and consider $d = 1$ battery. Let $N \geq 1$ be the number of scenarios considered. Each scenario has an associated probability $1/N$. Consider the problem:*

$$\begin{aligned} \min_{(u^+,u^-,E^+,E^-)} & \frac{1}{T+1} v_0(u_0^+, u_0^-) + \frac{1}{N(T+1)} \sum_{t=1}^T \sum_{n=1}^N v_{t,n}(u_{t,n}^+, u_{t,n}^-) + \frac{1}{N} \sum_{n=1}^N g(E_n^+, E_n^-) \\ s.t. \quad & \forall n \in \{1, ..., N\}, \quad E_n^+ = \frac{1}{T+1} u_0^+ + \frac{1}{T+1} \sum_{t=1}^T u_{t,n}^+, \\ & \forall n \in \{1, ..., N\}, \quad E_n^- = \frac{1}{T+1} u_0^- + \frac{1}{T+1} \sum_{t=1}^T u_{t,n}^-. \end{aligned}$$

Assuming ρ is a uniform upper bound on the non-convexities of the functions $(\rho(v_0), \rho(v_{t,n}))_{1 \leq t \leq T, 1 \leq n \leq N}$, then we obtain the a priori bound on the relaxation gap $\frac{2(N-1)}{N(T+1)}\rho$ to be compared with the naive bound ρ. An example of such a scenario tree with $N = 3$ is given in Figure 7.10.

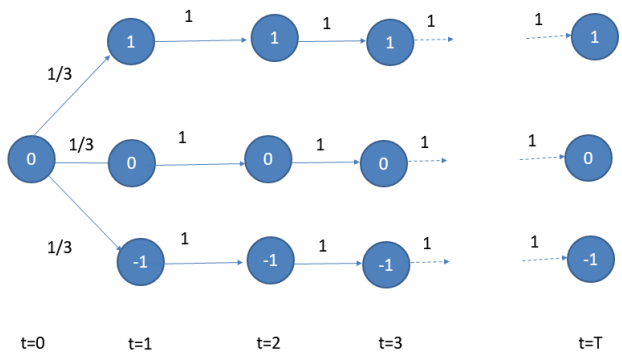


Figure 7.10: Scenario tree with comb structure

Example 7.5.13 (Three-stage stochastic dynamic problem)**.** *Consider the three-stage problem with B_1 nodes at time 1 with equal probability $1/B_1$ (i.e., the scenario tree has B_1 branches at time 1 and B_1B_2 nodes at time $t \geq 2$ with equal probability $1/(B_1B_2)$ (i.e., each node $(1,n)$ has B_2 successors with equal transition probability) :*

$$\min_{(u^+,u^-,E^+,E^-)} \frac{1}{T+1}\left(v_0(u_0^+,u_0^-) + \frac{1}{B_1}\sum_{n=1}^{B_1} v_{1,n}(u_{t,n}^+,u_{t,n}^-) + \frac{1}{B_1B_2}\sum_{t=2}^{T}\sum_{n=1}^{B_1B_2} v_{t,n}(u_{t,n}^+,u_{t,n}^-)\right) + \frac{1}{B_1B_2}\sum_{n=1}^{N} g(E_n^+,E_n^-)$$

$$s.t. \quad \forall n \in \{1,...,B_1B_2\}, \quad E_n^+ = \frac{1}{T+1}\left(u_0^+ + u_{pred_{1,2}}^+(2,n) + \sum_{t=2}^{T} u_{t,n}^+\right),$$

$$\forall n \in \{1,...,B_1B_2\}, \quad E_n^- = \frac{1}{T+1}\left(u_0^- + u_{pred_{1,2}}^-(2,n) + \sum_{t=2}^{T} u_{t,n}^-\right),$$

Assuming ρ is a uniform upper bound on the non-convexities of the functions $(\rho(v_0), \rho(v_{t,n}))_{1\leq t\leq T, 1\leq n\leq N}$, then we obtain the a priori bound on the relaxation gap $\frac{1}{T+1}\left(1 + \frac{B_1}{B_1} + \frac{2B_1B_2-(B_1+1)}{B_1B_2}\right)\rho = \frac{1}{T+1}\left(4 - \frac{B_1+1}{B_1B_2}\right)\rho$ to be compared with the naive bound ρ. An example of such a scenario tree with $B_1 = 3$ and $B_2 = 2$ is given in Figure 7.11.

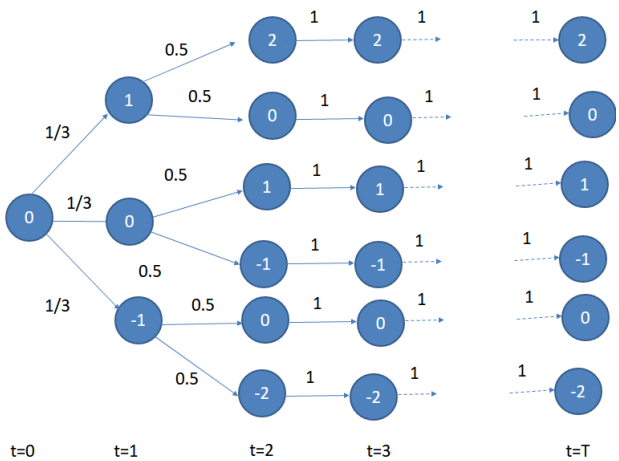


Figure 7.11: Scenario tree with 3 stages

The deterministic setting is the most favorable case as it has the lowest number of complicating constraints. Perhaps, the main limitation of our result is the need to work with the conservative problem $(\mathcal{Q})$ instead of $(\mathcal{P})$, which prevents the batteries to be extensively used. However, for high storage costs and high installed storage capacities, we may expect the value of $(\mathcal{Q})$ to be close to the value of $(\mathcal{P})$. A second limitation is the use of the metric ρ which may be infinite for problems with non-convex domains, as explained in the remark below.

Remark 7.5.14. *Of course, the Unit-Commitment and AC-OPF problems have non-convex domains in general, so that the corresponding non-convexity measures of static sub-problems $\rho(v_{t,n_t})$ and the bound in Corollary 7.5.10 is non-informative. However, in some cases, it can be proved that these functions have convex domain, see for instance results proving convexity of sub-injection regions of AC OPF problems [ZT12; LTZ13; Hua+16]. For problems with non-convex domains, the result still provides a hint that the duality gap may be reduced, due to a convexifying effect of batteries. Besides, for such problems, we believe that Starr's refinement (Theorem 7.5.4 [Sta69]) applied in the context of optimization (to epigraph of the functions appearing in the problem) could yield good results, with an appropriate non-convexity metric to tackle this issue. However, to our knowledge, no explicit formula for the value of this non-convexity metric exists in general.*

7.6 Conclusion

In this chapter, we have introduced a multi-stage stochastic AC OPF problem to account for both uncertainty of renewable production and dynamic constraints of storage systems. We have extended a result ensuring a vanishing duality gap under some realistic a priori conditions for the static deterministic AC OPF problem to the multi-stage stochastic setting. A similar procedure also yields an a posteriori bound on the relaxation gap of the problem. We have illustrated these results numerically on two realistic medium-size networks. Then, we have developed general theoretic results showing explicit a priori bounds on the duality gap of non-convex multi-stage stochastic problems with a few coupling constraints. In particular, we have identified the node-formulation of such multi-stage problems as particularly adapted to apply Shapley-Folkman-type results. These bounds on the duality gap provide a hint that energy storage can not only generate savings in network operations, but can also yield multi-stage stochastic problems with lower relative relaxation gap than their static counterparts, in conservative models where batteries are used in a limited way. Possible extensions of our work include considering unbalanced three-phase networks in a multi-stage stochastic setting in order to consider low-voltage networks.

Conclusion

In this PhD dissertation, we have focused on the impact of the uncertainty of renewable production on power balance and on electricity network management.

We have first considered a mechanism where an individual consumer (or smart grid, which aggregates several consumers/producers) commits in advance to a consumption profile on the electricity network, and controls its energy flexibilities in real-time so as to guarantee that its consumption tracks this commitment curve. We have formulated the underlying decision problem of choosing the commitment profile in advance and controlling in real-time the flexibilities as a Mean-Field Control problem. For this problem, we have obtained results such as necessary and sufficient optimality conditions, existence and uniqueness of a solution, under appropriate assumptions. We have also shown how to solve problems with a Linear-Quadratic structure and we have designed a numerical approximation method based on perturbation theory to solve more general problems.

Then we have formulated a Linear-Quadratic control problem of a large number of Thermostatically Controlled Loads (TCLs), such as heat pumps, hot water tanks, air conditioners..., aiming to track a power signal to provide ancillary services. After deriving the optimality system (FBSDE) for the high-dimensional control problem with common noise using the Stochastic Pontryagin principle, we have shown that this system can be decomposed into FBSDEs in smaller dimension, thereby solving the issue of the curse of dimensionality. Besides, the new FBSDEs characterize the equilibrium of a stochastic Stackelberg differential game, allowing for a decentralized implementation. A mean-field approximation has then been proposed to reduce the need for telecommunications between agents.

In the following chapter, we design a Newton method for stochastic control problems, which is based on successive linearizations of the FBSDE characterizing optimality and obtained by the stochastic Pontryagin principle. After showing that the linearized FBSDEs reduce to Riccati and linear BSDEs, strong convexity and regularity properties for the control problem have been obtained in an appropriate space, which have allowed us to prove global convergence of the Newton method combined with a line-search procedure specifically designed for the infinite dimensional setting. A fully implementable algorithm has been given and applied to the control problem of a large number of batteries providing ancillary services, using regression techniques to solve the BSDEs arising when computing the Newton steps.

Then, we have focused on network operations in an uncertain context created in particular by the random production of decentralized renewable energy sources. After introducing the modeling and state-of-the-art results for the Optimal Power Flow problem, we have formulated a multi-stage stochastic model Alternating Current Optimal Power Flow (AC OPF) problem. This formulation allows us to capture both stochastic (the random production of renewable energy sources) and dynamic aspects (storage) of the problem. We have given realistic a priori conditions ensuring a vanishing relaxation gap. When these assumptions do not hold, we have provided an a posteriori bound on the relaxation gap of the problem. Last, we have shown that the node formulation of multi-stage stochastic problems allows to apply Shapley-Folkman-type results to obtain a priori bounds on the duality gaps of non-convex problems in energy management.

Let us now mention perspectives of this PhD dissertation.

- A perspective of this work, related to Chapter 5, would be to use more advanced regression/machine-learning techniques in order to improve the implementation of the Newton method we proposed. This includes for instance automatic tuning of regression parameters and cross-validation to account for generalization errors.
- Another possible extension, linked to Chapters 6 and 7, would be to propose a stochastic model for the AC

OPF problem in a setting of three-phase unbalanced electricity networks in order to deal with low-voltage networks.

- More generally, convexification methods in the context of combined operations of multi-energy networks also seem very promising. For instance, polynomial optimization methods have recently been used in the context of electricity [Jos16; MH14; Mol+15; Mol+16] and heat networks [Hoh18; HWL19].

Appendix A

Acronyms and Abbreviations

AC OPF	Alternating Current Optimal Power Flow.
BFM	Branch Flow Model.
BIM	Bus Injection Model.
BSDE	Backward Stochastic Differential Equation.
DC OPF	Direct Current Optimal Power Flow.
DNO / DSO	Distribution Network Operator / Distribution System Operator.
EDF	Electricité de France.
FBSDE	Forward-Backward Stochastic Differential Equation.
LLSR	Linear Least-Squares Regression.
LQ	Linear-Quadratic.
MFG	Mean-Field Games.
MKV/MFC	McKean-Vlasov/Mean-Field Control.
ODE	Ordinary Differential Equation.
OPF	Optimal Power Flow.
RES	Renewable energy sources.
RMSE	Root Mean Square Error.
RTE	Réseau de Transport d'Electricité.
SD	Semi-Definite.
SDE	Stochastic Differential Equation.
SOC	Second-Order Cone.
TCL	Thermostatically Controlled Loads.
TNO / TSO	Transmission Network Operator / Transmission System Operator.

Bibliography

[ABVC19] B. Acciaio, J. Backhoff-Veraguas, and R. Carmona. “Extended mean field control problems: stochastic maximum principle and transport perspective”. In: *SIAM Journal on Control and Optimization* 57.6 (2019), pp. 3666–3693.

[Ack+18] W. van Ackooij, I. D. Lopez, A. Frangioni, F. Lacalandra, and M. Tahanan. “Large-scale Unit Commitment under uncertainty: An updated literature survey”. In: *Annals of Operations Research* 271.1 (2018), pp. 11–85.

[AA19] A. Agarkar and H. Agrawal. “A review and vision on authentication and privacy preservation schemes in smart grid network”. In: *Security and Privacy* 2.2 (2019), e62.

[ATM20] C. Alasseur, I. B. Taher, and A. Matoussi. “An extended mean field game for storage in smart grids”. In: *Journal of Optimization Theory and Applications* 184.2 (2020), pp. 644–670.

[AES08] M. H. Albadi and E. F. El-Saadany. “A summary of demand response in electricity markets”. In: *Electric power systems research* 78.11 (2008), pp. 1989–1996.

[ADS99] M. Alexiadis, P. Dokopoulos, and H. Sahsamanoglou. “Wind speed and power forecasting based on spatial correlation models”. In: *IEEE Transactions on Energy Conversion* 14.3 (1999), pp. 836–842.

[ANR17] I. F. Alves, C. Neves, and P. Rosário. “A general estimator for the right endpoint with an application to supercentenarian women’s records”. In: *Extremes* 20.1 (2017), pp. 199–237.

[Ang+19] A. Angiuli, C. Graves, H. Li, J.-F. Chassagneux, F. Delarue, and R. Carmona. “Cemracs 2017: numerical probabilistic approach to MFG”. In: *ESAIM: Proceedings and Surveys* 65 (2019), pp. 84–113.

[Arn+16] D. B. Arnold, M. Sankur, R. Dobbe, K. Brady, D. S. Callaway, and A. Von Meier. “Optimal dispatch of reactive power for voltage regulation and balancing in unbalanced distribution systems”. In: *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE. 2016, pp. 1–5.

[AE76] J.-P. Aubin and I. Ekeland. “Estimates of the duality gap in nonconvex optimization”. In: *Mathematics of Operations Research* 1.3 (1976), pp. 225–245.

[Bad+18] J. Badosa, E. Gobet, M. Grangereau, and D. Kim. “Day-ahead probabilistic forecast of solar irradiance: a Stochastic Differential Equation approach”. In: *Renewable Energy: Forecasting and Risk Management*. Ed. by P. Drobinski, M. Mougeot, D. Picard, R. Plougonven, and P. Tankov. Springer Proceedings in Mathematics & Statistics, 2018. Chap. 4, pp. 73–93.

[BQQ15] X. Bai, L. Qu, and W. Qiao. “Robust AC Optimal Power Flow for power networks with wind power generation”. In: *IEEE Transactions on Power Systems* 31.5 (2015), pp. 4163–4164.

[Bai+08] X. Bai, H. Wei, K. Fujisawa, and Y. Wang. “Semidefinite programming for Optimal Power Flow problems”. In: *International Journal of Electrical Power & Energy Systems* 30.6-7 (2008), pp. 383–392.

[BS11] K Balamurugan and D. Srinivasan. “Review of power flow studies on distribution network with distributed generation”. In: *2011 IEEE Ninth International Conference on Power Electronics and Drive Systems*. IEEE. 2011, pp. 411–417.

[BW89] M. Baran and F. F. Wu. “Optimal sizing of capacitors placed on a radial distribution system”. In: *IEEE Transactions on Power Delivery* 4.1 (1989), pp. 735–743.

[BP18] M. Basei and H. Pham. “A Weak Martingale Approach to Linear-Quadratic McKean–Vlasov Stochastic Control Problems”. In: *Journal of Optimization Theory and Applications* (2018), pp. 1–36.

[Bel+14] F. Bellini, B. Klar, A. M. uller, and E. R. Gianin. “Generalized quantiles as risk measures”. In: *Insurance: Mathematics and Economics* 54 (2014), pp. 41–48.

[BZ+08] C. Bender, J. Zhang, et al. “Time discretization and Markovian iteration for coupled FBSDEs”. In: *The Annals of Applied Probability* 18.1 (2008), pp. 143–177.

[Ben88] A. Bensoussan. *Perturbation methods in optimal control*. Wiley/Gauthier-Villars Series in Modern Applied Mathematics. Translated from the French by C. Tomson. Chichester: John Wiley & Sons Ltd., 1988.

[BFY+13] A. Bensoussan, J. Frehse, P. Yam, et al. *Mean field games and mean field type control theory*. Vol. 101. Springer, 2013.

[BSS05] B. Bibby, I. Skovgaard, and M. Sorensen. “Diffusion-type models with given marginal distribution and autocorrelation function”. In: *Bernoulli* 11.2 (2005), pp. 191–220.

[BCH14] D. Bienstock, M. Chertkov, and S. Harnett. “Chance-constrained Optimal Power Flow: Risk-aware network control under uncertainty”. In: *Siam Review* 56.3 (2014), pp. 461–495.

[BV19] D. Bienstock and A. Verma. “Strong NP-hardness of AC power flows feasibility”. In: *Operations Research Letters* 47.6 (2019), pp. 494–501.

[BAD18] C. Bingane, M. F. Anjos, and S. L. Digabel. “Tight-and-Cheap Conic Relaxation for the AC Optimal Power Flow Problem”. In: *IEEE Transactions on Power Systems* (2018), pp. 1–1. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2848965.

[Bis76] J.-M. Bismut. “Linear quadratic optimal stochastic control with random coefficients”. In: *SIAM Journal on Control and Optimization* 14.3 (1976), pp. 419–444.

[BZ03] J. F. Bonnans and H. Zidani. “Consistency of generalized finite difference schemes for the stochastic HJB equation”. In: *SIAM Journal on Numerical Analysis* 41.3 (2003), pp. 1008–1021.

[Bos+15] S. Bose, D. F. Gayme, K. M. Chandy, and S. H. Low. “Quadratically constrained quadratic programs on acyclic graphs with application to power flow”. In: *IEEE Transactions on Control of Network Systems* 2.3 (2015), pp. 278–287.

[Bos+14] S. Bose, S. H. Low, T. Teeraratkul, and B. Hassibi. “Equivalent relaxations of Optimal Power Flow”. In: *IEEE Transactions on Automatic Control* 60.3 (2014), pp. 729–742.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[Bre10a] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science+Business Media, 2010.

[Bre10b] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York, 2010. ISBN: 9780387709130.

[BM16] A. Bušić and S. Meyn. “Distributed randomized control for demand dispatch”. In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE. 2016, pp. 6964–6971.

[Cad02] A. Cadenillas. “A stochastic maximum principle for systems with jumps, with applications to finance”. In: *Systems & control letters* 47.5 (2002), pp. 433–444.

[Cal11] D. S. Callaway. “Can smaller loads be profitably engaged in power system services?” In: *2011 IEEE Power and Energy Society General Meeting*. IEEE. 2011, pp. 1–3.

[CH10] D. S. Callaway and I. A. Hiskens. “Achieving controllability of electric loads”. In: *Proceedings of the IEEE* 99.1 (2010), pp. 184–199.

[Cam+18a] N. Cammardella, J. Mathias, M. Kiener, A. Bušić, and S. Meyn. “Balancing California’s grid without batteries”. In: *2018 IEEE Conference on Decision and Control (CDC)*. IEEE. 2018, pp. 7314–7321.

[Cam+18b] N. J. Cammardella, R. W. Moye, Y. Chen, and S. P. Meyn. “An energy storage cost comparison: Li-ion batteries vs distributed load control”. In: *2018 Clemson University Power Systems Conference (PSC)*. IEEE. 2018, pp. 1–6.

[Cap+11] F. Capitanescu, J. M. Ramos, P. Panciatici, D. Kirschen, A. M. Marcolini, L. Platbrood, and L. Wehenkel. “State-of-the-art, challenges, and future trends in security constrained Optimal Power Flow”. In: *Electric Power Systems Research* 81.8 (2011), pp. 1731 –1741. ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2011.04.003. URL: http://www.sciencedirect.com/science/article/pii/S0378779611000885.

[Cap16] F. Capitanescu. “Critical review of recent advances and further developments needed in AC Optimal Power Flow”. In: *Electric Power Systems Research* 136.Supplement C (2016), pp. 57 –68. ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2016.02.008. URL: http://www.sciencedirect.com/science/article/pii/S0378779616300141.

[CD15] R. Carmona and F. Delarue. “Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics”. In: *Ann. Probab.* 43.5 (2015), pp. 2647–2700.

[CD18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018.

[CDL13] R. Carmona, F. Delarue, and A. Lachapelle. “Control of McKean-Vlasov dynamics versus mean field games”. In: *Math. Financ. Econ.* 7.2 (2013), pp. 131–166.

[Car62] J. Carpentier. “Contribution to the economic dispatch problem”. In: *Bulletin de la Societe Francoise des Electriciens* 3.8 (1962), pp. 431–447.

[Car+20] P. Carpentier, J.-P. Chancelier, M. De Lara, and F. Pacaud. “Mixed Spatial and Temporal Decompositions for Large-Scale Multistage Stochastic Optimization Problems”. In: *Journal of Optimization Theory and Applications* 186.3 (2020), pp. 985–1005.

[Car+19a] P. Carpentier, J.-P. Chancelier, M. De Lara, and T. Rigaut. “Algorithms for two-time scales stochastic optimization with applications to long term management of energy storage”. In: (2019).

[Car+19b] P. Carpentier, J.-P. Chancelier, M. de Lara, and F. Pacaud. “Upper and lower bounds for large scale multistage stochastic optimization problems: Decomposition methods”. In: *arXiv preprint arXiv:1912.10901* (2019).

[Cha+14] I. C. Change et al. “Mitigation of climate change”. In: *Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* 1454 (2014).

[CL82] F. Chernousko and A. Lyubushin. “Method of successive approximations for solution of optimal control problems”. In: *Optimal Control Applications and Methods* 3.2 (1982), pp. 101–114.

[CM09a] G. Chicco and P. Mancarella. “Distributed multi-generation: A comprehensive view”. In: *Renewable and sustainable energy reviews* 13.3 (2009), pp. 535–551.

[CM09b] G. Chicco and P. Mancarella. “Matrix modelling of small-scale trigeneration systems and application to operational optimization”. In: *Energy* 34.3 (2009), pp. 261–273.

[CR18] C. Coffrin and L. Roald. “Convex relaxations in power system optimization: A brief introduction”. In: *arXiv preprint arXiv:1807.07227* (2018).

[Com19] E. Commission. “Clean energy for all Europeans package”. In: (2019).

[DZG13] E. Dall’Anese, H. Zhu, and G. B. Giannakis. “Distributed Optimal Power Flow for smart microgrids”. In: *IEEE Transactions on Smart Grid* 4.3 (2013), pp. 1464–1475.

[Das+18] C. K. Das, O. Bass, G. Kothapalli, T. S. Mahmoud, and D. Habibi. “Overview of energy storage systems in distribution networks: Placement, sizing, operation, and power quality”. In: *Renewable and Sustainable Energy Reviews* 91 (2018), pp. 1205–1230.

[DP+19] A. De Paola, V. Trovato, D. Angeli, and G. Strbac. “A mean field game approach for distributed control of thermostatic loads acting in simultaneous energy-frequency response markets”. In: *IEEE Transactions on Smart Grid* 10.6 (2019), pp. 5987–5999.

[DG06] F. Delarue and G. Guatteri. “Weak existence and uniqueness for forward-backward SDEs.” In: *Stochastic Processes Appl.* 116.12 (2006), pp. 1712–1742.

[Din+19] T. Ding, R. Lu, Y. Yang, and F. Blaabjerg. “A Condition of Equivalence Between Bus Injection and Branch Flow Models in Radial Networks”. In: *IEEE Transactions on Circuits and Systems II: Express Briefs* (2019).

[DM16] K. Dvijotham and D. K. Molzahn. “Error bounds on the DC power flow approximation: A convex relaxation approach”. In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE. 2016, pp. 2411–2418.

[EPQ97] N. El Karoui, S. Peng, and M. Quenez. “Backward stochastic differential equations in finance”. In: *Math. Finance* 7.1 (1997), pp. 1–71.

[EH08] U. Eminoglu and M. H. Hocaoglu. “Distribution systems forward/backward sweep-based power flow algorithms: A review and comparison study”. In: *Electric Power Components and Systems* 37.1 (2008), pp. 91–110.

[Far+11] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy. “Inverter VAR control for distribution systems with renewables”. In: *2011 IEEE international conference on smart grid communications (SmartGridComm)*. IEEE. 2011, pp. 457–462.

[FL13] M. Farivar and S. H. Low. “Branch flow model: Relaxations and convexification—Part I”. In: *IEEE Transactions on Power Systems* 28.3 (2013), pp. 2554–2564.

[Far+12] M. Farivar, R. Neal, C. Clarke, and S. Low. “Optimal inverter VAR control in distribution systems with high PV penetration”. In: *2012 IEEE Power and Energy Society general meeting*. IEEE. 2012, pp. 1–7.

[FMHL19] A Fonseca-Morales and O Hernández-Lerma. “Stochastic differential games: the potential approach”. In: *Stochastics* (2019), pp. 1–14.

[For+10] D. Forner, T. Erseghe, S. Tomasin, and P. Tenti. “On efficient use of local sources in smart grids with power quality constraints”. In: *2010 First IEEE International Conference on Smart Grid Communications*. IEEE. 2010, pp. 555–560.

[FSR12a] S. Frank, I. Steponavice, and S. Rebennack. “Optimal Power Flow: a bibliographic survey I”. In: *Energy Systems* 3.3 (2012), pp. 221–258.

[FSR12b] S. Frank, I. Steponavice, and S. Rebennack. “Optimal Power Flow: a bibliographic survey II”. In: *Energy Systems* 3.3 (2012), pp. 259–289.

[Gan+12] L. Gan, N. Li, U. Topcu, and S. Low. “On the exactness of convex relaxation for Optimal Power Flow in tree networks”. In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE. 2012, pp. 465–471.

[Gan+14] L. Gan, N. Li, U. Topcu, and S. H. Low. “Exact convex relaxation of Optimal Power Flow in radial networks”. In: *IEEE Transactions on Automatic Control* 60.1 (2014), pp. 72–87.

[GL14] L. Gan and S. H. Low. “Convex relaxations and linear approximation for Optimal Power Flow in multiphase radial networks”. In: *2014 Power Systems Computation Conference*. IEEE. 2014, pp. 1–9.

[GT12] D. Gayme and U. Topcu. “Optimal Power Flow with large-scale storage integration”. In: *IEEE Transactions on Power Systems* 28.2 (2012), pp. 709–717.

[GA05] M. Geidl and G. Andersson. “A modeling and optimization approach for multiple energy carrier power flow”. In: *2005 IEEE Russia Power Tech*. IEEE. 2005, pp. 1–7.

[GA07] M. Geidl and G. Andersson. “Optimal power flow of multiple energy carriers”. In: *IEEE Transactions on power systems* 22.1 (2007), pp. 145–155.

[GKA13] S. Gill, I. Kockar, and G. W. Ault. “Dynamic Optimal Power Flow for active distribution networks”. In: *IEEE Transactions on Power Systems* 29.1 (2013), pp. 121–131.

[Gna+18] T. Gnann, T. S. Stephens, Z. Lin, P. Plötz, C. Liu, and J. Brokate. “What drives the market for plug-in electric vehicles?-A review of international PEV market diffusion models”. In: *Renewable and Sustainable Energy Reviews* 93 (2018), pp. 158–164.

[GT16a] E. Gobet and P. Turkedjiev. “Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions”. In: *Math. Comp.* 85.299 (2016), pp. 1359–1391.

[GT16b] E. Gobet and P. Turkedjiev. “Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions”. In: *Mathematics of Computation* 85.299 (2016), pp. 1359–1391.

[GHH17] T. G. Grandón, H. Heitsch, and R. Henrion. “A joint model of probabilistic/robust constraints for gas transport management in stationary networks”. In: *Computational Management Science* 14.3 (2017), pp. 443–460.

[GSGK18] E. Grover-Silva, R. Girard, and G. Kariniotakis. “Optimal sizing and placement of distribution grid connected battery systems through an SOCP Optimal Power Flow algorithm”. In: *Applied Energy* 219 (2018), pp. 385–393.

[GŠ09] I. Gyöngy and D. Šiška. “On finite-difference approximations for normalized Bellman equations”. In: *Applied Mathematics and Optimization* 60.3 (2009), p. 297.

[Håb19] M. Håberg. “Fundamentals and recent developments in stochastic unit commitment”. In: *International Journal of Electrical Power & Energy Systems* 109 (2019), pp. 38–48.

[Hae14] P. Haessig. “Dimensionnement et gestion d’énergie pour l’atténuation des incertitudes de production éolienne”. PhD thesis. École normale supérieure de Cachan-ENS Cachan, 2014.

[HPC18] L. Halilbašić, P. Pinson, and S. Chatzivasileiadis. “Convex relaxations and approximations of chance-constrained AC-OPF problems”. In: *IEEE Transactions on Power Systems* 34.2 (2018), pp. 1459–1470.

[HJW18] J. Han, A. Jentzen, and E Weinan. “Solving high-dimensional partial differential equations using deep learning”. In: *Proceedings of the National Academy of Sciences* 115.34 (2018), pp. 8505–8510.

[HL20] J. Han and J. Long. “Convergence of the deep BSDE method for coupled FBSDEs”. In: *Probability, Uncertainty and Quantitative Risk* 5.1 (2020), pp. 1–33.

[Hao+13] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent. “A generalized battery model of a collection of thermostatically controlled loads for providing ancillary service”. In: *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2013, pp. 551–558.

[Hao+14] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent. “Aggregate flexibility of thermostatically controlled loads”. In: *IEEE Transactions on Power Systems* 30.1 (2014), pp. 189–198.

[HRS06] H. Heitsch, W. Römisch, and C. Strugarek. “Stability of multistage stochastic programs”. In: *SIAM Journal on Optimization* 17.2 (2006), pp. 511–525.

[Hei19] H. Heitsch. “On probabilistic capacity maximization in a stationary gas network”. In: *Optimization* (2019).

[HR09] H. Heitsch and W. Römisch. “Scenario tree modeling for multistage stochastic programs”. In: *Mathematical Programming* 118.2 (2009), pp. 371–406.

[Hey+15] B. Heymann, J. Bonnans, P. Martinon, F. Silva, F. Lanas, and G. Jiménez-Estévez. “Continuous optimal control approaches to microgrid energy management”. In: *Energy Systems* (2015), pp. 1–19.

[Hey+16] B. Heymann, J. Bonnans, F. Silva, and G. Jimenez. “A stochastic continuous time model for microgrid energy management”. In: *Control Conference (ECC), 2016 European*. IEEE. 2016, pp. 2084–2089.

[Hoh18] M. Hohmann. *Polynomial Optimization in Energy Systems*. ETH Zurich, 2018.

[HWL19] M. Hohmann, J. Warrington, and J. Lygeros. “A two-stage polynomial approach to stochastic optimization of district heating networks”. In: *Sustainable Energy, Grids and Networks* 17 (2019), p. 100177.

[How60] R. A. Howard. “Dynamic Programming and Markov Processes.” In: (1960).

[HW01] K. Høyland and S. W. Wallace. “Generating scenario trees for multistage decision problems”. In: *Management science* 47.2 (2001), pp. 295–307.

[HP95] Y. Hu and S. Peng. “Solution of forward-backward stochastic differential equations”. In: *Probability Theory and Related Fields* 103.2 (1995), pp. 273–283.

[HMC+06] M. Huang, R. P. Malhamé, P. E. Caines, et al. “Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.

[Hua+16] S. Huang, Q. Wu, J. Wang, and H. Zhao. “A sufficient condition on convex relaxation of AC Optimal Power Flow in distribution networks”. In: *IEEE Transactions on Power Systems* 32.2 (2016), pp. 1359–1368.

[Hus+07] Z. Hussien, L. Cheung, M. Siam, and A. Ismail. “Modeling of sodium sulfur battery for power system applications”. In: *Elektrika Journal of Electrical Engineering* 9.2 (2007), pp. 66–72.

[IS87] M. Irving and M. Sterling. “Efficient Newton-Raphson algorithm for load-flow calculation in transmission and distribution networks”. In: *IEE Proceedings C (Generation, Transmission and Distribution)*. Vol. 134. 5. IET. 1987, pp. 325–330.

[IMM14] E. B. Iversen, J. M. Morales, and H. Madsen. “Optimal charging of an electric vehicle using a Markov decision process”. In: *Applied Energy* 123 (2014), pp. 1–12.

[Jab06] R. A. Jabr. “Radial distribution load flow using conic programming”. In: *IEEE transactions on power systems* 21.3 (2006), pp. 1458–1459.

[JKK14] R. A. Jabr, S. Karaki, and J. A. Korbane. “Robust multi-period OPF with storage and renewables”. In: *IEEE Transactions on Power Systems* 30.5 (2014), pp. 2790–2799.

[Jac+19] P. Jacquot, O. Beaude, P. Benchimol, S. Gaubert, and N. Oudjane. “A privacy-preserving disaggregation algorithm for non-intrusive management of flexible energy”. In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, pp. 890–896.

[Jac+18] P. Jacquot, O. Beaude, S. Gaubert, and N. Oudjane. “Analysis and implementation of an hourly billing mechanism for demand response management”. In: *IEEE Transactions on Smart Grid* 10.4 (2018), pp. 4265–4278.

[JW18] P. Jacquot and C. Wan. “Nonsmooth aggregative games with coupling constraints and infinitely many classes of players”. In: *arXiv preprint arXiv:1806.06230* (2018).

[Ji+20] S. Ji, S. Peng, Y. Peng, and X. Zhang. “Three algorithms for solving high-dimensional fully-coupled FBSDEs through deep learning”. In: *IEEE Intelligent Systems* (2020).

[Jos16] C. Josz. “Application of polynomial optimization to electricity transmission networks”. Thèse de doctorat dirigée par Gilbert, Jean Charles Mathématiques Appliquées Paris 6 2016. PhD thesis. 2016. URL: http://www.theses.fr/2016PA066352.

[Jov+15] S. Jovanović, S. Savić, M. Bojić, Z. Djordjević, and D. Nikolić. “The impact of the mean daily air temperature change on electricity consumption”. In: *Energy* 88 (2015), pp. 604–609.

[Kan48] L. V. Kantorovich. “Functional analysis and applied mathematics”. In: *Uspekhi Matematicheskikh Nauk* 3.6 (1948), pp. 89–185.

[KS98] I. Karatzas and S. E. Shreve. “Brownian motion”. In: *Brownian Motion and Stochastic Calculus*. Springer, 1998, pp. 47–127.

[KŠS20a] B. Kerimkulov, D. Šiška, and Ł. Szpruch. “A modified MSA for stochastic control problems”. In: *arXiv preprint arXiv:2007.05209* (2020).

[KŠS20b] B. Kerimkulov, D. Šiška, and L. Szpruch. “Exponential Convergence and Stability of Howard’s Policy Improvement Algorithm for Controlled Diffusions”. In: *SIAM Journal on Control and Optimization* 58.3 (2020), pp. 1314–1340.

[KM13] A. C. Kizilkale and R. P. Malhamé. “Mean field based control of power system dispersed energy storage devices for peak load relief”. In: *52nd IEEE Conference on Decision and Control*. IEEE. 2013, pp. 4971–4976.

[KM16] A. C. Kizilkale and R. P. Malhamé. “Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads”. In: *Control of Complex Systems*. Elsevier, 2016, pp. 559–584.

[KL19] E. de Klerk and M. Laurent. “A survey of semidefinite programming approaches to the generalized problem of moments and their error analysis”. In: *World Women in Mathematics 2018*. Springer, 2019, pp. 17–56.

[Koc+15] T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. *Evaluating gas network capacities*. SIAM, 2015.

[KDS16] B. Kocuk, S. S. Dey, and X. A. Sun. “Strong SOCP relaxations for the Optimal Power Flow problem”. In: *Operations Research* 64.6 (2016), pp. 1177–1196.

[KDS15] B. Kocuk, S. S. Dey, and X. A. Sun. “Inexactness of SDP relaxation and valid inequalities for Optimal Power Flow”. In: *IEEE Transactions on Power Systems* 31.1 (2015), pp. 642–651.

[Kry08] N. V. Krylov. *Controlled Diffusion Processes*. Vol. 14. Springer Science & Business Media, 2008.

[LL11] J. Lavaei and S. H. Low. “Zero duality gap in Optimal Power Flow problem”. In: *IEEE Transactions on Power Systems* 27.1 (2011), pp. 92–107.

[LTZ13] J. Lavaei, D. Tse, and B. Zhang. “Geometry of power flows and optimization in distribution networks”. In: *IEEE Transactions on Power Systems* 29.2 (2013), pp. 572–583.

[LGVH15] K. Lehmann, A. Grastien, and P. Van Hentenryck. “AC-feasibility on tree networks is NP-hard”. In: *IEEE Transactions on Power Systems* 31.1 (2015), pp. 798–801.

[LLW18] Y. Liu, J. Li, and L. Wu. “Coordinated optimal network reconfiguration and voltage regulator/DER control for unbalanced distribution systems”. In: *IEEE Transactions on Smart Grid* 10.3 (2018), pp. 2912–2922.

[Liu+17] Y. Liu, J. Li, L. Wu, and T. Ortmeyer. “Chordal relaxation based ACOPF for unbalanced distribution systems with DERs and voltage regulation devices”. In: *IEEE Transactions on Power Systems* 33.1 (2017), pp. 970–984.

[LS17] A. Lorca and X. A. Sun. “The adaptive robust multi-period alternating current Optimal Power Flow problem”. In: *IEEE Transactions on Power Systems* 33.2 (2017), pp. 1993–2003.

[Low14a] S. Low. “Convex Relaxation of Optimal Power Flow - Part 1: Formulations and Equivalence”. In: *IEEE Transactions on Control of Network Systems* 1.1 (2014), pp. 15–27. ISSN: 2325-5870. DOI: 10.1109/TCNS.2014.2309732.

[Low14b] S. Low. “Convex Relaxation of Optimal Power Flow - Part 2: Exactness”. In: *IEEE Transactions on Control of Network Systems* 1.2 (2014), pp. 177–189. DOI: 10.1109/TCNS.2014.2323634.

[MY99] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. A course on stochastic processes. Lecture Notes in Mathematics, 1702, Springer-Verlag, 1999.

[MPY94] J. Ma, P. Protter, and J. Yong. “Solving forward-backward stochastic differential equations explicitly—a four step scheme”. In: *Probability theory and related fields* 98.3 (1994), pp. 339–359.

[Ma+15] J. Ma, Z. Wu, D. Zhang, J. Zhang, et al. “On well-posedness of forward–backward SDEs—A unified approach”. In: *The Annals of Applied Probability* 25.4 (2015), pp. 2168–2214.

[MAL15] R. Madani, M. Ashraphijuo, and J. Lavaei. “Promises of conic relaxation for contingency-constrained Optimal Power Flow problem”. In: *IEEE Transactions on Power Systems* 31.2 (2015), pp. 1297–1307.

[MA+16] C. B. Martinez-Anido, B. Botor, A. R. Florita, C. Draxl, S. Lu, H. F. Hamann, and B.-M. Hodge. “The value of day-ahead solar power forecasting improvement”. In: *Solar Energy* 129 (2016), pp. 192–203.

[Mat+12] J Mathieu, M. Dyson, D. Callaway, and A Rosenfeld. “Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations”. In: *ACEEE Summer Study on Energy Efficiency in Buildings*. Citeseer. 2012, pp. 189–203.

[MMS19] A. Matoussi, A. Manai, and R. Salhi. “Mean-Field Backward-Forward SDE with Jumps and Storage problem in Smart Grids”. In: *arXiv preprint arXiv:1906.08525* (2019).

[Mol+15] D. Molzahn, C. Josz, I. Hiskens, and P. Panciatici. “Solution of Optimal Power Flow problems using moment relaxations augmented with objective function penalization”. In: *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE. 2015, pp. 31–38.

[MH14] D. K. Molzahn and I. A. Hiskens. “Moment-based relaxation of the optimal power flow problem”. In: *2014 Power Systems Computation Conference*. IEEE. 2014, pp. 1–7.

[MH+19] D. K. Molzahn, I. A. Hiskens, et al. *A survey of relaxations and approximations of the power flow equations*. Now Publishers, 2019.

[Mol+16] D. K. Molzahn, C. Josz, I. A. Hiskens, and P. Panciatici. “A Laplacian-based approach for finding near globally optimal solutions to OPF problems”. In: *IEEE Transactions on Power Systems* 32.1 (2016), pp. 305–315.

[Mor+14] J. Morales, A. Conejo, H. Madsen, P. Pinson, and M. Zugno. *Integrating Renewables in Electricity Markets*. Vol. 205. International Series in Operations Research & Management Science. Springer Science + Business Media New York, 2014.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[NT15] H. D. Nguyen and K. Turitsyn. “Voltage multistability and pulse emergency control for distribution system with power flow reversal”. In: *IEEE Transactions on Smart Grid* 6.6 (2015), pp. 2985–2996.

[NCP14] M. Nick, R. Cherkaoui, and M. Paolone. “Stochastic day-ahead optimal scheduling of Active Distribution Networks with dispersed energy storage and renewable resources”. In: *2014 IEEE Conference on Technologies for Sustainability (SusTech)*. IEEE. 2014, pp. 91–96.

[NW06] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[Ode+18] B. Odetayo, M. Kazemi, J. MacCormack, W. Rosehart, H. Zareipour, and A. R. Seifi. “A Chance Constrained Programming Approach to the Integrated Planning of Electric Power Generation, Natural Gas Network and Storage”. In: *IEEE Transactions on Power Systems* (2018), pp. 1–1. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2833465.

[ØZ12] B. Øksendal and T. Zhang. “Backward stochastic differential equations with respect to general filtrations and applications to insider finance”. In: *Communications on Stochastic Analysis* 6.4 (2012), p. 13.

[Ord+17] C. Ordoudis, S. Delikaraoglou, P. Pinson, and J. Kazempour. “Exploiting flexibility in coupled electricity and natural gas markets: A price-based approach”. In: *2017 IEEE Manchester PowerTech*. 2017, pp. 1–6. DOI: 10.1109/PTC.2017.7981047.

[OCS04] T. J. Overbye, X. Cheng, and Y. Sun. “A comparison of the AC and DC power flow models for LMP calculations”. In: *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*. IEEE. 2004, 9–pp.

[Pac+18] F. Pacaud, P. Carpentier, J.-P. Chancelier, and M. De Lara. “Stochastic optimal control of a domestic microgrid equipped with solar panel and battery”. In: *arXiv preprint arXiv:1801.06479* (2018).

[PMV02] A. Pardo, V. Meneu, and E. Valor. “Temperature and seasonality influences on Spanish electricity load”. In: *Energy Economics* 24.1 (2002), pp. 55–70.

[PT99] E. Pardoux and S. Tang. “Forward-backward stochastic differential equations and quasilinear parabolic PDEs”. In: *Probability Theory and Related Fields* 114.2 (1999), pp. 123–150. ISSN: 1432-2064. DOI: 10.1007/s004409970001. URL: https://doi.org/10.1007/s004409970001.

[Pen90] S. Peng. “A general stochastic maximum principle for optimal control problems”. In: *SIAM J. Control Optim.* 28.4 (1990), pp. 966–979.

[PW99] S. Peng and Z. Wu. “Fully coupled forward-backward stochastic differential equations and applications to optimal control”. In: *SIAM Journal on Control and Optimization* 37.3 (1999), pp. 825–843.

[PP12] G. C. Pflug and A. Pichler. “A distance for multistage stochastic optimization models”. In: *SIAM Journal on Optimization* 22.1 (2012), pp. 1–23.

[PP14] G. C. Pflug and A. Pichler. *Multistage stochastic optimization*. Springer, 2014.

[PP15] G. C. Pflug and A. Pichler. “Dynamic generation of scenario trees”. In: *Computational Optimization and Applications* 62.3 (2015), pp. 641–668.

[PW16] H. Pham and X. Wei. “Discrete time McKean–Vlasov control problem: a dynamic programming approach”. In: *Applied Mathematics & Optimization* 74.3 (2016), pp. 487–506.

[PW18] H. Pham and X. Wei. “Bellman equation and viscosity solutions for mean-field stochastic control problem”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 24.1 (2018), pp. 437–461.

[Pha09] H. Pham. *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Springer Science & Business Media, 2009.

[Pin+09] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl. “From probabilistic forecasts to statistical scenarios of short-term wind power production”. In: *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology* 12.1 (2009), pp. 51–62.

[Pro03] P. E. Protter. *Stochastic Integration and Differential Equations*. second. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. ISBN: 978-3-662-10061-5. DOI: 10.1007/978-3-662-10061-5_6. URL: https://doi.org/10.1007/978-3-662-10061-5_6.

[RA17] L. Roald and G. Andersson. “Chance-constrained AC Optimal Power Flow: Reformulations and efficient algorithms”. In: *IEEE Transactions on Power Systems* 33.3 (2017), pp. 2906–2918.

[Roa+16] L. Roald, S. Misra, T. Krause, and G. Andersson. “Corrective control to handle forecast uncertainty: A chance constrained Optimal Power Flow”. In: *IEEE Transactions on Power Systems* 32.2 (2016), pp. 1626–1637.

[Roa+13] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson. “Analytical reformulation of Security constrained Optimal Power Flow with probabilistic constraints”. In: *2013 IEEE Grenoble Conference*. IEEE. 2013, pp. 1–6.

[RMT17] L. A. Roald, D. K. Molzahn, and A. F. Tobler. “Power system optimization with uncertainty and AC power flow: Analysis of an iterative algorithm”. In: *10th IREP Symp. Bulk Power Syst. Dynamics Control*. 2017.

[Roc70] R. T. Rockafellar. *Convex analysis*. 28. Princeton university press, 1970.

[ROW14] U. Rubasheuski, J. Oppen, and D. L. Woodruff. “Multi-stage scenario generation by the combined moment matching and scenario reduction method”. In: *Operations Research Letters* 42.5 (2014), pp. 374–377.

[ST14] M. Sandhu and T. Thakur. “Issues, challenges, causes, impacts and utilization of renewable energy sources-grid integration”. In: *Int. Journal of Engineering Research and Applications* 4.3 (2014), pp. 636–643.

[SG16] M. J. Sanjari and H. Gooi. “Probabilistic forecast of PV power generation based on higher order Markov chain”. In: *IEEE Transactions on Power Systems* 32.4 (2016), pp. 2942–2952.

[San+16] M. D. Sankur, R. Dobbe, E. Stewart, D. S. Callaway, and D. B. Arnold. “A linearized power flow model for optimization in unbalanced distribution systems”. In: *arXiv preprint arXiv:1606.04492* (2016).

[Sch14] R. Schneider. *Convex bodies: the Brunn–Minkowski theory*. 151. Cambridge university press, 2014.

[Seg+20] A. Seguret, C. Alasseur, J. F. Bonnans, A. De Paola, N. Oudjane, and V. Trovato. “Decomposition of high dimensional aggregative stochastic control problems”. In: *arXiv preprint arXiv:2008.09827* (2020).

[ŞC14] N. Şen and P. E. Caines. “Mean field games with partially observed major player and stochastic mean field”. In: *53rd IEEE Conference on Decision and Control*. IEEE. 2014, pp. 2709–2715.

[Sha06] A. Shapiro. “On complexity of multistage stochastic programs”. In: *Operations Research Letters* 34.1 (2006), pp. 1–8.

[SDR14] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

[SL12] S. Sojoudi and J. Lavaei. “Physics of power networks makes hard optimization problems easy to solve”. In: *2012 IEEE Power and Energy Society General Meeting*. IEEE. 2012, pp. 1–8. DOI: 10.1109/PESGM.2012.6345272.

[SL11] S. Sojoudi and J. Lavaei. “Network topologies guaranteeing zero duality gap for Optimal Power Flow problem”. In: *submitted for publication* (2011).

[SM16] I. B. Sperstad and H. Marthinsen. “Optimal Power Flow methods and their application to distribution systems with energy storage: A survey of available tools and methods”. In: *SINTEF Energi. Rapport* (2016).

[Sta69] R. M. Starr. “Quasi-equilibria in markets with non-convex preferences”. In: *Econometrica: journal of the Econometric Society* (1969), pp. 25–38.

[SJA09] B. Stott, J. Jardim, and O. Alsaç. “DC power flow revisited”. In: *IEEE Transactions on Power Systems* 24.3 (2009), pp. 1290–1300.

[Suc+20] D. Suchet, A. Jeantet, T. Elghozi, and Z. Jehl. “Defining and Quantifying Intermittency in the Power Sector”. In: *Energies* 13.13 (2020), p. 3366.

[SSM16] C. Sun, F. Sun, and S. J. Moura. “Nonlinear predictive energy management of residential buildings with photovoltaics & batteries”. In: *Journal of Power Sources* 325 (2016), pp. 723–731.

[Sun+16] G. Sun, Y. Li, S. Chen, Z. Wei, S. Chen, and H. Zang. “Dynamic stochastic Optimal Power Flow of wind power and the electric vehicle integrated power system considering temporal-spatial characteristics”. In: *Journal of Renewable and Sustainable Energy* 8.5 (2016), p. 053309.

[Swa17] B. P. Swaminathan. “Operational Planning of Active Distribution Networks-Convex Relaxation under Uncertainty”. PhD thesis. Université Grenoble Alpes, 2017.

[TRY16] K. M. Tan, V. K. Ramachandaramurthy, and J. Y. Yong. “Integration of electric vehicles in smart grid: A review on vehicle to grid technologies and optimization techniques”. In: *Renewable and Sustainable Energy Reviews* 53 (2016), pp. 720–732.

[TTS15] S. H. Tindemans, V. Trovato, and G. Strbac. “Decentralized control of thermostatic loads for flexible demand response”. In: *IEEE Transactions on Control Systems Technology* 23.5 (2015), pp. 1685–1700.

[Tro+16] V. Trovato, I. M. Sanz, B. Chaudhuri, and G. Strbac. “Advanced control of thermostatic loads for rapid frequency response in Great Britain”. In: *IEEE Transactions on Power Systems* 32.3 (2016), pp. 2106–2117.

[TTS16] V. Trovato, S. H. Tindemans, and G. Strbac. “Leaky storage model for optimal multi-service allocation of thermostatic loads”. In: *IET Generation, Transmission & Distribution* 10.3 (2016), pp. 585–593.

[UB16] M. Udell and S. Boyd. “Bounding duality gap for separable problems with linear constraints”. In: *Computational Optimization and Applications* 64.2 (2016), pp. 355–378.

[UV+10] C. Unsihuay-Vila, J. W. Marangon-Lima, A. Z. De Souza, I. J. Perez-Arriaga, and P. P. Balestrassi. “A model to long-term, multiarea, multistage, and integrated expansion planning of electricity and natural gas systems”. In: *IEEE Transactions on Power Systems* 25.2 (2010), pp. 1154–1168.

[VB+17] I Van Beuzekom, M Gibescu, P. Pinson, and J. Slootweg. “Optimal planning of integrated multi-energy systems”. In: *2017 IEEE Manchester PowerTech*. IEEE. 2017, pp. 1–6.

[VBGS15] I Van Beuzekom, M Gibescu, and J. Slootweg. “A review of multi-energy system planning and optimization tools for sustainable urban development”. In: *2015 IEEE Eindhoven PowerTech*. IEEE. 2015, pp. 1–7.

[VCM20] A. Venzke, S. Chatzivasileiadis, and D. K. Molzahn. “Inexact convex relaxations for AC Optimal Power Flow: Towards AC feasibility”. In: *Electric Power Systems Research* 187 (2020), p. 106480.

[Ven+17] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis. “Convex relaxations of chance constrained AC Optimal Power Flow”. In: *IEEE Transactions on Power Systems* 33.3 (2017), pp. 2829–2841.

[Vra+13] M. Vrakopoulou, M. Katsampani, K. Margellos, J. Lygeros, and G. Andersson. “Probabilistic security-constrained AC Optimal Power Flow”. In: *2013 IEEE Grenoble Conference*. IEEE. 2013, pp. 1–6.

[WB19] X. Wang and A. Barnett. “The evolving value of photovoltaic module efficiency”. In: *Applied Sciences* 9.6 (2019), p. 1227.

[Wu+16] X. Wu, X. Hu, S. Moura, X. Yin, and V. Pickert. “Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array”. In: *Journal of Power Sources* 333 (2016), pp. 203–212.

[Yao86] A. C.-C. Yao. “How to generate and exchange secrets”. In: *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*. IEEE. 1986, pp. 162–167.

[Yon06] J. Yong. “Linear forward-backward stochastic differential equations with random coefficients”. In: *Probability theory and related fields* 135.1 (2006), pp. 53–83.

[Yon13] J. Yong. “Linear-quadratic optimal control problems for mean-field stochastic differential equations”. In: *SIAM journal on Control and Optimization* 51.4 (2013), pp. 2809–2838.

[Yon97] J. Yong. “Finding adapted solutions of forward–backward stochastic differential equations: Method of continuation”. In: *Probability Theory and Related Fields* 107.4 (1997), pp. 537–572.

[Yon99] J. Yong. “Linear Forward—Backward Stochastic Differential Equations”. In: *Applied Mathematics and Optimization* 39.1 (1999), pp. 93–119.

[ZT11] B. Zhang and D. Tse. “Geometry of feasible injection region of power networks”. In: *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2011, pp. 1508–1515.

[ZT12] B. Zhang and D. Tse. “Geometry of injection regions of power networks”. In: *IEEE Transactions on Power Systems* 28.2 (2012), pp. 788–797.

[ZDK16] B. Zhang, P. Dehghanian, and M. Kezunovic. “Spatial-temporal solar power forecast through use of Gaussian conditional random fields”. In: *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE. 2016, pp. 1–5.

[Zha17] J. Zhang. “Backward stochastic differential equations”. In: *Backward Stochastic Differential Equations*. Springer, 2017, pp. 79–99.

[ZSM16] Y. Zhang, S. Shen, and J. L. Mathieu. “Distributionally robust chance-constrained Optimal Power Flow with uncertain renewables and uncertain reserves provided by loads”. In: *IEEE Transactions on Power Systems* 32.2 (2016), pp. 1378–1388.

[Zoh+20] F. Zohrizadeh, C. Josz, M. Jin, R. Madani, J. Lavaei, and S. Sojoudi. “A survey on conic relaxations of Optimal Power Flow problem”. In: *European journal of operational research* (2020).

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Contrôle optimal de flexibilités énergétiques en contexte incertain

Mots clés : Réseaux intelligents, contrôle champs moyen, optimisation stochastique

Résumé : Dans cette thèse, nous utilisons des outils provenant du contrôle optimal stochastique et de l'optimisation stochastique et convexe afin de développer des mécanismes pour piloter des moyens de stockage énergétique permettant de gérer l'incertitude de production des sources d'énergie intermittentes (solaire et éolien). Tout d'abord, nous introduisons un mécanisme dans lequel un consommateur s'engage à suivre un profil de consommation sur le réseau, et contrôle ensuite ses systèmes de stockage pour suivre ce profil en temps réel. Nous modélisons cette situation par un problème de contrôle à champ moyen, pour lequel nous obtenons des résultats théoriques et numériques. Puis, nous introduisons un problème de contrôle d'un grand nombre d'unités de stockage thermique soumises à un bruit commun et fournissant des services au réseau. Nous montrons que ce problème de contrôle peut être remplacé par un problème de jeu différentiel stochastique. Ceci permet un schéma de contrôle décentralisé avec des garanties de performance, tout en préservant la confidentialité des données des consommateurs et en limitant les besoins en télécommunication. Ensuite, nous développons une méthode de Newton pour des problèmes de contrôle stochastique. Nous montrons que le pas de Newton peut être calculé en résolvant des Equations Différentielles Stochastiques Rétrogrades, puis nous proposons une méthode de recherche linéaire appropriée, et prouvons la convergence globale de la methode de Newton obtenue dans un espace adéquat. Sa performance numérique est illustrée sur un problème de contrôle d'un grand nombre de batteries fournissant des services au réseau. Enfin, nous étudions l'extension au cas stochastique multi-étapes du problème "Alternating Current Optimal Power Flow" afin de piloter un réseau électrique équipé de systèmes de stockage. Pour ce problème, nous donnons des conditions réalistes et vérifiables a priori garantissant l'absence de saut de relaxation, ainsi qu'une borne a posteriori sur celui-ci. Dans le cadre plus large de problèmes multi-étapes non-convexes avec une structure générique, nous établissons également des bornes a priori sur le saut de dualité, en nous basant sur des résultats liés au Théorème de Shapley-Folkman.

Title : Optimal control of energy flexibilities in a stochastic environment

Keywords : Smart grids, Mean-Field Control, Stochastic Optimization

Abstract : In this PhD dissertation, we use tools from stochastic optimal control, stochastic optimization and convex optimization to design mechanisms to control energy storage systems, to deal with the challenges created by the uncertain production of intermittent energy sources. First, we introduce a commitment mechanism where an individual consumer chooses a consumption profile, then controls its storage devices to track in real-time this profile. We formulate a Mean-Field Control problem to model this situation, for which we establish theoretic and numerical results. Second, we introduce a control problem for a large population of Thermostatically Controlled Loads (TCLs) subject to a common noise and providing ancillary services to the grid. We show that the centralized control problem can be replaced by a stochastic differential game with minimal information-sharing. This allows for a decentralized control scheme with performance guarantees, while preserving privacy of consumers and limiting telecommunication requirements. We then develop a Newton method for stochastic control problems. We show that the computation of the Newton step reduces to solving Backward Stochastic Differential Equations, then we design an appropriate line-search procedure and prove global convergence of the Newton method with line-search in an appropriate space. Its performance is illustrated on a problem of control of a large number of batteries providing services to the grid. Last, a multi-stage stochastic Alternating Current Optimal Power Flow problem is formulated in order to control a power network equipped with energy storage systems. A priori conditions ensuring a vanishing relaxation gap are derived and an easily computable a posteriori bound on the relaxation gap of the problem is given. Using Shapley-Folkman-type results, a priori bounds on the duality gap of non-convex multi-stage stochastic problems with a generic structure are derived.

Institut Polytechnique de Paris
91120 Palaiseau, France