NNT : 2019SACLX101

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PARIS-SACLAY

préparée à

L'ÉCOLE POLYTECHNIQUE

Centre de Mathématiques Appliquées (UMR 7641 CNRS)

ÉCOLE DOCTORALE N°574
Mathématiques Hadamard (EDMH)

Spécialité de doctorat: **Mathématiques appliquées**

par

Paulin JACQUOT

Game theory and Optimization Methods for Decentralized Electric Systems

---

Thèse présentée et soutenue à Palaiseau, le 5 décembre 2019

Composition du jury:

Sylvain SORIN	Professeur Émerite	Sorbonne Université	*Président du Jury*
Bruno GAUJAL	Directeur de Recherche	Inria	*Rapporteur*
Marco SCARSINI	Professeur	LUISS, Roma, Italy	*Rapporteur*
Stéphane GAUBERT	Directeur de Recherche	Inria, École polytechnique	*Directeur de thèse*
Nadia OUDJANE	Ingénieur-chercheuse sénior	EDF R&D	*Co-directrice de thèse*
Olivier BEAUDE	Ingénieur-chercheur	EDF R&D	*Co-directeur de thèse (invité)*
Emmanuel GOBET	Professeur	École polytechnique	*Examinateur*
Jason MARDEN	Professeur Associé	UC Santa Barbara	*Examinateur*
Mathias STAUDIGL	Professeur Associé	Maastricht University	*Examinateur*

December 6, 2019

Remerciements

Acknowledgements

Je remercie très sincèrement Bruno Gaujal et Marco Scarsini d'avoir mis à disposition leur grande expertise et leur temps pour rapporter cette thèse. Je souhaite également remercier vivement Jason Marden, Mathias Staudigl, Sylvain Sorin et Emmanuel Gobet de me faire l'honneur de prendre part au jury de cette thèse.

Je tiens à exprimer toute ma reconnaissance à mes encadrants Stéphane, Nadia et Olivier pour leurs conseils, leur expertise et leur soutien. Durant ces trois années, j'ai eu la chance de profiter de l'appui d'un trio exceptionnel. Merci pour la confiance que vous m'avez témoignée dès notre première réunion où nous avons discuté du sujet de cette thèse, ainsi que pour vos multiples recommandations et encouragements.

Je tiens à remercier tout particulièrement Cheng, pour nos travaux communs, et pour les centaines d'inégalités variationnelles qui ont noirci tableaux et papiers sous nos mains. Merci pour tes connaissances, tes idées et ta confiance.

Je remercie également Pascal Benchimol pour notre travail sur la désagrégation, ainsi qu'Hélène Le Cadre et Clémence Alasseur pour nos travaux sur les marchés pair-à-pair.

Je souhaite également remercier sincèrement Ghislain Hallez, chef du groupe R36 pour avoir soutenu ma candidature après notre premier entretien, ainsi que pour ses qualités d'écoute, sa bienveillance et son soutien dans mes différents choix.

Pendant ces trois ans, j'ai eu la chance d'être invité et accueilli par plusieurs personnes pour travailler sur différents projets : je remercie donc sincèrement Alejandro Jofré, Olivier Brun et Balakrishna Prabhu, et Guillaume Darmet ainsi que Scott Moura et Steven Low pour leurs invitations, leurs idées et conseils.

Je souhaite remercier l'ensemble des ingénieurs et chercheurs du département OSIRIS et en particulier du groupe R36 *Méthodes et outils d'optimisation* pour les différents échanges, pauses café et autres gâteaux. En particulier, merci à Riadh, ce fut un plaisir de partager ce bureau avec toi pendant ces trois années et de profiter de ton expérience en échange de quelques conseils et astuces LaTeX.

Je remercie l'ensemble des chercheurs et personnel administratif du CMAP et de l'Inria, en particulier je remercie vivement Hanadi Dib pour l'organisation de ma soutenance et des déplacements du jury.

Merci à mes partenaires de course à pied sur le plateau et ses jolis chemins : Olivier à nouveau, Hugo et Julien, et également aux membres et encadrants du club d'escalade de l'École polytechnique, qui m'ont permis de trouver une saine harmonie entre le travail de l'esprit et celui du corps.

Merci à l'ensemble des doctorants d'EDF avec qui j'ai passé de très bons moments : en particulier mes prédécesseurs Cécile, Rodolphe et Michaël, et mes successeurs Maxime, Adèle, Margaux, Benoît, Emma. Merci aussi aux doctorants du CMAP que j'ai pu croiser au Magnan ou dans les couloirs, en particulier merci à mes co-bureaux : Kevish, Benoît, Vianney, Felipe, Bastien et Clément, et à mes grands et petits frères académiques Jean-Bernard, Mateusz, Vianney, Nikolas, Marin, Omar, et Maxime à nouveau.

Merci à mes amis que j'ai pu voir régulièrement ces dernières années à Paris ou dans des régions plus sauvages : Thibault, Maud, Corentin, Nathan, Gabriel, Etienne, Antoine, Matthieu.

Maman, Papa, Lucien et Elisa, merci pour votre présence et votre soutien depuis toujours. Merci à ma famille. J'ai une pensée particulière pour mon grand-père Alain qui nous a quitté durant ma thèse, et à qui je suis reconnaissant de m'avoir témoigné confiance et estime.

Claire, merci pour tous ces magnifiques moments partagés, pour ton soutien et pour ton amour, pour toutes nos discussions passionnelles. L'épanouissement dans un travail de thèse émane des plaisirs de la curiosité et de la réflexion. Je te suis infiniment reconnaissant pour tout ce que tu m'as apporté depuis ces quelques années.

Contents

General Introduction

1 From a Centralized to a Decentralized Electric System

The major innovations and new regulations of the last decade have transformed electric systems across all continents.

Technological innovations and disruptive projects—the massive installation of smart meters in households [AY15] and associated communication technologies or the integration of electric vehicles in the grid [TRY16], along with smart charging and vehicle-to-grid technologies — have paved the way to the modern electric system.

Besides, new regulations—openings to competition, regulated markets, local renewable productions—emerging from institutional actors (governments, European energy commissions, competition authority) have painted a new, decentralized, electric landscape [AO16]. In France, the opening to competition of the production and distribution of electricity has started in 2000, after an historical monopoly of a unique centralized actor, EDF. The opening to competition for electricity supply to residential consumers followed later in 2014. Since then, the number of residential distribution companies in France has increased every year, to reach more than thirty in 2019. The number of decentralized, local renewable production installations has soared, as a consequence of the decrease in prices of photovoltaic panels and significant political incentives. In France in 2019, there are more than 36500 PV production sites [Rte].

Changes and innovations are fostered by ambitious objectives of environmental sustainability and reduction of green house gas emissions. For instance, with the Paris Agreement, the European Union has pledged to achieve greenhouse gas emission reductions of at least 40% by 2030 [Com19].

Aside from local renewable productions sources, other *distributed energy resources* (DERs) as *distributed storage* and load flexibility aggregators, new actors on the electric system, complete this decentralized picture.

Flexibility aggregators [GKS13] act as intermediaries between residential and commercial end-users and the operator of the electric system, possibly through a market. Their role is to *aggregate* a large number of negligible consumption flexibilities offered by end-users (originating from flexible usages such as charging of electric vehicles, smart washing machines and dishwashers, air conditioning, etc, see Figure 1) and valuate this lever on global load on the market or as a service offered to system operators (see Figure 2).

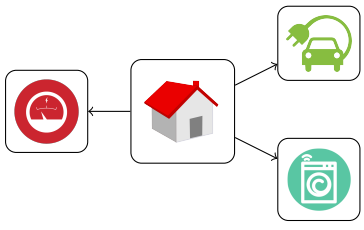


FIGURE 1: A flexible consumer is equipped with a *smart meter* and a consumption scheduler that can control one or several flexible appliances (electric vehicle, air conditioning, washing machine, etc.).

In the optimization of the electric system and the balance between demand and supply, these perspectives change the way we consider the load, from a fixed *parameter* in the former paradigm, to a *variable* in the new paradigm, which can be partially controlled, although in a decentralized way.

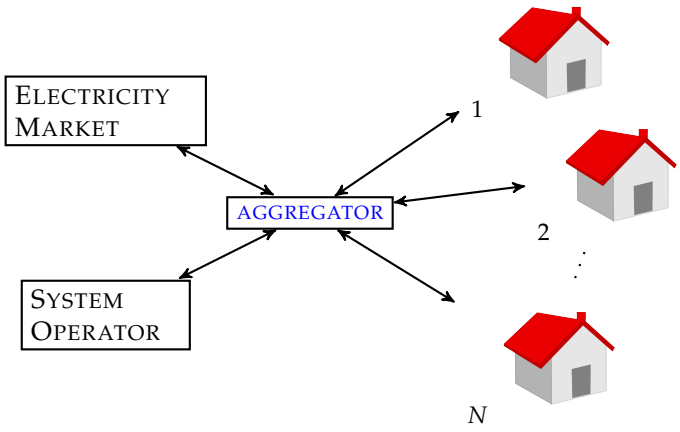


FIGURE 2: An aggregator is in charge of the consumption flexibilities of many individual electricity consumers and interacts with the system operator and electricity markets

In this context, *Demand Response* (DR) refers to the set of techniques used to exploit electricity consumers consumption flexibilities by giving them particular incentives, in order to achieve energy balancing or ancillary services to the grid as mentioned above [Sia14; Saa+12].

The exploitation of consumption flexibilities is considered a key lever to achieve the emissions reductions objectives, increase the share of intermittent renewable energy sources, limit infrastructure investments [All11] and ensure the network reliability and resilience [Sia14].

In this thesis, we address this problem of management of consumption flexibilities, by different methods based on optimization and game theory.

This topic has drawn the attention of a large community of researchers from electrical engineering, optimization and control, game theory and computer science: one can refer to [Den+15], [Sia14] and [VZV15] for surveys on the topic. Three fundamental aspects make this problem difficult. First, as we are modeling the flexibility of each individual electricity consumer, we are dealing with a very large number of variables. Second, the information concerning the constraints of consumers is not totally known to the central planning operator, but remains at the local consumer's level. It seems unrealistic to assume that a central operator would access to all the information, because of the quantity and variability of data involved and because of privacy issues. Indeed, consumers are not willing to provide critical and private information such as their time of presence at home. Last, decentralization entails the fact that a central operator can only have an indirect and partial control over the consumption profiles of end-consumers, for instance by the mean of incentives or signals, which may be received locally at the consumer's level by a smart meter linked to an autonomous consumption scheduling device.

The framework we consider, in particular in Part I and Part II, is the following: a central operator, for instance an isolated microgrid operator holding power production assets (generation plants, renewables, physical contracts, etc.) or an energy aggregator interacting with the electricity markets and other actors (see Figure 2), is in charge of the management of a pool of several consumers. Within the *smart grid* context, each consumer is equipped with an automatic *Energy Consumption Scheduler* (ECS), a device integrated inside a smart meter. This ECS is connected both to the central operator (via the power network or an ad-hoc communication network), from which it can receive and transmit signals, and to flexible electrical appliances (such as plug-in electric vehicles, air conditioning, heating, etc., see Figure 1), and is able to locally run algorithms to schedule the consumption profile of those appliances.

2 Mathematical framework: from Optimization to Games

Shifting from the centralized paradigm to the decentralized paradigm, as depicted in the previous section, requires to replace the standard optimization tools by game theory methods. We next review the main mathematical notions used in this context.

2.1 Distributed Optimization of a Centralized Problem

Let us adopt the point of view of a central operator, facing a cost function $f(\cdot)$ that depends on some variables $\boldsymbol{p} \in \mathcal{P}$, on which it has a direct control, but also depends on other variables $\boldsymbol{x} \stackrel{\text{def}}{=} (\boldsymbol{x}_n)_{n\in\mathcal{N}}$, where $\mathcal{N}$ denotes a set of actors different from the operator: for instance a set of flexible electricity consumers. Each individual variable $\boldsymbol{x}_n$, corresponding to agent n, is subject to some local constraints $\boldsymbol{x}_n \in \mathcal{X}_n$. If the individual variables $(\boldsymbol{x}_n)_n$ are controlled by the operator, we arrive at the centralized optimization problem:

$$\min_{\boldsymbol{p},\boldsymbol{x}} f(\boldsymbol{p},\boldsymbol{x}) \tag{1a}$$

$$\text{s.t. } \boldsymbol{p} \in \mathcal{P} \tag{1b}$$

$$\boldsymbol{x}_n \in \mathcal{X}_n,\ \forall n \in \mathcal{N}, \tag{1c}$$

$$\boldsymbol{h}(\boldsymbol{p},\boldsymbol{x}_1,\dots,\boldsymbol{x}_N) \leqslant 0, \tag{1d}$$

where the constraint (1d) couples the operator variable and the other actors variables. This coupling can be simple (for instance if $\boldsymbol{p}$ refers to the aggregate profile of a set of consumers, i.e. $\boldsymbol{p} = \sum_n \boldsymbol{x}_n$) but also very complex (for instance to integrate electricity network flow constraints).

A standard approach in optimization to solve problems of the form (1) taking advantage of the local character of the agents constraints one sees, is to rely on decomposition techniques [Coh78], [BT89], [BV04], in particular Lagrangian decomposition methods [PC06; XJB04], including dual subgradient methods [Ber99, Chapter 6] or ADMM [GM75]. Such approaches have also been investigated in the context of decentralized electric systems and Demand Response [Sag12], [Mol+17], [CBK17], [Shi+14], [LCL11], [Den+15]. The main drawback of these methods is that they usually rely on convexity hypotheses for the problem (1), which can be very restrictive in practice: for instance, the power production constraints [CA06] or the electricity network constraints are nonconvex. In Chapter 1, we provide an original method to solve particular instances of problem (1) in a decentralized manner, which does not rely on convexity of the function f or the set $\mathcal{P}$.

2.2 Game Models for Energy Management

The centralized problem (1) does not take into account the fact that the actors are *strategic*, in the sense that these actors will choose actions (the agent n has a control over her variable $\boldsymbol{x}_n$) and that their actions impact their individual objective (cost, bill, or comfort), which can be competing with the other agents objectives or with the objective function f of the operator. To model these strategic aspects in the context of DR, a large community of researchers has turned to the framework of *game theory*, e.g. [MR+10], [CK14], [Cha+14], [Saa+12], [Che+14], [Atz+13], [LCL11], [Bah+13].

A different but related framework one could use to address (1) while considering individual objective functions for agents, is *bilevel optimization* [CMS07]. However, to avoid the complexity of bilevel programming, we focus in this thesis on *single-level* models and problems, through the frameworks of distributed optimization and game theory.

A *game* [FT91] is a situation where a set of strategic agents, that we denote $\mathcal{N}$, called *players* (e.g. flexible consumers) can choose actions in *individual feasibility sets*, that we denote $(\mathcal{X}_n)_n$ (e.g. admissible electricity consumption profiles) and are in *interactions*.

This interaction is modeled through an individual *cost function* for each player $n \in \mathcal{N}$, $(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \mapsto f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \in \mathbb{R}$ that depends not only on her own action $\boldsymbol{x}_n$ but also on the actions of the others $\boldsymbol{x}_{-n} \stackrel{\text{def}}{=} (\boldsymbol{x}_m)_{m\neq n}$.

In the context of DR, we can model the interaction between energy consumers in this framework, with f_n representing an *electricity bill* minimized by agent n. The dependence of f_n on other agents consumption profiles $\boldsymbol{x}_{-n}$ results from the fact that the central operator is willing to avoid synchronization in the consumptions. Indeed, the load profile resulting from the consumers determines the sourcing cost of the energy: the higher the electric load is on a given time period, the more expensive it will be to produce and deliver electricity because of increasing marginal costs of production sources and of congestion effects.

In the game framework, we can consider a minimization problem for each agent n as:

$$\min_{\boldsymbol{x}_n} f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \tag{2a}$$

$$\text{s.t. } \boldsymbol{x}_n \in \mathcal{X}_n \, . \tag{2b}$$

In the framework of noncooperative games, a fundamental notion is brought by the concept of *Nash equilibrium* (NE) [Nas50]. This notion has emerged as the central solution concept in game theory, with extremely diverse applications [Nis+07, Sec.1.3.3]. A situation of Nash Equilibrium corresponds to action profiles $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_n)_{n\in\mathcal{N}}$ such that each player $n \in \mathcal{N}$, considering the actions of the other players $\hat{\boldsymbol{x}}_{-n}$ as fixed, has no interest to change her current action $\boldsymbol{x}_n$ (the player locally minimizes her cost function, and solves the problem (2)), that is:

$$\forall n \in \mathcal{N}, \forall \boldsymbol{x}_n \in \mathcal{X}_n, \;\; f_n(\hat{\boldsymbol{x}}_n, \hat{\boldsymbol{x}}_{-n}) \leqslant f_n(\boldsymbol{x}_n, \hat{\boldsymbol{x}}_{-n}) \, .$$

The Nash equilibrium captures the notion of a stable solution, from which no single player can individually decrease her cost function by deviating. When players are at such a solution, it is in each player's interest to stick to her current action.

2.3 Congestion Games

A class of games of particular importance in the context of DR are the so-called *congestion games*, introduced by [Ros73b]. In these games, each player chooses a subset of a set of shared resources $\mathcal{T}$. Each resource $t \in \mathcal{T}$ has a cost $c_t(.)$ given as an increasing function of the number of players selecting this resource.

When the action of each player n consists in selecting resources, that is, her action set is given as a subset $\mathcal{X}_n \subset \{0,1\}^{\mathcal{T}}$, we obtain a class of games called *unsplittable congestion games*. On the contrary, when each player n decides of a *load* on each resource $t \in \mathcal{T}$, that is, her action set is given as $\mathcal{X}_n \subset \mathbb{R}_+^{\mathcal{T}}$, we obtain a class of games called *splittable congestion games* [ORS93], [Wan12b, Sec. 1.3.3]. These games found their first applications in network routing problems [ORS93].

Splittable congestion games are particularly adapted in our context: if we consider that each agent chooses her *consumption profile* $\boldsymbol{x}_n = (x_{n,t})_{t\in\mathcal{T}} \in \mathcal{X}_n$ for a set of time periods $\mathcal{T} = \{1, \dots, T\}$, and that her cost function f_n corresponds to an energy bill with, for each period t, a *per-unit electricity price* $X_t \mapsto c_t(X_t)$ depending on the *aggregate* consumption $X_t = \sum_{n\in\mathcal{N}} x_{n,t}$ of users on time period t, we end up with congestion-like objective functions as:

$$f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) = \sum_{t\in\mathcal{T}} x_{n,t} c_t\Big(\sum_{m\in\mathcal{N}} x_{m,t}\Big). \tag{3}$$

Indeed, this framework corresponds to the splittable congestion games, or routing games, described in [ORS93], on a parallel network of T edges representing the time periods, and where each price function $c_t(.)$ corresponds to a *latency function* on edge t, as illustrated in Figure 3.

2.4 Efficiency of Equilibria

The evaluation of the *efficiency* of an equilibrium, or more generally of the outcome of a game, can follow several criteria depending on the situation: for instance, in our framework, an electricity operator could be interested in evaluating the total production costs or the total carbon emissions associated with the consumption outcome. However, a standard criterion

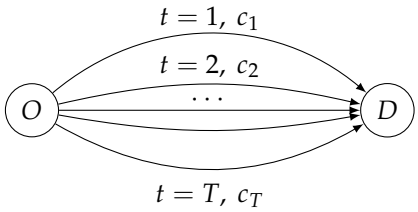


FIGURE 3: A congestion game for the energy consumption on T time periods can be interpreted as a routing congestion game on a parallel-arc network with T edges.

that has been widely adopted in the congestion game literature is what we refer to as the *social cost*, that is, the sum of the cost functions of all players, $\mathrm{SC}(\boldsymbol{x}) \overset{\text{def}}{=} \sum_{n\in\mathcal{N}} f_n(\boldsymbol{x})$.

Considering this criterion, a quantitative measure of the efficiency of equilibria in a game is the so-called *Price of Anarchy* (PoA) [KP99], given as the worst (maximal) social cost at a Nash equilibrium divided by the optimal (minimal) social cost that can be achieved (corresponding to a situation where a central operator could optimize and choose the actions for all players in order to minimize the social cost), that is:

$$\mathrm{PoA} = \frac{\max_{\hat{\boldsymbol{x}}\in\boldsymbol{\mathcal{X}}^{\mathrm{NE}}} \mathrm{SC}(\hat{\boldsymbol{x}})}{\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \mathrm{SC}(\boldsymbol{x})} \ ,$$

with $\boldsymbol{\mathcal{X}}^{\mathrm{NE}}$ denoting the set of possible NEs. As an equilibrium situation, corresponding to locally stable but selfish choices of actions for players, has no reason *a priori* to correspond to a minimizer of the social cost, the PoA is usually greater than one.

Different works in the congestion game literature focus on bounding the PoA in particular frameworks: for instance, [CBCS16] and [CB+17] have recently shown that in network congestion games, the PoA converges to one in the limit of infinitely large demands. The PoA has also been adopted as a reference in the smart grid literature to measure the efficiency of a game theoretic system, e.g. [CBK17], [Zhu+12], [CK14], [Saa+12], [NAC14], [HRD14].

As stated above, game theory offers a natural framework for distributed and decentralized optimization: for a central operator in charge of several agents (e.g. flexible electricity consumers), the problem is originally stated as an optimization problem (e.g. minimizing the system cost). Then, the question is: how to design a game (through incentives or prices) whose equilibria correspond to minimizers of the system cost ? In this context, game theory is used as a tool to operate distributed optimization [MRP13], [MS15]. For these mechanism design considerations, an important class of games are the so-called *potential games* [MS96]: a game is an *exact potential game* if there exist a *potential* function $\Phi : (\mathcal{X}_n)_{n\in\mathcal{N}} \to \mathbb{R}$ that all players have interest to minimize, that is:

$$\forall n \in \mathcal{N},\ \forall \boldsymbol{x} \in (\mathcal{X}_m)_{m\in\mathcal{N}},\ \forall \boldsymbol{x}'_n \in \mathcal{X}_n,\ f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) - f_n(\boldsymbol{x}'_n, \boldsymbol{x}_{-n}) = \Phi(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) - \Phi(\boldsymbol{x}'_n, \boldsymbol{x}_{-n}) \ .$$

In particular, if a central operator designs a game which has a potential [LM13], NEs of this game will correspond to local minima of the function Φ and, if Φ is convex, will be easy to compute. This approach has been followed in the context of DR, see e.g. [Tus+18], [Wu+11], [BW15], [YH16].

3 Contents of the manuscript

This dissertation is organized in four main parts. The works and models presented follow an order of increasing decentralization and degree of autonomy for the end-consumers.

• Part I addresses the problem of the management of consumption flexibilities through a distributed optimization approach. The point of view adopted in this part is the one of

a central operator in charge of decentralized resources (flexibilities) of many agents (consumers), and willing to optimize the distributed resources of agents while respecting their privacy. In Chapter 1, we start by defining the general framework of a distributed *resource allocation problem*—resources being, in our specific case, the energy on each time period—taking the point of view of the central operator in charge of multiple agents, each agent having individual and confidential constraints. We provide a method to compute an optimal solution of this problem while ensuring the privacy of the agents, and prove the validity and the computational efficiency of this method.

• Part II considers the same problem of the management of consumption flexibilities (Demand Response), but through a game theoretic approach. In Chapter 2 and Chapter 3, we use game theory as a tool for decentralized optimization: a *billing mechanism*, defined by an operator, corresponds to price signals sent to consumers and incentives for them to reach an equilibrium in the corresponding game. We compare two billing mechanisms in Chapter 2, and then, in Chapter 3, we focus on an hourly billing mechanism—corresponding to an atomic splittable congestion game—and its computational aspects. In Chapter 4, we consider the possibility for consumers to have local, divergent objectives, formulated as preferred consumption profiles, in addition to their energy bill set by the operator.

• Part III gives some theoretical results on the approximation of equilibria in large congestion games—where *large* refers to the cardinality of the players set. This can be applied in the case where players represent individual electricity consumers at a regional or national scale, where one obtains a game with hundreds of thousands of players. In Chapter 5, we consider a congestion game with a large number of players and coupling constraints, and we show that a Nash equilibrium can be approximated by the Wardrop equilibrium of an approximating game where similar players are clustered in populations. In Chapter 6, we consider the framework of *generalized nonatomic aggregative games* with an infinity of different types of players: this situation models, for instance, the interactions of a large population of consumers, described by a parametric distribution (their energy need, their time constraints). We show that, in this game, a *variational Wardrop equilibrium*, a notion we introduce, can be approximated by a Wardrop equilibrium of a population game of smaller dimension.

• Part IV goes further in terms of decentralization of the electric system, by considering a situation without any central operator and where consumers can exchange energy directly in peer to peer transactions. In Chapter 7, we propose an original model of a generalized game where each player is a flexible consumer owning a local renewable production source, and having preferences over the energy transactions she can engage to, in addition to her local objective. We study the generalized equilibria of this game and, in particular, we show that a variational equilibrium realizes the social optimum of the game.

4 Contributions of the Chapters

This work brings contributions of different types: theoretical, algorithmic, empirical and on modeling aspects. Indeed, this thesis introduces and analyzes different tools and models for the energy framework described in Section 1, but also brings several theoretical results in the fields of decentralized optimization and game theory, in particular in the framework of splittable congestion games. The main contributions are given below, in order of appearance.

Chapter 1. We propose an original algorithm (Algorithm 1.4) that computes an optimal aggregate resource allocation, solution of a nonconvex optimization problem and corresponding to the *aggregation* of feasible individual profiles for a set of agents having individual constraints. The algorithm is distributed and privacy-preserving, in the sense that the computation is done without revealing the individual constraints of each agent to a third party, either another agent or a central operator. In practice, this algorithm can be used for instance to optimize an electricity load profile aggregated from a pool of flexible consumers, while ensuring that the confidential information of each consumer remains private. The application of the method in this context has led to a patent application by EDF [Jac+18b].

As our algorithm relies on the method of alternate projections [VN50], [GPR67] on transportation polytopes, a key theoretical result for the algorithm is proven in Chapter 1: in Theorem 1.3 we show that, when the projections do not converge to a single point, we can build a polyhedral Hoffman cut [Hof60] in the network flow problem associated to a transportation polytope.

Last, in Theorem 1.4, relying on spectral graph theory arguments [CG97], we obtain an explicit bound on the speed of convergence of alternate projections on transportation polytopes. This result shows that the time complexity of the method proposed in this chapter evolves only linearly with the number of agents and can therefore be used at a large scale.

Chapter 2. We consider two DR mechanisms based on a potential game (referred to as *Daily Proportional* mechanism, introduced in [MR+10]) and based on an atomic splittable congestion game (referred to as *Hourly Proportional* mechanism, introduced in [Bah+13], see Section 2), and we study the Nash equilibria of these games through their existence and uniqueness. In particular, Theorem 2.1 gives a sufficient condition on the derivatives of the price functions (latency functions in a congestion game) for the uniqueness of an equilibrium, obtained from the standard strict monotonicity condition [Ros65] and using a result from perturbation theory in linear algebra. In the same chapter, we provide in Theorem 2.2 a theoretical bound on the Price of Anarchy (see the definition above) for this class of games, obtained from the (λ, μ)*-local smoothness* technique introduced in [RS15].

Chapter 3. We focus on the algorithmic aspects in atomic splittable congestion games. First, Theorem 3.2 extends the result of [ORS93] to prove the uniqueness of NE in a more general context. We give several results on the convergence of the *best response* (BR) algorithm [GM91], [DG16], [DGG19b] in particular frameworks of splittable congestion games. Theorem 3.4 proves the geometric convergence of the cyclic version of BR in a potential case, while Corollary 3.2 proves the geometric convergence of the randomized version of BR in specific cases. This proves the convergence of BR in a different context than the one considered in [Mer08]. We also study a projected gradient method, and show in Theorem 3.5 that, under strong monotonicity assumptions, this method converges geometrically to the unique equilibrium of the game. Using a result from perturbation theory in linear algebra, Proposition 3.1 gives a sufficient condition on the price functions $c_t(.)$ for strong monotonicity to hold. Last, we propose in this chapter an online DR procedure with receding horizons (Algorithm 3.4), in the spirit of Model Predictive Control [Wu+11], to take into account updated forecasts in a stochastic environment. Theorem 3.6 proves that the consumption profiles computed by this procedure correspond to the desired NE in the limit of perfect forecasts.

Chapter 4. We extend the energy consumers game model introduced in Chapter 2, by considering individual temporal preferences in the objective functions of electricity consumers. In the simplified framework of two time periods (*resources*) we give theoretical results on the impact of preferences at the equilibria in the two mechanisms introduced in Chapter 2. We compare the PoA in those games and the *price of efficiency*, a similar concept we introduce to measure the efficiency of the mechanism from the *system operator* side. In particular, Theorem 4.4 shows that the system costs are always smaller in the hourly proportional mechanism. This chapter also presents numerical results on a realistic test case, using consumption data from the *PecanStreet* project database [Pec].

Chapter 5. We consider atomic splittable congestion games with coupling constraints [Har91] and with a very large number of players, heterogeneous by their individual constraints or objective functions. Using the framework of variational inequalities [FP07], we show in Theorem 5.1 that Variational Nash Equilibria (VNE) are well approximated by Variational Wardrop Equilibria (VWE). We show in Theorem 5.2 that, regrouping similar players into homogeneous populations, we can define an approximating game whose equilibrium is close to an equilibrium of the initial game. Those results improve the work of [Pac+18], where the authors show that VNEs are close to VWEs in large aggregative games: Chapter 5 shows in addition that we can consider a small number of populations in the population game and that the VWE remains close to a VNE of the initial game. Those approximation results are very relevant for computational purposes, as the problems characterizing the equilibria in the approximations are of much smaller dimension than the initial problems.

Chapter 6. We define a new family of games, namely, nonatomic aggregative games with *infinitely many types* of players (Definition 6.1), where a player's type is defined by her feasible set of actions and her objective function, extending the framework considered in [MZ97]. These games emerge naturally when considering a population described by a parametric distribution. In the presence of coupling constraints, we define the notion of VWE for a game with infinitely many types (Definition 6.4), characterized by an infinite dimensional variational inequality. We prove in Theorem 6.5 an approximation result stating that an initial VWE can be approximated by a Wardrop equilibrium of a game with a finite number of homogeneous populations. This latter equilibrium is symmetric in the sense that all players within each population adopt the same equilibrium action. This result can also be used for computational purposes, as it establishes a low-dimension characterization of the equilibrium of the initial game.

Chapter 7. We present an original model of peer-to-peer exchanges within a community of electricity consumers with preferences. The model is defined as a *generalized game* [HP90], as the trade reciprocity conditions define coupling constraints between a consumer and her neighbors. This game model captures the strategic behavior of consumers, as opposed to the usual decomposition and distributed optimization approaches addressing this problem e.g. [SBP18]. A particular result we obtain is Proposition 7.6 which shows that the *variational Nash equilibria* of this game are socially optimal. We study numerically the efficiency of equilibria, through the PoA, by computing the generalized equilibria of the game on test cases and comparing the energy flows at equilibrium with the flows obtained in the centralized solution, corresponding to the social optimum.

Each of the seven chapters presented in this thesis is based on a conference proceeding reviewed and accepted, or on a journal article that has been either published or has been submitted for publication:

- Chapter 1 is based on the paper [Jac+19b] submitted for publication, and on the conference paper [Jac+19a] accepted in the *2019 IEEE 58th Conference on Decision and Control*;
- Chapter 2 is based on the conference paper [Jac+17b] presented at the *2017 IEEE Conference on Innovative Smart Grid Technologies (ISGT)*;
- Chapter 3 is based on the paper [Jac+19c], published in *IEEE Transactions on Smart Grid*;
- Chapter 4 is based on the conference paper [Jac+17a] presented at the *2017 IEEE Conference on Smart Grid Communications*;
- Chapter 5 is based on the paper [Jac+18a] submitted for publication;
- Chapter 6 is based on the preprint [JW18a], on the paper [JW19] submitted for publication, and on the conference paper [JW18b], presented at the *2018 IEEE 57th Conference on Decision and Control*;
- Chapter 7 is based on the journal paper [LC+19b], accepted for publication in *European Journal of Operational Research*.

Notation and Conventions

Below we define the mathematical notation and conventions used throughout the thesis. The papers presented in this thesis have been modified in order to have a coherent notation along the chapters.

- bold letters such as $\boldsymbol{x}$ are used to denote vectors, while normal letters x denote scalars;
- when $\boldsymbol{x} = (x_n)_{n\in\mathcal{N}}$ is a vector indexed by a set $\mathcal{N}$ of players or agents, the associated uppercase letter $\mathcal{X}$ denotes the *aggregate* vector $X = \sum_{n\in\mathcal{N}} x_n$;
- calligraphic letters such as $\mathcal{X}, \mathcal{T}, \mathcal{H}, \mathcal{A}$ are used to denote sets, except for $\mathcal{L}$ that is usually employed for a Lagrangian function;
- for a subset $\mathcal{T}_0$ of a set $\mathcal{T}$, the set $\mathcal{T}_0^c \stackrel{\text{def}}{=} \{t \in \mathcal{T} \setminus \mathcal{T}_0\}$ denotes the complementary set of $\mathcal{T}_0$;
- $\mathcal{U}([a,b])$ stands for the uniform distribution on $[a,b]$;
- $\mathrm{P}_{\mathcal{C}}$ denotes the (Euclidean) projection on a convex set $\mathcal{C}$;
- $\boldsymbol{x}^\top$ (resp. $\boldsymbol{A}^\top$) denotes the transpose of a vector $\boldsymbol{x}$ (resp. of a matrix $\boldsymbol{A}$);
- For $d \in \mathbb{N}$, $\mathbb{1}_d$ denotes the vector of ones $(1 \dots 1)^\top \in \mathbb{R}^d$;
- for a vector $(\boldsymbol{x}_n)_{n\in\mathcal{N}}$ and an index $n \in N$, the notation $\boldsymbol{x}_{-n}$ denotes $(\boldsymbol{x}_m)_{m\in\mathcal{N},m\neq n}$;
- we write $a \stackrel{\text{def}}{=} b$ when a is *equal to b by definition*.

Part I

Decentralized Management of Flexibilities and Optimization

In the management of energy consumption, the *aggregate* load (that is, the total load to deliver to a set of consumers) is a key metric: a central operator interacts with flexible consumers to optimize the aggregate production or the trades on the wholesale electricity market.

However, demand response management also involves a local, disaggregated scale at the consumers level: each flexible consumer has her own local and individual constraints (for instance, in the case of an electric vehicle, the energy need and the time periods available for charging), and the aggregate load—the main concern of the central operator—has to correspond to the aggregation of *feasible* individual profiles for the flexible consumers.

In this Part I, we consider a generalization of the DR problem described above, namely, the *distributed resource allocation problem*, involving the minimization of a cost as a function of an aggregate profile (the *central operator* cost function)and a constraint of *disaggregation* to individual profiles, each of these profiles being subject to individual constraints (see problem 1.1 in Chapter 1).

Owing to the high dimension of the problem and privacy issues concerning individual consumers constraints, the need for a decentralized optimization approach is a global consensus in the DR literature [VZV15].

A standard approach to solve resource allocation problems in a distributed way is to rely on Lagrangian decomposition methods (e.g. [XJB04]). These kind of approaches have also been considered in the context of DR [PC06; CBK17; Shi+14; LCL11; Den+15]: in that case, the central operator computes a vector of Lagrange multipliers (e.g. from the supply-demand balancing constraint) and sends it to consumers as energy prices for each time period.

However, there are two main drawbacks in the use of Lagrangian decomposition approaches. First, they require some convexity hypothesis of the problem considered, which do not hold in general in the energy context as, for instance, the problem of the optimization of production assets is nonconvex (see Section 1.5.2). Second, even in the convex case, the primal solution of the decomposition—which in our case would correspond to the consumption profiles—cannot be easily recovered (due to the absence of uniqueness of the responses to Lagrangian prices, see e.g. [CZ84]).

In the following Chapter 1, we consider a different decentralized optimization approach, and we provide an original privacy-preserving algorithm that does compute the optimal allocation of resources, avoiding each individual agent to reveal her private information (constraints and individual solution profile) neither to the central operator nor to a third party, and avoids the two drawbacks of Lagrangian methods mentioned above.

Our method relies on an aggregation procedure: we compute iteratively a global allocation of resources, and gradually ensure existence of a disaggregation, that is individual profiles satisfying agents' private constraints, by a protocol involving the generation of polyhedral cuts and Secure Multiparty Computation (SMC).

To obtain these cuts, we use an alternate projection method, which is implemented locally by each agent, preserving her privacy needs. We address especially the case in which the local and global constraints define a transportation polytope. Then, we provide theoretical convergence estimates together with numerical results, showing that the algorithm can be effectively used to solve the allocation problem in high dimension, while addressing privacy issues. We illustrate the procedure through the example of a microgrid operator minimizing production costs, resulting in a nonconvex optimization problem.

Chapter 1

A Privacy-preserving Disaggregation Method to Optimize Distributed Resource Allocation

This chapter is based on the paper [Jac+19b], submitted for publication, and on the conference paper [Jac+19a]. Appendix 1.C was added in the manuscript and has not been submitted for publication.

1.1 Introduction

1.1.1 Motivation

Consider an operator of an electricity microgrid optimizing the joint production schedules of renewable and thermal power plants in order to satisfy, at each time period, the consumption constraints of its consumers. To optimize power generation or market costs and the integration of renewable energies, this operator relies on demand response techniques, that is, taking advantage of the flexibilities of some of the consumers electric appliances—those which can be controlled without impacting the consumer's comfort, as electric vehicles or water heaters [Jac+19c]. However, for privacy reasons, consumers are not willing to provide neither their consumption constraints nor their consumption profiles to a central operator or any third party, as this information could be used to infer private information such as their presence at home.

The *global problem* of the operator is to find an allocation of power (aggregate consumption) $\boldsymbol{p} = (p_t)_t$ at each time period (*resource*) $t \in \mathcal{T}$, such that $\boldsymbol{p} \in \mathcal{P}$ (feasibility constraints of power allocation, induced by the power plants constraints). Besides, this aggregate allocation has to match an individual consumption profile $\boldsymbol{x}_n = (x_{n,t})_{t\in\mathcal{T}}$ for each of the consumer (agent) $n \in \mathcal{N}$ considered. The problem can be written as follows:

$$\min_{\boldsymbol{x}\in\mathbb{R}^{N\times T},\, \boldsymbol{p}\in\mathcal{P}} f(\boldsymbol{p}) \tag{1.1a}$$

$$\boldsymbol{x}_n \in \mathcal{X}_n,\ \forall n \in \mathcal{N} \tag{1.1b}$$

$$\sum_{n\in\mathcal{N}} x_{n,t} = p_t,\ \forall t \in \mathcal{T}\,. \tag{1.1c}$$

The (aggregate) allocation $\boldsymbol{p}$ can be made *public*, that is, revealed to all agents. However, the individual constraint set $\mathcal{X}_n$ and individual profiles $\boldsymbol{x}_n$ constitute *private* information of agent n, and should not be revealed to the operator or any third party.

It will be helpful to think of problem (1.1) as the combination of two interdependent subproblems:

i) given an aggregate allocation $\boldsymbol{p}$, the *disaggregation problem* consists in finding, for each agent n, an individual profile $\boldsymbol{x}_n$ satisfying her individual constraint (1.1b), so that constraint (1.1c) is satisfied, or equivalently, solve the problem:

$$\text{FIND } \boldsymbol{x} \in \mathcal{Y}_{\boldsymbol{p}} \cap \mathcal{X} \tag{1.2a}$$

$$\text{where } \mathcal{Y}_{\boldsymbol{p}} \stackrel{\text{def}}{=} \{\boldsymbol{y} \in \mathbb{R}^{NT} | \boldsymbol{y}^\top \mathbb{1}_N = \boldsymbol{p}\} \text{ and } \mathcal{X} \stackrel{\text{def}}{=} \prod_{n \in \mathcal{N}} \mathcal{X}_n \,. \tag{1.2b}$$

When (1.2) has a solution, we say that a *disaggregation* exists for $\boldsymbol{p}$;

ii) For a given subset $\mathcal{Q} \subset \mathcal{P}$, we define the *master problem*,

$$\min_{\boldsymbol{p} \in \mathcal{Q}} f(\boldsymbol{p}) \,. \tag{1.3}$$

When $\mathcal{Q}$ is precisely the set of aggregate allocations for which a disaggregation exists, the optimal solutions of the master problem correspond to the optimal solutions of (1.1).

Aside from the example above, *resource allocation problems* (optimizing common resources shared by multiple agents) with the same structure as (1.1), find many applications in energy [Mül+17; Jac+19c], logistics [LLC95], distributed computing [Ma+82], health care [RV11] and telecommunications [ZHS10]. In these applications, several entities or agents (e.g. consumers, stores, tasks) share a common resource (energy, products, CPU time, broadband) which has a global cost for the system. For large systems composed of multiple agents, the dimension of the overall problem can be prohibitive: a solution is to rely on decomposition and distributed approaches [BT89; PC06; XB06]. Besides, agents' individual constraints are often subject to privacy issues [HAF05]. These considerations have paved the way to the development of privacy-preserving, or non-intrusive methods and algorithms, e.g. [Zoh+12; JPW06].

In this work, except in Section 1.4, we consider that each agent $n \in \mathcal{N}$ has a global demand constraint (e.g. energy demand or product quantity), which confers to the disaggregation problem the particular structure of a transportation polytope [Bol72]: the sum over the agents is fixed by the aggregate solution $\boldsymbol{p}$, while the sum over the T resources are fixed by the agent global demand constraint. Besides, individual constraints can also include minimal and maximal levels on each resource, as for instance electricity consumers require, through their appliances, a minimal and maximal power at each time period.

1.1.2 Main Results

The main contribution of the chapter is to provide a non-intrusive and distributed algorithm (Algorithm 1.4) that computes an aggregate resource allocation $\boldsymbol{p}$, optimal solution of the—possibly nonconvex—optimization problem (1.1), along with feasible individual profiles $\boldsymbol{x}$ for agents, without revealing the individual constraints of each agent to a third party, either another agent or a central operator. The algorithm solves iteratively instances of *master problems* $\min_{\boldsymbol{p} \in \mathcal{P}^{(s)}} f(\boldsymbol{p})$ by constructing successive approximations $\mathcal{P}^{(s)} \subset \mathcal{P}$ of the aggregate feasible set of (1.1) for which a disaggregation exists, by adding to the set $\mathcal{P}^{(s)}$ a new constraint on $\boldsymbol{p}$ (i.e. a cutting plane), before solving the next master problem. We shall see that this cutting plane can be computed and added to the master problem without revealing any individual information on the agents.

More precisely, to identify whether or not disaggregation (1.2) is feasible and to add a new constraint in the latter case, our algorithm relies on the alternate projections method (APM) [VN50; GPR67] for finding a point in the intersection of convex sets. Here, we consider the two following sets: on the one hand, the affine space of profiles $\boldsymbol{x} \in \mathbb{R}^{NT}$ aggregating to a given resource allocation $\boldsymbol{p}$, and on the other hand, the set defined by all agents individual constraints (demands and bounds). As the latter is defined as a Cartesian product of each agent's feasibility set, APM can operate in a distributed fashion. The sequence constructed by the APM converges to a single point if the intersection of the convex sets is nonempty, and it converges to a periodic orbit of length 2 otherwise. If the APM converges to a periodic orbit, meaning that the disaggregation is not feasible, we construct from this

orbit a polyhedral *cut*, i.e. a linear inequality satisfied by all feasible solutions $\boldsymbol{p}$ of the global problem (1.1), but violated from the current resource allocation (Theorem 1.3). Adding this cut to the *master problem* (1.3) by updating $\mathcal{Q}$ to a specific subset, we can recompute a new resource allocation and repeat this procedure until disaggregation is possible. At this stage, the use of a cryptographic protocol, secure multiparty computation, allows us to preserve the privacy of agents. Another main result stated in this chapter is the explicit upper bound on the convergence speed of APM in our framework (Theorem 1.2), which is obtained by spectral graph theory methods, exploiting also geometric properties of transportation polytopes. This explicit speed shows a linear impact of the number of agents, which is a strong argument for the applicability of the method in large distributed systems.

1.1.3 Related Work

A standard approach (e.g. [PC06; XJB04; SMC06]) to solve resource allocation problems in a distributed way is to rely on a Lagrangian based decomposition technique: for instance dual subgradient methods [Ber99, Ch. 6] or ADMM [GM75]. Such techniques are generally used to decompose a large problem into several subproblems of small dimension. However, those methods often require global convexity hypothesis, which are not satisfied in many practical problems (e.g. MILP). We refer the reader to [Ber99, Chapter 6] for more background. On the contrary, our method can be used when the allocation problem (1.1) is not convex.

As developed in Section 1.4, the method proposed here can be related to Bender's decomposition [Ben62]. The difference with Bender's approach is in the way of generating a new cut to add in the master problem: instead of solving linear programs, we use APM and our theoretical results, which provides a decentralized, privacy-preserving and scalable procedure. In contrast, at each stage, Benders' algorithm requires to solve a linear program requiring the knowledge of the private constraints of each individual agent (see Section 1.4.1 for more details).

The problem of the aggregation of constraints has been studied in the field of energy, in the framework of smart grids [Mül+17; ALT18]. In [Mül+17], the authors study the management of energy flexibilities and propose to approximate individual constraints by zonotopic sets to obtain an aggregate feasible set. A centralized aggregated problem is solved via a subgradient method, and a disaggregation procedure of a solution computes individual profiles. In [ALT18], the authors propose to solve the economic power dispatch of a microgrid, subject to several agents private constraints, by using a Dantzig-Wolfe decomposition method.

The APM has been the subject of several works in itself [GPR67; BB93; BCW15]. The authors of [BLY14] provide general results on the convergence rate of APM for semi-algebraic sets. They show that the convergence is geometric for polyhedra. However, it is generally hard to compute explicitly the geometric convergence rate of APM, as this requires to bound the singular values of certain matrices arising from the polyhedral constraints. A remarkable example where an explicit convergence rate for APM has been established is in [NJJ14], where the authors consider a different class of polyhedra arising in submodular optimization. A common point with our results is the use of spectral graph theory arguments to estimate singular values.

1.1.4 Structure

Section 1.2 describes the class of resource allocation problems we address in this chapter, and formulate the idea of the decomposition with the *disaggregation* subproblems. In Section 1.3, we focus on APM, the subroutine used to solve the disaggregation subproblems. After stating results on the convergence of APM, In Section 1.3.1, we show the key result on which relies the proposed decomposition: how to generate a new cut to add in the master problem, from the output of APM. In Section 1.3.2, we show how to improve the privacy of the proposed procedure by using secure multiparty computation techniques. In Section 1.3.3 , we prove an explicit upper bound on the rate of convergence of APM in our case. In Section 1.4, we generalize part of our results and propose a modified algorithm in the case where agents constraints are polyhedral. Finally, in Section 1.5, we propose numerical examples of

the method: Section 1.5.1 gives an illustrative toy example in dimension $T = 4$, while in Section 1.5.2, we consider a larger scale, nonconvex example, coming from the microgrid application exposed at the beginning of the introduction.

1.2 Resource Allocation and Transportation Structure

1.2.1 A Decomposition based on Disaggregation

As stated in the introduction, we consider a centralized entity (e.g. an energy operator) interested in minimizing a possibly nonconvex cost function $\boldsymbol{p} \mapsto f(\boldsymbol{p})$, where $\boldsymbol{p} \in \mathbb{R}^T$ is the aggregate allocation of T dimensional resources (for example power production over T time periods). This resource allocation $\boldsymbol{p}$ is to be shared between a set $\mathcal{N}$ of N individual agents, each agent obtaining a part $\boldsymbol{x}_n \in \mathcal{X}_n$, where $\mathcal{X}_n$ denotes the individual feasibility set of agent n.

The global problem the operator wants to solve is described in (1.1). The idea behind the results of this chapter is that, in problem (1.1), the constraints set $\mathcal{X}_n$ and individual profile $\boldsymbol{x}_n$ are confidential to agent n and should not be disclosed to the central operator or to another agent.

Let us define the set $\mathcal{P}_{\mathrm{D}}$ of feasible aggregate allocations that are disaggregeable as:

$$\mathcal{P}_{\mathrm{D}} \stackrel{\text{def}}{=} \left\{ \boldsymbol{p} \in \mathcal{P} \mid \exists \boldsymbol{x} \in \mathcal{X} \; ; \; \boldsymbol{p} = \textstyle\sum_n \boldsymbol{x}_n \right\} . \tag{1.4}$$

Feasibility of problem (1.1) is equivalent to having $\mathcal{P}_{\mathrm{D}}$ not empty.

Constraints for each agent are composed of a global demand over the resources and lower and upper bounds over each resource, as given below:

Assumption 1.1. *For each $n \in \mathcal{N}$, there exists $E_n > 0$, $\overline{x}_n \in \mathbb{R}^T$, $\underline{x}_n \in \mathbb{R}^T$ such that :*

$$\mathcal{X}_n = \{\boldsymbol{x}_n \in \mathbb{R}^T : \textstyle\sum_{t \in \mathcal{T}} x_{n,t} = E_n \text{ and } \underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t}\} \neq \emptyset. \tag{1.5}$$

In particular, $\mathcal{X}_n$ is convex and compact. Given an allocation $\boldsymbol{p}$, the structure obtained on the matrix $(x_{n,t})_{n,t}$, where sums of coefficients along columns and along rows are fixed, is often referred to as the *transportation problem*. These problems found various applications (see e.g. [AN79; Mun57]). We focus on this case in Sections 1.2 and 1.3, while in Section 1.4, we shall give a generalization of some of our results in the general case where $\mathcal{X}_n$ is a polyhedron.

Given a particular allocation $\boldsymbol{p} \in \mathcal{P}$, the operator will be interested to know if this allocation is *disaggregeable*, that is, if there exists individual profiles $(\boldsymbol{x}_n)_{n \in \mathcal{N}} \in \prod_n \mathcal{X}_n$ summing to $\boldsymbol{p}$, or equivalently if the *disaggregation problem* (1.2) has a solution.

Following (1.2), the *disaggregate* profile refers to $\boldsymbol{x}$, while the *aggregate* profile refers to the allocation $\boldsymbol{p}$. Problem (1.2) may not always be feasible. Some necessary conditions for a disaggregation to exist, obtained by summing the individual constraints on $\mathcal{N}$, are the following *aggregate* constraints:

$$\boldsymbol{p}^\top \mathbb{1}_T = \boldsymbol{E}^\top \mathbb{1}_N \tag{1.6a}$$

$$\text{and } \underline{\boldsymbol{x}}^\top \mathbb{1}_N \leqslant \boldsymbol{p} \leqslant \overline{\boldsymbol{x}}^\top \mathbb{1}_N . \tag{1.6b}$$

Those conditions are not sufficient in general, as explained in the following section.

1.2.2 An equivalent flow problem and Hoffman conditions

The particular structure of the problem we consider implies that we can write it as a flow problem in a graph, as stated in Proposition 1.1. We refer the reader to the book [Coo+09, Chapter 3] for the terminology.

Definition 1.1. Consider a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertices $\mathcal{V}$ and edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, and demands $d : \mathcal{V} \to \mathbb{R}$ (where $d_v < 0$ means that v is a production node), edge lower capacities $\ell : \mathcal{E} \to \mathbb{R}_+$ and upper capacities $u : \mathcal{E} \to \mathbb{R}_+$. A flow on $\mathcal{G}$ is a function

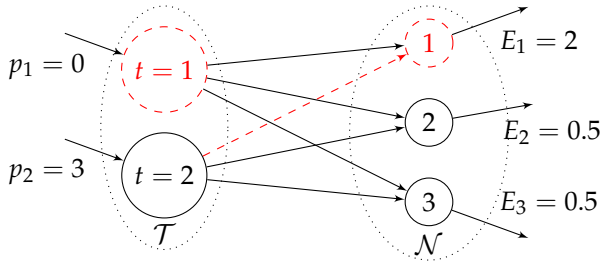


FIGURE 1.1: Example of a flow representation of the disaggregation problem ($T = 2$ and $N = 3, \underline{x} = 0, \overline{x} = 1$). *Here, the aggregate constraints* (1.6) *are verified, but condition* (1.7) *written with* $\mathcal{A} = \{t_1, n_1\}$ *(dashed nodes) does not hold.*

$x : \mathcal{E} \to \mathbb{R}_+$ such that x satisfies the capacity constraints, that is $\forall e \in \mathcal{E}, \ell_e \leqslant x_e \leqslant u_e$, and Kirchoff's law, that is, $\forall v \in \mathcal{V}, \sum_{e \in \delta_v^+} x_e = d_v + \sum_{e \in \delta_v^-} x_e$, where δ_v^+ (resp. δ_v^-) is the set of edges ending at (resp. departing from) vertex v.

The following proposition is immediate:

Proposition 1.1. *Consider the bipartite graph* $\mathcal{G}$ *with vertices* $\mathcal{V} = \mathcal{T} \cup \mathcal{N}$ *and with edges* $\mathcal{E} = \{(t,n)\}_{t \in \mathcal{T}, n \in \mathcal{N}}$. *Define demands on nodes* $\mathcal{T}$ *by* $d_t = -p_t$ *and demands on nodes* $\mathcal{N}$ *by* $d_n = E_n$. *Assign to each edge* (t,n) *an upper capacity* $u_{n,t} = \overline{x}_{n,t}$ *and lower capacity* $\ell_{n,t} = \underline{x}_{n,t}$. *Then, finding a solution* x *to* (1.2) *is equivalent to finding a feasible flow in* $\mathcal{G}$.

Hoffman [Hof60] gave a necessary and sufficient condition for the flow problem to be feasible. This generalizes a result of Gale (1957). The stated condition is intuitive: there cannot be a subset of nodes whose demand exceeds its "import capacity".

Theorem 1.1 ([Hof60])**.** *Given a digraph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *with demand* $d \in \mathbb{R}^{\mathcal{V}}$ *such that* $d(\mathcal{V}) = 0$ *and capacities* $\ell \in (\mathbb{R} \cup \{-\infty\})^{\mathcal{E}}$ *and* $u \in (\mathbb{R} \cup \infty)^{\mathcal{E}}$ *with* $\ell \leqslant u$, *there exists a feasible flow* $x \in \mathcal{E} \to \mathbb{R}_+$ *on* $\mathcal{G}$ *if and only if:*

$$\forall \mathcal{A} \subset \mathcal{V}, \quad \sum_{e \in \delta^+(\mathcal{A})} u_e \geqslant \sum_{v \in \mathcal{A}} d_v + \sum_{e \in \delta^+(\mathcal{A}^c)} \ell_e \,, \tag{1.7}$$

where $\delta_+(\mathcal{A}) \stackrel{\text{def}}{=} \{(u,v) \in \mathcal{E} | u \in \mathcal{A}^c, v \in \mathcal{A}\}$ *is the set of edges coming to set* $\mathcal{A}$ *and* $\mathcal{A}^c \stackrel{\text{def}}{=} \mathcal{V} \setminus \mathcal{A}$.

The following Proposition 1.2 translates Theorem 1.1 in our framework:

Proposition 1.2. *Disaggregation is possible iff:*

$$\forall \mathcal{T}_0 \subset \mathcal{T}, \forall \mathcal{N}_0 \subset \mathcal{N}, \quad \sum_{t \in \mathcal{T}_0} p_t - \sum_{n \in \mathcal{N}_0} E_n + \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} \underline{x}_{n,t} \leqslant \sum_{t \in \mathcal{T}_0, n \notin \mathcal{N}_0} \overline{x}_{n,t} \,. \tag{1.8}$$

Proof. We apply (1.7) with $\mathcal{A} \stackrel{\text{def}}{=} \mathcal{T}_0^c \cup \mathcal{N}_0^c$ and use the equality $d(\mathcal{V}) = 0 = \sum_{v \in} d_v + \sum_{v \in A^c} d_v$. □

From Theorem 1.1 or Proposition 1.2 above, one can observe that the aggregate constraints (1.6) are in general not sufficient to ensure that the disaggregation problem has a solution.

For a given set $\mathcal{T}_0$, there is a choice of $\mathcal{N}_0$ which leads to the strongest inequality (1.8), namely:

$$\sum_{t \in \mathcal{T}_0} p_t \leqslant \min_{\mathcal{N}_0 \subset \mathcal{N}} \left\{ \sum_{n \in \mathcal{N}_0} E_n - \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} \underline{x}_{n,t} + \sum_{t \in \mathcal{T}_0, n \notin \mathcal{N}_0} \overline{x}_{n,t} \right\}, \tag{1.9}$$

In this way, we get $2^T - 2$ inequalities corresponding to the proper subsets $\mathcal{T}_0 \subset \mathcal{T}$. Moreover, in general, these $2^T - 2$ inequalities are not redundant. Although this is not stated in [Hof60], this is a classical result whose proof is elementary.

1.3 Disaggregation based on APM

1.3.1 Generation of Hoffman's constraints with APM

In this section, we propose an algorithm that solves (1.1) while preserving the privacy of each agent constraints $\mathcal{X}_n$ and individual profile $\boldsymbol{x}_n \in \mathbb{R}^T$. To do this, the proposed algorithm is implemented in a decentralized fashion and relies on the alternate projections method (APM) to solve the disaggregation problem (1.2).

Let us consider the polyhedron enforcing the agents constraints:

$$\mathcal{X} \stackrel{\text{def}}{=} \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \,,$$

where

$$\mathcal{X}_n \stackrel{\text{def}}{=} \left\{ \boldsymbol{x}_n \in \mathbb{R}_+^T \mid \textstyle\sum_{t\in\mathcal{T}} x_{n,t} = E_n \text{ and } \forall t,\ \underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t} \right\} . \tag{1.10}$$

Besides, given an allocation $\boldsymbol{p} \in \mathcal{P}$, we consider the set of profiles aggregating to $\boldsymbol{p}$:

$$\mathcal{Y}_{\boldsymbol{p}} \stackrel{\text{def}}{=} \left\{ \boldsymbol{x} \in \mathbb{R}^{NT} \mid \forall t \in \mathcal{T}, \textstyle\sum_{n\in\mathcal{N}} x_{n,t} = p_t \right\} .$$

Note that $\mathcal{Y}_{\boldsymbol{p}}$ is an affine subspace of $\mathbb{R}^{NT}$ (to be distinguished from $\mathcal{P}$ which is a subset of $\mathbb{R}^T$), and that $\mathcal{Y}_{\boldsymbol{p}} \cap \mathcal{X}$ is empty iff $\boldsymbol{p} \notin \mathcal{P}_{\mathrm{D}}$, according to the definition of $\mathcal{P}_{\mathrm{D}}$ in (1.4). The idea of the proposed algorithm is to build a finite sequence of decreasing subsets $(\mathcal{P}^{(s)})_{0\leqslant s\leqslant S}$ such that:

$$\mathcal{P} = \mathcal{P}^{(0)} \supset \mathcal{P}^{(1)} \supset \cdots \supset \mathcal{P}^{(S)} \supset \mathcal{P}_{\mathrm{D}} \,.$$

At each iteration, a new aggregate resource allocation $\boldsymbol{p}^{(s)}$ is obtained by solving an instance of the master problem introduced in (1.3) with $\mathcal{Q} = \mathcal{P}^{(s)}$:

$$\min_{\boldsymbol{p}\in\mathbb{R}^T} f(\boldsymbol{p}) \tag{1.11a}$$

$$\text{s.t. } \boldsymbol{p} \in \mathcal{P}^{(s)} \,. \tag{1.11b}$$

In the remaining of the chapter, we will refer to (1.11) as an instance of *master problem*. Our procedure relies on the following immediate observation:

Proposition 1.3. *If $\boldsymbol{p}^{(s)}$ is a solution of* (1.11)*, and $\mathcal{Y}_{\boldsymbol{p}^{(s)}} \cap \mathcal{X} \neq \emptyset$ and $\boldsymbol{x} \in \mathcal{Y}_{\boldsymbol{p}^{(s)}} \cap \mathcal{X}$, then $(\boldsymbol{p}^{(s)}, \boldsymbol{x})$ is an optimal solution of the initial problem* (1.1)*.* □

Having in hands a solution $\boldsymbol{p}^{(s)}$, we can check if $\mathcal{Y}_{\boldsymbol{p}^{(s)}} \cap \mathcal{X} \neq \emptyset$ using APM on $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}^{(s)}}$, as described in Algorithm 1.1 below (where $\mathcal{Y} = \mathcal{Y}_{\boldsymbol{p}}$).

Algorithm 1.1 Alternate Projections Method (APM)

Require: Start with $\boldsymbol{y}^{(0)}$, $k = 0$, ε_{cvg}, a norm $\|.\|$ on $\mathbb{R}^{NT}$

1: **repeat**
2: $\quad \boldsymbol{x}^{(k+1)} \leftarrow P_{\mathcal{X}}(\boldsymbol{y}^{(k)})$
3: $\quad \boldsymbol{y}^{(k+1)} \leftarrow P_{\mathcal{Y}}(\boldsymbol{x}^{(k+1)})$
4: $\quad k \leftarrow k+1$
5: **until** $\left\| \boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)} \right\| < \varepsilon_{\text{cvg}}$

The idea of using cyclic projections to compute a point in the intersection of two sets comes from Von Neumann [VN50], where the idea was applied for affine subspaces. Convergence of APM is proved by Theorem 1.2:

Theorem 1.2 ([GPR67])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two closed convex sets with $\mathcal{X}$ bounded, and let $(\boldsymbol{x}^{(k)})_k$ and $(\boldsymbol{y}^{(k)})_k$ be the two infinite sequences generated by APM on $\mathcal{X}$ and $\mathcal{Y}$ (Algorithm 1.1) with*

$\varepsilon_{\text{cvg}} = 0$. *Then there exists* $\boldsymbol{x}^\infty \in \mathcal{X}$ *and* $\boldsymbol{y}^\infty \in \mathcal{Y}$ *such that:*

$$\boldsymbol{x}^{(k)} \underset{k\to\infty}{\longrightarrow} \boldsymbol{x}^\infty\,, \quad \boldsymbol{y}^{(k)} \underset{k\to\infty}{\longrightarrow} \boldsymbol{y}^\infty; \tag{1.12a}$$

$$\|\boldsymbol{x}^\infty - \boldsymbol{y}^\infty\|_2 = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}\in\mathcal{Y}} \|\boldsymbol{x}-\boldsymbol{y}\|_2\,. \tag{1.12b}$$

The convergence theorem is illustrated in Figure 1.2 in the case where $\mathcal{X}\cap\mathcal{Y} = \emptyset$, that is, when the disaggregation problem (1.2) is not feasible. The idea of the algorithm proposed

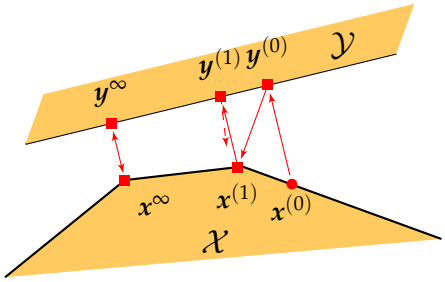


FIGURE 1.2: Alternate projections method (APM) on two sets $\mathcal{X}$ and $\mathcal{Y}$. *When* $\mathcal{X}\cap\mathcal{Y} = \emptyset$, *APM cycles over two points* $\boldsymbol{x}^\infty$ *and* $\boldsymbol{y}^\infty$.

in this chapter is, in the case where $\mathcal{Y}_{\boldsymbol{p}^{(s)}} \cap \mathcal{X} = \emptyset$, to use the resulting vectors $\boldsymbol{x}^\infty$ and $\boldsymbol{y}^\infty$ to construct a new subset $\mathcal{P}^{(s+1)}$ by adding a constraint of type (1.8) to $\mathcal{P}^{(s)}$: indeed, from Proposition 1.2, we know that, if $\mathcal{Y}_{\boldsymbol{p}^{(s)}} \cap \mathcal{X} = \emptyset$, there exists at least one inequality (1.8) violated.

The difficulty is to guess one of this violated inequality among the set of 2^T possible inequalities. It turns out that, using the output of APM, we can build such an inequality.

Suppose that we obtain $\boldsymbol{x}^\infty \neq \boldsymbol{y}^\infty$ as defined in Theorem 1.2: we get a periodic cycle of the APM, that is, we have $\boldsymbol{x}^\infty = P_{\mathcal{X}}(\boldsymbol{y}^\infty)$ and $\boldsymbol{y}^\infty = P_{\mathcal{Y}}(\boldsymbol{x}^\infty)$, and the couple $(\boldsymbol{x}^\infty, \boldsymbol{y}^\infty)$ is the solution of the following optimization problem:

$$\min_{\boldsymbol{x},\boldsymbol{y}} \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 \tag{1.13a}$$

$$\forall n \in \mathcal{N},\ \sum_{t\in\mathcal{T}} x_{n,t} = E_n \qquad (\lambda_n) \tag{1.13b}$$

$$\forall n \in \mathcal{N}, \forall t \in \mathcal{T},\ \underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t} \qquad (\underline{\mu}_{n,t}, \overline{\mu}_{n,t}) \tag{1.13c}$$

$$\forall t \in \mathcal{T},\ \sum_{n\in\mathcal{N}} y_{n,t} = p_t \qquad (\nu_t)\,, \tag{1.13d}$$

where $\lambda_n \in \mathbb{R}$, $\underline{\mu}_{n,t}, \overline{\mu}_{n,t} \in \mathbb{R}_+$ and $\nu_t \in \mathbb{R}$ are the Lagrangian multipliers associated to the constraints (1.13b),(1.13c),(1.13d), with the associated Lagrangian function:

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 - \boldsymbol{\lambda}^\top(\textstyle\sum_t x_{n,t} - E_n)_n - \underline{\boldsymbol{\mu}}^\top(\boldsymbol{x}-\underline{\boldsymbol{x}}) - \overline{\boldsymbol{\mu}}^\top(\overline{\boldsymbol{x}}-\boldsymbol{x}) - \boldsymbol{\nu}^\top(\textstyle\sum_n \boldsymbol{y}_n - \boldsymbol{p})\,.$$

We notice that the stationarity condition of the Lagrangian with respect to the variable $y_{n,t}$ yields:

$$\forall n \in \mathcal{N}, \forall t \in \mathcal{T},\ \nu_t = x_{n,t} - y_{n,t}\,. \tag{1.14}$$

Let us consider the sets $\mathcal{T}_0 \subset \mathcal{T}$ and $\mathcal{N}_0 \subset \mathcal{N}$ defined from the output of APM on $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$ as:

$$\mathcal{T}_0 \overset{\text{def}}{=} \left\{t \in \mathcal{T} \mid \exists n \in \mathcal{N},\ y_{n,t} > \overline{x}_{n,t}\right\} \tag{1.15}$$

$$\text{and } \mathcal{N}_0 \overset{\text{def}}{=} \left\{n \in \mathcal{N} \mid E_n - \textstyle\sum_{t\notin\mathcal{T}_0} \underline{x}_{n,t} - \sum_{t\in\mathcal{T}_0} \overline{x}_{n,t} < 0\right\}. \tag{1.16}$$

In Theorem 1.3 below, we show that applying the inequality (1.8) with the sets $\mathcal{T}_0$ and $\mathcal{N}_0$ defined in (1.16) defines a valid inequality for the disaggregation problem violated by the current allocation $\boldsymbol{p}$.

The intuition behind the definition of $\mathcal{T}_0$ and $\mathcal{N}_0$ in (1.16) is the following: $\mathcal{T}_0$ is the subset of resources for which there is an over supply (which overcomes the upper bound for at least one agent). Once $\mathcal{T}_0$ is defined, $\mathcal{N}_0$ is the associated subset of $\mathcal{N}$ minimizing the right hand side of (1.9). Indeed, (1.9) can be rewritten as:

$$\sum_{t\in\mathcal{T}_0} p_t \leqslant \min_{\mathcal{N}_0\subset\mathcal{N}} \left\{ \sum_{n\in\mathcal{N}_0} \left(E_n - \textstyle\sum_{t\notin\mathcal{T}_0} \underline{x}_{n,t} - \sum_{t\in\mathcal{T}_0} \overline{x}_{n,t}\right) \right\} + \sum_{t\in\mathcal{T}_0, n\in\mathcal{N}} \overline{x}_{n,t}.$$

The following Theorem 1.3 is the key result on which relies the algorithm proposed in this chapter.

Theorem 1.3. *Consider the sequence of iterates $(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)})_{k\in\mathbb{N}}$ generated by the APM on $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$ (see Algorithm 1.1). Then one of the following holds:*

(i) if $\mathcal{X}\cap\mathcal{Y}_{\boldsymbol{p}} \neq \emptyset$, then $\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)} \underset{k\to\infty}{\longrightarrow} \boldsymbol{x}^{\infty} \in \mathcal{X}\cap\mathcal{Y}_{\boldsymbol{p}}$;

(ii) else, if $\mathcal{X}\cap\mathcal{Y}_{\boldsymbol{p}} = \emptyset$, then $\boldsymbol{x}^{(k)} \underset{k\to\infty}{\longrightarrow} \boldsymbol{x}^{\infty} \in \mathcal{X}$ and $\boldsymbol{y}^{(k)} \underset{k\to\infty}{\longrightarrow} \boldsymbol{y}^{\infty} \in \mathcal{Y}_{\boldsymbol{p}}$. Then, considering the sets $\mathcal{T}_0$ and $\mathcal{N}_0$ as defined in (1.16) *gives an inequality of Hoffman* (1.8) *violated by $\boldsymbol{p}$, that is:*

$$\sum_{n\in\mathcal{N}_0} E_n - \sum_{t\in\mathcal{T}_0} p_t + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} \overline{x}_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} < 0\,. \tag{1.17}$$

Moreover, this Hoffman inequality can be written as a function of $\boldsymbol{x}^{\infty}$ as:

$$A_{\mathcal{T}_0}(\boldsymbol{x}^{\infty}) < \sum_{t\in\mathcal{T}_0} p_t \text{ with } A_{\mathcal{T}_0}(\boldsymbol{x}^{\infty}) \overset{\text{def}}{=} \sum_{t\in\mathcal{T}_0}\sum_{n\in\mathcal{N}} x^{\infty}_{n,t}\,. \tag{1.18}$$

Before giving the proof of Theorem 1.3, we need to show some technical properties on the sets $\mathcal{T}_0, \mathcal{N}_0$. For simplicity of notations, we use $\boldsymbol{x}$ and $\boldsymbol{y}$ to denote $\boldsymbol{x}^{\infty}$ and $\boldsymbol{y}^{\infty}$ in Proposition 1.4 and the proof of Theorem 1.3.

Proposition 1.4. *With $\boldsymbol{x} \neq \boldsymbol{y}$ solutions of problem* (1.13) *(outputs of the APM on $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$ with $\varepsilon_{\text{cvg}} = 0$), the following assertions hold:*

(i) $\forall t \in \mathcal{T}_0, \forall n \notin \mathcal{N}_0,\ y_{n,t} \geqslant \overline{x}_{n,t}$ and $x_{n,t} = \overline{x}_{n,t}$;

(ii) $\mathcal{T}_0 = \{t \mid \nu_t > 0\} = \{t \mid p_t > \sum_n x_{n,t}\}$, where ν_t is the optimal Lagrangian multiplier associated to (1.13d)*;*

(iii) $\forall n \in \mathcal{N}_0, \lambda_n < 0$;

(iv) $\forall t \notin \mathcal{T}_0, \forall n \in \mathcal{N}_0,\ x_{n,t} = \underline{x}_{n,t}$;

(v) the sets $\mathcal{T}_0, \mathcal{T}_0^c, \mathcal{N}_0$ and $\mathcal{N}_0^c$ are nonempty.

The proof of Proposition 1.4 is technical and given in Section 1.A. With Proposition 1.4, we are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. We have:

$$\begin{aligned}
&\sum_{n\in\mathcal{N}_0} E_n + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} \overline{x}_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} - \sum_{t\in\mathcal{T}_0} p_t \\
&= \sum_{n\in\mathcal{N}_0} \sum_t x_{n,t} + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} \overline{x}_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} - \sum_{t\in\mathcal{T}_0} p_t \qquad \text{(from (1.13b))} \\
&= \sum_{n\in\mathcal{N}_0} \Big(\sum_{t\notin\mathcal{T}_0} \underline{x}_{n,t} + \sum_{t\in\mathcal{T}_0} x_{n,t} \Big) + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} x_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} - \sum_{t\in\mathcal{T}_0} p_t \quad \text{(from Prop.1.4 (i)-(iv))} \\
&= \sum_{t\in\mathcal{T}_0} \left(\sum_{n\in\mathcal{N}_0} x_{n,t} + \sum_{n\notin\mathcal{N}_0} x_{n,t} \right) - \sum_{t\in\mathcal{T}_0} p_t = \sum_{t\in\mathcal{T}_0} \left(\sum_{n\in\mathcal{N}} x_{n,t} - \sum_{n\in\mathcal{N}} y_{n,t} \right) \\
&= \sum_{t\in\mathcal{T}_0} \left(\sum_{n\in\mathcal{N}} -\nu_t \right) = \sum_{t\in\mathcal{T}_0} \left(- \sum_{n\in\mathcal{N}} |x_{n,t} - y_{n,t}| \right)
\end{aligned}$$

using the stationarity conditions (1.14) and for all $t \in \mathcal{T}_0, \nu_t > 0$ by Prop.1.4 (ii). Moreover, using:

$$\sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} (x_{n,t} - y_{n,t}) = \sum_{n\in\mathcal{N}} E_n - \sum_{t\in\mathcal{T}} p_t = 0\,, \tag{1.19}$$

we see that:

$$\sum_{t\in\mathcal{T}_0} \Big(- \sum_{n\in\mathcal{N}} |x_{n,t} - y_{n,t}| \Big) = -(\|\boldsymbol{x} - \boldsymbol{y}\|_1)/2 < 0\,, \tag{1.20}$$

which shows (1.17). We now show that inequality (1.18) is obtained by a rewriting of (1.17), indeed:

$$\begin{aligned}
&\sum_{n\in\mathcal{N}_0} E_n + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} \overline{x}_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} \\
&= \sum_{n\in\mathcal{N}_0} \sum_{t\in\mathcal{T}} x_{n,t} + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} x_{n,t} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} x_{n,t} \qquad \text{(from Prop.1.4 (i) and (iv))} \\
&= \sum_{t\in\mathcal{T}_0, n\in\mathcal{N}_0} x_{n,t} + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} x_{n,t} = \sum_{t\in\mathcal{T}_0} \sum_{n\in\mathcal{N}} x_{n,t} \; = A_{\mathcal{T}_0}(\boldsymbol{x}).
\end{aligned}$$

□

Suppose, as before, that the two sequences generated by the APM on $\mathcal{X}$ and $\mathcal{Y}$ converge to two distinct points $\boldsymbol{x}^\infty$ and $\boldsymbol{y}^\infty$. Then, at each round k, we can define from (1.40) and considering any $n \in \mathcal{N}$, the multiplier $\boldsymbol{\nu}^{(k)} = \boldsymbol{y}_n^{(k)} - \boldsymbol{x}_n^{(k)}$, which converges to $\boldsymbol{y}_n^\infty - \boldsymbol{x}_n^\infty \stackrel{\text{def}}{=} \boldsymbol{\nu}^\infty$. The set $\mathcal{T}_0$ of Theorem 1.3 is:

$$\mathcal{T}_0^\infty \stackrel{\text{def}}{=} \{t \in \mathcal{T} \mid 0 < \nu_t^\infty\}, \tag{1.21}$$

which raises an issue for practical computation, as $\boldsymbol{\nu}^\infty$ is only obtained *ultimately* by APM, possibly in infinite time. To have access to $\mathcal{T}_0^\infty$ *in finite time*, that is, from one of the iterates $(\boldsymbol{\nu}^{(k)})_k$, we consider the set:

$$\mathcal{T}_0^{(K)} \stackrel{\text{def}}{=} \{t \in \mathcal{T} \mid B\varepsilon_{\text{cvg}} < \nu_t^{(K)}\},$$

where ε_{cvg} is the tolerance for convergence of APM as defined in Algorithm 1.1, $B > 0$ is a constant, and K (depending on ε_{cvg}) is the first integer such that the profile $\boldsymbol{x}^{(k)}$ satisfies $\left\|\boldsymbol{x}^{(K)} - \boldsymbol{x}^{(K-1)}\right\| < \varepsilon_{\text{cvg}}$.

We next show that we can choose B to ensure that $\mathcal{T}_0^{(K)} = \mathcal{T}_0^\infty$ for ε_{cvg} small enough. We rely on the geometric convergence rate of APM on polyhedra [BLY14; NJJ14]:

Proposition 1.5. *[NJJ14] If $\mathcal{X}$ and $\mathcal{Y}$ are polyhedra, there exists $\rho \in (0,1)$ such that the sequence $(\boldsymbol{x}^{(k)})_k$ and $(\boldsymbol{y}^{(k)})_k$ generated by APM verify for all $k \geqslant 1$:*

$$\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\right\|_2 \leqslant \rho \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\|_2 \quad \textit{and} \quad \left\|\boldsymbol{y}^{(k+1)} - \boldsymbol{y}^{(k)}\right\|_2 \leqslant \rho \left\|\boldsymbol{y}^{(k)} - \boldsymbol{y}^{(k-1)}\right\|_2.$$

Proposition 1.5 applies to any polyhedra $\mathcal{X}$ and $\mathcal{Y}$. In Section 1.3.3 we shall give an explicit upper bound on the constant ρ in the specific transportation case given by (1.1c) and (1.5).

From the previous proposition, we can quantify the distance to the limits in terms of ρ:

Lemma 1.1. *Consider an integer K such that the sequence $(\boldsymbol{x}^{(k)})_{k\geqslant 0}$ generated by APM satisfies $\left\|\boldsymbol{x}^{(K)}-\boldsymbol{x}^{(K-1)}\right\| \leqslant \varepsilon_{\text{cvg}}$, then we have for any $K' \geqslant K-1$:*

$$\left\|\boldsymbol{x}^{\infty}-\boldsymbol{x}^{(K')}\right\| \leqslant \tfrac{\varepsilon_{\text{cvg}}}{1-\rho}\,.$$

Proof. From Proposition 1.5, we have for any $k \geqslant K$:

$$\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(K')}\right\| \leqslant \sum_{s=0}^{k-K'}\left\|\boldsymbol{x}^{(K+s+1)}-\boldsymbol{x}^{(K+s)}\right\| \leqslant \sum_{s=0}^{k-K'}\rho^s\left\|\boldsymbol{x}^{(K'+1)}-\boldsymbol{x}^{(K')}\right\| \leqslant \frac{1}{1-\rho}\varepsilon_{\text{cvg}}\,,$$

so that, by taking the limit $k\to\infty$, one obtains $\left\|\boldsymbol{x}^{\infty}-\boldsymbol{x}^{(K')}\right\| \leqslant \frac{\varepsilon_{\text{cvg}}}{1-\rho}$. □

With this previous lemma, we can state the condition on B ensuring the desired property:

Proposition 1.6. *Define $\underline{\nu} \stackrel{\text{def}}{=} \min\{|\nu_t^{\infty}| > 0\}$ (least nonzero element of $\boldsymbol{\nu}^{\infty}$). If the constants B and $\varepsilon_{\text{cvg}} > 0$ are chosen such that $B > \frac{1}{1-\rho}$ and $\varepsilon_{\text{cvg}} \times 2B < \underline{\nu}$, and Algorithm 1.1 stops at iteration K, then we have:*

$$\mathcal{T}_0^{(K)} = \mathcal{T}_0^{\infty}\,.$$

Proof. Let $t \in \mathcal{T}_0^{\infty}$, that is $\nu_t^{\infty} > 0$ which is equivalent to $\nu_t^{\infty} \geqslant \underline{\nu}$ by definition of $\underline{\nu}$. We have:

$$\begin{aligned}\nu_t^{(K)} = \frac{1}{N}(p_t - \sum_n x_{n,t}^{(K)}) &= \frac{1}{N}(p_t - \sum_n x_{n,t}^{\infty}) + \frac{1}{N}(\sum_n x_{n,t}^{\infty} - \sum_n x_{n,t}^{(K)}) \\ &> \nu_t^{\infty} - \tfrac{\varepsilon_{\text{cvg}}}{1-\rho} \geqslant \underline{\nu} - \tfrac{\varepsilon_{\text{cvg}}}{1-\rho} > \varepsilon_{\text{cvg}}(2B - \tfrac{1}{1-\rho})\,,\end{aligned}$$

and this last quantity is greater than $B\varepsilon_{\text{cvg}}$ as soon as $B \geqslant \frac{1}{1-\rho}$, thus $t \in \mathcal{T}_0^{(K)}$.

Conversely, if $t \in \mathcal{T}_0^{(K)}$, then:

$$\begin{aligned}\nu_t^{\infty} = \frac{1}{N}(p_t - \sum_n x_{n,t}^{\infty}) &= \frac{1}{N}(p_t - \sum_n x_{n,t}^{(K)}) - \frac{1}{N}(\sum_n x_{n,t}^{\infty} - \sum_n x_{n,t}^{(K)}) \\ \geqslant \nu_t^{(K)} - \tfrac{B}{1-\rho} \;&>\; \nu_t^{(K)} - B\varepsilon_{\text{cvg}} \geqslant (B-B)\varepsilon_{\text{cvg}} \geqslant 0\,,\end{aligned}$$

so that $t \in \mathcal{T}_0^{\infty}$. Furthermore, the "approximated" cut $\sum_{t\in\mathcal{T}_0}\left(\sum_{n\in\mathcal{N}} x_{n,t}^{(K)} - p_t\right) \geqslant 0$ is violated by the current value of $\boldsymbol{p}$ (or $\boldsymbol{p}^{(s)}$ at iteration s) in the algorithm as:

$$\begin{aligned}\sum_{t\in\mathcal{T}_0}\left(\sum_{n\in\mathcal{N}} x_{n,t}^{(K)} - p_t\right) &\leqslant \sum_{t\in\mathcal{T}_0}\left(\sum_{n\in\mathcal{N}} x_{n,t}^{(K)} - x_{n,t}^{\infty}\right) + \sum_{t\in\mathcal{T}_0}\left(\sum_{n\in\mathcal{N}} x_{n,t}^{\infty} - p_t\right) \\ &\leqslant \left\|\boldsymbol{x}^{(K)} - \boldsymbol{x}^{\infty}\right\|_1 - \frac{1}{2}\left\|\boldsymbol{x}^{\infty} - \boldsymbol{y}^{\infty}\right\|_1\end{aligned}$$

using (1.19) and (1.20). This last quantity is negative as soon as $\left\|\boldsymbol{x}^{(K)} - \boldsymbol{x}^{\infty}\right\|_1 < \frac{1}{2}\left\|\boldsymbol{x}^{\infty} - \boldsymbol{y}^{\infty}\right\|_1$ which holds in particular if $B\varepsilon_{\text{cvg}} < \frac{1}{2}\left\|\boldsymbol{x}^{\infty} - \boldsymbol{y}^{\infty}\right\|_1$. □

This second proposition shows a surprising result: even if we do not have access to the limit $\boldsymbol{x}^{\infty}$, we can compute *in finite time* the *exact* left hand side term $A_{\mathcal{T}_0}(\boldsymbol{x}^{\infty})$ of the cut (1.18):

Proposition 1.7. *Under the hypotheses of Proposition 1.6, we have:*

$$A_{\mathcal{T}_0}(\boldsymbol{x}^{(K)}) = \sum_{t\in\mathcal{T}_0}\sum_{n\in\mathcal{N}} x_{n,t}^{(K)} = A_{\mathcal{T}_0}(\boldsymbol{x}^{\infty})\,.$$

Proof. We start by showing some technical properties similar to Proposition 1.4:

Lemma 1.2. *The iterate $\boldsymbol{x}^{(K)}$ satisfies the following properties:*

(i) $\forall t \in \mathcal{T}_0, \forall n \notin \mathcal{N}_0, x_{n,t}^{(K)} = x_{n,t}^{\infty} = \overline{x}_{n,t}$;

(ii) $\forall t \notin \mathcal{T}_0, \forall n \in \mathcal{N}_0,\ x_{n,t}^{(K)} = x_{n,t}^{\infty} = \underline{x}_{n,t}$.

The proof of Lemma 1.2 is similar to Proposition 1.4 and is given in Section 1.B. Then, having in mind that $\mathcal{T}_0^{(K)} = \mathcal{T}_0^{\infty}$ from Proposition 1.6, and $\mathcal{N}_0$ is obtained from $\mathcal{T}_0^{\infty}$ by (1.16), we obtain:

$$\begin{aligned} A_{\mathcal{T}_0}(\boldsymbol{x}^{(K)}) &= \sum_{n\in\mathcal{N}_0} \Big(\sum_{t\notin\mathcal{T}_0} \underline{x}_{n,t} + \sum_{t\in\mathcal{T}_0} x_{n,t}^{(K)} \Big) - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} x_{n,t}^{(K)} \\ &\leqslant \sum_{n\in\mathcal{N}_0} \sum_{t\in\mathcal{T}} x_{n,t}^{(K)} - \sum_{t\notin\mathcal{T}_0, n\in\mathcal{N}_0} \underline{x}_{n,t} + \sum_{t\in\mathcal{T}_0, n\notin\mathcal{N}_0} \overline{x}_{n,t} \quad \text{(from Lemma 1.2)} \end{aligned}$$

which equals to $A_{\mathcal{T}_0}(\boldsymbol{x}^{\infty})$ as we have $\sum_{t\in\mathcal{T}} x_{n,t}^{(K)} = E_n$ for each $n \in \mathcal{N}$. □

Before presenting our algorithm using this last result, we focus on the technique of multiparty secure computation (SMC) which will be used here to ensure the privacy of agent's constraints and profiles while running the APM.

1.3.2 Privacy-preserving Projections through SMC

APM, as described in Algorithm 1.1, enables a distributed implementation in our context, by the structure of the algorithm itself: the operator computes the projection on $\mathcal{Y}_{\boldsymbol{p}}$ while each agent n can compute, possibly in parallel, the projection on $\mathcal{X}_n$ of the new profile transmitted by the operator. This enables each agent (as well as the operator) to keep her individual constraint and not reveal it to the operator or other agents. However, each agent has to transmit back her newly computed individual profile to the operator for the next iteration.

Using a secure multi-party computation (SMC) protocol (see [Yao86]), we can avoid this communication of individual profiles and perform APM without revealing the sequence of agent profiles $\boldsymbol{x}$ to the aggregator.

For this, we use the fact that $\mathcal{Y}_{\boldsymbol{p}}$ is an affine subspace and thus the projection on $\mathcal{Y}_{\boldsymbol{p}}$ can be obtained explicitly component-wise. Indeed, summing (1.14) on $\mathcal{T}$, we immediatly obtain:

$$\forall n \in \mathcal{N},\ [\mathcal{P}_{\mathcal{Y}_{\boldsymbol{p}}}(\boldsymbol{x})]_n = \boldsymbol{x}_n + \tfrac{1}{N}(\boldsymbol{p} - \textstyle\sum_{m\in\mathcal{N}} \boldsymbol{x}_m)\ . \tag{1.22}$$

Thus, having access to the *aggregate* profile $\boldsymbol{S} \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} \boldsymbol{x}_n$, each agent can compute locally the component of the projection on $\mathcal{Y}_{\boldsymbol{p}}$ of her profile, instead of transmitting the profile to the operator for computing the projection in a centralized way.

Using SMC, the sum $\boldsymbol{S}$ can be computed in a non-intrusive manner, by several communications between agents and the operator, as described in Algorithm 1.2. The main idea of SMC is that, instead of sending her profile $\boldsymbol{x}_n$, agent n *splits* $x_{n,t}$ for each t into N random parts $(s_{n,t,m})_m$, according to an uniform distribution and summing to $\boldsymbol{x}_{n,t}$ (Lines 2-3). Thus, each part $s_{n,t,m}$ taken individually does not reveal any information on $\boldsymbol{x}_n$ nor on $\mathcal{X}_n$, and can be sent to agent m. Once all exchanges of parts are completed (Line 5), and n has herself received the parts from other agents, agent n computes a new aggregate quantity $\boldsymbol{\sigma}_n$ (Line 7), which does not contain either any information about any of the agents, and sends it to the operator (Line 8). The operator can finally compute the quantity $\boldsymbol{S} = \boldsymbol{x}^\top \mathbb{1}_N = \boldsymbol{\sigma}^\top \mathbb{1}_N$.

Remark 1.1. *As $\boldsymbol{\sigma}_n$, and $\boldsymbol{s}_n$ are random by construction, an eavesdropper aiming to learn the profile $\boldsymbol{x}_n$ of n has no choice but to intercept all the communications of n to all other agents (to learn $(s_{n,t,m})_{m\neq n}$ and $(s_{m,t,n})_{m\neq n}$) and to the operator (to learn σ_n).*

We sum up in Algorithm 1.3 below the procedure of generating a new constraint as stated in Theorem 1.3 from the output of APM in finite time (see Proposition 1.6) and in a privacy-preserving way using SMC.

Algorithm 1.2 SMC of Aggregate (SMCA) $\sum_{n\in\mathcal{N}} \boldsymbol{x}_n$

Require: A profile $\boldsymbol{x}_n$ for each agent $n \in \mathcal{N}$
for each agent $n \in \mathcal{N}$ **do**
  Draw $\forall t, (s_{n,t,m})_{m=1}^{N-1} \in \mathcal{U}([0,A]^{N-1})$
  and set $\forall t, s_{n,t,N} \stackrel{\text{def}}{=} x_{n,t} - \sum_{m=1}^{N-1} s_{n,t,m}$
  Send $(s_{n,t,m})_{t\in\mathcal{T}}$ to agent $m \in \mathcal{N}$
done
for each agent $n \in \mathcal{N}$ **do**
  Compute $\forall t, \sigma_{n,t} = \sum_{m\in\mathcal{N}} s_{m,t,n}$
  Send $(\sigma_{n,t})_{t\in\mathcal{T}}$ to operator
done
Operator computes $\boldsymbol{S} = \sum_{n\in\mathcal{N}} \boldsymbol{\sigma}_n$ (and broadcasts it to agents)

Algorithm 1.3 Non-intrusive APM (NI-APM)

Require: Start with $\boldsymbol{y}^{(0)}$, $k=1$, $\varepsilon_{\text{cvg}}, \varepsilon_{\text{dis}}$, norm $\|.\|$ on $\mathbb{R}^{NT}$
1: **repeat**
2: **for** each agent $n \in \mathcal{N}$ **do**
3: $\boldsymbol{x}_n^{(k)} \leftarrow P_{\mathcal{X}_n}(\boldsymbol{y}_n^{(k-1)})$
4: **done**
5: Operator obtains $\boldsymbol{S}^{(k)} \leftarrow$SMCA$(\boldsymbol{x}^{(k)})$ (*cf* Algo. 1.2)
6: and sends $\boldsymbol{\nu}^{(k)} \stackrel{\text{def}}{=} \frac{1}{N}(\boldsymbol{p} - \boldsymbol{S}^{(k)}) \in \mathbb{R}^T$ to agents $\mathcal{N}$
7: **for** each agent $n \in \mathcal{N}$ **do**
8: Compute $\boldsymbol{y}_n^{(k)} \leftarrow \boldsymbol{x}_n^{(k)} + \boldsymbol{\nu}^{(k)}$ ▷ *from* (1.14) *and* (1.22), $\boldsymbol{y}_n^{(k)} = [P_{\mathcal{Y}_{\boldsymbol{p}}}(\boldsymbol{x}^{(k)})]_n$
9: **done**
10: $k \leftarrow k+1$
11: **until** $\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\| < \varepsilon_{\text{cvg}}$
12: **if** $\left\|\boldsymbol{x}^{(k)} - \boldsymbol{y}^{(k)}\right\| \leqslant \varepsilon_{\text{dis}}$ **then** ▷ *found a* ε_{dis}*-solution of the disaggregation problem*
13: Each agent adopts profile $\boldsymbol{x}_n^{(k)}$
14: **return** DISAG $\leftarrow$ TRUE
15: **else** ▷ *have to find a valid constraint violated by* $\boldsymbol{p}$
16: Operator computes $\mathcal{T}_0 \leftarrow \{t \in \mathcal{T} \mid B\varepsilon_{\text{cvg}} < \nu_t^{(k)}\}$
17: Operator computes $A_{\mathcal{T}_0} \leftarrow$ SMCA$(\ (\boldsymbol{x}_t^{(k)})_{t\in\mathcal{T}_0})$
18: **if** $A_{\mathcal{T}_0} - \sum_{t\in\mathcal{T}_0} \boldsymbol{p}_t < 0$ **then**
19: **return** DISAG $\leftarrow$ FALSE, $\mathcal{T}_0$, $A_{\mathcal{T}_0}$
20: **else** ▷ *need to run APM with higher precision*
21: Return to Line 1 with $\varepsilon_{\text{cvg}} \leftarrow \varepsilon_{\text{cvg}}/2$
22: **end**
23: **end**

To choose B and ε_{cvg} satisfying the conditions of Proposition 1.6 *a priori*, one has to know the value of $\underline{\nu}$. Although a conservative lower bound could be obtained by Diophantine arguments if we consider rationals as inputs of the algorithm, in practice it is easier and more efficient to proceed in an iterative manner for the value of ε_{cvg}. Indeed, one can start with ε_{cvg} arbitrary large so that APM will converge quickly, and then check if the cut obtained is violated by the current value of $\boldsymbol{p}$ (Line 18): if it is not the case, we can continue the iterations of APM with convergence precision improved to $\varepsilon_{\text{cvg}}/2$ (Line 21). Proposition 1.6 ensures that this loop terminates in finite time.

The parameter $\varepsilon_{\text{dis}} > 0$ (Line 12 of Algorithm 1.3) has to be chosen a priori by the operator, depending on the precision required. In general in APM, $\boldsymbol{x}^\infty = \boldsymbol{y}^\infty$ will only be achieved in infinite time, so choosing ε_{dis} strictly positive is required.

We end this section by summarizing in Algorithm 1.4 the global iterative procedure to compute an optimal and disaggregable resource allocation $\boldsymbol{p}$, solution of the initial problem

(1.1), using iteratively NI-APM (Algorithm 1.3) and adding constraints as stated in Theorem 1.3.

Algorithm 1.4 Non-intrusive Optimal Disaggregation

Require: $s = 0$, $\mathcal{P}^{(0)} = \mathcal{P}$; DISAG= FALSE
1: **while** Not DISAG **do**
2: Solve $\min_{\boldsymbol{p} \in \mathcal{P}^{(s)}} f(\boldsymbol{p})$
3: **if** problem infeasible **then**
4: Exit
5: **else**
6: Compute $\boldsymbol{p}^{(s)} = \arg\min_{\boldsymbol{p} \in \mathcal{P}^{(s)}} f(\boldsymbol{p})$
7: **end**
8: DISAG$\leftarrow$ NI-APM($\boldsymbol{p}^{(s)}$) (Algo. 1.3)
9: **if** DISAG **then**
10: Operator adopts $\boldsymbol{p}^{(s)}$
11: **else**
12: Obtain $\mathcal{T}_0^{(s)}$, $A_{\mathcal{T}_0}^{(s)}$ from NI-APM($\boldsymbol{p}^{(s)}$)
13: $\mathcal{P}^{(s+1)} \leftarrow \mathcal{P}^{(s)} \cap \{\boldsymbol{p} | \sum_{t \in \mathcal{T}_0^{(s)}} p_t \leqslant A_{\mathcal{T}_0}^{(s)}\}$
14: **end**
15: $s \leftarrow s + 1$
16: **done**

Algorithm 1.4 iteratively calls NI-APM (Algorithm 1.3) and in case disaggregation is not possible (Line 11), a new constraint is added (Line 13), obtained from the quantity $A_{\mathcal{T}_0}$ defined in (1.18), to the feasible set of resource allocations $\mathcal{P}^{(s)}$ in problem (1.11). This constraint is an inequality on $\boldsymbol{p}$ and thus does not reveal significant individual information to the operator. The algorithm stops when disaggregation is possible (Line 9). The termination of Algorithm 1.4 is ensured by the following property and the form of the constraints added (1.17):

Proposition 1.8. *Algorithm 1.4 stops after a finite number of iterations, as at most* $2^T - 2$ *constraints (Line 13) can be added to the master problem (Line 2).*

The following Proposition 1.9 shows the correctness of our Algorithm 1.4.

Proposition 1.9. *Let* B *and* ε_{cvg} *satisfy the conditions of Proposition 1.6 and 1.7. Then:*

- *if the problem* (1.1) *has no solution, Algorithm 1.4 exits at Line 4 after at most* $2^T - 2$ *iterations;*

- *else, Algorithm 1.4 computes, after at most* $s \leqslant 2^T - 2$ *iterations, an aggregate solution* $\boldsymbol{p}^{(s)} \in \mathcal{P}$*, associated to individual profiles* $(\boldsymbol{x}^*)_n$ *= NI-APM*$(\boldsymbol{p}^{(s)})$ *such that:*

$$\boldsymbol{p}^{(s)} \in \mathcal{P}, \quad \forall n \in \mathcal{N}, \boldsymbol{x}_n^* \in \mathcal{X}_n, \quad \left\| \sum_{n \in \mathcal{N}} \boldsymbol{x}_n^* - \boldsymbol{p}^{(s)} \right\| \leqslant \varepsilon_{\text{dis}} \quad \text{and } f(\boldsymbol{p}^{(s)}) \leqslant f^* \,,$$

where f^* *is the optimal value of problem* (1.1).

Proof. The proof is immediate from Theorem 1.3, Proposition 1.6 and Proposition 1.7. □

Remark 1.2. *The upper bound on the number of constraints added has no dependence on* N *because, as stated in* (1.9)*, once a subset of* $\mathcal{T}$ *is chosen, the constraint we add in the algorithm is found by taking the minimum over the subsets of* $\mathcal{N}$*.*

Although there exist some instances with an exponential number of independent constraints, this does not jeopardize the proposed method: in practice, the algorithm stops after a very small number of constraints added. Intuitively, we will only add constraints "supporting" the optimal allocation $\boldsymbol{p}$. Thus, Algorithm 1.4 is a method which enables the operator to compute a resource allocation $\boldsymbol{p}$ and the N agents to adopt profiles $(\boldsymbol{x}_n)_n$, such that $(\boldsymbol{x}, \boldsymbol{p})$ solves the global problem (1.1), and the method ensures that both agent constraints

(upper bounds $\overline{x}_n$, lower bounds $\underline{x}_n$, demand E_n); and disaggregate (individual) profile $\boldsymbol{x}_n$ (as well as the iterates $(\boldsymbol{x}^{(k)})_k$ and $(\boldsymbol{y}^{(k)})_k$ in NI-APM) are kept confidential by agent n and can not be induced by a third party (either the operator or any other agent $m \neq n$).

Remark 1.3. *A natural approach to adress problem* (1.1) *in a distributed way, assuming that both the cost function $\boldsymbol{p} \mapsto f(\boldsymbol{p})$ and the feasibility set $\mathcal{P}$ are convex, is to rely on Lagrangian based decomposition techniques. Examples of such methods are Dual subgradient methods [Ber99, Chapter 6], auxiliary problem principle method [CZ84], ADMM [GM75],[YN17] or bundle methods [LNN95].*

One can think of a privacy-preserving implementation of those techniques, where Lagrangian multipliers associated to the (relaxed) aggregation constraint $\sum_n \boldsymbol{x}_n = \boldsymbol{p}$ would be updated using the SMC technique as described in Algorithm 1.2. However, those techniques usually ask for strong convexity hypothesis: for instance, in ADMM, in order to keep the decomposition structure in agent by agent, a possibility is to use multi-blocs ADMM with $N+1$ blocs (N agents and the operator), which is known to converge in the condition that strong convexity of the cost function in at least N of the $N+1$ variables holds [Den+17]. The complete study of privacy-preserving implementations of Lagrangian decomposition methods is left for further work.

The advantage of Algorithm 1.4 proposed here is that convergence is ensured (see Proposition 1.8) even if the cost function $\boldsymbol{p} \mapsto f(\boldsymbol{p})$ and the feasibility set $\mathcal{P}$ are not convex, which is the case in many practical situations (see Section 1.5.2).

In the next section, we focus on the convergence rate of APM in the particular case of transportation constraints, precising the geometric rate stated in Theorem 1.2.

1.3.3 Complexity Analysis of APM in the Transportation Case

In this section we analyze the speed of convergence of the alternate projections method (APM) described in Algorithm 1.1 on the sets $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$ defined in Section 1.2.

A general result in [BLY14] gives an upper bound on the convergence of the sequences generated by APM on $\mathcal{X}$ and $\mathcal{Y}$ if these two sets are semi-algebraic. In particular, it establishes the geometric convergence for polyhedral sets. However, as stated in [NJJ14], given two particular polyhedral sets $\mathcal{X}$ and $\mathcal{Y}$, it is not straightforward to deduce an explicit rate of convergence from their result.

The authors in [NJJ14] established in a particular case a geometric convergence with an explicit upper bound on the convergence rate. They consider APM on two sets P and Q, where P is a linear subspace and Q is a product of base polytopes of submodular functions.

In this section, we also establish an explicit upper bound on the convergence rate of APM in the transportation case, that is with $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$ defined in (1.5) and (1.2b):

Theorem 1.4. *For the two sets $\mathcal{X}$ and $\mathcal{Y}_{\boldsymbol{p}}$, the sequence of alternate projections converges to $\boldsymbol{x}^* \in \mathcal{X}$, $\boldsymbol{y}^* \in \mathcal{X}^P$ satisfying $\|\boldsymbol{x}^* - \boldsymbol{y}^*\| = \inf_{\boldsymbol{x}\in\mathcal{X}, \boldsymbol{y}\in\mathcal{Y}_{\boldsymbol{p}}} \|\boldsymbol{x} - \boldsymbol{y}\|$, at the geometrical rate:*

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\| \leqslant 2\left\|\boldsymbol{x}^{(0)} - \boldsymbol{x}^*\right\| \times \left(1 - \tfrac{4}{N(T+1)^2(T-1)}\right)^k,$$

and the symmetric inequalities hold for $(\boldsymbol{y}^{(k)})_k$.

For the remaining of this section, we will just use $\mathcal{Y}$ to denote $\mathcal{Y}_{\boldsymbol{p}}$, as $\boldsymbol{p}$ remains fixed during APM. For the result stated in Theorem 1.4 above, we use several partial results of [NJJ14].

Proof. First, we use the fact stated in [NJJ14] that APM on subspaces U and V converge with geometric rate $c_F(U,V)^2$, where the rate is given by the square of the cosine of the Friedrichs angle between U and V, given by:

$$c_F(U,V) \stackrel{\text{def}}{=} \sup\{u^T v \mid u \in U \cap (U\cap V)^\perp, v \in V \cap (U\cap V)^\perp, \|u\| \leqslant 1, \|v\| \leqslant 1\}.$$

An intuitive generalization of this result for polyhedra $\mathcal{X}$ and $\mathcal{Y}$, considering all affine subspaces supporting the faces of $\mathcal{X}$ and $\mathcal{Y}$ is given in [NJJ14]:

Lemma 1.3 ([NJJ14]). *For APM on polyhedra $\mathcal{X}$ and $\mathcal{Y}$ in $\mathbb{R}^D$, the convergence is geometric with rate bounded by the square of the maximal cosine of Friedrichs angle between subspaces supporting faces of $\mathcal{X}$ and $\mathcal{Y}$:*

$$\max_{\boldsymbol{x},\boldsymbol{y}} c_F\big(\operatorname{aff}_0(\mathcal{X}_{\boldsymbol{x}}), \operatorname{aff}_0(\mathcal{Y}_{\boldsymbol{y}})\big), \tag{1.23}$$

where, for any $\boldsymbol{x} \in \mathbb{R}^D$, $\mathcal{X}_{\boldsymbol{x}} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{v}\in\mathcal{X}} \boldsymbol{x}^\top \boldsymbol{v}$ is the face of $\mathcal{X}$ generated by direction $\boldsymbol{x}$ and $\operatorname{aff}_0(C) = \operatorname{aff}(C) - \boldsymbol{c}$ for some $\boldsymbol{c} \in C$ denotes the subspace supporting the affine hull of C, for $C = \mathcal{X}_{\boldsymbol{x}}$ or $C = \mathcal{Y}_{\boldsymbol{y}}$.

In the remaining of the proof, we bound the quantity (1.23) for our polyhedra $\mathcal{X}$ and $\mathcal{Y}$.

For this, we use the space $\mathbb{R}^{NT} = \mathbb{R}^T \times \cdots \times \mathbb{R}^T$, where the $(n-1)T+1$ to nT entries correspond to the profile of agent n, for $1 \leqslant n \leqslant N$. As in [NJJ14], we use a result connecting angles between subspaces and the eigenvalues of matrices giving the directions of these spaces:

Lemma 1.4 ([NJJ14]). *If A and B are matrices with orthonormal rows with same number of columns, then:*

- *if all singular values of $AB^\top$ are equal to one, then $c_F\big(\operatorname{Ker} A,\ \operatorname{Ker} B\big) = 0$;*
- *else, $c_F\big(\operatorname{Ker} A,\ \operatorname{Ker} B\big)$ is equal to the largest singular value of $AB^\top$ among those that are smaller than one.*

We are left with finding a matricial representation of the faces of polyhedra $\mathcal{X}$ and $\mathcal{Y}$ and, then, bounding the corresponding singular values.

In our case, the polyhedra $\mathcal{Y}$ is an affine subspace $\mathcal{Y} = \{\boldsymbol{x} \in \mathbb{R}^{NT} \,|A\boldsymbol{x} = \sqrt{N}^{-1}\boldsymbol{p}\}$ where:

$$A \stackrel{\text{def}}{=} \sqrt{N}^{-1} J_{1,N} \otimes I_T,$$

where $\otimes$ denotes the Kronecker product. The matrix A has orthonormal rows and the linear subspace associated to $\mathcal{Y}$ is equal to $\operatorname{Ker}(A)$.

Obtaining a matricial representation of the faces of $\mathcal{X}$ is more complex. The faces of $\mathcal{X}$ are obtained by considering, for each $n \in \mathcal{N}$, subsets of the time periods that are at lower or upper bound (respectively $\underline{\mathcal{T}}_n$ and $\overline{\mathcal{T}}_n$, with $\underline{\mathcal{T}}_n \cap \overline{\mathcal{T}}_n = \emptyset$). Considering a collection of such subsets, a face of $\mathcal{X}$ can be written as:

$$\mathcal{A}_{(\overline{\mathcal{T}}_n,\underline{\mathcal{T}}_n)_n} \stackrel{\text{def}}{=} \Big\{(\boldsymbol{x})_{n,t} \,|\forall n,\ \textstyle\sum_t x_{n,t} = E_n \text{ and } \forall t \in \underline{\mathcal{T}}_n, x_{n,t} = \underline{x}_{n,t}, \text{ and } \forall t \in \overline{\mathcal{T}}_n, x_{n,t} = \overline{x}_{n,t}\Big\}.$$

For some particular collection of subsets $(\overline{\mathcal{T}}_n, \underline{\mathcal{T}}_n)_n$, the set $\mathcal{A}_{(\overline{\mathcal{T}}_n,\underline{\mathcal{T}}_n)_n}$ might be empty. The linear subspace associated to $\mathcal{A}_{(\overline{\mathcal{T}}_n,\underline{\mathcal{T}}_n)_n}$ is given by $\{\boldsymbol{x} \in \mathbb{R}^{NT} |\ B\boldsymbol{x} = 0\} = \operatorname{Ker}(B)$, where the N first rows of B, corresponding to the constraints $\sum_t x_{n,t} = E_n$, are given before orthonormalization by:

$$\sqrt{T}^{-1} I_N \otimes J_{1,T}\ ,$$

and the matrix B has $b \stackrel{\text{def}}{=} \sum_n |\mathcal{T}_n|$ more rows, where $\mathcal{T}_n \stackrel{\text{def}}{=} \underline{\mathcal{T}_n} \cup \overline{\mathcal{T}_n}$, corresponding to the saturated bounds. Each of this row is given by the unit vector $\boldsymbol{e}_{n,t}^\top \in \mathbb{R}^{NT}$ for $n \in \mathcal{N}$, $t \in \mathcal{T}_n$, which gives already an orthonormalized family of (unit) vectors. Therefore, a simple orthonormalized matrix B giving the direction of $\mathcal{A}_{(\overline{\mathcal{T}}_n,\underline{\mathcal{T}}_n)_n}$ is given by:

$$B \stackrel{\text{def}}{=} \begin{pmatrix} \sqrt{T-|\mathcal{T}_1|}^{-1}\mathbf{1}_{\mathcal{T}_1^c}^\top & 0 & \dots & 0 \\ 0 & \sqrt{T-|\mathcal{T}_2|}^{-1}\mathbf{1}_{\mathcal{T}_2^c}^\top & \dots & 0 \\ 0 & 0 & \dots & \sqrt{T-|\mathcal{T}_N|}^{-1}\mathbf{1}_{\mathcal{T}_N^c}^\top \\ B_{\mathcal{T}_1} & 0 & \dots & 0 \\ 0 & B_{\mathcal{T}_2} & \dots & 0 \\ 0 & & \dots & B_{\mathcal{T}_N} \end{pmatrix},$$

where $\mathbf{1}_{\mathcal{T}_n^c} \in \mathbb{R}^T$ is the vector where the indices in $\mathcal{T}_n^c$ are equal to 1 and 0 otherwise, and $B_{\mathcal{T}_n} \stackrel{\text{def}}{=} \sum_{\substack{1 \leqslant k \leqslant |\mathcal{T}_n| \\ \mathcal{T}_n = \{t_1, \ldots, t_{|\mathcal{T}_n|}\}}} E_{kt_k}$ is the matrix $|\mathcal{T}_n| \times T$ with indices of $\mathcal{T}_n$. We obtain the double product:

$$
\begin{aligned}
(AB^\top)(BA^\top) &= \frac{1}{N}\left(\sum_n \frac{\mathbb{1}_{k \notin \mathcal{T}_n \wedge \ell \notin \mathcal{T}_n}}{T - |\mathcal{T}_n|}\right)_{1 \leqslant k,\ell \leqslant T} + \frac{1}{N}\sum_n B_{\mathcal{T}_n}^\top B_{\mathcal{T}_n} \\
&= \frac{1}{N}\left(\sum_n \frac{\mathbb{1}_{\{k,\ell\} \subset \mathcal{T}_n^c}}{T - |\mathcal{T}_n|}\right)_{1 \leqslant k,\ell \leqslant T} + \frac{1}{N}\sum_{1 \leqslant t \leqslant T} \left(\sum_n \mathbb{1}_{t \in \mathcal{T}_n}\right) E_{t,t} \, .
\end{aligned}
$$

We observe that:

- if $t_0 \in \bigcap_{i=1}^N \mathcal{T}_n$, then $\boldsymbol{e}_{t_0}$ is an eigenvector associated to eigenvalue $\lambda_{t_0} = 1$;
- the vector $\mathbf{1}_{\bar{\mathcal{T}}} \stackrel{\text{def}}{=} (\mathbb{1}_{t \notin \cap_n \mathcal{T}_n})_{t \in \mathcal{T}} \in \mathbb{R}^T$, where $\bar{\mathcal{T}} \stackrel{\text{def}}{=} \cup_n \mathcal{T}_n^c$, is an eigenvector associated to eigenvalue $\lambda = 1$. Indeed, if we denote by $\mathcal{N}_\theta = \{n \in \mathcal{N} | \theta \in \mathcal{T}_n\}$, then $[\mathbf{1}_{\bar{\mathcal{T}}}]_\theta = 1 \Leftrightarrow \mathcal{N}_\theta^c \neq \emptyset$, and for each $\theta \in \bar{\mathcal{T}}$:

$$
\begin{aligned}
[(AB^\top)(BA^\top)]_\theta \mathbf{1}_{\bar{\mathcal{T}}} &= \frac{1}{N}\left(\sum_{i \in \mathcal{N}_\theta^c} \sum_t \frac{\mathbb{1}_{t \notin \mathcal{T}_n}}{T - |\mathcal{T}_n|}[\mathbf{1}_{\bar{\mathcal{T}}}]_t + \sum_n \mathbb{1}_{\theta \in \mathcal{T}_n}[\mathbf{1}_{\bar{\mathcal{T}}}]_\theta\right) \\
&= \frac{1}{N}\left(\sum_{i \in \mathcal{N}_\theta^c} \frac{T - |\mathcal{T}_n|}{T - |\mathcal{T}_n|} 1 + \sum_{i \in \mathcal{N}_\theta} 1 \times [\mathbf{1}_{\bar{\mathcal{T}}}]_\theta\right) = \frac{|\mathcal{N}_\theta^c| + |\mathcal{N}_\theta|[\mathbf{1}_{\bar{\mathcal{T}}}]_\theta}{N} = [\mathbf{1}_{\bar{\mathcal{T}}}]_\theta \, .
\end{aligned}
$$

To bound the other eigenvalues of the matrix $(AB^\top)(BA^\top)$, we rely on spectral graph theory arguments. Consider the weighted graph $\mathcal{G} = (\mathcal{T}, \mathcal{E})$ whose vertices are the time periods $\mathcal{T}$ and each edge $(k, \ell) \in \mathcal{T} \times \mathcal{T}$ with $k \neq \ell$ has a weight given by $S_{k,\ell} = \frac{1}{N}\sum_n \frac{\mathbb{1}_{\{k,\ell\} \subset \mathcal{T}_n^c}}{T - |\mathcal{T}_n|}$ (if this quantity is zero, then there is no edge between k and ℓ).

The matrix $P \stackrel{\text{def}}{=} I_T - (AB^\top)(BA^\top)$ verifies for each $k \in \mathcal{T}$:

$$
\begin{aligned}
\sum_{\ell \neq k} -P_{k,\ell} &= \sum_{\ell \neq k} \tfrac{1}{N}\sum_n \frac{\mathbb{1}_{\{k,\ell\} \subset \mathcal{T}_n^c}}{T - |\mathcal{T}_n|} = \tfrac{1}{N}\sum_n \frac{\mathbb{1}_{k \in \mathcal{T}_n^c}(T - |\mathcal{T}_n| - 1)}{T - |\mathcal{T}_n|} \\
&= \tfrac{1}{N}\sum_n (1 - \mathbb{1}_{k \in \mathcal{T}_n}) - \tfrac{1}{N}\sum_n \frac{\mathbb{1}_{k \in \mathcal{T}_n^c}}{T - |\mathcal{T}_n|} = P_{kk} \, ,
\end{aligned}
$$

which shows that P is the Laplacian matrix of graph $\mathcal{G}$. As $\mathrm{Sp}(AB^\top BA^\top) = 1 - \mathrm{Sp}(P)$, we want to have a lower bound on the least eigenvalue of P greater than 0, that we denote by λ_1.

By rearranging the indices of $\mathcal{T}$ in two blocs $\bar{\mathcal{T}}$ and $\bar{\mathcal{T}}^c$, we observe that P can be written as a block diagonal matrix $P = \mathrm{diag}(P_{\bar{\mathcal{T}}}, 0_{\bar{\mathcal{T}}^c})$. As we are only interested in the positive eigenvalues of $\mathcal{P}$, we can therefore study the linear application associated to P restricted to the subspace $\mathrm{Vect}(e_t)_{t \in \mathcal{T}}$.

As $\mathbf{1}_{\bar{\mathcal{T}}}$ is an eigenvector of P associated to $\lambda_0 = 0$, from the minmax theorem, we have:

$$
\lambda_1 = \min_{u \perp \mathbf{1}_{\mathcal{T}}, u \neq 0} \frac{u^\top P u}{u^\top u} \, . \tag{1.24}
$$

Let us consider an eigenvector u realizing (1.24). Let $t^* \in \arg\max_t u_t$ and $s^* \in \arg\min_t u_t$ and let d_{s^*,t^*} be the distance between s^* and t^* in $\mathcal{G}$, and let (s^*-t^*) denote a shortest path

from s^* to t^* in $\mathcal{G}$. As P is a Laplacian matrix, we have:

$$\begin{aligned} u^\top P u &= \frac{1}{2} \sum_{k,\ell \in \mathcal{T}} -P_{k,\ell}(u_k - u_\ell)^2 \\ &\geqslant \frac{1}{2} \sum_{\{k,\ell\} \in (s^*\text{-}t^*)} -P_{k,\ell}(u_k - u_\ell)^2 \\ &\geqslant \min_{k,\ell \in (s^*\text{-}t^*)} (-P_{k,\ell}) \frac{(u_{t^*} - u_{s^*})^2}{d_{s^*,t^*}}, \end{aligned} \tag{1.25}$$

where the last inequality is obtained from Cauchy-Schwartz inequality.

Let us write the path $(s^*\text{-}t^*) = (t_0, t_1, \ldots, t_d)$. As $(s^*\text{-}t^*)$ is a shortest path, for each $k \in \{0, d-1\}$, the edge (t_k, t_{k+1}) exists so there exists $n \in \mathcal{N}$ such that $\{t_k, t_{k+1}\} \subset \mathcal{T}_n^c$. Moreover, for each n, we have $\mathcal{T}_n^c \cap \{t_0, \ldots, t_{k-1}, t_{k+2}, \ldots, t_d\} = \emptyset$, otherwise we could "shortcut" the path $(s^*\text{-}t^*)$, thus we have $|\mathcal{T}_n| \geqslant d-1$. We obtain:

$$-P_{t_k,t_{k+1}} = \frac{1}{N} \sum_n \frac{\mathbb{1}_{\{t_k,t_{k+1}\} \subset \mathcal{T}_n^c}}{T - |\mathcal{T}_n|} \geqslant \frac{1}{N(T-d+1)}.$$

On the other hand, we have $(u_{t^*} - u_{s^*}) \geqslant u_{t^*} + \frac{u_{t^*}}{T-1} = \frac{T}{T-1} u_{t^*} \geqslant \frac{T}{(T-1)\sqrt{T}} \|u\|_2$.

Using these bounds and (1.25), we obtain:

$$\begin{aligned} u^\top P u &\geqslant \frac{(u_{t^*} - u_{s^*})^2}{N(T - d_{s^*,t^*} + 1) d_{s^*,t^*}} \\ &\geqslant \frac{4T}{N(T+1)^2(T-1)^2} \|u\|_2^2 \geqslant \frac{4}{N(T+1)^2(T-1)} \|u\|_2^2. \end{aligned}$$

Therefore, $\lambda_1 \geqslant \frac{4}{N(T+1)^2(T-1)} \stackrel{\text{def}}{=} \kappa_{N,T}$ and the greatest singular value lower than one of $(AB^\top)(A^\top B)$ is $1 - \kappa_{N,T}$. We conclude by applying successively Lemma 1.4 and Lemma 1.3, to obtain the convergence rate stated in Theorem 1.4. □

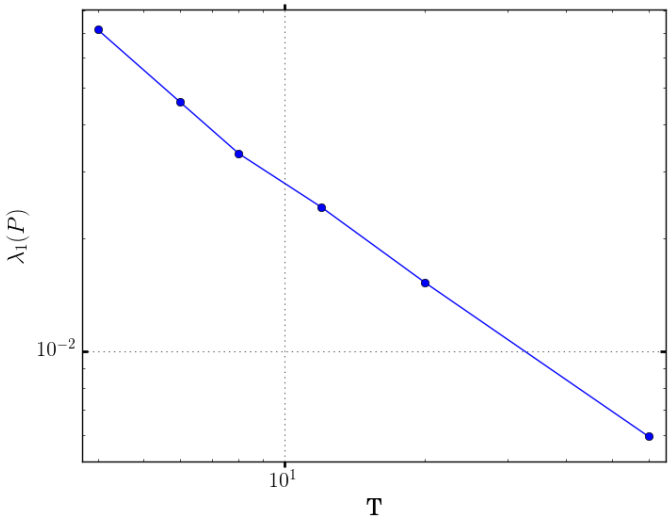


FIGURE 1.3: Evolution of the convergence rate, given as $\lambda_1(P)$ (lowest nonzero eigenvalue of P), with $N = 6$ and $T \in \{4, 6, 8, 12, 20, 60\}$. *The worst convergence rate is evaluated by taking* $100 \times T$ *random draws of the sets* $\mathcal{T}_n \subset \mathcal{T}$ *for each* n*, and evaluating the eigenvalue of the matrix. The slope is around -0.93, which indicates that in practice the convergence rate is* $\mathcal{O}(T^{-1})$*, faster than the upper bound in* $\mathcal{O}(T^{-3})$ *established in Theorem 1.4.*

1.4 Generalization to Polyhedral Agents Constraints

In this section, we extend our results to a more general framework where for each $n \in \mathcal{N}$, $\mathcal{X}_n$ is an arbitrary polyhedron, instead of having the particular structure given in (1.5). Let us now consider that $(\mathcal{X}_n)_n$ are polyhedra with, for each n:

$$\mathcal{X}_n = \{\boldsymbol{x}_n \in \mathbb{R}^T | A_n \boldsymbol{x}_n \leqslant \boldsymbol{b}_n\}\,, \tag{1.26}$$

with $A_n \in \mathcal{M}_{T,k_n}(\mathbb{R})$ with $k_n \in \mathbb{N}$. The disaggregation problem (1.2), *with* $\boldsymbol{p} \in \mathcal{P}$ *fixed*, writes:

$$\min_{\boldsymbol{x} \in \mathbb{R}^{NT}} 0 \tag{1.27a}$$

$$\text{s.t. } A_0 \boldsymbol{x} = B\boldsymbol{p} \quad (\boldsymbol{\lambda}_0) \tag{1.27b}$$

$$A_n \boldsymbol{x}_n \leqslant \boldsymbol{b}_n,\ \forall n \in \mathcal{N} \quad (\boldsymbol{\lambda}_n)\,. \tag{1.27c}$$

where $A_0 = J_{1,N} \otimes I_T$, $B = I_T$, (such that (1.27b) corresponds to the aggregation constraint $\sum_n \boldsymbol{x}_n = \boldsymbol{p}$) and $\boldsymbol{\lambda}_0 \in \mathbb{R}_+^T$, $(\boldsymbol{\lambda}_n)_{n\in\mathcal{N}} \in \mathbb{R}^{\sum_n k_n}$ are the Lagrangian multipliers associated to (1.27b) and (1.27c).

With the polyhedral constraints (1.26), the graph representation of the disaggregation problem, as illustrated in Figure 1.1 is no longer valid. Consequently, one can not directly apply Hoffman's theorem (Theorem 1.1) to obtain a characterization of disaggregation feasibility by inequalities on $\boldsymbol{p}$. However, using duality theory, Proposition 1.10 below also gives a characterization of disaggregation:

Proposition 1.10. *A profile $\boldsymbol{p} \in \mathcal{P}$ is disaggregeable iff:*

$$\forall (\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, \dots \boldsymbol{\lambda}_N) \in \Lambda,\ \ \boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \textstyle\sum_{n\in\mathcal{N}} \boldsymbol{\lambda}_n^\top \boldsymbol{b}_n \geqslant 0\,, \tag{1.28}$$

where $\Lambda \stackrel{\text{def}}{=} \{\boldsymbol{\lambda}_0 \in \mathbb{R}^{k_0}, \forall n \in \mathcal{N}, \boldsymbol{\lambda}_n \in \mathbb{R}_+^{k_n} \mid A_0^\top \boldsymbol{\lambda}_0 + A^\top (\boldsymbol{\lambda}_n)_n = 0\}$, *with* $A \stackrel{\text{def}}{=} diag(A_n)_{n\in\mathcal{N}}$.

Proof. From strong duality, we have:

$$\min_{\boldsymbol{x}\in\mathbb{R}^{NT}} \max_{\boldsymbol{\lambda}_0\in\mathbb{R}^{k_0}, \boldsymbol{\lambda}_n \in \mathbb{R}_+^{k_n}} \boldsymbol{\lambda}_0^\top (A_0\boldsymbol{x} - B\boldsymbol{p}) + \sum_n \boldsymbol{\lambda}_n^\top (A_n \boldsymbol{x}_n - \boldsymbol{b}_n) \tag{1.29}$$

$$= \max_{\substack{\boldsymbol{\lambda}_0\in\mathbb{R}^{k_0}, \boldsymbol{\lambda}_n \in \mathbb{R}_+^{k_n} \\ \text{s.t. } \boldsymbol{\lambda}_0^\top A_0 + (\boldsymbol{\lambda}_n)_n^\top A = 0}} -\boldsymbol{\lambda}_0^\top B\boldsymbol{p} - \sum_n \boldsymbol{\lambda}_n^\top \boldsymbol{b}_n\,. \tag{1.30}$$

If the polytope $\mathcal{Y}_{\boldsymbol{p}} \cap \mathcal{X}$ given by the constraints of (1.27) is empty, then there is an infeasibility certificate $\boldsymbol{\lambda}^\top = (\boldsymbol{\lambda}_0^\top\ \boldsymbol{\lambda}_1^\top \dots \boldsymbol{\lambda}_N^\top) \in \mathbb{R}^T \times \prod_n \mathbb{R}_+^{k_n}$ such that:

$$\boldsymbol{\lambda}_0^\top A_0 + (\boldsymbol{\lambda}_n^\top)_n A = 0 \tag{1.31a}$$

$$\boldsymbol{\lambda}_0^\top B\boldsymbol{p} + (\boldsymbol{\lambda}_n)_n^\top \boldsymbol{b} < 0\,. \tag{1.31b}$$

On the other hand, if $\mathcal{Y}_{\boldsymbol{p}} \cap \mathcal{X}$ is nonempty, then a solution to the dual problem (1.30) is bounded, which implies that $\forall \boldsymbol{\lambda} \stackrel{\text{def}}{=} (\boldsymbol{\lambda}_0, (\boldsymbol{\lambda}_n)_n) \in \Lambda,\ \ \boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \sum_n \boldsymbol{\lambda}_n^\top \boldsymbol{b}_n \geqslant 0$.

□

As opposed to Hoffman circulation's theorem where disaggregation is characterized by a finite number of inequalities, Proposition 1.10 involves a priori an infinite number of inequalities.

However, we know that the polyhedral cone Λ can be represented by a finite number of generators (edges), that is, there exists $\Lambda^* \stackrel{\text{def}}{=} \{\boldsymbol{\lambda}^{*(1)}, \dots, \boldsymbol{\lambda}^{*(d)}\}$ such that:

$$\Lambda = \left\{ \textstyle\sum_{1\leqslant i\leqslant d} \alpha_i \boldsymbol{\lambda}^{*(i)} \mid (\alpha_i)_i \in \mathbb{R}_+^d \right\}. \tag{1.32}$$

Thus, we obtain the following corollary to Proposition 1.10:

Corollary 1.1. *There exists a finite set $\Lambda^* \subset \Lambda$ such that, for any $\boldsymbol{p} \in \mathcal{P}$, $\boldsymbol{p}$ is disaggregeable iff:*

$$\forall (\boldsymbol{\lambda}_0, (\boldsymbol{\lambda}_n)_n) \in \Lambda^*, \;\; \boldsymbol{\lambda}_0 B\boldsymbol{p} + \textstyle\sum_n \boldsymbol{\lambda}_n \boldsymbol{b}_n \geqslant 0\,. \tag{1.33}$$

Remark 1.4. *In the transportation case* (1.5)*, we can write each agent constraints in the form $A_n \boldsymbol{x}_n \leqslant \boldsymbol{b}_n$ (writing the equality $\sum_t x_{n,t} = E_n$ is written as two inequalities), and Hoffman conditions* (1.8) *can be written in the form* (1.31)*. Moreover, Theorem 1.3 ensures that one possibility for Λ^* of Corollary 1.1 is to consider the collection of 2^T multipliers corresponding to the subsets $\mathcal{T}_0 \subset \mathcal{T}$ and $\mathcal{N}_0$ minimizing* (1.9)*. We skip the details here for brevity.*

As in the first part of the chapter, we want to use APM to decompose problem (1.1) and, in the case where disaggregation is not possible, use the result of APM to generate an inequality (1.28) violated by the current profile $\boldsymbol{p}$.

In the case of impossible disaggregation, the APM converges to the orbit $(\boldsymbol{y}^\infty, \boldsymbol{x}^\infty)$ and if we define $\boldsymbol{\mu} \stackrel{\text{def}}{=} \boldsymbol{y}^\infty - \boldsymbol{x}^\infty$, it defines a separating hyperplan $\bar{\boldsymbol{x}} + \boldsymbol{\mu}^\perp$, where $\bar{\boldsymbol{x}} = \frac{\boldsymbol{y}^\infty + \boldsymbol{x}^\infty}{2}$, that satisfies, with $a \stackrel{\text{def}}{=} \bar{\boldsymbol{x}}.\boldsymbol{\mu}$ (note that $\bar{\boldsymbol{x}}$ can be replaced by any $\boldsymbol{y} \in [\boldsymbol{y}^\infty, \boldsymbol{x}^\infty]$):

$$\forall \boldsymbol{x} \in \mathcal{Y}_{\boldsymbol{p}};\; \boldsymbol{\mu}^\top \boldsymbol{x} > a \qquad\qquad \forall \boldsymbol{x} \in \mathcal{X};\; a > \boldsymbol{\mu}^\top \boldsymbol{x},$$

which give lower bounds on the following linear problems (problem (1.35) is decomposed because A is a block-diagonal matrix, but it can also be written in one single problem):

$$\begin{array}{ll} \min\limits_{\boldsymbol{x} \in \mathbb{R}^{NT}} \boldsymbol{\mu}^\top \boldsymbol{x} & \max\limits_{\boldsymbol{\lambda}_0 \in \mathbb{R}^{k_0}} -\boldsymbol{\lambda}_0^\top B\boldsymbol{p} \\ A_0 \boldsymbol{x} = B\boldsymbol{p} \;(\boldsymbol{\lambda}_0) & \boldsymbol{\mu} = -A_0^\top \boldsymbol{\lambda}_0 \end{array} \tag{1.34}$$

$$\text{and} \quad \forall n \in \mathcal{N}, \quad \begin{array}{ll} \max\limits_{\boldsymbol{x} \in \mathbb{R}^{NT}} \boldsymbol{\mu}_n \boldsymbol{x}_n & \min_{\boldsymbol{\lambda}_n \in \mathbb{R}^{k_n}_+} \boldsymbol{b}_n^\top \boldsymbol{\lambda}_n \\ A_n \boldsymbol{x}_n \leqslant \boldsymbol{b}_n \;(\boldsymbol{\lambda}_n) & \boldsymbol{\mu}_n = \boldsymbol{\lambda}_n^\top A_n \end{array}\,. \tag{1.35}$$

Strong duality on these problems implies that there exist $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}$ such that:

$$\boldsymbol{\mu} = -A_0^\top \boldsymbol{\lambda}_0 \text{ and } a < -\boldsymbol{\lambda}_0^\top B\boldsymbol{p} \qquad\qquad \boldsymbol{\mu} = \boldsymbol{\lambda}^\top A \text{ and } a > \boldsymbol{b}^\top \boldsymbol{\lambda}\,. \tag{1.36}$$

Thus, we obtain $(\boldsymbol{\lambda}_0, \boldsymbol{\lambda}) \in \Lambda$ satisfying (1.31), that is, $\boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \boldsymbol{b}^\top \boldsymbol{\lambda} < 0$, and we can use this to add a new valid additional inequality on $\boldsymbol{p}$ of form (1.28), that will change the current profile $\boldsymbol{p}$:

$$\boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \boldsymbol{\lambda}^\top \boldsymbol{b} \geqslant 0\,. \tag{1.37}$$

In Algorithm 1.5 below, we summarize the proposed decomposition of problem (1.1). This is a generalization of the decomposition principle used for Algorithm 1.4.

Algorithm 1.5 Non-intrusive optimal disaggregation with polyhedral constraints

Require: Start with $\Lambda^{(0)} = \{\}$, $k = 0$, DISAG= *false*
1: **while** not DISAG **do**
2: get solution $\boldsymbol{p}^{(k)}$ of problem $\min_{\boldsymbol{p} \in \mathcal{P}} \{f(\boldsymbol{p}) \mid \boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \boldsymbol{\lambda}^\top \boldsymbol{b} \geqslant 0,\; \forall \boldsymbol{\lambda} \in \Lambda^{(k)}\}$
3: get $\boldsymbol{\mu}^{(k)} = \boldsymbol{y}^\infty - \boldsymbol{x}^\infty \leftarrow \text{APM}(\mathcal{Y}_{\boldsymbol{p}^{(k)}}, \mathcal{X})$
4: **if** $\boldsymbol{\mu}^{(k)} \neq 0$ **then**
5: obtain $\boldsymbol{\lambda}_0^{(k)} \leftarrow \max_{\boldsymbol{\lambda}_0 \in \mathbb{R}^{k_0}} \{-\boldsymbol{\lambda}_0^\top B\boldsymbol{p}^{(k)} \mid \boldsymbol{\mu}^{(k)} = -A_0^\top \boldsymbol{\lambda}_0\}$
6: obtain for each n, $\boldsymbol{\lambda}_n^{(k)} \leftarrow \max_{\boldsymbol{\lambda}_n \geqslant 0} \{\boldsymbol{b}_n^\top \boldsymbol{\lambda}_n \mid \boldsymbol{\mu}_n^{(k)} = \boldsymbol{\lambda}_n^\top A_n\}$
7: add $\Lambda^{(k+1)} = \Lambda^{(k)} \cup \{(\boldsymbol{\lambda}_0^{(k)}, \boldsymbol{\lambda}^{(k)})\}$
8: **else**
9: Return DISAG $=$ *true*, $\boldsymbol{p}^{(k)}$ as optimal solution
10: **end**
11: $k \leftarrow k + 1$
12: **done**

Remark 1.5. *We use the fact that $\boldsymbol{\mu} = \boldsymbol{y}^\infty - \boldsymbol{x}^\infty$ although, as before, we only have an* approximation *of this quantity. The approximation has to be precise enough to ensure that the solution obtained verifies $\boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \boldsymbol{b}^\top\boldsymbol{\lambda} < 0$. In practice, one can proceed as in the transportation case and Algorithm 1.3: start with a large ε_{cvg} as stopping criterion in APM, then compute $(\boldsymbol{\lambda}_0, \boldsymbol{\lambda}) \in \Lambda$ and check if $\boldsymbol{\lambda}_0^\top B\boldsymbol{p} + \boldsymbol{b}^\top\boldsymbol{\lambda} < 0$. If this is not the case, restart with $\varepsilon_{\text{cvg}} = \varepsilon_{\text{cvg}}/2$.*

Remark 1.6. *When $\mathcal{Y}_{\boldsymbol{p}} = \{\boldsymbol{x} \in \mathbb{R}^{NT} | A_0\boldsymbol{x} = B_0\boldsymbol{p}\} = \{\boldsymbol{x} | \sum_n \boldsymbol{x}_n = \boldsymbol{p}\}$, we can obtain a non-intrusive version of APM on $\mathcal{Y}_{\boldsymbol{p}}$ and $\mathcal{X}$, similar to Algorithm 1.3. In this case,* (1.36) *ensures that we have $\boldsymbol{\mu}_{n,t} = -[\boldsymbol{\lambda}_0]_t$ for each $n \in \mathcal{N}$, and $\boldsymbol{\lambda}_0$ is fixed by $\boldsymbol{\mu}$). The only difference with the transportation case for a non-intrusive APM in the general polyhedral case, is in the way of computing the valid constraint violated by $\boldsymbol{p}$. Thus, Lines 16 to 19 of Algorithm 1.3 have to be replaced by Algorithm 1.6.*

Algorithm 1.6 Modification of Lines 16-19 of Algorithm 1.3 for NI-APM with polyhedral constraints

16: **for** each agent $n \in \mathcal{N}$ **do**
17: compute M_n optimal value of (1.35).
18: **done**
19: Operator computes $M \leftarrow \text{SMCA}((M_n)_n)$
20: **if** $-\boldsymbol{\nu}.\boldsymbol{p} + M < 0$ **then**
21: **return** DISAG $\leftarrow$ FALSE, $-\boldsymbol{\nu}, M$

1.4.1 Link with Benders' decomposition

In this generalized case, we obtain an algorithm related to Benders' decomposition [Ben62] (recall that in our specific case (1.27), the cost function does not involve the variable $\boldsymbol{x}$ but only variable $\boldsymbol{p}$).

The difference between the proposed Algorithm 1.5 and Benders' decomposition is on the way of generating the new cut. Benders' decomposition would directly solve the dual problem (1.30): $\max_{\boldsymbol{\lambda}}\{-\boldsymbol{\lambda}_0^\top B_0\boldsymbol{p} - (\boldsymbol{\lambda}_n)_n^\top\boldsymbol{b} \mid \boldsymbol{\lambda}_0 A_0 + (\boldsymbol{\lambda}_n)_n A = 0\}$ and obtain a cut if it is unbounded. However, this problem involves the constraints of all users (through A and $\boldsymbol{b}$), and it is not straightforward to obtain a method to solve this problem in a decentralized and efficient way.

1.5 Numerical examples

1.5.1 An illustrative example with T=4

In this section we illustrate the iterations of the method proposed in this chapter on an example with $T = 4$ and $N = 3$. Assuming that we have to satisfy the aggregate constraint $\sum_t p_t = \sum_n E_n$, we can use the projections on this affine space of solutions of master problems $(\boldsymbol{p}^{(s)})_s$ to visualize them in dimension 3.

One can wonder if, in the transportation case, applying Algorithm 1.4 or Algorithm 1.5 will always lead to the same cuts and solutions: the answer is no, as shown by the instance considered in this section, for which Algorithm 1.4 converges in 3 iterations and Algorithm 1.5 needs 4 iterations.

We consider the problem (1.1) with agents constraints (1.5) with parameters $\underline{\boldsymbol{x}} \stackrel{\text{def}}{=} 0$ and:

$$\overline{\boldsymbol{x}} \stackrel{\text{def}}{=} \begin{matrix}[0.8, 0.2, 0.7, 0.1] \\ [0.5, 0.1, 0.3, 0.6] \\ [0.1, 0.1, 0.7, 0.2]\end{matrix} \begin{matrix} E_1 = 1.8 \\ E_2 = 0.4\,, \\ E_3 = 1.1 \end{matrix} \quad \forall \boldsymbol{p} \in \mathbb{R}^4, f(\boldsymbol{p}) \stackrel{\text{def}}{=} \sum_{1 \leqslant t \leqslant 4} 0.8 \times p_t + 0.1 \times p_t^2\,. \tag{1.38}$$

Considering the aggregate equality constraint $\sum_{1 \leqslant t \leqslant 4} p_t = \sum_{1 \leqslant n \leqslant 3} E_n = 3.3$, we use the canonical projection of 4 dimensional vectors into the 3 dimensional space (p_1, p_2, p_3) to

visualize the cuts and solutions. We first derive on this example the $2^T - 2 = 14$ Hoffman inequalities characterizing disaggregation from Theorem 1.1:

$$\begin{array}{llll} p_1 \leqslant 1.3 & p_2 \leqslant 0.4 & p_3 \leqslant 1.7 & p_4 \leqslant 0.7 \\ p_1 + p_2 \leqslant 1.6 & p_1 + p_3 \leqslant 2.7 & p_1 + p_4 \leqslant 1.6 & p_2 + p_3 \leqslant 2.1 \\ p_2 + p_4 \leqslant 1.0 & p_3 + p_4 \leqslant 2.1 & p_1 + p_2 + p_3 \leqslant 3.0 & p_1 + p_2 + p_4 \leqslant 1.9 \\ p_2 + p_3 + p_4 \leqslant 2.4 & p_1 + p_3 + p_4 \leqslant 3.0 & & \end{array}$$

The projection of the obtained polytope $\mathcal{P}_{\mathrm{D}}$, as defined in (1.4), is represented in Figure 1.4a. One can remark that this polytope has only 6 facets. Our Algorithm 1.4 applied on this instance with $\varepsilon_{\mathrm{dis}} = 10^{-3}$ and $\varepsilon_{\mathrm{cvg}} = 10^{-5}$ converges in 3 iterations, with successive solutions of the master problem (1.11) and cuts added:

$$\begin{array}{lll} \boldsymbol{p}^{(1)} = [1., 0.4, 1., 0.9] & \xrightarrow{\text{cut}} & p_1 + p_2 + p_4 \leqslant 1.9 \\ \boldsymbol{p}^{(2)} = [0.75, 0.4, 1.4, 0.75] & \xrightarrow{\text{cut}} & p_2 + p_3 + p_4 \leqslant 2.4 \\ \boldsymbol{p}^{(3)} = [0.9, 0.4, 1.4, 0.6]\ . & & \end{array}$$

Figure 1.4b represents in the projection space the three successive solutions and the two generated cuts (in red for each iteration).

On the other hand, applying Algorithm 1.5 with the same parameters $(\varepsilon_{\mathrm{dis}}, \varepsilon_{\mathrm{cvg}})$, there are 3 cuts generated and 4 resolutions of master problem needed for convergence, given by (we refer the reader to Remark 1.5 for the numerical precision obtained in the values):

$$\begin{array}{lll} \hat{\boldsymbol{p}}^{(1)} = [1., 0.4, 1., 0.9] & & \\ & \xrightarrow{\text{cut}} & -0.25p_1 - 0.25p_2 + 1.0p_3 - 0.5p_4 \geqslant 0.75 \\ \hat{\boldsymbol{p}}^{(2)} = [0.8097, 0.4, 1.3984, 0.6919] & & \\ & \xrightarrow{\text{cut}} & 1.0p_1 - 0.509p_2 + 0.018p_3 - 0.509p_4 \geqslant 0.4161 \\ \hat{\boldsymbol{p}}^{(3)} = [0.9062, 0.4, 1.3823, 0.6115] & & \\ & \xrightarrow{\text{cut}} & -0.333p_1 - 0.333p_2 + 1.0p_3 - 0.333p_4 \geqslant 0.7666 \\ \hat{\boldsymbol{p}}^{(4)} = [0.9, 0.4, 1.4, 0.6]\ . & & \end{array}$$

Owing to the strict convexity of the cost function $\boldsymbol{p} \mapsto f(\boldsymbol{p})$, the global problem (1.1) has a unique minimizer $\boldsymbol{p}$, and Algorithms 1.4 and 1.5 both lead to this solution, albeit with a different number of cuts. The 4 successive solutions and the 3 added cuts are represented in the three dimensional space on Figure 1.4c: we observe that the last cut needed to obtain the convergence of Algorithm 1.5, corresponds to the first one added with Algorithm 1.4.

1.5.2 A nonconvex example: management of a microgrid

In this section, we illustrate the proposed method on a larger scale practical example from energy. We consider an electricity microgrid [Kat+08] composed of N electricity consumers with flexible appliances (such as electric vehicles or water heaters), a photovoltaic (PV) power plant and a conventional generator. The operator of the microgrid aims at satisfying the demand constraints of consumers over a set of time periods $\mathcal{T} = \{1, \dots, T\}$, while minimizing the energy cost for the community. We have the following characteristics:

- the PV plant generates a nondispatchable power profile $(p_t^{\mathrm{PV}})_{t \in \mathcal{T}}$ at marginal cost zero;
- the conventional generator has a starting cost C^{ST}, minimal and maximal power production $\underline{p}^g, \overline{p}^g$, and piecewise-linear and continuous generation cost function $p^g \mapsto f(p^g)$:
$$f(p^g) = \alpha_k + c_k p^g, \text{ if } p^g \in \mathcal{I}_k \stackrel{\text{def}}{=} [\theta_{k-1}, \theta_k[,\ k = 1 \dots K,$$
where $\theta_0 \stackrel{\text{def}}{=} 0$ and $\theta_K \stackrel{\text{def}}{=} \overline{p}^g$;
- each agent $n \in \mathcal{N}$ has some flexible appliances which require a global energy demand E_n on $\mathcal{T}$, and has consumption constraints on the total household consumption, on each time

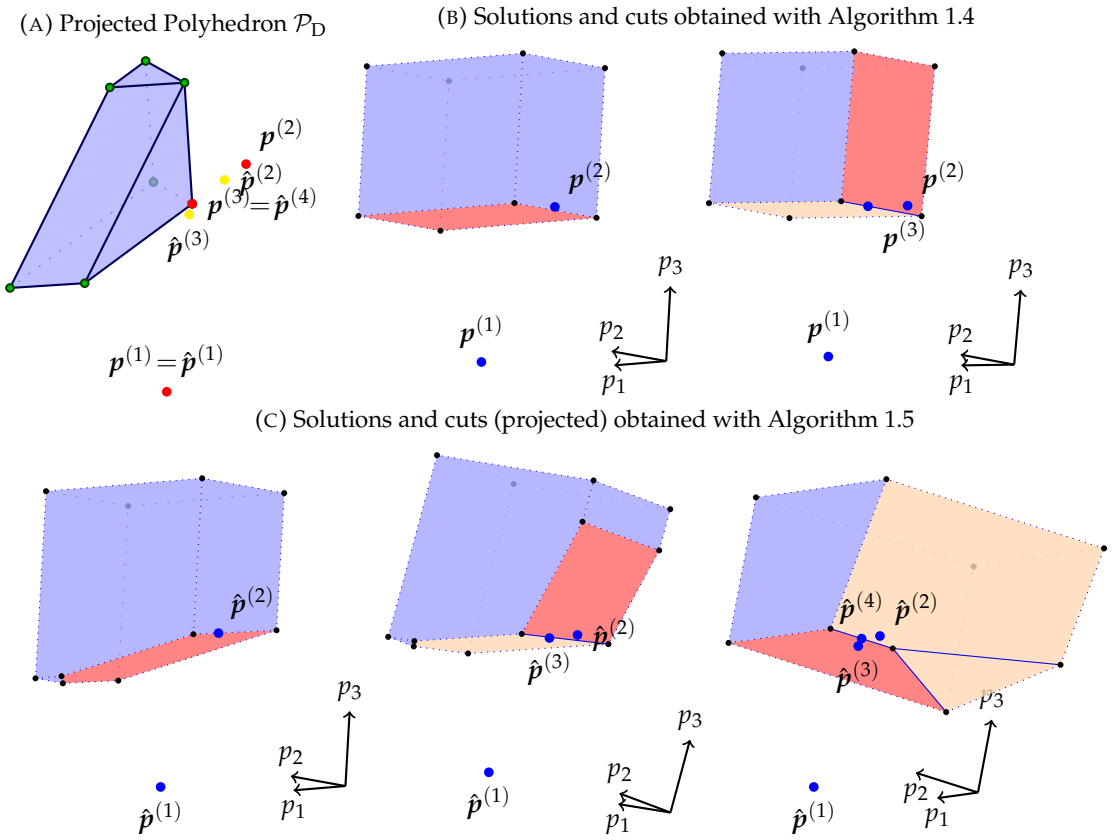


FIGURE 1.4: Illustration of the iterations of the proposed decomposition method. *The cut $p_3 \geqslant 1.4$, which is added at first for Algorithm 1.4 , is only added at the third iteration of Algorithm 1.5*

period $t \in \mathcal{T}$, that are formulated with $\underline{x}_n, \overline{x}_n$. These parameters are confidential because they could for instance contain some information on agent n habits.

The master problem (1.11) can be written as the following MILP (1.39):

$$\min_{\boldsymbol{p},\boldsymbol{p}^g,(\boldsymbol{p}_k^g),(\boldsymbol{b}_k),\boldsymbol{b}^{\mathrm{ON}},\boldsymbol{b}^{\mathrm{ST}}} \sum_{t\in\mathcal{T}} \left(\alpha_1 b_t^{\mathrm{ON}} + \sum_k c_k p_{k,t}^g + C^{\mathrm{ST}} b_t^{\mathrm{ST}}\right) \tag{1.39a}$$

$$p_t^g = \textstyle\sum_{k=1}^K p_{k,t}^g \,,\ \forall t \in \mathcal{T} \tag{1.39b}$$

$$b_{k,t}(\theta_k - \theta_{k-1}) \leqslant p_{k,t}^g \leqslant b_{k-1,t}(\theta_k - \theta_{k-1}),\ \forall 1 \leqslant k \leqslant K,\ \forall t \in \mathcal{T} \tag{1.39c}$$

$$b_t^{\mathrm{ST}} \geqslant b_t^{\mathrm{ON}} - b_{t-1}^{\mathrm{ON}},\ \forall t \in \{2,\dots,T\} \tag{1.39d}$$

$$\underline{p}^g b_t^{\mathrm{ON}} \leqslant p_t^g \leqslant \overline{p}^g b_t^{\mathrm{ON}},\ \forall t \in \mathcal{T} \tag{1.39e}$$

$$b_t^{\mathrm{ON}}, b_t^{\mathrm{ST}}, b_{1,t}, \dots, b_{K-1,t} \in \{0,1\},\ \forall t \in \mathcal{T} \tag{1.39f}$$

$$\boldsymbol{p} \leqslant \boldsymbol{p}^{\mathrm{PV}} + \boldsymbol{p}^g \tag{1.39g}$$

$$\boldsymbol{p}^\top \mathbb{1}_T = \boldsymbol{E}^\top \mathbb{1}_N \tag{1.39h}$$

$$\underline{\boldsymbol{x}}^\top \mathbb{1}_N \leqslant \boldsymbol{p} \leqslant \overline{\boldsymbol{x}}^\top \mathbb{1}_N \,. \tag{1.39i}$$

In this formulation (1.39b-1.39c), where $b_{0,t} \overset{\text{def}}{=} 1$ and $b_{K,t} \overset{\text{def}}{=} 0$, are a mixed integer formulation of the generation cost function f. One can show that the Boolean variable $b_{k,t}$ is equal to one *iff* $p_t^g \geqslant \theta_k$ for each $k \in \{1,\dots,K-1\}$. Note that only α_1 appears in (1.39a) because of the continuity of f.

Constraints (1.39d-1.39e) ensure the on/off and starting constraints of the power plant, (1.39g) ensures that the power allocated to consumption is not above the total production, and (1.39h-1.39i) are the aggregated feasibility conditions already referred to in (1.6). The

nonconvexity of (1.39) comes from the existence of starting costs and constraints of minimal power, which makes necessary to use Boolean state variables $b^{\mathrm{ST}}, b^{\mathrm{ON}}$.

We simulate the problem described above for different values of $N \in \{2^4, 2^5, 2^6, 2^7, 2^8\}$ and one hundred instances with random parameters for each value of N. A scaling factor $\kappa_N = N/20$ is applied on parameters to ensure that production capacity is large enough to meet consumers demand. The parameters are chosen as follows:

- $T = 24$ (hours of a day);
- production costs: $K = 3$, $\theta = [0, 70, 100, 300]\kappa_N$, $\boldsymbol{c} = [0.2, 0.4, 0.5]$, $\underline{p}^g = 50\kappa_N$, $\overline{p}^g = 300\kappa_N$, $\alpha_1 = 4$ and $C^{\mathrm{ST}} = 15$;
- photovoltaic: $p_t^{\mathrm{PV}} = \left[50(1-\cos(\frac{(t-6)2\pi}{16}) + \mathcal{U}([0,10])\right]\kappa_N$ for $t \in \{6, \dots, 20\}$, $p_t^{\mathrm{PV}} = 0$ otherwise;
- consumption parameters are drawn randomly with: $\underline{x}_{n,t} \sim \mathcal{U}([0,10])$, $\overline{x}_{n,t} \sim \mathcal{U}([0,5]) + \underline{x}_{n,t}$ and $E_n \sim \mathcal{U}([\mathbb{1}_T^\top \underline{\boldsymbol{x}}_n, \mathbb{1}_T^\top \overline{\boldsymbol{x}}_n])$, so that individual feasibility ($\mathcal{X}_n \neq \emptyset$) is ensured.

$N =$	2^4	2^5	2^6	2^7	2^8
# master	193.6	194.1	225.5	210.9	194.0
# projs.	9507	15367	24319	26538	26646

TABLE 1.1: number of subproblems solved (average on 100 instances)

We implement Algorithm 1.4 using Python 3.5. The MILP (1.39) is solved using Cplex Studio 12.6 and Pyomo interface. Simulations are run on a single core of a cluster at 3GHz. For the convergence criteria (see Lines 11 and 12 of Algorithm 1.3), we use $\varepsilon_{\mathrm{dis}} = 0.01$ with the operator norm defined by $|||\boldsymbol{x}||| = \max_{n\in\mathcal{N}} \sum_t |x_{n,t}|$ (to avoid the $\sqrt{N}$ factor in the convergence criteria appearing with $\|.\|_2$), and starts with $\varepsilon_{\mathrm{cvg}} = 0.1$. The largest instances took around 10 minutes to be solved in this configuration and without parallel implementation. As the CPU time needed depends on the cluster load, it is not a reliable indicator of the influence on N on the complexity of the problems. Moreover, one advantage of the proposed method is that the projections in APM can be computed locally by each agent in parallel, which could not be implemented here for practical reasons.

Table 1.1 gives the number of master problems solved and the total number of projections computed, on average over the hundred instance for each value of N.

One observes that the number of master problems (1.39) solved (or equivalently, the number of "cuts" added), remains almost constant when N increases. In all instances, this number is way below the upper bound of $2^{24} > 1,6 \times 10^7$ possible constraints (see Proposition 1.8), which suggests that only a limited number of constraints are added in practice. The average total number of projections computed for each instance (total number of iterations of the **while** loop of Algorithm 1.3, Line 1 over all calls of APM in the instance) increases in a sublinear way which is even better that one could expect from the upper bound given in Theorem 1.4.

1.6 Conclusion

We provided a non-intrusive algorithm that enables to compute an optimal resource allocation, solution of a–possibly nonconvex–optimization problem, and to affect to each agent an individual profile satisfying a global demand and lower and upper bounds constraints. Our method uses local projections and works in a distributed fashion. Hence, the resolution of the problem is still efficient even in the case of a very large number of agents. The method is also privacy-preserving, as agents do not need to reveal any information on their constraints or their individual profile to a third party.

Several extensions and generalizations can be considered for this work.
Section 1.4 generalizes the procedure to arbitrary polyhedral constraints for agents. However, the number of constraints (cuts) added to the master problem is not proved to be finite

as done in the transportation case. Proving that only a finite number of constraints can be added (maybe up to a refinement procedure of the current constraint obtained) will enable to have a termination result for the algorithm in the general polyhedral case. In the transportation case, we showed the geometric convergence of APM with a rate linear in the number of agents. Moreover, the number of cuts added in the procedure is finite but the upper bound that we have remains exponential. In practice however, the number of constraints to consider remains small, as seen in Section 1.5. A thinner upper bound on the number of cuts added in the algorithm in this case would constitute an interesting result.

Appendix

1.A Proof of Proposition 1.4

Proof of Item (i). Let us write the stationarity conditions associated to problem (1.13):

$$\forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad 0 = (x_{n,t} - y_{n,t}) - \lambda_n - \underline{\mu}_{n,t} + \overline{\mu}_{n,t} \quad \text{and} \quad y_{n,t} = x_{n,t} + \nu_t \,. \tag{1.40}$$

By summing the preceding equalities on $\mathcal{T}$ and $\mathcal{N}$, we obtain the three equalities:

$$\sum_t \nu_t = \sum_t y_{n,t} - E_n, \ \forall n \in \mathcal{N} \qquad p_t = \sum_n x_{n,t} + N\nu_t, \ \forall t \in \mathcal{T} \tag{1.41}$$

$$|\mathcal{T}_n^\circ| \lambda_n = E_n - \sum_{t \in \underline{\mathcal{T}}_n} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} y_{n,t} - \sum_{t \in \overline{\mathcal{T}}_n} \overline{x}_{n,t}, \forall n \in \mathcal{N} \,, \tag{1.42}$$

where we define for each $n \in \mathcal{N}$:

$$\mathcal{T}_n^\circ \stackrel{\text{def}}{=} \{t \mid \underline{x}_{n,t} < x_{n,t} < \overline{x}_{n,t}\}, \quad \underline{\mathcal{T}}_n = \{t \mid x_{n,t} = \underline{x}_{n,t}\} \text{ and } \overline{\mathcal{T}}_n = \{t \mid x_{n,t} = \overline{x}_{n,t}\} \,.$$

From (1.41) and the aggregate equality $\sum_n E_n = \sum_t p_t$, we obtain $\sum_t \nu_t = 0$ and:

$$\forall n \in \mathcal{N}, \ \textstyle\sum_{t \in \mathcal{T}} y_{n,t} = E_n. \tag{1.43}$$

Suppose that Item (i) is false: there exists $n \notin \mathcal{N}_0$ and $\hat{t} \in \mathcal{T}_0$ such that $x_{n,\hat{t}} < \overline{x}_{n,\hat{t}}$. We have:

$$x_{n,\hat{t}} \geqslant y_{n,\hat{t}} + \lambda_n = x_{n,\hat{t}} + \nu_{\hat{t}} + \lambda_n \implies \lambda_n < 0 \,.$$

Immediatly, we have $\overline{\mathcal{T}}_n \subset \mathcal{T}_0$: indeed, for $t \in \overline{\mathcal{T}}_n$, we have:

$$y_{n,t} + \lambda_n \geqslant \overline{x}_{n,t} \implies y_{n,t} \geqslant \overline{x}_{n,t} - \lambda_n > \overline{x}_{n,t} \implies t \in \mathcal{T}_0 \,.$$

From the condition (1.43) and from $\nu_t > 0$ for each $t \in \overline{\mathcal{T}}_n$ because $\overline{\mathcal{T}}_n \subset \mathcal{T}_0$, we get:

$$\begin{aligned} 0 = \sum_{t \in \mathcal{T}} (y_{n,t} - x_{n,t}) &= \sum_{t \in \underline{\mathcal{T}}_n} (y_{n,t} - \underline{x}_{n,t}) + \sum_{t \in \mathcal{T}_n^\circ} (-\lambda_n) + \sum_{t \in \overline{\mathcal{T}}_n} \nu_t \\ \iff \quad \sum_{t \in \underline{\mathcal{T}}_n} (y_{n,t} - \underline{x}_{n,t}) &= \sum_{t \in \mathcal{T}_n^\circ} \lambda_n - \sum_{t \in \overline{\mathcal{T}}_n} \nu_t, \end{aligned}$$

which is strictly negative: this implies that there exists $t' \in \underline{\mathcal{T}}_n$ such that $y_{n,t'} < \underline{x}_{n,t'}$. Necessarily, $t' \notin \mathcal{T}_0$ because $\nu_{t'} = y_{n,t'} - x_{n,t'} < \underline{x}_{n,t'} - \underline{x}_{n,t'} = 0$. Then, as we have $\sum_{m \in \mathcal{N}} y_{m,t'} = p_{t'} \geqslant \sum_m \underline{x}_{m,t'}$, there exists $m \in \mathcal{N}$ such that $y_{m,t'} > \underline{x}_{m,t'}$. If $\lambda_m \leqslant 0$, and as $x_{m,t'} = y_{m,t'} - \nu_{t'} > \underline{x}_{m,t'}$, we get:

$$x_{m,t'} = \min(\overline{x}_{m,t'}, y_{m,t'} + \lambda_m) \leqslant y_{m,t'} + \lambda_m \leqslant y_{m,t'} = x_{m,t'} + \nu_{t'} < x_{m,t'} \,,$$

which is impossible, thus $\lambda_m > 0$. Now, we observe that $\mathcal{T}_n^\circ \subset \mathcal{T}_0$. Indeed, otherwise, if $t'' \in \mathcal{T}_n^\circ \cap \mathcal{T}_0^c$, we have $\nu_{t''} = -\lambda_n > 0$ and $x_{m,t''} = y_{m,t''} - \nu_{t''} < y_{m,t''} < \overline{x}_{m,t''}$, thus we get:

$$x_{m,t''} = \max(\underline{x}_{m,t''}, y_{m,t''} + \lambda_m) \geqslant y_{m,t''} + \lambda_m > y_{m,t''} = x_{m,t''} + \nu_{t''} + \lambda_m > x_{m,t''} \,,$$

which is impossible, thus $\mathcal{T}_n^\circ \subset \mathcal{T}_0$.

Finally, since $\overline{\mathcal{T}}_n^c \neq \emptyset$, consider $t_0 \in \arg\min_{t \notin \overline{\mathcal{T}}_n} \{\overline{x}_{n,t} - y_{n,t}\}$. By (1.42), we obtain:

$$y_{n,t_0} + \lambda_n < \overline{x}_{n,t_0} \iff E_n - \sum_{t \in \underline{\mathcal{T}}_n} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} y_{n,t} - \sum_{t \in \overline{\mathcal{T}}_n} \overline{x}_{n,t} < |\mathcal{T}_n^\circ|(\overline{x}_{n,t_0} - y_{n,t_0}) \tag{1.44}$$

and thus:

$$\begin{aligned}
&E_n - \sum_{t \in \mathcal{T}_0^c} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_0} \overline{x}_{n,t} \\
&= E_n - \sum_{t \in \underline{\mathcal{T}}_n} \underline{x}_{n,t} + \sum_{t \in \underline{\mathcal{T}}_n \cap \mathcal{T}_0} \underline{x}_{n,t} - \sum_{t \in \overline{\mathcal{T}}_n \cup \mathcal{T}_n^\circ} \overline{x}_{n,t} - \sum_{t \in \underline{\mathcal{T}}_n \cap \mathcal{T}_0} \overline{x}_{n,t} \quad (\text{as } \overline{\mathcal{T}}_n \cup \mathcal{T}_n^\circ \subset \mathcal{T}_0) && (1.45)\\
&< \sum_{t \in \mathcal{T}_n^\circ} (\overline{x}_{n,t_0} - y_{n,t_0}) - (\overline{x}_{n,t} - y_{n,t}) + \sum_{\underline{\mathcal{T}}_n \cap \mathcal{T}_0} (\underline{x}_{n,t} - \overline{x}_{n,t}) \quad (\text{from (1.44)}) && (1.46)\\
&\leqslant 0 \quad (\text{from the definition of } t_0 \text{ and } \underline{x}_{n,t} \leqslant \overline{x}_{n,t}), && (1.47)
\end{aligned}$$

which contradicts $n \notin \mathcal{N}_0$ and terminates the proof for Item (i). □

Proof of Item (ii). To prove (ii), we see that if t is such that $\nu_t > 0$, then all the facts said before are true for $n \notin \mathcal{N}_0$ if we consider $\mathcal{T}_0' \stackrel{\text{def}}{=} \{t | \nu_t > 0\}$ instead of $\mathcal{T}_0$. In that case we will have $\lambda_n < 0$. However we cannot have $t'' \in \mathcal{T}_n^\circ \cap \mathcal{T}_0'^c$ because this would mean $\nu_{t''} = -\lambda_n > 0$ but we have $\nu_{t''} \leqslant 0$ because $t'' \notin \mathcal{T}_0'$. Thus $\mathcal{T}_n^\circ$ is necessarily empty, and if there is $t \in \underline{\mathcal{T}}_n \cap \mathcal{T}_0'$, the same serie of inequalities as (1.46-1.47) show a contradiction. Consequently, for each $t \in \mathcal{T}_0'$, $x_{n,t} = \overline{x}_{n,t}$ and $y_{n,t} = x_{n,t} + \nu_t > \overline{x}_{n,t}$, thus $t \in \mathcal{T}_0$ and $\mathcal{T}_0' \subset \mathcal{T}_0$. The other inclusion is immediate. □

Proof of Item (iii). Suppose on the contrary that there exists $n \in \mathcal{N}_0$ such that $\lambda_n \geqslant 0$. For $t \in \mathcal{T}_n^\circ$, we have $\nu_t = -\lambda_n \leqslant 0$, thus, $\mathcal{T}_n^\circ \subset \mathcal{T}_0^c$. Then, if $t \in \mathcal{T}_0$ and if $x_{n,t} < \overline{x}_{n,t}$, we would have:

$$x_{n,t} = \max(\underline{x}_{n,t}, y_{n,t} + \lambda_n) \geqslant x_{n,t} + 0 + \nu_t > x_{n,t}\ ,$$

which is impossible, thus $x_{n,t} = \overline{x}_{n,t}$, and $\mathcal{T}_0 \subset \overline{\mathcal{T}}_n$. As we show independently in Item (v) that $\mathcal{T}_0 \neq \emptyset$, we know $\overline{\mathcal{T}}_n \neq \emptyset$. Let us consider $t_0 \in \arg\min_{t \notin \underline{\mathcal{T}}_n} \{y_{n,t} - \underline{x}_{n,t}\}$. By (1.42), we obtain:

$$y_{n,t_0} + \lambda_n > \underline{x}_{n,t_0} \iff E_n - \sum_{t \in \underline{\mathcal{T}}_n} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} y_{n,t} - \sum_{t \in \overline{\mathcal{T}}_n} \overline{x}_{n,t} > |\mathcal{T}_n^\circ|(\underline{x}_{n,t_0} - y_{n,t_0}) \tag{1.48}$$

and thus:

$$\begin{aligned}
&E_n - \sum_{t \in \mathcal{T}_0^c} \underline{x}_{n,t} - \sum_{\mathcal{T}_0} \overline{x}_{n,t} \\
&= E_n - \sum_{t \in \underline{\mathcal{T}}_n} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_0^c \cap \overline{\mathcal{T}}_n} \underline{x}_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} \underline{x}_{n,t} - \sum_{t \in \overline{\mathcal{T}}_n} \overline{x}_{n,t} + \sum_{t \in \mathcal{T}_0^c \cap \overline{\mathcal{T}}_n} \overline{x}_{n,t} \quad (\text{as } \mathcal{T}_0 \subset \overline{\mathcal{T}}_n) && (1.49)\\
&> \sum_{t \in \mathcal{T}_n^\circ} \big((y_{n,t} - \underline{x}_{n,t}) - (y_{n,t_0} - \underline{x}_{n,t_0})\big) + \sum_{t \in \mathcal{T}_0^c \cap \overline{\mathcal{T}}_n} \overline{x}_{n,t} - \underline{x}_{n,t} \quad (\text{from (1.48)}) && (1.50)\\
&\geqslant 0 \quad (\text{from the definition of } t_0 \text{ and } \underline{x}_{n,t} \leqslant \overline{x}_{n,t}), && (1.51)
\end{aligned}$$

which contradicts $n \in \mathcal{N}_0$ and terminates the proof of Item (iii). □

Proof of Item (iv). From (ii), we know that $\mathcal{T}_0^c = \{t | \nu_t \leqslant 0\}$, thus, if $t \notin \mathcal{T}_0$ and $n \in \mathcal{N}_0$, if $x_{n,t} > \underline{x}_{n,t}$ then we would have $x_{n,t} \leqslant y_{n,t} + \lambda_n = x_{n,t} + \nu_t + \lambda_n < x_{n,t}$, which is a contradiction. □

Proof of Item (v). From $\sum_t \nu_t = 0$, we see that if $\mathcal{T}_0 = \emptyset$, then this means that $\nu_t = 0$ for all $t \in \mathcal{T}$, and thus $\boldsymbol{y} = \boldsymbol{x}$ which is a contradiction. Thus there exists t_0 such that $\nu_{t_0} > 0$ and for the same reason, there exists t_0' such that $\nu_{t_0'} < 0$.

If $\mathcal{N}_0 = \emptyset$, then using (i), we would have for all n, $y_{n,t_0} > \overline{x}_{n,t_0}$ and thus $p_{t_0} > \sum_{n \in \mathcal{N}} \overline{x}_{n,t_0}$, which contradicts the aggregate upper bound constraint $p_t \leqslant \sum_{n \in \mathcal{N}} \overline{x}_{n,t}$ for each $t \in \mathcal{T}$.

If $\mathcal{N}_0^c = \emptyset$, then using (iv), we would have for all n, $y_{n,t'_0} < \underline{x}_{n,t'_0}$ and thus $p_{t'_0} < \sum_{n\in\mathcal{N}} \underline{x}_{n,t'_0}$, which contradicts the aggregate lower bound constraint $\forall t,\ p_t \geqslant \sum_{n\in\mathcal{N}} \underline{x}_{n,t}$. □

1.B Proof of Lemma 1.2

Proof of Item (i). From $\boldsymbol{x}^{(K)} = P_{\mathcal{X}}(\boldsymbol{y}^{(K-1)})$ and $\boldsymbol{y}^{(K)} = P_{\mathcal{Y}}(\boldsymbol{x}^{(K)})$, we obtain, similarly to (1.40):

$$\begin{aligned} &\forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad 0 = (x_{n,t}^{(K)} - y_{n,t}^{(K-1)}) - \lambda_n^{(K)} - \underline{\mu}_{n,t}^{(K)} + \overline{\mu}_{n,t}^{(K)} \\ &\text{and} \quad y_{n,t}^{(K)} = x_{n,t}^{(K)} + \nu_t^{(K)} \ . \end{aligned}$$

where the Lagrangian multipliers $\lambda_n^{(K)}, \underline{\mu}_{n,t}^{(K)}, \overline{\mu}_{n,t}^{(K)}$ (resp. $\nu_t^{(K)}$) are associated to the quadratic problem characterizing the projections $P_{\mathcal{X}}(\boldsymbol{y}^{(K-1)})$ (resp. $\boldsymbol{y}^{(K)} = P_{\mathcal{Y}}(\boldsymbol{x}^{(K)})$). We obtain equalities similar to (1.41, 1.42). We proceed as for Proposition 1.4(i) and suppose that there exists $n \notin \mathcal{N}_0$ and $\hat{t} \in \mathcal{T}_0$ such that $x_{n,\hat{t}} < \overline{x}_{n,\hat{t}}$. Then, as $\left\|\boldsymbol{y}^{(K)} - \boldsymbol{y}^{\infty}\right\| \leqslant \frac{\varepsilon_{\text{cvg}}}{1-\rho}$ and $\sum_{t\in\mathcal{T}} \boldsymbol{y}_t^{(K)} = \sum_{t\in\mathcal{T}} \boldsymbol{y}_t^{\infty}$, we have for each $n \in \mathcal{N}, t \in \mathcal{T}$, $|y_{n,t}^{(K)} - y_{n,t}^{\infty}| \leqslant \frac{\varepsilon_{\text{cvg}}}{2(1-\rho)}$, and thus we get:

$$\begin{aligned} &\overline{x}_{n,\hat{t}} \geqslant x_{n,\hat{t}}^{(K)} \geqslant y_{n,\hat{t}}^{(K-1)} + \lambda_n^{(K)} \geqslant y_{n,\hat{t}}^{\infty} - \tfrac{\varepsilon_{\text{cvg}}}{2(1-\rho)} + \lambda_n^{(K)} = x_{n,\hat{t}}^{\infty} + \nu_{\hat{t}}^{\infty} - \tfrac{\varepsilon_{\text{cvg}}}{2(1-\rho)} + \lambda_n^{(K)} \\ &\implies \lambda_n^{(K)} < \tfrac{\varepsilon_{\text{cvg}}}{2(1-\rho)} v - \nu_{\hat{t}}^{\infty} < \tfrac{B\varepsilon_{\text{cvg}}}{2} - 2B\varepsilon_{\text{cvg}} = -\tfrac{3}{2}B\varepsilon_{\text{cvg}} \ , \end{aligned} \tag{1.52}$$

as $\nu_{\hat{t}}^{\infty} \geqslant \underline{\nu} > 2B\varepsilon_{\text{cvg}}$. Let us now consider $t' \in \mathcal{T}_n^{\circ(K)} \cup \overline{\mathcal{T}}_n^{(K)}$, then:

$$\nu_{t'}^{(K)} = y_{n,t'}^{(K)} - x_{n,t'}^{(K)} \geqslant y_{n,t'}^{(K)} - y_{n,t'}^{(K)} - \lambda_n^{(K)} \tag{1.53}$$

$$> -\tfrac{\varepsilon_{\text{cvg}}}{2} + \tfrac{3}{2}B\varepsilon_{\text{cvg}} > B\varepsilon_{\text{cvg}} + \tfrac{\varepsilon_{\text{cvg}}}{2}(B-1) \geqslant B\varepsilon_{\text{cvg}} \ , \tag{1.54}$$

which shows that $t' \in \mathcal{T}_0^{(K)} = \mathcal{T}_0^{\infty}$ and thus $\mathcal{T}_n^{\circ(K)} \cup \overline{\mathcal{T}}_n^{(K)} \subset \mathcal{T}_0$. Then, the same serie of inequalities as (1.44, 1.46, 1.47) applied to $\boldsymbol{y}^{(K-1)}$ gives a contradiction to $n \notin \mathcal{N}_0$. □

Proof of Item (ii). The proof of Item (ii) is symmetric to the one of Item (i): if we suppose that there exists $n \in \mathcal{N}_0$ and $\hat{t} \notin \mathcal{T}_0$ such that $x_{n,\hat{t}}^{(K)} > \underline{x}_{n,\hat{t}}$, we obtain, symmetrically to (1.52), that $\lambda_n^{(K)} \geqslant -\frac{\varepsilon_{\text{cvg}}}{2(1-\rho)}$. Then, considering $t' \in \underline{\mathcal{T}}_n^{(K)} \cup \mathcal{T}_n^{\circ(K)}$, we show, symmetrically to (1.53), that $\nu_{t'}^{(K)} < B\varepsilon_{\text{cvg}}$ i.e. $t' \notin \mathcal{T}_0$ and thus $\underline{\mathcal{T}}_n^{(K)} \cup \mathcal{T}_n^{\circ(K)} \subset \mathcal{T}_0^c$. We conclude by obtaining a contradiction to $n \in \mathcal{N}_0$ by the same serie of inequalities as (1.48, 1.50, 1.51). □

1.C Fast Projection on a Boxed Simplex

In general, the Euclidean projection onto a convex set $\mathcal{X}$ can be computed with a mathematical program with a quadratic objective. Indeed the projection $P_{\mathcal{X}}(\boldsymbol{y})$ of a vector $\boldsymbol{y}$ will be the solution of:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \ . \tag{1.55}$$

However, in some cases, the projection can be computed much more efficiently: for instance in the case of a box, the $\mathcal{L}_1$-ball or the simplex $\mathcal{X} = \{\boldsymbol{x} \in \mathbb{R}_+^p \ ; \ \sum_i x_i = 1\} \subset \mathbb{R}^p$ for which there is an algorithm in $O(p\log(p))$ where p is the dimension of the space [CY11; Con16; Mic86; WCP13].

This happens to be also the case for the polytope (1.10) that we consider in this paper. Indeed, we give here a generalization of the algorithm studied in [WCP13] for the projection

onto a simplex, adapted to a polytope $\mathcal{X}$ of the form (1.10), that is defined as:

$$\mathcal{X} \stackrel{\text{def}}{=} \{(x_i)_i \in \mathbb{R}^p \; ; \; \sum_{i=1}^{p} x_i = E \text{ and } \forall i, \; \underline{x}_i \leqslant x_i \leqslant \overline{x}_i\} \; . \tag{1.56}$$

In this Section 1.C, we drop the index n to simplify the notation, as we consider on individual projection for a fixed agent $n \in \mathcal{N}$.

Of course, as the algorithm for projection onto a simplex in [WCP13] is of complexity $\mathcal{O}(p \log p)$, and can be easily adapted to lower bounds $x_i \geqslant \underline{x}_i$ (instead of $x_i \geqslant 0$), we can easily get the projection on our set $\mathcal{X}$ in time $\mathcal{O}(p^2)$. For that, one can sort the indices and consider adding them to the set $\mathcal{I}_{\max} = \{i | x_i = \overline{x}_i\}$ one by one, and calling each time (at most p times) the algorithm [WCP13] which is of order p once the sorting has been done. This is the idea the authors of [WCP13] exposed in a later paper [WL15].

Some algorithms (see [Bru84; HKL80] and the references therein) with running time of order $\mathcal{O}(p \log p)$ (or even of order $\mathcal{O}(p)$ in average complexity for [Bru84]) have been found to solve, more generally, convex quadratic programs with a unique equality constraint, that is, problems having the structure:

$$\min_{\boldsymbol{x} \in \mathbb{R}^p} \sum_{1 \leqslant i \leqslant p} a_i x_i + \tfrac{b_i}{2} x_i^2 \tag{1.57a}$$

$$\sum_{1 \leqslant i \leqslant p} b_i x_i = b_0 \tag{1.57b}$$

$$\underline{x}_i \leqslant x_i \leqslant \overline{x}_i, \; \forall i, \tag{1.57c}$$

with $b_i \geqslant 0, \forall i$, for which (1.55)-(1.56) is a particular case.

Algorithm 1.7 presented here has the advantages of being very simple to implement and of having a simple proof of validity, which relies on the optimality conditions of the quadratic optimization problem (1.55). In fact, one can easily extend Algorithm 1.7 and the proof of validity to generalized convex quadratic programs as (1.57), by an affine transformation of the coefficients. This extension is easy and left to readers: we prefer to give and prove the validity of the algorithm in the case of (1.56) for simplicity and as it is directly linked to the framework of this thesis (Chapter 1 and Chapter 3).

The Algorithm 1.7 presented here relies on the following result: on the resulting projection of $\boldsymbol{y}$ on $\mathcal{X}$, $\boldsymbol{x} \stackrel{\text{def}}{=} \mathrm{P}_{\mathcal{X}}(\boldsymbol{y})$, some coordinates will be at the upper bound $\overline{x}_i$, some at the lower bound $\underline{x}_i$ and some within the two bounds. If we denote by $\mathcal{I}_{\max}, \mathcal{I}_{\min}$ and $\mathcal{I}_{\text{int}}$ the respective set of indices, then a result from the Karush-Kuhn-Tucker conditions (see in the proof below) is that:

$$\begin{aligned} \forall i \in \mathcal{I}_{\text{int}}, \; x_i = y_i + \lambda \text{ with } \lambda &= \frac{1}{|\mathcal{I}_{\text{int}}|} \sum_{i \in \mathcal{I}_{\text{int}}} (x_i - y_i) \\ &= \frac{1}{|\mathcal{I}_{\text{int}}|} \left(E - \sum_{i \in \mathcal{I}_{\text{int}}} y_i - \sum_{i \in \mathcal{I}_{\min}} \underline{x}_i - \sum_{i \in \mathcal{I}_{\max}} \overline{x}_i \right). \end{aligned}$$

Thus, the complexity of the computation of the projection $\boldsymbol{x} = \mathrm{P}_{\mathcal{X}}(\boldsymbol{y})$ lies in finding the sets $\mathcal{I}_{\max}$ and $\mathcal{I}_{\min}$: once these sets are obtained, we can immediatly obtain λ and $\boldsymbol{x}$ by the above formula. The algorithm presented here computes these two sets $\mathcal{I}_{\max}$ and $\mathcal{I}_{\min}$ by considering the indices $i \in \mathcal{I}$ sorted in two different orders.

Termination and Complexity We first show that Algorithm 1.7 terminates if the instance is feasible, that is $\mathcal{X}$ is nonempty. One can notice that in the first **while** (line 6) we run over the list Lmax and as soon as the condition on line 18 is not satisfied, we break on Line 22, thus we can enter the **while** loop at most cardLmax) $= p$ times.

The second **while** loop (line 7) runs over the list Lmin. At each entering in the loop we get a new index as the last element from Lmin (this action can be executed on a list in constant time). If we put it back in Lmin afterwards (line 22), then we exit the **while** loop and Lmax

Algorithm 1.7 Fast Projection on Boxed Simplex Algorithm

Require: a point $\boldsymbol{y} \in \mathbb{R}^p$

1: compute sorting σ such that $(y_{\sigma(1)} - \underline{x}_{\sigma(1)}) \leqslant (y_{\sigma(2)} - \underline{x}_{\sigma(2)}) \leqslant \ldots \leqslant (y_{\sigma(p)} - \underline{x}_{\sigma(p)})$
2: compute sorting τ such that $(\overline{x}_{\tau(1)} - y_{\tau(1)}) \geqslant (\overline{x}_{\tau(2)} - y_{\tau(2)}) \geqslant \ldots \geqslant (\overline{x}_{\tau(p)} - y_{\tau(p)})$
3: Lmin $= [\sigma(1), \sigma(2), \ldots, \sigma(p)]$ ▷ List of indices that are at $\underline{x}$
4: Lmax $= [\tau(1), \tau(2), \ldots, \tau(p)]$ ▷ List of indices that are NOT at $\overline{x}$
5: $S \leftarrow E - \sum_{i=1}^{p} \underline{x}_i$ $N_{\text{int}} \leftarrow 0$ ▷ number of resulting indices in $]\underline{x}, \overline{x}[$
6: **while** Lmax $\neq \emptyset$ **do**
7: **while** Lmin $\neq \emptyset$ **do**
8: $i \leftarrow$ get and remove last element of Lmin
9: **if** $y_i + \frac{S + \underline{x}_i - y_i}{N_{\text{int}} + 1} > \underline{x}_i$ **then**
10: $S \leftarrow S + \underline{x}_i - y_i$
11: $N_{\text{int}} \leftarrow N_{\text{int}} + 1$
12: **else**
13: append i at the end of Lmin
14: **break**
15: **end**
16: **done**
17: $i \leftarrow$ get and remove last element of Lmax
18: **if** $y_i + \frac{S}{N_{\text{int}}} > \overline{x}_i$ **then**
19: $S \leftarrow S + y_i - \overline{x}_i$
20: $N_{\text{int}} \leftarrow N_{\text{int}} - 1$
21: **else**
22: Lmax.append(i)
23: **break**
24: **end**
25: **done**
26: **return** $\left[\min\left(\overline{x}_i, \max\left(\underline{x}_i, y_i + \frac{S}{N_{\text{int}}}\right)\right)\right]_{1 \leqslant i \leqslant p}$

decreases (line 17): this can happen only p times, so finally we can enter this second **while** loop line 7 only $2p$ times in total.

Thus, the algorithm terminates and the lines 6 to 25 are run within $\mathcal{O}(p)$ operations (affectations and additions). As the sortings on lines 1 and 2 takes $\mathcal{O}(p \log p)$ operations, the overall complexity of Algorithm 1.7 is of $\mathcal{O}(p \log p)$.

Proof of Validity. The proof extends the one of [WCP13]. Let us consider the problem:

$$\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \tag{1.58a}$$

$$\text{s.t.} \sum_{i=1}^{p} x_i = E \tag{1.58b}$$

$$\forall i,\ \underline{x}_i \leqslant x_i \leqslant \overline{x}_i\,, \tag{1.58c}$$

for which the solution is the projection $\mathrm{P}_{\mathcal{X}}(\boldsymbol{y})$ and for which the Lagrangian function is:

$$\mathcal{L}(x) = \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda(E - \sum_{i=1}^{p} x_i) + \sum_i \overline{\nu}_i (x_i - \overline{x}_i) + \sum_i \underline{\nu}_i (x_i - \underline{x}_i)$$

with $\lambda \in \mathbb{R}$ the Lagrangian multiplier associated to (1.58b) and $\overline{\nu}, \underline{\nu} \geqslant 0$ the multipliers associated to (1.58c). The KKT stationarity conditions give:

$$\forall i \in \mathcal{I},\ (x_i - y_i) = \lambda - \overline{\nu}_i + \underline{\nu}_i \tag{1.59}$$

along with the completarity conditions:

$$\forall i \in \mathcal{I},\ 0 \leqslant \overline{\nu}_i \ \perp\ (\overline{x}_i - x_i) \geqslant 0 \ \text{ and } \ 0 \leqslant \underline{\nu}_i \perp (x_i - \underline{x}_i) \geqslant 0\,.$$

Thus, if $\overline{\nu}_i > 0$ then $x_i = \overline{x}_i$ and if $\underline{\nu}_i > 0$ then $x_i = \underline{x}_i$. Denote by $\mathcal{I} \stackrel{\text{def}}{=} \{1, \ldots, p\}$ the set of indices, $\mathcal{I}_{\max} \stackrel{\text{def}}{=} \{i \in \mathcal{I}\ ;\ x_i = \overline{x}_i\}$, $\mathcal{I}_{\min} \stackrel{\text{def}}{=} \{i \in \mathcal{I}\,; x_i = \underline{x}_i\}$ and $\mathcal{I}_{\text{int}} \stackrel{\text{def}}{=} \{i \in \mathcal{I}\ ;\ \underline{x}_i < x_i < \overline{x}_i\}$ the subsets of indices to the max , min and in-between at the solution, and denote by $k_{\text{int}} \stackrel{\text{def}}{=} |\mathcal{I}_{\text{int}}|$ the cardinal of $\mathcal{I}_{\text{int}}$ at the solution. Then we get by summing (1.59) on $\mathcal{I}_{\text{int}}$ that:

$$k_{\text{int}}\lambda = \sum_{in \in \mathcal{I}_{\text{int}}} (x_i - y_i) = E - \sum_{i \in \mathcal{I}_{\text{int}}} y_i - \sum_{i \in \mathcal{I}_{\min}} \underline{x}_i - \sum_{i \in \mathcal{I}_{\max}} \overline{x}_i$$

where the second equality is obtained from $\sum_{i=1}^{p} x_i = E$. Hence, because we have $\overline{\nu}_i > 0 \Leftrightarrow \overline{x}_i < y_i + \lambda$ and $\underline{\nu}_i > 0 \Leftrightarrow \underline{x}_i > y_i + \lambda$, the solution is given by:

$$\forall i\ , x_i = \min\left(\overline{x}_i, \max\left(\underline{x}_i, y_i + \lambda\right)\right)\ ,$$

and thus, to show the validity of Algorithm 1.7, it is sufficient to show that at the end of the algorithm, $\frac{S}{N_{\text{int}}}$ is equal to λ, or equivalently, that $S = E - \sum_{i \in \mathcal{I}_{\text{int}}} y_i - \sum_{i \in \mathcal{I}_{\min}} \underline{x}_i - \sum_{i \in \mathcal{I}_{\max}} \overline{x}_i$ and that N_{int} corresponds to k_{int}, the cardinality of $\mathcal{I}_{\text{int}}$.

For that, we first notice that the coordinates are at $\underline{x}_i$ in the order of σ, that is for $i, j \in \mathcal{I}$:

$$\left[y_i - \underline{x}_i \leqslant y_j - \underline{x}_j \text{ and } x_j = \underline{x}_j\right] \Rightarrow\ x_i = \underline{x}_i\,. \tag{1.60}$$

Inded, if not we have from (1.59) $x_i - y_i = \lambda - \overline{\nu}_i \leqslant \lambda$ and $\underline{x}_j - y_j = \lambda + \underline{\nu}_j \geqslant \lambda$. Thus:

$$\underline{x}_j - y_j \geqslant \lambda \geqslant x_i - y_i > \underline{x}_i - y_i\,,$$

which is impossible by hypothesis. Symmetrically, we have:

$$\left[\overline{x}_i - y_i \leqslant \overline{x}_j - y_j \text{ and } x_j = \overline{x}_j\right] \Rightarrow\ x_i = \overline{x}_i\,. \tag{1.61}$$

Let us denote by $k_{\min}$ and $k_{\max}$ the cardinals of $\mathcal{I}_{\min}$ and $\mathcal{I}_{\max}$ at the solution. Then we know from (1.60) that $\mathcal{I}_{\min} = \{\sigma(1), \ldots, \sigma(k_{\min})\}$ and from (1.61) that $\mathcal{I}_{\max} = \{\tau(p - k_{\max} + 1), \ldots, \tau(p)\}$.

If i is taken out of Lmax (line 17) and condition line 18 is True, then i was already taken out of Lmin before because $\underline{x}_i \leqslant \overline{x}_i$ (feasibility). At any time of the algorithm, we have:

$$\begin{aligned}
&\text{Lmax} = [\tau(1), \ldots \tau(p - N_{\max})]\\
&\text{Lmin} = [\sigma(1), \ldots, \sigma(N_{\min})] \text{ with } N_{\min} = p - (N_{\text{int}} + N_{\max})\\
&S = E - \sum_{i=1}^{p} \underline{x}_i + \sum_{i=1}^{N_{\text{int}}+N_{\max}} (\underline{x}_{\sigma(i)} - y_{\sigma(i)}) + \sum_{i=p-N_{\max}+1}^{p} (y_{\tau(i)} - \overline{x}_{\tau(i)})\\
&\ = E - \sum_{i \in \text{Lmin}} \underline{x}_i - \sum_{i \notin \text{Lmax}} \overline{x}_i - \sum_{i \in (\text{Lmax} \setminus \text{Lmin})} y_i\,.
\end{aligned}$$

Taking the values of Lmin, Lmax, N_{int}, $N_{\min}$, $N_{\max}$ and S *at the end of the algorithm*, we have:

$$y_{\sigma(N_{\min}+1)} + \frac{S}{N_{\text{int}}} > \underline{x}_{\sigma(N_{\min}+1)} \tag{1.62}$$

$$y_{\sigma(N_{\min})} + \frac{S + \underline{x}_{\sigma(N_{\min})} - y_{\sigma(N_{\min})}}{N_{\text{int}} + 1} \leqslant \underline{x}_{\sigma(N_{\min})} \tag{1.63}$$

$$y_{\tau(p-N_{\max})} + \frac{S}{N_{\text{int}}} \leqslant \overline{x}_{\tau(p-N_{\max})}\,. \tag{1.64}$$

For $i \notin \text{Lmin}$, we want to show that $y_i - \underline{x}_i + \frac{S}{N_{\text{int}}} \geqslant 0$. We know that $y_i - \underline{x}_i \geqslant y_{\sigma(N_{\min}+1)} - \underline{x}_{\sigma(N_{\min}+1)}$. From (1.62) we get:

$$S + N_{\text{int}}(y_i - \underline{x}_i) = S + N_{\text{int}}(y_{\sigma(N_{\min}+1)} - \underline{x}_{\sigma(N_{\min}+1)}) + N_{\text{int}}\Big(y_i - \underline{x}_i) - (y_{\sigma(N_{\min}+1)} - \underline{x}_{\sigma(N_{\min}+1)})\Big)$$

and both parts are nonnegative. For $i \in \text{Lmin} \Leftrightarrow \sigma^{-1}(i) \leqslant N_{\min}$, we have $y_i - \underline{x}_i \leqslant (y_{\sigma(N_{\min})} - \underline{x}_{\sigma(N_{\min})})$ and from (1.63) we get:

$$\begin{aligned} & N_{\text{int}}(y_{\sigma(N_{\min})} - \underline{x}_{\sigma(N_{\min})}) + S \leqslant 0 \\ \implies S + N_{\text{int}}(y_i - \underline{x}_i) &= \Big(S + N_{\text{int}}(y_{\sigma(N_{\min})} - \underline{x}_{\sigma(N_{\min})})\Big) \\ &\quad + N_{\text{int}}\Big((y_i - \underline{x}_i) - (y_{\sigma(N_{\min})} - \underline{x}_{\sigma(N_{\min})})\Big) \end{aligned}$$

and both parts are nonpositive, so that $y_i + \frac{S}{N_{\text{int}}} \leqslant 0$.

Now, let us consider $i \in \text{Lmax} \Leftrightarrow \tau^{-1}(i) \leqslant p - N_{\max}$ so that we have $\overline{x}_i - y_i \geqslant \overline{x}_{\tau(p-N_{\max})} - y_{\tau(p-N_{\max})}$ and from (1.64):

$$y_i - \overline{x}_i + \frac{S}{N_{\text{int}}} \leqslant y_{\tau(p-N_{\max})} - \overline{x}_{\tau(p-N_{\max})} + \frac{S}{N_{\text{int}}} \leqslant 0$$

so that $y_i + \frac{S}{N_{\text{int}}} \leqslant \overline{x}_i$.

It remains to show that for $i \notin \text{Lmax} \Leftrightarrow \tau^{-1}(i) \geqslant p - N_{\max} + 1$, we have $y_i + \frac{S}{N_{\text{int}}} \geqslant \overline{x}_i$.

We begin with the last index popped out of Lmax, $a \stackrel{\text{def}}{=} \tau(p - N_{\max} + 1)$. We know that when a was popped out of Lmax at line 17, there was S' and N'_{int} such that:

$$y_a + \frac{S'}{N'_{\text{int}}} > \overline{x}_a \,. \tag{1.65}$$

After that, we went into the while loop Line 7 only once before the algorithm terminated. There was a set of r (possibly 0) indices $k_1, \ldots, k_r$ popped out of Lmin during this last step, starting with k_1 which gives (1.66). As k_1 was not popped out of Lmin before adding a to the indices at max, we have also (1.67):

$$y_{k_1} + \frac{S' + \underline{x}_{k_1} - y_{k_1}}{N'_{\text{int}} + 1} \leqslant \underline{x}_{k_1} \;\Leftrightarrow\; N'_{\text{int}}(y_{k_1} - \underline{x}_{k_1}) + S' \leqslant 0 \tag{1.66}$$

$$y_{k_1} + \frac{S' + y_a - \overline{x}_a + \underline{x}_{k_1} - y_{k_1}}{N'_{\text{int}} - 1 + 1} > \underline{x}_{k_1} \tag{1.67}$$

These two equations (1.66) and (1.67) imply that:

$$y_a - \overline{x}_a + \underline{x}_{k_1} - y_{k_1} > -(N'_{\text{int}}(y_{k_1} - \underline{x}_{k_1}) + S') \geqslant 0 \,. \tag{1.68}$$

As at the end of the algorithm, we have $S = S' + y_a - \overline{x}_a + \sum_{j=1}^{r}(\underline{x}_{k_j} - y_{k_j})$ and $N_{\text{int}} = N'_{\text{int}} - 1 + r$, we get:

$$\begin{aligned} N_{\text{int}}\Big(y_a - \overline{x}_a + \frac{S}{N_{\text{int}}}\Big) &= N_{\text{int}}\,(y_a - \overline{x}_a) + S \\ &= N'_{\text{int}}(y_a - \overline{x}_a) + (r-1)(y_a - \overline{x}_a) + S' + y_a - \overline{x}_a + \sum_{j=1}^{r}(\underline{x}_{k_j} - y_{k_j}) \\ &= \big[N'_{\text{int}}(y_a - \overline{x}_a) + S'\big] + \sum_{j=1}^{r}(y_a - \overline{x}_a) + (\underline{x}_{k_j} - y_{k_j}) \geqslant 0 \end{aligned}$$

where the positivity of the first term comes from (1.65) and the sum is positive componentwise because $(y_a - \overline{x}_a) + (\underline{x}_{k_j} - y_{k_j}) \geqslant y_a - \overline{x}_a + \underline{x}_{k_1} - y_{k_1} > 0$ by (1.68). Hence we get that

$y_a + \frac{S}{N_{\text{int}}} > \overline{x}_a$. By induction, the next elements of the list $\tau(k)$ for $k > p - N_{\max} + 1$ also satisfy $y_{\tau(k)} + \frac{S}{N_{\text{int}}} > \overline{x}_{\tau(k)}$.

Finally, we have shown that the solution $(\boldsymbol{x}, \overline{\boldsymbol{\nu}}, \underline{\boldsymbol{\nu}}, \lambda)$ defined by:

$$\forall i \in \text{Lmin}, x_i = \underline{x}_i, \quad \forall i \notin \text{Lmax}, x_i = \overline{x}_i, \quad \text{and } x_i = y_i + \frac{S}{N_{\text{int}}} \text{ else}$$

and with the optimal Lagrangian multipliers:

$$\lambda = \frac{S}{N_{\text{int}}}, \quad \forall i \in \mathcal{I}, \underline{\nu}_i = \begin{cases} 0 \text{ if } i \notin \text{Lmin} \\ \underline{x}_i - y_i - \frac{S}{N_{\text{int}}} \text{ else} \end{cases}, \overline{\nu}_i = \begin{cases} 0 \text{ if } i \in \text{Lmax} \\ \frac{S}{N_{\text{int}}} - (\overline{x}_i - y_i) \text{ else} \end{cases},$$

satisfy the KKT conditions. Hence it corresponds to the unique solution of the system, which is the projection $\mathrm{P}_{\mathcal{X}}(\boldsymbol{y})$. □

It is an open question to know if the proposed Algorithm 1.7, based on a sorting of the coordinates, could be extended to other family of polyhedra: for instance to polyhedra defined by *two* linear equalities instead of a single one. Such a generalization, along with the impact on the time complexity of the algorithm, is not straightforward from our results and is out of the scope of this paper, but would be an interesting avenue for further work.

Part II

Decentralized Management of Flexibilities and Game Theory

Demand Response has been the subject of a blooming literature in the last decade. One can refer to [Sia14], [VZV15] and [Saa+12] for surveys on the topic. Various aspects of DR have been investigated such as consumers personal utilities and discomfort related to their electricity consumption, consumers privacy, network and power flow constraints, often leading to complex optimization problems.

As stated in Part I, the need for a decentralized optimization approach is a consensus and, among the different decentralized approaches considered in the DR literature, several works rely on a dual decomposition of the optimization problem of a central operator (e.g. [CBK17; Shi+14; LCL11; Den+15]). In those works, an iterative algorithm between the operator (updating the Lagrange multipliers) and the consumers (adjusting their consumption) is run until convergence of the consumption profiles (decomposition-coordination).

In Chapter 1, we considered another decentralized optimization approach, which can be applied to solve the central optimization of flexibilities subject to individual and confidential constraints for each consumer. In the method proposed in Chapter 1, as well as in the dual decomposition approaches, it is considered that each consumer (agent) is indifferent between the profiles that satisfy her individual constraints (Chapter 1 focuses on the privacy issues in the procedure used to manage the flexibilities).

In these decentralized approaches, agents are not *strategic*, in the sense that they do not have individual and possibly competing objective cost functions to optimize.

To answer to these strategic aspects, different game-theoretic frameworks have been proposed in the smart grid literature, e.g. [Cha+14; Saa+12; Che+10; Che+14; Atz+13; CK14; Bah+13], following the seminal paper [MR+10].

This Part II of the thesis follows this path and focuses on game-theoretic approaches for demand response. Among the different game theory models studied in the literature, we focus in this thesis on the *daily billing* mechanism proposed by [MR+10], that we compare in Chapter 2 and Chapter 4 with the *hourly billing mechanism* considered in [Bah+13; Atz+13; Che+14; CK14]. In the latter, a consumer pays for each time period (e.g. each hour) proportionally to the energy consumed on this period. This billing mechanism has the structure of a *routing congestion game* [ORS93]. We show in Chapter 2, in addition to [Bah+13], that this latter model has desirable properties of fairness, incentives compatibility and efficiency, while remaining simple and intuitive. These properties are also confirmed by the analysis conducted in Chapter 4, considering consumers individual preferences. In Chapter 3, we focus on the hourly billing mechanism and analyze its implementation and algorithmic aspects.

This part of the thesis is divided in three chapters:

• Chapter 2 is based on the conference paper [Jac+17b]. In this chapter, we introduce and compare the two DR mechanisms, introduced respectively by [MR+10] and [Bah+13], in terms of efficiency and fairness. Each mechanism defines a game where the consumers optimize their flexible consumption to reduce their electricity bills. Mohsenian-Rad *et al* propose a daily mechanism for which they prove the social optimality. Baharlouei *et al* propose a hourly billing mechanism for which we give theoretical results: we prove the uniqueness of an equilibrium in the associated game and give an upper bound on its price of anarchy, quantifying its distance to the social optimum. Section 2.4.3 and Section 2.5.3 give an explicit expression of the equilibrium profiles under each of the two billing mechanisms in a particular case, and have not been submitted for publication.

• Chapter 3 is based on the journal paper [Jac+19c]. As the hourly billing mechanism has been shown in Chapter 2 to be of special interest as an intuitive and fair mechanism, we focus on this model and answer to several theoretical and practical questions. After precising some theoretical results on the uniqueness of the equilibrium profiles in the game defined, this chapter focuses on the computational aspects. We provide results on the convergence rates of two decentralized algorithms to compute the equilibrium: the cycling best response dynamics and a projected gradient descent method. Last, we simulate the proposed demand response framework in a stochastic environment where the parameters depend on forecasts. We show numerically the relevance of an online demand response procedure which reduces the impact of inaccurate forecasts in comparison to a standard offline procedure.

- Chapter 4 is based on the conference paper [Jac+17a]. In Demand Response programs, price incentives might not be sufficient to modify residential consumers load profile. In this work, we consider that each consumer has a preferred profile and a discomfort cost when deviating from it. Consumers can value this discomfort at a varying level that we take as a parameter. As in Chapter 2, we study a Demand Response environment as a game between consumers, and consider now their profile preferences. We compare the equilibria of the games associated to the two different dynamic pricing mechanisms defined and studied in Chapter 2. We give new results about equilibria as functions of the preference level in the case of quadratic system costs and prove that, whatever the preference level, system costs are smaller with the hourly mechanism.

Chapter 2

Two billing mechanisms for Demand Response: an Efficiency and Fairness Tradeoff

This chapter is based on the conference paper [Jac+17b], except for Section 2.4.3 and Section 2.5.3, which give explicit expressions of the equilibrium profiles under each of the two billing mechanisms studied in this chapter, in a specific case (quadratic costs and symmetric users).

2.1 Introduction

The implementation of a Demand Response (DR) mechanism raises several difficulties. Owing to the huge number of variables and constraints and to the impossibility for an aggregator to collect all the consumption constraints because of privacy concerns, the optimization has to rely on a decentralized algorithm that minimizes the information exchanged with the users. Obviously, efficiency is also a desirable property. One wishes that, as in [MR+10], the scheduling process leads to an optimal or close to optimal consumption profile for the global system cost, while respecting all the users constraints. Another important feature, more discussed in [BNMR12], is fairness: the payment model should penalize consumers imposing costly constraints for the system, while rewarding flexible ones. This point is essential to ensure the acceptability of the process and encourage users to stay in a DR program. Besides, this feature has a merit in terms of incentives. Indeed, such fair billing models would encourage consumers to modify their constraints so that available flexibility for the system increases.

The contributions of this chapter are twofold. First, we give new theoretical results associated with the hourly billing mechanism [BNMR12] by proving the uniqueness of a Nash Equilibrium, and by specifying an explicit upper bound on its price of anarchy. Next, we present numerical results, based on simulations using real consumption data, that compare the two billing mechanisms [MR+10] and [BNMR12] in terms of efficiency and fairness. The results show that the hourly mechanism achieves a very small price of anarchy and an important fairness property.

The chapter is organized as follows. In Section 2.2 we introduce the consumption game model. In Section 2.3 we recall some notions from game theory and define quantitative indicators to measure the efficiency and the fairness of a given billing mechanism. In Section 2.4 we introduce the daily proportional billing considered in [MR+10] and recall its main properties. In Section 2.5, we focus on the hourly proportional billing proposed in [BNMR12] and present our new theoretical results. Last, Section 2.6 is devoted to numerical experiments, based on real consumption data, from which we compare the fairness and efficiency of the two different billing mechanisms.

2.2 Energy Consumption Game

We consider an autonomous network composed of a unique electricity provider and a set $\mathcal{N}$ of N electricity consumers. We use a model similar to Mohsenian-Rad *et al* [MR+10].

2.2.1 Consumers constraints

Each user n has a set of electric appliances $\mathcal{A}_n$. For each $a \in \mathcal{A}_n$, this user—itself or through an Energy Consumption Scheduler (ECS)—can set the power $x_{na,t}$ allowed to a at each time period t in $\mathcal{T}_{na} = \{\alpha_{na}, \dots, \beta_{na}\} \subset \mathcal{T}$, where $\mathcal{T}$ is the set of time periods considered over a day. We consider the following constraints:

$$\sum_{t \in \mathcal{T}} x_{na,t} = E_{na}, \ \forall a \in \mathcal{A}_n \ , \tag{2.1a}$$

$$\underline{x}_{na,t} \leqslant x_{na,t} \leqslant \overline{x}_{na,t}, \ \forall a \in \mathcal{A}_n, \forall t \in \mathcal{T} \ . \tag{2.1b}$$

Each electric appliance $a \in \mathcal{A}_n$ requires a fixed daily amount of energy (2.1a). Due to physical limits, the power set to appliance a is bounded from below and above (2.1b). If the appliance a can not be used in the time period t, we set $\underline{x}_{na,t} = \overline{x}_{na,t} = 0$. The set of available time periods for $a \in \mathcal{A}_n$ is therefore given by $\mathcal{T}_{na} = \{t : \ \overline{x}_{na,t} > 0\}$, and thus, without loss of generality, we can consider $\mathcal{T} = \mathcal{T}_{n,a}$ for each n, a.

We will denote more compactly by $\mathcal{X}_n$ the set of feasible loads $(x_{na,t})_{a,t}$ that respect the constraints given by (2.1).

2.2.2 System cost functions

We denote by $C_t(X_t)$ the system cost for providing to users the total *aggregate* load

$$X_t \stackrel{\text{def}}{=} \sum_{a,n} x_{na,t}$$

at time t. It is widely accepted that marginal production costs increase with demand. Hence we assume that cost functions $C_t(.)$ are *increasing* and *strictly convex* [MR+10]. These functions can depend on the period t as it is more expensive to produce energy on peak hours or for instance, when the renewable production is low. Practically, in our model, $(C_t)_{t \in \mathcal{T}}$ can be the actual costs for the provider but can also be an artificial signal that is sent to each user's ECS in order to make him perform a decentralized optimization of his consumption. We will mostly consider quadratic cost functions as done in [MR+10]:

$$C_t(X_t) = \beta_t (X_t)^2 + \alpha_t X_t + \delta_t \ . \tag{2.2}$$

We assume for simplicity that the total system cost, denoted by $\mathcal{C} \stackrel{\text{def}}{=} \sum_t C_t$ is divided among consumers, in a way defined by the provider. If we denote by b_n the bill paid by user n for the day, then we have $\mathcal{C} = \sum_n b_n$. In general, b_n can depend on the induced costs $(C_t)_t$ and the load vector of each user $(x_{n,t})_{n,t}$, where $x_{n,t} \stackrel{\text{def}}{=} \sum_a x_{na,t}$. In this work we are interested in two different billing models given in Sections 2.4 and 2.5.

Finally, each user will try to minimize his bill b_n while respecting his constraints (2.1), by solving the problem:

$$\min_{\boldsymbol{x}_n \in \mathcal{X}_n} b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \tag{2.3}$$

where $\boldsymbol{x}_n \stackrel{\text{def}}{=} (x_{na,t})_{a,t}$ and $\boldsymbol{x}_{-n} = (\boldsymbol{x}_m)_{m \neq n}$ stands for the consumption vector of all users but n. As b_n depends on both $\boldsymbol{x}_n$ and $\boldsymbol{x}_{-n}$, this problem can be stated in the framework of game theory. We refer the reader to [FT91] for background. With $\mathcal{N}$ denoting the set of players, $\mathcal{X} \stackrel{\text{def}}{=} \prod_{n \in \mathcal{N}} \mathcal{X}_n$ the set of pure strategies and $(b_n)_n$ the vector of bills, the game is formulated under normal form by $\mathcal{G} = (\mathcal{N}, \mathcal{X}, (b_n)_n)$.

2.3 Measuring efficiency and fairness

2.3.1 Efficiency and the Price of Anarchy

We define the social cost of a load solution $\boldsymbol{x} = (\boldsymbol{x}_n)_{n\in_{\mathcal{N}}}$ as the sum of the bills in the population, that is:

$$\mathrm{SC}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \,. \tag{2.4}$$

Since we assume that the total system cost $\mathcal{C}$ is shared among the users, we also have the equality:

$$\mathcal{C}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} C_t(X_t) = \mathrm{SC}(\boldsymbol{x}) \,, \tag{2.5}$$

with $X_t \stackrel{\text{def}}{=} \sum_{a,n} x_{na,t}$ the total load at time t.

In game-theoretic models, a desirable stability property is when each player n has no interest to deviate unilaterally from her current profile $\boldsymbol{x}_n$. This corresponds to the notion of Nash Equilibrium (NE), that is:

Definition 2.1 (Nash [Nas50])**. Nash Equilibrium (NE).**
$(\hat{\boldsymbol{x}}_n)$ is a NE of the minimization game $\mathcal{G} = (\mathcal{N}, \mathcal{X}, (b_n)_n)$ iff for any player $n \in \mathcal{N}$:

$$\forall \boldsymbol{x}_n \in \mathcal{X}_n,\ b_n(\hat{\boldsymbol{x}}_n, \hat{\boldsymbol{x}}_{-n}) \leqslant b_n(\boldsymbol{x}_n, \hat{\boldsymbol{x}}_{-n}) \,.$$

An NE situation does not necessarily correspond to an optimal situation in terms of social cost. The efficiency of a mechanism is usually measured in game theory by the ratio of the optimal social cost of the system $\mathrm{SC}^* \stackrel{\text{def}}{=} \inf_{\boldsymbol{x}\in\mathcal{X}} \mathrm{SC}(\boldsymbol{x})$ and the social cost of the worst Nash Equilibrium:

Definition 2.2 ([KP99])**.** *Price of Anarchy.*
Given a game $\mathcal{G}$ and $\mathcal{X}_{\mathcal{G}}^{\mathrm{NE}}$ its set of Nash Equilibria, the price of anarchy of $\mathcal{G}$ is given as:

$$\mathrm{PoA}(\mathcal{G}) \stackrel{\text{def}}{=} \frac{\sup_{\boldsymbol{x}\in\mathcal{X}_{\mathcal{G}}^{\mathrm{NE}}} \mathrm{SC}\,(\boldsymbol{x})}{\mathrm{SC}^*} \,, \tag{2.6}$$

where $\mathrm{SC}^* \stackrel{\text{def}}{=} \inf_{\boldsymbol{x}\in\mathcal{X}} \mathrm{SC}(\boldsymbol{x})$ is the optimal social cost of the system.

The notion of Price of Anarchy has been widely studied in congestion and routing games (see [Rou16],[JMT05],[JT06]). Theoretical bounds have been established in particular frameworks (e.g. congestion games in [CK05],[Rou06],[Rou15]).

2.3.2 Fairness and Marginal Cost Pricing

To design a fair mechanism, the bill b_n paid by each user n should reflect the cost user n induces to the system, what we call the externality of n. Precisely, we denote by:

$$\mathcal{C}^*_{\mathcal{M}} \stackrel{\text{def}}{=} \inf_{(\boldsymbol{x}_m)_{m\in\mathcal{M}}} \textstyle\sum_{t\in\mathcal{T}} C_t\left(\sum_{m\in\mathcal{M}} x_{m,t}\right) \tag{2.7}$$

the optimal system cost that can be achieved with the users in the set $\mathcal{M} \subset \mathcal{N}$ while respecting their constraints. The externality of user n is the difference between the optimal system cost achieved with n in the population and the optimal system cost that can be achieved without n, that is,

$$V_n \stackrel{\text{def}}{=} \mathcal{C}^*_{\mathcal{N}} - \mathcal{C}^*_{\mathcal{N}\setminus\{n\}} \,.$$

The quantity V_n is not necessarily proportional to the total energy the user asks per day, as the load distribution between peak and off-peak hours also impacts the system cost.

To be fair, the bill of user n should be proportional to V_n [BNMR12]. This motivates the introduction of the following mechanism:

$$b_n^{\mathrm{VCG}}(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} C_t\left(\sum_{m\in\mathcal{N}} x_{m,t}\right) - \mathcal{C}^*_{\mathcal{N}\setminus\{n\}} \tag{2.8}$$

which, as noticed in [Sam+12], corresponds to a Vickrey-Clarke-Groves mechanism (VCG, see [Cla71]). In particular, it minimizes the system cost, which implies that in this model at an equilibrium $\boldsymbol{x}^{\mathrm{NE}}$ we will have $\forall n, b_n^{\mathrm{VCG}}(\boldsymbol{x}^{\mathrm{NE}}) = V_n$. The authors in [MS01] defined, in a more general framework, this pricing as Marginal Cost Pricing, and showed (Prop. 3) that it is the unique VCG mechanism that satisfies reasonable conditions. However, the mechanism (2.8) does not recover the system cost $\mathcal{C}^*_{\mathcal{N}}$, and should be renormalized as $b_n^{\mathrm{F}} \stackrel{\text{def}}{=} \frac{b_n^{\mathrm{VCG}}}{\sum_m V_m}\mathcal{C}^*_{\mathcal{N}}$. Although being centralized and hardly tractable, the billing mechanism b_n^{F} is efficient (PoA=1) and fair ($b_n^{\mathrm{F}} \propto V_n$) and we take it as a reference, following [Bah+13],[BH14], to define a fairness measure of any billing mechanism:

Definition 2.3 ([BH14])**.** *Fairness Index.*

The fairness index of a billing mechanism $(b_n)_n$ is its maximal normalized distance to $(V_n)_n$ (or equivalently to $(b_n^{\mathrm{F}})_n$) at a Nash Equilibrium:

$$F \stackrel{\text{def}}{=} \sup_{\boldsymbol{x}\in\mathcal{X}_{\mathcal{G}}^{\mathrm{NE}}} \left[\sum_{n\in\mathcal{N}} \left| \frac{V_n}{\sum_{m\in\mathcal{N}} V_m} - \frac{b_n(\boldsymbol{x})}{\sum_{m\in\mathcal{N}} b_m(\boldsymbol{x})} \right|\right]. \tag{2.9}$$

In [BH14], the authors notice the link between V_n and the notion of Shapley Value [Sha53] defined for cooperative games. However, since the Shapley Value is given by a combinatorial formula involving all possible coalitions within $\mathcal{N}$, it becomes quickly untractable as the cardinality of $\mathcal{N}$ grows. It is therefore more appropriate to use $(V_n)_n$ as reference.

2.4 Daily Proportional Billing: social optimality

2.4.1 Daily Proportional (DP) billing: definition

In this section, we recall the standard billing mechanism of [MR+10]. Consumers share the total cost of the system proportionally to their total consumption over the day. More precisely, if we denote by $E_n = \sum_{a\in\mathcal{A}_n} E_{na}$ the total energy needed by n, the bill of this user is:

$$b_n^{\mathrm{DP}}(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \stackrel{\text{def}}{=} \frac{E_n}{\sum_{m\in\mathcal{N}} E_m} \sum_{t\in\mathcal{T}} C_t(X_t)\,. \tag{2.10}$$

2.4.2 Properties

As all users minimize $\mathrm{SC}(\boldsymbol{x})$ up to a constant factor, several properties follow (detailed proofs are in [MR+10]). First, [MR+10, Thm. 1] ensures that a Nash Equilibrium (NE) exists and, as cost functions $(C_t)_t$ are assumed strictly convex, it is unique in terms of aggregated load $(X_t)_t$. The NE minimizes the social cost SC ([MR+10, Thm. 2]). In order to compute the NE in the game, the authors in [MR+10] consider the implementation of Best Response Dynamic, that can be defined as follows:

Definition 2.4. *Best Response Dynamic (BRD).*

In BRD, at each iteration k, a user n is randomly chosen, and solves his local optimization problem (2.3) with knowledge of the load of others $\boldsymbol{x}_{-n}^{(k-1)}$, taken as a parameter. The resulting load $\boldsymbol{x}_n^*$ is used to update $\boldsymbol{x}_n^{(k)} = \boldsymbol{x}_n^*$ and $\boldsymbol{x}_{-n}^{(k)} = \boldsymbol{x}_{-n}^{(k-1)}$.

In practice, we only need the aggregated load $\sum_{m\neq n} x_{m,t}$ to solve user n's problem, which can help to protect the privacy of the users (we refer the reader to Chapter 1 for more insight). Again, due to the proportionality of the users objectives, the BRD will converge to the NE [MR+10, Thm. 3]. Finally, [MR+10, Thm. 4] ensures that no user n can reduce his bill b_n by giving wrong information about his load $(x_{n,t})_t$ during the process.

2.4.3 Explicit DP solution without power bounds and quadratic costs

In this section, we focus on a particular case where we can compute explicitly the equilibrium profile of each player. For that, we consider the total energy demand for all appliances of

each consumer, and consider that there is no power upper bound in the constraints of each consumer. We consider that the cost function associated to each time period is quadratic. More precisely, the social cost minimization problem is given by:

$$\min_{\boldsymbol{x}_n} \sum_n b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) = \sum_t C_t\left(X_t\right) \tag{2.11a}$$

$$0 \leqslant x_{n,t},\ \forall n, \forall t \tag{2.11b}$$

$$\sum_{t\in\mathcal{T}} x_{n,t} = E_n,\ \forall n \tag{2.11c}$$

where $C_t(X_t) = \alpha_t X_t + \beta_t {X_t}^2$ (no constant term, $\forall t, \delta_t = 0$).

We know from [MR+10] that the equilibrium profiles correspond to the solution of (2.11). The optimal hourly load can be computed explicitly as given in Proposition 2.1:

Proposition 2.1. *Let us consider quadratic cost functions: $\forall t \in \mathcal{T}, C_t(X_t) = \alpha_t X_t + \beta_t {X_t}^2$ with $\beta_t > 0$, and, as in Proposition 2.1, w.l.o.g, let us assume that $\alpha_1 \leqslant \alpha_2 \ldots \leqslant \alpha_T$ and define, for all t:*

$$\kappa_t \stackrel{\text{def}}{=} \frac{1}{2} \sum_{k<t} \frac{\alpha_t - \alpha_k}{\beta_k}\ , \tag{2.12}$$

then $0 \stackrel{\text{def}}{=} \kappa_1 \leqslant \kappa_2 \leqslant \ldots \leqslant \kappa_T < \kappa_{T+1} \stackrel{\text{def}}{=} \infty$. Let $\bar{t}$ the unique integer in $\{1, \ldots, T\}$ such that $E \in [\kappa_{\bar{t}}, \kappa_{\bar{t}+1}]$. Then the support of the optimal aggregate profile $\boldsymbol{X}^*$ is:

$$\mathcal{T}_1 = \{1, 2, \ldots, \bar{t}\}\ ,$$

and the optimal aggregate profile $(X_t^*)_{t\in\mathcal{T}}$ is given, for all $t \in \mathcal{T}_1$ by:

$$X_t^* = \frac{1}{\sum_{k\in\mathcal{T}_1} \frac{\beta_t}{\beta_k}} \left[\frac{1}{2} \left(\sum_{k\in\mathcal{T}_1\setminus\{t\}} \frac{\alpha_k - \alpha_t}{\beta_k} \right) + E \right]\ . \tag{2.13}$$

Proof of Proposition 2.1. We denote by $\nu_{n,t}$ the Lagrangian multiplier associated to the constraint (2.11b) and λ_n the one associated with constraint (2.11c). The Lagrangian function associated to (2.11) is:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_t C_t\left(X_t\right) - \sum_n \lambda_n \Big(\sum_{t\in\mathcal{T}} x_{n,t} - E_n \Big) - \sum_{n,t} x_{n,t}\nu_{n,t}\ . \tag{2.14}$$

The KKT conditions ensure that for all t and n:

$$\frac{\partial \mathcal{L}}{\partial x_{n,t}} = C_t'\left(X_t\right) - \lambda_n - \nu_{n,t} = 0. \tag{2.15}$$

Let us start by the simple case where $x_{n,t} > 0$ for each $n \in \mathcal{N}, t \in \mathcal{T}$. Thus we have $\nu_{n,t} = 0$ and $\forall t \in \mathcal{T}, n \in \mathcal{N}$, $C_t'(X_t)$, is equal to λ_n. Thus, for all $t, k \in \mathcal{T}$:

$$\alpha_t + 2\beta_t X_t = \alpha_k + 2\beta_k X_k\ , \tag{2.16}$$

and $\boldsymbol{X}^*$ is the solution of $\boldsymbol{A}\boldsymbol{X}^* = \boldsymbol{b}$ with:

$$\boldsymbol{A} \stackrel{\text{def}}{=} \begin{pmatrix} \beta_1 & -\beta_2 & \ldots & & \\ 0 & \beta_2 & -\beta_3 & & \\ & & & & \\ 0 & & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \beta_{T-1} & -\beta_T \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix},\ \boldsymbol{b} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\alpha_2 - \alpha_1}{2} \\ \vdots \\ \frac{\alpha_T - \alpha_{T-1}}{2} \\ E \end{pmatrix} \tag{2.17}$$

We can compute $\det(A) = \sigma_{T-1}(\boldsymbol{\beta})$ with σ_k denoting the k^{th} symmetric function $\sigma_{T-1}(\boldsymbol{\beta}) = \sum_{t\in\mathcal{T}} \prod_{t'\neq t} \beta_{t'}$. By inverting the matrix $\boldsymbol{A}$, we get the expression (2.13).

We remark that the positiveness of X_t^* in (2.13) is verified if:

$$\left(\sum_{\substack{k\in\mathcal{T}\\k\neq t}}\left(\prod_{\substack{k'\in\mathcal{T}\\k'\neq t,k}}\beta_{k'}\right)\alpha_k\right)+\left(\prod_{\substack{k\in\mathcal{T}\\k\neq t}}\beta_k\right)2E>\sigma_{T-2}\big((\beta_k)_{k\neq t}\big)\,\alpha_t\,. \tag{2.18}$$

In the general case, let $\bar{t}\in\mathcal{T}$ be such that $E\in[\kappa_{\bar{t}},\kappa_{\bar{t}+1}]$ and let $\bar{\mathcal{T}}\overset{\text{def}}{=}\{1,\dots,\bar{t}\}$. Let us show that the support of $\boldsymbol{X}^*$ is given by $\mathcal{T}_1=\bar{\mathcal{T}}$ as stated in Proposition 2.1. Let us define the profile:

$$\forall t\in\bar{\mathcal{T}},\ X_t^*\overset{\text{def}}{=}\frac{1}{\sum_{k\in\mathcal{T}_1}\frac{\beta_t}{\beta_k}}\left[\frac{1}{2}\left(\sum_{k\in\mathcal{T}_1\setminus\{t\}}\frac{\alpha_k-\alpha_t}{\beta_k}\right)+E\right],\ \text{and}\ \forall t\notin\bar{\mathcal{T}},\ X_t^*\overset{\text{def}}{=}0\,, \tag{2.19}$$

and check that the KKT conditions and complementarity conditions hold.

For t such that $X_t^*>0$, the multiplier λ associated to total energy is given by the marginal cost at t:

$$\lambda=C_t'(X_t^*)=\alpha_t+2\beta_tX_t^*=\left(\sum_{k\in\mathcal{T}_1}\frac{1}{\beta_k}\right)^{-1}\left[2E+\sum_{k\in\mathcal{T}_1}\frac{\alpha_k}{\beta_k}\right], \tag{2.20}$$

and besides, we have to ensure the positivity and complementarity of multipliers:

$$\forall t\in\bar{\mathcal{T}},\ \nu_t^*=0,\ \forall t\notin\bar{\mathcal{T}},\ \nu_t=C_t'(0)-\lambda=\alpha_t-\left(\sum_{k\in\mathcal{T}_1}\frac{1}{\beta_k}\right)^{-1}\left[2E+\sum_{k\in\mathcal{T}_1}\frac{\alpha_k}{\beta_k}\right]\geqslant 0\,.$$

First, we observe that for any $t,h\in\mathcal{T}$ with $t<h$, we have:

$$2\kappa_h=\sum_{k<h}\frac{\alpha_h-\alpha_k}{\beta_k}=\sum_{k<t}\frac{\alpha_h-\alpha_k}{\beta_k}+\sum_{t\leqslant k<h}\frac{\alpha_h-\alpha_k}{\beta_k}\geqslant\sum_{k<t}\frac{\alpha_h-\alpha_k}{\beta_k}\geqslant\sum_{k<t}\frac{\alpha_t-\alpha_k}{\beta_k}=2\kappa_t\,.$$

Then, for $h\in\bar{\mathcal{T}}$ and by the definition of $\bar{h}$, we have:

$$\frac{1}{2}\left(\sum_{k\in\bar{\mathcal{T}}\setminus\{h\}}\frac{\alpha_k-\alpha_t}{\beta_k}\right)+E\geqslant 0\,,$$

such that for $t\in\bar{\mathcal{T}}$, X_t^* as defined in (2.19) is positive. Besides, for any $t\notin\bar{\mathcal{T}}$ (that is $t\geqslant\bar{t}+1$), we have:

$$\begin{aligned}
\nu_t&\overset{\text{def}}{=}\alpha_t-\left(\sum_{k\in\bar{\mathcal{T}}}\frac{1}{\beta_k}\right)^{-1}\left[2E+\sum_{k\in\bar{\mathcal{T}}}\frac{\alpha_k}{\beta_k}\right]\\
&\geqslant\left(\sum_{k\in\bar{\mathcal{T}}}\frac{1}{\beta_k}\right)^{-1}\left[-2E+\sum_{k\in\bar{\mathcal{T}}}\frac{\alpha_t-\alpha_k}{\beta_k}\right]\\
&\overset{\alpha_t\geqslant\alpha_{\bar{t}+1}}{\geqslant}\left(\sum_{k\in\bar{\mathcal{T}}}\frac{1}{\beta_k}\right)^{-1}\left[-2E+\sum_{k\in\bar{\mathcal{T}}}\frac{\alpha_{\bar{t}+1}-\alpha_k}{\beta_k}\right]\\
&=\left(\sum_{k\in\bar{\mathcal{T}}}\frac{1}{\beta_k}\right)^{-1}\left[-2E+2\kappa_{\bar{t}+1}\right]\geqslant 0\,,
\end{aligned}$$

which terminates the proof of Proposition 2.1. □

2.5 Hourly Proportional Billing: Fairness

2.5.1 Hourly Proportional (HP) billing: definition

Here, the total cost is divided between consumers at each time period, according to the load they asked at this time period. Intuitively, this enables to bring to each user the real cost of its demand, in particular during peak hours. More formally, the bill of user n is:

$$b_n^{\mathrm{HP}}(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} \frac{x_{n,t}}{\sum_{m\in\mathcal{N}} x_{m,t}} C_t(X_t)\ . \tag{2.21}$$

This billing mechanism was already formulated in [BNMR12], [Bah+13], [BH14]. In the latter, the authors show numerically that it is fairer (according to indicator F (2.9)) than the billing mechanism (2.10).

2.5.2 Properties

With payments (2.21), the game $\mathcal{G} = \left(\mathcal{N}, \mathcal{X}, (b_n^{\mathrm{HP}})_n\right)$ has the following properties.

Theorem 2.1. *Let $c_t(X_t) \stackrel{\text{def}}{=} \frac{1}{X_t} C_t(X_t)$ be the per-unit price of electricity. If $c_t' \geqslant 0$, i.e. prices are increasing with global load, then a Nash Equilibirum exists. If, in addition:*

$$\forall t, \frac{(X_t)^2}{\sum_n (x_{n,t})^2} > \left(\frac{X_t c_t''(X_t)}{2c_t'(X_t)}\right)^2 \tag{2.22}$$

then the Nash Equilibrium is unique.

Note that if the load is close to uniform, we have $\frac{x_{n,t}}{X_t} \simeq \frac{1}{N}$ so ${X_t}^2/\sum {x_{n,t}}^2 \simeq N$, and Assumption (2.22) is satisfied as soon as the network has enough users.

We refer the reader to Theorem 3.1 in Chapter 3 for the proof.

In general, the Nash Equilibrium does not achieve social optimality. However, the following result provides a bound on the Price of Anarchy (Definition 2.2).

Theorem 2.2. *In the quadratic case (2.2) with no constant coefficient ($\forall t, \delta_t = 0$), the price of anarchy is bounded:*

$$\mathrm{PoA} \leqslant 1 + \frac{3}{4} \sup_{t\in\mathcal{T}} \frac{1}{1 + \alpha_t/(\beta_t \overline{X}_t)}\ , \tag{2.23}$$

where $\overline{X}_t \stackrel{\text{def}}{=} \sum_n \overline{x}_{n,t}$.

We refer the reader to Chapter 3 and Theorem 3.3 where this result is recalled for the proof.

2.5.3 Explicit HP solution without power bounds and quadratic costs

As in Section 2.4.3, we can compute explicitly the equilibrium when the costs are quadratic and with the additional assumption that agents have the *same energy demand*, as shown in proposition Proposition 2.2 below. In the general case, it is difficult to obtain such an explicit expression of the equilibrium profiles, due to the combinatorial choices of the different cases of positivity of the multipliers.

Proposition 2.2. *Let us consider quadratic cost functions: $\forall t \in \mathcal{T}, C_t(X_t) = \alpha_t X_t + \beta_t {X_t}^2$ with $\beta_t > 0$, and, as in Proposition 2.1, w.l.o.g, let us assume that $\alpha_1 \leqslant \alpha_2 \ldots \leqslant \alpha_T$.*

Let assume that all agents have the same energy demand, that is $E_n = E_m$ for each $n, m \in \mathcal{N}$. Then, with $(\kappa_t)_t$ as defined in (2.12) above, let $\bar{t}$ be the unique integer in $\{1, \ldots, T\}$ such that $E \in [\frac{2N}{N+1}\kappa_{\bar{t}}, \frac{2N}{N+1}\kappa_{\bar{t}+1}]$. Then the support of the aggregate equilibrium profile $\hat{\boldsymbol{X}}$ is:

$$\mathcal{T}_1 = \{1, 2, \ldots, \bar{t}\}\ ,$$

and this HP aggregatepn equilibrium profile $\hat{\boldsymbol{X}}$ is given for all $t \in \mathcal{T}_1$ by:

$$\hat{X}_t = \frac{1}{\sum_{k\in\mathcal{T}_1} \frac{\beta_t}{\beta_k}} \left[\frac{N}{N+1} \left(\sum_{k\in\mathcal{T}_1\setminus\{t\}} \frac{\alpha_k - \alpha_t}{\beta_k} \right) + E \right] . \tag{2.24}$$

Proof. First, we know from [Nas50] that, as there is a unique Nash Equilibrium, it is necessarily symmetric for players. At the equilibrium, each user solves the quadratic program:

$$\min_{\boldsymbol{x}_n} b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) = \sum_t x_{n,t} c_t(X_t) \tag{2.25a}$$

$$0 \leqslant x_{n,t}, \ \forall t \tag{2.25b}$$

$$\sum_{t\in\mathcal{T}} x_{n,t} = E_n \ . \tag{2.25c}$$

We still denote by $\nu_{n,t}$ the Lagrangian multiplier associated to (2.25b) and λ_n the one associated with constraint (2.25c). The Lagrangian function associated to (2.25) is:

$$\mathcal{L}_n(\boldsymbol{x}_n, \boldsymbol{\nu}_n, \lambda_n) = \sum_t x_{n,t} c_t(X_t) - \lambda_n \left(\sum_{t\in\mathcal{T}} x_{n,t} - E_n \right) - \sum_{n,t} x_{n,t} \nu_{n,t}. \tag{2.26}$$

The KKT conditions state that, for all $t \in \mathcal{T}$, we have:

$$\frac{\partial \mathcal{L}_n}{\partial x_{n,t}} = x_{n,t} c_t'(X_t) + c_t(X_t) - \lambda_n - \nu_{n,t} = 0. \tag{2.27}$$

Besides, from the complementarity conditions $\forall n, t, \ x_{n,t} > 0 \iff \nu_{n,t} = 0$.

To simplify, let us make the assumption that the equilibrium is an interior point, in the sense that $\forall n \in \mathcal{N}, \ \forall t \in \mathcal{T}, \ x_{n,t} > 0$ (in this case, we have $\mathcal{T}_1 = \mathcal{T}$).

We then deduce that for all $t, k \in \mathcal{T}$ and for all $n \in \mathcal{N}$:

$$\begin{aligned} & x_{n,t} c_t'(X_t) + c_t(X_t) = x_{n,k} c_k'(X_k) + c_k(X_k) \\ \iff & \beta_t(2x_{n,t} + X_{-n,t}) - \beta_k(2x_{n,k} + X_{-n,k}) = \alpha_k - \alpha_t \ , \end{aligned}$$

where $X_{-n,t} \overset{\text{def}}{=} \sum_{m\neq n} x_{m,t}$. The equilibrium point $(\hat{x}_{n,t})_{n,t}$ is therefore given as the solution of the linear system:

$$\boldsymbol{A}_{N,T} \boldsymbol{x} = \boldsymbol{b}_{N,T}$$

with $\boldsymbol{A}_{N,T}$ the $NT \times NT$ matrix given by:

$$\begin{pmatrix} \beta_1(I_N + J_N) & -\beta_2(I_N + J_N) & 0 & \dots & 0 \\ 0 & \beta_2(I_N + J_N) & -\beta_3(I_N + J_N) & 0 & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \beta_{T-1}(I_N + J_N) & -\beta_T(I_N + J_N) \\ I_N & I_N & \dots & I_N & I_N \end{pmatrix}$$

where I_N is the identity matrix and J_N the matrix full of ones, of size $N \times N$. The vector $\boldsymbol{b}_{N,T}$ is given by:

$$\boldsymbol{b}_{N,T} \overset{\text{def}}{=} \begin{pmatrix} (\alpha_2 - \alpha_1)\mathbb{1}_N \\ (\alpha_3 - \alpha_2)\mathbb{1}_N \\ \vdots \\ (\alpha_T - \alpha_{T-1})\mathbb{1}_N \\ (E_n)_{n\in\mathcal{N}} \end{pmatrix}^T \text{ with } \mathbb{1}_N eqd \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^N.$$

Using the results of Section 2.4.3, we can compute that:

$$\det(\boldsymbol{A}_{N,T}) = (N+1)^{T-1} \left(\sigma_{T-1}(\boldsymbol{\beta})\right)^N .$$

By inverting the matrix $\boldsymbol{A}_{N,T}$, we finally get that, for all n, t:

$$\hat{x}_{n,t} = \frac{(N+1)^{T-2}\left(\sigma_{T-1}(\boldsymbol{\beta})\right)^{N-1}}{(N+1)^{T-1}\left(\sigma_{T-1}(\boldsymbol{\beta})\right)^{N}}\Bigg[\left(\sum_{k\in\mathcal{T}\setminus\{t\}}\left(\prod_{k'\in\mathcal{T}\setminus\{t,k\}}\beta_{k'}\right)\alpha_k\right) + \left(\prod_{k\in\mathcal{T}\setminus\{t\}}\beta_k\right)(N+1)E_n - \sigma_{T-2}\big((\beta_k)_{k\neq t}\big)\,\alpha_t\Bigg], \tag{2.28}$$

leading to:

$$\hat{x}_{n,t} = \frac{1}{(N+1)\sigma_{T-1}(\boldsymbol{\beta})}\Bigg[\left(\sum_{k\in\mathcal{T}\setminus\{t\}}\left(\prod_{k'\in\mathcal{T}\setminus\{t,k\}}\beta_{k'}\right)\alpha_k\right) + \left(\prod_{k\in\mathcal{T}\setminus\{t\}}\beta_k\right)(N+1)E_n - \sigma_{T-2}\big((\beta_k)_{k\neq t}\big)\,\alpha_t\Bigg] \tag{2.29}$$

$$= \frac{1}{\sum_{k\in\mathcal{T}}\frac{\beta_t}{\beta_k}}\left[\frac{1}{N+1}\left(\sum_{k\in\mathcal{T}\setminus\{t\}}\frac{\alpha_k-\alpha_t}{\beta_k}\right)+E_n\right]. \tag{2.30}$$

In the case where $\mathcal{T}_1$ is a proper subset of $\mathcal{T}$, we can adapt the proof to obtain an expression similar to (2.30) where $\mathcal{T}$ is replaced by $\mathcal{T}_1$. Then, as for the proof of Proposition 2.1, we can conclude that the KKT optimality conditions of (2.25) are verified by the obtained profiles, and finally obtain the aggregate load $\hat{X}_t$ at period t at the equilibrium as stated in Proposition 2.2. □

Comparing the HP equilibrium expression (2.24) with the optimal load (DP) computed in (2.13), we get immediately the two following results:

Proposition 2.3. *With quadratic costs, the equilibrium associated to HP is socially optimal if and only if the coefficients of the linear part of the costs functions $(\alpha_t)_t$ are uniform.*

Proposition 2.4. *If the provider gives modified coefficients $(\hat{\alpha}_k)_k$, defined as:*

$$\hat{\alpha}_k \overset{\text{def}}{=} \frac{N+1}{2N}\alpha_k,$$

to the users, then the equilibrium achieved in the modified game corresponds to the social optimum in the initial game.

We can wonder if the two propositions still hold in the case where bounding constraints are active at the equilibrium.

Proposition 2.3 still holds because, if the linear part is uniform (that is $\forall t, \alpha_t$ is a constant) then the quantity in (2.30) is still positive, hence it still defines an equilibrium.

However, the optimality given by Proposition 2.4 does not hold in general, as shown in the Example 2.1 below.

Example 2.1. Let $\mathcal{T} = \{P, O\}$ and $\mathcal{N} = \{1, 2\}$ (2 players) with $E_1 = 0.5$ and $E_2 = 2$. Let us define the costs:

$$\alpha_O = 0,\ \alpha_P = 1,\ \beta_O = \beta_P = 1\,.$$

The modified costs are the same except $\hat{\alpha}_P = \frac{N+1}{2N}\alpha_P = \frac{3}{4}$.
The optimal profile is given by (2.13) and corresponds to the loads:

$$X_P^* = 1,\ X_O^* = \frac{3}{2}\,. \tag{2.31}$$

The expression of the equilibrium profile cannot be given by (2.30), otherwise $x_{1,P}$ would be negative. Let us make the assumption that $x_{1,O} = E_1,\ x_{1,P} = 0$ and $x_{2,P}, x_{2,O} > 0$. Then,

from $\frac{\partial \mathcal{L}_2}{\partial x_{2,P}} = \frac{\partial \mathcal{L}_2}{\partial x_{2,O}} = 0$, we obtain:

$$\begin{pmatrix} 2\beta_O & -2\beta_P \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{2,O} \\ x_{2,P} \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_P - \hat{\alpha}_O - \beta_O E_1 \\ E_2 \end{pmatrix}$$
$$\Leftrightarrow \begin{pmatrix} x_{2,O} \\ x_{2,P} \end{pmatrix} = \begin{pmatrix} \left(\frac{\beta_O}{\beta_P} + 1\right)^{-1} \left[\frac{1}{2}\frac{\hat{\alpha}_P - (\hat{\alpha}_O + \beta_O E_1)}{\beta_P} + E_2\right] \\ \left(\frac{\beta_P}{\beta_O} + 1\right)^{-1} \left[\frac{1}{2}\frac{(\hat{\alpha}_O + \beta_O E_1) - \hat{\alpha}_P}{\beta_O} + E_2\right] \end{pmatrix} .$$

This expression is similar to (2.30) where α_O is replaced by $\hat{\alpha}_O + \beta_O E_1$. With the values given above, we get the equilibrium values $\hat{x}_{2,O} = \frac{17}{16}$ and $\hat{x}_{2,P} = \frac{15}{16}$.

We show that $\boldsymbol{x}_1 = (E_1, 0)$ is the equilibrium profile for player 1 by verifying that the multiplier $\nu_{1,P}$ associated to constraint $x_{1,P} \geqslant 0$ is positive:

$$\begin{aligned} \nu_{1,P} = \lambda_1 - c_P(\hat{X}_P) &= E_1 c_O'(\hat{X}_O) + c_O(\hat{X}_O) - c_P(\hat{X}_P) \\ &= E_1\beta_O + \hat{\alpha}_O + \beta_O(E_1 + \hat{x}_{2,O}) - \hat{\alpha}_P - \beta_P \hat{x}_{2,P} = \frac{3}{8} \geqslant 0 . \end{aligned}$$

Finally, we obtain the aggregate load corresponding to the equilibrium:

$$\hat{X}_O = \frac{17}{16} + \frac{1}{2} , \hat{X}_P = \frac{15}{16} ,$$

which differs from the optimal profile (2.31). Moreover, as the expressions of $\hat{\boldsymbol{x}}_2$ depends on E_1, one can not find some cost coefficients $\hat{\alpha}_P, \hat{\alpha}_O$ that will ensure that the equilibrium will correspond to the optimal profile for any $E_1 > 0$!

Note that in Example 2.1, we took $E_2 > E_1$ and this is indeed necessary to obtain a counter example to Proposition 2.4. If all customers have the same energy demand (E_n is a constant), and there are no power bounds, then Proposition 2.2 shows that the property of optimality with the modified costs stated in Proposition 2.4 holds.

We can go further and obtain an explicit PoA for this specific case of interior equilibrium. To begin with, we can easily analyze the difference between the equilibrium (2.24) and the social optimum profile (2.13), as for all t:

$$\hat{X}_t - X_t^* = \frac{N-1}{2N+2} \frac{1}{\sum_{k\in\mathcal{T}} \frac{\beta_t}{\beta_k}} \left(\sum_{k\in\mathcal{T}\setminus\{t\}} \frac{\alpha_k - \alpha_t}{\beta_k} \right) , \tag{2.32}$$

$$\left|\hat{X}_t - X_t^*\right| \leqslant \frac{N-1}{2N+2} \frac{\sigma_{T-2}((\boldsymbol{\beta}_k)_{k\neq t})}{\sigma_{T-1}(\boldsymbol{\beta})} \sup_k |\alpha_k - \alpha_t| . \tag{2.33}$$

To compute the PoA $= \mathrm{SC}(\hat{\boldsymbol{x}})/\mathrm{SC}^*$, we start by obtaining simple expressions for $\mathrm{SC}(\hat{\boldsymbol{x}})$ and SC^*. We compute for the cross term in ${X_t}^2$:

$$\sum_t \beta_t \left(\frac{\prod_{k\in\mathcal{T}\setminus\{t\}} \beta_k}{\sigma_{T-1}(\boldsymbol{\beta})} \right)^2 \left(\sum_{k\in\mathcal{T}\setminus\{t\}} \frac{\alpha_k - \alpha_t}{\beta_k} \right) = \frac{\prod_{t\in\mathcal{T}} \beta_t}{\sigma_{T-1}(\boldsymbol{\beta})} \sum_{t,k\in\mathcal{T}^2} \frac{\alpha_k - \alpha_t}{\beta_t \beta_k} = 0 \tag{2.34}$$

and:

$$\begin{aligned}\sum_t \prod_{k\neq t}\beta_k \left(\sum_{k\in\mathcal{T}} \frac{\alpha_k-\alpha_t}{\beta_k}\right)^2 &= \left(\prod_t \beta_t\right)\sum_t\sum_{k,\ell}\frac{(\alpha_k-\alpha_t)(\alpha_\ell-\alpha_t)}{\beta_k\beta_\ell\beta_t}\\ &= \left(\prod_t \beta_t\right)\left(\sum_k \frac{\alpha_k}{\beta_k}\sum_{\ell,h}\frac{\alpha_\ell-\alpha_t}{\beta_\ell\beta_t} - \sum_k\frac{1}{\beta_k}\sum_{\ell,h}\alpha_t\frac{\alpha_\ell-\alpha_t}{\beta_\ell\beta_t}\right)\\ &= 0 - \left(\prod_t \beta_t\right)\sum_k\frac{1}{\beta_k}\sum_{\ell,h}\alpha_\ell\frac{\alpha_t-\alpha_\ell}{\beta_\ell\beta_t}\\ &= \sigma_{T-1}(\boldsymbol{\beta})\frac{1}{2}\sum_{\ell,h}\frac{(\alpha_t-\alpha_\ell)^2}{\beta_\ell\beta_t}\,.\end{aligned} \tag{2.35}$$

With $\hat{\phi} = \frac{N}{N+1}$ and $\phi^* = \frac{1}{2}$, and $\sharp$ denoting either "$\hat{}$"(equilibrium profile) or "*" (social optimum), we obtain the expression of the social cost as:

$$\mathrm{SC}^\sharp = \frac{1}{\sum_{k\in\mathcal{T}}\frac{1}{\beta_k}}\left(\frac{{\phi^\sharp}^2-\phi^\sharp}{2}\sum_{k,t\in\mathcal{T}^2}\frac{(\alpha_k-\alpha_t)^2}{\beta_k\beta_t} + \sum_t\frac{\alpha_t}{\beta_t}E + E^2\right)\,. \tag{2.36}$$

With ${\phi^*}^2-\phi^* = -1/4$ and $\hat{\phi}^2-\hat{\phi} = -N/(N+1)^2$, we get the price of Anarchy:

$$\mathrm{PoA} = 1 + \frac{\frac{1}{2}\left(\frac{1}{4}-\frac{N}{(N+1)^2}\right)\sum_{k,t\in\mathcal{T}^2}\frac{(\alpha_k-\alpha_t)^2}{\beta_k\beta_t}}{-\frac{1}{8}\sum_{k,t\in\mathcal{T}^2}\frac{(\alpha_k-\alpha_t)^2}{\beta_k\beta_t} + \sum_t\frac{\alpha_t}{\beta_t}E + E^2} \tag{2.37}$$

$$= 1 + \frac{\left(1-\frac{4N}{(N+1)^2}\right)V}{-V + 8\left(\sum_t\frac{\alpha_t}{\beta_t}E + E^2\right)}\,, \tag{2.38}$$

with $V \stackrel{\text{def}}{=} \sum_{k,t\in\mathcal{T}^2}\frac{(\alpha_k-\alpha_t)^2}{\beta_k\beta_t}$.

2.6 Numerical Experiments

We compare numerically the billing mechanisms DP (2.10) and HP (2.21), based on the two criterias of efficiency (Def. 2.2) and fairness (Def. 2.3). We extracted a set $\mathcal{N}$ of 30 users from the database *PecanStreet Inc.* [Pec], which gathers hundreds of disaggregated residential consumption profiles in Texas, U.S. We use hourly timesteps so that $\mathcal{T} = \{0,1,\ldots,23\}$. We consider that each day, just before midnight, the flexible consumption of each user for the next day is computed as an equilibrium strategy of the game, implementing a BRD algorithm (Definition 2.4), relying on a communication infrastructure between consumers and the operator for exchanging the information $\boldsymbol{x}_{-n}$. We run simulations day by day on the set of 30 days [1] $\mathcal{D} \stackrel{\text{def}}{=} \{02/01/2016,\ldots,31/01/2016\}$.

2.6.1 Flexible Appliances: Electric Vehicles and Heating

We study a population of residential consumers owning electric vehicles (EV) and electrical heating systems (furnace). EVs present an important flexibility [Bea+16] since an EV remains plugged in while it is parked, and a smart charging can be automated without impacting the user. Similarly, the initial consumption profile of a heating system can be modified without strong impact on the comfort of the household.

In our simulations, we consider a first case where EVs are the only flexible appliances, accounting for 20.4% of an average daily global energy of 1014kWh, and a second case where furnaces are also considered as flexible appliances, increasing the part of flexible energy to 25.8%. The remaining of each user's consumption is nonflexible, which we denote by $(x_{n,t}^{\mathrm{NF}})_t$.

[1] To start simulations with a working day, we dismissed January, 1st.

Figure 2.1(top) shows the repartition between flexible and nonflexible load on a typical day. The nonflexible load is more important on some hours than others, so that even with hourly uniform system cost (2.39), these hours will have bigger marginal costs.

The users constraints (2.1) are evaluated as follows: we consider two types of days: $\mathcal{D}_1$ for weekdays (Monday to Friday) and $\mathcal{D}_2$ for weekend days (Saturday and Sunday). For each type $\mathcal{D}_k$ of day and each user n, we suppose that appliance a can be used at t if it exists a day of type $\mathcal{D}_k$ where a was on at t. More precisely, for a day of type $\mathcal{D}_k$, $\mathcal{T}_{na} = \bigcup_{d\in\mathcal{D}_k}\{t : x^{\text{data}}_{n,a,d,t} > 0\}$. For simplicity, we took the min power $\underline{x}_{na,t}$ equal to 0 and the max power $\overline{x}_{na,t}$ equal to the maximal nonnegative value found on the data set $\overline{x}_{na,t} = \max_{t,d\in\mathcal{T}\times\mathcal{D}} x^{\text{data}}_{n,a,d,t}$, if $t \in \mathcal{T}_{na}$, and 0 otherwise.

2.6.2 System Cost

We consider that the provider costs are gieven as functions $(\tilde{C}_t)_t$ of the total load $X^{\text{tot}}_t \stackrel{\text{def}}{=} X^{\text{NF}}_t + X_t$, where X^{NF}_t denotes the aggregate *nonflexible* load at time period t, while X_t denotes the flexible part.

Moreover, we assume that those costs are quadratic and are uniform over time periods, and given in dollar cents by:

$$\forall t \in \mathcal{T}, \tilde{C}_t(X) = \tilde{C}(X) \stackrel{\text{def}}{=} 0.1 + 8X + 0.04(X)^2 \,. \tag{2.39}$$

The average hourly nonflexible load on all days in $\mathcal{D}$ is:

$$\langle X^{\text{NF}}_t \rangle \stackrel{\text{def}}{=} \frac{1}{|\mathcal{D}| \times |\mathcal{T}|} \sum_{d,t\in\mathcal{D}\times\mathcal{T}} \left(\sum_{n\in\mathcal{N}} x^{\text{NF}}_{n,d,t} \right) = 31.3\text{kWh}.$$

Coefficients in (2.39) are chosen arbitrarily but such that the per-unit price $\tilde{C}_t(\langle X^{\text{NF}}_t \rangle)/\langle X^{\text{NF}}_t \rangle$ given by (2.39) matches the electricity price proposed by the distributor *CoServ* [Cos] of $8.5\boldsymbol{c}$/kWh for base contracts.

We assume that the nonflexible load $(x^{\text{NF}}_{n,t})_t$ is billed in a separate process (for instance, according to a baseline contract as defined in the next subsection). We apply the proposed billing mechanisms DP (2.10) and HP (2.21) on the flexible part x_t only. Although the system costs $(\tilde{C}_t)_t$ are hourly uniform, the variation in the nonflexible load X^{NF}_t over the hours induces a variation on the *cost of flexible load* $C_t(X_t)$ over the hours, as we have:

$$\begin{aligned} \forall t \in \mathcal{T}, C_t(X_t) &\stackrel{\text{def}}{=} \tilde{C}(X^{\text{NF}}_t + X_t) - \tilde{C}(X^{\text{NF}}_t) \\ &= (8 + 0.08X^{\text{NF}}_t)X_t + 0.04(X_t)^2 \,. \end{aligned} \tag{2.40}$$

In practice, the operator could rely on a forecast of X^{NF}_t for the next day to compute the functions $(\tilde{C}_t)_t$ before sending the signal to consumers.

2.6.3 Two Reference Non-Game Theoretic Billing Models

In order to compare the formulated game-theoretic models (2.10) and (2.21) to existent non game-theoretic billing models, we also consider the two following standard models as references:

1) **Baseline billing.** No information on the global load is sent to the users, who know *a priori* that they will pay a fixed price per kWh p. Each user consumes energy without any optimization of the system costs and we consider that the consumption profile is given by the original (observed) profile of each user $(x_{n,t})_t$. The bill of a user n with a total consumption $E_n = \sum_t \sum_a x_{na,t}$ will be $b^{\text{base}}_n(\boldsymbol{x}) = b^{\text{base}}_n(\boldsymbol{x}_n) \stackrel{\text{def}}{=} p \times E_n$. As both the PoA (2.6) and F (2.9) are normalized, the choice of p has no influence at all on the values of those indicators.

2) **Peak/Offpeak billing**. This kind of contract already exists in many countries and many of the Texas electricity distributors are proposing it. The provider defines *a priori* a fixed set of peak hours $\mathcal{T}_P$ on which the prices are higher. We consider that users avoid

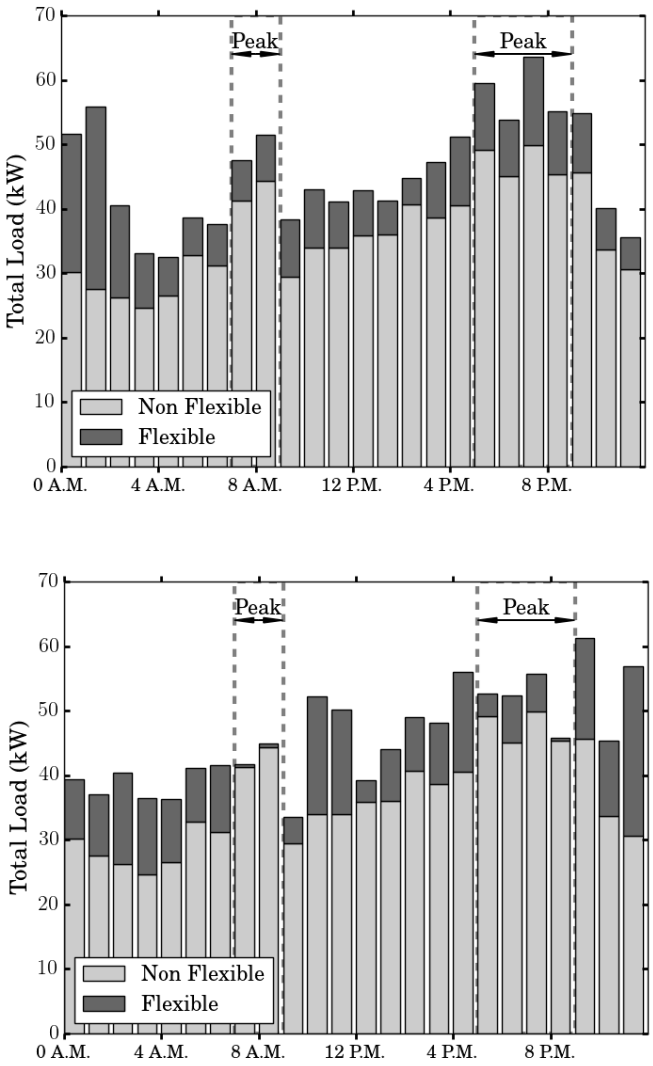


FIGURE 2.1: Aggregated load of 30 users on January 10, 2016.
top: *Observed profile from the data.* ***bottom:*** *The profile is modified to avoid peak hours in the billing model* (2.41).

peak hours as soon as their constraints enable it, by applying a simple greedy algorithm: recursively, a random offpeak hour t^{off} is chosen and the onpeak load of an appliance a is moved to t^{off} until $\overline{x}_{na,t^{\text{off}}}$ is reached. The resulting consumption profile is denoted by $(\tilde{x}_t)_t$ (see Figure 2.1(bottom) for an example). In our simulations, we define $\mathcal{T}_P$ as the set of hours where the nonflexible load is the higher on average, which gives $\mathcal{T}_P = [7\text{A.M.-}9\text{A.M.}] \cup [5\text{P.M.-}9\text{P.M.}]$. We keep the same price ratio $r^{\text{peak}} = p^{\text{peak}}/p^{\text{off}} = 2.84$ than the Texan distributor *Coserv* [Cos]. The bill of user n is:

$$b_n^{\text{Peak/Off}}(\boldsymbol{x}_n) \stackrel{\text{def}}{=} r^{\text{peak}} p^{\text{off}} \sum_{t \in \mathcal{T}_P} \tilde{x}_{n,t} + p^{\text{off}} \sum_{t \in \mathcal{T} \setminus \mathcal{T}_P} \tilde{x}_{n,t} \,. \tag{2.41}$$

As explained above, the choice of p^{off} has no influence at all on the value of the PoA and F. However, the ratio r^{peak} has a direct impact on the fairness indicator (2.9).

Flexible Items	EV only		EV + furnace	
Billing	PoA-1 (%)	F (%)	PoA-1 (%)	F (%)
HP	0.0830 (0.0772)	0.999 (0.286)	0.0886 (0.104)	1.17 (0.302)
DP	0.0 (0.0)	3.18 (1.38)	0 (0)	3.36 (1.57)
Baseline	18.8 (5.12)	3.19 (1.38)	18.5 (5.42)	3.36 (1.57)
Peak/Off	13.3 (3.17)	3.20 (1.34)	12.8 (3.69)	3.27 (1.04)

TABLE 2.1: Mean (and standard deviation) of inefficiency (PoA-1) and unfairness (F) over the days $\mathcal{D}$ and users $\mathcal{N}$.

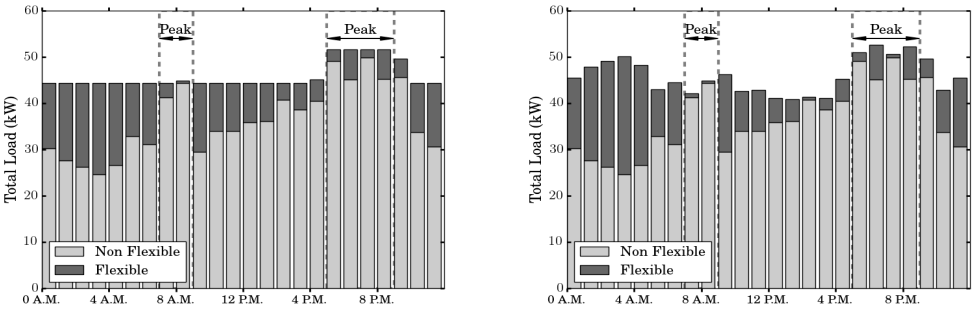


FIGURE 2.2: Equilibrium profiles for DP *(left)* and HP *(right)*, on January 10, 2016. *In DP, marginal costs are equal on all hours if the flexible load is sufficient. The HP equilibrium profile remains close to the optimal DP profile.*

2.6.4 Results

For each day in $\mathcal{D}$, we obtain the equilibrium profile with HP billing (2.21) by running the BRD (Definition 2.4). In most cases, a hundred iterations (that is, around three optimizations of (2.3) per user) were sufficient to converge to the equilibrium. The optimal profile (corresponding to the equilibrium in DP) is obtained as the solution of a quadratic program solved with the solver Cplex 12.5. The simulations were implemented in Python 3.5 and run on a single core IntelCore i7-6600U@2.6Ghz with 7.7GB of RAM. The BRD process takes around 50 seconds in average for each simulated day in $\mathcal{D}$, resulting in a total simulation time of around 20 minutes.

The inefficiency (PoA-1) and unfairness (F) induced by the four billing mechanisms, that is, DP (2.10), HP (2.21), baseline billing and Peak/OffPeak billing (2.41), are represented in Figure 2.4 for each day in $\mathcal{D}$. The precise values (mean and variance) are given in Table 2.1. In practice, the PoA of HP is one up to 10^{-3}: this billing mechanism almost reaches the optimal social cost. The equilibrium profile is not very far from the optimal load profile. Figure 2.2 shows the equilibria of the two mechanisms DP (optimal) an HP. The optimal load profile is very flat. Indeed, due to Kuhn-Tucker conditions of optimality, marginal costs are equal on all hours where constraints (2.1b) are not tight. Therefore, if the part of flexible load is large enough, the total load X_t will be the same for all hours. The equilibrium of the mechanism HP is not as flat, but it remains close to the optimal profile, due to its limited PoA.

Because of the absence of coordination among users in the non-game theoretic Peak/ Off-Peak billing mechanism, some offpeak hours become congested, as seen on Figure 2.1(bottom), resulting in high system costs. This efficiency loss is avoided by using a game mechanism as HP and DP.

We can see both from Figure 2.4 and Table 2.1 that the HP mechanism achieves an important fairness property in comparison with the other mechanisms. The associated standard deviation of F of 0.3% indicates that its fairness is also more robust than the other models. Indeed, Figure 2.3 shows the evolution of the indicator F when we relax the constraint of max power (2.1b) by scaling the value $\overline{x}_{na,t}$ chosen in Section 2.6.1 by a factor in $[0.5, 3]$. The unfairness induced by DP decreases when the constraints are relaxed, and it gets closer and closer to HP. Therefore, the HP mechanism will be much more interesting when the constraints are tight.

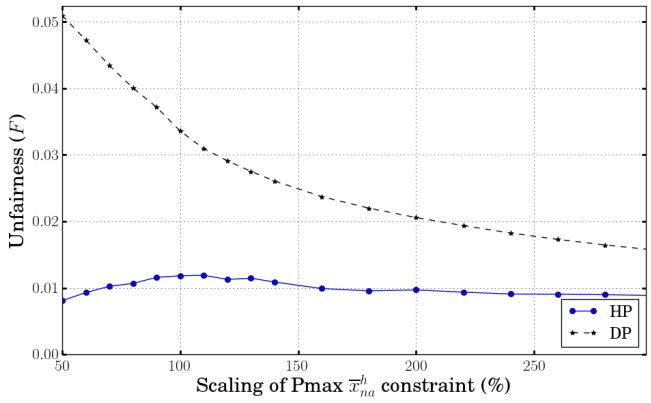


FIGURE 2.3: Evolution of unfairness in HP and DP with constraints. *When constraints* (2.1b) *are tight, the DP mechanism has a large unfairness and gets fairer when the constraints are relaxed.*

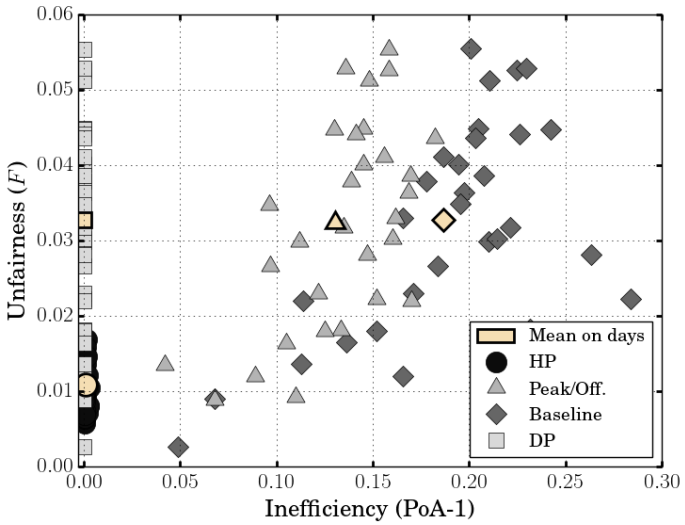


FIGURE 2.4: Comparison of billing mechanisms on the 30 days in $\mathcal{D}$ with EV charging considered as flexible. *HP has a PoA of one up to* 10^{-3}. *The non-game theoretic billings* Baseline *and* Peak/Offpeak *are dominated on average. Results are similar when we also consider heating as a flexible appliance.*

2.7 Conclusion

We gave theoretical results ensuring that the hourly proportional billing mechanism has a unique equilibrium and that its price of anarchy is bounded. Experimental results revealed that this mechanism achieves an important fairness property with our quantitative indicator (2.9), while being very close to the social optimum (up to 0.08%). We have seen that the fairness indicator of the hourly mechanism was three times smaller than the other mechanisms, associated with a low variance. As this fairness difference increases with the level of constraints, using the hourly mechanism in practice will be really interesting if the consumption is highly constrained. If we consider some utility functions in each user optimization (as in [Sam+12]), the daily proportional billing mechanism has no reason to conserve its property of social optimality, as the reader will see in Chapter 4. Therefore, the small efficiency loss of the hourly mechanism should not have any influence in practice.

Chapter 3

Analysis and Algorithms for an Hourly Billing Mechanism for Demand Response Management

This Chapter is based on the published paper [Jac+19c], except for Section 3.3.3, added in this chapter and which has not been submitted for publication. In this Section 3.3.3, we show the convergence of a randomized version of the best response algorithm in atomic splittable congestion games in two particular cases: the case of linear costs, and the case of two resources and general convex and increasing costs.

As Chapter 2 has shown that the hourly proportional billing mechanism was a valuable model in fairness and efficiency, we focus on this model in this Chapter 3. This chapter recalls some results and concepts already enounced in Chapter 2: the bound on the price of anarchy Theorem 3.3 is the same result than Theorem 2.2 above. Chapter 3 also presents some convergence results concerning the Best Response algorithm introduced in Definition 2.4.

3.1 Introduction

In game-theoretic models for Demand Response , a major issue is to define an effective procedure to compute and reach the consumption *equilibrium* associated with the game. Several papers [BNH17; Atz+13; Che+14] have investigated the complexity and algorithmic aspects associated to the notion of equilibrium.

In [Atz+13] and [Che+14], the authors consider the same billing mechanism as the one studied in this chapter, and propose decentralized methods to compute the Nash equilibrium (an iterative proximal best response in [Atz+13] and a proximal-point algorithm in [Che+14]).

In this chapter, we investigate the theoretical properties and computational aspects of the hourly billing mechanism and discuss its practical implementation.
This chapter brings five theoretical and practical contributions:

1. We give a new result on the uniqueness of the equilibrium (Theorem 3.2) under a convexity assumption. This result extends [Che+14, Prop. 1]—which relies on the general results of [Ros65]—where uniqueness is given for a particular class of price functions (of the form $c_t(x) = \alpha_t + \beta_t x^{b_t}$ with $\alpha_t \geqslant 0, \beta_t > 0$ and $b_t \geqslant 1$). In contrast, our uniqueness theorem applies to any convex and strictly increasing price functions. It extends [ORS93, Thm. 1] to a more general model of constraints where we consider upper and lower bounds on the load at each time period;

2. We give a new result on the induced *Price of Anarchy* (PoA). This result (Theorem 3.3) gives an evaluation of the equilibrium efficiency in terms of social cost. The PoA is numerically close to one but not one. To our knowledge, there are very few existing results on the PoA for this framework. A related but different result is [CK14], where the authors consider a maximization game with individual utilities, and consider the same hourly proportional billing. A bound on the PoA is obtained assuming that players individual utilities are large enough compared to the system cost. Our bound applies to the minimization game without

utilities, but is tighter asymptotically. Another related work is [Bea+16] where the authors prove that the PoA converges to one in the asymptotic case of an infinite number of players;

3. We bound the convergence rate of the Best Response (BR) algorithm in the case of affine prices (Theorem 3.4). In that case, convergence is known but, to our knowledge, no bound on the rate has ever been given. The convergence has been conjectured more generally for any convex prices [Mer08; BPS13].

4. We introduce a different algorithm: SIRD, based on a simultaneous projected gradient descent (Algorithm 3.2), and show its geometric convergence (Theorem 3.5) with a condition on the price functions only. To our knowledge, those results are also new. The proposed algorithms (BR and SIRD) and their convergence rates are compared numerically with the algorithms proposed in [Che+14] and [Atz+13]. In the case of SIRD, we allow a fix step-size and we do not need a proximal term so the convergence is faster;

5. Last, we introduce an online DR procedure with receding horizons (Algorithm 3.4), in the spirit of Model Predictive Control [Wu+11], to take into account updated forecasts in a stochastic environment. We prove that the consumption profiles computed by this procedure correspond to the desired NE in the limit of perfect forecasts (Theorem 3.6). We show numerically, based on real consumption data, that this procedure can achieve significant savings compared to an offline procedure.

This chapter is organized as follows: Section 3.2 gives the mathematical model of the DR framework and the associated billing mechanism, under the form of a game. In Section 3.3, we define two decentralized algorithms that enable to compute the equilibrium consumption profiles. We prove the convergence of those algorithms and provide upper bounds on their convergence rates. We present a numerical study of the given algorithms and compare them to two others algorithms from [Che+14] and [Atz+13]. Finally, in Section 3.4, we define an *online* DR procedure and simulate it with historical consumption data of consumers with electric vehicles as flexible consumptions. We compare the performance of this online DR scheme to the *offline* version and other consumption scenarios.

3.2 Consumption Game with Hourly Billing

3.2.1 District of flexible consumers

We consider a set $\mathcal{N} = \{1, \dots, N\}$ of residential consumers linked to a local *aggregator*. Each household is equipped with a smart meter enabling two-way communication of information with the aggregator. We assume that each household electricity consumption can be divided into two parts: one which is *nonflexible* (lights, cooking appliances, TVs) and one which is *flexible* (Electric Vehicle charging, water heating, etc). Moreover, each smart meter is linked to an *Electricity Consumption Scheduler* (ECS) that can automatically optimize and schedule the consumption profile of the consumer's flexible appliances, given the constraints set by the consumer and the physical constraints of each appliance.

3.2.2 From individual to aggregated consumption profiles

In the DR program, we determine a consumption profile for each consumer on a finite time horizon T. In this study, we take $\mathcal{T}$ as a discrete set of time periods $\mathcal{T} = \{1, \dots T\}$. In the simulations, $\mathcal{T}$ will correspond to one day, and each time period t to one hour. The aggregated flexible load profile on the set of consumers is obtained as:

$$\boldsymbol{X} = (X_t)_{t\in\mathcal{T}} \in \mathbb{R}^T \text{ with } \forall t \in \mathcal{T}, X_t \stackrel{\text{def}}{=} \sum_n x_{n,t} \,, \tag{3.1}$$

where $x_{n,t}$ denotes the flexible consumption of consumer n on time period t.

3.2.3 Aggregator objective from the aggregated consumption

The aggregator is himself linked to electricity providers and we consider that he faces a per-unit (of energy) price function $X_t \mapsto c_t(X_t)$ associated with each time period $t \in \mathcal{T}$ for the flexible electricity demand X_t given in (3.1). The total system cost for providing the flexible profile $(X_t)_{t\in\mathcal{T}}$ is then:

$$\mathcal{C}(\boldsymbol{X}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} X_t \times c_t(X_t) \,, \tag{3.2}$$

a quantity that should be minimized by the aggregator. In particular, we make the assumption that the system cost $\mathcal{C}$ is time-separable. The prices $(c_t)_t$ can either correspond to real prices or be abstract functions revealing the objective function of the aggregator, up to an additive or multiplicative constant, as seen in the three practical examples below.

Example 3.1. The aggregator has taken positions $(X_t^{\text{DA}})_t$ on the Day-Ahead market. Then he is facing penalties on the balancing market, and wants to minimize the distance to its bid profile:

$$\mathcal{C}(\boldsymbol{X}) = \left\|\boldsymbol{X}^{\text{DA}} - \boldsymbol{X}\right\|_2^2 = \sum_{t\in\mathcal{T}} (X_t^{\text{DA}})^2 + \sum_{t\in\mathcal{T}} X_t \times (X_t - 2X_t^{\text{DA}}) \,.$$

Example 3.2. The aggregator owns a source of renewable energy and forecasts a production profile $(\hat{G}_t)_t$. He wants to maximize the flexible consumption when $\hat{G}_t$ is the most important [WMRH12], and can therefore minimize:

$$\mathcal{C}(\boldsymbol{X}) = \left\|\hat{\boldsymbol{G}} - \boldsymbol{X}\right\|_2^2 = \sum_{t\in\mathcal{T}} \hat{G}_t^2 + \sum_{t\in\mathcal{T}} X_t \times (X_t - 2\hat{G}_t) \,.$$

Example 3.3. The aggregator has his own production facilities with convex and increasing production cost $C(X^{\text{tot}})$ where X^{tot} is the total power to be provided. If the set of consumers has a total aggregated nonflexible profile $\boldsymbol{X}^{\text{NF}}$, then at each time the total demand is $X_t^{\text{tot}} = X_t^{\text{NF}} + X_t$. The additional cost for the flexible load is $C(X_t^{\text{NF}} + X_t) - C(X_t^{\text{NF}})$ and the aggregator will minimize:

$$\mathcal{C}(\boldsymbol{X}) = \sum_t X_t \times \left(\frac{C(X_t^{\text{NF}} + X_t) - C(X_t^{\text{NF}})}{X_t}\right) \,,$$

where the term between parentheses can be set as the price signal $c_t(X_t)$ to be sent to consumers for coordination.

Note that in Examples 3.1 and 3.2, the functions (c_t) are not directly related to real prices but act rather as signals for coordination between consumers.

In our framework, we will consider the following different assumptions on the price functions $(c_t)_t$.

Assumption 3.1. *For each $t \in \mathcal{T}$, c_t is twice differentiable, convex and strictly increasing.*

Assumption 3.2. *For each $t \in \mathcal{T}$, c_t is twice differentiable, convex and strictly increasing. Moreover, there exists $a > 0$ s.t. for any t and admissible $\boldsymbol{x}$:*

$$2c_t'(X_t)\left(1 - \left(\frac{c_t''(X_t)}{2c_t'(X_t)}\right)^2 \|\boldsymbol{x}_t\|_2^2\right) \geqslant a. \tag{3.3}$$

Assumption 3.3. *For each $t \in \mathcal{T}$, c_t is affine, positive and increasing:*

$$\forall t \in \mathcal{T},\ c_t(x) = \alpha_t + \beta_t x \text{ with } \alpha_t, \beta_t \in (\mathbb{R}_+^*)^2 \,.$$

The three latter assumptions are more and more restrictive: Assumption 3.3 implies Assumption 3.2 with $a = 2\min_t \beta_t$, and Assumption 3.2 implies Assumption 3.1. Note that Assumption 3.3 provides a practical case for which all our results hold, and is verified in the case of Examples 3.1 and 3.2.

Remark 3.1. *For $a = 0$, inequality* (3.3) *in Assumption 3.2 simplifies to the condition:* $\|\boldsymbol{x}_t\|_2^{-1} \geqslant \left|\frac{c_t''(X_t)}{2c_t'(X_t)}\right|$. *For each t, c_t'' has to be small relatively to c_t'.*

Assumption 3.1 is standard in the congestion games literature and corresponds to "type-B" functions in the seminal paper [ORS93]. This assumption is also made in most of the papers dealing with game-theoretic DR models as [Bah+13]. Indeed, it is justified by the fact that marginal costs of producing and providing electricity are increasing. Assumption 3.3 is also a standard assumption made in [BH14] because it enables fast computation of NE (see Section 3.3), but it is restrictive, although several papers as [MR+10; Atz+13] simply consider linear price functions $c_t(x) = \beta_t x$. Last, Assumption 3.2 is not very explicit but is an in-between condition that comprises a larger set of functions than linear functions and for which our main results hold. For instance, the assumption holds for the family of polynomial functions considered in [Che+14]: $c_t(x) = \alpha + \beta x^{\nu_t}$ with $\alpha \geqslant 0, \beta > 0$, $1 \leqslant \nu_t < 3$ and if $X_t > 0$ for each $t \in \mathcal{T}$. More generally, this condition will be verified if c_t'' is small enough compared to c_t'.

Whatever the assumption retained, the objective of the aggregator is to send the right incentives to consumers through a *billing mechanism* in order to minimize his costs. A billing mechanism does not refer to a real billing system but more generally to a *signal* sent in order to coordinate consumers. It is given as a tuple of billing functions $(b_n)_{n\in\mathcal{N}}$ chosen by the aggregator to recover the global system cost $\mathcal{C}(\boldsymbol{X}) = \sum_{t\in\mathcal{T}} X_t c_t(X_t)$. As a result, the billing functions b_n are chosen such that $\mathcal{C} = \sum_{n\in\mathcal{N}} b_n$. Of course, one can always consider a profit ratio κ if the billing functions are used to design real consumer bills (the bill of n is set to κb_n). The function b_n depends of course on n's flexible consumption profile $\boldsymbol{x}_n$ but also depends on the load of the other consumers through the aggregated load $\boldsymbol{X}$.

3.2.4 Consumer's Optimization Problems

In this chapter, following our studies in [Jac+17b; Jac+17a], we will use an hourly proportional billing mechanism, where each consumer $n \in \mathcal{N}$ minimizes her bill:

$$b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} x_{n,t} c_t(X_t) = \sum_{t\in\mathcal{T}} x_{n,t} c_t\big(x_{n,t} + s_{n,t}\big), \tag{3.4}$$

where $\boldsymbol{x}_{-n} \stackrel{\text{def}}{=} (\boldsymbol{x}_m)_{m\neq n}$ denotes the consumption of all consumers but n and $s_{n,t} \stackrel{\text{def}}{=} \sum_{m\neq n} x_{m,t}$.

This billing mechanism was shown to have interesting fairness properties and is also adequate when considering consumers' utility functions (representing, e.g., temporal preferences for flexible consumption) [BH14; Jac+17b; Jac+17a]. This mechanism gives a particular *aggregative* structure, where the dependency to the others is only through the aggregated load $X_t = x_{n,t} + s_{n,t}$.

Through her ECS, each consumer will adjust her flexible consumption profile $\boldsymbol{x}_n \in \mathbb{R}^T$ to minimize her bill, which corresponds to the following optimization problem:

$$\min_{\boldsymbol{x}_n\in\mathcal{X}_n} b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \tag{3.5}$$

where $\mathcal{X}_n \subset \mathbb{R}^T$ is the set of consumer n feasible profiles. In the remaining of the chapter, we assume the following:

Assumption 3.4. *For each $n \in \mathcal{N}$, $\mathcal{X}_n$ is compact and convex.*

Problem (3.5) is a convex nonlinear mathematical program for which efficient methods of resolution exist [BTN01]. Most of the results given in this chapter hold without any further assumptions than Assumption 3.4, but we will focus on feasibility sets of the form (3.6), also considered in [MR+10; LCL11; Che+14; Bah+13; BH14; Sam+12].

Main Example 3.1. Deferrable load with fixed energy demand:

$$\mathcal{X}_n \stackrel{\text{def}}{=} \Big\{ \boldsymbol{x}_n \in \mathbb{R}^T \text{ s.t. } \textstyle\sum_{t\in\mathcal{T}} x_{n,t} = E_n \,, \tag{3.6a}$$

$$\underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t}, \forall t \in \mathcal{T} \Big\}. \tag{3.6b}$$

Constraint (3.6a) ensures that the total energy given to n satisfies her daily flexible energy demand over $\mathcal{T}$, denoted by E_n, that we assume fixed and deterministic[1]. Constraint (3.6b) takes into account the physical power constraints and the personal scheduling constraints (supposed given by the user to her ECS). Note that taking $\underline{x}_{n,t} = \overline{x}_{n,t} = 0$ forces $x_{n,t} = 0$ so that constraint (3.6b) includes in particular unavailability during some time periods. Constraints (3.6) give a simple model for deferrable loads such as water heaters (energy to heat a quantity of cold water between refill and usage periods) or electric vehicles (energy to be charged in the battery during parking period) [CBK17].

Note that (3.6) gives a generalization of routing "atomic splittable" congestion games [ORS93], well studied in the game theory literature, where the feasibility sets generally considered are $\mathcal{X}_n \stackrel{\text{def}}{=} \{\boldsymbol{x}_n \in (\mathbb{R}_+)^T \text{ s.t } \sum_t x_{n,t} = E_n\}$ where $x_{n,t}$ represents the flow of n on arc t. The addition of time-dependent bounding constraints (3.6b), also considered in [Che+14; Sam+12], gives a more accurate model for electrical loads.

Another important practical example that fits in our context, considered for instance in [CBK17; LCL11] is given below.

Example 3.4. Thermostatically controlled load :

$$\mathcal{X}_n \stackrel{\text{def}}{=} \big\{\underline{\theta}^{\text{comf}}_{n,t} \leqslant \theta_{n,t} \leqslant \overline{\theta}^{\text{comf}}_{n,t}, \forall t \in \mathcal{T}, \tag{3.7a}$$

$$\theta_{n,t} = \theta_{n,t-1} + \rho_n(\theta^{\text{out}}_{n,t} - \theta_{n,t-1}) + \varepsilon_n x_{n,t}, \forall t \in \mathcal{T}\big\}. \tag{3.7b}$$

Constraints (3.7) offer a model for thermostatically controlled loads such as fridges or air conditioning. Here, (3.7a) ensures that the temperature remains within the comfort range $[\underline{\theta}^{\text{comf}}_{n,t}, \underline{\theta}^{\text{comf}}_{n,t}]$. The temperature evolves through the linear equation (3.7b) according to the efficiency parameters ρ_n and ε_n, and to the exterior temperature $\theta^{\text{out}}_{n,t}$ (see [LCL11] for details). Using (3.7b), one can rewrite (3.7) only with the variables $(\boldsymbol{x}_{n,t})_{t\in\mathcal{T}}$.

Remark 3.2. *Owing to the convexity of $\mathcal{X}_n$ (Assumption 3.4), we do not consider appliances that require a fixed consumption profile but for which the starting time can be optimized (e.g. washing machines). In this case, one can use a (nonconvex) mixed-integer formulation, as in [BWH17].*

We denote by $\mathcal{X} \stackrel{\text{def}}{=} \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ the Cartesian product of the feasible sets. As b_n depends both on $\boldsymbol{x}_n$ and $\boldsymbol{x}_{-n}$, we get a N-person minimization game that we write in the standard form [FT91] as $\mathcal{G} \stackrel{\text{def}}{=} (\mathcal{N}, \mathcal{X}, (b_n)_n)$.

3.2.5 Equilibrium Analysis and Efficiency

We are interested in the equilibrium situations of the game $\mathcal{G}$. The concept of Nash Equilibrium (NE) (see Definition 2.1), is accepted as a desirable and plausible outcome in N-person noncooperative games, as it implies by definition individual stability (each player has no interest to change her action unilaterally).

It is known that an NE may not exist or may not be unique, even in routing congestion games [ORS93]. In our framework however, both properties are ensured, as stated below.

Theorem 3.1. *Under Assumption 3.1, there exists an NE of $\mathcal{G}$.*

Proof. This is a corollary of Rosen [Ros65] as $\mathcal{G}$ is convex. □

To ensure the uniqueness of the NE, a common approach, adopted in [Che+14], is to verify that a game is "diagonally strictly convex" [Ros65]. We will see further from Remark 3.8 and Proposition 3.1 that this property holds with Assumption 3.2. However, the uniqueness results based on this approach ask for more demanding conditions on the price functions $c_t(.)$ than Assumption 3.1. Here, in the case of feasibility sets of the form (3.6), Theorem 3.2 ensures the uniqueness for arbitrary convex and strictly increasing prices (Assumption 3.1).

Theorem 3.2. *Under Assumption 3.1 and if, for each $n \in \mathcal{N}$, $\mathcal{X}_n$ is of the form* (3.6)*, then $\mathcal{G}$ has a unique NE.*

[1] E_n can be set by the consumer, induced by the physical parameters of her appliances (battery capacity), or computed by learning the consumer's habits.

Proof: See Appendix 3.A. This proof extends the uniqueness theorem given in [ORS93] in presence of the constraint (3.6b).

As said above, an NE is a very interesting situation in practice because of its stability: each player will only increase her bill if she changes her profile. However, an NE does not necessarily minimize the *social cost*

$$\mathrm{SC}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_n b_n(\boldsymbol{x}) \,. \tag{3.8}$$

Remark 3.3. *With the billing equation* (3.4), $\mathrm{SC}(\boldsymbol{x})$ *is equal to the total* system cost, *that is,* $\sum_t X_t c_t(X_t)$*, a quantity that the aggregator should minimize. In general, the system cost can differ from the social cost of consumers, for instance if we consider that the aggregator makes a positive profit, or if we consider consumers utility functions as done in [Jac+17a].*

In general games, an NE can be suboptimal in terms of SC. To measure the inefficiency of Nash Equilibria in terms of social cost, a standard quantity is the Price of Anarchy:

Definition 3.1 (Koutsoupias *et al*, 1999)**. Price of Anarchy (PoA).**
Given a N-player game $\mathcal{G} = (\mathcal{N}, \mathcal{X}, (b_n)_n)$ and $\mathcal{X}_{\mathrm{NE}}$ its set of Nash equilibria, the PoA is defined as the following ratio:

$$\mathrm{PoA}(\mathcal{G}) = \frac{\sup_{\boldsymbol{x}\in\mathcal{X}_{\mathrm{NE}}} \mathrm{SC}(\boldsymbol{x})}{\inf_{\boldsymbol{x}\in\mathcal{X}} \mathrm{SC}(\boldsymbol{x})} \,.$$

Note that, from Definition 3.1, as $\mathcal{X}_{\mathrm{NE}} \subset \mathcal{X}$, the PoA is always greater than one. Furthermore, finding an upper bound on the PoA ensures that the social cost at any NE will be relatively close to the minimal social cost. Bounding the PoA is a hard theoretical question in general congestion games [Rou15; JMT05]. In [RS15], the authors give an upper bound if the price functions are polynomial with bounded degree and positive coefficients. With degree one (affine prices, Assumption 3.3) the bound is PoA $\leqslant 1.5$, which means that the NE profile can induce costs as much as 50% higher than the optimal costs: implementing such a framework would not be worthwhile for the aggregator, as uncoordinated consumers will probably perform better (in our simulations, the uncoordinated profiles induce costs 16% higher than the optimal costs, see Table 3.2). However, the results in [RS15] are worst-case bounds and these bounds are only approached asymptotically[2]: in our simulations with affine prices, the PoA was always much lower than 1.5 (around 1.017 from Table 3.2). One of the reasons is that in [RS15] the model does not consider the power constraints (3.6b), and a PoA of 1.5 might be reached in our case only if the constraints (3.6b) are coarse enough. To further explain the low PoA in our instances, we found the following theorem by applying the (λ, μ) local smoothness technique of [RS15]:

Theorem 3.3. *Under Assumption 3.3, define for any* $t \in \mathcal{T}$, $\varphi_t = (1 + \frac{\alpha_t}{\beta_t \overline{X}_t})^2$, *where* $\overline{X}_t = \sum_n \overline{x}_{n,t}$ *and* $t_0 \stackrel{\text{def}}{=} \arg\min_t \frac{\alpha_t}{\beta_t \overline{X}_t}$*. Assuming that, for all* $t \in \mathcal{T}$*:*

$$\varphi_t \leqslant \varphi_{t_0} + 2 + \sqrt{1 + \varphi_{t_0}} \,, \tag{3.9}$$

the following inequality holds:

$$\mathrm{PoA}(\mathcal{G}) \leqslant \frac{1}{2}\left(1 + \sqrt{1 + \varphi_{t_0}^{-1}} + \frac{1}{2}\varphi_{t_0}^{-\frac{1}{2}}\right) \,. \tag{3.10}$$

Proof: See Appendix 3.B.

Remark 3.4. *Using the inequality* $\forall x \geqslant 0,\ \sqrt{1+x^2} \leqslant 1 + x$*, inequality* (3.10) *implies the following simplified—but coarser—bound:*

$$\mathrm{PoA}(\mathcal{G}) \leqslant 1 + \tfrac{3}{4} \sup_{t\in\mathcal{T}} \left(1 + \frac{\alpha_t}{\beta_t \overline{X}_t}\right)^{-1} \,. \tag{3.11}$$

[2]Meaning that there exists a sequence of games $(\mathcal{G}_\nu)_{\nu\geqslant 0}$, with parameters depending on ν, and affine price functions c_t such that $\mathrm{PoA}(\mathcal{G}_\nu) \underset{\nu\to\infty}{\longrightarrow} 1.5$.

The assumption (3.9) in Theorem 3.3 ensures that price functions (c_t) cannot differ too much from one time period to another. This is verified for instance if the price functions are uniform over $\mathcal{T}$ (*i.e.* $\forall t, c_t = c$). One can see that, according to Theorem 3.3, the PoA converges to one when $\alpha_t/(\beta_t \overline{X}_t)$ diverges to infinity for each t: the PoA can be arbitrarily close to one if we choose the coefficients α_t large enough. This result is indeed intuitive: when the prices are constant ($\beta_t = 0$), they do not depend on the aggregates $\boldsymbol{X}$ and there is no congestion effect; the optimal profile is obtained by each consumer choosing the time periods with lowest prices, independently of $\boldsymbol{x}_{-n}$. Another interesting result is that the PoA also converges to one when the total load is low ($\forall t, \overline{X}_t \to 0$). Note that the right-hand-side of inequality (3.10) is decreasing with φ_0 and is equal to $(\frac{1+\sqrt{2}}{2})^2 \approx 1.457$ for $\varphi_0 = 1$ so our result is always tighter than the bound of 1.5 given in [RS15]. However, in our simulations with linear prices, the PoA was still lower than the bound (3.10), even when assumption (3.9) does not hold: the inequality (3.10) gives PoA $\leqslant 1.271$ (average on the simulated days), while the PoA on mean values from Table 3.2 is 1.017. In this regards, getting a tighter bound or generalizing Theorem 3.3 to more general price functions could be the subject of future work.

3.3 Fast Computation of the Nash Equilibrium

Since we have shown that the NE is a good decentralized optimization target, the next question we address is the computation of the NE consumption profiles. This question is a central problem in game theory [FPT04]. Furthermore, this computation has to be done in a short time to be implemented in practice. In this section, we provide two algorithms, we prove their convergence to the NE and we give a guarantee on their convergence rate in our specific setting.

3.3.1 Two Decentralized Algorithms

Given a profile $\boldsymbol{x}_{-n}$ of the others, consumer n chooses the profile $\boldsymbol{x}_n$ corresponding to a minimizer of (3.5), which is called her *Best Response*[3]. It is denoted by

$$\mathrm{BR}_n : \boldsymbol{s}_n \mapsto \operatorname*{argmin}_{x_n \in \mathcal{X}_n} \sum_t x_{n,t} c_t(s_{n,t} + x_{n,t}) \,, \tag{3.12}$$

which only depends on the sum of the load of the others $\boldsymbol{s}_n \overset{\text{def}}{=} \sum_{m \neq n} \boldsymbol{x}_m \in \mathbb{R}^T$ because of the "aggregated" structure: in a general setting, BR_n would be a function of $\boldsymbol{x}_{-n}$.

Remark 3.5. *Here, the* arg min *function in the definition* (3.12) *of* BR_n *is* single-valued *because of the strict monotonicity assumption (Assumption 3.1). In general,* $\mathrm{BR}_n(\boldsymbol{s}_n)$ *can be multivalued. In that case, we can still use Algorithm 3.1 by arbitrarily choosing any element of* $\mathrm{BR}_n(\boldsymbol{s}_n)$ *at Line 5.*

A natural algorithm to compute an NE is to iterate best responses and update the strategies, cycling over the set of users until convergence. This procedure, referred to as *Cycling Best-Response Dynamics* (CBRD) [GM91] is described by Algorithm 3.1.

Algorithm 3.1 Cycling Best Response Dynamics (CBRD)

Require: $\boldsymbol{x}^{(0)}$, stopping criterion
1: $k \leftarrow 0$
2: **while** stopping criterion not true **do**
3: **for** $n = 1$ to N **do**
4: $\boldsymbol{s}_n^{(k)} = \sum_{m<n} \boldsymbol{x}_m^{(k+1)} + \sum_{m>n} \boldsymbol{x}_m^{(k)}$
5: $\boldsymbol{x}_n^{(k+1)} \leftarrow \mathrm{BR}_n(\boldsymbol{s}_n^{(k)})$
6: **done**
7: $k \leftarrow k + 1$
8: **done**

Standard stopping criteria that can be used in Algorithm 3.1 are a maximum number of iterations $k_{\max}$, a maximum CPU time, an objective on the distance between iterates

[3] As player n chooses her best profile to the fixed profiles of the others; she responds to them.

$\left\|\boldsymbol{x}^{(k-1)} - \boldsymbol{x}^{(k)}\right\| \leqslant \varepsilon_{\text{stop}}$, or the satisfaction of the KKT conditions of optimality for each user's convex optimization problem (3.5) up to an absolute error tolerance.

The only computationally demanding step in Algorithm 3.1 is the computation of the best response $\text{BR}_n(\boldsymbol{s}_n)$ on Line 5. Its complexity differs according to the price functions c_t and the feasibility set $\mathcal{X}_n$. In general, there is no explicit expression of $\text{BR}_n(\boldsymbol{s}_n)$ but, as $\mathcal{X}_n$ is convex and $\boldsymbol{x}_n \mapsto b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n})$ is convex, techniques of nonlinear convex optimization can be used to find an approximating solution [BTN01]. The problem simplifies if prices are affine (Assumption 3.3) and $\mathcal{X}_n$ is given by (3.6a)-(3.6b) and none of the bounding constraints (3.6b) is active. In that case, an explicit expression of $\text{BR}_n(\boldsymbol{s}_n)$ can be found [Alt+02] so Line 5 can be executed in constant time. In the general case of feasibility sets of the form (3.6) (bounding constraints (3.6b) can be active), we are still in a specific case of quadratic programming where an exact solution can be computed in $\mathcal{O}(T)$ with [Bru84]. When $\mathcal{X}_n$ is a general polytope given as a set of linear inequalities (as in [CBK17]), convex quadratic programming [BTN01] can be used to compute the solution.

Remark 3.6. *The **for** loop in Algorithm 3.1 (Line 3) implements sequential updates and cycles over the set of players in the arbitrary order $1, 2, \ldots, N$ in a Gauss-Seidel manner [PC06]. Choosing a "good" order of the* BR *in the **for** loop might accelerate the convergence of the algorithm. A* simultaneous *version of Algorithm 3.1 (without Line 4 and with Line 5 executed by all players in parallel) could also improve the speed of Algorithm 3.1, but we observed that doing so can prevent its convergence.*

Another natural algorithm to compute the equilibrium is to emulate the projected gradient descent, well-known in convex optimization [CM87], by considering the gradient of each objective function of the players, as described in Algorithm 3.2.

Algorithm 3.2 Simultaneous Improving Response Dynamics (SIRD)

Require: $\boldsymbol{x}^{(0)}$, γ, stopping criterion
1: $k \leftarrow 0$
2: **while** stopping criterion not true **do**
3: **for** $n = 1$ to N **do**
4: $\boldsymbol{x}_n^{(k+1)} \leftarrow \text{P}_{\mathcal{X}_n}\left(\boldsymbol{x}_n^{(k)} - \gamma \nabla_n b_n(\boldsymbol{x}_n^{(k)}, \boldsymbol{x}_{-n}^{(k)})\right)$
5: **done**
6: $k \leftarrow k + 1$
7: **done**

At Line 4 of Algorithm 3.2, $\text{P}_{\mathcal{X}_n}$ denotes the projection on the feasibility set $\mathcal{X}_n$ of consumer n and $\nabla_n b_n = (\partial b_n / \partial x_{n,t})_{t \in \mathcal{T}}$ denotes the gradient with respect to variable $\boldsymbol{x}_n$. The same stopping criteria listed below Algorithm 3.1 can be used for Algorithm 3.2. The chosen denomination *improving response* recalls that, at each iteration of Algorithm 3.2, player n *improves* her profile $\boldsymbol{x}_n$ by performing a projected gradient step (Line 4), but in general does not choose the best improvement as in Algorithm 3.1.

Note that from Algorithm 3.1 to Algorithm 3.2, only the instructions within the **for** loop are changed: the update of $\boldsymbol{s}_n$ and computation of BR_n (Lines 4 and 5 of Algorithm 3.1) are replaced with the gradient step (Line 4 of Algorithm 3.2).

Remark 3.7. *Both Algorithm 3.1 and Algorithm 3.2 can be implemented in a "decentralized" procedure: Lines 4 and 5 in Algorithm 3.1 and Line 4 in Algorithm 3.2 can be performed locally by each consumer's ECS. In this way, consumers' privacy is respected as they do not send any information about their constraints to the aggregator. On the other hand, they only receive information on the aggregated load $\boldsymbol{s}_n^{(k)}$ and can hardly deduce the individual consumption profiles $\boldsymbol{x}_{-n}$ of the other consumers.*

The computational complexity of an iteration of Algorithm 3.2 (within the **for** loop) is equivalent to the complexity of the projection $\text{P}_{\mathcal{X}_n}$, Apart from specific cases—as the Euclidean projection on a p-dimensional simplex which can be computed in $O(p\ln(p))$, see [CY11]—this projection requires in general to solve a Quadratic Program (QP), so it is of the

same order of complexity (see [BTN01, Lecture 4]) as one iteration (within the **for** loop) of algorithm CBRD. However, note that, as we do not update sequentially the load of the others $\boldsymbol{x}_{-n}$ in Algorithm 3.2, the projected gradient step within the **for** loop can be computed simultaneously and can be parallelized.

3.3.2 Game Stability and Convergence of Algorithms BRD and SIRD

In this section, we provide theoretical convergence rates of the two algorithms proposed in Section 3.3.1. We first recall the notion of *stability*, and prove (Proposition 3.1) that the energy consumption game $\mathcal{G}$ defined above is *strongly stable* under Assumption 3.2. The notion of stability was introduced in [HS09] in order to study different game dynamics in continuous time and their convergence to NE. We extend this property to a "strong" version (symmetrically to the concept of strong monotonicity for operators).

Definition 3.2 (Hofbauer and Sandholm [HS09])**. Stable Game.**
A minimization game $\mathcal{G} = (\mathcal{N}, \mathcal{X}, (b_n)_n)$ is stable iff

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X},\ (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{T}}.\left(F(\boldsymbol{x}') - F(\boldsymbol{x})\right) \geqslant 0\,, \tag{3.13}$$

with $F(\boldsymbol{x}) \stackrel{\text{def}}{=} (\nabla_n b_n(\boldsymbol{x}))_{n\in\mathcal{N}}$.
Moreover, $\mathcal{G}$ is a-strongly stable, with a constant $a > 0$, iff:

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X},\ (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{T}}.\left(F(\boldsymbol{x}') - F(\boldsymbol{x})\right) \geqslant a\left\|\boldsymbol{x} - \boldsymbol{x}'\right\|^2\,. \tag{3.14}$$

Remark 3.8. *The condition of stability in* (3.13) *is equivalent to the condition of strict diagonal convexity in [Ros65], which implies uniqueness of* NE *[Ros65, Thm.2].*

Definition 3.2 gives an abstract condition on an operator that depends on the objective functions of the players. In our case, players objectives (b_n) depend linearly on price functions $(c_t)_t$ through (3.4), so it is interesting to translate the condition of Definition 3.2 directly on the price functions, as stated in Proposition 3.1.

Proposition 3.1. *Let $a > 0$ such that Assumption 3.2 holds. Then, the game $\mathcal{G}$ is a-strongly stable.*

Proof: See Appendix Section 3.C.

This property will be used to show the convergence of Algorithm 3.2 in Theorem 3.5. Concerning Algorithm 3.1, the approach is different and the convergence is established only in the specific case of Assumption 3.3. In general games, CBRD might not converge [HB16] or might take an exponential time to converge [AR08]. In atomic splittable congestion games on a parallel network, as in our case, the convergence and the speed of Algorithm 3.1 has been studied previously in [Mer08] and [BPS13], where the authors show by different methods that there is a geometric convergence in the case of $N = 2$ players and convex and strictly increasing price functions (Assumption 3.1). However, to the best of our knowledge, the convergence in this setting and for more players $N > 2$ is still an open question.

In our case, simulations show a geometric convergence rate for any instance of $\mathcal{G}$ satisfying Assumption 3.3 and for any $N \in \mathbb{N}$, as illustrated in Figure 3.1. In [BPS13], it is conjectured that this geometric convergence may also hold under Assumption 3.1. Restricting ourselves to affine price functions, we notice that game $\mathcal{G}$ is a potential game [Jac+17a; MS96] and we get the following guarantee on the rate of convergence of Algorithm 3.1:

Theorem 3.4. *Under Assumption 3.3, the sequence of iterates $\left(\boldsymbol{x}^{(k)}\right)_{k\geqslant 0}$ of Algorithm CBRD converges to the unique* NE *$\hat{\boldsymbol{x}}$ of $\mathcal{G}$. Moreover, the convergence rate satisfies:*

$$\forall k \geqslant 0,\ \left\|\hat{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\right\|_2 \leqslant C\sqrt{\frac{M}{a}} \times \frac{N}{\sqrt{k}}\,, \tag{3.15}$$

where C depends on $x^{(0)}$ and the billing functions, $M \stackrel{\text{def}}{=} 2\max_t \beta_t$ and $a \stackrel{\text{def}}{=} 2\min_t \beta_t$.

Proof: See Appendix 3.D. The result is implied by convergence of alternating block coordinate minimization method [Hon+17].

The proof of Theorem 3.4 uses the fact that $M = \max_n M_n$ where M_n is a Lipschitz constant of $\nabla_n b_n$, and a is a strong convexity (and a-strong stability) constant. To the best of our knowledge, the question to know if Theorem 3.4 holds for general price functions is open; it can be an avenue for future research.

It is easier to get a strong guarantee on the convergence rate of Algorithm 3.2 for general price functions, as stated in Theorem 3.5:

Theorem 3.5. *Denote by M_n a Lipschitz constant of $\nabla_n b_n$ and $M \stackrel{\text{def}}{=} \max_n M_n$. Under Assumption 3.2 (a- strong stability), for a step $\gamma \stackrel{\text{def}}{=} a/(NM^2)$, SIRD converges to the NE. Moreover:*

$$\forall k \geqslant 0, \left\| \hat{\boldsymbol{x}} - \boldsymbol{x}^{(k)} \right\|_2 \leqslant \left(1 - \tfrac{a^2}{NM^2}\right)^k \left\| \hat{\boldsymbol{x}} - \boldsymbol{x}^{(0)} \right\|_2 . \tag{3.16}$$

Proof: See Appendix 3.E.

Under Assumption 3.3, as stated in Theorem 3.5 we have $M \stackrel{\text{def}}{=} 2\max_t \beta_t$ and $a \stackrel{\text{def}}{=} 2\min_t \beta_t$, which gives the explicit contraction ratio $\eta = 1 - \frac{\max_t \beta_t}{N \min_t \beta_t}$.

The bound given in Theorem 3.5 shows that the convergence of Algorithm 3.2 is slower when the number of consumers N increases.

3.3.3 Randomized Best Response and Convergence in Splittable Congestion Games

As said above in Section 3.3.2, there are few results concerning the convergence of Best Response (Algorithm 3.1) in splittable congestion games. Although the algorithm seems to converge in practice in general splittable congestion games, to our limited knowledge, the only existing theoretical results are the works of [Mer08] and [BPS13] showing, by two different methods, the convergence in the case of $N = 2$ players and with strictly convex costs.

In this section, we consider a *randomized* variant of Algorithm 3.1, and we give in Corollary 3.2 the geometric convergence for two specific cases different from [Mer08]: the case of linear cost functions, and the case with $T = 2$ resources and arbitrary cost functions.

We are interested in the randomized version of Best Response, where at each iteration k, a player n_k is drawn out uniformly among $\mathcal{N}$, and its strategy is updated to its best response BR_{n_k}, as explained in Algorithm 3.3.

Algorithm 3.3 Randomized Alternate Best Response

Require: $\boldsymbol{x}^{(0)}$, stopping criterion
1: $k \leftarrow 0$
2: **while** stopping criterion not true **do**
3: Draw $n_k \in \mathcal{N}$ uniformly
4: Update $\boldsymbol{x}_{n_k}^{(k+1)} \in \mathrm{BR}(\boldsymbol{x}_{-n_k}^{(k)})$, leave other players strategy unchanged
5: Go to next iteration $k \leftarrow k + 1$
6: **done**

Recall that, from the specific structure of a congestion game on a parallel network, the best response of a player $n \in \mathcal{N}$ to other players strategies $\boldsymbol{x}_{-n}$ is:

$$\mathrm{BR}_n(\boldsymbol{x}_{-n}) \in \operatorname*{argmin}_{\boldsymbol{x}_n \in \mathcal{X}_n} \sum_t x_{n,t} c_t \left(\sum_{m \neq n} x_{m,t} + x_{n,t} \right) \stackrel{\text{def}}{=} \mathcal{B}_n \left(\sum_{m \neq n} x_{m,t} \right) , \tag{3.17}$$

where $\mathcal{B}_n(.) : \mathbb{R}^T \to \mathbb{R}^T$ is a function of the vector of the sum of load of other players.

For this analysis, we make the assumption that, for each $t \in \mathcal{T}$ and $s \in \mathbb{R}_+$, the function $x \mapsto x c_t(s + x)$ is strictly convex. An immediate consequence is that, given any strategy profiles $(\boldsymbol{x}_m)_{m \in \mathcal{N}}$ and any player n, the best response $\mathrm{BR}_n(\boldsymbol{x}_{-n})$ is unique, and a Nash equilibrium (Definition 2.1), if it exists, is also unique.

To a point $\boldsymbol{x} \in \mathcal{X}$, we associate an extended vector $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x} & \sum_{n\in\mathcal{N}} \boldsymbol{x}_n \end{pmatrix} \in \mathcal{Y} \stackrel{\text{def}}{=} \mathcal{X} \times \mathbb{R}^T$. Let us introduce the operator

$$\sigma : \left\{ \begin{array}{l} \mathcal{Y} \to \mathbb{R}^T \\ (\boldsymbol{x}, \boldsymbol{s}) \mapsto \boldsymbol{s} - \sum_{n\in\mathcal{N}} \boldsymbol{x}_n \end{array} \right. ,$$

as well as the coordinate-wise Best Response operators on $y \in \mathcal{Y}$ by:

$$Q_n : \left\{ \begin{array}{l} \mathcal{Y} \to \mathcal{Y} \\ (\boldsymbol{x}, \boldsymbol{s}) \mapsto \left(\boldsymbol{x}_1, \ldots, \mathcal{B}_n(\boldsymbol{s} - \boldsymbol{x}_n), \ldots, \boldsymbol{x}_N, \boldsymbol{s} - \boldsymbol{x}_n + \mathcal{B}_n(\boldsymbol{s} - \boldsymbol{x}_n)\right) \end{array} \right. \tag{3.18}$$

for $n \in \mathcal{N}$. Then Algorithm 3.3 can be rewritten by a sequence of iterates $\boldsymbol{y}^{(k+1)} = Q_{n_k}(\boldsymbol{y}^{(k)})$, and we have the following property:

Lemma 3.1. *Let $\boldsymbol{y} \in \mathcal{Y}$ such that $\sigma(\boldsymbol{y}) = 0$ and $n \in \mathcal{N}$. Then $\sigma\left(Q_n(\boldsymbol{y})\right) = 0$.*

Proof. Let $b_n = \mathcal{B}_n(\boldsymbol{s} - \boldsymbol{x}_n)$. We have, for any t, $\sigma\left(Q_n(y)\right)_t = s_t - x_{n,t} + b_{n,t} - \sum_{m\neq n} x_{m,t} - b_{n,t} = s_t - \sum_{m\in\mathcal{N}} x_{m,t} = 0$. □

Next, we use the norm on $\mathcal{Y}$: $\|\boldsymbol{y}\|_A \stackrel{\text{def}}{=} \left(\sum_n \|\boldsymbol{x}_n\|_2^2 + A\,\|\boldsymbol{s}\|_2^2\right)^{\frac{1}{2}}$, where the norm $\|.\|_2$ on $\mathbb{R}^T$ is the euclidean norm and $A > 0$ is a chosen constant. We have the following result:

Proposition 3.2. *Assume that, for all $n \in \mathcal{N}$, the sum best-response function introduced in* (3.17) *$\mathcal{B}_n : \mathbb{R}^T \to \mathcal{X}_n$ is α-contracting and that $\mathrm{Id}_n + \mathcal{B}_n : \mathbb{R}^T \to \mathbb{R}^T$ is β-contracting, and that $\alpha^2 + \beta^2 < 1$. Denote by n a random variable with uniform law over $\mathcal{N}$. Then there exists $\rho < 1$ such that, for all y and $y' \in \mathcal{Y} \cap \mathrm{Ker}(\sigma)$:*

$$\mathbb{E}_{I\sim\mathcal{U}(\mathcal{N})}\left[\left\|Q_I(\boldsymbol{y}) - Q_I(\boldsymbol{y}')\right\|_A^2\right] \leqslant \rho\left\|\boldsymbol{y} - \boldsymbol{y}'\right\|_A^2 . \tag{3.19}$$

Proof.

$$\begin{aligned}
&\mathbb{E}_{I\sim\mathcal{U}(\mathcal{N})}\left[\left\|Q_I(\boldsymbol{y}) - Q_I(\boldsymbol{y}')\right\|_A^2\right] = \frac{1}{N}\left(\sum_{n\in\mathcal{N}} \left\|Q_n(\boldsymbol{y}) - Q_n(\boldsymbol{y}')\right\|_A^2\right) \\
&= \frac{1}{N}\sum_{n\in\mathcal{N}}\Big(\sum_{m\neq n}\left\|\boldsymbol{x}_m - \boldsymbol{x}'_m\right\|_2^2 + \left\|\mathcal{B}_n(\boldsymbol{s} - x_n) - \mathcal{B}_n(\boldsymbol{s}' - x'_n)\right\|_2^2 \\
&+ A\left\|(\mathrm{Id} + \mathcal{B}_n)(\boldsymbol{s} - \boldsymbol{x}_n) - (\mathrm{Id} + \mathcal{B}_n)(\boldsymbol{s}' - x'_n)\right\|_2^2\Big) \\
&\leqslant \left(\frac{N-1}{N}\left\|\boldsymbol{x} - \boldsymbol{x}'\right\|^2\right) + \frac{\alpha^2 + A\beta^2}{N}\sum_{n\in\mathcal{N}}\left\|(\boldsymbol{s} - \boldsymbol{x}_n) - (\boldsymbol{s}' - \boldsymbol{x}'_n)\right\|_2^2 .
\end{aligned}$$

Then, we observe that:

$$\begin{aligned}
\sum_{n\in\mathcal{N}}\left\|(\boldsymbol{s} - \boldsymbol{x}_n) - (\boldsymbol{s}' - \boldsymbol{x}'_n)\right\|_2^2 &= \sum_{n\in\mathcal{N}}\left\|\boldsymbol{s} - \boldsymbol{s}'\right\|_2^2 + \left\|\boldsymbol{x}_n - \boldsymbol{x}'_n\right\|_2^2 - 2\left\langle \boldsymbol{s} - \boldsymbol{s}' \,,\, \boldsymbol{x}_n - \boldsymbol{x}'_n\right\rangle \\
&= N\left\|\boldsymbol{s} - \boldsymbol{s}'\right\|_2^2 + \left\|\boldsymbol{x} - \boldsymbol{x}'\right\|_2^2 - 2\left\langle \boldsymbol{s} - \boldsymbol{s}' \,,\, \boldsymbol{s} - \boldsymbol{s}'\right\rangle \\
&= (N-2)\left\|\boldsymbol{s} - \boldsymbol{s}'\right\|_2^2 + \left\|\boldsymbol{x} - \boldsymbol{x}'\right\|_2^2 ,
\end{aligned}$$

so that we obtain:

$$\begin{aligned}
\mathbb{E}\left[\left\|Q_I(y) - Q_I(y')\right\|_A^2\right] &\leqslant \tfrac{N-1+\alpha^2+A\beta^2}{N}\left\|\boldsymbol{x} - \boldsymbol{x}'\right\|_2^2 + \tfrac{(N-2)(\alpha^2+A\beta^2)}{N}\left\|\boldsymbol{s} - \boldsymbol{s}'\right\|_2^2 \\
&\leqslant \rho\left\|\boldsymbol{y} - \boldsymbol{y}'\right\|_A ,
\end{aligned}$$

where $\rho \stackrel{\text{def}}{=} \max\left(\tfrac{N-1+\alpha^2+A\beta^2}{N} \,,\, \tfrac{(N-2)(\alpha^2+A\beta^2)}{AN}\right) < 1$. □

The two assumptions on $\mathcal{B}_n$ and $\mathrm{Id} - \mathcal{B}_n$ in Proposition 3.2 might seem very demanding. However, we observe that the first assumption holds coarsely in the case of affine costs functions c_t:

Lemma 3.2. *Let $n \in \mathcal{N}$ and costs on arcs be affine: for all $t \in \mathcal{T}, c_t(x) = a_t + b_t x$, with $b_t > 0$. Then the functions:*

- $\mathcal{B}_n : \mathbb{R}^T \ni \boldsymbol{s} \mapsto \underset{\boldsymbol{x}_n \in \boldsymbol{x}_n}{\operatorname{argmin}} \sum_t x_{n,t} c_t(s_t + x_{n,t})$,
- $(\mathrm{Id} + \mathcal{B}_n) : \mathbb{R}^T \ni \boldsymbol{s} \mapsto \boldsymbol{s} + \mathcal{B}_n(\boldsymbol{s})$

are $\frac{1}{2}$-contracting for $\|.\|_2$ on $\mathbb{R}^T$.

Proof. The proof is given in Appendix 3.G.

Combining this Lemma 3.2 and Proposition 3.2, we immediately obtain:

Corollary 3.1. *Assume that the cost functions on resources $(c_t)_{t\in\mathcal{T}}$ are affine. Then Algorithm 3.3 is contracting in expectation, that is, for all $\boldsymbol{y}$ and $\boldsymbol{y}' \in \mathcal{Y} \cap \mathrm{Ker}(\sigma)$:*

$$\forall \boldsymbol{y}, \boldsymbol{y}' \in \mathcal{Y},\ \mathbb{E}_{I\sim\mathcal{U}(\mathcal{N})}\left[\left\|Q_I(\boldsymbol{y}) - Q_I(\boldsymbol{y}')\right\|_A^2\right] \leqslant \rho \left\|\boldsymbol{y} - \boldsymbol{y}'\right\|_A^2 , \tag{3.20}$$

where $\rho < 1$ is defined in the proof of Proposition 3.2 with $\alpha = \beta = \frac{1}{2}$.

One idea to show a contraction property of the operator Q_I is to analyze the partial derivatives of the best response function. It turns out that, with some calculus, we can compute explicitly those partial derivatives, as shown below.

For $n \in \mathcal{N}$ and $\boldsymbol{s} = \sum_{m\neq n} x_m$, let us denote the support of Best Response $\mathcal{B}_n$ by

$$\mathcal{S}_n(\boldsymbol{s}) = \left\{t \in \mathcal{T} \mid [\mathcal{B}_n(\boldsymbol{s})]_t \in (\underline{x}_{n,t}, \overline{x}_{n,t})\right\} .$$

For a given $\boldsymbol{s} \in \mathbb{R}^T$, the optimal Lagrangian multiplier $\lambda_n(\boldsymbol{s})$ defined in (3.40) and associated to the solution of $\mathcal{B}_n(\boldsymbol{s})$, is uniquely determined [Ber99, Prop.3.3.2] if $\mathcal{B}_n(\boldsymbol{s})$ is *regular*, that is if $\mathcal{S}_n(\boldsymbol{s}) \neq \emptyset$. From (3.41), for each $t \in \mathcal{T}$, we have the implication:

$$t \in \mathcal{S}_n(\boldsymbol{s}) \implies \gamma_{n,t}\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) = \lambda_n(\boldsymbol{s}) ,$$

but the reverse implication may not be true (a resource t can have marginal cost $\gamma_{n,t}$ equal to $\lambda_n(\boldsymbol{s})$ and be either at lower bound or at upper bound).

From [BPS13], we know that $\mathcal{B}_n$ is differentiable and that $\mathcal{S}_n(\boldsymbol{s})$ is locally stable for the points $\boldsymbol{s}$ for which this reverse implication is true, that is:

$$\mathrm{Adm}_n \overset{\text{def}}{=} \left\{\boldsymbol{s} \in \mathbb{R}^T \mid \gamma_{n,t}\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) = \lambda_n(\boldsymbol{s}) \Leftrightarrow t \in \mathcal{S}_n(\boldsymbol{s})\right\} ,$$

which is the set of points for which the resources $t \in \mathcal{T}$ that have their marginal cost equal to $\lambda_n(\boldsymbol{s})$ are exactly those such that $\mathcal{B}_n(\boldsymbol{s})_t \in (\underline{x}_{n,t}, \overline{x}_{n,t})$.

For any point $\boldsymbol{s} \in \mathrm{Adm}_n$, we can explicitly compute the Jacobian of $\mathcal{B}_n$, as stated below:

Lemma 3.3. *For a player $n \in \mathcal{N}$, the partial derivatives of the BR function $\boldsymbol{s} \mapsto \mathcal{B}_n(\boldsymbol{s})$ (function of the sum of load of others $\boldsymbol{s} \overset{\text{def}}{=} \sum_{m\neq n} x_{m,t}$) are equal to:*

$$\forall t, u \in \mathcal{T}, \quad \frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}) = \theta_u(\boldsymbol{s}) \times (r_t(\boldsymbol{s}) - \delta_{tu}) ,$$

where δ_{tu} is the standard Kronecker symbol equal to one iff $t = u$ and:

$$\theta_u(\boldsymbol{s}) \overset{\text{def}}{=} \frac{[\mathcal{B}_n(\boldsymbol{s})]_u\, c_u''(s_u) + c_u'(s_u)}{[\mathcal{B}_n(\boldsymbol{s})]_u\, c_u''(s_u) + 2c_u'(s_u)} \ \text{ and } \ r_t(\boldsymbol{s}) \overset{\text{def}}{=} \frac{\left([\mathcal{B}_n(\boldsymbol{s})]_t c_t''(s_t) + 2c_t'(s_t)\right)^{-1}}{\sum_{v\in\mathcal{S}_n(s)}\left([\mathcal{B}_n(\boldsymbol{s})]_v c_v''(s_v) + 2c_v'(s_v)\right)^{-1}}.$$

Proof. The proof is given in Appendix 3.H.

For instance, in the case of affine costs $c_t(x) = a_t + b_t x$, we get:

$$\theta_u(\boldsymbol{s}) = \frac{1}{2}, \ \ r_t(\boldsymbol{s}) = \frac{b_t^{-1}}{\sum_{v\in\mathcal{S}_n(\boldsymbol{s})} b_v^{-1}} .$$

In the case of two arcs $|\mathcal{T}| = 2$ that we denote $\mathcal{T} = \{t, \tilde{t}\}$ the problem is easier as, for each player n, the constraint $x_{n,t} + x_{n,\tilde{t}}$ can be considered directly in b_n as:

$$b_n(\boldsymbol{x}_n, \boldsymbol{s}) = x_{n,t} c_t(x_{n,t} + s_t) + (E_n - x_{n,t}) c_{\tilde{t}}(E - (x_{n,t} + s_t)) \tag{3.21}$$

which only depends on x_t, the load on arc t. One can see immediately that the bound constraints reduce to the two bounds on $x_{n,t}$ given by:

$$\ell_{n,t} \leqslant x_{n,t} \leqslant u_{n,t} \text{ where } \ell_{n,t} \stackrel{\text{def}}{=} \max(\underline{x}_{n,t}, E - \overline{x}_{n,\tilde{t}}) \text{ and } u_{n,t} \stackrel{\text{def}}{=} \min(\overline{x}_{n,t}, E - \underline{x}_{n,\tilde{t}}) \,. \tag{3.22}$$

We obtain the following result:

Proposition 3.3. *Assume that $\mathcal{T} = \{t, \tilde{t}\}$, and that cost functions $c_t, c_{\tilde{t}}$ are convex and strictly increasing. Then, for each player $n \in \mathcal{N}$, there exists two points $\underline{s}_{n,t} \leqslant \overline{s}_{n,t} \in \mathbb{R}$ such that:*

i) $\mathcal{B}_n(\boldsymbol{s}) = (u_{n,t}, E_n - u_{n,t})$ for any $\boldsymbol{s} = (s_t, s_{\tilde{t}})$ such that $s_t \leqslant \underline{s}_{n,t}$;

ii) $\mathcal{B}_n(\boldsymbol{s}) = (\ell_{n,t}, E_n - \ell_{n,t})$ for any $\boldsymbol{s} = (s_t, s_{\tilde{t}})$ such that $s_t \geqslant \overline{s}_{n,t}$;

iii) there exists $\alpha_n \in [0, \frac{1}{2})$ for which $\mathcal{B}_n$ is α_n contracting and $(\mathrm{Id} + \mathcal{B}_n)$ is $(1 - \alpha_n)$ contracting on $\mathcal{S}^\circ \stackrel{\text{def}}{=} \{\boldsymbol{s} = (s_t, E - E_n - s_t) \mid s_t \in [\underline{s}_{n,t}, \overline{s}_{n,t}]\}$.

Proof. The proof is given in Appendix 3.I.

The contraction in expectation stated in Proposition 3.2 leads to the following convergence property of the randomized BRD Algorithm 3.3.

Proposition 3.4. *Let $\boldsymbol{y}^{(0)} \stackrel{\text{def}}{=} (\boldsymbol{x}^{(0)}, \sum_n \boldsymbol{x}_n^{(0)})$ a starting point, $k \in \mathbb{N}^*$, and $I_1, \ldots, I_k$ be k independent random variables, uniformly distributed over $\mathcal{N}$, and $\boldsymbol{y}^*$ be the unique stationary point of the algorithm. Then, if Proposition 3.2 holds, we have:*

$$\mathbb{E}_{\boldsymbol{I} \overset{iid}{\sim} \mathcal{U}(\mathcal{N})} \left[\left\| \boldsymbol{y}^* - \boldsymbol{y}_{\boldsymbol{I}}^{(k)} \right\|_A^2 \right] \leqslant \rho^k \left\| \boldsymbol{y}^* - \boldsymbol{y}^{(0)} \right\|_A \,, \tag{3.23}$$

where $\boldsymbol{y}_{\boldsymbol{I}}^{(k)}$ denotes the random variable $Q_{I_k} \circ \cdots \circ Q_{I_1}(\boldsymbol{y}^{(0)})$, and $\rho \in (0, 1)$ is defined in Proposition 3.2.

Proof. To obtain the result, we just apply recursively the contraction property and use the independence of $I_1, \ldots, I_k$:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{I} \overset{\text{iid}}{\sim} \mathcal{U}(\mathcal{N})} \left[\left\| \boldsymbol{y}^* - \boldsymbol{y}_{\boldsymbol{I}}^{(k)} \right\|_A^2 \right] &= \mathbb{E}_{I_k} \mathbb{E}_{I_{k-1}} \ldots \mathbb{E}_{I_1} \left[\left\| \boldsymbol{y}^* - Q_{I_k} \circ \cdots \circ Q_{I_1}(\boldsymbol{y}^{(0)}) \right\|_A^2 \right] \\
&= \mathbb{E}_{I_{k-1}} \ldots \mathbb{E}_{I_1} \mathbb{E}_{I_k} \left[\left\| Q_{I_k}(\boldsymbol{y}^*) - Q_{I_k} \circ \cdots \circ Q_{I_1}(\boldsymbol{y}^{(0)}) \right\|_A^2 \right] \\
&\leqslant \rho \mathbb{E}_{I_{k-1}} \ldots \mathbb{E}_{I_1} \left[\left\| \boldsymbol{y}^* - Q_{I_{k-1}} \circ \cdots \circ Q_{I_1}(\boldsymbol{y}^{(0)}) \right\|_A^2 \right] \,,
\end{aligned}$$

which, recursively, gives the result of Proposition 3.4. □

Combining Proposition 3.4, Proposition 3.2, Proposition 3.3 and Lemma 3.2, we obtain immediately the convergence of Algorithm 3.3 in the two specific cases:

Corollary 3.2. *For a congestion game with costs functions $(c_t)_{t \in \mathcal{T}}$ on resources $\mathcal{T}$, the randomized Best response converges geometrically in the sense of Equation* (3.23) *in the two following cases:*

1. *if $\mathcal{T} = \{t, \tilde{t}\}$ and $c_t, c_{\tilde{t}}$ are strictly increasing and convex;*
2. *if $(c_t)_t$ are affine and strictly increasing functions.*

To our limited knowledge, it is an open question to know if the results of Corollary 3.2 can be extended to more general cases. Still considering convex and increasing resource cost functions, numerical tests show that Algorithm 3.3 also converges in the case of a set $\mathcal{T}$ of arbitrary length, and that the cyclic version of the algorithm (see Algorithm 3.1) also

converges geometrically. Extending our techniques to prove the convergence in those cases is not straightforward and would constitute interesting results.

3.3.4 Numerical Convergence and Comparisons

In this section, we present a numerical comparison of the two algorithms CBRD and SIRD given above. We also add two other algorithms from related papers in the comparison:

- the distributed iterative proximal-point algorithm [Che+14, Algo. 1], referred to as *it-ProxPt*. This algorithm is analogous to the SIRD algorithm proposed here, but with a decreasing time-step (as opposed to the constant step γ) and a regularization proximal term. Line 4 of Algorithm 3.2 is replaced with:

$$\boldsymbol{x}_n^{(k+1)} \leftarrow \mathrm{P}_{\mathcal{X}_n}\left[\boldsymbol{x}_n^{(k)} - \gamma_k\left(\nabla_n b_n(\boldsymbol{x}_n^{(k)}, \boldsymbol{x}_{-n}^{(k)}) + \theta(\boldsymbol{x}_n^{(k)} - \boldsymbol{x}_n^{(k-1)})\right)\right].$$

In the numerical results below, we choose $\gamma^{(k)} = k^{-0.52}$ (to ensure the convergence criterion of $\sum \gamma_k < \infty$ and $\sum \gamma_k^2 = +\infty$ while keeping sufficiently large steps) and a regularization weight $\theta = 0.5$ (according to our tests, this latter parameter does not have a significant impact on the speed of convergence);

- the proximal decomposition algorithm [Atz+13, Algo. 1] referred to as *proxBR*. This algorithm is analogous to the CBRD (Algorithm 3.1) proposed here, with a proximal regularization term and an additional loop to update this proximal term. Namely, the authors introduced a "regularized" game where each player's objective is replaced by:

$$f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) = b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) + \tfrac{\tau}{2}\left\|\boldsymbol{x}_n - \bar{\boldsymbol{x}}_n\right\|_2^2, \tag{3.24}$$

where $\bar{\boldsymbol{x}}_n$ is the "centroid" updated in an additional loop. The idea of the algorithm is to compute the NE of this regularized game, and update the centroid to the computed NE. Of course, the NE of the regularized game can only be computed approximately. In the numerical results below, we choose to update the centroid when the distance between the iterates of two BR cycles $\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\right\|_2$ is lower than 10^{-4}. The regularization parameter is taken to $\tau = 3(N-1)\max_t c_t$, just high enough to ensure the given condition of convergence (choosing a higher parameter τ slows the convergence). Note that CBRD corresponds to the case $\tau = 0$, a case dismissed by the proposed convergence conditions [Atz+13, Thm. 2].

We consider numerical instances with $T = 10$, and feasibility sets of the form (3.6), constructed as follows:

1. in instances **I1**, functions (c_t) are affine and uniform *i.e.* $\forall t, c_t(X) = c(X) = \alpha + \beta X$ where α (resp. β) is drawn uniformly from $[0, 4]$ (resp. $[1, 4]$). For each $n \in \mathcal{N}$, E_n is drawn uniformly from $[1, 10]$. The lower bounds are all set to $\underline{x}_{n,t} = 0$. A subset $\mathcal{T}_n$ of consecutive time periods of length $T_n \geqslant 4$ is drawn randomly from $\mathcal{T}$, and the upper bounds are set to $\overline{x}_{n,t} = E_n$ if $t \in \mathcal{T}_n$ and to $\overline{x}_{n,t} = 0$ if $t \notin \mathcal{T}_n$;

2. in instances **I2**, price functions are affine and time dependent *i.e.* $\forall t, c_t(X) = \alpha_t + \beta_t X$ where α_t (resp. β_t) is drawn uniformly from $[0, 4]$ (resp. $[1, 4]$). For each $n \in \mathcal{N}$, E_n is drawn uniformly from $[1, 10]$. A subset $\mathcal{T}_n$ of consecutive time periods of length $T_n \geqslant 4$ is drawn randomly from $\mathcal{T}$. For $t \notin \mathcal{T}_n$, we take $\underline{x}_{n,t} = \overline{x}_{n,t} = 0$. For $t \in \mathcal{T}_n$, $\underline{x}_{n,t}$ is drawn uniformly from $[0, E_n/T_n]$, and $\overline{x}_{n,t}$ is drawn uniformly from $[E_n/T_n, E_n]$, ensuring that $\mathcal{X}_n \neq \emptyset$.

The four algorithms considered above are implemented in Python 3.5 and run on an Intel i7 @2.6GHz on a single core and with 8GB of RAM. To solve the quadratic programs (QP) subproblems of each algorithm (Line 5 of CBRD and Line 4 of SIRD), we use the Brucker algorithm [Bru84]. Note that if $\mathcal{X}_n$ is a general polytope, any quadratic programming solver can be used instead to solve those subproblems.

Figure 3.1 shows the convergence of the four algorithms to the NE. The results are given on average on a set of ten instances **I1**. The convergence speed of the algorithms decreases with the number of users, as given in Theorem 3.5. We observe that, in spite of the weaker

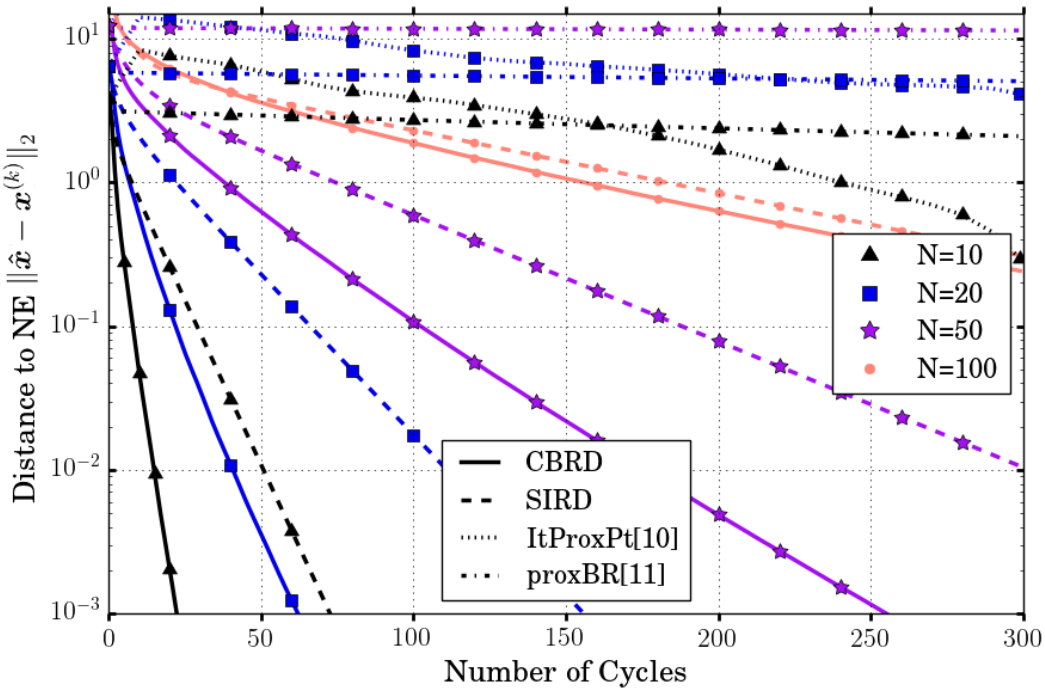


FIGURE 3.1: Average convergence rate of the four implemented algorithms on ten instances **I1**. *When the number of players N increases, the convergence rate of both algorithms decreases. CBRD is faster than SIRD.*

theoretical result for CBRD (see (3.15) compared to (3.16)), the convergence seems also geometric, and even faster than Algorithm SIRD. We also observe that the convergence rate of the comparison algorithms *itProxPt* and *proxBR* is not comparable with the convergence rate of CBRD and SIRD. This can be explained by the addition of the regularization term which slows the convergence and, also, by the diminishing step for ItProxPoint (instead of a constant step γ for SIRD) and by the additional loop for proxBR compared to CBRD.

N=	**10**	**20**	**50**	**100**
BR	0.15	1.4	23.98	112.1 TL(97)
SIRD	0.47	4.11	41.97	TL(100)
ItProxPt[Che+14]	2.07	17.18	89.3 TL(83)	TL(100)
proxBR[Atz+13]	57.68 TL(2)	98.4 TL(94)	TL(100)	TL(100)

(A) **I1** Uniform affine prices, bounds $\underline{x}_{n,t} = 0$, $\overline{x}_{n,t} \in \{0, E_n\}$

N=	**10**	**20**	**50**	**100**
BR	0.08	0.36	3.45	54.17 TL(3)
SIRD	2.24	12.08	95.5 TL(86)	TL(100)
ItProxPt[Che+14]	0.77	4.83	65.1 TL(30)	TL(100)
proxBR[Atz+13]	75.28 TL(4)	TL(100)	TL(100)	TL(100)

(B) **I2** Random affine prices, random bounds $\underline{x}_{n,t}, \overline{x}_{n,t}$

TABLE 3.1: Average CPU time (sec.) for NE computation for one hundred instances **I1** (a) and one hundred instances **I2** (b). *xx-TL(k) indicates that the time limit (120 sec.) was reached for k instances, while the remaining instances took an average CPU time of xx seconds.*

Table 3.1 shows the CPU time needed to compute the NE at a given precision: the stopping criterion considered here is the satisfaction of the KKT conditions for each problem (3.5) with an absolute error of 10^{-2}, with a time limit (TL) of two minutes per instance.

We observe that CBRD is the fastest method to compute the NE. The algorithm *ProxBR* reaches the time limit even for a relatively small number of consumers ($N = 20$). We notice that SIRD is slower on the heterogeneous instances **I2**, which can be easily explained from the step γ chosen in Theorem 3.5. On the contrary, CBRD and *itProxPt* are slower on homogeneous instances **I1**, which can be explained by the importance of symmetries in those instances. We can see that for the bigger instances ($N = 100$), the time limit is reached for half of the instances for CBRD (and for all instances for other algorithms). The time limit considered here was only of two minutes: it could be extended in a practical implementation of a DR program. However, the computational time can be limiting, for instance if the equilibrium needs to be recomputed in case of a change of parameters (see Section 3.4). Thus, using those methods for a larger system (thousands of consumers) might be prohibited, in particular if we allow a more complex description of users constraints than (3.6).

However, if we stay within the proposed order of magnitude ($N \leqslant 100$), the simulations show that CBRD (and SIRD in most cases) needs only a few seconds to compute the NE. It enables to consider an online procedure, where the equilibrium can be recomputed at each hour, as explained below.

3.4 Simulation of Online Demand Response

In this section, we propose a practical procedure to implement the DR framework described above. We assume that, as described in Section 3.2.3 Example 3.3, the aggregator prices come from a generic cost function $C(.)$ that depends on the total load $X_t^{\mathrm{NF}} + X_t$ (nonflexible plus flexible) at each time period. The flexible consumption is considered as an additional load and its price is set for each time period $t \in \mathcal{T}$ to:

$$c_t(X_t) \stackrel{\text{def}}{=} \frac{1}{X_t} \left(C(X_t^{\mathrm{NF}} + X_t) - C(X_t^{\mathrm{NF}}) \right) . \tag{3.25}$$

The equilibrium consumption for the flexible consumption profiles have to be computed before real-time consumption. As a result, the nonflexible demand $\boldsymbol{X}^{\mathrm{NF}}$ has to be estimated in order to evaluate price functions $(c_t)_t$ by injecting the estimation $\tilde{\boldsymbol{X}}_{\mathrm{NF}}$ in (3.25). To minimize the impact of forecast errors made on $\boldsymbol{X}^{\mathrm{NF}}$, we consider an online procedure in which, at each hour, an updated forecast $\tilde{\boldsymbol{X}}_{\mathrm{NF}}$ is taken into account. The equilibrium profiles for the flexible consumption is then re-computed for the hours ahead to the end of the optimization horizon T, using Algorithm 3.1 or Algorithm 3.2.

3.4.1 Online Demand Response Procedure

The initial time horizon $\mathcal{T}$ that we consider for the planning via DR starts each day at noon ($t = 1$) and stops at noon the day after ($t = T$), with an hourly time step. The "online" procedure computes the DR equilibrium flexible consumption profiles on time horizon $\{1, \dots T\}$ for each day. As the price functions c_t depend on the nonflexible load through (3.25), and as the accuracy of forecast of this load improves when approaching from real-time, we recompute the equilibrium using updated forecasts at each time period, as described below.

Algorithm 3.4 Online Demand Response Procedure

1: Start at $t = 1$
2: **while** $t \leqslant T$ **do**
3: Set new horizon $\mathcal{T}^{(t)} = \{t, t+1, \dots, T\}$
4: Get $\boldsymbol{X}^{\mathrm{NF}}$ forecast on $\mathcal{T}^{(t)}$: $\tilde{\boldsymbol{X}}^{(t)}_{\mathrm{NF}} \stackrel{\text{def}}{=} \left(\tilde{X}^{(t)}_{\mathrm{NF},s}\right)_{t \leqslant s \leqslant T}$
5: Re-compute prices $c_t(.)$ for $t \in \mathcal{T}^{(t)}$ from (3.25)
6: Run Algo. SIRD or BRD to compute NE $\boldsymbol{x}^{(t)}$ on $\mathcal{T}^{(t)}$
7: **for** each user $n \in \mathcal{N}$ **do**
8: Realize computed profile on time t, $x^{(t)}_{n,t}$
9: Update $\mathcal{X}_n^{(t+1)} \stackrel{\text{def}}{=} \left\{ (x_{n,s})_{s>t} | (x^{(t)}_{n,t}, [x_{n,s}]_{s>t}) \in \mathcal{X}_n^{(t)} \right\}$
10: **done**
11: Wait for $t+1$
12: **done**

Remark 3.9. *If one considers sets $(\mathcal{X}_n)_n$ of the form* (3.6)*, then the updating step on Line 9 only consists in updating the energy demand for the remaining time: $E_n^{(t+1)} \stackrel{\text{def}}{=} E_n^{(t)} - x^{(t)}_{n,t}$.*

Remark 3.10. *In practice, the NE profile $\boldsymbol{x}^{(t)}$ has to be computed before period t to begin consumption at time t (Line 8). If τ is an upper bound on the computation time of the NE profile (Line 6), then, as we want to use the latest available forecast, Lines 3-5 would be run just before $t - \tau$, Line 6 is run in the interval $[t - \tau, t]$ and Line 8 is executed through $[t, t+1]$.*

Observe that in Algorithm 3.4, considering an updated forecast at Line 4 leads to updated price functions $(c_t)_t$ (Line 5), according to equation (3.25). In turn, the updated price functions modify the objective function of user n, b_n, used in Line 6.

The difference of Algorithm 3.4 with an "offline" version is that we recompute the equilibrium consumption (Line 6) at each time for all the time periods ahead. In an offline DR, we would compute the equilibrium consumption for all the horizon $\mathcal{T} = \{1, \dots, T\}$ *only once*, just before $t = 1$.

Proceeding with this "online" version has two main advantages. First, it enables to rely on updated forecasts with new information acquired on the nonflexible load $\boldsymbol{X}^{\mathrm{NF}}$ (Line 4). Second, it also enables to cope with local issues as disconnection of an user or a communication bug: in that case, lines 8 and 9 would not be executed for the involved user, and this user will have the same energy demand for the next round at $t+1$. With this kind of online procedure, it is also important to ensure that the final realized profile $(x^{(t)}_{n,t})_{t \in \mathcal{T}}$ is globally consistent: in the limit of perfect forecasts, it has to correspond to an equilibrium of the initial game.

Theorem 3.6. *Suppose that either Assumption 3.2 with $a = 0$ holds (strict stability from Proposition 3.1) or that Assumption 3.1 holds and the sets $(\mathcal{X}_n)_n$ are of the form* (3.6)*. Then the online DR procedure of Algorithm 3.4 is consistent: if forecasts are perfect (i.e. $\forall t \in \mathcal{T}, \forall t' \in \mathcal{T}^{(t)}, \tilde{X}^{(t)}_{\mathrm{NF},t'} = X^{\mathrm{NF}}_{t'}$), then for any $t_2 > t_1$, the NE profile $\boldsymbol{x}^{(t_1)}$ computed at t_1 with forecast $\tilde{\boldsymbol{X}}^{(t_1)}_{\mathrm{NF}}$ is equal on $\{t_2, \dots, T\}$ to the NE profile $\boldsymbol{x}^{(t_2)}$ computed at t_2 with forecast $\tilde{\boldsymbol{X}}^{(t_2)}_{\mathrm{NF}}$.*

Proof: See Appendix 3.F.

Theorem 3.6 states a *dynamic programming principle* adapted to our game-theoretic framework. To quantify the value of this online procedure in the more realistic case of imperfect forecasts, we simulate it on a set of consumers and parameters taken from real data, defined below.

3.4.2 Consumers

We consider a set of $N = 30$ users owning an Electric Vehicle (EV) from the database of Texan residential consumers *PecanStreet Inc.* [Pec]. We consider that the charging of the EV is the only flexible appliance of each consumer managed through the DR program, while the remaining of the user's consumption is nonflexible and is taken as in the data. We denote by

$\mathcal{D} \stackrel{\text{def}}{=} \{16/01/01, \ldots, 16/01/31\}$ the set of the 31 days of January 2016 for which we simulate the DR program and we index a parameter by $d \in \mathcal{D}$ when it is specific to day d. For constraints (3.6a-3.6b), we take, for each day $d \in \mathcal{D}$, the total flexible demand of user n, $E_{n,d}$ as the total observed consumption for the EV of n on the time set $\mathcal{T} = \{1, \ldots, T\}$, taken as the twenty-four hours from day d 12PM to day $d+1$ 11AM (including the regular EV residential night charging period). The power lower bound is always taken to zero $\underline{x}_{n,d,t} = 0$. For the power upper bound $\overline{x}_{n,d,t}$, we consider two cases: if a positive power was given at d, t in the data, we set $\overline{x}_{n,d,t}$ to the maximum power given to n's EV on all time periods in the data in the set $\mathcal{D}$. If the power given to the EV is 0 at d, t in the data, we take $\overline{x}_{n,d,t} = 0$ (*i.e.* we consider that the EV of n was not available to charge on period d, t).

3.4.3 Price Functions and Forecasts of the Nonflexible Load

As in Chapter 2, we consider that the aggregator has a providing cost for the global demand at time t, $X_t^{\text{tot}} \stackrel{\text{def}}{=} (X_t^{\text{NF}} + X_t)$, that does not depend on the time, and given (in \$) by

$$\tilde{C}(X_t^{\text{tot}}) \stackrel{\text{def}}{=} 0.711 - 0.0417 X_t^{\text{tot}} + 0.00295 (X_t^{\text{tot}})^2 \,, \tag{3.26}$$

where the coefficients are chosen to replicate the cost function of a real residential electricity provider[4]. For this, we computed the average, minimum and maximum values of X_t^{NF} over all the hours of the 31 days of January 2016 on our set of 30 consumers and interpolate the three values (avg X_t^{NF}, min X_t^{NF}, max X_t^{NF}) to three respective prices proposed by the Texan distributor *Coserv* [Cos] so that the per-unit price $\tilde{c}(D) \stackrel{\text{def}}{=} \tilde{C}(D)/D$ verifies:

- $\tilde{c}(\text{avg } X_t^{\text{NF}}) = 0.080\$/\text{kWh}$ (price for "base" contracts);
- $\tilde{c}(\min X_t^{\text{NF}}) = 0.055\$/\text{kWh}$ (price for Off-Peak hours in Time-of-Use contracts);
- $\tilde{c}(\max X_t^{\text{NF}}) = 0.14\$/\text{kWh}$ (price for Peak hours).

Following (3.25), the price for the flexible load is given by:

$$c_t(X_t) = (-4.17 + 0.590 X_t^{\text{NF}}) + 0.295 X_t \,, \tag{3.27}$$

hence, Assumption 3.3 holds.

As prices depend on the nonflexible load, the aggregator has to compute a forecast $\tilde{\boldsymbol{X}}_{\text{NF}}^{(t)} \stackrel{\text{def}}{=} (\tilde{X}_{\text{NF},t}^{(t)}, \ldots \tilde{X}_{\text{NF},T}^{(t)})$ at each time t so that the equilibrium consumption for time periods $\{t, \ldots, T\}$ can be computed using Algorithm 3.1 or Algorithm 3.2. To simulate the forecasts, we assume that the forecast made at time t for period $t' \geqslant t$, $\tilde{X}_{\text{NF},t'}^{(t)}$ has no bias, that is $\mathbb{E}[X_{t'}^{\text{NF}} | \sigma(\mathcal{F}_t)] = \tilde{X}_{\text{NF},t'}^{(t)}$ (where $\mathcal{F}_t$ is the natural filtration over $(X_t^{\text{NF}})_t$), and that we have perfect information at time t, that is: $\tilde{X}_{\text{NF},t}^{(t)} = X_t^{\text{NF}}$. Considering that $X_t^{\text{NF}} = P_t e^{X_t}$ where X_t follows an Ornstein-Uhlenbeck [UO30] process with mean reverting coefficient m and volatility σ, and P_t a seasonality factor that depends on the hour of the week (1st hour to 168th hour), we get for any $t \leqslant t'$:

$$\tilde{X}_{\text{NF},t'}^{(t)} = P_{t'} \left(\frac{X_t^{\text{NF}}}{P_t} \right)^{e^{-m(t'-t)}} \exp\left(\frac{\sigma^2}{4m} (1 - e^{-2m(t'-t)}) \right).$$

Using a least-squares regression on the observed data from years 2014 and 2015, we compute $m \simeq 0.198\ \text{h}^{-1}$ and $\sigma \simeq 0.117\ \text{h}^{-1/2}$. An example of the simulated forecasts made at four different time periods is given in Figure 3.2.

3.4.4 Gains with the Online DR Procedure

For each day of January 2016, we run the online DR Procedure described in Section 3.4.1 to get the flexible consumption profile of each user $\boldsymbol{x}_n$, and the associated social cost on the

[4] Although we still consider a quadratic cost model, the calibration of the price coefficients used here is more sophisticated than the one used in Chapter 2.

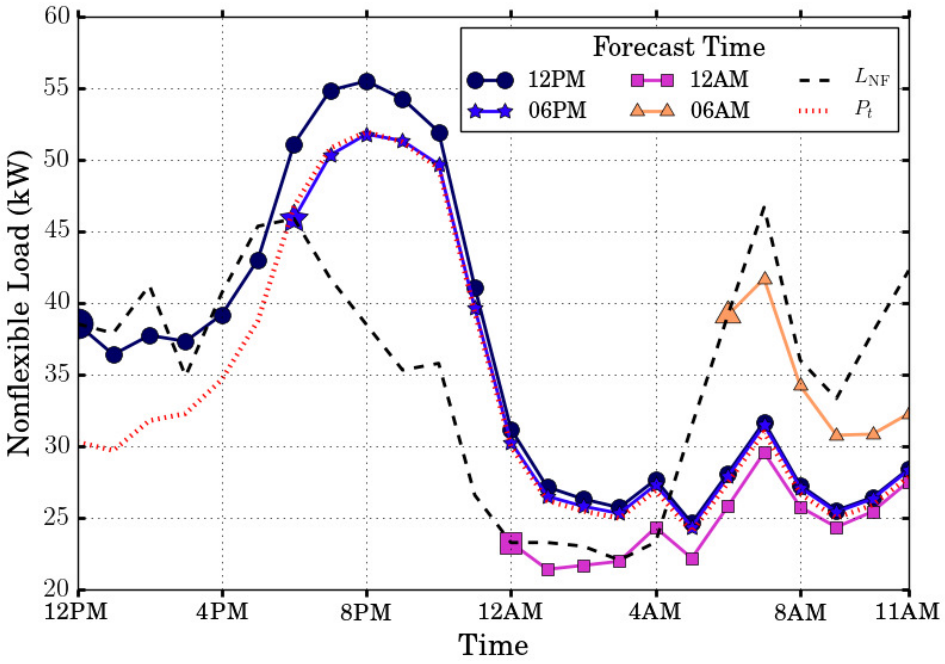


FIGURE 3.2: Forecasts of the nonflexible load $\tilde{X}_{\text{NF}}^{(t)}$ evolving in time. *We assume a perfect forecast at time t for t. Forecasting performance increases when approaching real time.*

DR horizon $\{1, \dots, T\}$. We compare the total social cost over the set $\mathcal{D}$ of simulated days obtained via the DR online procedure with the total social costs obtained with the four other consumption scenarios below:

1. *uncoordinated* case: no DR is implemented to control or incentivize consumers flexibility; the flexible consumption profiles are taken as the observed values in the data;

2. *offline DR*: the equilibrium is computed only once at $t = 1$ and for the whole time horizon $\{1, \dots, T\}$ considering the first forecast $\tilde{\boldsymbol{X}}_{\text{NF}}^{(1)}$ available at $t = 1$;

3. *perfect forecast DR*: offline DR, where we take $\tilde{\boldsymbol{X}}_{\text{NF}}^{(1)} = \boldsymbol{X}^{\text{NF}}$. With Theorem 3.6, it is useless to recompute the profiles at each time period;

4. *optimal* scenario: a centralized entity (with perfect forecasts) computes the flexible consumption profile $\boldsymbol{x}$ that minimizes the system cost $\sum_t X_t c_t(X_t)$ (also equal to the social cost, Remark 3.3).

For the online DR and the comparison scenarios 2) and 3), NE are computed by implementing Algorithm 3.1 (CBRD) with the same configuration than in Section 3.3.4. For the comparison scenario 4), we compute the *optimal* consumption profile satisfying all users constraints (3.6a-3.6b), for each simulated day (from 12PM to 11AM) in $\mathcal{D}$. The associated problems are convex QPs with linear constraints, that are solved easily with the solver Cplex 12.6 in 0.31seconds on average.

Table 3.2 summarizes the numerical results: it gives the total costs on the 31 days of January 2016 and compares the gains of the different flexible consumption scenarios relatively to the *uncoordinated* one. We first observe that performances of perfect forecasts DR are close to the optimal scenario. This confirms the theoretical results provided on the efficiency of the NE in Theorem 3.3. We see on this table that the online DR procedure achieves significant savings compared to the offline version for which the performance is really low on average: using the offline DR decreases the system costs by 2% relatively to the uncoordinated profile, that is, when consumers behave without any incentives (comparison scenario

Cons. Scenario	Social Cost	Avg. Price	Gain
Uncoordinated	$ 1257.2	0.200 $/kWh	—
Offline DR	$ 1231.6	0.195 $/kWh	2.036%
Online DR	**$ 1131.1**	0.180 $/kWh	**10.03%**
Perfect forecast DR	$ 1075.2	0.171 $/kWh	14.47%
Optimal scenario	$ 1056.8	0.169 $/kWh	15.94%

TABLE 3.2: Social Costs, average prices and relative gain to the uncoordinated consumption scenario on January 2016 .

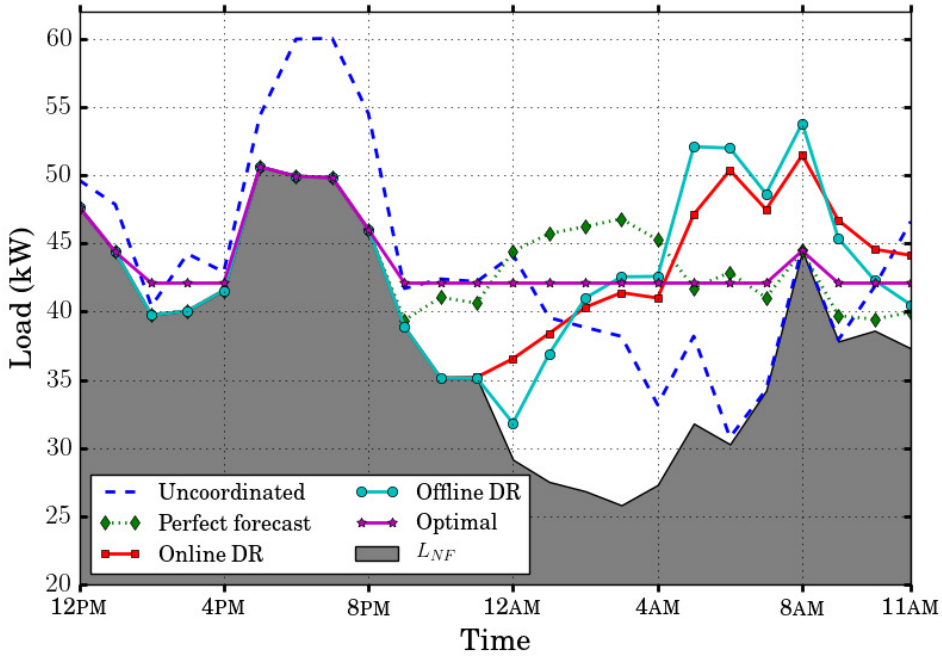


FIGURE 3.3: Consumption profiles on a typical day, with the different scenarios listed in Section 3.4.4. *The optimal profile flattens the consumption. The online DR procedure of Algorithm 3.4 gets closer to the Perfect forecast (offline) DR profile.*

1)). Implementing this offline DR program might not be worthy as it still involves a sophisticated communication and automation structure and it adds more constraints for consumers. This low performance is directly linked to our simple and naive model for the nonflexible load forecasts, which results in inaccurate forecasts for the last hours, as seen in Figure 3.2. Even if more advanced forecasting methods (see [AN02]) could improve this accuracy, we cannot get rid of the high variance due to the small number of consumers ($N = 30$ in our example, and several hundreds for an aggregator at the scale of a typical low-voltage station). The online DR procedure seems to bring a solution to this issue: even with our simple forecast model, we achieve more than 10% of savings, with a gap of only 4% from the scenario with perfect forecasts. These results show that implementing the given online DR procedure, even without very accurate forecasts, is worthwhile for the aggregator.

3.5 Conclusion

In this chapter, we developed a game-theoretic model for a residential demand response program, and we addressed several issues both on the theoretical and practical aspects. We gave several new theoretical results about the uniqueness and existence of a Nash equilibrium consumption profile for which the price of anarchy is theoretically bounded. We proved

that the two proposed algorithms CBRD and SIRD provide approximations of the NE at an arbitrary accuracy in finite time. We introduced and simulated an online procedure that recomputes the NE profiles at each time period to take into account new information, for example updated forecasts. We showed numerically that this online procedure achieves a small price of anarchy when the parameters are fixed but also when the demand is uncertain. Our simulations show that the online procedure reduces the impact of inaccurate forecasts on the social cost by 8%.

Several extensions of this work can be undertaken. First, our online procedure can be directly applied in the presence of other sources of stochasticity such as market prices or local renewable production sources. The aggregator objective can also be generalized to take into account the distance to a reference load profile or to maximize consumption during renewable production peaks or to take into account market prices. Also, two main theoretical questions are still open. First, the result on the PoA bound could be improved to be tighter to the numerical results, and generalized to a larger set of functions. Second, the convergence theorem for the Best Response Dynamics (CBRD) could also be improved, as the observed convergence rate is faster than the given bound, and the convergence is numerically observed for a larger set of prices than affine functions.

Appendix

3.A Proof of Theorem 3.2: Uniqueness of NE in $\mathcal{G}$

The proof follows the one of [ORS93], extending it to the constrained case with constraints of the form (3.6b).

We denote by $\lambda_n \in \mathbb{R}$ the Lagrange multiplier associated to (3.6a), along with $\underline{\mu}_{n,t} \geqslant 0$ (resp. $\overline{\mu}_{n,t} \geqslant 0$) the multiplier associated to $\underline{x}_{n,t} \leqslant x_{n,t}$ (resp. to $x_{n,t} \leqslant \overline{x}_{n,t}$).

Note that the KKT conditions give that, at optimality:

$$\gamma_{n,t}(x_{n,t}, X_t) = \lambda_n + \underline{\mu}_{n,t} - \overline{\mu}_{n,t} \,, \tag{3.28}$$

where $\gamma_{n,t}(x_{n,t}, X_t) \stackrel{\text{def}}{=} c_t(X_t) + x_{n,t} c_t'(X_t)$ is the marginal cost of n. Let us consider $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ two NEs. From (3.28), we get:

$$\begin{aligned} x_{n,t} < \overline{x}_{n,t} \Rightarrow \overline{\mu}_{n,t} = 0 \Rightarrow \gamma_{n,t}(x_{n,t}, X_t) \geqslant \lambda_n \\ \text{and } x_{n,t} > \underline{x}_{n,t} \Rightarrow \underline{\mu}_{n,t} = 0 \Rightarrow \gamma_{n,t}(x_{n,t}, X_t) \leqslant \lambda_n \end{aligned}$$

and the same inequalities hold for $\hat{x}$. First note that:

$$\left(\hat{\lambda}_n \leqslant \lambda_n \text{ and } \hat{X}_t \geqslant X_t\right) \Rightarrow \hat{x}_{n,t} \leqslant x_{n,t} \,, \tag{3.29}$$

$$\left(\hat{\lambda}_n \geqslant \lambda_n \text{ and } \hat{X}_t \leqslant X_t\right) \Rightarrow \hat{x}_{n,t} \geqslant x_{n,t} \,. \tag{3.30}$$

Let us show (3.29). If $\hat{x}_{n,t} = \underline{x}_{n,t}$ or $x_{n,t} = \overline{x}_{n,t}$, then $\hat{x}_{n,t} \leqslant x_{n,t}$ is clear. Else, $\hat{x}_{n,t} > \underline{x}_{n,t}$ and $x_{n,t} < \overline{x}_{n,t}$ so:

$$\gamma_{n,t}(\hat{x}_{n,t}, \hat{X}_t) \leqslant \hat{\lambda}_n \leqslant \lambda_n \leqslant \gamma_{n,t}(x_{n,t}, X_t) \leqslant \gamma_{n,t}(x_{n,t}, \hat{X}_t) \tag{3.31}$$

as $\gamma_{n,t}$ is increasing in X_t. As $c_t'(\hat{X}_t) > 0$ from Assumption 3.1, $\gamma_{n,t}$ is increasing in $x_{n,t}$ and we deduce that $x_{n,t} \geqslant \hat{x}_{n,t}$.

Now, let us consider $\mathcal{T}_1 = \{t \,:\, \hat{X}_t > X_t\}$ along with $\mathcal{T}_2 = \mathcal{T} \setminus \mathcal{T}_1 = \{t \,:\, \hat{X}_t \leqslant X_t\}$ and $\mathcal{N}_0 = \{n \,:\, \hat{\lambda}_n > \lambda_n\}$. Suppose $\mathcal{T}_1 \neq \emptyset$. From constraint (3.6a) and from (3.30), we have:

$$\forall n \in \mathcal{N}_0, \; \sum_{t \in \mathcal{T}_1} \hat{x}_{n,t} = E_n - \sum_{t \in \mathcal{T}_2} \hat{x}_{n,t} \leqslant E_n - \sum_{t \in \mathcal{T}_2} x_{n,t} = \sum_{t \in \mathcal{T}_1} x_{n,t} \,.$$

On the other hand, considering for $t \in \mathcal{T}_1$ and $n \notin \mathcal{N}_0$, we have from (3.29) that $\hat{x}_{n,t} \leqslant x_{n,t}$, and thus:

$$\sum_{t\in\mathcal{T}_1} \hat{X}_t = \sum_{t\in\mathcal{T}_1}\sum_{n\in\mathcal{N}_0} \hat{x}_{n,t} + \sum_{t\in\mathcal{T}_1}\sum_{n\notin\mathcal{N}_0} \hat{x}_{n,t} \leqslant \sum_{t\in\mathcal{T}_1} X_t \,, \tag{3.32}$$

which is in contradiction with the definition of $\mathcal{T}_1$. Thus $\mathcal{T}_1 = \emptyset$ and $\forall t,\ \hat{X}_t = X_t$. We can now precise (3.29) with:

$$\big[\hat{\lambda}_n{<}\lambda_n \text{ and } \hat{X}_t{=}X_t\big] \Longrightarrow \big[\hat{x}_{n,t}{<}x_{n,t} \text{ or } \hat{x}_{n,t}{=}x_{n,t}{=}\underline{x}_{n,t} \text{ or } \hat{x}_{n,t}{=}x_{n,t}{=}\overline{x}_{n,t}\big] \tag{3.33}$$

and similarly for (3.30). Indeed, if $\hat{x}_{n,t} = \underline{x}_{n,t}$ (resp. if $x_{n,t} = \overline{x}_{n,t}$) then the implication holds because $x_{n,t} \geqslant \underline{x}_{n,t}$ (resp. $x_{n,t} \leqslant \overline{x}_{n,t}$). Else, $\hat{x}_{n,t} > \underline{x}_{n,t}$ and $x_{n,t} < \overline{x}_{n,t}$, and the same sequence of inequalities as in (3.31) gives $\gamma_{n,t}(\hat{x}_{n,t}, X_t) < \gamma_{n,t}(x_{n,t}, X_t)$, implying that $\hat{x}_{n,t} < x_{n,t}$.

Finally, suppose that there exists n s.t. $\hat{\lambda}_n < \lambda_n$. If only the two latter cases in (3.33) happen, then $x_{n,t} = \hat{x}_{n,t},\ \forall t$. Else, there is at least one t for which $\hat{x}_{n,t} < x_{n,t}$, so $E_n = \sum_t \hat{x}_{n,t} < \sum_t x_{n,t} = E_n$ which cannot happen. Thus, $\hat{\lambda}_n = \lambda_n$ for all n and (3.29) and (3.30) imply that $x_{n,t} = \hat{x}_{n,t}$ for all n and t.

3.B Proof of Theorem 3.3: PoA upper bound

The proof relies on the notion of *local smoothness* introduced in [RS15]. The idea is to get a tighter bound than [RS15] by specifying the parameters of the affine price functions $(c_t)_t$ and by using the upper bound on X_t instead of looking at the worst possible cases as done in [RS15].

Let $\kappa_t \overset{\text{def}}{=} \alpha_t/\beta_t$ so that $c_t(X) = \beta_t(X + \kappa)$. From [RS15], we know that if there exist $\lambda, \mu > 0$ and a profile $\boldsymbol{y} \in \mathcal{X}$ satisfying for each $t \in \mathcal{T}$ and each $\boldsymbol{x} \in \mathcal{X}$:

$$Y_t(X_t + \kappa_t) + \frac{Y_t^2}{4} \leqslant \lambda Y_t(Y_t + \kappa_t) + \mu X_t(X_t + \kappa_t), \tag{3.34}$$

where $Y_t = \sum_n y_{n,t}$ and $X_t = \sum_n x_{n,t}$, then $\mathcal{G}$ is locally λ, μ-smooth for $\boldsymbol{y}$, *i.e.* for any admissible profile $\boldsymbol{x} \in \mathcal{X}$:

$$\sum_{n\in\mathcal{N}} b_n(\boldsymbol{x}) + \nabla_n b_n(\boldsymbol{x})^\top(\boldsymbol{y}_n - \boldsymbol{x}_n) \leqslant \lambda \mathrm{SC}(\boldsymbol{y}) + \mu \mathrm{SC}(\boldsymbol{x}) \,,$$

where $\mathrm{SC}(\boldsymbol{x}) = \sum_n b_n(\boldsymbol{x})$. In that case, it follows from [RS15] that the PoA is bounded by $\lambda/(1-\mu)$. We want to find the best possible λ, μ such that (3.34) holds for each $t \in \mathcal{T}$. For the remaining of the proof, we fix t and omit subscript t in the notations. As done in [RS15], we introduce:

$$\phi_{xy}(\mu) \overset{\text{def}}{=} \frac{y(x+\kappa) + \frac{y^2}{4} - \mu x(x+\kappa)}{y(y+\kappa)},\ \lambda^*(\mu) \overset{\text{def}}{=} \sup_{x,y\geqslant 0} \phi_{xy}(\mu) \,.$$

$\lambda^*(\mu)$ is the minimum value of $\lambda > 0$ such that (3.34) holds with values (λ, μ). Let us compute an explicit expression of $\lambda^*(\mu)$. If $x = 0$, $\phi_{0,y}(\mu) = \frac{y+4b}{4(y+\kappa)}$ and $\dfrac{\partial \phi_{0,y}}{\partial y} < 0$ so $\sup_{x,y} \phi_{x,y}$ would be attained with $y = 0$ and is $\phi_{0,0} = 1$. Otherwise:

$$0 = \frac{\partial \phi}{\partial x} \Leftrightarrow \frac{1}{y(y+\kappa)}(y - 2\mu x - \mu\kappa) \Rightarrow x = \frac{y - \kappa\mu}{2\mu}$$

but as $x \geqslant 0$, this supposes that $y \geqslant \mu\kappa$. We compute:

$$\phi_{\frac{y-\kappa\mu}{2\mu},y} = \frac{1}{y(y+\kappa)4\mu}\left((y+\kappa\mu)^2 + \mu y^2\right) \overset{\text{def}}{=} h(y) \,.$$

We can see that h' vanishes on $\mathbb{R}_+$ at $y_+ \stackrel{\text{def}}{=} \frac{\kappa\mu^2+\kappa\mu\sqrt{\mu^2+1-\mu}}{1-\mu}$ that gives a min of h so h is decreasing then increasing. At the lower bound $y = \kappa\mu$, we get:

$$\phi = \frac{\kappa\mu + 4b}{4(\kappa\mu+\kappa)} = \frac{\mu+4}{4(\mu+1)} = \frac{1}{4} + \frac{3}{4(\mu+1)} < 1\,,$$

which is not the maximal value as $\phi_{0,0} = 1$. At the upper bound $y = \overline{X}$, we have:

$$h(\overline{X}) = \frac{(\overline{X}+\kappa\mu)^2 + \mu\overline{X}^2}{\overline{X}(\overline{X}+\kappa)4\mu} = \lambda^*(\mu)\,.$$

Last, to compute the best bound $\inf_\mu \lambda^*(\mu)/(1-\mu)$, let us consider:

$$g(\mu) \stackrel{\text{def}}{=} 4\overline{X}(\overline{X}+\kappa)\frac{\lambda^*(\mu)}{1-\mu} = \frac{(\overline{X}+\kappa\mu)^2+\mu\overline{X}^2}{\mu(1-\mu)}\,.$$

If we denote $\varphi \stackrel{\text{def}}{=} (1+r)^2$ and $r \stackrel{\text{def}}{=} \kappa/\overline{X}$, $g(\mu)$ is minimal at $\mu^* \stackrel{\text{def}}{=} (-1+\sqrt{1+\varphi})/\varphi$. We finally get our PoA bound as:

$$\begin{aligned}\frac{\lambda^*(\mu^*)}{1-\mu^*} &= \frac{(3+2r)+2\sqrt{1+\varphi}}{4(1+r)} = \frac{1}{2}\left(1+\sqrt{1+\frac{1}{\varphi}}+\frac{1}{2\sqrt{\varphi}}\right)\\ &= \tfrac{1}{2}\left(1+\sqrt{1+(1+r)^{-2}}+(2(1+r))^{-1}\right) \leqslant 1+\tfrac{3}{4(1+r)}\,,\end{aligned}$$

where the last inequality, giving a more explicit bound, is obtained from the usual inequality:

$$\sqrt{a^2+b^2} \leqslant a+b \text{ for any } a,b \geqslant 0\,.$$

Next, following [RS15], for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ (admissible solutions):

$$\begin{aligned}&\textstyle\sum_n b_n(\boldsymbol{x}) + \nabla_n b_n(\boldsymbol{x})^\top(\boldsymbol{y}-\boldsymbol{x})\\ &= \textstyle\sum_n \sum_{t\in\mathcal{T}} x_{n,t}\cdot c_t(X_t) + (y_{n,t}-x_{n,t})\,(c_t(X_t)+x_{n,t}c_t'(X_t))\\ &= \textstyle\sum_t Y_t\cdot c_t(X_t) + c_t'(X_t)\sum_n \left(y_{n,t}x_{n,t} - {x_{n,t}}^2\right)\\ &\leqslant \textstyle\sum_t Y_t\cdot c_t(X_t) + c_t'(X_t)\cdot\frac{{X_t}^2}{4}\\ &= \textstyle\sum_t \beta_t\left[Y_t(X_t+\kappa_t)+\frac{{X_t}^2}{4}\right] \qquad (3.35)\\ &\leqslant \textstyle\sum_t \beta_t\left[\lambda Y_t(Y_t+\kappa_t)+\mu X_t(X_t+\kappa_t)\right] \qquad (3.36)\\ &= \lambda\mathrm{SC}(\boldsymbol{y}) + \mu\mathrm{SC}(\boldsymbol{x})\,,\end{aligned}$$

where (3.36) is valid if (λ,μ) is chosen such that:

$$\forall t\in\mathcal{T},\ \lambda \geqslant \frac{(\overline{X}_t+\kappa_t\mu)^2+\mu\overline{X}_t^2}{\overline{X}_t(\overline{X}_t+\kappa_t)4\mu} \stackrel{\text{def}}{=} \lambda^*_{\kappa_t}(\mu)\,.$$

Let us denote $t_0 \stackrel{\text{def}}{=} \operatorname*{argmin}_t \kappa_t$ and choose $\mu^* \stackrel{\text{def}}{=} \mu^*_{t_0}$, $\lambda^* \stackrel{\text{def}}{=} \lambda^*_{\kappa_{t_0}}(\mu^*)$ (the optimal (λ,μ) for t_0), then we have to check that for all $t\in\mathcal{T}$, $\lambda^* \geqslant \lambda^*_{\kappa_t}(\mu^*)$. For that, it is sufficient to show that $r \mapsto \lambda^*_r(\mu^*)$ is decreasing on $[r_{t_0}, r_t]$, which is true if

$$r_t < -1+\sqrt{1+\frac{1-\mu^*}{{\mu^*}^2}} \iff \varphi_{r_t} < \varphi_{r_{t_0}} + 2 + \sqrt{1+\varphi_{r_{t_0}}}\,,$$

where $\varphi_r = (1+r)^2$, which gives condition (3.9) stated in Theorem 3.3.

3.C Proof of Proposition 3.1: Strong Stability of $\mathcal{G}$

We denote by $G(\boldsymbol{x}) \stackrel{\text{def}}{=} J_F(\boldsymbol{x})$ the Jacobian of operator $F = (\nabla_n b_n)_{n\in\mathcal{N}}$. Since functions b_n are twice differentiable, condition (3.13) is equivalent to having the matrix $G(\boldsymbol{x}) + G^\top(\boldsymbol{x})$ positive definite for all $\boldsymbol{x} \in \mathcal{X}$, that is, $G(\boldsymbol{x}) + G(\boldsymbol{x})^\top \succ 0$.

As $b_n = \sum_t b_{n,t}$, with $b_{n,t}(\boldsymbol{x}_t) \stackrel{\text{def}}{=} x_{n,t} c_t(X_t)$, is separable in t, we can re-index the matrix $G(\boldsymbol{x})$ to have a diagonal block hourly matrix $G(\boldsymbol{x}) = \text{diag}(G_1, ...G_T)$, with $G_t(\boldsymbol{x}_t) \stackrel{\text{def}}{=} \left(\frac{\partial^2 b_{n,t}}{\partial x_{n,t}\partial x_{m,t}}(\boldsymbol{x}_t)\right)_{n,m\in\mathcal{N}^2}$ and we get for each $t \in \mathcal{T}$:

$$G_t(\boldsymbol{x}_t) + G_t(\boldsymbol{x}_t)^\top = \left(\frac{\partial^2 b_{n,t}(\boldsymbol{x}_t)}{\partial x_{m,t}\partial x_{n,t}} + \frac{\partial^2 b_{m,t}(\boldsymbol{x}_t)}{\partial x_{n,t}\partial x_{m,t}}\right)_{n,m}.$$

Let $t \in \mathcal{T}$ and $\boldsymbol{x} \in \mathbb{R}^N \setminus \{0\}$. Furthermore, let $[\sigma(\boldsymbol{x},\boldsymbol{x})]_t \stackrel{\text{def}}{=} \boldsymbol{x}_t^\top \left(G_t(\boldsymbol{x}_t) + G_t^\top(\boldsymbol{x}_t)\right)\boldsymbol{x}_t$. For notation simplicity, let us forget the index t and the argument (X) in functions c_t. We have:

$$\begin{aligned}[\sigma(\boldsymbol{x},\boldsymbol{x})]_t &= \sum_{n=1}^{N} 2x_n^2(x_n c''+2c') + 2\sum_{n<m} x_n x_m\left((x_n+x_m)c'' + 2c'\right)\\ &= \sum_{n=1}^{N} 2x_n^2\left(r_n\gamma + (1-r_n)a\right) + 2\sum_{n<m} x_n x_m\left((r_n+r_m)\gamma + (1-r_n-r_m)a\right)\end{aligned}$$

with $r_n \stackrel{\text{def}}{=} x_n/X$, $a = 2c'(X)$ and $\gamma \stackrel{\text{def}}{=} 2c'(X) + Xc''(X)$. Then:

$$\sigma = a\sum_n x_n^2 + a\Big(\sum_n \left(1 - r_n\left(1-\tfrac{\gamma}{a}\right)\right)x_n\Big)^2 - \tfrac{(a-\gamma)^2}{a}\sum_{n,m} r_n r_m x_n x_m$$

which is the sum of three quadratic forms: $q_1(\boldsymbol{x}) = a\sum x_n^2$ which has one eigenvalue a of multiplicity N, $q_2(\boldsymbol{x}) = a\left(\boldsymbol{x}^\top \boldsymbol{v}^\top \boldsymbol{v}\boldsymbol{x}\right)$ with $v_n \stackrel{\text{def}}{=} \sum_n 1 - \frac{x_n}{X}\left(1-\frac{\gamma}{a}\right)$ of rank one whose nonzero eigenvalue is $a||\boldsymbol{v}||_2^2$, and a negative form of rank one $q_3(\boldsymbol{x}) = -\frac{1}{a}(a-\gamma)^2\left(\sum_{n,m}\frac{x_n}{X}\frac{x_m}{X}x_n x_m\right)$ whose nonzero eigenvalue is $-\frac{1}{a}(a-\gamma)^2\sum_n\left(\frac{x_n}{X}\right)^2$.

We deduce that the quadratic form $q_1 + q_2$ is positive definite, and that its eigenvalues are a with multiplicity $N-1$ and $a(1+||\boldsymbol{v}||_2^2)$ with multiplicity 1. Next, we use the following result from perturbation theory:

Theorem 3.7 (Horn and Johnson,p. 367)**.** *Let $A, E \in \mathcal{M}_n$ be two Hermitian matrices and let $\lambda_1^M \leqslant ... \leqslant \lambda_n^M$ denotes the (real) ordered eigenvalues of an Hermitian matrix M. Then the following inequalities hold:*

$$\begin{aligned}\forall k = 1\ldots n,\ &\lambda_1^E \leqslant \lambda_k^{A+E} - \lambda_k^A \leqslant \lambda_n^E\\ &\text{and}\ \left|\lambda_k^{A+E} - \lambda_k^A\right| \leqslant \rho(E) = |||E|||_2\,.\end{aligned}$$

Applying this theorem with $A = q_1 + q_2$ and perturbation $E = q_3$ we get that the smallest eigenvalue λ_1^{A+E} of σ verifies:

$$\begin{aligned}\lambda_1^{A+E} &\geqslant \min\{\text{Sp}(q_1+q_2)\} - \tfrac{(a-\gamma)^2}{a}\sum_n r_n^2\\ &= a\Big(1-\big(1-\tfrac{\gamma}{a}\big)^2\textstyle\sum_n r_n^2\Big).\end{aligned}$$

Replacing a and γ, we can get the condition of Assumption 3.2.

3.D Proof of Theorem 3.4: Convergence of CBRD

The key of the proof is that, under Assumption 3.3, the game is an exact potential game [MS96] with convex potential:

$$\Phi(\boldsymbol{x}) = \textstyle\sum_{t\in\mathcal{T}} a_t X_t + \frac{\beta_t}{2}(X_t^2 + \sum_n x_{n,t}^2)\,,$$

that is, for any $\boldsymbol{x} \in \mathcal{X}$ and any n, $\nabla_n \Phi(\boldsymbol{x}) = \nabla_n b_n(\boldsymbol{x})$.

Thus, the NE corresponds to the minimum of Φ and we have, for any $\boldsymbol{x} \in \mathcal{X}$,

$$\operatorname*{argmin}_{\boldsymbol{x}_n\in\mathcal{X}_n} b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) = \operatorname*{argmin}_{\boldsymbol{x}_n\in\mathcal{X}_n} \Phi(\boldsymbol{x}_n, \boldsymbol{x}_{-n}).$$

Therefore, running Algorithm 3.1 is equivalent to performing an alternating block coordinate minimization on Φ. According to [Hon+17, Thm. 6.1]:

$$\Phi(\boldsymbol{x}^{(k)}) - \Phi(\hat{\boldsymbol{x}}) \leqslant \tfrac{1}{k} \times 2MN^2R^2\Omega\,, \tag{3.37}$$

with $M = \max_n M_n = 2\max_t \beta_t$ (max of Lipschitz constants of $\nabla_n b_n = \nabla_n \Phi$), $R = \max_{\boldsymbol{x}}\{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \;; \Phi(\boldsymbol{x}) \leqslant \Phi(\boldsymbol{x}^{(0)})\}$ and $\Omega = \max\{\frac{2}{MN^2R^2} - 2,\ \Phi(\boldsymbol{x}^{(1)}) - \Phi(\hat{\boldsymbol{x}}),\ 2\}$. But Φ is also strongly convex, that is, for any $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$:

$$\Phi(\boldsymbol{x}) - \Phi(\boldsymbol{x}') \geqslant \langle \nabla\Phi(\boldsymbol{x}'), \boldsymbol{x} - \boldsymbol{x}'\rangle + \tfrac{a}{2}\,\|\boldsymbol{x} - \boldsymbol{x}'\|^2 \tag{3.38}$$

with $a = 2\min_t \beta_t$. Also, the minimality of $\hat{\boldsymbol{x}}$ on the convex set $\mathcal{X}$ implies that for any $\boldsymbol{x} \in \mathcal{X}$, $\langle \nabla\Phi(\hat{\boldsymbol{x}}), \boldsymbol{x} - \hat{\boldsymbol{x}}\rangle \geqslant 0$. Then from (3.38) , we get for any $k \geqslant 0$:

$$\frac{a}{2}\left\|\boldsymbol{x}^{(k)} - \hat{\boldsymbol{x}}\right\|^2 \leqslant \Phi(\boldsymbol{x}^{(k)}) - \Phi(\hat{\boldsymbol{x}}) + \langle \nabla\Phi(\boldsymbol{x}^{\mathrm{NE}}), \hat{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\rangle \qquad \leqslant \Phi(\boldsymbol{x}^{(k)}) - \Phi(\hat{\boldsymbol{x}})\,,$$

and from (3.37) we get the convergence result of Theorem 3.4.

3.E Proof of Theorem 3.5: Convergence of SIRD

We analyze the convergence of the sequence $(\mathrm{T}_\gamma^k(\boldsymbol{x}))_k$ where

$$[\mathrm{T}_\gamma(\boldsymbol{x})]_n \stackrel{\text{def}}{=} \mathrm{P}_{\mathcal{X}_n}\left(\boldsymbol{x}_n - \gamma\nabla_n b_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n})\right).$$

First notice that the unique NE of the game $\boldsymbol{x}^{\mathrm{NE}}$ is the unique fixed point of T_γ i.e. $\boldsymbol{x}^{\mathrm{NE}} = \mathrm{T}_\gamma(\boldsymbol{x}^{\mathrm{NE}})$. The idea is then to prove that T_γ is a η-contraction, for a given norm $\|\cdot\|$ which will imply the convergence rate $\|T^k(\boldsymbol{x}^{(0)}) - \boldsymbol{x}^{\mathrm{NE}}\| \leqslant \eta^k\|\boldsymbol{x}^{(0)} - \boldsymbol{x}^{\mathrm{NE}}\|$, for any initial condition $\boldsymbol{x}^{(0)} \in \mathbb{R}^{N\times T}$.

Let $\|\cdot\|$ denote the Euclidean norm on $\mathbb{R}^d$ for any positive integer d. As the projection on a convex set is nonexpansive [RW09, Corollary 12.20], we get for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$:

$$\begin{aligned}
&\|\mathrm{T}_\gamma(\boldsymbol{x}) - \mathrm{T}_\gamma(\boldsymbol{y})\|^2 = \textstyle\sum_{n=1}^N \|[\mathrm{T}_\gamma(\boldsymbol{x})]_n - [\mathrm{T}_\gamma(\boldsymbol{y})]_n\|^2\\
&= \textstyle\sum_{n=1}^N \|\mathrm{P}_{\mathcal{X}_n}(\boldsymbol{x}_n - \gamma\nabla_n b_n(\boldsymbol{x})) - \mathrm{P}_{\mathcal{X}_n}(\boldsymbol{y}_n - \gamma\nabla_n b_n(\boldsymbol{y}))\|^2\\
&\leqslant \textstyle\sum_{n=1}^N \|\boldsymbol{x}_n - \boldsymbol{y}_n + \gamma(\nabla_n b_n(\boldsymbol{y}) - \nabla_n b_n(\boldsymbol{x}))\|^2\\
&= \textstyle\sum_{n=1}^N \|\boldsymbol{x}_n - \boldsymbol{y}_n\|^2 + \gamma^2\|\nabla_n b_n(\boldsymbol{x}) - \nabla_n b_n(\boldsymbol{y})\|^2 - 2\gamma\,\langle \nabla_n b_n(\boldsymbol{x}) - \nabla_n b_n(\boldsymbol{y}), \boldsymbol{x}_n - \boldsymbol{y}_n\rangle\,.
\end{aligned}$$

Since $\nabla_n b_n$ is M_n-Lipschitz for each n, we have

$$\sum_{n=1}^N |\nabla_n b_n(\boldsymbol{x}) - \nabla_n b_n(\boldsymbol{y})|^2 \leqslant NM^2\|\boldsymbol{x} - \boldsymbol{y}\|^2\,,$$

where $M \stackrel{\text{def}}{=} \max_n M_n$. From a-strong stability (Definition 3.2), we get:

$$\|\mathrm{T}_\gamma(\boldsymbol{x}) - \mathrm{T}_\gamma(\boldsymbol{y})\|^2 \leqslant \eta \|\boldsymbol{x} - \boldsymbol{y}\|^2 ,$$

with $\eta \stackrel{\text{def}}{=} 1 + NM^2\gamma^2 - 2\gamma\alpha$. Minimizing on $\gamma > 0$ gives:

$$\gamma = \frac{\alpha}{NM^2} \text{ and } \eta = 1 - \frac{\alpha^2}{NM^2} < 1,$$

which shows that T_γ is a contraction.

3.F Proof of Theorem 3.6: Consistency of DR Procedure

First, we observe that an NE is unique in any "subgame" $\mathcal{G}^{(t)}$ played on the subset of time periods $\mathcal{T}^{(t)} = \{t, \dots, \mathcal{T}\}$ (considered at t in the procedure). Indeed, in the case where we assume that the operator F (*cf* Definition 3.2) is strictly monotone on $\mathcal{X}$, then the operator $F^{(t)} : \boldsymbol{x}^{(t)} \mapsto [\nabla_{\boldsymbol{x}_n^{(t)}} b_n(\boldsymbol{x}^{(t)})]_{n\in\mathcal{N}}$ restricted to the set $\mathcal{T}^{(t)}$, is also strictly monotone on $\mathcal{X}^{(t)} = \prod_n \mathcal{X}_n^{(t)}$. In the case where we consider Assumption 3.1 and that the sets $(\mathcal{X}_n)_n$ have the structure (3.6), this structure is inherited for the sets $(\mathcal{X}_n^{(t)})_n$ so Theorem 3.2 can be applied on the game $\mathcal{G}^{(t)}$ to ensure the uniqueness.

Let $t_0 \in \{1, \dots, T-1\}$ and $\mathcal{G}^{(t_0)}$ the DR-game played at t_0. Let $\boldsymbol{x}^{(t_0)}$ be the unique NE of $\mathcal{G}^{(t_0)}$. From the variational inequality characterization of an NE, we have:

$$\langle F^{(t_0)}(\boldsymbol{x}^{(t_0)}), \boldsymbol{\lambda}^{(t_0)} - \boldsymbol{x}^{(t_0)} \rangle \geqslant 0, \ \forall \boldsymbol{\lambda}^{(t_0)} \in \mathcal{X}^{(t_0)} . \tag{3.39}$$

Let $\mathcal{G}^{(t_0+1)}$ the DR-game on hours $\{t_0 + 1, \dots T\}$ with updated strategy sets

$$\mathcal{X}_n^{(t_0+1)} \stackrel{\text{def}}{=} \left\{ (x_{n,s})_{s>t_0} | (x_{n,t_0}^{(t_0)}, [x_{n,s}]_{s>t_0}) \in \mathcal{X}_n^{(t_0)} \right\}$$

for each n. Let $\boldsymbol{\lambda}^{(t_0+1)} \in \mathcal{X}^{(t_0+1)}$, then $\boldsymbol{\lambda}^{(t_0)} \stackrel{\text{def}}{=} (\boldsymbol{x}_{t_0}^{(t_0)}, \boldsymbol{\lambda}^{(t_0+1)}) \in \mathcal{X}^{(t_0)}$.

$$\begin{aligned} 0 &\leqslant \langle F^{(t_0)}(\boldsymbol{x}^{(t_0)}), \boldsymbol{\lambda}^{(t_0)} - \boldsymbol{x}^{(t_0)} \rangle \\ &= 0 + \langle F^{(t_0+1)}((\boldsymbol{x}_s^{(t_0)})_{s>t_0}), \boldsymbol{\lambda}^{(t_0+1)} - (\boldsymbol{x}_s^{(t_0)})_{s>t_0} \rangle , \end{aligned}$$

which shows, from (3.39), that $(\boldsymbol{x}_s^{(t_0)})_{s>t_0}$ is an NE of the game $\mathcal{G}^{(t_0+1)}$. From the uniqueness of the NE in $\mathcal{G}^{(t_0+1)}$, we finally conclude that $(\boldsymbol{x}_s^{(t_0)})_{s>t_0} = \boldsymbol{x}^{(t_0+1)}$.

3.G Proof of Lemma 3.2: Contraction of BR operator with Linear Costs

Let $\boldsymbol{s}$ and $\boldsymbol{s}' \in \mathbb{R}^T$. Let us denote for simplicity by $\boldsymbol{x}_{n,t} \stackrel{\text{def}}{=} [\mathcal{B}_n(\boldsymbol{s})]_t$ the component t of the best response of n to $\boldsymbol{s}$, and $x'_{n,t} \stackrel{\text{def}}{=} [\mathcal{B}_n(\boldsymbol{s}')]_t$ similarly for $\boldsymbol{s}'$. We consider the Lagrangian function $\mathcal{L}(\boldsymbol{x}_n, \lambda_n, \overline{\boldsymbol{\mu}}_n, \underline{\boldsymbol{\mu}}_n)$ associated to the minimization problem defining $\mathcal{B}_n(\boldsymbol{s})$, given by:

$$\sum_t x_{n,t} c_t(s_t + x_{n,t}) - \lambda_n \Big(\sum_{t\in\mathcal{T}} x_{n,t} - E_n \Big) - \sum_{t\in\mathcal{T}} \Big(\overline{\mu}_{n,t}(x_{n,t} - \overline{x}_{n,t}) + \underline{\mu}_{n,t}(\underline{x}_{n,t} - x_{n,t}) \Big) , \tag{3.40}$$

where, following the notations introduced before, $\lambda_n, \overline{\mu}_{n,t}, \underline{\mu}_{n,t}$ (resp. $\lambda'_n, \overline{\mu}'_{n,t}, \underline{\mu}'_{n,t}$) are the Lagrangian multipliers associated to the energy, upper and lower bound constraints on $\boldsymbol{x}_n$ (resp $\boldsymbol{x}'_n$).

Recall that the KKT conditions of optimality of $\boldsymbol{x}_n$ ensure that for each $t \in \mathcal{T}$:

$$\gamma_{n,t}(x_{n,t}, s_t) = \lambda_n + \overline{\mu}_{n,t} - \underline{\mu}_{n,t} , \tag{3.41}$$

and symmetrically for $\boldsymbol{x}'_n$ and where $\gamma_{n,t}$ is the marginal cost associated to period t given by:

$$\gamma_{n,t}(x_{n,t}, s_t) = a_t + b_t(2x_{n,t} + s_t)\,, \quad \text{with } s_t = \sum_{m \neq n} x_{m,t}.$$

According to the equality (3.41) on $\gamma_{n,t}$ and the definitions of multipliers, we have:

$$x_{n,t} < \overline{x}_{n,t} \;\Rightarrow\; \gamma_{n,t}(x_{n,t}, s_t) \geqslant \lambda_n \;\text{ and }\; x_{n,t} > \underline{x}_{n,t} \;\Rightarrow\; \gamma_{n,t}(x_{n,t}, s_t) \leqslant \lambda_n\,. \tag{3.42}$$

Let us assume, w.l.o.g, that $\lambda_n \geqslant \lambda'_n$. Then, consider the sets:

$$\mathcal{T}_- \overset{\text{def}}{=} \{t \in \mathcal{T} \,:\, x_{n,t} < x'_{n,t}\} \text{ and } \mathcal{T}_+ \overset{\text{def}}{=} \{t \in \mathcal{T} \,:\, x_{n,t} \geqslant x'_{n,t}\} = \mathcal{T} \setminus \mathcal{T}_-\,,$$

and symmetrically consider the sets:

$$\mathcal{S}_- \overset{\text{def}}{=} \{t \in \mathcal{T} \,:\, s_t < {s_t}'\} \text{ and } \mathcal{S}_+ \overset{\text{def}}{=} \{t \in \mathcal{T} \,:\, s_t \geqslant {s_t}'\} = \mathcal{T} \setminus \mathcal{S}_-\,.$$

Suppose that there exists $t \in \mathcal{T}$ such that $s_t \leqslant {s_t}'$ and $\overline{x}_{n,t} \geqslant x'_{n,t} > x_{n,t} \geqslant \underline{x}_{n,t}$. Then according to (3.42), we have:

$$\lambda_n \leqslant \gamma_{n,t}(x_{n,t}, s_t) < \gamma_{n,t}(x'_{n,t}, {s_t}') \leqslant \lambda'_n$$

which is impossible because we assumed $\lambda_n \geqslant \lambda'_n$. Therefore $\mathcal{T}_+ = \mathcal{S}_-$ and $\mathcal{T}_- = \mathcal{S}_+$.

If $\mathcal{T}_- = \emptyset$, then $x_{n,t} \geqslant x'_{n,t}$ for all t and since $\sum_t x_{n,t} = \sum_t x'_{n,t} = E_n$, we have $x_{n,t} = x'_{n,t}$ for all t.

Now assume that $\mathcal{T}_- \neq \emptyset$ and let $t \in \mathcal{T}_-$ so that $\overline{x}_{n,t} \geqslant x'_{n,t} > x_{n,t} \geqslant \underline{x}_{n,t}$ so that

$$\begin{aligned}
&\gamma_{n,t}(x_{n,t}, s_t) \geqslant \lambda_n \geqslant \lambda'_n \geqslant \gamma_{n,t}(x'_{n,t}, {s_t}') \\
\Leftrightarrow\; &a_t + b_t(2x_{n,t} + s_t) \geqslant a_t + b_t(2x'_{n,t} + {s_t}') \\
\Leftrightarrow\; &\frac{1}{2}({s_t} - {s_t}') \geqslant x'_{n,t} - x_{n,t} \geqslant 0\,,
\end{aligned} \tag{3.43}$$

as $b_t > 0$ and similarly for $t \in \mathcal{T}_+$. Finally, we obtain:

$$\begin{aligned}
\left\|\boldsymbol{x}'_n - \boldsymbol{x}_n\right\|_2^2 &= \sum_{t \in \mathcal{T}_-} (x'_{n,t} - x_{n,t})^2 + \sum_{t \in \mathcal{T}_+} (x'_{n,t} - x_{n,t})^2 \\
&\leqslant \frac{1}{4} \sum_{t \in \mathcal{T}_-} ({s_t}' - s_t)^2 + \frac{1}{4} \sum_{t \in \mathcal{T}_+} ({s_t}' - s_t)^2 \\
&= \frac{1}{4} \times \sum_{t \in \mathcal{T}} ({s_t}' - s_t)^2\,,
\end{aligned} \tag{3.44}$$

which concludes for $\mathcal{B}_n$. For $\mathrm{Id} + \mathcal{B}_n$, one can observe that, adding $({s_t}' - s_t)$ on both sides of (3.43) we get for $t \in \mathcal{T}_-$:

$$\frac{1}{2}({s_t}' - s_t) \geqslant ({s_t}' + x'_{n,t}) - (s_t - x_{n,t}) \geqslant 0\,,$$

and the symmetric inequality for $t \in \mathcal{T}_+$. We conclude as in (3.44).

3.H Proof of Lemma 3.3: Expression of the Jacobian of BR function

Let us fix a player n. Assuming $\dfrac{\partial s_t}{\partial s_u} = 0$ if $t \neq u$, we have for any $t \neq u \in \mathcal{T}$:

$$\frac{\partial \gamma_{n,t}\big(s_t, [\mathcal{B}_n(\boldsymbol{s})]_t\big)}{\partial s_t} = (D_2\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) \times 0 + (D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) \times \frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}),$$

and:

$$\frac{\partial \gamma_{n,t}\big(s_t, [\mathcal{B}_n(\boldsymbol{s})]_t\big)}{\partial s_u} = (D_2\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) \times 1 + (D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) \times \frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}).$$

As we have $\sum_{t\in\mathcal{S}_n(\boldsymbol{s})}[\mathcal{B}_n(\boldsymbol{s})]_t$ fixed, we get $\sum_{t\in\mathcal{S}_n(\boldsymbol{s})}\frac{\partial \mathcal{B}_{n,t}}{\partial s_u}(\boldsymbol{s}) = 0$ for any $u \in \mathcal{T}$. Moreover, as for any $t \in \mathcal{S}_n(\boldsymbol{s})$, the marginal cost is $h_t(\boldsymbol{s}) = \lambda_n(\boldsymbol{s})$, we get for any $t, u \in T$:

$$\begin{aligned}
&\frac{\partial \lambda_n}{\partial s_u}(\boldsymbol{s}) = \delta_{tu} \times (D_2\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) + (D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big) \times \frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}) \\
\Leftrightarrow\ &\frac{1}{(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)} \times \frac{\partial \lambda_n}{\partial s_u}(\boldsymbol{s}) = \delta_{tu}\frac{(D_2\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)}{(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)} + \frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}) \\
\overset{\sum_{t\in\mathcal{S}_n}}{\Longrightarrow}\ &\frac{\partial \lambda_n}{\partial s_u}(\boldsymbol{s}) = \left(\sum_{t\in\mathcal{S}_n(\boldsymbol{s})}\frac{1}{(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)}\right)^{-1} \times \frac{(D_2\gamma_{n,u})\big([\mathcal{B}_n(\boldsymbol{s})]_u, s_u\big)}{(D_1\gamma_{n,u})\big([\mathcal{B}_n(\boldsymbol{s})]_u, s_u\big)}\,.
\end{aligned}$$

Thus we get:

$$\begin{aligned}
\frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}) &= \frac{1}{(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)} \times \frac{\partial \lambda_n}{\partial s_u}(\boldsymbol{s}) - \delta_{tu}\frac{(D_2\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)}{(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)} \\
&= \frac{(D_2\gamma_{n,u})\big([\mathcal{B}_n(\boldsymbol{s})]_u, s_u\big)}{(D_1\gamma_{n,u})\big([\mathcal{B}_n(\boldsymbol{s})]_u, s_u\big)}\left(\frac{\big[(D_1\gamma_{n,t})\big([\mathcal{B}_n(\boldsymbol{s})]_t, s_t\big)\big]^{-1}}{\sum_{v\in\mathcal{S}_n(\boldsymbol{s})}\big[(D_1\gamma_{n,v})\big([\mathcal{B}_n(\boldsymbol{s})]_v, s_v\big)\big]^{-1}} - \delta_{tu}\right) \\
&= \theta_u(\boldsymbol{s}) \times \big(r_t(\boldsymbol{s}) - \delta_{tu}\big)\,.
\end{aligned} \tag{3.45}$$

With $\gamma_{n,t}(s_t, x_{n,t}) = x_{n,t}c_t'(s_t) + c_t(s_t)$ we have:

$$\begin{aligned}
D_2\gamma_{n,t}(s_t, x_{n,t}) &= x_{n,t}c_t''(s_t) + c_t'(s_t) \\
D_1\gamma_{n,t}(s_t, x_{n,t}) &= x_{n,t}c_t''(s_t) + 2c_t'(s_t)\,.
\end{aligned}$$

and we obtain finally:

$$\frac{\partial [\mathcal{B}_n]_t}{\partial s_u}(\boldsymbol{s}) = \frac{[\mathcal{B}_n(\boldsymbol{s})]_u\ c_u''(s_u) + c_u'(s_u)}{[\mathcal{B}_n(\boldsymbol{s})]_u\ c_u''(s_u) + 2c_u'(s_u)}\left(\frac{\big([\mathcal{B}_n(\boldsymbol{s})]_t c_t''(s_t) + 2c_t'(s_t)\big)^{-1}}{\sum_{v\in\mathcal{S}_n(\boldsymbol{s})}\big([\mathcal{B}_n(\boldsymbol{s})]_v c_v''(s_v) + 2c_v'(s_v)\big)^{-1}} - \delta_{tu}\right).$$

3.I Proof of Proposition 3.3: Contraction of BR operator for T=2

Let $n \in \mathcal{N}$. We omit the indices n and t on $x_{n,t}$ and t on s, as we only look at player n and arc t for variables x and s. From KKT conditions, in the case where the bounding constraints (3.22) are not active, we get, with $x = \mathcal{B}_n(s) \in (\ell_{n,t}, u_{n,t})$:

$$0 = \frac{\mathrm{d}b_n}{\mathrm{d}x} = \big(c_t(x+s) + xc_t'(x+s) - c_{\tilde{t}}(E-x-s) - (E_n - x)c_{\tilde{t}}'(E-x-s)\big) \tag{3.46}$$

and

$$\frac{\mathrm{d}^2 b_n}{\mathrm{d}x^2} = 2c_t'(x+s) + xc_t''(x+s) + 2c_{\tilde{t}}'(E-x-s) + (E_n - x)c_{\tilde{t}}''(E-x-s) > 0\,,$$

thus there exists at most one value such that $\frac{\mathrm{d}b_n}{\mathrm{d}x}(\boldsymbol{x}_n, \boldsymbol{s}) = 0$, and, from KKT conditions, we obtain that if $\frac{\mathrm{d}b_n}{\mathrm{d}x}(\boldsymbol{x}_n, \boldsymbol{s})$ stays positive for $x \in (\ell_{n,t}, u_{n,t})$, we will get $\mathcal{B}_n(\boldsymbol{s}) = (\ell_{n,t}, E_n -$

$\ell_{n,t})$. If $\dfrac{\mathrm{d}b_n}{\mathrm{d}x}(\boldsymbol{x}_n, \boldsymbol{s})$ stays negative for $x \in (\ell_{n,t}, u_{n,t})$, we will get $\mathcal{B}_n(\boldsymbol{s}) = (u_{n,t}, E_n - u_{n,t})$. Combined with the fact that $\dfrac{\mathrm{d}b_n}{\mathrm{d}x}(\boldsymbol{x}_n, \boldsymbol{s})$ is increasing in $\boldsymbol{s}$, we obtain Proposition 3.3 i), ii).

Then, taking the derivative of (3.46) with respect to s, we obtain:

$$\begin{aligned}\frac{\partial x}{\partial s} &= -\frac{c_t'(x+s) + c_{\tilde{t}}'(E-x-s) + x\left(c_t''(x+s) - c_{\tilde{t}}''(E-x-s)\right) + E_n c_{\tilde{t}}''(E-x-s)}{2\left(c_t'(x+s) + c_{\tilde{t}}'(E-x-s)\right) + x\left(c_t''(x+s) - c_{\tilde{t}}''(E-x-s)\right) + E_n c_{\tilde{t}}''(E-x-s)} \\ &= -\frac{c_t'(X_t) + c_{\tilde{t}}'(X_{\tilde{t}}) + x_{n,t} c_t''(X_t) + x_n^{\tilde{t}} c_{\tilde{t}}''(X_{\tilde{t}})}{2\left(c_t'(X_t) + c_{\tilde{t}}'(X_{\tilde{t}})\right) + x_{n,t} c_t''(X_t) + x_n^{\tilde{t}} c_{\tilde{t}}''(X_{\tilde{t}})} \stackrel{\text{def}}{=} -\alpha_n(\boldsymbol{x})\end{aligned}$$

We can see that $\alpha_n(\boldsymbol{x})$ is always in $[0,1]$ as $c' > 0, c'' \geqslant 0$ and that:

$$\alpha_n(\boldsymbol{x}) \leqslant 1 \Leftrightarrow\ 0 \leqslant c_t'(X_t) + c_{\tilde{t}}'(X_{\tilde{t}})\ .$$

We have also that the derivative of $\mathrm{Id} + \mathcal{B}_n$ is given by:

$$\frac{\partial[\mathrm{Id} + \mathcal{B}_n]_t}{\partial s_t}(\boldsymbol{s}) = 1 + \frac{\partial x}{\partial s} = \frac{c_t'(X_t) + c_{\tilde{t}}'(X_{\tilde{t}})}{2\left(c_t'(X_t) + c_{\tilde{t}}'(X_{\tilde{t}})\right) + x_{n,t} c_t''(X_t) + x_n^{\tilde{t}} c_{\tilde{t}}''(X_{\tilde{t}})}$$

and thus:

$$\left(\frac{\partial[\mathcal{B}_n]_t}{\partial s_t}(\boldsymbol{s})\right)^2 + \left(\frac{\partial[Id + \mathcal{B}_n]_t}{\partial s_t}(\boldsymbol{s})\right)^2 = 1 - 2\alpha_n + 2\alpha_n^2 = 1 - 2\alpha_n(1-\alpha_n) < 1\ .$$

Chapter 4

Impact of Consumers Temporal Preferences in Demand Response

This Chapter is based on the conference paper [Jac+17a]. Here, as an extension of Chapter 2 and Chapter 3, we consider a noncooperative setting where consumers have some preferences over their energy consumption profiles, in addition to their energy bills. We analyze the influence of preferences over the equilibria, in the two different billing mechanisms considered in Chapter 2.

4.1 Introduction

Individual and personal constraints will influence the behaviour of flexible consumers. In turn, many papers have addressed the question of modeling individual effort of consumers in the context of DSM. Several works, considering thermostatic control loads (AC, heating), relate this individual effort to the desired indoor temperature. For instance, [NL14] considers a distance between "desired indoor temperature" profile and effective one, weighted by an occupancy variable. Another standard model consists in penalizing the delay between possible operation time (e.g., the starting period of availability) and effective one of a flexible electrical appliance [YM13; CKS11; MRLG10]. In this case, the cost is generally linear with the waiting time (and sometimes weighted by the power of considered task as in [MRLG10]). A different approach in [Ban+14] considers as a metric for uncomfort of residential consumers the colour quality of a "smart lighting". Note that in all of the aforementioned metrics, the total flexible energy consumed is fixed, and consumption flexibility consists only in temporal scheduling of this fixed amount. Individual effort made by consumers consists then of a temporal preference for consumption. As a result, these metrics can be formulated as particular cases of the framework that we will propose in this chapter.

Other works such as [LCL11; FA01; Sam+12] rather consider an individual utility term that depends on the total flexible energy consumed; a standard representation of this utility is made with an increasing and concave function of total amount of flexible energy (a quadratic function with a saturation threshold is often used, as mentioned in [Den+15]). In this model, consumers can receive no energy at all and it is assumed that their satisfaction increases with the volume of energy they consume.

Whatever the metrics considered for the individual consumer preference, to the best of our knowledge no study has been made on the impact of the weight given to this preference on the consumers behaviour. More precisely, this weight will influence the induced equilibrium in the associated consumption energy game, which will impact the system efficiency. This is precisely the issue addressed in this chapter.

In this work, we will distinguish standard metrics of efficiency in a game: the system cost and the social cost. We will study as well the Price of Anarchy (PoA) [KP99], a standard measure of efficiency in a game, and a measure called *Price of Efficiency* defined to be similar to the Price of Anarchy on the system operator side. While the study of such indicators in energy consumption games has been done previously (see [MR+10] or [Bea+16] which exhibit games where $\text{PoA} = 1$), the analysis of the evolution of these indicators with respect to the weight on individual effort term is a novelty addressed here.

This chapter brings several contributions. We extend the standard model of an energy consumption game among consumers, studied in [Jac+17b], by adding individual temporal preferences. Next, we give theoretical results in this extended framework about the impact of preferences on the equilibria of the game and the efficiency of those equilibria. For that, we analyse the induced social cost and system cost. Last, we present numerical results on a realistic test case, using consumption data from *PecanStreet* database [Pec]. In particular, we show that the equilibrium induced by the hourly billing mechanism [Bah+13] is robust to the level given to preferences.

This chapter is organized as follows: Section 4.2 introduces the notion of consumers temporal preferences and defines the energy consumption game model. In Section 4.3, we define the main metrics of our study: we recall the definition of the Price of Anarchy and define the Price of Efficiency. In Section 4.4, we give theoretical results and properties on the formulated model. We present explicit results on the equilibria in a simplified framework. Last, Section 4.5 is devoted to numerical experiments on a realistic framework, where we simulate the equilibria among thirty Texan residential consumers in January, 2016.

4.2 Context and Energy Consumption Game

The model of this work falls within the class of DSM studies where the interaction of individual consumers is coordinated introducing an energy consumption game, as in [MR+10]. While all the proposed results could be applied to numerous operational frameworks, the one described here consists in the interaction between a provider and its set $\mathcal{N} = \{1, \cdots, N\}$ of consumers in a given day. As opposed to [MR+10] which does not show any preference for consumers and suppose that they are indifferent to any consumption schedule as soon as it satisfies their constraints, here we will focus on the integration of individual preferences of consumers into their objectives.

Indeed, consumers tend to have a "natural" or preferred consumption profile, and asking them to deviate from it might be inconvenient or decrease their comfort. Individual utility functions have been previously used through different models. A common approach (see for instance [FA01] and [Sam+12]) is to consider that a consumer's utility can be modeled as an increasing function of the total energy he receives. Here, on the contrary, we keep the assumption made in [MR+10] that consumers have flexible appliances that need a fixed quantity of energy per day, and this demand must be satisfied each day. However, we assume that consumers are not indifferent to the time they can use electricity and therefore use their appliances.

4.2.1 Introducing users temporal preferences

From the provider's point of view, only the load profile asked by a user n, $(x_{n,t})_{t\in\mathcal{T}}$ matters, where $\mathcal{T}$ is the discrete set of time periods considered. However, that may not be the case for users: for instance, one would like to charge an Electric Vehicle (EV) battery as soon as possible in case of unscheduled need (Plug-and-Charge), or one would like to turn on the heating system in a household at precise time periods ([NL14]), etc. We denote user n's *preferred* or *desirable* consumption profile by the vector:

$$\boldsymbol{x}^\sharp_{na} = (x^\sharp_{na,t})_t \in \mathbb{R}^T \,,$$

for his flexible appliance a. As a result, user n would like to receive the power profile $(x^\sharp_{n,t})_t \stackrel{\text{def}}{=} (\sum_a x^\sharp_{na,t})_t$ and, if he has no incentives to do otherwise, this profile will be his actual one. Deviating from the profile $\boldsymbol{x}^\sharp_n$ decreases the comfort or utility of consumer n. To model this fact, we introduce the individual utility $u_n(\boldsymbol{x}_n)$ of consumer n as the opposite of the squared distance[1] between the actual consumption profile of consumer n, $\boldsymbol{x}_n$, and his

[1] In general, one could use $u_n(\boldsymbol{x}_n) \stackrel{\text{def}}{=} -\omega_n d(\boldsymbol{x}_n, \boldsymbol{x}^\sharp_n)$ where $d(.,.)$ is an arbitrary metric. For simplicity and computational purposes, we use $d = \|.\|^2$.

preferred profile $\boldsymbol{x}_n^\sharp$:

$$u_n(\boldsymbol{x}_n) \stackrel{\text{def}}{=} -\omega_n \textstyle\sum_t (x_{n,t} - x^\sharp{}_{n,t})^2 \,, \tag{4.1}$$

where the weight ω_n indicates how much user n values the distance to his preference. As some users will give more importance to their electricity bills (defined below) and some others to their utility, in a general framework we assume that each user could set a different weight ω_n in this model.

4.2.2 Users billing mechanism

As done in [MR+10; Jac+17b], we suppose that the providing costs on each period t are represented as a quadratic function $\tilde{C}(X_t^{\text{tot}})$ of the total load $X_t^{\text{tot}} = X_t^{\text{NF}} + X_t$, where X_t^{NF} denotes the aggregated nonflexible load at period t and X_t the flexible part:

$$\tilde{C}(X_t^{\text{tot}}) \stackrel{\text{def}}{=} \tilde{\delta} + \tilde{\alpha} X_t^{\text{tot}} + \tilde{\beta} (X_t^{\text{tot}})^2. \tag{4.2}$$

The surplus cost induced by the flexible part of the load X_t, at time t, denoted by $C_t(X_t)$, can be deduced as:

$$C_t(X_t) \stackrel{\text{def}}{=} \tilde{C}(X_t^{\text{tot}}) - \tilde{C}(X_t^{\text{NF}}) = \alpha_t X_t + \beta (X_t)^2 \tag{4.3}$$

with $\alpha_t \stackrel{\text{def}}{=} \left(\tilde{\alpha} + 2\tilde{\beta} X_t^{\text{NF}}\right)$ and $\beta = \tilde{\beta}$. Even if the cost function for the provider (4.2) does not depend on time, the nonflexible load profile X_t^{NF} induces a difference in the costs $(C_t)_{t\in\mathcal{T}}$ between the different time periods $t \in \mathcal{T}$.

In this chapter, we consider that the nonflexible part of the load is managed and billed in a distinct process, e.g., in a standard contract. We focus on DR billing mechanisms for the *flexible* part of the load. Through this study, we will consider two different billing mechanisms which, in practice, would require a "two-way" communication system [IA09], which enables the system operator to send its price functions $(C_t)_t$ and aggregated load $(X_t)_t$ and users to send back their consumption profile $\boldsymbol{x}_n$. First, we consider the *Daily Proportional* (DP) billing mechanism introduced in [MR+10]: we assume that the system cost $C_t(X_t)$ induced by the flexible load at time t:

$$X_t \stackrel{\text{def}}{=} \textstyle\sum_n x_{n,t} \tag{4.4}$$

are shared among users proportionaly to their total flexible consumption on the entire day $E_n = \sum_{t\in\mathcal{T}} x_{n,t}$. Formally, each consumer will pay the daily bill:

$$b_n^{\text{DP}}(\boldsymbol{x_n}, \boldsymbol{x_{-n}}) = \frac{E_n}{E} \sum_{t\in\mathcal{T}} C_t(X_t) \,, \tag{4.5}$$

where $\boldsymbol{x_{-n}} = (\boldsymbol{x}_m)_{m\neq n}$ and $E \stackrel{\text{def}}{=} \sum_n E_n$. We will compare its efficiency to the natural "congestion" *Hourly Proportional* (HP) billing mechanism introduced in [Bah+13], where system costs on each period are shared among consumers respectively to their consumption on this period. Formally, the daily bill b_n^{HP} of user n for his flexible consumption is[2]:

$$b_n^{\text{HP}}(\boldsymbol{x_n}, \boldsymbol{x_{-n}}) = \sum_{t\in\mathcal{T}} \frac{x_{n,t}}{X_t} C_t(X_t) \,. \tag{4.6}$$

Intuitively, with b_n^{HP}, users are more impacted by their actions to use expensive/cheap time periods than with b_n^{DP} where the costs induced by actions are "averaged" over the day. This property helps to interpret the results of Section 4.4 and Section 4.5.

[2]Introducing per-unit prices $c_t \stackrel{\text{def}}{=} C_t(X_t)/X_t$, the bill of n can also be formulated in the "congestion" form: $b_n^{\text{HP}}(\boldsymbol{x}) = x_{n,t} c_t(x_t)$ analyzed in [ORS93].

4.2.3 Energy consumption game

To analyze the impact of the importance given to users' temporal preferences, we consider through this work the parametrized users' objective functions:

$$f_n^\theta(\boldsymbol{x_n}, \boldsymbol{x_{-n}}) \stackrel{\text{def}}{=} (1-\theta) b_n(\boldsymbol{x}) - \theta u_n(\boldsymbol{x}) \tag{4.7}$$

where the *preference factor* $\theta \in [0,1]$ indicates the weight given to user n's preference[3] in comparison to his bill b_n. We get the optimization problem for user n, as already defined in Chapter 2, but now indexed by θ:

$$\min_{\boldsymbol{x}_n \in \mathbb{R}^T} f_n^\theta(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \tag{4.8a}$$

$$\text{s.t.} \quad \textstyle\sum_{t\in\mathcal{T}} x_{n,t} = E_n, \tag{4.8b}$$

$$\underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t}, \forall t \in \mathcal{T}\ . \tag{4.8c}$$

Constraint (4.8b) expresses that a fixed daily amount of energy is required for the flexible appliances of user n (EV battery, washing machine...). Due to physical limits of his electrical items or personal constraints, the power given to n is bounded (4.8c). We denote more compactly by $\mathcal{X}_n$ the feasible set of user n, given by the polytope (4.8b), (4.8c), and $\mathcal{X} \stackrel{\text{def}}{=} \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$. When $\theta = 0$, user's preference $\boldsymbol{x}_n^\sharp$ has no influence on his behaviour, while when $\theta = 1$, the user gives no importance to b_n and only wants to minimize $-u_n$: his resulting load profile will be exactly his preference $\boldsymbol{x}_n^\sharp$.

As f_n^θ depends on the consumption of n but also on other users, this induces a game between users [FT91] denoted by $\mathcal{G}_\theta \stackrel{\text{def}}{=} \big(\mathcal{N}, \mathcal{X}, (f_n^\theta)_n\big)$. We will use the notations $\mathcal{G}_\theta^{\text{DP}}$ and $\mathcal{G}_\theta^{\text{HP}}$ when we specify the billing mechanism according to the DP rule (4.5) or HP rule (4.6). The importance that each user gives to his utility function u_n in comparison to his bill b_n, through the parameter θ, will change the set of Nash Equilibria (NE) of the game $\mathcal{G}_\theta$ given by:

$$\mathcal{X}_\theta^{\text{NE}} \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{X} : \forall n, \forall \boldsymbol{x}_n' \in \mathcal{X}_n,\ f_n^\theta(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \leqslant f_n^\theta(\boldsymbol{x}_n', \boldsymbol{x}_{-n})\}.$$

4.3 Social cost versus System cost

An Independent System Operator is interested in both an efficient electricity network and the welfare of the consumers. Starting with the latter, we define the *social cost* of the set of consumers as the sum of their objective functions:

$$\text{SC}_\theta(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} f_n^\theta(\boldsymbol{x}) = (1-\theta) \sum_{n\in\mathcal{N}} b_n(\boldsymbol{x}) - \theta \sum_{n\in\mathcal{N}} u_n(\boldsymbol{x}_n)\ . \tag{4.9}$$

To quantify the efficiency of a billing mechanism in a game, we consider the standard notion of Price of Anarchy (PoA) introduced by Koutsoupias and Papadimitriou [KP99] and which we recall in Definition 2.2 in Chapter 2.

The PoA measures the gap between the minimal social cost (4.9), and the social cost induced by the worst equilibrium of the game.

From the provider's point of view, only the costs induced for the system, without users personal utilities, matters: we denote by $\mathcal{C}$ the total system cost function, defined as the providing costs induced by an aggregate consumption profile $(X_t)_{t\in\mathcal{T}}$:

$$\mathcal{C}(\boldsymbol{x}) \stackrel{\text{def}}{=} \textstyle\sum_{t\in\mathcal{T}} C_t(X_t)\ . \tag{4.10}$$

Note that in the particular billing mechanisms considered in (4.5) and (4.6), we assume that the provider costs are shared among users[4], so that we have the equality $\mathcal{C}(\boldsymbol{x}) = \sum_n b_n(\boldsymbol{x})$.

[3] We could extend this study by using different $(\theta_n)_n$ for different users.

[4] This assumption could be relaxed, as done in [MR+10], by adding a ratio profit $\kappa > 1$ for the provider, so that we have the equality $\sum_n b_n(\boldsymbol{x}) = \kappa \mathcal{C}(\boldsymbol{x})$.

We introduce a measure similar to the PoA (2.6), but that will be more relevant for a provider that is more interested in the system cost $\mathcal{C}$ and does not have access to the utility functions $(u_n)_n$ of its users. In the original paper [KP99], the authors introduce the notion of PoA for an arbitrary *social cost* function. However, for clarity in this thesis, we will use the term PoA to refer only to the quantity (2.6) (where SC to refers to the sum of players cost functions, as in (4.9)) and we will use the term PoE to refer to the symmetric quantity considering the global cost $\mathcal{C}$ instead of SC.

Definition 4.1. Price of Efficiency (PoE).

Given a game $\mathcal{G}$, $\mathcal{X}_{\mathcal{G}}^{\mathrm{NE}}$ its set of Nash Equilibria and its minimal feasible system cost $\mathcal{C}^* \stackrel{\text{def}}{=} \min_{\boldsymbol{x}\in\mathcal{X}} \mathcal{C}(\boldsymbol{x})$, the price of efficiency of $\mathcal{G}$ is given as:

$$\mathrm{PoE}(\mathcal{G}) \stackrel{\text{def}}{=} \left(\sup_{\boldsymbol{x}\in\mathcal{X}_{\mathcal{G}}^{\mathrm{NE}}} \mathcal{C}\left(\boldsymbol{x}\right)\right) / \mathcal{C}^* \,. \tag{4.11}$$

Observe that PoA $\geqslant 1$ and PoE $\geqslant 1$. Following (4.7), one can notice that for $\theta = 0$, $\mathrm{PoE}(\mathcal{G}_\theta) = \mathrm{PoA}(\mathcal{G}_\theta)$. In general, the PoA and PoE will be different as shown below.

4.4 Properties

4.4.1 Potential property, existence and uniqueness of NE

We start by showing that the considered games have the property of potential (see [MS96]).

Theorem 4.1. *$\mathcal{G}_\theta^{DP}$ is a weighted potential game with potential:*

$$W_\theta^{DP} = (1-\theta)\sum_{t\in\mathcal{T}} C_t(X_t) - \theta\sum_n \frac{E}{E_n} u_n(\boldsymbol{x}_n) \,. \tag{4.12}$$

Proof: $\forall n$, $\nabla_n f_n^\theta = \frac{E_n}{E}\nabla_n W_\theta^{DP}$, *we conclude from [MS96].*

With the billing mechanism HP, we need an additional assumption on the system cost functions to get a similar result:

Theorem 4.2. *If we consider quadratic costs* (4.2), *$\mathcal{G}_\theta^{HP}$ is an exact potential game with potential:*

$$W_\theta^{HP} = (1-\theta)\left[\sum_{t\in\mathcal{T}} \frac{\beta_t}{2}\Big((X_t)^2 + \textstyle\sum_n (x_{n,t})^2\Big) + \alpha_t X_t\right] - \theta\sum_{n\in\mathcal{N}} u_n(\boldsymbol{x}) \,. \tag{4.13}$$

Proof: Similarly to the proof of Theorem 4.1, $\forall n$, $\nabla_n f_n^\theta = \nabla_n W_\theta^{HP}$.

From the fact that W_θ^{DP} and W_θ^{HP} are strictly convex and from [MS96], we can deduce the existence and uniqueness of NE:

Corollary 4.1. *In the game $\mathcal{G}_\theta^{DP}$ (resp. $\mathcal{G}_\theta^{HP}$), there exists a unique Nash Equilibrium, corresponding to the minimum argument of W_θ^{DP} (resp. W_θ^{HP}) over the set $\mathcal{X}$.*

Corollary 4.1 extends the results of [MR+10] which gives the uniqueness of NE in the particular case of θ=0 with the DP billing.

A natural algorithm to compute an NE is to run the Best Response Dynamics (BRD), as defined below.

Definition 4.2 ([GM91])**.** Best Response Dynamics (BRD).
At each iteration k, a user n_k is randomly chosen and solves problem (4.8) to optimum $\boldsymbol{x}_{n_k}^*$, with load of others $x_{-n_k}^{(k)}$ fixed (n_k best responses to the others). We update $\boldsymbol{x}_{n_k}^{(k+1)} = \boldsymbol{x}_{n_k}^*$.

From Theorem 4.1 and Theorem 4.2, we deduce the convergence of BRD:

Corollary 4.2. *In $\mathcal{G}_\theta^{DP}$ and $\mathcal{G}_\theta^{HP}$, BRD is equivalent to a block coordinate minimization of the potential function. Hence, it converges to the unique NE of the game (see [BT13]).*

4.4.2 Theoretical results on a simplified framework

In this section, we consider that the set $\mathcal{T}$ is reduced to two time periods $\mathcal{T} \stackrel{\text{def}}{=} \{P, O\}$ which represent for instance the *Peak* and *Offpeak* times. For computational purposes, we consider that the system cost is reduced to a quadratic term:

$$\forall t \in \mathcal{T},\ C_t(X_t) = (X_t)^2\,, \tag{4.14}$$

and there is no nonflexible part as in the general case described in Section 4.2.2. Each consumer n has a preference weight $\omega_n = 1$, a preferred profile $(x^\sharp_{n,P}, x^\sharp_{n,O})$, satisfying $x^\sharp_{n,P} + x^\sharp_{n,O} = E_n$ as in (4.8b). Without loss of generality, we assume $X^\sharp_P \geqslant \frac{E}{2} \geqslant X^\sharp_O$. Power constraints (4.8c) are replaced by positivity $x_{n,t} \geqslant 0$.

Nash equilibrium

From the KKT conditions of optimality, we get the following result:

Theorem 4.3. *Assume that for all $n \in \mathcal{N}$, we have:*

$$\frac{x^\sharp_{n,P}}{E_n} + \frac{1}{2} \geqslant \frac{X^\sharp_P}{E}\,, \tag{4.15}$$

then, for $\theta \in (0,1]$, the unique NE of $\mathcal{G}^{DP}_\theta$ is given by:

$$x_{n,P} = x^\sharp_{n,P} + \tfrac{E_n}{E}\tfrac{1-\theta}{2}(X^\sharp_O - X^\sharp_P)\,, \tag{4.16}$$

with symmetric expression for $x_{n,O}$. For $\theta = 0$, the KKT system is degenerated, and any $(\boldsymbol{x}_{n,t})_n$ satisfying $X_P = E/2$ is an NE.

Assume that for all $n \in \mathcal{N}$, we have:

$$2(N-1)x^\sharp_{n,P} \geqslant (X^\sharp_P - X^\sharp_O) - E_n\,, \tag{4.17}$$

then, for $\theta \in [0,1]$, the unique NE of $\mathcal{G}^{HP}_\theta$ is given by:

$$x_{n,P} = x^\sharp_{n,P} + \tfrac{1-\theta}{2(1+\theta)}\left(\phi(\theta)(X^\sharp_O - X^\sharp_P) + (x^\sharp_{n,O} - x^\sharp_{n,P})\right), \tag{4.18}$$

with symmetric expression holding for $x_{n,O}$, and with:

$$\phi(\theta) \stackrel{\text{def}}{=} \tfrac{2\theta}{(1+\theta)+(1-\theta)N} \in [0,1]\,. \tag{4.19}$$

One can check that the positivity of the offpeak load $x_{n,O}$ in (4.16) and in (4.18) is always verified. The positivity of the peak load $x_{n,P}$ is a consequence of assumptions (4.15) and (4.17).

We consider that (4.15) and (4.17) hold through all this Section.

Corollary 4.3. *The aggregated load at the NE is given by:*

$$\text{for } \mathcal{G}^{DP}_\theta,\ X_P = \tfrac{E}{2} + \theta\tfrac{(X^\sharp_P - X^\sharp_O)}{2}\,, \tag{4.20}$$

$$\text{for } \mathcal{G}^{HP}_\theta,\ X_P = \tfrac{E}{2} + \phi(\theta)\tfrac{(X^\sharp_P - X^\sharp_O)}{2}\,. \tag{4.21}$$

With HP and DP, the aggregated load evolves to the preferred profile when θ goes to one but, with the HP mechanism, this evolution is influenced by the number of players N through $\phi(\theta)$.

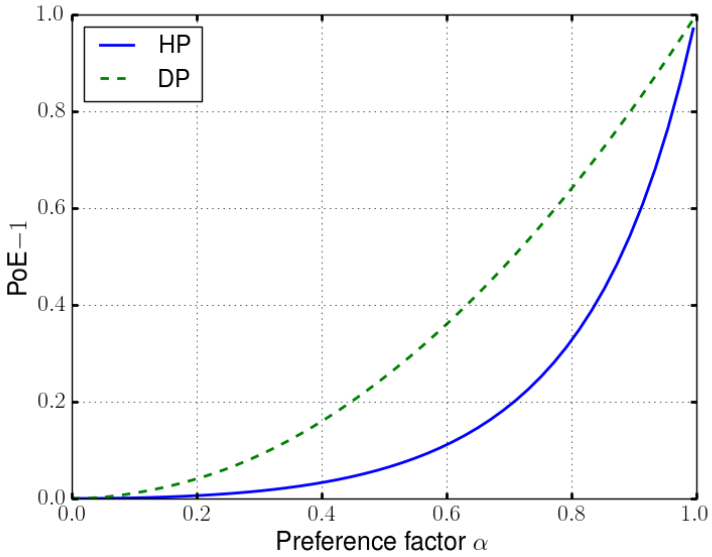


FIGURE 4.1: Evolution of PoE-1 with costs (4.14) and $N = 5$ users.
For all θ, HP billing is more efficient for the system than DP.

System Costs

The total system costs $\mathcal{C}(\boldsymbol{X}) = (X_P)^2 + (X_O)^2$ at the equilibrium are given from (4.20) and (4.21) by:

$$\text{for } \mathcal{G}_\theta^{\text{DP}}, \mathcal{C}^{\text{DP}}(\theta) \stackrel{\text{def}}{=} \tfrac{1}{2}\left(E^2 + \theta^2(X_P^\sharp - X_O^\sharp)^2\right) , \tag{4.22}$$

$$\text{for } \mathcal{G}_\theta^{\text{HP}}, \mathcal{C}^{\text{HP}}(\theta) \stackrel{\text{def}}{=} \tfrac{1}{2}\left(E^2 + \phi^2(\theta)(X_P^\sharp - X_O^\sharp)^2\right) . \tag{4.23}$$

On Figure 4.1, we see that the PoE is increasing with θ in both cases (the proof is straightforward from (4.22) and (4.23)), and that it is always smaller with the HP billing, as shown below:

Theorem 4.4. *The system costs induced by the equilibrium with HP are always smaller than with DP, or equivalently:*

$$\forall \theta \in [0,1],\ PoE(\mathcal{G}_\theta^{HP}) \leqslant PoE(\mathcal{G}_\theta^{DP}) \tag{4.24}$$

and the inequality is strict for $\theta \in (0,1)$.

Proof: First, note that in (4.11), the minimal system cost $\mathcal{C}^*$ does not depend on θ, so that $\mathcal{C}$ and PoE are proportional. From the expressions (4.22) and (4.23), we get that:

$$\mathcal{C}^{\text{DP}}(\theta) - \mathcal{C}^{\text{HP}}(\theta) = \frac{\theta^2(X_P^\sharp - X_O^\sharp)^2}{2}\left(1 - \tfrac{4}{(N(1-\theta)+(1+\theta))^2}\right) > 0$$
$$\text{because } \forall \theta \in (0,1), \frac{4}{(N(1-\theta)+(1+\theta))^2} < 1 \iff N > 1 .$$

Figure 4.1 shows the evolution of the PoE induced by the NE of $\mathcal{G}_\theta$, in the case of $N = 5$ players that have a flexible energy $E_n = 1$ that they prefer loading totally on peak hour ($x_{n,P}^\sharp = 1$).

Social Cost

If users do not care about their bills but only on their utility ($\theta = 1$), they choose their preferred profile $(x_{n,P}^\sharp, x_{n,O}^\sharp)$. As a result, the social cost will be exactly zero. On the opposite,

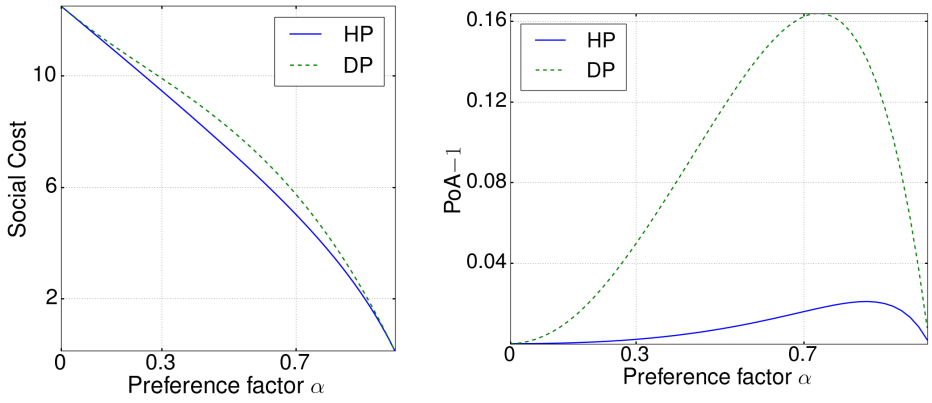


FIGURE 4.2: Evolution of Social Cost and PoA-1 with costs (4.14). *Social Cost always decreases with θ. The PoA is unimodal and with DP, it reaches its maximum at a critical level of $\theta \simeq 0.72$, where it is more than 10% larger than with HP.*

if consumers only care about their bills ($\theta = 0$), [MR+10] shows that users will reach the optimal system cost in $\mathcal{G}_0^{\mathrm{DP}}$ (the potential W_0^{DP} is equal to the system costs) while [Jac+17b] shows that the equilibrium in $\mathcal{G}_0^{\mathrm{HP}}$ will stay close to the social optimum (it is even optimal in the framework of this section, as seen in (4.21)). However, it is not clear how the social cost evolves with $\theta \in [0,1]$. Figure 4.2 shows numerically that, for both the DP and HP mechanism, the social cost is a decreasing function of θ. We prove this in Theorem 4.5 for the DP mechanism.

To this end, considering the expressions of the equilibrium in $\mathcal{G}_\theta^{\mathrm{DP}}$ from (4.16), we get the induced social cost:

$$\mathrm{SC}_\theta^{\mathrm{DP}} = (1-\theta)\left[\tfrac{E^2}{2} + \tfrac{D^2}{2}(\theta^2 + V_E(1-\theta)\theta)\right] \tag{4.25}$$

with $E \overset{\text{def}}{=} \sum_{n\in\mathcal{N}} E_n$, $D \overset{\text{def}}{=} (X_P^\sharp - X_O^\sharp)$ and $V_E \overset{\text{def}}{=} \sum_n \frac{E_n^2}{E^2}$.

Theorem 4.5. *SC_θ^{DP} is a decreasing function of θ.*

Proof. $\partial_\theta \mathrm{SC}_\theta^{\mathrm{DP}} = \frac{D^2}{2}\left[-3(1-V_E)\theta^2 + 2(1-2V_E)\theta\right] + D^2V_E - E^2$ is always negative. Details are omitted here for brevity.

We did not manage to prove a symmetric result for $\mathrm{SC}_\theta^{\mathrm{HP}}$.

4.5 Numerical experiments

In this Section, we present numerical results on the sensitivity of the equilibria of $\mathcal{G}_\theta^{\mathrm{DP}}$ and $\mathcal{G}_\theta^{\mathrm{HP}}$ to θ, in a realistic framework. We simulate the games $\mathcal{G}_\theta^{\mathrm{DP}}$ and $\mathcal{G}_\theta^{\mathrm{HP}}$ and the convergence to the equilibria day by day on the set of the thirty one days of January 2016, which we denote by $\mathcal{D}$, each day being decomposed by a hourly timeset $\mathcal{T} = \{0, 1, \ldots, 23\}$.

4.5.1 Parameters

Consumers We extracted $N = 30$ residential consumption profiles of Electric Vehicles (EV) owners from *PecanStreet Inc.* [Pec], a database of residential consumers in Texas (U.S.). Each consumer has a nonflexible consumption $(x_{n,t}^{\mathrm{NF}})_t$ (lights, cooking, TV...) and we consider EV charging as the flexible usage. We take the EV historical profile of user n as its preferred profile $(x_{n,t}^\sharp)_t$, and assume it corresponds to its flexible energy need $E_n \overset{\text{def}}{=} \sum_t x_{n,t}^\sharp$. For power

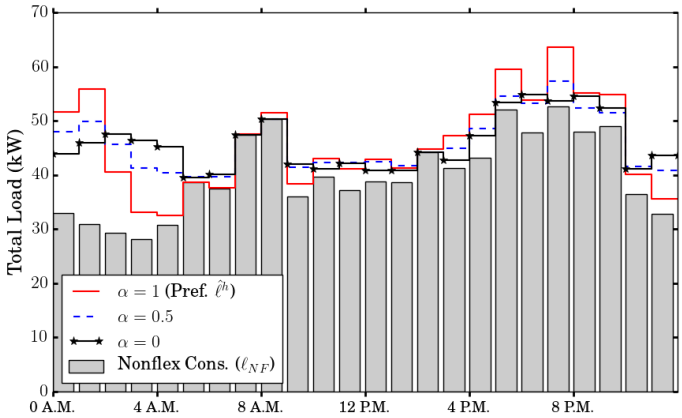


FIGURE 4.3: Aggregated consumption profiles at equilibrium $\mathcal{G}_\theta^{\text{HP}}$ for $\theta = 0,\ 0.5$ and 1 on day 10/01/2016 with 30 users.
The equilibrium profile converges to $(X_t^\sharp)_t$ when $\theta \to 1$.

constraints (4.8c), we take $\underline{x}_{n,t} = 0$ and $\overline{x}_{n,t}$ equal to the max observed value if hour t was ever used by n and $\overline{x}_{n,t} = 0$ otherwise.

System Costs As explained in Section 4.2.2, we suppose that system costs are, for each time t, function of the total load $X_t^{\text{tot}} = X_t^{\text{NF}} + X_t$, and are given in dollar cents as:

$$\tilde{C}(X_t^{\text{tot}}) = 71.1 - 4.17 X_t^{\text{tot}} + 0.295 (X_t^{\text{tot}})^2 .$$

To compute those coefficients, we make an interpolation based on three load values and three corresponding prices $(C_t(X_t)/X_t)$. The three load values are the mean (33.8kW), min (17.8kW) and max (58.9kW) values of the nonflexible load per hour aggregated over the set $\mathcal{N}$ of consumers in all hours of January, 2016. The three corresponding prices are those proposed by the Texan provider Coserv ([Cos]): 8.0 ***c*** /kWh for base contracts, 14.0 ***c*** /kWh (Peak) and 5.5 ***c*** /kWh (OffPeak) in Time-of-Use contracts. From (4.3), the cost of flexible load is given by:

$$C_t(X_t) \stackrel{\text{def}}{=} (-4.17 + 0.590 X_t^{\text{NF}}) X_t + 0.295 (X_t)^2 . \tag{4.26}$$

To ensure that b_n and u_n are of the same order of magnitude, we use a common factor in (4.1) of $\omega_n = \omega \stackrel{\text{def}}{=} \dfrac{\mathcal{C}^*}{\sum_n \left\| \boldsymbol{x}_n^* - \boldsymbol{x}_n^\sharp \right\|^2} = 49.1\ \boldsymbol{c}/\text{kWh}^2$ with $\mathcal{C}^* = \mathcal{C}(x^*) = \min_x \mathcal{C}(x)$ the optimal costs.

Note that for $\theta = 1$, $\text{SC}_\theta^* = 0 = \text{SC}_\theta^{\text{DP}} = \text{SC}_\theta^{\text{HP}}$ so the PoA is not defined, but Figure 4.4 shows that $\lim_{\theta \to 1} \text{PoA}(\mathcal{G}_\theta) = 1$.

4.5.2 Results

For both mechanisms HP and DP, we compute the NE by playing a BRD (Definition 4.2) with a limiting number of 150 BR iterations, which in practice was sufficient for convergence.

The optimization problem (4.8) is a quadratic program that we solve with the optimization solver Cplex 12.6. Playing the BRD takes around 2.5sec. for each of the 50 values of θ and each day in $\mathcal{D}$. The total simulation time was 3160 sec. with an Intel Xeon CPU E3-1240v3@3.4GHz×8 run on 5 threads.

Figure 4.3 shows the different aggregated profile $(X_t)_t = (\sum_n x_{n,t})_t$ at the equilibrium of $\mathcal{G}_\theta^{\text{HP}}$ on January, 10, chosen arbitrarily in $\mathcal{D}$. We can see a significant variation (more than 15%) on the aggregated load when θ changes.

From Figure 4.4, we see that at $\theta = 0$, there is a very small PoA for the HP mechanism (see [Jac+17b] for a deeper analysis) while DP achieves optimality. However, when θ grows the HP mechanism becomes much more efficient than DP in terms of PoA (Figure 4.4) and

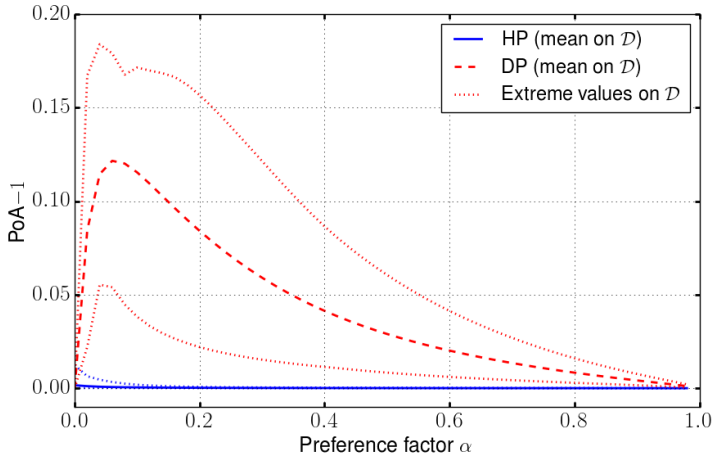


FIGURE 4.4: Evolution of PoA-1 (mean on days) with θ.
HP is more robust and has a smaller PoA than DP.

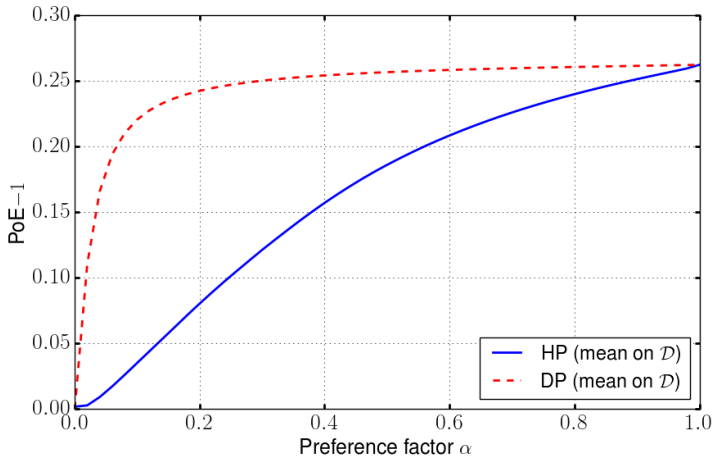


FIGURE 4.5: Evolution of PoE-1 (mean on days) with θ.
For $\theta > 3 \cdot 10^{-4}$, the equilibrium induced by the HP billing is also more efficient in terms of system cost.

PoE (Figure 4.5), as already seen in the simplified framework of Section 4.4.2. We observe that, in Section 4.4.2 as in this realistic case, $\text{PoA}(\mathcal{G}_\theta^{\text{DP}})$ is an unimodal function of θ.

Figure 4.4 shows that the PoA induced by the equilibrium of $\mathcal{G}_\theta^{\text{HP}}$ remains very low (it is maximal at $\theta = 0$ with PoA=1.0015 and then decreases) while the PoA of DP reaches a maximum of 1.122 at θ=0.06: this billing mechanism is much less robust to consumers' preferences. This lack of robustness is underlined by the important discrepancy between the minimal and maximal PoA values over our set of 31 days. Figure 4.5 shows that the PoE, as a function of θ, is much more concave for the DP mechanism, resulting in a larger system cost on a wide range of θ (the two curves intersect at $\theta \simeq 3 \cdot 10^{-4}$). As a result, the HP mechanism will also be more interesting for the provider. For $\theta \leqslant 3 \cdot 10^{-4}$ the system cost is larger for HP than for DP because of the small PoA mentioned before [Jac+17b].

4.6 Conclusion

We considered a game theoretic model to study the behavior of residential consumers in a DR program. We formulated an energy consumption game with a temporal preference

term in each user's cost function. We gave several theoretical results on a simplified test case and showed by simulations that those results still hold in a realistic framework where consumers have a nonflexible load. Without consumers preferences, the Daily Proportional billing reaches the optimal social cost and is more efficient than the Hourly Proportional billing which is not exactly optimal. When we add the temporal preference term, the Hourly Proportional billing becomes much more advantageous than the Daily Proportional mechanism in terms of social cost and in terms of costs induced for the provider.

Several extensions of this work could be considered. First, the theoretical results could be extended to take into account a nonflexible part, or considering general functions instead of a quadratic model for the system cost. Besides, we could study a dynamic population of users who have the choice to remain in the demand response program or not: if consumers are not satisfied with the program, they might consider another kind of contract or suscribing a more competitive provider.

Part III

Efficient Estimation of Equilibria in Large Games: from Nash to Wardrop

In Part II, we have seen that the decentralized management of electricity consumers flexibility (demand response management) can be modeled as an atomic splittable congestion game, where consumers are the players optimizing their electricity bill or a price signal. In those chapters, we have also explained that a Nash equilibrium of the considered game is a desirable outcome to achieve, both in terms of stability and in terms of efficiency (small price of anarchy). In the energy applications, we can imagine the situation of an operator (an electricity aggregator) in charge of the consumption flexibilities of several thousands of electricity consumers.

Despite the decentralized aspects of the proposed approaches, there is still a need for the operator to anticipate the outcome of the system. The reasons for anticipation are multiple: for instance, to optimize some price signal, to forecast an aggregate consumption to use on forward electricity markets, or to propose quantitative services to the system operator.

However, as discussed in Part II, computing a Nash equilibrium in a splittable congestion game with a very large number of players is computationally difficult: in this Part III of the thesis, we propose some methods and results related to the approximation of a Nash equilibrium with a very large number of players.

In this Part III, we are considering a more general setting than in Part II, as we consider the notion of *coupling constraints* and the associated concept of *generalized equilibrium*. Coupling constraints are constraints impacting several players in the game considered: in the model of an electric system, these kind of constraints appear when considering bounds on the aggregate consumption profile, global capacity constraints or ramp constraints on production generators. We refer the reader to Section 5.4 in Chapter 5 for more details.

While these chapters are more theoretical than the aspects developed in Part II, the work exposed in Chapters 5 and 6 is motivated by the use of large congestion games in the modeling of demand response. This application is developed and used to illustrate the proposed results in both Section 5.4 and Section 6.4.

The approximation results developed in Chapters 5 and 6 are based on the characterization of equilibria as solutions to *variational inequalities.* Even if the point of view adopted in each chapter is different from the one adopted in the other, the reader will observe some symmetry in the derived results:

- in Chapter 5 we consider the case of generalized *atomic* splittable congestion games, with coupling constraints, and with a very large number of players (see discussion on demand response aggregator above) that are heterogeneous through their action sets and their utility functions. We obtain an approximation of the variational Nash equilibria—a subset of generalized Nash equilibria in the presence of coupling constraints—of a large atomic congestion game by a Wardrop equilibrium of an auxiliary population game, where each population corresponds to a group of atomic players of the initial game. Because the variational inequalities characterizing the equilibrium of the auxiliary game have smaller dimension than the original problem, this approach enables the fast computation of an estimation of equilibria in a large congestion game with thousands of heterogeneous players. This chapter is based on the work [Jac+18a];

- in Chapter 6, the point of view adopted is different. We consider modeling the interactions of a population described by continuous parametric density functions: in this case, one obtains a *nonatomic* game with an infinity of different players types (as opposed to the *atomic* framework in Chapter 5). In this work, we consider the notion of nonatomic *aggregative games*, which generalizes congestion games. We define and analyze the notion of variational Wardrop equilibrium for nonatomic aggregative games with an infinity of players types. These equilibria are characterized through an infinite-dimensional variational inequality. We show, under monotonicity conditions, a convergence theorem which enables to approximate such an equilibrium with arbitrary precision. To this end, we introduce a sequence of nonatomic games with a finite number of players types, which approximates the initial game. We show the existence of a symmetric Wardrop equilibrium in each of these games. We prove that the sequence of symmetric equilibria converges to an equilibrium of the infinite-type game, and that the symmetric equilibria can be computed as solutions of finite-dimensional variational inequalities. The model is illustrated through an example

from smart grids: the description of a large population of electricity consumers by a parametric distribution gives a nonatomic game with an infinity of different players types, with actions subject to coupling constraints. This chapter is based on the paper [JW19], submitted for publication.

Chapter 5

Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players

This chapter is based on the paper [Jac+18a], submitted for publication.

5.1 Introduction

Motivation Congestion games form a class of noncooperative games [Nis+07]. In a congestion game, each player chooses a certain quantity of each of the available resources, and pays a cost for each resource obtained by the per-unit cost of that resource multiplied by the quantity she has chosen. A congestion game is said to be atomic if there is a finite number of players, and nonatomic if there is a continuum of infinitesimal players. The particularity of congestion games is that the per-unit cost of each resource depends only on its total demand. Congestion games find practical applications in various fields such as traffic management [Zie+08], communications [Scu+12; Alt+06] and more recently in electrical systems [MR+10; Jac+19c].

The concept of Nash equilibrium (NE) [Nas50] has emerged as the most credible outcome in the theory of noncooperative games. However, it has been shown in different frameworks that the computation of a NE is a hard problem: for instance, [FPT04] shows that computing a pure NE in finite two-players matricial games is PPAD-complete. [KW18] shows that the problem of computing of an NE in splittable congestion games with *player-specific affine* latency functions is also PPAD-complete.

In splittable congestion games, NEs can be characterized as solutions of variational inequalities. Therefore, the efficiency of the computation of NEs depends on the dimension of the variational inequalities in question, hence on the number of players and the number of constraints. For the description of practical situations (transport, energy, etc), one may consider several thousands of heterogeneous agents. At this scale, computing a Nash equilibrium can be intractable. In the case where the model involves coupling constraints, the computation of generalized Nash equilibria [Har91] is even harder to solve. Meanwhile, coupling constraints commonly exist in practice: for instance, in transportation, roads and communication channels have a limited capacity that should be considered. In the energy domain, production plants are also limited in the magnitude of variations of power, inducing some "ramp constraints" [CA06].

However, estimating the outcome situation—assumed to be an equilibrium—is often a priority for the operator of the system. For instance, the operator controls some variables such as physical or managerial parameters, of a communication or transport network and wishes to optimize the performance of the system. The computation of equilibria or their approximation is also a key aspect in bi-level programming [CMS07], where the lower level corresponds to a usually large scale game, and the upper level corresponds to a decision problem of an operator choosing optimal parameters. These parameters, such as prices or taxes, are to be applied in the low level game, with the aim of maximizing the revenue in various industrial sectors and public economics, such as highway management, urban traffic

control, air industry, freight transport and radio network [LMS98; Bro+00; Bro+01; CMS03; Eli+13].

In this chapter, we consider atomic congestion games with a finite but large number of players. We propose a method to compute an approximation of variational Nash equilibria (VNEs) [Har91]. In the presence of coupling constraints VNEs form a subset of the NEs while, without coupling constraints, all NEs are VNEs. The main idea is to reduce the dimension of the variational inequalities characterizing VNEs. The players are divided into groups with similar characteristics. Then, each group is replaced by a homogeneous population of nonatomic players. To provide an estimation of the equilibria of the original game, we compute a variational Wardrop equilibrium (VWE) [War52] in the approximating nonatomic population game.The quality of the estimation depends on how well the characteristics, such as action set and cost function, of each homogeneous population approximate those of the atomic players it replaces.

In bi-level programming, the uniqueness of a low level equilibrium allows for clear-cut comparative statics and sensitivity analysis at the high level. Last, from a computational point of view, different algorithms have been proposed to solve monotone variational inequalities corresponding to VNE or VWE, such as [Coh88; Fuk86; ZM93; FP07; FK10], and more recently [YP18; YP17; PGL17; TK18] and the references therein.

Related works The relation between VNEs in large games and VWEs has been studied in Gentile et al. [Pac+18]. In their paper, the authors also consider atomic congestion games with coupling constraints and show, using a variational inequalities approach, that the distance between an VNE and a VWE converges to zero when the number of players tends to infinity. Their VWE corresponds to an equilibrium of the game where each atomic player is replaced by a population, as done in Section 5.3.1 of this chapter. The objective of the present work is different. We look for an approximation of VNEs by *reducing the dimension* of the original game. To this end, in Section 5.3.2, we regroup many players into few homogeneous populations. This latter perspective, along with the subdifferentiable case we consider, are not considered in [Pac+18].

In [JW18b], Jacquot and Wan show that, in congestion games with a continuum of heterogeneous players, the WE can be approximated by an NE of an approximating game with a finite number of players. In [JW18a], those results are extended to aggregative games, a more general class of games including congestion games, furthermore with nonsmooth cost functions.

The approach developed in the present chapter is actually the inverse of the one taken in [JW18b] and [JW18a]: here, the WE in the auxiliary game serves as an approximation of an NE of the original large game.

Main contributions The contributions of this chapter are the following.

- We define an approximating population game (Section 5.3.2). The idea is that the auxiliary game has smaller dimension but is close enough to the original large game—quantified through the Hausdorff distance between action sets and between subgradients of players' objective functions.

- We show that a particular variational Wardrop equilibrium (VWE) of the approximating population game is close to any variational Nash equilibria (VNE) of the original game with or without coupling constraints, while the computation of the former is much faster than the latter because of the dimension reduction. We provide an explicit expression of the error bound of the approximating VWE (Theorem 5.2).

- We give auxiliary results on variational equilibria: we show that when the number of players is large, VNEs are close to each other (Corollary 5.1) and that VNEs are close to the approximating VWE (Corollary 5.3). This last theorem extends [Pac+18, Thm. 1] to the case of nondifferentiable cost functions (in the framework of congestion games, but can be easily extended to aggregative games).

• Last, we provide a numerical illustration of our results (Section 5.4) based on a practical application: the decentralized charging of electric vehicles through a demand response mechanism [PD11]. This example illustrates the nondifferentiable case through piece-wise linear electricity prices ("block rates tariffs"), with coupling constraints of capacities and limited variations on the aggregate load profile between time periods. This example shows that the proposed method is implementable and that it reduces the time needed to compute an equilibrium by computing its approximation (six times faster for an approximation with a normalized distance to the actual equilibrium of less than 2%).

Structure The remainder of this chapter is organized as follows: Section 5.2 specifies the framework of congestion games with coupling constraints, and recalls the notions of variational equilibria and monotonicity for variational inequalities, as well as several results on the existence and uniqueness of equilibria. Section 5.3 formulates the main results: Section 5.3.1 shows that a VWE approximates VNEs in large games and then, Section 5.3.2 formulates the approximating population game with the approximation measures, and gives an upper bound on the distance between the VWE profile of the approximating game to an original VNE profile. Section 5.4 presents a numerical illustration in the framework of demand response for electric vehicle smart charging.

5.2 Congestion Games with Coupling Constraints

5.2.1 Model and equilibria

The original game throughout this chapter is an *atomic splittable* congestion game: a finite set of resources is shared among finitely many players, and each resource incurs a cost increasing with the aggregate demand for it. In an atomic game, there are finitely many players whose actions are not negligible on the aggregate profile and on the objectives of other players. The term "atomic" is opposed to "nonatomic": in the latter case, players have an infinitesimal weight [Nis+07]. The term "splittable" refers to the infinite number of choices of pure actions $\boldsymbol{x}_n \in \mathcal{X}_n$ for each player n, as opposed to the unsplittable case where each player can only choose one action in a finite subset of $2^{\mathcal{T}}$ [Ros73a]. The formal definition is as follows.

Definition 5.1. An *atomic splittable congestion* game $\mathcal{G}$ is defined by:

- a finite set of players: $\mathcal{N} = \{1, \dots, n, \dots, N\}$,
- a finite set of resources: $\mathcal{T} = \{1, \dots, t, \dots, T\}$,
- for each resource t, a cost function $c_t : \mathbb{R}_+ \to \mathbb{R}$,
- for each player n, a set of feasible choices: $\mathcal{X}_n \subset \mathbb{R}_+^T$, an element $\boldsymbol{x}_n = (x_{n,t})_{t\in\mathcal{T}} \in \mathcal{X}_n$ signifies that n has demand $x_{n,t}$ for resource t,
- for each player n, an individual utility function $u_n : \mathcal{X}_n \to \mathbb{R}$,
- a coupling constraint set $\mathcal{A} \subset \mathbb{R}^T$.

We denote by $\boldsymbol{\mathcal{X}} \stackrel{\text{def}}{=} \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ the product set of action profiles. An action profile $\boldsymbol{x} = (\boldsymbol{x}_n)_{n\in\mathcal{N}} \in \boldsymbol{\mathcal{X}}$ induces a profile of average demand for the resources, denoted by $\overline{\boldsymbol{X}} = (\overline{X}_t)_{t\in\mathcal{T}} \stackrel{\text{def}}{=} (\frac{1}{N}\sum_{n\in\mathcal{N}} x_{n,t})_{t\in\mathcal{T}}$. We denote the set of feasible average demand profiles by:

$$\overline{\mathcal{X}} \stackrel{\text{def}}{=} \{\overline{\boldsymbol{X}} \in \mathbb{R}^T : \forall n \in \mathcal{N}, \exists \boldsymbol{x}_n \in \mathcal{X}_n \text{ s.t. } \tfrac{1}{N}\textstyle\sum_{n\in\mathcal{N}} \boldsymbol{x}_n = \overline{\boldsymbol{X}}\} .$$

With the coupling constraint $\mathcal{A}$, the set of feasible average demand profiles is $\overline{\mathcal{X}} \cap \mathcal{A}$, and the set of feasible action profiles is denoted by:

$$\boldsymbol{\mathcal{X}}(\mathcal{A}) = \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \tfrac{1}{N}\textstyle\sum_{n\in\mathcal{N}} \boldsymbol{x}_n \in \overline{\mathcal{X}} \cap \mathcal{A}\} .$$

Let the vector of cost functions be denoted by $\boldsymbol{c}(\overline{\boldsymbol{X}}) = (c_t(\overline{X}_t))_{t\in\mathcal{T}}$, where $c_t(\overline{X}_t)$ is the (per-unit of demand) cost of resource t when the average demand for it is $\overline{X}_t$.

Player n's cost function $f_n : \mathcal{X}_n \times \overline{\mathcal{X}} \to \mathbb{R}$ is defined for each $\boldsymbol{x}_n \in \mathcal{X}_n$ and $\overline{\boldsymbol{Y}} \in \overline{\mathcal{X}}$ by:

$$f_n(\boldsymbol{x}_n, \overline{\boldsymbol{Y}}) = \textstyle\sum_{t\in\mathcal{T}} x_{n,t} c_t(\overline{Y}_t) - u_n(\boldsymbol{x}_n), \quad . \tag{5.1}$$

Given $\boldsymbol{x}_{-n} \in \prod_{m\neq n} \mathcal{X}_m$ and $\overline{\boldsymbol{X}}_{-n} \stackrel{\text{def}}{=} \frac{1}{N}\sum_{m\neq n} \boldsymbol{x}_m$, player n's cost is $f_n(\boldsymbol{x}_n, \frac{1}{N}\boldsymbol{x}_n + \overline{\boldsymbol{X}}_{-n})$, composed of the network costs and her individual utility.

This atomic congestion game with coupling constraints is written as the tuple:

$$\mathcal{G} = (\mathcal{N}, \mathcal{T}, \boldsymbol{\mathcal{X}}, \mathcal{A}, \boldsymbol{c}, (u_n)_{n\in\mathcal{N}}) \, .$$

Atomic splittable congestion games are particular cases of aggregative games [Pac+18]: each player's cost function depends on the actions of the others only through the average profile $\overline{\boldsymbol{X}}$.

Remark 5.1. *In this chapter, following [Pac+18], the dependency of cost functions in other players is expressed through the average action $\overline{\boldsymbol{X}}$ rather than the aggregate action $\boldsymbol{X} \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} \boldsymbol{x}_n$. Considering a game with N fixed, the two formulations are equivalent, by a linear scaling of the cost functions. However, the former formulation enables an intrinsic scaling of the game and of cost functions with N, such that we can consider the limit $N \to \infty$ and avoid diverging quantities.*

In this chapter, we adopt the following standard assumptions:

Assumption 5.1.

i) For each player $n \in \mathcal{N}$, the set $\mathcal{X}_n$ is a convex and compact subset of $\mathbb{R}^T$ with nonempty relative interior.

ii) For each resource $t \in \mathcal{T}$, function c_t is convex (hence continuous) and non-decreasing on $(-\eta, +\infty)$ for a positive $\eta > 0$.

iii) For each player $n \in \mathcal{N}$, individual utility function u_n is concave (hence continuous) in $\boldsymbol{x}_n$ on $\mathcal{X}_n$.

iv) $\mathcal{A}$ is a convex closed set of $\mathbb{R}^T$, and $\overline{\mathcal{X}} \cap \mathcal{A}$ is not empty.

Example 5.1 (individual utility functions)**.** A common example of utility function is the distance between a player's choice and her preference $\boldsymbol{y}_n \in \mathcal{X}_n$: $u_n(\boldsymbol{x}_n) = -\omega_n \|\boldsymbol{x}_n - \boldsymbol{y}_n\|^2$, where $\omega_n > 0$ is the value that the player attaches to her preference. Another common example is $u_n(\boldsymbol{x}_n) = \omega_n \log(1 + \sum_t x_{n,t})$, which measures the satisfaction of a player according to the total demand she receives.

Finally, in congestion games, *aggregate constraints* are very common. For example, in routing games, there can be a capacity constraint linked to each arc. In energy consumption games, due to the operational constraints of the power grid, there can be both minimum and maximum consumption level for each time slot, and ramp constraints on the variation of energy consumption between time slots. This is why congestion games with aggregate constraints are of particular interest.

To separate player n's choice from those of the other players in her cost function, define

$$\hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{Y}}_{-n}) \stackrel{\text{def}}{=} f_n(\boldsymbol{x}_n, \overline{\boldsymbol{Y}}_{-n} + \tfrac{1}{N}\boldsymbol{x}_n)$$

for $\boldsymbol{x}_n \in \mathcal{X}_n$ and $\overline{\boldsymbol{Y}}_{-n} \in \overline{\mathcal{X}}_{-n} = \{\frac{1}{N}\sum_{m\in\mathcal{N}\setminus\{n\}} \boldsymbol{x}_m : \boldsymbol{x}_m \in \mathcal{X}_m, \forall m\}$.

Since $\boldsymbol{c}$ and u_n's are not necessarily differentiable, let us define the subdifferential of the players' utilities w.r.t. their actions for the characterization of an equilibrium.

Let us consider the two correspondences, $\hat{H}$ and H, from $\boldsymbol{\mathcal{X}}$ to $\mathbb{R}^{NT}$: for any $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$,

$$\begin{aligned}\hat{H}(\boldsymbol{x}) &\stackrel{\text{def}}{=} \{(\boldsymbol{h}_n)_{n\in\mathcal{N}} \in \mathbb{R}^{NT} : \boldsymbol{h}_n \in \partial_1 \hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n}),\ \forall n \in \mathcal{N}\} \\ &= \textstyle\prod_{n\in\mathcal{N}} \partial_1 \hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n})\ ; \\ H(\boldsymbol{x}) &\stackrel{\text{def}}{=} \{(\boldsymbol{h}_n)_{n\in\mathcal{N}} \in \mathbb{R}^{NT} : \boldsymbol{h}_n \in \partial_1 f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}),\ \forall n \in \mathcal{N}\} \\ &= \textstyle\prod_{n\in\mathcal{N}} \partial_1 f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}})\ ,\end{aligned}$$

where ∂_1 signifies the partial differential w.r.t. the first variable of the function. The interpretation of $\hat{H}(\boldsymbol{x})$ is clear: $\boldsymbol{h}_n$ is a subgradient of player n's utility function $\hat{f}_n$ w.r.t. her action $\boldsymbol{x}_n$. Let us leave the interpretation of $H(\boldsymbol{x})$ until Definition 5.3. For the moment, let us write the explicit expression of $\hat{H}$ and H:

Lemma 5.1. *For each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$, we have:*

- *$\boldsymbol{h} \in \hat{H}(\boldsymbol{x})$ if and only if there are $\mathbf{g}'_n \in \partial(-u_n)(\boldsymbol{x}_n)$ and $\boldsymbol{a}_n \in \prod_{t\in\mathcal{T}} \partial c_t(\overline{X}_t)$ $\forall n$ s.t.:*

$$\boldsymbol{h}_n = \boldsymbol{c}(\overline{\boldsymbol{X}}) + (\tfrac{1}{N} x_{n,t}\, a_{n,t})_{t\in\mathcal{T}} + \mathbf{g}'_n\ , \quad \forall n \in \mathcal{N}\ ;$$

- *$\boldsymbol{h}' \in H(\boldsymbol{x})$ if and only if there is $\mathbf{g}'_n \in \partial(-u_n)(\boldsymbol{x}_n)$ $\forall n$ s.t.:*

$$\boldsymbol{h}'_n = \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}'_n\ , \quad \forall n \in \mathcal{N}\ .$$

where $\partial(-u_n)(\boldsymbol{x}_n)$ is the subdifferential of convex function $-u_n$ at $\boldsymbol{x}_n$ and $\partial c_t(\overline{X}_t)$ the subdifferential of c_t at $\overline{X}_t$ (w.r.t. $\overline{X}_t$).

Proof. See Section 5.A. □

In the presence of coupling constraints, the notion of Nash equilibrium (NE) [Nas50] is replaced by the one of Generalized Nash Equilibrium (GNE). A profile $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ is a GNE if, for each player n, $\hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n}) \leqslant \hat{f}_n(\boldsymbol{y}_n, \overline{\boldsymbol{X}}_{-n})$ for all $\boldsymbol{y}_n$ s.t. $\boldsymbol{y}_n + \overline{\boldsymbol{X}}_{-n} \in \overline{\mathcal{X}} \cap \mathcal{A}$. For atomic games, a special class of GNE is called Variational Nash Equilibria (VNE) [Har91; KS12], which enjoys some symmetric properties: in some sense, the burden of constraint $\overline{\boldsymbol{X}} \in \mathcal{A}$ is shared symmetricaly by players (we refer to [Har91] for more details). VNEs, in the subdifferentiable case, are characterized as the solution of generalized Variational Inequalities (GVI) (5.2) stated below.

Definition 5.2 (Variational Nash Equilibrium (VNE), [Har91])**.** A VNE is a solution $\hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ to the following GVI problem:

$$\exists\, \mathbf{g} \in \hat{H}(\hat{\boldsymbol{x}}) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \geqslant 0,\ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A}). \tag{5.2}$$

In particular, if $\overline{\mathcal{X}} \subset \mathcal{A}$, a VNE is an NE.

In this work, we adopt the notion of VNE as the equilibrium notion in the presence of aggregate constraints.

As the first step of approximation, let us define a nonatomic congestion game $\mathcal{G}(\mathcal{A})'$ associated to $\mathcal{G}(\mathcal{A})$. Let each player n be replaced by a continuum of identical nonatomic players, represented by interval $[0, 1]$ with each point thereon corresponding to a nonatomic player. Each player in population n has action set $\mathcal{X}_n$ and individual utility function u_n.

Definition 5.3. A *symmetrical variational Wardrop equilibrium* (SVWE) of $\mathcal{G}(\mathcal{A})'$ is a solution $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ to the GVI:

$$\exists \mathbf{g} \in H(\boldsymbol{x}^*) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \boldsymbol{x}^* \rangle \geqslant 0,\ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})\ . \tag{5.3}$$

For the definition of variational Wardrop equilibrium (VWE) and further discussion, we refer to [JW18a]. In particular, a VWE is characterized by an infinite dimensional variational inequality. Here, we consider only those VWE where all the nonatomic players in population n *take the same action* $\boldsymbol{x}_n$. Such a SVWE exists because the players are identical in the same population.

Aside from the interpretation of each individual player as a population, a second interpretation of SVWE is that, when the number of players is very large, the individual contribution of each player on the aggregate action $\boldsymbol{X}$ is almost negligible, the term $\frac{1}{N} x_{n,t}\, a_{n,t}$ in $\boldsymbol{h} \in \hat{H}(\boldsymbol{x})$ is so small that $\hat{H}(\boldsymbol{x})$ can be approximated by $H(\boldsymbol{x})$ (cf. Lemma 5.1). This is the interpretation adopted in [Pac+18]. However, a SVWE of $\mathcal{G}(\mathcal{A})'$ is not an equilibrium of $\mathcal{G}(\mathcal{A})$ in the sense of a "stable state" for the atomic congestion game, as the individual profile $\boldsymbol{x}_n^*$ affected to a player n at a SVWE does not necessarily minimize her cost function $\hat{f}_n(., \overline{\boldsymbol{X}}_{-n})$.

The existence of equilibria defined in Definitions 5.2 and 5.3 are obtained under mild assumptions:

Proposition 5.1 (Existence of equilibria)**.** *Under Assumption 5.1, $\mathcal{G}(\mathcal{A})$ (resp. $\mathcal{G}(\mathcal{A})'$) admits a VNE (resp. SVWE).*

Proof. : We observe that $\hat{f}_n(\cdot, \overline{\boldsymbol{X}}_{-n})$ is convex on $\mathcal{X}_n$ for all $\overline{\boldsymbol{X}}_{-n}$ in $\overline{\mathcal{X}}_{-n}$, $f_n(\cdot, \overline{\boldsymbol{X}})$ is convex on $\mathcal{X}_n$ for all $\overline{\boldsymbol{X}} \in \overline{\mathcal{X}}$, and $\hat{H}$ and H are nonempty, convex, compact valued, upper hemicontinuous correspondences.

From [RW09, Proposition 8.7], we know that the subgradient mapping of a convex function is upper hemicontinuous. Thus, $\hat{H}$ and H are also upper hemicontinuous as linear mappings of subgradients.

Then [CP82, Corollary 3.1] shows that the GVI problems (5.2) and (5.3) both admit a solution on the finite dimensional convex compact $\boldsymbol{\mathcal{X}}(\mathcal{A})$.

Before discussing the uniqueness of equilibria, let us recall some relevant definition of monotonicity for correspondences:

Definition 5.4. A correspondence $\Gamma : \boldsymbol{\mathcal{X}} \rightrightarrows \mathbb{R}^T$ is:

i) *monotone* if for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}, \mathbf{g} \in \Gamma(\boldsymbol{x}), \mathbf{h} \in \Gamma(\boldsymbol{y})$:

$$\textstyle\sum_{n\in\mathcal{N}} \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant 0\,; \tag{5.4}$$

ii) *strictly monotone* if the equality in (5.4) holds *iff* $\boldsymbol{x} = \boldsymbol{y}$;

iii) *aggregatively strictly monotone* if the equality in (5.4) holds *iff* $\sum_n \boldsymbol{x}_n = \sum_n \boldsymbol{y}_n$;

iv) *α-strongly monotone* if $\alpha > 0$ and, for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$:

$$\sum_{n\in\mathcal{N}} \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant \alpha \|\boldsymbol{x} - \boldsymbol{y}\|^2,\ \forall \mathbf{g} \in \Gamma(\boldsymbol{x}), \mathbf{h} \in \Gamma(\boldsymbol{y})\,; \tag{5.5}$$

v) *β-aggregatively strongly monotone* on $\boldsymbol{\mathcal{X}}$ if $\beta > 0$ and, for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$ with $\overline{\boldsymbol{X}} = \frac{1}{N}\sum_n \boldsymbol{x}_n$, $\overline{\boldsymbol{Y}} = \frac{1}{N}\sum_n \boldsymbol{y}_n$:

$$\sum_{n\in\mathcal{N}} \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant N\beta \|\overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}}\|^2,\ \forall \mathbf{g} \in \Gamma(\boldsymbol{x}), \mathbf{h} \in \Gamma(\boldsymbol{y})\,. \tag{5.6}$$

If $T = 1$, "monotone" corresponds to "increasing". Besides, (aggregatively) strict monotonicity implies monotonicity, while strong (resp. aggregatively strong) monotonicity implies strict (resp. aggregatively strict) monotonicity.

In Proposition 5.2 below, we recall some existing results concerning the uniqueness of VNE and SVWE, according to the monotonicity of $\hat{H}$ and H:

Proposition 5.2 (Uniqueness of equilibria)**.** *Under Assumption 5.1:*

i) if $\hat{H}$ (resp. H) is strictly monotone, then $\mathcal{G}(\mathcal{A})$ (resp. $\mathcal{G}(\mathcal{A})'$) has a unique VNE (resp. SVWE);

ii) if $\hat{H}$ (resp. H) is aggregatively strictly monotone, then all VNE (resp. SVWE) of $\mathcal{G}(\mathcal{A})$ (resp. $\mathcal{G}(\mathcal{A})'$) have the same aggregate profile;

iii) in the absence of aggregative constraint, if $\hat{H}$ (resp. H) is aggregatively strictly monotone and, for each $n \in \mathcal{N}$, $u_n(\boldsymbol{x})$ is strictly concave, then then $\mathcal{G}$ (resp. $\mathcal{G}'$) has a unique NE (resp. WE).

Proof. We prove only the uniqueness of SVWE only as the proof for the uniqueness of VNE is symmetric. Suppose that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ are both SVWE, with $\overline{\boldsymbol{X}} = \frac{1}{N}\sum_n \boldsymbol{x}_n$ and $\overline{\boldsymbol{Y}} = \frac{1}{N}\sum_n \boldsymbol{y}_n$. According to the definition of SVWE, there is $\mathbf{g} \in H(\boldsymbol{x})$ an $\mathbf{h} \in H(\boldsymbol{y})$ such that:

$$\textstyle\sum_n \langle \mathbf{g}_n, \boldsymbol{y}_n - \boldsymbol{x}_n \rangle \geqslant 0 \text{ and } \sum_n \langle \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant 0.$$

Adding up these two inequalities yields:

$$\textstyle\sum_n \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{y}_n - \boldsymbol{x}_n \rangle \geqslant 0.$$

i) If H is strictly monotone, then $\sum_n \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle = 0$ and thus $\boldsymbol{x} = \boldsymbol{y}$.
ii-iii) If H is aggregatively strictly monotone, then $\sum_n \langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle = 0$ and thus $\overline{\boldsymbol{X}} = \overline{\boldsymbol{Y}}$. If there is no aggregative constraint and u_n is strictly concave, then $\boldsymbol{x}_n$ (resp. $\boldsymbol{y}_n$) is the unique minimizer of $f_n(\cdot, \overline{\boldsymbol{X}})$ (resp. $f_n(\cdot, \overline{\boldsymbol{Y}})$). Since $\overline{\boldsymbol{X}} = \overline{\boldsymbol{Y}}$, one has $\boldsymbol{x}_n = \boldsymbol{y}_n$. □

Proposition 5.3 below gives sufficient conditions for the (strong) monotonicity of H to hold.

Proposition 5.3 (Monotonicity of H)**.** *Under Assumption 5.1,*

i) H is monotone;

ii) if for each $n \in \mathcal{N}$, u_n is α_n-strongly concave, then H is α-strongly monotone with $\alpha \stackrel{\text{def}}{=} \min_{n\in\mathcal{N}} \alpha_n$;

iii) if for each $t \in \mathcal{T}$, c_t is β_t-strictly increasing, then H is β-aggregatively strongly monotone with $\beta \stackrel{\text{def}}{=} \min_{t\in\mathcal{T}} \beta_t$.

Proof. See Section 5.B.

As opposed to the monotonicity of H shown in Proposition 5.3, $\hat{H}$ is rarely monotone (except in some particular cases, e.g. with $\boldsymbol{c}$ linear [ORS93; RS07]): even in the case where $\boldsymbol{c}$ is piece-wise linear, the Example 5.2 below shows that $\hat{H}$ can be non monotone.

Example 5.2. Let $N = 2$ and $T = 1$, $\mathcal{X}_1 = \mathcal{X}_2 = [0, 4]$. Consider the cost function $c(X) = X$ for $X \leqslant 4$ and $c(X) = 3X - 8$ for $X \geqslant 4$. Assumption 5.1 holds. Consider the profiles $x_1 = 3, x_2 = 1$ and $y_1 = 4, y_2 = 0$, then:

$$\begin{aligned} \mathbf{g} &\stackrel{\text{def}}{=} (c(4) + 3x_1, c(4) + 3x_2) \in \hat{H}(\boldsymbol{x}) \\ \text{and } \boldsymbol{h} &\stackrel{\text{def}}{=} (c(4) + y_1, c(4) + y_2) \in \hat{H}(\boldsymbol{y}) \,, \end{aligned}$$

$$\text{but: } \textstyle\sum_{n\in\{1,2\}} \langle \mathbf{g}_n - \boldsymbol{h}_n, x_n - y_n \rangle = -2 < 0 \,.$$

In view of Proposition 5.2, the absence of monotonicity of $\hat{H}$ can result in multiple VNEs: we refer the reader to [Bha+09] for counter examples. In [Jac+19c], a particular case with parallel arc network is shown to have a unique NE. In the next section, we shall prove that, when the number of players is very large, even if we cannot ensure their uniqueness, all VNEs are close to each other and, moreover, they are well approximated by an SVWE.

5.3 Approximating VNEs of a Large Game

5.3.1 Considering SVWE instead of VNE

The approximation of VNEs is done in two steps. The first step consists in replacing VNE by SVWE. According to the second interpretation of SVWE, the SVWE should be close to the VNEs in a large game. Now, let us formulate this idea and bound the distance between the two.

Let us denote by $\mathcal{X}_0 \subset \mathbb{R}_+^T$ the convex closed hull of $\bigcup_{n\in\mathcal{N}} \mathcal{X}_n$, and consider the radius:

$$R \stackrel{\text{def}}{=} \max_{\boldsymbol{x}\in\mathcal{X}_0} \max_{t\in\mathcal{T}} |x_t| \,.$$

Besides, let us denote the upper bound on the subgradients of $\boldsymbol{c}$ by:

$$C \stackrel{\text{def}}{=} \sup\{b \in \mathbb{R} : b \in \partial c_t(\overline{X}_t), \overline{\boldsymbol{X}} \in \overline{\mathcal{X}}, t \in \mathcal{T}\} .$$

The first step of approximation is based upon the following theorem which gives an upper bound on the distance between a VNE and an SVWE.

Theorem 5.1 (SVWE is close to VNE)**.** *Under Assumption 5.1, let $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ be a VNE of $\mathcal{G}(\mathcal{A})$ and $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ a SVWE of $\mathcal{G}(\mathcal{A})'$, then:*

i) if for each $n \in \mathcal{N}$, u_n is a α_n-strongly concave, then $\boldsymbol{x}^*$ is unique and:

$$\|\boldsymbol{x} - \boldsymbol{x}^*\| \leqslant \tfrac{RC}{\alpha}\sqrt{\frac{T}{N}}, \quad \text{with } \alpha \stackrel{\text{def}}{=} \min_{n\in\mathcal{N}} \alpha_n; \tag{5.7}$$

$$\text{besides,} \quad \tfrac{1}{N}\sum_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\| \leqslant \tfrac{RC}{\alpha}\frac{\sqrt{T}}{N} \quad \text{and} \quad \|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\| \leqslant \tfrac{RC}{\alpha}\frac{\sqrt{T}}{N}; \tag{5.8}$$

ii) if for each $t \in \mathcal{T}$, c_t is β_t-strictly increasing, then $\boldsymbol{X}^*$ is unique and:

$$\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\| \leqslant R\sqrt{\tfrac{2TC}{\beta N}}, \quad \text{with } \beta \stackrel{\text{def}}{=} \min_{t\in\mathcal{T}} \beta_t . \tag{5.9}$$

Proof. By (5.2), there are $\mathbf{g}_n \in \partial_1 \hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n})$ and $\mathbf{h}_n^* \in \partial_1 f_n(\boldsymbol{x}_n^*, \overline{\boldsymbol{X}}_{-n}^*)$ for each n such that:

$$\begin{aligned} \mathbf{g}_n &= \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}'_n + (\tfrac{1}{N} a_{n,t}\, x_{n,t})_t \quad \text{and } \mathbf{g} \stackrel{\text{def}}{=} (\mathbf{g}_n)_n \in \hat{H}(\boldsymbol{x}) \\ \mathbf{h}_n &= \boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{h}'_n \text{ and } \mathbf{g} \stackrel{\text{def}}{=} (\mathbf{g}_n)_n \in H(\boldsymbol{x}^*) , \end{aligned}$$

where $\mathbf{g}'_n \in \partial(-u_n)(\boldsymbol{x}_n)$, $\mathbf{h}'_n \in \partial(-u_n)(\boldsymbol{x}_n^*)$, $a_{n,t} \in \partial c_t(\overline{X}_t)$ for all t, such that:

$$\langle \mathbf{g}, \boldsymbol{x}^* - \boldsymbol{x}\rangle \geqslant 0 \text{ and } \langle \mathbf{h}, \boldsymbol{x} - \boldsymbol{x}^*\rangle \geqslant 0 .$$

Summing up these two inequalities yields:

$$\begin{aligned} 0 &\leqslant \langle \mathbf{g} - \mathbf{h}, \boldsymbol{x}^* - \boldsymbol{x}\rangle \\ &= \textstyle\sum_n \langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \boldsymbol{x}_n^* - \boldsymbol{x}_n\rangle + \sum_n \langle \mathbf{g}'_n - \mathbf{h}'_n, \boldsymbol{x}_n^* - \boldsymbol{x}_n\rangle + \tfrac{1}{N}\langle (a_{n,t}x_{n,t})_{n,t}, \boldsymbol{x}^* - \boldsymbol{x}\rangle . \end{aligned}$$

Therefore, rearranging the terms and using Cauchy-Schwartz inequality:

$$N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\rangle + \textstyle\sum_n \langle \mathbf{g}'_n - \mathbf{h}'_n, \boldsymbol{x}_n - \boldsymbol{x}_n^*\rangle \leqslant \left\| \tfrac{1}{N}(a_{n,t}x_{n,t})_{n,t}\right\| \|\boldsymbol{x}^* - \boldsymbol{x}\| . \tag{5.10}$$

Besides, we have:

$$\left\| \tfrac{1}{N}(a_{n,t}x_{n,t})_{n,t}\right\|^2 = \tfrac{1}{N^2}\sum_{n,t}(a_{n,t}x_{n,t})^2 \leqslant \tfrac{1}{N^2}NTC^2R^2 = \tfrac{1}{N}TC^2R^2 \tag{5.11}$$

i) Since $\boldsymbol{c}$ is monotone and so are ∂u_n's because u_n's are concave, we have:

$$\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\rangle \geqslant 0 \text{ and } \sum_n \langle \mathbf{g}'_n - \mathbf{h}'_n, \boldsymbol{x}_n - \boldsymbol{x}_n^*\rangle \geqslant 0 .$$

If for each n, u_n's are α_n-strongly concave, then we obtain by using (5.10) and (5.11):

$$\alpha \sum_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\|^2 \leqslant \sum_n \alpha_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\|^2 \leqslant \sum_n \langle \mathbf{g}'_n - \mathbf{h}'_n, \boldsymbol{x}_n - \boldsymbol{x}_n^*\rangle \leqslant \sqrt{\tfrac{T}{N}}CR\, \|\boldsymbol{x}^* - \boldsymbol{x}\|$$

so that we obtain the desired inequality $\|\boldsymbol{x} - \boldsymbol{x}^*\| \leqslant \frac{RC}{\alpha}\sqrt{\frac{T}{N}}$. Besides, by Cauchy-Scharz inequality, we have:

$$(\sum_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\|)^2 \leqslant N\sum_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\|^2 \leqslant \frac{R^2C^2}{\alpha^2} N\frac{T}{N} = \frac{R^2C^2}{\alpha^2}T$$

which implies that $\frac{1}{N}\sum_n \|\boldsymbol{x}_n - \boldsymbol{x}_n^*\| \leqslant \frac{RC}{\alpha}\frac{\sqrt{T}}{N}$. Similarly, we have:

$$N\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\| = \|\sum_n (\boldsymbol{x}_n - \boldsymbol{x}_n^*)\| \leqslant \sqrt{N\|\boldsymbol{x} - \boldsymbol{x}^*\|^2} \leqslant \frac{RC}{\alpha}\sqrt{T}$$

hence $\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\| \leqslant \frac{RC}{\alpha}\frac{\sqrt{T}}{N}$.
ii) If for each t, c_t is β_t-strictly increasing, combining (5.10) and (5.11) and the bound $\|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \leqslant NT(2R)^2$, we obtain:

$$N\beta\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\|^2 \leqslant N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\rangle \leqslant \sqrt{\tfrac{T}{N}}CR\sqrt{NT}(2R) = 2TCR^2$$

thus $\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\| \leqslant R\sqrt{\frac{2TC}{N\beta}}$. □

Proposition 5.2 and Proposition 5.3 show that, in general, VNEs are not unique. However, applying the triangle inequality to the results of Theorem 5.1, we can obtain upper bounds on the distance between two VNEs when the set of players is large, as stated in Corollary 5.1 below.

Corollary 5.1 (VNEs are close to each other)**.** *Under Assumption 5.1, let $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\boldsymbol{\mathcal{X}}(\mathcal{A})$ be two distinct VNEs of $\mathcal{G}(\mathcal{A})$. Then:*

i) if for each $n \in \mathcal{N}$, u_n is α_n-strongly concave, then:

$$\|\boldsymbol{x} - \boldsymbol{y}\| \leqslant 2\tfrac{RC}{\alpha}\sqrt{\frac{T}{N}}, \quad \text{with } \alpha \stackrel{\text{def}}{=} \min_{n\in\mathcal{N}} \alpha_n;$$

$$\text{besides, } \tfrac{1}{N}\sum_n \|\boldsymbol{x}_n - \boldsymbol{y}_n\| \leqslant 2\tfrac{RC}{\alpha}\frac{\sqrt{T}}{N} \quad \text{and} \quad \|\overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}}\| \leqslant 2\tfrac{RC}{\alpha}\frac{\sqrt{T}}{N};$$

ii) if for each $t \in \mathcal{T}$, c_t is β_t-strictly increasing, then:

$$\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}}\| \leqslant 2R\sqrt{\tfrac{2TC}{\beta N}}, \quad \text{with } \beta \stackrel{\text{def}}{=} \min_{t\in\mathcal{T}} \beta_t\,.$$

Theorem 5.1 shows that, if the number of players N is large, then the (aggregate) SVWE will provide a good approximation of a (aggregate) VNE of $\mathcal{G}(\mathcal{A})$. Similar results are obtained in [Pac+18] for the differentiable case. However, this does not reduce the dimension of the GVI to resolve: the GVI characterizing the VNE and those characterizing the SVWE have the same dimension. For this reason, the second step of approximation consists in regrouping similar populations.

5.3.2 Classification of populations

In this subsection, we shall regroup the populations in $\mathcal{G}(\mathcal{A})'$ with similar strategy sets $\mathcal{X}_n$ and utility subgradients $\partial(-u_n)$, into larger populations, endow them with a common strategy set and a common utility function, so that the SVWE of this new population game approximates the SVWE of $\mathcal{G}(\mathcal{A})'$. The similarity between sets is measured through the Hausdorff distance:

Definition 5.5. The Hausdorff distance between two sets $\mathcal{X}$ and $\mathcal{Y}$ is defined as:

$$d_H(\mathcal{X}, \mathcal{Y}) \stackrel{\text{def}}{=} \max\left(\max_{\boldsymbol{x}\in\mathcal{X}} d(\boldsymbol{x}, \mathcal{Y}), \max_{\boldsymbol{y}\in\mathcal{Y}} d(\boldsymbol{y}, \mathcal{X})\right).$$

At the SVWE of the new population game with a reduced dimension, all the nonatomic players in the same population play the same action, by the definition of SVWE. Therefore, in order for this new SVWE to well approximate the SVWE in $\mathcal{G}(\mathcal{A})'$, we must ensure that populations with similar characteristics in $\mathcal{G}(\mathcal{A})'$ do play similar actions at the SVWE of $\mathcal{G}(\mathcal{A})'$. Proposition 5.4 formulates this results in the case without coupling constraint.

Let us define the compact set

$$\mathcal{M} \stackrel{\text{def}}{=} [0, R+\delta]^T ,$$

where $\delta \geqslant \max_{n,m\in\mathcal{N}} d_H(\mathcal{X}_n, \mathcal{X}_m)$. Without loss of generality, we assume that for each $n \in \mathcal{N}$, u_n can be extended to a neighborhood of $\mathcal{M}$, and is bounded on $\mathcal{M}$. Let us introduce the notations for upper bounds:

$$\begin{aligned} B_{\mathbf{u_n}} &\stackrel{\text{def}}{=} \sup\{\|\mathbf{g}'_n\| : \mathbf{g}'_n \in \partial(-u_n)(\boldsymbol{x}_n), \boldsymbol{x}_n \in \mathcal{X}_n\} \\ \text{and } B_{\boldsymbol{c}} &\stackrel{\text{def}}{=} \sup\{\|\boldsymbol{c}(\overline{\boldsymbol{X}})\| : \overline{\boldsymbol{X}} \in \overline{\mathcal{X}}\} . \end{aligned}$$

Proposition 5.4. *Under Assumption 5.1, let $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ be a SVWE of $\mathcal{G}'(\mathcal{A})$ with $\mathcal{A} = \mathbb{R}^T$ (no coupling constraint). For two populations $n, m \in \mathcal{N}$, if u_n is α_n-strongly concave, $d_H(\mathcal{X}_n, \mathcal{X}_m) \leqslant \delta$, and $\sup_{\boldsymbol{x}_m\in\mathcal{X}_m} \sup_{\mathbf{g}'_m\in\partial(-u_m)(\boldsymbol{x}_m)} d(\mathbf{g}'_m, \partial(-u_n)(\boldsymbol{x}_m)) \leqslant d$, then*

$$\|\boldsymbol{x}^*_n - \boldsymbol{x}^*_m\|^2 \leqslant \tfrac{1}{\alpha_n}\big((B_{\mathbf{u_n}} + B_{\mathbf{u_m}} + 2B_{\boldsymbol{c}})d + 2\delta TR\big).$$

Proof. : See Section 5.C.

In the case with coupling constraints, the proof for a similar result is more complicated, and we leave it to Corollary 5.2.

Let us now present the regrouping procedure. Denote by $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$ an auxiliary game, with a set $\mathcal{I}$ of p populations. Each population $i \in \mathcal{I}$ corresponds to a subset $\mathcal{N}_i$ of populations in the game $\mathcal{G}(\mathcal{A})'$, such that $\bigcup_{i\in\mathcal{I}} \mathcal{N}_i = \mathcal{N}$ and for any $i, j \in \mathcal{I}, \mathcal{N}_i \cap \mathcal{N}_j = \emptyset$, i.e $(\mathcal{N}_i)_{i\in\mathcal{I}}$ forms a partition of $\mathcal{N}$. Denote $p_i = |\mathcal{N}_i|$ the number of original populations now included in i. By abuse of notations, let i also denote the interval $[0, p_i]$, so that each nonatomic player in population i is represented by a point $\theta \in [0, p_i]$. The common action set of each nonatomic player in i is a compact convex subset of $\mathbb{R}^T$, denoted by $\mathcal{X}_i$.

Each player θ in each population i having chosen action $\boldsymbol{x}_\theta$, let

$$\overline{\boldsymbol{X}} \stackrel{\text{def}}{=} \tfrac{1}{N} \sum_{i\in\mathcal{I}} \int_{\theta\in i} \boldsymbol{x}_\theta \, \mathrm{d}\theta$$

denotes the aggregate action profile. The aggregate action-profile set in $\mathbb{R}^T$ is then:

$$\overline{\mathcal{X}}^{\mathcal{I}} \stackrel{\text{def}}{=} \{\tfrac{1}{N}\textstyle\sum_{i\in\mathcal{I}} \int_{\theta\in i} \boldsymbol{x}_\theta \, \mathrm{d}\theta : \boldsymbol{x}_\theta \in \mathcal{X}_i, \forall\theta \in i, \forall i \in \mathcal{I}\} .$$

The cost function of player θ in population $i \in \mathcal{I}$ is:

$$f_i(\boldsymbol{x}_\theta, \overline{\boldsymbol{X}}) \stackrel{\text{def}}{=} \langle \boldsymbol{x}_\theta, \boldsymbol{c}(\overline{\boldsymbol{X}})\rangle - u_i(\boldsymbol{x}_\theta),$$

where the common individual utility function u_i for all the players in i is concave on a neighborhood of $\mathcal{X}_i$.

We are only interested in *symmetric* action profiles, i.e. where all the nonatomic players in the same population i play the same action. Denote the set of symmetric action profiles by $\boldsymbol{\mathcal{X}}^{\mathcal{I}} = \prod_{i\in\mathcal{I}} \mathcal{X}_i$. Let us point out that a symmetric action profile happens as a specific case in the non-cooperative game, without any coordination between the players within a population. Besides, considering the coupling constraint $\boldsymbol{X} \in \mathcal{A}$, we define

$$\boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A}) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}} : \overline{\boldsymbol{X}} = \tfrac{1}{N} \sum_{i\in\mathcal{I}} p_i \boldsymbol{x}_i \in \mathcal{A}\} .$$

Let us introduce two indicators to "measure" the quality of the clustering of $\tilde{\mathcal{G}}^{\mathcal{I}}$:

- $\overline{\delta} = \max_{i\in\mathcal{I}} \delta_i$, where
$$\delta_i \stackrel{\text{def}}{=} \max_{n\in\mathcal{N}_i} d_H\left(\mathcal{X}_n, \mathcal{X}_i\right) \ , \tag{5.12}$$
- $\overline{d} = \max_{i\in\mathcal{I}} d_i$, where
$$d_i \stackrel{\text{def}}{=} \max_{n\in\mathcal{N}_i} \sup_{\boldsymbol{x}\in\mathcal{X}_i} d_H\left(\partial(-u_i)(\boldsymbol{x}), \partial(-u_n)(\boldsymbol{x})\right). \tag{5.13}$$

The quantity δ_i measures the heterogeneity in strategy sets of populations within the group $\mathcal{N}_i$, while d_i measures the heterogeneity in the subgradients in the group $\mathcal{N}_i$.

Since the auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}$ is to be used to compute an approximation of an equilibrium of the large game $\mathcal{G}$, the indicators δ_i and d_i should be minimized when defining $\tilde{\mathcal{G}}^{\mathcal{I}}$. Thus, we assume that $(\mathcal{X}_i)_i$ and $(u_i)_i$ are chosen such that the following holds:

Assumption 5.2. *For each $i \in \mathcal{I}$, we have:*

i) $\mathcal{X}_i$ is in the convex hull of $\bigcup_{n\in\mathcal{N}_i} \mathcal{X}_n$, so that $\max_{\boldsymbol{x}\in\mathcal{X}_i} \max_{t\in\mathcal{T}} |x_t| \leqslant R$. Moreover, for each $n \in \mathcal{N}_i$, $\operatorname{aff}\mathcal{X}_n \subset \operatorname{aff}\mathcal{X}_i$, where $\operatorname{aff} S$ denotes the affine hull of set S;

ii) similarly, u_i is such that $\partial(-u_i)(\boldsymbol{x})$ is contained in the convex hull of $\bigcup_{n\in\mathcal{N}_i} \partial(-u_n)(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{X}_i$, so that $\|\partial(-u_i)\|_\infty \leqslant \max_{n\in\mathcal{N}_i} B_{\mathbf{u_n}}$.

An interesting case in the perspective of minimizing the quantities $\overline{\delta}$ and $\overline{d}$ is when $\mathcal{N}$ can be divided into homogeneous populations, as in Example 5.3 below.

Example 5.3. The player set $\mathcal{N}$ can be divided into a small number p of subsets $(\mathcal{N}_i)_{1\leqslant i\leqslant p}$, with homogeneous players inside each subset $\mathcal{N}_i$ (i.e., for each i and $n, m \in \mathcal{N}_i$, $\mathcal{X}_n = \mathcal{X}_m$ and $u_n = u_m$). In that case, consider an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}$ with i populations and, for each $i \in \mathcal{I}$ and $n \in \mathcal{N}_i$, $\mathcal{X}_i \stackrel{\text{def}}{=} \mathcal{X}_n$ and $u_i \stackrel{\text{def}}{=} u_n$. Then, $\overline{\delta} = \overline{d} = 0$.

In order to approximate the SVWE of $\mathcal{G}(\mathcal{A})'$ by the SVWE of an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}$, let us first state the following result on the geometry of the action sets for technical use.

Lemma 5.2. *Under Assumption 5.1, there exists a strictly positive constant ρ and an action profile $\boldsymbol{z} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ such that, $d(\boldsymbol{z}_n, \operatorname{rbd}\mathcal{X}_n) \geqslant \rho$ for all $n \in \mathcal{N}$, where* rbd *stands for the relative boundary.*

Proof. See App. 5.D. □

Lemma 5.2 ensures the existence of a profile $\boldsymbol{z}$ such that $\boldsymbol{z}_n$ has uniform distance to the relative boundary of $\mathcal{X}_n$ for all n and that $\boldsymbol{z}$ satisfies the coupling constraint.

Recall that we are only considering the *symmetric* action profiles in population games $\mathcal{G}(\mathcal{A})'$ and $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$. Given a symmetric action profile $\boldsymbol{x}^{\mathcal{I}} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$ in the auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, we can define a corresponding symmetric action profile of $\mathcal{G}(\mathcal{A})'$ such that all the nonatomic players in the populations regrouped in $\mathcal{N}_i$ play the same action $\boldsymbol{x}_i^{\mathcal{I}}$. (It is allowed that $\boldsymbol{x}_i^{\mathcal{I}}$ be not in $\mathcal{X}_n$. Recall that we can extend u_n to a neighborhood of $\mathcal{M}$ such that u_n is bounded on $\mathcal{M}$). Formally, we define the map $\psi : \mathbb{R}^{pT} \to \mathbb{R}^{NT}$:

$$\forall \boldsymbol{x}^{\mathcal{I}} \in \mathbb{R}^{pT},\ \psi(\boldsymbol{x}^{\mathcal{I}}) = (\boldsymbol{x}_n)_{n\in\mathcal{N}} \text{ where } \boldsymbol{x}_n = \boldsymbol{x}_i^{\mathcal{I}} \ ,\ \forall n \in \mathcal{N}_i \, .$$

Conversely, for a symmetric action profile $\boldsymbol{x}$ in $\mathcal{G}(\mathcal{A})'$, we define a corresponding symmetric action profile in the auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$ by the following map $\overline{\psi} : \mathbb{R}^{NT} \to \mathbb{R}^{pT}$:

$$\forall \boldsymbol{x} \in \mathbb{R}^{NT},\ \overline{\psi}(\boldsymbol{x}) = (\boldsymbol{x}_i^{\mathcal{I}})_{i\in\mathcal{I}} \text{ where } \boldsymbol{x}_i^{\mathcal{I}} = \tfrac{1}{p_i} \textstyle\sum_{n\in\mathcal{N}_i} \boldsymbol{x}_n.$$

Theorem 5.2 below is the main result of this subsection. It gives an upper bound on the distance between the SVWE of the population game $\mathcal{G}(\mathcal{A})'$, which has the same dimension as the original atomic game $\mathcal{G}(\mathcal{A})$, and that of an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, which has a reduced dimension.

Theorem 5.2 (SVWE of $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$ is close to SVWE of $\mathcal{G}(\mathcal{A})'$)**.** *Under Assumptions 5.1 and 5.2, in an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, $\overline{\delta}$ and $\overline{d}$ are defined by Equations* (5.12) *and* (5.13)*, with $\overline{\delta} < \frac{\rho}{2}$. Let $\boldsymbol{x}$ be a SVWE of $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, and $\boldsymbol{x}^*$ a SVWE of $\mathcal{G}(\mathcal{A})'$. Then:*

i) if H is strongly monotone with modulus α, then both $\boldsymbol{x}$ and $\boldsymbol{x}^*$ are unique and

$$\|\psi(\boldsymbol{x}) - \boldsymbol{x}^*\| \leqslant \sqrt{N\tfrac{K(\overline{\delta},\overline{d})}{\alpha}}\,; \tag{5.14}$$

$$\textit{besides}\quad \tfrac{1}{N}\sum_n \|\psi_n(\boldsymbol{x}) - \boldsymbol{x}_n^*\| \leqslant \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\alpha}} \quad \textit{and} \quad \left\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\right\| \leqslant \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\alpha}}; \tag{5.15}$$

ii) if H is aggregatively strongly monotone with modulus β, then both $\overline{\boldsymbol{X}} = \frac{1}{N}\sum_{i\in\mathcal{I}} \boldsymbol{x}_i$ and $\overline{\boldsymbol{X}}^* = \frac{1}{N}\sum_{n\in\mathcal{N}} \boldsymbol{x}_n^*$ are unique, and

$$\left\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\right\| \leqslant \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\beta}}\,, \tag{5.16}$$

where $K(\overline{\delta},\overline{d}) \underset{\overline{\delta},\overline{d}\to 0}{\longrightarrow} 0$ is a quantity defined as:

$$K(\overline{\delta},\overline{d}) \overset{\text{def}}{=} 2TR\big(3\tfrac{L_{\mathbf{f}}}{\rho}\overline{\delta} + \overline{d}\big) \quad \textit{with } L_{\mathbf{f}} \overset{\text{def}}{=} B_{\boldsymbol{c}} + \max_{n\in\mathcal{N}} B_{\mathbf{u_n}}\,. \tag{5.17}$$

Proof. See App. 5.E. □

Remark 5.2. *One can observe in Theorem 5.2 that the bound* (5.14) *given on the* individual profiles *diverges with the number of players N. This is a consequence of the fact that individual errors, $\|\boldsymbol{x}_n - \boldsymbol{x}_i^*\|$ for each n within a population $\mathcal{N}_i$, may accumulate, which is captured by the euclidean norm, as $\|\mathbb{1}_N\| = \sqrt{N}$.*

We have pointed out that the approximation error depends on how the populations are clustered according to $\mathcal{I}$, and is related to the heterogeneity of players in $\mathcal{N}$ rather than their number. In particular, in the case of Example 5.3, Theorem 5.2 states that the (aggregate) SVWE of the auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$ is exactly equal to the (aggregate) SVWE of the large game $\mathcal{G}(\mathcal{A})'$.

A direct corollary of Theorem 5.2-(1) is that two populations in $\mathcal{G}(\mathcal{A})'$ with similar characteristics have similar behavior at a SVWE there. This is the extension of Proposition 5.4 in the presence of coupling constraints.

Corollary 5.2. *Let $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ be a SVWE of game $\mathcal{G}(\mathcal{A})'$. Under Assumption 5.1, for two populations n and m in $\mathcal{N}$, if $d_H(\mathcal{X}_n, \mathcal{X}_m) \leqslant \delta$, $\sup_{\boldsymbol{x}\in\mathcal{M}/p} d_H(\partial(-u_m)(\boldsymbol{x}), \partial(-u_n)(\boldsymbol{x})) \leqslant d$, and u_n (resp. u_m) is α_n- (resp. α_m-)strongly concave, then*

$$\|\boldsymbol{x}_n^* - \boldsymbol{x}_m^*\| \leqslant \left(\tfrac{1}{\sqrt{\alpha_n}} + \tfrac{1}{\sqrt{\alpha_m}}\right)\, K(\delta, d)^{1/2}\,.$$

5.3.3 Combining the two steps to approximate a VNE of $\mathcal{G}(\mathcal{A})$

The following theorem is the main result of the chapter, which is immediately obtained as the combination of the two steps of approximation given in Theorem 5.1 and in Theorem 5.2 in the computation of a VNE of the original game $\mathcal{G}(\mathcal{A})$.

Corollary 5.3 (SVWE of $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$ is close to VNEs of $\mathcal{G}(\mathcal{A})$)**.** *Under Assumptions 5.1 and 5.2, in an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, $\overline{\delta}$ and $\overline{d}$ are defined by Equations* (6.12) *and* (6.13)*, with $\overline{\delta} < \frac{\rho}{2}$. Let $\boldsymbol{x}^*$ be a SVWE of $\tilde{\mathcal{G}}^{\mathcal{I}}(\mathcal{A})$, $\hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ be a VNE of $\mathcal{G}(\mathcal{A})$, $\overline{\boldsymbol{X}}^* = \frac{1}{N}\sum_{i\in\mathcal{I}} p_i \boldsymbol{x}_i^*$, $\overline{\boldsymbol{X}}^{\wedge} = \frac{1}{N}\sum_{n\in\mathcal{N}} \hat{\boldsymbol{x}}_n$, and $K(\overline{\delta},\overline{d})$ the constant given by* (5.17)*.*

i) if u_n is α_n-strongly concave for each $n \in \mathcal{N}$, with $\alpha \overset{\text{def}}{=} \min_n \alpha_n$, then $\boldsymbol{x}^*$ is unique and

$$\|\psi(\boldsymbol{x}^*) - \hat{\boldsymbol{x}}\| \leqslant \tfrac{RC}{\alpha}\sqrt{\tfrac{T}{N}} + \sqrt{N\tfrac{K(\overline{\delta},\overline{d})}{\alpha}}$$
$$\tfrac{1}{N}\sum_n \|\psi_n(\boldsymbol{x}^*) - \hat{\boldsymbol{x}}_n\| \leqslant \tfrac{RC}{\alpha}\tfrac{\sqrt{T}}{N} + \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\alpha}},$$
$$\left\|\overline{\boldsymbol{X}}^{\wedge} - \overline{\boldsymbol{X}}^*\right\| \leqslant \tfrac{RC}{\alpha}\tfrac{\sqrt{T}}{N} + \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\alpha}};$$

ii) if c_t is β_t-strictly increasing for each $t \in \mathcal{T}$, with $\beta \overset{\text{def}}{=} \min_t \beta_t$, then $\mathbf{X}^{\wedge}$ is unique and

$$\left\|\overline{\boldsymbol{X}}^{\wedge} - \overline{\boldsymbol{X}}^*\right\| \leqslant R\sqrt{\tfrac{2TC}{N\beta}} + \sqrt{\tfrac{K(\overline{\delta},\overline{d})}{\beta}}.$$

Given the large game $\mathcal{G}(\mathcal{A})$ and a certain $p \in \mathbb{N}^*$, Corollary 5.3 suggests that we should find the auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}$ with $\mathcal{I} = \{1, \dots, p\}$ that minimizes $K(\overline{\delta}, \overline{d})$ in order to have the best possible approximation of the equilibria. This would correspond to a "clustering problem" given as follows:

$$\min_{(\mathcal{N}_i)_i \in \mathcal{P}^p(\mathcal{N})} \min_{(\mathcal{X}_i)_i} \min_{(u_i)_i} K(\overline{\delta}, \overline{d}), \tag{5.18}$$

where $\mathcal{P}^p(\mathcal{N})$ denotes the set of all partitions of $\mathcal{N}$ of cardinal p, while $(\mathcal{X}_i)_i$ and $(u_i)_i$ are chosen according to Assumption 5.2.

The value of the optimal solutions of problem (5.18), and thus of the quality of the approximation in Corollary 5.3, depends on the homogeneity of the N players in $\mathcal{N}$ in terms of action sets and utility functions. The "ideal" case is given in Example 5.3 where $\mathcal{N}$ is composed of a small number p of homogeneous populations and thus $K(\overline{\delta}, \overline{d}) = 0$.

In general, solving (5.18) is a hard problem in itself. It is indeed a generalization of the k-means clustering problem [Llo82] (with $k = p$ and considering a function of Hausdorff distances), which is itself NP-hard [GJW82]. In Section 5.4, we illustrate how we use directly the k-means algorithm to compute efficiently an approximate solution $(\mathcal{N}_i, \mathcal{X}_i, u_i)_{i \in \mathcal{I}}$ in the parametric case.

Finally, the number p in the definition of the auxiliary game should be chosen a priori as a trade-off between the minimization of $K(\overline{\delta}, \overline{d})$ and a sufficient minimization of the dimension. Indeed, with $\mathcal{I} = \mathcal{N}$, $\mathcal{X}_n = \mathcal{X}_i$ and $u_i = u_n$, we get $\overline{d} = \overline{\delta} = 0$. However, the aim of Corollary 5.3 is to find an auxiliary game $\tilde{\mathcal{G}}^{\mathcal{I}}$ with $p \ll N$ so that the dimension of the GVIs characterizing the equilibria (and thus the time needed to compute their solutions) is significantly reduced, while ensuring a relatively small error, measured by $\overline{d}$ and $\overline{\delta}$.

5.4 Application to Demand Response and Electricity Flexibilities

Demand response (DR) [IA09] refers to a set of techniques to influence, control or optimize the electric consumption of agents in order to provide some services to the grid, e.g. reduce production costs and CO_2 emissions or avoid congestion [Jac+19c]. The increasing number of electric vehicles (EV) offers a new source of flexibility in the optimization of the production and demand, as electric vehicles require a huge amount of energy and present sufficiently flexible charging schemes (the EV can be charged whenever it is parked). Because of the privacy of each consumer or EV owner's information and the decentralized aspects of the DR problem, many relevant works adopt a game theoretical approach by considering consumers as players minimizing a cost function and an utility [Saa+12].

In this section, we consider the consumption associated to electric vehicle charging on a set of 24-hour time-periods $\mathcal{T} = \{1, \dots, T\}$, with $T = 24$, indexing the hours from 10 PM to 9PM the day after (including the night time periods where EVs are usually parked at home).

5.4.1 Price functions: block rates energy prices

As in the framework described in [Jac+19c], we consider a centralized entity, called the aggregator, who manages the aggregate flexible consumption. The aggregator interacts with the electricity market and energy producers, with his own objectives such as minimizing his cost or achieving a target aggregate demand profile.

The aggregator imposes electricity prices on each time-period. We consider that prices take the specific form of inclining block-rates tariffs (IBR tariffs, [WP17]), i.e. a piece-wise affine function $c(\cdot)$ which depends on the aggregate demand at time period t, that is, $X_t = \sum_{n\in\mathcal{N}} x_{n,t}$, and is defined as follows:

$$\begin{aligned} c(X) &= 1 + 0.1X \text{ if } X \leqslant 500 \\ c(X) &= -49 + 0.2X \text{ if } 500 \leqslant X \leqslant 1000 \\ c(X) &= -349 + 0.5X \text{ if } 1000 \leqslant X\,. \end{aligned} \tag{5.19}$$

This function c is continuous and convex. Those price functions are transmitted by the aggregator to each consumer or EV owner. Thus, each consumer n minimizes, with respect to her decision variable $\boldsymbol{x}_n$, a cost function of the form (5.1) with an energy cost determined by (5.19) and a utility function u_n defined below. An equilibrium gives a stable situation where each consumer minimizes her objective and has no interest to deviate from her current consumption profile.

5.4.2 Consumers' constraints and parameters

We simulate the consumption of $N = 2000$ consumers who have demand constraints of the form:

$$\mathcal{X}_n \stackrel{\text{def}}{=} \{\boldsymbol{x}_n \in \mathbb{R}_+^T : \textstyle\sum_t x_{n,t} = E_n \text{ and } \underline{x}_{n,t} \leqslant x_{n,t} \leqslant \overline{x}_{n,t}\} \tag{5.20}$$

where E_n is the total energy needed by n, and $\underline{x}_{n,t}, \overline{x}_{n,t}$ the (physical) bounds on the power allowed to her at time t. The utility functions have the form

$$u_n(\boldsymbol{x}_n) \stackrel{\text{def}}{=} -\omega_n \left\|\boldsymbol{x}_n - \boldsymbol{y}_n\right\|^2 .$$

The parameters are chosen as follows:

- E_n is drawn uniformly between 1 and 30 kWh, which corresponds to a typical charge of a residential electric vehicle.

- $\underline{\boldsymbol{x}}_n, \overline{\boldsymbol{x}}_n$: First, we generate, in two steps, a continual set of charging time-periods $\mathcal{T}_n = \{h_n - \frac{\tau_n}{2}, \ldots, h_n + \frac{\tau_n}{2}\}$;
  - the duration τ_n is uniformly drawn from $\{4, \ldots, T\}$;
  - h_n is then uniformly drawn from $\{1 + \frac{\tau_n}{2}, \ldots, T - \frac{\tau_n}{2}\}$.

Next, for $t \notin \mathcal{T}_n$, let $\underline{x}_{n,t} = \overline{x}_{n,t} = 0$.
Finally, for $t \in \mathcal{T}_n$, $\underline{x}_{n,t}$ (resp. $\overline{x}_{n,t}$) is drawn uniformly from $[0, \frac{E_n}{\tau_n}]$ (resp. $[\frac{E_n}{\tau_n}, E_n]$).

- ω_n is drawn uniformly from $[1, 10]$.

- $\boldsymbol{y}_{n,t}$ is taken equal to $\overline{x}_{n,t}$ on the first time periods of $\mathcal{T}_n$ (first available time periods) until reaching E_n (which corresponds to a profile "the sooner the better" or "plug and charge").

5.4.3 Coupling constraints on capacities and limited variations

We consider the following *coupling* constraints on the aggregate demand $\boldsymbol{X}$ which are often encountered in energy applications:

$$-50 \leqslant X_T - X_1 \leqslant 50 \tag{5.21}$$

$$X_t \leqslant 1400, \quad \forall t \in \mathcal{T} \tag{5.22}$$

Here, constraint (5.21) imposes that the demand X_T at the very end of the time horizon is relatively close to the first aggregate X_1, so that the demand response profiles computed for the finite time set $\mathcal{T}$ can be applied on a day-to-day, periodical basis.

Constraint (5.22) is a capacity constraint, induced by the maximal capacity of the electrical lines or by the generation capacities of electricity producers. These linear coupling constraints can be written in the closed form:

$$\boldsymbol{A}\boldsymbol{X} \leqslant \boldsymbol{b}\,, \tag{5.23}$$

where $\boldsymbol{A}$ is a real matrix of size $(T+2) \times T$ and $\boldsymbol{b} \in \mathbb{R}^{T+2}$.

5.4.4 Computing populations with k-means

Since N is very large, determining an exact VNE is computationally demanding. Thus, we apply the clustering procedure described in Section 5.3.2 to regroup the players.

We use the k-means algorithm [Llo82], where "k"$=\ p$ is the number of populations (groups) to replace the large set of N players. For each player $n \in \mathcal{N}$, we define her parametric description vector:

$$\boldsymbol{v}_n = [\omega_n, \boldsymbol{y}_n, E_n, \underline{\boldsymbol{x}}_n, \overline{\boldsymbol{x}}_n] \in \mathbb{R}^{3T+2}\,. \tag{5.24}$$

Then, the k-means algorithm finds an approximate solution of finding a partition $(S_i)_{1\leqslant i\leqslant p}$ of $\mathcal{N}$ into p clusters. The algorithm solves the combinatorial minimization problem:

$$\min_{S_1,\dots,S_p} \sum_{1\leqslant i\leqslant p} \sum_{\boldsymbol{v}\in S_i} \left\|\mathbb{E}_{S_i}(\boldsymbol{v}) - \boldsymbol{v}\right\|^2 = \min_{S_1,\dots,S_p} \sum_{1\leqslant i\leqslant p} |S_i| \mathbb{V}\mathrm{ar}(S_i),$$

where $\mathbb{E}_{S_i}(\boldsymbol{v}) = \frac{1}{|S_i|}\sum_{n\in S_i} \boldsymbol{v}_n$ denotes the average value of $\boldsymbol{v}$ over the set S_i. These average values are taken to be w_i, $\boldsymbol{y}_i$, E_i, $\underline{\boldsymbol{x}}_i$ and $\overline{\boldsymbol{x}}_i$.

The simulations are run with different population numbers, with $p \ll N$ chosen among $\{5, 10, 20, 50, 100\}$.

Since the k-means algorithm minimizes the squared distance of the average vector of parameters in S_i to the vectors of parameters of the points in S_i, the clustered populations obtained can be sub-optimal in terms of $K(\overline{\delta}, \overline{d})$. As explained above, choosing the optimal populations $\mathcal{I}$, as formulated in problem (5.18), is a complex problem in itself which deserves further research. Our example shows that the k-means algorithm gives a practical and efficient way to compute a heuristic solution in the case where u_n and $\mathcal{X}_n$ are parameterized.

5.4.5 Computation methods

We compute a VNE (Definition 5.2) with the original set of N players and the approximating SVWE (Definition 5.3) as solutions of the associated GVI (5.2).

However, to our limited knowledge and outside of the *strongly monotone* case [Coh88] and the *maximal monotone* case [AT09], there is no convergence result on a simple algorithm that would find the solution of a monotone GVI.

For our example, we employ a projection descent algorithm, as exposed in [Pac+18, Algo. 2] in the differentiable case, adapted in Algorithm 5.1 to the subdifferentiable case. In particular, the fixed step τ used in [Pac+18] is replaced by a variable step $\tau^{(k)} = 1/k$, in the spirit of subgradient algorithms (e.g. [Coh88; AT09]).

The principle of this algorithm is to relax the coupling constraint (5.23) and to consider the Lagrangian multipliers $\boldsymbol{\lambda} \in \mathbb{R}_+^{T+2}$ associated to these constraints as extra variables, and to consider the *extended* operator $T : \boldsymbol{\mathcal{X}}^{\mathcal{I}} \times \mathbb{R}_+^{T+2} \rightrightarrows \mathbb{R}^{IT} \times \mathbb{R}^{T+2}$ defined as:

$$T(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{pmatrix} \left(\partial_1 f_i(\boldsymbol{x}_i, \boldsymbol{X}) + \boldsymbol{\lambda}^\top \boldsymbol{A}\right)_{i\in\mathcal{I}} \\ -(\boldsymbol{A}\boldsymbol{X} - \boldsymbol{b}) \end{pmatrix}$$

on which we can apply a projected subgradient algorithm. The advantage is that we can perform the projections on the sets $(\mathcal{X}_n)_{n\in\mathcal{N}}$ and on $\mathbb{R}_+^{T+2}$ as shown below:

Algorithm 5.1 Projected Descent Algorithm

Require: $\boldsymbol{x}^{(0)}, \boldsymbol{\lambda}^{(0)}$, stopping criterion
1: $k \leftarrow 0$
2: **while** stopping criterion not true **do**
3: **for** $i = 1$ to p **do**
4: take $g_i^{(k)} \in \partial_1 f_i(\boldsymbol{x}_i^{(k)}, \boldsymbol{X}^{(k)})$
5: $\boldsymbol{x}_i^{(k+1)} \leftarrow \mathrm{P}_{\mathcal{X}_i}\left(\boldsymbol{x}_i^{(k)} - \tau^{(k)}(g_i^{(k)} + \boldsymbol{\lambda}^{(k)\top}\boldsymbol{A})\right)$
6: **done**
7: $\boldsymbol{\lambda}^{(k+1)} \leftarrow \left(\boldsymbol{\lambda}^{(k)} - \tau^{(k)}(b - 2\boldsymbol{A}\boldsymbol{X}^{(k+1)} + \boldsymbol{A}\boldsymbol{X}^{(k)})\right)^+$
8: $k \leftarrow k + 1$
9: **done**

The stopping criterion that we adopt here is the distance between two iterates: the algorithm stops when $\|(\lambda^{(k+1)}, \boldsymbol{x}^{(k+1)}) - (\lambda^{(k)}, \boldsymbol{x}^{(k)})\|_2 \leqslant 10^{-3}$. Although the algorithm converges for this criterion in practice in our numerical experiments, the general convergence of Algorithm 5.1 is not proven theoretically. Proving the convergence of a projected subgradient algorithm for a general game, under Assumption 5.1 and assuming $\boldsymbol{x} \mapsto \left(\partial_1 f_i(\boldsymbol{x}_i, \boldsymbol{X})\right)_{i\in\mathcal{I}}$ to be (strongly) monotone, is out of the scope of this work, but would constitute an interesting path for further research.

Due to the form of the strategy sets considered (5.20), the projection steps (Line 5) can be computed efficiently and exactly in $\mathcal{O}(T)$ with the Brucker algorithm [Bru84]. However, if we consider more general strategy sets (arbitrary convex sets), this projection step can be costly: in that case, other algorithms such as [Fuk86] would be more efficient.

5.4.6 A trade-off between precision and computation time

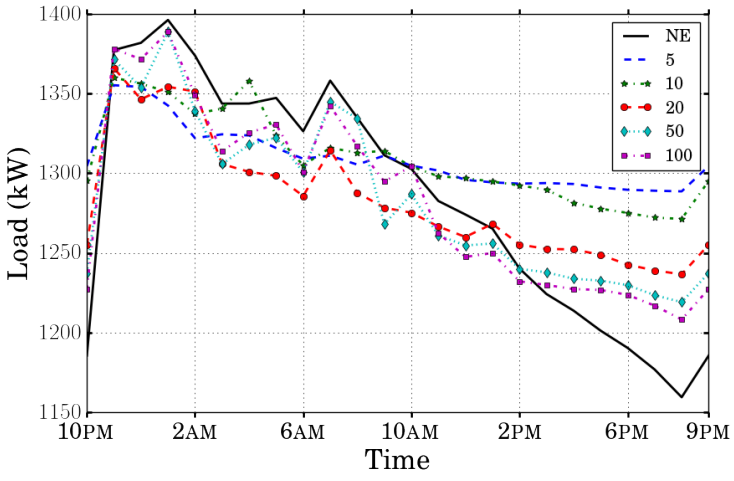


FIGURE 5.1: Convergence of the aggregate SVWE profile of auxiliary games to a VNE profile of the original game.

Simulations were run using Python on a single core Intel Xeon @3.4Ghz and 16GB of RAM.

Figure 5.1 shows the different aggregate SVWE profiles $\boldsymbol{X}^{*\mathcal{I}}$ obtained for sets $\mathcal{I}$ of different sizes, as well as a VNE of the original game for comparison. Thanks to the specific form

of the strategy sets (5.20)—which enables a fast projection—we are able to compute a VNE (and the corresponding aggregate profile $\boldsymbol{X}^{\wedge}$) of the original game with $N = 2000$ players.

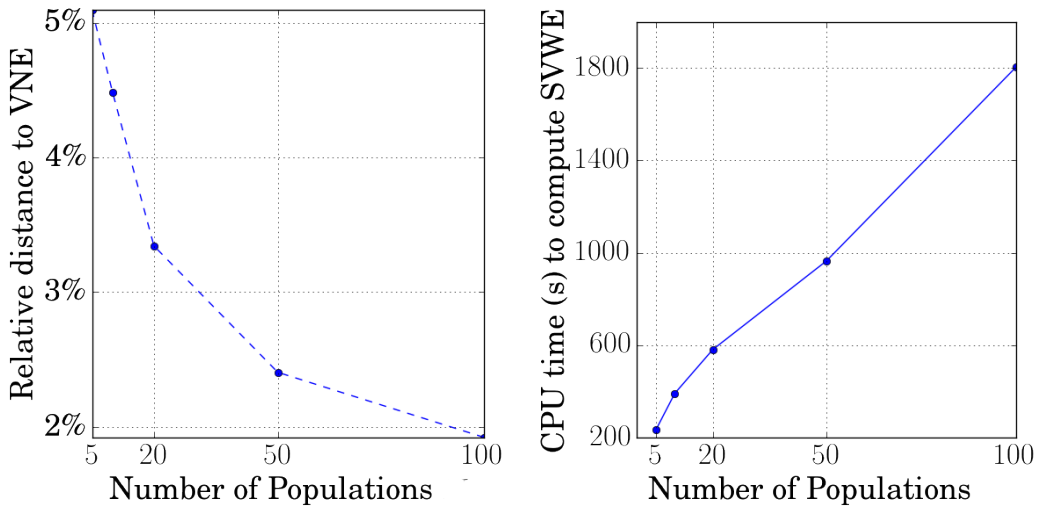


FIGURE 5.2: (a) Relative error to actual VNE ; (b) Time to compute SVWE. *The time needed to compute SVWE (b) increases roughly linearly with N, at a faster rate than the error on the approximation of the VNE decreases (a).*

Figure 5.2.a and Figure 5.2.b show the two main metrics to consider to choose a relevant number of populations N: the precision of the SVWE approximating the equilibrium (measured by the relative distance of the aggregate SVWE profile $\boldsymbol{X}^*$ to the aggregate VNE profile $\boldsymbol{X}^{\wedge}$ computed along, $\|\boldsymbol{X}^* - \boldsymbol{X}^{\wedge}\| \,/\, \|\boldsymbol{X}^*\|$), and the CPU time needed to compute the SVWE.

First notice on Figure 5.2.a that the distance between the aggregate equilibrium profile and its estimation decreases with N at a sublinear rate. This is partially explained in light of Theorem 5.2 and in addition with the following remarks:

- the Hausdorff distance of two parameterized polyhedral sets is Lipschitz continuous w.r.t their parameter vectors (see Theorem 5.3 in App. 5.F), which ensures that there is $K > 0$ s.t. for all n:

$$\delta_i^{\mathcal{I}} = \max_{n \in \mathcal{N}_i} d_H\left(\mathcal{X}_n, \mathcal{X}_i^{\mathcal{I}}\right) \leqslant K \max_{n \in \mathcal{N}_i} \left\| \begin{pmatrix} E_i \\ \underline{x}_i \\ \overline{x}_i \end{pmatrix} - \begin{pmatrix} E_n \\ \underline{x}_n \\ \overline{x}_i \end{pmatrix} \right\| ;$$

- similarly, as subgradients of utility functions are reduced to a point, one has, for all n:

$$\begin{aligned} d_i &= \max_{n \in} \max_{\boldsymbol{x} \in \mathcal{M}} 2 \left\| \omega_i(\boldsymbol{x} - \boldsymbol{y}_i) - \omega_n(\boldsymbol{x} - \boldsymbol{y}_n) \right\| \\ &= \mathcal{O}\left(\max_{n \in \mathcal{N}_i} |\omega_i - \omega_n| + \|\boldsymbol{y}_i - \boldsymbol{y}_n\| \right) . \end{aligned}$$

Figure 5.2.b shows the CPU time needed to compute the SVWE with our stopping criterion. Computing a solution of the clustering problem with the k-means algorithm takes, for each value of N, less than ten seconds. This time is negligible in comparison to the time needed for convergence of Algorithm 5.1.

As a reference time, to compute a VNE of the original game (observed on Figure 5.1) with the same stopping criterion and the same CPU configuration, 3 hours and 26 minuteswe were needed, more than six times as long as the CPU time to compute the SVWE with one hundred populations.

On this figure, we see that the CPU time evolves linearly with the number of populations p. This is explained by the structure of Algorithm 5.1, as each iteration k is executed in a time proportional to p due to the **for** loop.

Last, one observes from Figure 5.2.a that, in our example, the error between the aggregate demand profile at equilibrium and its approximation is between 2% and 5%, which remains significant. However, as pointed out in Section 5.3, the quality of the approximation depends on the heterogeneity of the set of players $\mathcal{N}$. In the example of this section, as the parameters are drawn uniformly (see Section 5.4.2), the set of players $\mathcal{N}$ presents a large variance so that it is a "worst" case (as opposed to the case of Example 5.3 which offers an optimal situation).

5.5 Conclusion

This chapter shows that equilibria in splittable congestion games with a very large number of atomic players can be approximately computed with a Wardrop equilibrium of an auxiliary population game of smaller dimension. Our results give explicit bounds on the distance of this approximating equilibrium to the equilibria of the original large game. These theoretical results can be used in practice to solve by an iterative method, complex nonconvex bilevel programs where the lower level is the equilibrium of a large congestion game, for instance, to optimize tariffs or tolls for the operator of a network. A detailed analysis of such a procedure would constitute an interesting extension of the present work.

Appendix

5.A Proof of Lemma 5.1: Expressions of Subgradients

Recall that $\hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n}) \stackrel{\text{def}}{=} f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n} + \frac{1}{N}\boldsymbol{x}_n)$. According to [BC11, Proposition 16.6]:

$$\partial \hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n}) \subset \{(I_T, I_T)\mathbf{g} : \mathbf{g} \in \partial\psi_n(\boldsymbol{x}_n)\},$$

with I_T the identity matrix of size T, and where $\partial\psi_n(\boldsymbol{x}_n)$ is the subdifferential of $\psi_n(\cdot) \stackrel{\text{def}}{=} f_n(\cdot, \overline{\boldsymbol{X}}_{-n} + \frac{1}{N}\cdot)$, at $\boldsymbol{x}_n$. On the other hand, according to [BC11, Proposition 16.7], $\partial\psi_n(\boldsymbol{x}_n)$ is a subset of:

$$\Big\{(\mathbf{g}_{n,1}, \mathbf{g}_{n,2}) : \mathbf{g}_{n,1} \in \partial_1 f_n(\boldsymbol{w}_n, \overline{\boldsymbol{Y}})|_{\boldsymbol{w}_n = \boldsymbol{x}_n, \overline{\boldsymbol{Y}} = \overline{\boldsymbol{X}}_{-n} + \frac{1}{N}\boldsymbol{x}_n}, \\ \mathbf{g}_{n,2} \in \partial_2 f_n(\boldsymbol{w}_n, \overline{\boldsymbol{Y}})|_{\boldsymbol{w}_n = \boldsymbol{x}_n, \overline{\boldsymbol{Y}} = \overline{\boldsymbol{X}}_{-n} + \frac{1}{N}\boldsymbol{x}_n}\Big\}.$$

Therefore, $\partial_1 \hat{f}_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}_{-n})$ is a subset of:

$$\begin{aligned}&\{\boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}'_{n,1} + \mathbf{g}_{n,2} : \mathbf{g}'_{n,1} \in \partial(-u_n)(\boldsymbol{x}_n), \mathbf{g}_{n,2} \in \partial_2 f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}})\} \\ &= \{\boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}'_{n,1} + (\tfrac{1}{N} a_{n,t}\, x_{n,t})_t : \mathbf{g}'_{n,1} \in \partial(-u_n)(\boldsymbol{x}_n),\, a_{n,t} \in \partial c_t(\overline{X}_t)\ \forall t \in T\}.\end{aligned}$$

By the definition of subdifferential, it is easy to show that:

$$\big\{\boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}'_{n,1} + (\tfrac{1}{N} a_{n,t} x_{n,t})_t : \mathbf{g}'_{n,1} \in \partial(-u_n)(\boldsymbol{x}_n),\, a_{n,t} \in \partial c_t(\overline{X}_t)\ \forall t \in T\big\} \subset \partial_1 \hat{f}_n(\boldsymbol{x}_n, \boldsymbol{X}_{-n}).$$

The proof for $\partial_1 f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}})$ is similar.

5.B Proof of Proposition 5.3: monotonicity of H

i) Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$ and $\overline{\boldsymbol{X}} = \frac{1}{N}\sum_n \boldsymbol{x}_n$, $\overline{\boldsymbol{Y}} = \frac{1}{N}\sum_n \boldsymbol{y}_n$. Recall that

$$\begin{aligned}\partial_1 f_n(\boldsymbol{x}_n, \overline{\boldsymbol{X}}) &= \{\boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}_n : \mathbf{g} \in \partial(-u_n)(\boldsymbol{x}_n)\}, \\ \partial_1 f_n(\boldsymbol{y}_n, \overline{\boldsymbol{Y}}) &= \{\boldsymbol{c}(\overline{\boldsymbol{Y}}) + \mathbf{h}_n : \mathbf{h} \in \partial(-u_n)(\boldsymbol{y}_n)\}.\end{aligned}$$

Let $\mathbf{g}_n \in \partial(-u_n)(\boldsymbol{x}_n)$ and $\mathbf{h}_n \in \partial(-u_n)(\boldsymbol{y}_n)$. One has $\langle \mathbf{g}_n - \mathbf{h}_n, \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant 0$ because u_n is concave so that $\partial(-u_n)$ is monotone on $\mathcal{X}_n$. Then we get:

$$\begin{aligned}&\textstyle\sum_n \langle (\boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{g}_n) - (\boldsymbol{c}(\overline{\boldsymbol{Y}}) - \mathbf{h}_n), \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \\ =&N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{Y}}), \overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}} \rangle + \textstyle\sum_n \langle \mathbf{g}_n(\boldsymbol{x}_n) - \mathbf{h}_n(\boldsymbol{y}_n), \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant 0\end{aligned}$$

because $\boldsymbol{c}$ is monotone. Hence H is monotone.
ii) By the definition of α_n-strong concavity:

$$\textstyle\sum_n \langle \mathbf{g}_n(\boldsymbol{x}_n) - \mathbf{h}_n(\boldsymbol{y}_n), \boldsymbol{x}_n - \boldsymbol{y}_n \rangle \geqslant \sum_n \alpha_n \|\boldsymbol{x}_n - \boldsymbol{y}_n\|^2 \geqslant \alpha \|\boldsymbol{x} - \boldsymbol{y}\|^2 ,$$

hence H is α-strongly monotone.
iii) By the definition of β_t-strictly increasing:

$$\begin{aligned}N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) - \boldsymbol{c}(\overline{\boldsymbol{Y}}), \overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}} \rangle &= N \textstyle\sum_{t \in \mathcal{T}} \langle c_t(\overline{X}_t) - c_t(\overline{Y}_t), \overline{X}_t - \overline{Y}_t \rangle \\ &\geqslant N \textstyle\sum_t \beta_t \|\overline{X}_t - \overline{Y}_t\|^2 \geqslant N\beta \|\overline{\boldsymbol{X}} - \overline{\boldsymbol{Y}}\|^2 ,\end{aligned}$$

hence H is β-aggregatively strongly monotone.

5.C Proof of Proposition 5.4: SWE profiles for similar players

Let $\mathbf{g}'_n(\boldsymbol{x}^*_n) \in \partial(-u_n)(\boldsymbol{x}^*_n)$ be s.t., for all $\boldsymbol{x}_n \in \mathcal{X}_n$, $\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{g}'_n(\boldsymbol{x}^*_n), \boldsymbol{x}^*_n - \boldsymbol{x}_n \rangle \leqslant 0$. Let $\mathbf{h}'_n(\boldsymbol{x}^*_m) \in \partial(-u_n)(\boldsymbol{x}^*_m)$ be such that $\|\mathbf{h}'_n(\boldsymbol{x}^*_m) - \mathbf{g}'_m(\boldsymbol{x}^*_m)\| \leqslant d$. Then, by the strong concavity of u_n:

$$\begin{aligned}\alpha_n \|\boldsymbol{x}^*_n - \boldsymbol{x}^*_m\|^2 &\leqslant \langle \mathbf{g}'_n(\boldsymbol{x}^*_n) - \mathbf{h}'_n(\boldsymbol{x}^*_m), \boldsymbol{x}^*_n - \boldsymbol{x}^*_m \rangle \\ &= \langle \mathbf{g}'_n(\boldsymbol{x}^*_n) - \mathbf{g}'_m(\boldsymbol{x}^*_m) + \mathbf{g}'_m(\boldsymbol{x}^*_m) - \mathbf{h}'_n(\boldsymbol{x}^*_m), \boldsymbol{x}^*_n - \boldsymbol{x}^*_m \rangle \\ &\leqslant \langle \mathbf{g}'_n(\boldsymbol{x}^*_n) - \mathbf{g}'_m(\boldsymbol{x}^*_m), \boldsymbol{x}^*_n - \boldsymbol{x}^*_m \rangle + 2dTR .\end{aligned}$$

Making use of the inequalities on $\mathbf{g}'_n(\boldsymbol{x}^*_n)$ and $\mathbf{g}'_m(\boldsymbol{x}^*_m)$, we obtain that this is equal to:

$$\begin{aligned}&\langle \mathbf{g}'_n(\boldsymbol{x}^*_n) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \boldsymbol{x}^*_n - \boldsymbol{x}^*_m \rangle + \langle \mathbf{g}'_m(\boldsymbol{x}^*_m) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \boldsymbol{x}^*_m - \boldsymbol{x}^*_n \rangle + 2dTR \\ =&\langle \mathbf{g}'_n(\boldsymbol{x}^*_n) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \boldsymbol{x}^*_n - \mathrm{P}_n(\boldsymbol{x}^*_m) + \mathrm{P}_n(\boldsymbol{x}^*_m) - \boldsymbol{x}^*_m \rangle \\ &\quad + \langle \mathbf{g}'_m(\boldsymbol{x}^*_m) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \boldsymbol{x}^*_m - \mathrm{P}_m(\boldsymbol{x}^*_n) + \mathrm{P}_m(\boldsymbol{x}^*_n) - \boldsymbol{x}^*_n \rangle + 2dTR \\ \leqslant&\langle \mathbf{g}'_n(\boldsymbol{x}^*_n) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \mathrm{P}_n(\boldsymbol{x}^*_m) - \boldsymbol{x}^*_m \rangle + \langle \mathbf{g}'_m(\boldsymbol{x}^*_m) + \boldsymbol{c}(\overline{\boldsymbol{X}}^*), \mathrm{P}_m(\boldsymbol{x}^*_n) - \boldsymbol{x}^*_n \rangle + 2dTR \\ \leqslant&(B_{\mathbf{u_n}} + B_{\mathbf{u_m}} + 2B_{\boldsymbol{c}})\delta + 2dTR .\end{aligned}$$

where P_n (resp. P_m) is the projector on $\mathcal{X}_n$ (resp. $\mathcal{X}_m$).

5.D Proof of Lemma 5.2: Existence of interior profile

Let $\check{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}$ be s.t. for all n, $d(\check{\boldsymbol{x}}_n, \mathrm{rbd}\ \mathcal{X}_n) = \max_{\boldsymbol{x} \in \mathcal{X}_n} d(\boldsymbol{x}, \mathrm{rbd}\ \mathcal{X}_n) \stackrel{\text{def}}{=} \eta_n$. Denote $\check{\overline{\boldsymbol{X}}} = \frac{1}{N}\sum_n \check{\boldsymbol{x}}_n$ and $\eta = \min_n \eta_n > 0$.

Let $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ and $\overline{\boldsymbol{Y}} = \frac{1}{N}\sum_n \boldsymbol{y}_n$ be such that

$$d(\overline{\boldsymbol{Y}}, \mathrm{rbd}\ \mathcal{A}) = \max_{\overline{\boldsymbol{X}} \in \overline{\mathcal{X}} \cap \mathcal{A}} d(\overline{\boldsymbol{X}}, \mathrm{rbd}\ \mathcal{A}) .$$

Let us denote $t = d(\overline{\boldsymbol{Y}}, \mathrm{rbd}\ \mathcal{A})/3TR$ and let us define $\boldsymbol{z} = \boldsymbol{y} - t(\boldsymbol{y} - \check{\boldsymbol{x}}) \in \boldsymbol{\mathcal{X}}$ and $\overline{\boldsymbol{Z}} = \frac{1}{N}\sum_n \boldsymbol{z}_n$. We obtain:

$$\|\overline{\boldsymbol{Y}} - \overline{\boldsymbol{Z}}\| = t\|\overline{\boldsymbol{Y}} - \check{\overline{\boldsymbol{X}}}\| \leqslant t2TR \leqslant \tfrac{2}{3} d(\overline{\boldsymbol{Y}}, \mathrm{rbd}\ \mathcal{A}) ,$$

hence $\overline{\boldsymbol{Z}} \in \overline{\mathcal{X}} \cap \text{ri } \mathcal{A}$, where ri means the relative interior. Besides, for any n, $\boldsymbol{z}_n = \boldsymbol{y}_n - t(\boldsymbol{y}_n - \check{\boldsymbol{x}}_n)$. Since $d(\check{\boldsymbol{x}}_n, \text{rbd } \mathcal{X}_n) \geqslant \eta$, $\boldsymbol{y}_n \in \mathcal{X}_n$, and $\mathcal{X}_n$ is convex, we have:

$$d(\boldsymbol{z}_n, \text{rbd } \mathcal{X}_n) \geqslant \eta t = \tfrac{\eta}{3TR} d(\overline{\boldsymbol{Y}}, \text{rbd } \mathcal{A}) \,.$$

Finally, define $\rho \stackrel{\text{def}}{=} \frac{\eta}{3TR} d(\overline{\boldsymbol{Y}}, \text{rbd } \mathcal{A})$ to obtain the desired result.

5.E Proof of Theorem 5.2: approximation of SVWE

Lemma 5.3.

i) For each $i \in \mathcal{I}$ and $\boldsymbol{x} \in \mathcal{X}_i$, if $d(\boldsymbol{x}, \text{rbd } \mathcal{X}_i) > \delta_i$, then $\boldsymbol{x} \in \mathcal{X}_n$ for each $n \in \mathcal{N}_i$.

ii) For each $i \in \mathcal{I}$, $n \in \mathcal{N}_i$ and $\boldsymbol{x} \in \mathcal{X}_n$, if $d(\boldsymbol{x}, \text{rbd } \mathcal{X}_n) > \delta_i$, then $\boldsymbol{x} \in \mathcal{X}_i$.

Proof of Lemma 5.3. i) Suppose $\boldsymbol{x} \notin \mathcal{X}_n$. Let $\boldsymbol{y} \stackrel{\text{def}}{=} \text{P}_{\mathcal{X}_n}(\boldsymbol{x}) \neq \boldsymbol{x}$. As $\boldsymbol{y} \in \text{aff } \mathcal{X}_n \subset \text{aff } \mathcal{X}_i$, then $\boldsymbol{x} - \boldsymbol{y} \in \text{aff } \mathcal{X}_i$. Let $\boldsymbol{z} \stackrel{\text{def}}{=} \boldsymbol{x} + \delta_i \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|}$. Then, $\boldsymbol{z} \in \mathcal{X}_i$ because $\|\boldsymbol{z} - \boldsymbol{x}\| \leqslant \delta_i$. By the convexity of $\mathcal{X}_n$ and the definition of $\boldsymbol{y}$, we have $d(\boldsymbol{z}, \mathcal{X}_n) = d(\boldsymbol{x}, \mathcal{X}_n) + \delta_i > \delta_i$, contradicting the fact that $\delta_i \geqslant d_H(\mathcal{X}_n, \mathcal{X}_i)$. The proof of ii) is symmetric. □

Lemma 5.4. *Under Assumption 5.1, if $\overline{\delta} < \frac{\rho}{2}$, then:*

i) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$, there is $\boldsymbol{w} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ such that $\|\boldsymbol{w}_n - \psi_n(\boldsymbol{x})\| \leqslant 4TR\frac{\overline{\delta}}{\rho}$ for each $n \in \mathcal{N}$;

ii) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$, there is $\boldsymbol{w} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$ such that $\|\boldsymbol{w}_i - \overline{\psi}_i(\boldsymbol{x})\| \leqslant 2TRN_i\frac{\overline{\delta}}{\rho}$ for each $i \in \mathcal{I}$.

Proof of Lemma 5.4.
i) For $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$, define $\boldsymbol{w} \in \boldsymbol{\mathcal{X}}$ as follows: $\forall i \in \mathcal{I}$, $\forall n \in \mathcal{N}_i$, let $\boldsymbol{w}_n \stackrel{\text{def}}{=} \boldsymbol{x}_i + t(\boldsymbol{z}_n - \boldsymbol{x}_i)$ where $\boldsymbol{z}$ is defined in Lemma 5.2, with $t \stackrel{\text{def}}{=} 2\overline{\delta}/\rho < 1$. On the one hand:

$$\begin{aligned} & \forall i \in \mathcal{I},\ \forall n \in \mathcal{N}_i,\ d(\boldsymbol{z}_n, \text{rbd } \mathcal{X}_i) \geqslant \rho - \overline{\delta} \\ \text{implies} \quad & d(\boldsymbol{w}_n, \text{rbd } \mathcal{X}_i) \geqslant t(\rho - \overline{\delta}) > t\rho/2 = \overline{\delta} \,, \end{aligned}$$

because each point in the ball with radius $t(\rho - \overline{\delta})$ centered at $\boldsymbol{w}_n$ is on the segment linking $\boldsymbol{x}_i$ and some point in the ball with radius $\rho - \overline{\delta}$ centered at $\boldsymbol{z}_n$ which is contained in $\mathcal{X}_n$.

Thus, $\boldsymbol{w}_n \in \mathcal{X}_n$ $\forall n \in \mathcal{N}_i$ according to Lemma 5.3.i). On the other hand, the linear mapping $S : \mathbb{R}^{IT} \ni \boldsymbol{v} \mapsto \frac{1}{N}\sum_{n\in\mathcal{N}} \boldsymbol{v}_n$ maps the segment linking $\psi(\boldsymbol{x})$ and $\boldsymbol{z}$ in $\boldsymbol{\mathcal{X}}(\mathcal{A})$ to a segment linking $\overline{\boldsymbol{X}} = \frac{1}{N}\sum_i N_i \boldsymbol{x}_i$ and $\overline{\boldsymbol{Z}}$ in the convex $\mathcal{A}$. Hence, we get:

$$\tfrac{1}{N}\textstyle\sum_{n\in\mathcal{N}} \boldsymbol{w}_n = t\overline{\boldsymbol{Z}} + (1-t)\overline{\boldsymbol{X}} \in \mathcal{A}$$

as well. Therefore, $\boldsymbol{w} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$. Finally, $\|\boldsymbol{w}_n - \psi_n(\boldsymbol{x})\| = t\|\boldsymbol{z}_n - \psi_n(\boldsymbol{x})\| \leqslant t2TR = 4TR\frac{\overline{\delta}}{\rho}$.

ii) For $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$, let $\boldsymbol{y} \stackrel{\text{def}}{=} \boldsymbol{x} + t(\boldsymbol{z} - \boldsymbol{x})$ with $t \stackrel{\text{def}}{=} \frac{\overline{\delta}}{\rho}$. Then, by similar arguments as above, $d(\boldsymbol{y}_n, \text{rbd } \mathcal{X}_n) \geqslant \overline{\delta}$ hence $\boldsymbol{y}_n \in \mathcal{X}_i$ and $\overline{\psi}(\boldsymbol{y}) \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}$. Besides, from the convexity of $\mathcal{A}$, we have:

$$\tfrac{1}{N}\textstyle\sum_n \boldsymbol{y}_n = t\overline{\boldsymbol{Z}} + (1-t)(\frac{1}{N}\sum_n \boldsymbol{x}_n) \in \mathcal{A} \,.$$

Hence $\boldsymbol{w} \stackrel{\text{def}}{=} \overline{\psi}(\boldsymbol{y}) \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$. Finally, $\|\boldsymbol{w}_i - \psi_i(\boldsymbol{x})\| = t\|\sum_{n\in N_i}(\boldsymbol{z}_n - \boldsymbol{x}_n)\| \leqslant 2TRN_i\frac{\overline{\delta}}{\rho}$. □

Let $\boldsymbol{w} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ be s.t. $\forall n \in \mathcal{N}$, $\|\boldsymbol{w}_n - \psi_n(\boldsymbol{x})\| \leqslant 4TR\overline{\delta}/\rho$ (cf. Lemma 5.4). Since $\boldsymbol{x}^*$ is a SVWE in $\mathcal{G}(\mathcal{A})'$, for each $n \in \mathcal{N}$ there is $\mathbf{g}'_n \in \partial(-u_n)(\boldsymbol{x}^*_n)$ s.t.:

$$\textstyle\sum_n \langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{g}'_n,\ \boldsymbol{x}^*_n - \boldsymbol{w}_n\rangle \leqslant 0 \,.$$

Secondly, since $\boldsymbol{x}$ is a SVWE in $\tilde{\mathcal{G}}(\mathcal{A})$, then for each $i \in \mathcal{I}$ there is $\mathbf{h}'_i \in \partial(-u_i)(\boldsymbol{x}_i)$ s.t.:

$$\textstyle\sum_i N_i \langle \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{h}'_i, \boldsymbol{x}_i - \boldsymbol{y}_i\rangle \leqslant 0 \,,$$

for all $\boldsymbol{y} \in \mathcal{X}^{\mathcal{I}}(\mathcal{A})$. Thirdly, $\forall i, \forall n \in \mathcal{N}_i$, by the definition of d_n, there is $\boldsymbol{r}'_n \in \partial(-u_n)(\boldsymbol{x}_i)$ such that $\|\boldsymbol{r}'_n - \mathbf{h}'_i\| \leqslant d_n$.

The above results and $|\hat{x}_{i,t}| \leqslant R, \forall i, \forall t$ imply:

$$\begin{aligned} &N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) - \boldsymbol{c}(\overline{\boldsymbol{X}}),\ \overline{\boldsymbol{X}}^* - \overline{\boldsymbol{X}}\rangle + \textstyle\sum_n \langle \mathbf{g}'_n - \boldsymbol{r}'_n, \boldsymbol{x}^*_n - \boldsymbol{x}_i\rangle \\ &= N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) - \boldsymbol{c}(\overline{\boldsymbol{X}}), \overline{\boldsymbol{X}}^* - \overline{\boldsymbol{X}}\rangle + \textstyle\sum_{i,n\in\mathcal{N}_i}\langle \mathbf{g}'_n - \boldsymbol{r}'_n, \boldsymbol{x}^*_n - \boldsymbol{x}_i\rangle \end{aligned}$$

$$= \textstyle\sum_{i,n\in\mathcal{N}_i}\big[\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{g}'_n, \boldsymbol{x}^*_n - \boldsymbol{w}_n\rangle + \langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{g}'_n, \boldsymbol{w}_n - \boldsymbol{x}_i\rangle\big] + \textstyle\sum_{i,n\in\mathcal{N}_i}\big[\langle \boldsymbol{r}'_n - \mathbf{h}'_i, \boldsymbol{x}_i - \boldsymbol{x}^*_n\rangle + \langle \boldsymbol{c}(\boldsymbol{X}) + \mathbf{h}'_i, \boldsymbol{x}_i - \boldsymbol{x}^*_n\rangle\big]$$

$$\begin{aligned} &\leqslant 0 + \textstyle\sum_{i,n\in\mathcal{N}_i}\left\|\boldsymbol{c}(\overline{\boldsymbol{X}}^*) + \mathbf{g}'_n\right\| \|\boldsymbol{w}_n - \boldsymbol{x}_i\| + \textstyle\sum_{i,n\in\mathcal{N}_i}\|\boldsymbol{r}'_n - \mathbf{h}'_i\|\ \|\boldsymbol{x}_i - \boldsymbol{x}^*_n\| + J \\ &\leqslant L_{\mathbf{f}}\, 4TNR\tfrac{\overline{\delta}}{\rho} + 2TNR\overline{d} + J \end{aligned} \tag{5.25}$$

where $J \stackrel{\text{def}}{=} \sum_{i,n\in\mathcal{N}_i}\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{h}'_i,\ \boldsymbol{x}_i - \boldsymbol{x}^*_n\rangle$. Next, for the SVWE $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(\mathcal{A})$, let $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}^{\mathcal{I}}(\mathcal{A})$ be s.t. $\forall i$, $\|\boldsymbol{y}_i - \overline{\psi}_i(\boldsymbol{x}^*)\|_{\mathcal{I}} \leqslant 2TRN_i\overline{\delta}/\rho$ (cf. Lemma 5.4). Then we obtain:

$$\begin{aligned} J &= \textstyle\sum_{i\in\mathcal{I}}\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{h}'_i,\ \boldsymbol{x}_i - \overline{\psi}_i(\boldsymbol{x}^*)\rangle \\ &= \textstyle\sum_{i\in\mathcal{I}}\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{h}'_i,\ \boldsymbol{x}_i - \boldsymbol{y}_i\rangle \quad + \sum_{i\in\mathcal{I}}\langle \boldsymbol{c}(\overline{\boldsymbol{X}}) + \mathbf{h}'_i,\ \boldsymbol{y}_i - \overline{\psi}_i(\boldsymbol{x}^*)\rangle \\ &\leqslant 0 + \textstyle\sum_{i\in\mathcal{I}} L_{\mathbf{f}}\|\overline{\psi}_i(\boldsymbol{x}^*) - \boldsymbol{y}_i\| \\ &\leqslant L_{\mathbf{f}}\, 2\, TNR\tfrac{\overline{\delta}}{\rho}\,, \end{aligned} \tag{5.26}$$

Let us summarize by combining (5.25) and (5.26), to get:

$$N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) - \boldsymbol{c}(\overline{\boldsymbol{X}}), \overline{\boldsymbol{X}}^* - \overline{\boldsymbol{X}}\rangle + \textstyle\sum_n\langle \mathbf{g}'_n - \boldsymbol{r}'_n, \boldsymbol{x}^*_n - \boldsymbol{x}_i\rangle \ \leqslant 2TNR\left(3\tfrac{L_{\mathbf{f}}}{\rho}\overline{\delta} + \overline{d}\right)\,.$$

Hence, if H is strongly monotone with modulus α, then

$$\alpha\,\|\psi(\boldsymbol{x}) - \boldsymbol{x}^*\|^2 \leqslant \textstyle\sum_n\langle \mathbf{g}'_n - \boldsymbol{r}'_n, \boldsymbol{x}^*_n - \boldsymbol{x}_i\rangle \leqslant 2TNR\big(3\tfrac{L_{\mathbf{f}}}{\rho}\overline{\delta} + \overline{d}\big)\,;$$

and the other two inequalities can be proved as (5.8) and (5.9). If H is aggregatively strongly monotone with modulus β, then

$$N\beta\|\bar{\boldsymbol{X}} - \overline{\boldsymbol{X}}^*\|^2 \leqslant N\langle \boldsymbol{c}(\overline{\boldsymbol{X}}^*) - \boldsymbol{c}(\overline{\boldsymbol{X}}), \overline{\boldsymbol{X}}^* - \overline{\boldsymbol{X}}\rangle \leqslant 2TNR\big(3\tfrac{L_{\mathbf{f}}}{\rho}\overline{\delta} + \overline{d}\big)\,.$$

5.F Bounding the perturbation of parameterized polyhedra

Let us consider a parameterized family of *polytopes* $(\Lambda_{\boldsymbol{b}})_{\boldsymbol{b}\in\mathcal{B}}$ given by the generic form:

$$\Lambda_{\boldsymbol{b}} \triangleq \{\boldsymbol{x} \in \mathbb{R}^T \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\},$$

where $\boldsymbol{A}$ is a $p \times T$ real constant matrix, with $p \in \mathbb{N}^*, T \in \mathbb{N}^*$, and $\boldsymbol{b} \in \mathcal{B}$ is a p-dimensional real vector, and $\mathcal{B} \subset \mathbb{R}^p$ is (possibly) infinite arbitrary set.

In particular, $\Lambda_{\boldsymbol{b}}$ is a convex and compact subset of $\mathbb{R}^T$.

This situation corresponds for instance to the *heterogeneous* action sets of atomic or non-atomic players, decribed by the parameter $\boldsymbol{b}$, as given in the example of electricity consumers (6.22). In this example, $\boldsymbol{b}$ would represent the parameters of energy demand and lower and upper bounds on the energy power needed at each time period.

Let us make the following practical assumption:

Assumption 5.3. *The parameter vector $\boldsymbol{b}$ defining $\Lambda_{\boldsymbol{b}}$ lies in a compact set $\mathcal{B}$ such that Λ_θ is a nonempty polytope, that is:*

$$\forall \boldsymbol{b} \in \mathcal{B},\ \Lambda(\boldsymbol{b}) \stackrel{\text{def}}{=} \{\boldsymbol{x}\ :\ \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\} \neq \emptyset\,.$$

When parameter vectors $\boldsymbol{b}$ and $\boldsymbol{b}' \in \mathcal{B}$ are close to each other, we expect that the corresponding polytopes $\Lambda_{\boldsymbol{b}}, \Lambda_{\boldsymbol{b}'}$ not to differ too much. This appendix quantify and proves this idea using the Hausdorff distance d_H between sets, as stated in Theorem 5.3 given below.

Theorem 5.3. *The Hausdorff distance between polyhedra $\Lambda_{\boldsymbol{b}}$, $\boldsymbol{b} \in \mathcal{B}$, is linearly bounded: that is, there exists a constant $D > 0$ such that:*

$$\forall \boldsymbol{b}, \boldsymbol{b}' \in \mathcal{B},\ d_H(\Lambda_{\boldsymbol{b}}, \Lambda_{\boldsymbol{b}'}) \leqslant D\|\boldsymbol{b} - \boldsymbol{b}'\| .$$

Proof. The proof follows in several parts [Bat87], but we extend the result on the compact set $\mathcal{B}$, and drop the irredundance assumption made in [Bat87].

For each $\boldsymbol{b} \in \mathcal{B}$, let $\mathcal{V}(\boldsymbol{b})$ denotes the set of vertices of polyhedron $\Lambda_{\boldsymbol{b}}$. Under Assumption 5.3, $\mathcal{V}(\boldsymbol{b})$ is nonempty for any $\boldsymbol{b} \in \mathcal{B}$.

First, as $\Lambda_{\boldsymbol{b}}$ is a polyhedron, we have $\Lambda_{\boldsymbol{b}} = \mathrm{conv}(\mathcal{V}(\boldsymbol{b}))$ where $\mathrm{conv}(\mathcal{X})$ is the convex hull of a set $\mathcal{X}$. As the function $\boldsymbol{x} \mapsto d(\boldsymbol{x}, \Lambda_{\boldsymbol{b}'})$ defined over $\Lambda_{\boldsymbol{b}}$ is continuous and convex, by the maximum principle, its maximum over the polyhedron $\Lambda_{\boldsymbol{b}}$ is achieved on $\mathcal{V}(\boldsymbol{b})$. Thus, we have:

$$\begin{aligned} d_H(\Lambda_{\boldsymbol{b}}, \Lambda_{\boldsymbol{b}'}) &= \max[\max_{\boldsymbol{x}\in\Lambda_{\boldsymbol{b}}} d(\boldsymbol{x}, \Lambda_{\boldsymbol{b}'}) , \max_{\boldsymbol{x}\in\Lambda_{\boldsymbol{b}'}} d(\Lambda_{\boldsymbol{b}}, \boldsymbol{x})] \\ &= \max[\max_{\boldsymbol{x}\in\mathcal{V}(\boldsymbol{b})} d(\boldsymbol{x}, \Lambda_{\boldsymbol{b}'}) , \max_{\boldsymbol{x}\in\mathcal{V}(\boldsymbol{b}')} d(\Lambda_{\boldsymbol{b}}, \boldsymbol{x})] \\ &\leqslant \max[\max_{\boldsymbol{x}\in\mathcal{V}(\boldsymbol{b})} d(\boldsymbol{x}, V(\boldsymbol{b}')) , \max_{\boldsymbol{x}\in\mathcal{V}(\boldsymbol{b}')} d(\mathcal{V}(\boldsymbol{b}), \boldsymbol{x})] = d_H\left(\mathcal{V}(\boldsymbol{b}), \mathcal{V}(\boldsymbol{b}')\right) . \end{aligned}$$

For $i \in \{1, \dots, p\}$, let us consider the following hyperplane and associated half-spaces:

$$\mathcal{H}_i(\boldsymbol{b}) \stackrel{\text{def}}{=} \{\boldsymbol{x}|\ \boldsymbol{A}_i\boldsymbol{x} = b_i\}, \quad \mathcal{H}_i^-(\boldsymbol{b}) \stackrel{\text{def}}{=} \{\boldsymbol{x}|\boldsymbol{A}_i\boldsymbol{x} \leqslant b_i\}, \quad \mathcal{H}_i^+(\boldsymbol{b}) \stackrel{\text{def}}{=} \{\boldsymbol{x}|\boldsymbol{A}_i\boldsymbol{x} \geqslant b_i\}.$$

Then, another definition of $\Lambda_{\boldsymbol{b}}$ is given as $\Lambda_{\boldsymbol{b}} = \bigcap_{i\in[1,p]} \mathcal{H}_i^-(\boldsymbol{b})$.

Now fix $\boldsymbol{b}_0 \in \mathcal{B}$ and consider $\boldsymbol{v}_0 \in \mathcal{V}(\boldsymbol{b}_0)$. By definition, $\boldsymbol{v}_0$ is the intersection of hyperplanes i.e. there exists $K_0 \subset \{1, \dots, p\}$, maximal for the inclusion, such that $\{\boldsymbol{v}_0\} = \bigcap_{i\in K_0} \mathcal{H}_i(\boldsymbol{b}_0)$. We have $k \stackrel{\text{def}}{=} \mathrm{card}(K_0) \geqslant n$ otherwise $\boldsymbol{v}$ is not a vertex.

For $J \subset \{1, \dots, p\}$, let $\boldsymbol{A}_J$ denote the submatrix of $\boldsymbol{A}$ obtained by considering the rows $\boldsymbol{A}_j$ for $j \in J$.

More generally, for an aribitrary subset of rows $K \subset \{1, \dots, p\}$ and a point $\boldsymbol{b} \in \mathcal{B}$, let us define the sets of *derived points* of K as:

$$\mathcal{V}_K(\boldsymbol{b}) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{R}^n\ ; \exists J \subset K\ ;\ \boldsymbol{A}_J \text{ is invertible and } \boldsymbol{x} = \boldsymbol{A}_J^{-1}\boldsymbol{b}\} .$$

By definition, we have $\mathcal{V}_{K_0}(\boldsymbol{b}_0) = \{\boldsymbol{v}_0\}$ and, for any $\boldsymbol{b} \in \mathcal{B}$, $\mathcal{V}_K(\boldsymbol{b})$ is a set of at most $\binom{k}{n}$ elements. For any $\boldsymbol{b} \in \mathcal{B}$ and $\boldsymbol{v}' \stackrel{\text{def}}{=} \mathcal{V}_K(\boldsymbol{b})$, there exists $J \subset K$ such that $\boldsymbol{v}' = \boldsymbol{A}_J^{-1}\boldsymbol{b}$ such that we have:

$$\|\boldsymbol{v}_0 - \boldsymbol{v}'\| = \|\boldsymbol{A}_J^{-1}\boldsymbol{b}_0 - \boldsymbol{A}_J^{-1}\boldsymbol{b}\| \leqslant \|\boldsymbol{A}_J^{-1}\| \times \|\boldsymbol{b}_0 - \boldsymbol{b}\| \leqslant \alpha\|\boldsymbol{b}_0 - \boldsymbol{b}\| \tag{5.27}$$

where $\alpha \stackrel{\text{def}}{=} \max_{\boldsymbol{A}_J \text{ invertible}} \|\boldsymbol{A}_J^{-1}\|$.

By definition of K maximal, we have $\eta \stackrel{\text{def}}{=} \min_{j\in[1,p]\setminus K} d\left(\boldsymbol{v}_0, \mathcal{H}_j^+(\boldsymbol{b}_0)\right) > 0$. Moreover, as for any $j \in \{1, \dots, p\}$, the function $\boldsymbol{x} \mapsto d(\boldsymbol{x}, \mathcal{H}_j^+(\boldsymbol{b}_0))$ is continuous, and from (5.27) , there exists $\delta > 0$ such that:

$$\|\boldsymbol{b} - \boldsymbol{b}_0\| \leqslant \delta \quad \Longrightarrow \quad \forall \boldsymbol{v}' \in \mathcal{V}_K(\boldsymbol{b}), \min_{j\in[1,p]\setminus K} d\left(\boldsymbol{v}', \mathcal{H}_j^+(\boldsymbol{b})\right) > 0 .$$

We are going to show that, for $\boldsymbol{b}$ such that $\|\boldsymbol{b} - \boldsymbol{b}_0\| \leqslant \delta$, there exists $\boldsymbol{v}' \in \mathcal{V}_K(\boldsymbol{b}) \cap \mathcal{V}(\boldsymbol{b})$, i.e. a derived point of K which is a vertex. We proceed by induction on $k - n \geqslant 0$.

If $k = n$, then $\boldsymbol{v} = \boldsymbol{A}_K^{-1}\boldsymbol{b}_0$ and for any $\boldsymbol{b}$ in the ball $\mathcal{S}_\delta(\boldsymbol{b}_0)$, $\mathcal{V}_K(\boldsymbol{b}) = \{\boldsymbol{A}_K^{-1}\boldsymbol{b}\}$. Thus $\boldsymbol{v}' = \boldsymbol{A}_K^{-1}\boldsymbol{b}$ verifies $\boldsymbol{A}_K\boldsymbol{v}' = \boldsymbol{b}_K$, and $\boldsymbol{A}_j\boldsymbol{v}' < b_j$ for all $j \notin K$, thus $\boldsymbol{v}'$ belongs to $\mathcal{V}(\boldsymbol{b})$.

If $k = n + t$ with $t \geqslant 1$, there exists $j_0 \in K$ such that with $K' = K \setminus \{j_0\}$, $\mathcal{V}_{K'}(\boldsymbol{b}_0) = \{\boldsymbol{v}\}$. Consider the polyhedron $\mathcal{P} = \bigcap_{i\in K'} \mathcal{H}_i^-(\boldsymbol{b})$. By induction, there exists $J \subset K'$ such that $\boldsymbol{A}_J^{-1}\boldsymbol{b}$ is a vertex of $\mathcal{P}$. If it satisfies also $\boldsymbol{A}_{j_0}\boldsymbol{x} \leqslant \boldsymbol{b}_{j_0}$ then it is an element of $\mathcal{V}(\boldsymbol{b})$. Else, consider a vertex $\boldsymbol{v}'$ of the polyhedron $\mathcal{P} \cap \mathcal{H}_{j_0}^-(\boldsymbol{b})$ on the facet associated with $\mathcal{H}_{j_0}(\boldsymbol{b})$. Then, $\boldsymbol{v}' \in \mathcal{V}_K(\boldsymbol{b})$ and, as $\boldsymbol{b} \in \mathcal{S}_\delta(\boldsymbol{b}_0)$, it verifies $\boldsymbol{A}_j\boldsymbol{v}' < \boldsymbol{b}_j$ for all $j \notin K$, thus $\boldsymbol{v}' \in \mathcal{V}(\boldsymbol{b})$, which terminates the induction.

Thus, in any case and for $\boldsymbol{b} \in \mathcal{S}_\delta(\boldsymbol{b}_0)$, there exists $\boldsymbol{v}' \in \mathcal{V}(\boldsymbol{b}) \cap \mathcal{V}_K(\boldsymbol{b})$ such that:

$$d(\boldsymbol{v}, \mathcal{V}(\boldsymbol{b})) \leqslant \|\boldsymbol{v} - \boldsymbol{v}'\| \leqslant \alpha\|\boldsymbol{b}_0 - \boldsymbol{b}\|$$

and thus, finally, $d_H\left(V(\boldsymbol{b}_0), \mathcal{V}(\boldsymbol{b})\right) \leqslant \alpha\|\boldsymbol{b}_0 - \boldsymbol{b}\|$.

The collection $\left\{\mathcal{S}_{\delta_{\boldsymbol{b}_0}(\boldsymbol{b}_0)}\right\}_{\boldsymbol{b}_0\in\mathcal{B}}$ defines a cover of open sets of the compact set $\mathcal{B}$, thus there exists a finite subcollection, of cardinal denoted by r, that also covers $\mathcal{B}$, from wich we deduce that there exists $D \leqslant \max(r\alpha)$ such that:

$$\forall \boldsymbol{b}, \boldsymbol{b}' \in \mathcal{B},\ d\left(\mathcal{V}(\boldsymbol{b}'), \mathcal{V}(\boldsymbol{b})\right) \leqslant D\|\boldsymbol{b}' - \boldsymbol{b}\|\ ,$$

which terminates the proof. □

Chapter 6

Nonatomic Aggregative Games with Infinitely Many Types

This chapter is based on the paper [JW19], submitted for publication. In this chapter, we consider nonatomic aggregative games: the set of players is given as a continuum, as opposed to a finite set as in Chapter 5. Here, there can be an infinite number of different players types, a player's type being defined by her action set and cost function. As in Chapter 5, we consider the presence of aggregative coupling constraints. After defining the notion of variational Wardrop equilibrium (VWE) we show that a sequence of symmetric VWE associated to auxiliary games with a finite number of types, converge to a VWE of the initial game.

6.1 Introduction

A motivating example Consider the example of an energy operator studying the flexibility potential between peak and off-peak periods in a large population of energy consumers, for instance all households in France.

The operator considers that each household n has a certain quantity of energy E_n that can be balanced between consumption on peak period $x_{P,n}$ and consumption on off-peak periods $x_{O,n}$, such that $x_{O,n} + x_{P,n} = E_n$, depending on the cost (per unit of energy) $c_P(X_P)$ and $c_O(X_O)$ associated with the peak and off peak periods. The total on-peak consumption $X_P = \sum_n x_{P,n}$ and off-peak consumption $X_O = \sum_n x_{O,n}$ affect the prices on the energy market and, therefore, change the costs $c_P(X_P)$ and $c_O(X_O)$ set by the operator.

The operator wants to compute an equilibrium of this game (for instance to design tariffs). For practical and privacy reasons, it is impossible to have access to the flexibility potential E_n of the thirty millions of French households. However, the operator may have an easier access to a precise parametric, continuous distribution function of the flexibility potential among the French households.

Then, using the inverse transform sampling method, the game is replicated by modeling the population of households as a continuum $\Theta = [0,1]$ and associating to each $\theta \in \Theta$ the flexible energy quantities $E_\theta = F_E^{-1}(\theta)$ from the inverse of the cumulative distribution function F_E. As the distribution is continuous, there is an infinity of different energy quantity E_θ i.e. an infinity of *players types* in the obtained game, where a *type* refers to the definition of a set of feasible actions and a payoff function. The operator has two questions: 1) how to characterize an equilibrium of this nonatomic game with an infinity of players types ? and 2) how to compute such an equilibrium ? This chapter provides answers to those two questions.

The game described above belongs to the class of aggregative games. In such a game, a player's payoff is determined by her own action and the aggregate of all the players' actions [Cor94]. The setting of aggregative games is particularly relevant to the study of nonatomic games [Sch73], i.e. games with a continuum of players. There, a player has an interaction with the other players only via an aggregate-level profile of their actions, while she has no interest or no way to know the behavior of any particular player or the identity of the player making a certain choice.

Nonatomic games are readily adapted to many situations in industrial engineering or public sectors where a huge number of users, such as traffic commuters and electricity consumers, are involved. These users have no direct interaction except through the aggregate congestion or consumption to which they are contributing collectively. These situations can often be modeled as a congestion game, a special class of aggregative games, both in nonatomic version and finite-player version. The latter, called atomic congestion game, was formally formulated by Rosenthal in 1973 [Ros73b], while related research work in transportation and traffic analysis, mostly in the nonatomic version, appeared much earlier [War52; BMW56]. The theory of congestion games has also found numerous applications in telecommunications [ORS93], distributed computing [AKH02], energy management [Atz+13], and so on.

Nonatomic games are mathematical tools adapted to the modeling of interactions between a very large number of agents. Practical cases exist where a nonatomic model is intuitive and straightforward as when the modeler has an easier access to a description of the population through a parametric distribution of the types, as illustrated in the example.

As many distributions used in practice (e.g. normal distribution) are continuous, this implies that the nonatomic game obtained using these distributions will have an infinite number of players' types.

The concept of equilibrium in nonatomic games is captured by the so called Wardrop equilibrium (WE) [War52]. A nonatomic player neglects the impact of her deviation on the aggregate profile of the whole population's actions, in contrast to a finite player.

For the computation of WE, existing results are limited to particular classes of nonatomic games, such as population games [MS82; HS98; San11], where only a finite number of types of players are considered, each type sharing the same finite number of pure actions and the same payoff function.

The objective of this chapter is to provide a model of nonatomic aggregative games with *infinitely many compact convex pure-action sets and infinitely many payoff functions*—in general a specific action set and a specific payoff function for each nonatomic player— then introduce a general form of coupling aggregative constraints into these games, define an appropriate notion of equilibrium, study the properties such as existence and uniqueness of these equilibria and, finally, their computation through an approximation.

Main results After defining a pure-action profile in a nonatomic game where players have specific compact convex pure-action sets lying in $\mathbb{R}^T$, and specific cost functions, convex in their own action variable, Theorem 6.1 characterizes a WE as a solution to an infinite-dimensional variational inequality (IDVI).

Using the IDVI formulation, we extend this equilibrium notion to the case of a game with coupling aggregative constraints, by defining variational Wardrop equilibrium (VWE). Theorem 6.3 proves the existence of WE and VWE in monotone nonatomic games by showing the existence of solutions to the characteristic IDVI.

In Theorem 6.4, we establish the uniqueness of WE and VWE in case of strictly monotone or aggregatively strictly monotone games. The definition of monotone games is an extension of the stable games [HS09], also called dissipative games [SW15], in population games with a finite types of nonatomic players to the case with infinitely many types.

In the case where the nonatomic aggregative game has only a finite number of types of players, we define the notion of symmetric action profiles and symmetric VWE (SVWE), describing situations where all players of the same type play the same action. Proposition 6.2 shows that SVWEs are characterized as solutions of a *finite*-dimensional VI. Besides, Proposition 6.3 shows that, under monotonicity assumptions, there always exists an SVWE.

Theorem 6.5 is the main result of this chapter. It shows that, for a sequence of finite-type approximating games, if the finite number of pure-action sets and cost functions converge to those of the players of a monotone nonatomic aggregative game, and if the aggregative constraint converges to the aggregative constraint of the infinite nonatomic game, then any sequence of SVWE associated to the sequence of approximating games converges in pure-action profile or in aggregate action profile to the VWE of the infinite-type game. We provide an upper bound on the distance between the approximating SVWE and the VWE, specified

as a function of the parameters of the approximating finite-type games and the initial infinite nonatomic game.

This result allows the construction of a finite-type approximating game sequence and associated SVWEs so as to approximate the infinite-dimensional VWE in the special class of strongly or aggregatively strongly monotone nonatomic aggregative games, with or without aggregative constraints. Since finding solutions of finite-dimensional variational inequalities —characterizing SVWEs—is a computationally tractable problem (see e.g. [FP07]), it follows from our results that a VWE of a nonatomic game with infinitely many types can be approximated with arbitrary precision.

Section 6.3.3 shows how to construct a finite-type approximating games sequence for two general classes of nonatomic games. Appendix 6.A gives the main ideas to extend all our results to the case where players have nonsmooth subdifferentiable cost functions: to make the presentation of the key ideas easier, we focus on the smooth case in the body of the chapter. Appendix 6.B explains how we can use the same arguments to show the convergence of a sequence of Nash equilibria associated to *atomic finite-player games* (instead of nonatomic finite-type games) to a VWE of a nonatomic game.

Related work Extensive research has been conducted on WE in nonatomic congestion games via their formulation with variational inequalities [MP07]. In addition to their existence and uniqueness, the computational and dynamical aspects of equilibria as solutions to variational inequalities have also been studied [Smi84a; ZM94; ZN97; CPP02]. However, in most cases, the variational inequalities involved have finite dimensions, as opposed to the case of WE in this chapter. Marcotte and Zhu [MZ97] consider nonatomic players with continuous types (leading to a characterization of the WE as an IDVI) and studied the equilibrium in an aggregative game with nonatomic players differentiated through a linear parameter in their cost function.

Convergence of some dynamical systems describing the evolution of pure-action distribution in the population of a nonatomic game has been established for some particular equilibria in some particular classes such as linear games [TJ78], potential games [BMW56; San01] and stable games [Smi84b; HS09]. Algorithms corresponding to discretized versions of such dynamical systems for the computation of WE have been studied, in particular for congestion games [Fri+94; ZN97].

In engineering applications of nonatomic games such as the management of traffic flow or energy consumption, individual commuters or consumers often have specific choice sets due to individual constraints, and specific payoff functions due to personal preferences. Also, unlike for a transportation user who usually chooses a single path, an electricity consumer as modeled in the example above faces a resource allocation problem where she has to split the consumption of a certain quantity of energy over different time periods. Hence, her pure-action set is no longer a finite, discrete set as a commuter but a compact convex set in $\mathbb{R}^T$ where T is the total number of time periods. Few results exist for the computation of pure-action WE in the case where players have continuous action sets and in the case where there are infinitely many different types (i.e. action sets and payoff functions) of players. For example, [Sch73] shows the existence of equilibrium in nonatomic games with finite action sets. Mas-Colell [MC84] and Carmona and Podzeck [CP09] consider compact strategy sets and show the existence of mixed strategy equilibria, and do not consider the case of aggregative games and pure-strategy equilibria. In their model, all players share the same actions set. Besides, most of the existing work assumes smooth cost functions of players which is somewhat restrictive in applications, as for instance electricity tariffs or tolls are usually not continuous.

Similarly, the subject of nonatomic games involving (aggregative) coupling constraints has only been partially adressed. Coupling constraints at an aggregative level are to be considered in many of the above-mentioned applications, as also mentioned in [Gra17]: for instance, when modeling the electricity consumption (see above example), some capacity constraints of the network or ramping constraints on the variation of total energy consumption between time periods are natural to consider from an engineering point of view. As seen in this chapter, the presence of coupling constraints is not a simple artifact, as it adds non

trivial difficulties in the analysis of WE and their computation. Indeed, an appropriate definition of equilibrium is already not obvious. An analog to the so-called generalized Nash equilibrium [Har91] for finite-player games does not exist for nonatomic games because a nonatomic player's behavior has no impact on the aggregate profile. Moreover, dynamical systems and algorithms used to compute WE in population games cannot be straightforwardly extended to this case. Indeed, in these dynamics and algorithms, players adapt their strategies unilaterally in their respective strategy spaces, which can well lead to a new strategy profile violating the coupling constraint.

Several works have quantified the relationship between Nash and Wardrop equilibria, a subject close to the present work, as shown in Appendix 6.B. Haurie and Marcotte [HM85] show that in a sequence of atomic splittable games where atomic splittable players are replaced by smaller and smaller equal-size players with constant total weight, Nash equilibria converge to the WE of a nonatomic game. Their proof is based on the convergence of variational inequalities corresponding to the sequence of Nash equilibria, a technique similar to the one used in the present work. Wan [Wan12a] generalizes this result to composite games where nonatomic players and atomic splittable players coexist, by allowing the atomic players to replace themselves by players with heterogeneous sizes.

Gentile et al. [Pac+18] consider a specific class of finite-player aggregative games with linear coupling constraints. They use the variational inequality formulations for the unique generalized Nash equilibrium and the unique generalized Wardrop-type equilibrium (which consists in letting each finite player act as if she was nonatomic) of the same finite-player game to show that, when the number of players grows, the former can be approximated by the latter. Their results are different from ours, as we consider nonatomic games with players of infinitely-many different types instead of finite-player games. Consequently, we consider VWE and symmetric VWE instead of generalized equilibria, which do not exist in nonatomic games. In contrast to generalized equilibria, a variational equilibrium is not characterized by a best reply condition for each of the finite or nonatomic players, as shown in Section 6.3. We also consider a general form of coupling constraints, and extend our results to nonsmooth cost functions, as shown in Appendix 6.A (we focus on the differentiable case in the body of the chapter).

Milchtaich [Mil00] studies finite and nonatomic crowding games (similar to nonatomic aggregative games), where players have finitely many pure actions, and shows that, if each player in an n-person game is replaced by m identical replicas with constant total weight, pure Nash equilibria generically converge to the unique equilibrium of the limit nonatomic game as m goes to infinity. His proof is not based on a variational inequality formulation.

Structure The remaining of the chapter is organized as follows. Section 6.2 introduces the definitions of nonatomic aggregative games with and without aggregative constraints. After defining the WE and VWE notions in the case of an infinite number of players' types, we show, under monotonicity assumptions, the existence and uniqueness of equilibria via generalized IDVIs. In the case of finite-type games, we define the notions of symmetric profiles and SVWE and show their characterization through finite-dimensional VIs and their existence. In Section 6.3, we give the definition of a sequence of finite-type approximating games for a nonatomic aggregative game with or without coupling constraints, and present the main theorem of the chapter on the convergence of the sequence of S(V)WE, associated to the sequence of finite-type approximating games, to the (V)WE of the nonatomic game. In Section 6.3.3, the construction of sequences of finite-type approximating games is shown for two important classes of nonatomic aggregative games. In Section 6.4, we step back to the flexible energy example given above, and derive our results to the computation of an SVWE in this framework.

Last, in Appendix 6.A, we show how our results extend to the case of nonsmooth cost functions and in Appendix 6.B, we show how the results can be adapted to prove the convergence of Nash equilibria to a VWE of a nonatomic game.

Notation The closed ball in a metric space, centered at $\boldsymbol{x}$ and of radius η, is denoted by $B_\eta(\boldsymbol{x})$.

For a nonempty convex set C in a Hilbert space $\mathcal{H}$ (over $\mathbb{R}$),

- $T_C(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathcal{H} : \boldsymbol{y} = 0 \text{ or } \exists (\boldsymbol{x}_k)_k \text{ in } C \text{ s.t. } \boldsymbol{x}_k \neq \boldsymbol{x}, \boldsymbol{x}_k \to \boldsymbol{x}, \frac{\boldsymbol{x}_k - \boldsymbol{x}}{\|\boldsymbol{x}_k - \boldsymbol{x}\|} \to \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}\}$ is the *tangent cone* of C at $\boldsymbol{x} \in C$;
- $\operatorname{span} C = \{\sum_{i=1}^k \alpha_i \boldsymbol{x}_i : k \in \mathbb{N}, \alpha_i \in \mathbb{R}, \boldsymbol{x}_i \in C\}$ is the *linear span* of C;
- $\operatorname{aff} C = \{\sum_{i=1}^k \alpha_i \boldsymbol{x}_i : k \in \mathbb{N}, \alpha_i \in \mathbb{R}, \sum_i \alpha_i = 1, \boldsymbol{x}_i \in C\}$ is the *affine hull* of C;
- $\operatorname{ri} C = \{\boldsymbol{x} \in C : \exists \eta > 0 \text{ s.t. } B_\eta(\boldsymbol{x}) \cap \operatorname{aff} C \subset C\}$ is the *relative interior* of C;
- $\operatorname{rbd} C$ is the *relative boundary* of C in $\mathcal{H}$, i.e. the boundary of C in $\operatorname{span} C$.

The *inner product* of two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in any Euclidean space $\mathbb{R}^T$ is denoted by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^T x_i y_i$. The *$l^2$-norm* of $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\| \stackrel{\text{def}}{=} \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2}$.

We denote by $L^2([0,1], \mathbb{R}^T)$ the Hilbert space of measurable functions from $[0,1]$ to $\mathbb{R}^T$ w.r.t. the Lebesgue measure μ and which are square integrable with respect to the Lebesgue measure.

The *inner product* of two vector functions F and G in $L^2([0,1], \mathbb{R}^T)$ is denoted by $\langle F, G \rangle_2 = \int_0^1 \langle F(\theta), G(\theta) \rangle \, \mathrm{d}\theta$.

The Hilbert space $L^2([0,1], \mathbb{R}^T)$ is endowed with *L^2-norm*: $\|F\|_2 = \langle F, F \rangle_2^{1/2}$.

We denote $d_m(\boldsymbol{x}, \mathcal{A}) \stackrel{\text{def}}{=} \inf_{\boldsymbol{y} \in \mathcal{A}} \|\boldsymbol{x} - \boldsymbol{y}\|_m$ the distance between a point $\boldsymbol{x}$ and a set $\mathcal{A}$, where m is omitted or is equal to 2, depending on whether we consider an Euclidean space or $L^2([0,1], \mathbb{R}^T)$.

Similarly, $d_{H,m}(\mathcal{A}, \mathcal{B})$ denotes the *Hausdorff distance* between two sets $\mathcal{A}$ and $\mathcal{B}$, which is defined as

$$d_{H,m}(\mathcal{A}, \mathcal{B}) \stackrel{\text{def}}{=} \max\{\sup_{\boldsymbol{x} \in \mathcal{A}} d_m(\boldsymbol{x}, \mathcal{B}), \sup_{\boldsymbol{y} \in \mathcal{B}} d_m(\boldsymbol{y}, \mathcal{A})\} .$$

For a function $(\boldsymbol{x}, \boldsymbol{X}) \mapsto f(\boldsymbol{x}, \boldsymbol{X})$ of two explicit variables, convex in $\boldsymbol{x}$, we denote by $\nabla_1 f(\boldsymbol{x}, \boldsymbol{X})$ the differential of function $f(\cdot, \boldsymbol{X})$ for any fixed $\boldsymbol{X}$, except in Appendix 6.A in which ∂ (resp. ∂_1) is used to denote the (resp. partial) subdifferential.

6.2 Monotonicity, Coupling Constraints and Symmetric Equilibrium

6.2.1 Nonatomic aggregative games

In nonatomic aggregative games considered here, players have compact pure-action sets, and heterogeneous pure-action sets as well as heterogeneous cost function. This model is in line with Schmeidler's seminal paper [Sch73]. It differs from most of the population games studied in game theory [HS98; San11], in which nonatomic players are grouped into several populations, with players in the same population having the same finite pure-action set and the same cost function.

Definition 6.1 (Nonatomic aggregative game)**.** A *nonatomic aggregative game* G is defined by:

i) a continuum of players represented by the points on the real interval $\Theta = [0,1]$ endowed with Lebesgue measure;

ii) a set of feasible pure actions $\mathcal{X}_\theta \subset \mathbb{R}^T$ for each player $\theta \in \Theta$, with $T \in \mathbb{N}^*$ a constant;

iii) a cost function $\mathcal{X}_\theta \times \mathbb{R}^T \to \mathbb{R} : (\boldsymbol{x}_\theta, \boldsymbol{X}) \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$ for each player θ, where $\boldsymbol{X} = (X_t)_{t=1}^T$ and $X_t \stackrel{\text{def}}{=} \int_0^1 \boldsymbol{x}_{\theta',t} \, \mathrm{d}\theta'$ refers to an aggregate-action profile, given the action profile $(\boldsymbol{x}_{\theta'})_{\theta' \in \Theta}$ for the population Θ.

The set of feasible pure-action profiles is defined by:

$$\boldsymbol{\mathcal{X}} \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in L^2([0,1], \mathbb{R}^T) \ : \ \forall \, \theta \in \Theta, \boldsymbol{x}_\theta \in \mathcal{X}_\theta\right\}.$$

Denote the game by $G = (\Theta, \boldsymbol{\mathcal{X}}, (f_\theta)_{\theta \in \Theta})$.

Remark 6.1. *The definition of a nonatomic aggregative game asks the pure-action profile $\boldsymbol{x}$ to be a measurable and integrable function on Θ instead of simply being a collection of $\boldsymbol{x}_\theta \in \mathcal{X}_\theta$ for $\theta \in \Theta$. In other words, a coupling constraint is inherent in the definition of nonatomic aggregative games and the notion of WE. This is in contrast with finite-player games.*

The set of feasible aggregate actions is defined as:

$$\overline{\mathcal{X}} \overset{\text{def}}{=} \{\boldsymbol{X} \in \mathbb{R}^T : \exists\, \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \text{ s.t. } \textstyle\int_0^1 \boldsymbol{x}_\theta \,\mathrm{d}\theta = \boldsymbol{X}\}\,.$$

Further assumptions are needed for $\boldsymbol{\mathcal{X}}$ to be nonempty and for the existence of an equilibrium.

Assumption 6.1 (Nonatomic pure-action sets)**.** *The correspondence $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^T, \theta \mapsto \mathcal{X}_\theta$ has nonempty, convex, compact values. Moreover, for all $\theta \in \Theta$, $\mathcal{X}_\theta \subset B_R(\mathbf{0})$, with $R > 0$ a constant.*

Under Assumption 6.1, a sufficient condition for $\boldsymbol{x}$ to be in $L^2([0,1], \mathbb{R}^T)$ is that $\boldsymbol{x}$ is measurable.

Notation We denote by $\mathcal{M} = [0, R+1]^T$ the hypercube in $\mathbb{R}^T$ of edge $R+1$.

Assumption 6.2 (Measurability)**.** *The correspondence $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^T, \theta \mapsto \mathcal{X}_\theta$ has a measurable graph $Gr_{\mathcal{X}} = \{(\theta, \boldsymbol{x}_\theta) \in \mathbb{R}^{T+1} : \theta \in \Theta, \boldsymbol{x}_\theta \in \mathcal{X}_\theta\}$, i.e. $Gr_{\mathcal{X}}$ is a Borel subset of $\mathbb{R}^{T+1}$. The function $Gr_{\mathcal{X}} \to \mathbb{R}^T : (\theta, \boldsymbol{x}_\theta) \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ is measurable for each $\boldsymbol{Y} \in \mathbb{R}^T$.*

Assumption 6.3 (Nonatomic convex cost functions)**.** *For all θ, f_θ is defined on $(\mathcal{M}')^2$, where $\mathcal{M}'$ is a neighborhood of $\mathcal{M}$, and:*

i) for each $\theta \in \Theta$, function f_θ is continuous. In particular, f_θ is bounded on $\mathcal{M}^2$;

ii) for each $\theta \in \Theta$ and each aggregate profile $\boldsymbol{Y} \in \mathcal{M}$, $\boldsymbol{x} \mapsto f_\theta(\boldsymbol{x}, \boldsymbol{Y})$ is differentiable and convex on $\mathcal{M}'$;

iii) there is $L_{\mathbf{f}} > 0$ such that $\|\nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})\| \leqslant L_{\mathbf{f}}$ for each $\boldsymbol{x}_\theta \in \mathcal{M}$, each $\boldsymbol{Y} \in \mathcal{M}$, and each $\theta \in \Theta$.

Remark 6.2. *Assumption 6.3.iii) implies that $f_\theta(\cdot,\cdot)$'s are Lipschitz in the first variable with a uniform Lipschitz constant $L_{\mathbf{f}}$ on $\mathcal{M}^2$ for all θ.*

We also need the continuity of the derivative of cost functions in the second (aggregate) variable:

Assumption 6.4. *For each $\theta \in \Theta$ and each $\boldsymbol{x}_\theta \in \mathcal{M}$, the function $\boldsymbol{Y} \mapsto \nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ is continuous on $\mathcal{M}$.*

Wardrop equilibrium extends the notion of Nash equilibrium in the framework of nonatomic games, where a single player of measure zero has a negligible impact on the others.

Definition 6.2 (Wardrop Equilibrium (WE), [War52])**.** A pure-action profile $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ is a pure *Wardrop equilibrium* of nonatomic aggregative game G if we have, with $\boldsymbol{X}^* = \int_{\theta\in\Theta} \boldsymbol{x}^*_\theta \,\mathrm{d}\theta$:

$$f_\theta(\boldsymbol{x}^*_\theta, \boldsymbol{X}^*) \leqslant f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}^*), \quad \forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta,\ \forall\, a.e.\, \theta \in \Theta\,.$$

All the actions and equilibria in this chapter are pure, hence from now on, we no longer emphasize it.

Before characterizing WE by infinite-dimensional VI (IDVI), we introduce some notions and a technical assumption ensuring that the IDVI is well-defined.

Lemma 6.1. *For all $\boldsymbol{x} \in L^2([0,1], \mathcal{M})$, the function $\mathbf{g}_{\boldsymbol{x}}$ defined from Θ to $\mathbb{R}^T$ by*

$$\mathbf{g}_{\boldsymbol{x}}(\theta) \overset{\text{def}}{=} \nabla_1 f_\theta(\boldsymbol{x}_\theta, \textstyle\int \boldsymbol{x})\,, \quad \forall \theta \in \Theta\,,\ \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}) \tag{6.1}$$

is measurable on Θ. Consequently, $\mathbf{g}_\cdot$ is a mapping from $L^2([0,1], \mathcal{M})$ to $L^2([0,1], \mathbb{R}^T)$.

Proof of Lemma 6.1. For $n \in \mathbb{N}$ large enough,

$$\mathbf{g}_{\boldsymbol{x},n} \stackrel{\text{def}}{=} \theta \mapsto n\big(f_\theta(\boldsymbol{x}_\theta + \frac{1}{n}, \int \boldsymbol{x}) - f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x})\big)$$

is well defined. It is measurable according to Assumption 6.2. Thus $\mathbf{g}_{\boldsymbol{x}} = \lim_n \mathbf{g}_{\boldsymbol{x},n}$ is also measurable as a limit of measurable functions.

□

Theorem 6.1 (IDVI formulation of WE)**.** *Under Assumptions 6.1 to 6.3, $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ is a WE of nonatomic aggregative game G if and only if either of the following two equivalent conditions is true:*

$$\forall\, a.e.\, \theta \in \Theta, \quad \langle \nabla_1 f_\theta(\boldsymbol{x}^*_\theta, \boldsymbol{X}^*), \boldsymbol{x}_\theta - \boldsymbol{x}^*_\theta \rangle \geqslant 0\,, \quad \forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta\,, \tag{6.2a}$$

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta \geqslant 0\,, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}\,. \tag{6.2b}$$

Proof of Theorem 6.1. Given $\boldsymbol{X}^*$, (6.2a) is a necessary and sufficient condition for $\boldsymbol{x}^*_\theta$ to minimize the convex function $f_\theta(., \boldsymbol{X}^*)$ on $\mathcal{X}_\theta$. Condition (6.2a) implies condition (6.2b) because of Lemma 6.1.

For the converse, suppose that $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ satisfies condition (6.2b) but not (6.2a). Then there must be a subset Θ' of Θ with strictly positive measure such that:

$$\forall \theta \in \Theta',\ \boldsymbol{x}^*_\theta \notin \mathcal{Y}_\theta \stackrel{\text{def}}{=} \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{X}^*)\,.$$

In particular, for any $\boldsymbol{y}_\theta \in \mathcal{Y}_\theta$:

$$\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta),\, \boldsymbol{y}_\theta - \boldsymbol{x}^*_\theta \rangle < f_\theta(\boldsymbol{y}_\theta, \boldsymbol{X}^*) - f_\theta(\boldsymbol{x}^*_\theta, \boldsymbol{X}^*) < 0\,.$$

A consequence of Assumptions 6.2 and 6.3 is that the function $\Theta \times \mathcal{M} \to \mathbb{R} : (\theta, \boldsymbol{z}) \mapsto f_\theta(\boldsymbol{z}, \boldsymbol{X}^*)$ is a Carathéodory function, that is, (i) $f_\cdot(\boldsymbol{z}, \boldsymbol{X}^*)$ is measurable on Θ for each $\boldsymbol{z} \in \mathcal{M}$, and (ii) $f_\theta(\cdot, \boldsymbol{X}^*)$ is continuous on $\mathcal{M}$ for each $\theta \in \Theta$. Thus, according to the measurable maximum theorem [AB06, Thm. 18.19] applied to $f_\cdot(\cdot, \boldsymbol{X}^*)$, there exists a selection $\boldsymbol{y}_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{X}^*)$ for $\theta \in \Theta'$ such that $\Theta' \to \mathbb{R}^T : \theta \mapsto \boldsymbol{y}_\theta$ is a measurable function. By defining $\boldsymbol{y}_\theta = \boldsymbol{x}^*_\theta$ for $\theta \notin \Theta'$, one has $\Theta \to \mathbb{R}^T : \theta \mapsto \boldsymbol{y}_\theta$ is measurable and hence belongs to $\boldsymbol{\mathcal{X}}$.

However, we have:

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta = \int_{\Theta'} \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta < 0\,,$$

contradicting (6.2b). □

Remark 6.3. *Condition* (6.2a) *is equivalent to $\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta \rangle \geqslant 0$ for all $\boldsymbol{y}_\theta \in T_{\mathcal{X}_\theta}(\boldsymbol{x}^*_\theta)$ for each θ. It means that no unilateral deviation is profitable. However, since each nonatomic player has measure zero, when considering a deviation in the action profile, one must let players in a set of strictly positive measure deviate:* (6.2b) *means that the collective deviation of players of any set of strictly positive measure increases their cost.*

The existence of WE is obtained by an equilibrium existence theorem for nonatomic games.

Theorem 6.2 (Existence of a WE, [Rat92])**.** *Under Assumption 6.1, Assumption 6.2 and Assumption 6.3.i), if for all θ and all $\boldsymbol{Y} \in \mathcal{M}$, $f_\theta(\cdot, \boldsymbol{Y})$ is continuous on $\mathcal{M}$, then the nonatomic aggregative game G admits a WE.*

Proof. The conditions required in Remark 8 in Rath's 1992 paper [Rat92] on the existence of WE in aggregate games are satisfied. □

Remark 6.4. *No convexity of $f_\theta(\cdot, \boldsymbol{Y})$'s are needed for the existence.*

6.2.2 Monotone Nonatomic Aggregative Games

For the uniqueness of WE and the existence of equilibrium notion to be introduced in the next subsection for the case with coupling constraints, we consider the different *monotonicity* assumptions given Definition 6.3 below. One can observe the similarity with the monotonicity notions defined in Definition 5.4 in Chapter 5 for the atomic case.

Monotonicity in nonatomic aggregative games, is also sometimes referred to as *stability*.

Definition 6.3. With notation $\mathbf{g}_{\boldsymbol{x}}(\theta) = \nabla_1 f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x})$, for any $\theta \in \Theta$ and for any $\boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M})$, we say that the nonatomic aggregative game G is

i) *monotone* if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geqslant 0, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \, . \tag{6.3}$$

ii) *strictly monotone* if the equality in (6.3) holds if and only if $\boldsymbol{x} = \boldsymbol{y}$ almost everywhere.

iii) *aggregatively strictly monotone* if the equality in (6.3) holds if and only if $\int \boldsymbol{x} = \int \boldsymbol{y}$.

iv) *strongly monotone* with modulus α if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geqslant \alpha \|\boldsymbol{x} - \boldsymbol{y}\|_2^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \, . \tag{6.4}$$

v) *aggregatively strongly monotone* with modulus β if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geqslant \beta \| \int \boldsymbol{x} - \int \boldsymbol{y} \|^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \, . \tag{6.5}$$

Remark 6.5. *A recent paper of Hadikhanloo [Had17] generalizes the notion of stability in population games [HS09] to monotonicity in* anonymous games*, an extension of population games with players having heterogeneous compact action sets but the same payoff function. He defines the notion of monotonicity directly on the distribution of actions among the players instead of action profile as done here. The two approaches are compatible.*

Example of public products games. An interesting example of aggregative games is given by cost functions of the form:

$$f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) = \langle \boldsymbol{x}_\theta, \boldsymbol{c}(\boldsymbol{X}) \rangle - u_\theta(\boldsymbol{x}_\theta) \, , \tag{6.6}$$

where $\boldsymbol{c}(\boldsymbol{X})$ specifies the per-unit cost (or negative of per-unit utility) of each of the T "public products", which is a function of the aggregative contribution $\boldsymbol{X}$ to each of the "public products". Player θ's cost (resp. negative of utility) associated to these products is scaled by her own contribution $\boldsymbol{x}_\theta$. The function $u_\theta(\boldsymbol{x}_\theta)$ measures the private utility of player θ (resp. negative of private cost) for the contribution $\boldsymbol{x}_\theta$.

For instance, in a public goods game, $-c_t(X_t)$ is the common per-unit payoff for using public good t, determined by the total contribution X_t, while $-u_\theta(\boldsymbol{x}_\theta)$ is player θ's private cost of supplying $\boldsymbol{x}_\theta$ to the public goods; in a Cournot competition, $-c_t(X_t)$ is the common market price for product t, determined by its total supply X_t, while $-u_\theta(\boldsymbol{x}_\theta)$ is player θ's private cost of producing $x_{\theta,t}$ unit of product t; in a congestion game, $c_t(X_t)$ is the common per-unit cost for using arc t in a network, determined by the aggregate load X_t on arc t, while $u_\theta(\boldsymbol{x}_\theta)$ is player θ's private utility of her routing or energy consuming choice $\boldsymbol{x}_\theta$.

Proposition 6.1. *Under Assumptions 6.1 to 6.3, in a* public products *game G (i.e. with cost functions of form* (6.6)*), assume that $\boldsymbol{c}$ is monotone on $\mathcal{M}$ and, for each θ, u_θ is a concave function on $\mathcal{M}$. Then:*

i) G is a monotone game.

ii) If u_θ is strictly concave on $\mathcal{M}$ for all $\theta \in \Theta$, then G is a strictly monotone game.

iii) If $\boldsymbol{c}$ is strictly monotone on $\mathcal{M}$, then G is an aggregatively strictly monotone game.

iv) If u_θ is strongly concave on $\mathcal{M}$ with modulus α_θ for each $\theta \in \Theta$ and $\inf_{\theta\in\Theta} \alpha_\theta = \alpha > 0$, then G is a strongly monotone game with modulus α.

v) If $\boldsymbol{c}$ is strongly monotone on $\mathcal{M}$ with β, then G is an aggregatively strongly monotone game with modulus β.

Proof. The proof is symmetric to the atomic case given in Section 5.B. □

In particular, if $\boldsymbol{c}(\boldsymbol{X}) = (c_t(X_t))_{t\in T}$, then $\boldsymbol{c}$ is monotone if functions c_t, $t \in \mathcal{T}$ are all nondecreasing, and $\boldsymbol{c}$ is strongly monotone if functions c_t, $t \in \mathcal{T}$ are all strictly increasing.

6.2.3 Aggregate constraints and VWE

Let us consider the *aggregative constraint* in nonatomic aggregative game G:

$$\boldsymbol{X} \in \mathcal{A}\,,$$

where $\mathcal{A}$ is a convex compact subset of $\mathbb{R}^T$ such that $\mathcal{A} \cap \overline{\mathcal{X}} \neq \emptyset$. Let $\boldsymbol{\mathcal{X}}(\mathcal{A})$ be a subset of $\boldsymbol{\mathcal{X}}$ defined by

$$\boldsymbol{\mathcal{X}}(\mathcal{A}) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \boldsymbol{X} = \int \boldsymbol{x} \in \mathcal{A}\}\,.$$

Let us denote the nonatomic aggregative game with aggregative constraint $\boldsymbol{X} \in \mathcal{A}$ by $G(\mathcal{A})$.

A notion of generalized WE similar to the one of generalized Nash equilibrium in finitely-many-player games—where each player does the best she can while not violating the coupling constraints given the choices of the others [Har91]—is not well-defined in a nonatomic game. Indeed, since the impact of a nonatomic player's choice on the aggregate profile is negligible, the feasible action set of a nonatomic player θ facing the choices of the others $\boldsymbol{x}_{-\theta}$ in a game with coupling constraint is not a well-established notion: either $\int \boldsymbol{x}_{-\theta} \in \mathcal{A}$ then $\mathcal{X}_\theta = \mathcal{X}$, or $\int \boldsymbol{x}_{-\theta} \notin \mathcal{A}$ then $\mathcal{X}_\theta = \emptyset$. Departing from an action profile in $\boldsymbol{\mathcal{X}}(\mathcal{A})$, simultaneous unilateral deviations by the players can lead to any profile in $\boldsymbol{\mathcal{X}}$. If only profiles in $\boldsymbol{\mathcal{X}}(\mathcal{A})$ are allowed to be attained, then one lands on a notion similar to the so-called variational Nash equilibrium in finite-many-player games [Har91]. Indeed, the most natural notion of equilibrium with the presence of aggregative constraint is the notion of variational Wardrop equilibrium, where feasible deviations are defined on a collective basis.

Definition 6.4 (Variational Wardrop Equilibrium (VWE))**.** A solution to the following IDVI problem:

$$\text{Find } \boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(\mathcal{A}) \text{ s.t. } \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta \geqslant 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A}), \tag{6.7}$$

is called a *variational Wardrop equilibrium* of $G(\mathcal{A})$.

Remark 6.6 (VWE in the literature)**.** *In the literature of congestion games, the equilibrium notion characterized by VI of form* (6.7) *but in finite dimension and with smooth cost functions has long been studied. For example, see [LP99; MNS04; CSSM04; Zho+11] and references therein.*

The following facts are needed for later use.

Lemma 6.2. *Under Assumptions 6.1 and 6.2:*

i) $\boldsymbol{\mathcal{X}}$ is a nonempty, convex, closed and bounded subset of $L^2([0,1], \mathbb{R}^T)$;

ii) $\boldsymbol{\mathcal{X}}(\mathcal{A})$ is a nonempty, convex and closed subset of $\boldsymbol{\mathcal{X}}$;

iii) $\overline{\mathcal{X}}$ and $\mathcal{A} \cap \overline{\mathcal{X}}$ are nonempty, convex and compact subsets of $\mathbb{R}^T$.

We omit the proof and only point out that $\boldsymbol{\mathcal{X}}$ and $\overline{\mathcal{X}}$ are nonempty because of Assumption 6.1 and the measurable selection theorem of Aumann [Aum69], while aggregate-action set $\overline{\mathcal{X}}$ is compact by [Aum65, Theorem 4].

Theorem 6.3 shows the existence of VWE via the VI approach.

Theorem 6.3 (Existence of VWE)**.** *Under Assumptions 6.1 to 6.4, if a nonatomic aggregative game with coupling constraint $G(\mathcal{A})$ is monotone on $\boldsymbol{\mathcal{X}}(\mathcal{A})$, then a VWE exists.*

Proof. We can apply [DT96, Corollary 2.1] which shows that Equation (6.7) has a solution, as:

- $\boldsymbol{\mathcal{X}}(\mathcal{A})$ is bounded, closed and convex in $L^2([0,1],\mathbb{R}^T)$;
- $\mathbf{g}_\cdot : L^2([0,1],\mathcal{M}) \rightrightarrows L^2([0,1],\mathbb{R}^T)$ is a monotone correspondence which is upper hemi-continuous from the line segments in $\boldsymbol{\mathcal{X}}(\mathcal{A})$ to the weak* topology of $L^2([0,1],\mathbb{R}^T)$. Notice that $\mathbf{g}_\cdot$ has closed values. Let us do the proof in the general nonsmooth case.

Take $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\boldsymbol{\mathcal{X}}(\mathcal{A})$, consider sequence $(\boldsymbol{x}^{(k)})_k$ with $\boldsymbol{x}^{(k)} \in [\boldsymbol{x},\boldsymbol{y}]$ with $\boldsymbol{x}^{(k)} \to \boldsymbol{x}$, and sequence $(\mathbf{g}^{(k)})_k$ such that $\mathbf{g}^{(k)} \in \mathbf{g}_{\boldsymbol{x}^{(k)}}$ and $\mathbf{g}^{(k)} \overset{*}{\rightharpoonup} \mathbf{g}^\infty$ with $\mathbf{g}^\infty \in L^2([0,1],\mathbb{R}^T)$. Let us show that $\mathbf{g}^\infty \in \mathbf{g}_{\boldsymbol{x}}$. We have $\boldsymbol{X} = \int \boldsymbol{x}$ converging to $\boldsymbol{X}^{(k)} = \int \boldsymbol{x}^{(k)}$ in l^2-norm.

By definition of $\mathbf{g}_\cdot$ and convexity, for each $\boldsymbol{z} \in \mathcal{M}$ and each θ, we have

$$f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}_\theta^{(k)}) \geqslant f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}^{(k)}) + \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle \,.$$

Since f_θ is continuous in both variables, we get:

$$f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}^{(k)}) \underset{k\to\infty}{\longrightarrow} f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}) \;\text{ and }\; f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}^{(k)}) \underset{k\to\infty}{\longrightarrow} f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) \,. \tag{6.8}$$

Besides, we have:

$$\begin{aligned} \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle &= \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle + \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle \,, \\ \text{and}\quad \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle &\to \langle \mathbf{g}_\theta^{\infty}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle \,, \end{aligned}$$

because $\mathbf{g}^{(k)} \overset{*}{\rightharpoonup} \mathbf{g}^\infty$, while $\langle \mathbf{g}_\theta^{(k)}, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle \to 0$ because $\mathbf{g}_\theta^{(k)}$'s are uniformly bounded by $L_{\mathbf{f}}$. Therefore,

$$f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}) \geqslant f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}) + \langle \mathbf{g}_\theta, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle \,,$$

so that $\mathbf{g}_\theta^\infty \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$. Since the limit of measurable functions is measurable, $\mathbf{g}$ is measurable. Hence $\mathbf{g}^\infty \in \mathbf{g}_{\boldsymbol{x}}$ (and $\mathbf{g}^\infty = \mathbf{g}_{\boldsymbol{x}}$ in the smooth case), which concludes.

□

Theorem 6.4 (Uniqueness of VWE)**.** *Under Assumptions 6.1 to 6.3:*

i) if $G(\mathcal{A})$ is strictly monotone on $\boldsymbol{\mathcal{X}}(\mathcal{A})$, then it has at most one VWE;

ii) if $G(\mathcal{A})$ is aggregatively strictly monotone on $\boldsymbol{\mathcal{X}}(\mathcal{A})$, then all VWE of $G(\mathcal{A})$ have the same aggregative profile;

iii) if G (without aggregative constraint) is only aggregatively strictly monotone but, for each $\theta \in \Theta$ and all $\boldsymbol{Y} \in \mathcal{M}$, $f_\theta(\boldsymbol{x}, \boldsymbol{Y})$ is strictly convex in $\boldsymbol{x}$, then there is at most one WE.

Proof. Suppose that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ are both VWE. Let $\boldsymbol{X} = \int \boldsymbol{x}$ and $\boldsymbol{Y} = \int \boldsymbol{y}$. According to Theorem 6.1, we have $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \,\mathrm{d}\theta \geqslant 0$ and $\int_\Theta \langle \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \,\mathrm{d}\theta \geqslant 0$. Adding up these two inequalities yields $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \,\mathrm{d}\theta \geqslant 0$.

i) If $G(\mathcal{A})$ is a strictly monotone game, then $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \,\mathrm{d}\theta = 0$ and thus $\boldsymbol{x} = \boldsymbol{y}$ almost everywhere.

ii)-iii) If $G(\mathcal{A})$ is an aggregatively strictly monotone game, then we have:

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \,\mathrm{d}\theta = 0 \,,$$

and thus $\boldsymbol{X} = \boldsymbol{Y}$. If there is no aggregative constraint and $f_\theta(\cdot, \boldsymbol{Z})$ is strictly convex for all $\boldsymbol{Z} \in \mathcal{M}$, then for all θ, $\boldsymbol{x}_\theta$ (resp. $\boldsymbol{y}_\theta$) is the unique minimizer of $f_\theta(\cdot, \boldsymbol{X})$ (resp. $f_\theta(\cdot, \boldsymbol{Y})$). Since $\boldsymbol{X} = \boldsymbol{Y}$, one has $\boldsymbol{x}_\theta = \boldsymbol{y}_\theta$. □

6.2.4 Symmetric VWE with a finite number of types

A particular class of nonatomic aggregative games is those with only a finite number of types of players, that is, when the sets $\{\mathcal{X}_\theta\}_\theta$ and $\{f_\theta\}_\theta$ are both finite. Consider a nonatomic aggregative game with a set of I types $\mathcal{I} = \{1, \dots, I\}$. The player set Θ is divided into I measurable subsets $\Theta_1, \dots, \Theta_I$ such that each nonatomic player belonging to Θ_i is of type i. Let us denote the common action set of players in Θ_i by $\mathcal{X}_i$ and their common cost functions by f_i.

Let us consider a particular class of action profiles in these finite-type nonatomic aggregative games, called *symmetric action profiles*:

Definition 6.5 (Symmetric action profile and symmetric variational Wardrop equilibrium (SVWE))**.** The set of symmetric action profiles, denoted by $\boldsymbol{\mathcal{X}}_S$, is the set of action profiles where players of the same type play the same action:

$$\boldsymbol{\mathcal{X}}_S \stackrel{\text{def}}{=} \{x \in \boldsymbol{\mathcal{X}} : \boldsymbol{x}_\theta = \boldsymbol{x}_\xi, \forall \theta, \xi \in \Theta_i, \forall i \in \mathcal{I}\}$$

The set of symmetric action profiles satisfying the aggregative constraint is denoted by

$$\boldsymbol{\mathcal{X}}_S(\mathcal{A}) \stackrel{\text{def}}{=} \boldsymbol{\mathcal{X}}_S \cap \boldsymbol{\mathcal{X}}(\mathcal{A})\,. \tag{6.9}$$

A symmetric variational Wardrop equilibrium is a VWE that is symmetric.

For any symmetric action profile $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S$, let the common action of players of type $i \in \mathcal{I}$ be denoted by $\boldsymbol{x}_i$, so that the action profile can be specified by $(\boldsymbol{x}_i)_{i\in\mathcal{I}}$. Obviously, for $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S$, for each type $i \in \mathcal{I}$, $\mathbf{g}_{\boldsymbol{x}}(\theta) = \mathbf{g}_{\boldsymbol{x}}(\xi) = \nabla_1 f_i(\boldsymbol{x}_i, \int \boldsymbol{x})$ for all $\theta, \xi \in \mathcal{I}$. Let us abusively denote this quantity by $\mathbf{g}_{\boldsymbol{x}}(i)$.

Proposition 6.2. *In a finite-type nonatomic aggregative game $G(\mathcal{A})$ with an aggregative constraint, a VWE is a symmetric one if and only if it is a solution to the following VI:*

$$\text{Find } \hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}_S(\mathcal{A}) \text{ s.t. } \textstyle\sum_{i\in\mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \mu_i \boldsymbol{x}_i - \mu_i \hat{\boldsymbol{x}}_i\rangle \geqslant 0,\ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S(\mathcal{A})\,, \tag{6.10}$$

where μ_i is the Lebesgue measure of Θ_i.

Proof. Since $\boldsymbol{\mathcal{X}}_S(\mathcal{A}) \subset \boldsymbol{\mathcal{X}}(\mathcal{A})$, it is clear that a SVWE, characterized as a solution to the IDVI (6.7), is a solution to (6.10).

Conversely, suppose that $\hat{\boldsymbol{x}}$ is a solution to the VI problem (6.10), let us show that it also solves the IDVI (6.7). Indeed, for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$,

$$\begin{aligned}\int_\Theta \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(\theta), \boldsymbol{x}_\theta - \hat{\boldsymbol{x}}_\theta\rangle \,\mathrm{d}\theta &= \textstyle\sum_{i\in\mathcal{I}} \int_{\Theta_i} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(\theta), \boldsymbol{x}_\theta - \hat{\boldsymbol{x}}_\theta\rangle \,\mathrm{d}\theta = \sum_{i\in\mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \textstyle\int_{\Theta_i} \boldsymbol{x}_\theta - \mu_i \hat{\boldsymbol{x}}_i\rangle \\ &= \textstyle\sum_{i\in\mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \mu_i \frac{\int_{\Theta_i} \boldsymbol{x}_\theta}{\mu_i} - \mu_i \hat{\boldsymbol{x}}_i\rangle \geqslant 0\end{aligned}$$

as for all $\theta \in \Theta_i$, $\boldsymbol{x}_\theta \in \mathcal{X}_i$ which is convex, hence $\frac{\int_{\Theta_i} \boldsymbol{x}_\theta}{\mu_i} \in \mathcal{X}_i$, so that (6.10) can be applied. □

Proposition 6.3 (Existence of SVWE)**.** *Under Assumptions 6.1, 6.3 and 6.4, a finite-type nonatomic aggregative game $G(\mathcal{A})$ admits a SVWE.*

Proof. First note that the VI problem (6.10) is equivalent to a finite dimension VI:

$$\text{Find } \hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}_S(\mathcal{A})' \text{ s.t. } \textstyle\sum_{i\in\mathcal{I}} \langle \nabla_1 f_i(\hat{\boldsymbol{x}}_i, \sum_{j\in\mathcal{I}} \mu_j \hat{\boldsymbol{x}}_j), \mu_i \boldsymbol{x}_i - \mu_i \hat{\boldsymbol{x}}_i\rangle \geqslant 0,\ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S(\mathcal{A})'\,, \tag{6.11}$$

where $\boldsymbol{\mathcal{X}}_S(\mathcal{A})'$ is the finite-dimensional set

$$\boldsymbol{\mathcal{X}}_S(\mathcal{A})' \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{R}^{IT} : \boldsymbol{x} \in \textstyle\prod_{i\in\mathcal{I}} \mathcal{X}_i,\ \sum_{i\in\mathcal{I}} \mu_i \boldsymbol{x}_i \in \mathcal{A}\}\,.$$

As the mapping $(\boldsymbol{x}_i)_{i\in\mathcal{I}} \mapsto \left(\nabla_1 f_i(\boldsymbol{x}_i, \sum_j \mu_j \hat{\boldsymbol{x}}_j)\right)_{i\in\mathcal{I}}$ is continuous from Assumption 6.4, then [HS66, Lemma 3.1] implies that the VI (6.11) has a solution on the finite dimensional convex compact $\boldsymbol{\mathcal{X}}_S(\mathcal{A})'$. □

In this chapter, only SVWE are considered for finite-type nonatomic aggregative games. For such equilibria, the distribution of different types on Θ is not relevant, since the equilibrium behavior of each player is only determined by the finite dimensional VI (6.11). Therefore, we shall specify a finite-type nonatomic aggregative game only by the tuple $\big((\mu_i)_{i\in\mathcal{I}}, (\mathcal{X}_i)_{i\in\mathcal{I}}, (f_i)_{i\in\mathcal{I}}, A\big)$. In particular, a symmetric action profile in such a game shall be specified by a vector $(x_i)_{i\in\mathcal{I}} \in \mathbb{R}^{IT}$, and the set of symmetric action profiles is nothing else but $\boldsymbol{\mathcal{X}}_S(\mathcal{A})'$.

6.3 Approximating Nonatomic Aggregative Games with an Infinity of types

6.3.1 Finite-type approximating game sequence

After introducing (V)WE in nonatomic aggregative games and SVWE in finite-type nonatomic aggregative games with coupling constraints, we study the relation between these notions. As WE is a particular case of VWE when the aggregate constraint set is any subset of $\mathbb{R}^T$ containing $\overline{\mathcal{X}}$, we can only consider the case of VWE and SVWE.

This section shows the following result: considering a sequence of equilibria of "approximating" finite-type nonatomic aggregative games $(G^\nu(\mathcal{A}^\nu))_\nu$ of a nonatomic aggregative game $G(\mathcal{A})$, where each type of players in $G^\nu(\mathcal{A}^\nu)$ corresponds to a collection of nonatomic players who are similar in their types, a sequence of SVWE in $(G^\nu(\mathcal{A}^\nu))_\nu$ converges to the VWE of $G(\mathcal{A})$ when this one is (aggregatively) strongly monotone.

The particularity of SVWE is that it can be characterized by a finite dimensional VI. As opposed to the case of infinite dimensional ones, there is a large literature on algorithms for computing solutions of finite dimensional VI (e.g. [FP07] and references therein). Therefore, the result stated above can be practically used to compute a VWE, solution of an IDVI, with arbitrary precision.

In this section, we always suppose that Assumptions 6.1 to 6.4 hold.

Let us consider the following definition of an approximating game sequence, where the indicators to measure the quality of approximation are similar to the ones introduced in (5.12) and (5.13) in Chapter 5, Section 5.3.2.

Definition 6.6. Finite-type Approximating Games Sequence
A sequence of finite-type nonatomic aggregative games

$$\{G^\nu(\mathcal{A}^\nu) = \big((\mu_i^\nu)_{i\in\mathcal{I}^\nu}, (\mathcal{X}_i^\nu)_{i\in\mathcal{I}^\nu}, (f_i^\nu)_{i\in\mathcal{I}^\nu}, \mathcal{A}^\nu\big) : \nu \in \mathbb{N}^*\}$$

with aggregative constraints is a *finite-type approximating game sequence* for the nonatomic aggregative game $G(\mathcal{A}) = \big(\Theta, \boldsymbol{\mathcal{X}}, (f_\theta)_\theta, A\big)$ with an aggregative constraint if, for each $\nu \in \mathbb{N}^*$, there exists a partition $(\Theta_0^\nu, \Theta_1^\nu, \dots, \Theta_{I^\nu}^\nu)$ of the set Θ, with $\mathcal{I}^\nu \overset{\text{def}}{=} \{1, \dots, I^\nu\}$, such that the Lebesgue measure of Θ_0^ν is $\mu_0^\nu = 0$, and if, for each $i \in \mathcal{I}^\nu$, the Lebesgue measure of Θ_i^ν is μ_i^ν while the collection of nonatomic players in Θ_i^ν are getting close to the nonatomic players of type $i \in \mathcal{I}^\nu$ in the sense that, as $\nu \to +\infty$:

i) $\overline{\delta}^\nu \overset{\text{def}}{=} \max_{i\in\mathcal{I}^\nu} \delta_i^\nu \longrightarrow 0$, where δ_i^ν is the Hausdorff distance between the action sets of nonatomic players in Θ_i^ν in $G(\mathcal{A})$ and the action set of nonatomic players of type $i \in \mathcal{I}^\nu$ in $G^\nu(\mathcal{A}^\nu)$:

$$\delta_i^\nu \overset{\text{def}}{=} \sup_{\theta\in\Theta_i^\nu} d_H\left(\mathcal{X}_\theta, \mathcal{X}_i^\nu\right), \tag{6.12}$$

and span $\mathcal{X}_i^\nu = \text{span } \mathcal{X}_\theta$, $\forall\theta \in \Theta_i^\nu$.

ii) $\overline{d}^\nu \overset{\text{def}}{=} \max_{i\in\mathcal{I}^\nu} d_i^\nu \longrightarrow 0$, where d_i^ν measures the distance between the differential of nonatomic players' cost functions in $G(\mathcal{A})$ and that of their corresponding players' cost functions in $G^\nu(\mathcal{A}^\nu)$:

$$d_i^\nu \overset{\text{def}}{=} \sup_{\theta\in\Theta_i} \sup_{(\boldsymbol{x},\boldsymbol{Y})\in\mathcal{M}^2} \left\|\nabla_1 f_i^\nu(\boldsymbol{x}_i, \boldsymbol{Y}) - \nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})\right\|. \tag{6.13}$$

iii) $D^\nu \longrightarrow 0$, where $D^\nu \overset{\text{def}}{=} d_H\left(\mathcal{A}^\nu, \mathcal{A}\right)$ is the Hausdorff distance between the aggregative constraint set $\mathcal{A}^\nu \subset \mathbb{R}^T$ and the aggregative constraint set $\mathcal{A} \subset \mathbb{R}^T$. Besides, span $\mathcal{A} =$ span $\mathcal{A}^\nu$ for all $\nu \in \mathbb{N}^*$.

Roughly speaking, along a sequence of finite-type approximating games, for each nonatomic player θ in Θ, the difference between her type and her corresponding type i in the approximating game $G^\nu(\mathcal{A})$ (in the sense that $\theta \in \Theta_i^\nu$) is disappearing as ν goes to infinity. Also, the aggregate-profile constraint sets of the sequence of approximating games converge to the one in $G(\mathcal{A})$.

Note that, except the last condition on D^ν, the other conditions are independent of the constraint sets $(\mathcal{A}^\nu)_\nu$ and $\mathcal{A}$.

Remark 6.7. *The assumption* span $\mathcal{X}_i^\nu =$ span $\mathcal{X}_\theta$, $\forall \theta \in \Theta_i^\nu$ *is needed for our proofs because of the existence of coupling constraints. It implies in particular that the nonatomic infinite game considered is such that* $\{\text{span } \mathcal{X}_\theta\}_{\theta\in\Theta}$ *has a finite number of elements. This assumption is natural as, in many models,* span $\mathcal{X}_\theta$ *will be the same for all* $\theta \in \Theta$ *(see example of Section 6.4).*

Remark 6.8. *Without loss of generality, we assume* $\text{ri}\left(\mathcal{A} \cap \overline{\mathcal{X}}\right) \neq \emptyset$ *in this section. Indeed, if the nonempty convex compact set* $\mathcal{A} \cap \overline{\mathcal{X}}$ *has an empty relative interior, then it is reduced to a point hence the problem becomes trivial.*

In Section 6.3.3, we will construct a sequence of finite-type approximating games for two fairly general cases of nonatomic aggregative games.

In order to compare symmetric action profiles in the approximating games and action profiles in the original game, we introduce the following linear mappings which define an equivalent action profile for a symmetric action profile in an approximating game, and vice versa.

First, define the set:

$$L_{\text{S}}^{\mathcal{I}^\nu} \overset{\text{def}}{=} \left\{\boldsymbol{x} \in L^2([0,1], \mathcal{M}) : \boldsymbol{x}_\theta = \boldsymbol{x}_\xi, \forall \theta, \xi \in \Theta_i, \forall i \in \mathcal{I}^\nu\right\},$$

and the mapping $\psi^\nu : L^2([0,1], \mathcal{M}) \to L_{\text{S}}^{\mathcal{I}^\nu}$ for each $\nu \in \mathbb{N}^*$ by

$$\begin{aligned} \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}^T),\ \psi^\nu(\boldsymbol{x}) &= \left(\psi_\theta^\nu(\boldsymbol{x})\right)_{\theta\in\Theta}, \\ \text{where } \forall i \in \mathcal{I}^\nu,\ \forall \theta \in \Theta_i^\nu,\ \psi_\theta^\nu(\boldsymbol{x}) &= \frac{\int_{\Theta_i^\nu} \boldsymbol{x}_\xi \mathrm{d}\xi}{\mu_i^\nu}. \end{aligned} \tag{6.14}$$

The interpretation of ψ^ν is that a nonatomic player $\theta \in \Theta_i$ (type i) adopts the average behavior of players in Θ_i^ν.

In the following, we assume that Assumptions 6.1, 6.3 and 6.4 also hold for each game G^ν of a sequence of finite-type approximating games: this appears naturally in many cases if $(G^\nu)_\nu$ is built from G, as seen in Section 6.3.3. Finally, let us make the following additional assumption for this section.

Assumption 6.5. *There is a strictly positive constant* η *and an action profile* $\bar{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}$ *such that, for almost all* $\theta \in \Theta$, $d(\bar{\boldsymbol{x}}_\theta, \text{rbd } \mathcal{X}_\theta) > \eta$.

It means that the action space of each player has an (aggregatively) nonempty relative interior and that the relative interior is not vanishing along any sequence of players.

6.3.2 Convergence of equilibrium and aggregate equilibrium profiles

The following Theorem 6.5 gives the main result of this chapter. It shows that a VWE in a strongly monotone nonatomic aggregative game can be approximated by SVWE of a finite-type approximating games sequence, both in the case with and without aggregative constraints.

According to Theorem 6.4, a strongly monotone game is strictly monotone, hence the VWE is unique, while an aggregatively strongly monotone game is aggregatively strictly monotone, hence the aggregate-action profile at VWE is unique.

Theorem 6.5 (Convergence of SVWE to VWE)**.** *Under Assumptions 6.1 to 6.5, let* $(G^\nu(\mathcal{A}^\nu))_\nu$ *be a sequence of finite-type approximating games for the nonatomic aggregative game* $G(\mathcal{A})$ *with an aggregative constraint* $\mathcal{A}$*. Let* $\boldsymbol{x}^*$ *be the VWE of* $G(\mathcal{A})$*,* $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}^\nu(\mathcal{A}^\nu)$ *an SVWE of* $G^\nu(\mathcal{A}^\nu)$ *for each* $\nu \in \mathbb{N}^*$*, and* $\boldsymbol{X}^*$*,* $\boldsymbol{X}^\nu$ *their respective aggregate-action profiles. Then, there exists a constant* $\rho > 0$ *such that the following results hold with* $K_{\mathcal{A}} \overset{\text{def}}{=} \frac{R+1}{\rho}$*:*

i) If G *is aggregatively strongly monotone with modulus* β*,* $(\boldsymbol{X}^\nu)_\nu$ *converges to* $\boldsymbol{X}^*$*: for all* $\nu \in \mathbb{N}^*$ *such that* $\max(\overline{\delta}^\nu, D^\nu) < \rho$*,*

$$\|\boldsymbol{X}^\nu - \boldsymbol{X}^*\|^2 \leqslant \frac{1}{\beta}\Big((4L_{\mathbf{f}}+1)K_{\mathcal{A}}\max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu\Big). \tag{6.15}$$

ii) If G *is strongly monotone with modulus* α*, then* $(hxx^\nu)_\nu$ *, converges to* $\boldsymbol{x}^*$ *in* L^2*-norm: for all* $\nu \in \mathbb{N}^*$ *such that* $\max(\overline{\delta}^\nu, D^\nu) < \rho$*,*

$$\|\boldsymbol{x} - \boldsymbol{x}^*\|_2^2 \leqslant \frac{1}{\alpha}\Big((4L_{\mathbf{f}}+1)K_{\mathcal{A}}\max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu\Big)\,. \tag{6.16}$$

If there are no aggregate constraints, one can replace $K_{\mathcal{A}}$ *(resp.* D^ν*) by* $\frac{1}{2}$ *(resp. 0) in* (6.15) *and* (6.16)*, and Assumption 6.5 is no longer required.*

Some notions and a series of lemmas are needed for the proof of Theorem 6.5.

Notation Let $\mathrm{P}_i^\nu(\cdot)$ denote the (Euclidean) projection function onto $\mathcal{X}_i^\nu$ for $i \in \mathcal{I}^\nu$ and $\mathrm{P}_\theta(\cdot)$ the projection function onto $\mathcal{X}_\theta$ for $\theta \in \Theta$.

Let P^ν denote the (Euclidean) projection function onto $\boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu) \subset L_{\mathrm{S}}^{I^\nu}$, and P the projection function onto $\boldsymbol{\mathcal{X}}(\mathcal{A}) \subset L^2([0,1], \mathbb{R}^T; \mu)$.

Since $\mathcal{X}_i^\nu$'s, $\mathcal{X}_\theta$'s, $\boldsymbol{\mathcal{X}}(\mathcal{A})$ and $\boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$'s (as defined by (6.9)) are all convex and closed in their respective Hilbert spaces, the projection functions onto these sets are well defined.

The following Lemma 6.3 shows that the players become infinitesimal along a sequence of finite-type approximating games.

Lemma 6.3. *Under Assumption 6.1, for all* $\nu \in \mathbb{N}^*$*,* $\|\boldsymbol{x}^\nu\|_2 \leqslant \overline{\delta}^\nu + R$ *for all* $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$*.*

Proof. Let $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$ and $\theta \in \Theta_i^\nu$. By definition of δ_i^ν,

$$\left\|\boldsymbol{x}_i^\nu - \mathrm{P}_\theta\left(\boldsymbol{x}_i^\nu\right)\right\| \leqslant \delta_i^\nu\,,$$

so that we have:

$$\|\boldsymbol{x}_i^\nu\| \leqslant \left(\delta_i^\nu + \left\|\mathrm{P}_\theta(\boldsymbol{x}_i^\nu)\right\|\right) \leqslant (\delta_i^\nu + R)\,.$$

Then, $\|\boldsymbol{x}^\nu\|_2^2 = \sum_{i=1}^{I^\nu}\int_{\Theta_i}\|\boldsymbol{x}_i^\nu\|^2\,\mathrm{d}\theta = \sum_{i=1}^{I^\nu}\mu_i^\nu\|\boldsymbol{x}_i^\nu\|^2 \leqslant \sum_{i=1}^{I^\nu}\mu_i^\nu(\delta_i^\nu + R)^2 \leqslant (\overline{\delta}^\nu + R)^2$. □

The following lemma shows that the convergence of each type of action set in finite-type game G^ν to that of her corresponding nonatomic player in G, assumed by Equation (6.12), implies the convergence of the product action sets in $L^2([0,1], \mathcal{M})$.

Lemma 6.4 (Convergence of $\boldsymbol{\mathcal{X}}_S^\nu$ to $\boldsymbol{\mathcal{X}}$)**.** *Under Assumption 6.1, for all* $\nu \in \mathbb{N}^*$*,*

i) for each $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$*,* $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}) \leqslant \overline{\delta}^\nu$*;*

ii) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$*,* $d_2(\psi^\nu(\boldsymbol{x}), \boldsymbol{\mathcal{X}}_S^\nu) \leqslant \overline{\delta}^\nu$*;*

iii) for each $i \in \mathcal{I}^\nu$ *and each* $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$*, if* $d(\boldsymbol{x}_i^\nu, \mathrm{rbd}\ \mathcal{X}_i^\nu) > \delta_i^\nu$*, then* $\boldsymbol{x}_i^\nu \in \mathcal{X}_\theta$ *for all* $\theta \in \Theta_i^\nu$*;*

iv) for each $i \in \mathcal{I}^\nu$*, each* $\theta \in \Theta_i^\nu$*, and each* $\boldsymbol{x}_\theta \in \mathcal{X}_\theta$*, if* $d(\boldsymbol{x}_\theta, \mathrm{rbd}\ \mathcal{X}_\theta) > \delta_i^\nu$*, then* $\boldsymbol{x}_\theta \in \mathcal{X}_i^\nu$*.*

Proof. i) Let $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define

$$\boldsymbol{y}_\theta = \mathrm{P}_\theta(\boldsymbol{x}^\nu) \in \mathcal{X}_\theta \text{ so that } \|\boldsymbol{y}_\theta - \boldsymbol{x}^\nu\| \leqslant \delta_i^\nu \,.$$

Let us show that $\boldsymbol{y}$ is measurable on each Θ_i^ν, hence measurable on Θ so that $\boldsymbol{y} \in \mathcal{X}$. For that, define κ_i on $\Theta_i^\nu \times \mathcal{M}^T$ by $\kappa_i : (\theta, \boldsymbol{w}) \mapsto \|\boldsymbol{x}_i^\nu - \boldsymbol{w}\|$. Then, κ_i is a Carathéodory function. Since the correspondence $\Theta_i^\nu \ni \theta \mapsto \mathcal{X}_\theta$ is measurable, according to the measurable maximum theorem [AB06, Thm. 18.19], there is a measurable selection of $\boldsymbol{x}_\theta \in \arg\min_{\mathcal{X}_\theta} \kappa_i(\theta, \cdot)$. The minimum of $\kappa_i(\theta, \cdot)$ on $\mathcal{X}_\theta$ is unique and is just $\boldsymbol{y}_\theta$, hence $\boldsymbol{y}$ is measurable on Θ_i^ν.

Then, $\|\boldsymbol{x}^\nu - \boldsymbol{y}\|_2 \leqslant \overline{\delta}^\nu$, which shows that $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}) \leqslant \overline{\delta}^\nu$.

ii) Let $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$. For each $i \in \mathcal{I}^\nu$, $\theta \in \Theta_i^\nu$, $\|\boldsymbol{x}_\theta - \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\| \leqslant \delta_i^\nu$. Since $\mathcal{X}_i^\nu$ is a convex subset in $\mathbb{R}^T$, $\frac{1}{\mu_i^\nu} \int_{\Theta_i^\nu} \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\, \mathrm{d}\theta \in \mathcal{X}_i^\nu$.

Define $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}^\nu$ with, for each $\theta \in \Theta_i$, for $i \in \mathcal{I}_i^\nu$:

$$\boldsymbol{y}_\theta = \boldsymbol{y}_i \overset{\text{def}}{=} \frac{1}{\mu_i^\nu} \int_{\Theta_i^\nu} \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\, \mathrm{d}\theta \in \mathcal{X}_i^\nu \,.$$

Then, from Cauchy-Schwartz inequality:

$$\begin{aligned}
\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2^2 &= \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu \|\psi_i^\nu(\boldsymbol{x}) - \frac{1}{\mu_i^\nu} \textstyle\int_{\Theta_i^\nu} \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\, \mathrm{d}\theta\|^2 \\
&= \sum_{i \in \mathcal{I}^\nu} \frac{1}{\mu_i^\nu} \| \int_{\Theta_i^\nu} (\boldsymbol{x}_\theta - \mathrm{P}_i^\nu(\boldsymbol{x}_\theta))\, \mathrm{d}\theta\|^2 \\
&\leqslant \sum_{i \in \mathcal{I}^\nu} \frac{1}{\mu_i^\nu} \mu_i^\nu \int_{\Theta_i} \|\boldsymbol{x}_\theta - \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\|^2\, \mathrm{d}\theta \\
&= \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i} \|\boldsymbol{x}_\theta - \mathrm{P}_i^\nu(\boldsymbol{x}_\theta)\|^2\, \mathrm{d}\theta \leqslant \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu (\delta_i^\nu)^2 \leqslant (\overline{\delta}^\nu)^2,
\end{aligned}$$

so that $\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2 \leqslant \overline{\delta}^\nu$. This concludes the proof.

iii) Fix $\nu \in \mathbb{N}^*$, $i \in \mathcal{I}^\nu$ and $\theta \in \Theta_i^\nu$. Consider $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$ such that $d(\boldsymbol{x}_i^\nu, \mathrm{rbd}\ \mathcal{X}_i^\nu) > \eta$ for some $\eta > \delta_i^\nu$. Assume that $\boldsymbol{x}_i^\nu \notin \mathcal{X}_\theta$ i.e. $\left\|\boldsymbol{x}_i^\nu - \mathrm{P}_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\| > 0$ and define:

$$\boldsymbol{y}_i^\nu = \boldsymbol{x}_i^\nu + \eta \frac{\boldsymbol{x}_i^\nu - \mathrm{P}_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)}{\left\|\boldsymbol{x}_i^\nu - \mathrm{P}_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\|} \in \mathcal{X}_i^\nu \,,$$

as span $\mathcal{X}_\theta \subset$ span $\mathcal{X}_i^\nu$.

Since $\mathcal{X}_i^\nu$ is convex, we have

$$d(\boldsymbol{y}_i^\nu, \mathcal{X}_\theta) = \left\|\boldsymbol{x}_i^\nu - \mathrm{P}_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\| + \left\|\boldsymbol{y}_i^\nu - \boldsymbol{x}_i^\nu\right\| > \eta > \delta_i^\nu \,,$$

which contradicts the fact that $d(\mathcal{X}_\theta, \boldsymbol{x}_i^\nu) \leqslant \delta_i^\nu$. Hence $\boldsymbol{x}_i^\nu \in \mathcal{X}_\theta$.

iv) The proof is similar to that of iii).

□

Note that, because of the convexity assumptions (Assumption 6.1), the sets of aggregate action profiles in the finite-type game $G^\nu(\mathcal{A}^\nu)$, obtained by considering symmetric or by considering non symmetric profiles are the same, that is:

$$\overline{\mathcal{X}}^\nu \overset{\text{def}}{=} \left\{ \textstyle\int_\Theta \boldsymbol{x}_\theta\, \mathrm{d}\theta \mid \boldsymbol{x} \in \boldsymbol{\mathcal{X}}^\nu \right\} = \left\{ \textstyle\int_\Theta \boldsymbol{x}_\theta\, \mathrm{d}\theta \mid \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S^\nu \right\}, \tag{6.17}$$

and the same equality holds for $\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu$ when considering aggregate constraint $\mathcal{A}^\nu$.

The sequence of sets of aggregate action profiles $(\overline{\mathcal{X}}^\nu)_\nu$ in games $(G^\nu(\mathcal{A}^\nu))_\nu$ *with aggregative constraints*, converges to the set of aggregate-action profiles of the nonatomic aggregative game $G(\mathcal{A})$ *with an aggregative constraint*, as the following lemma says.

Lemma 6.5. *Under Assumption 6.1, for* $\nu \in \mathbb{N}^*$,

i) $d_H(\overline{\mathcal{X}}^\nu, \overline{\mathcal{X}}) \leqslant \overline{\delta}^\nu$;

ii) for $\boldsymbol{X} \in \text{ri } \overline{\mathcal{X}}$, *if* $d(\boldsymbol{X}, \text{rbd } \overline{\mathcal{X}}) > \overline{\delta}^\nu$, *then* $\boldsymbol{X} \in \overline{\mathcal{X}}^\nu$;
for $\boldsymbol{X}^\nu \in \text{ri } \overline{\mathcal{X}}^\nu$, *if* $d(\boldsymbol{X}, \text{rbd } \overline{\mathcal{X}}^\nu) > \overline{\delta}^\nu$, *then* $\boldsymbol{X} \in \overline{\mathcal{X}}$;

iii) for $\boldsymbol{X} \in \text{ri } \mathcal{A}$, *if* $d(\boldsymbol{X}, \text{rbd } \mathcal{A}) > D^\nu$, *then* $\boldsymbol{X} \in \mathcal{A}^\nu$;
for $\boldsymbol{X}^\nu \in \text{ri } \mathcal{A}^\nu$, *if* $d(\boldsymbol{X}^\nu, \text{rbd } \mathcal{A}^\nu) > D^\nu$, *then* $\boldsymbol{X}^\nu \in \mathcal{A}$;

iv) for $\boldsymbol{X} \in \text{ri } (\overline{\mathcal{X}} \cap \mathcal{A})$, *if* $d(\boldsymbol{X}, \text{rbd } (\overline{\mathcal{X}} \cap \mathcal{A})) > \max(\overline{\delta}^\nu, D^\nu)$, *then* $\boldsymbol{X} \in \overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu$;
for $\boldsymbol{X}^\nu \in \text{ri } (\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu)$, *if* $d(\boldsymbol{X}^\nu, \text{rbd } (\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu)) > \max(\overline{\delta}^\nu, D^\nu)$, *then* $\boldsymbol{X}^\nu \in \overline{\mathcal{X}} \cap \mathcal{A}$.

Proof. i) Fix $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$. Consider $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}_S^\nu$ such that $\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2 \leqslant \overline{\delta}^\nu$ (cf. Lemma 6.4). Then, from Cauchy-Schwarz inequality:

$$\|\int \psi^\nu(\boldsymbol{x}) - \int \boldsymbol{y}\|^2 \leqslant \|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2^2 \leqslant (\overline{\delta}^\nu)^2 \,.$$

Hence $d(\int \boldsymbol{x}, \overline{\mathcal{X}}^\nu) \leqslant \|\int \psi^\nu(\boldsymbol{x}) - \int \boldsymbol{y}\| \leqslant \overline{\delta}^\nu$.

On the other hand, let $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$, thus $\boldsymbol{X}^\nu \overset{\text{def}}{=} \int \boldsymbol{x}^\nu \in \overline{\mathcal{X}}^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define $\boldsymbol{y}_\theta = \mathrm{P}_\theta(\boldsymbol{x}_i^\nu) \in \mathcal{X}_\theta$, so that $\|\boldsymbol{x}_\theta^\nu - \boldsymbol{y}_\theta\| \leqslant \delta_i^\nu$. Then, we get:

$$\|\int \boldsymbol{x}^\nu - \int \boldsymbol{y}\| \leqslant \int \|\boldsymbol{x}_\theta^\nu - \boldsymbol{y}_\theta\| \,\mathrm{d}\theta \leqslant \overline{\delta}^\nu \,,$$

which shows that $d\left(\boldsymbol{X}^\nu, \overline{\mathcal{X}}\right) \leqslant \overline{\delta}^\nu$ for all $\boldsymbol{X}^\nu \in \overline{\mathcal{X}}^\nu$.

ii-iii) The proof is similar to the one of Lemma 6.4.iii).

iv) These are corollaries of ii) and iii). □

For the proof of the main theorem, we need to rely on nonatomic profiles that are away from the boundary of the feasible domain, giving us some margin. The existence of such profiles is ensured by the following lemma.

Lemma 6.6. *Under Assumptions 6.1 and 6.5, there is a strictly positive constant* ρ^* *and a nonatomic action profile* $\boldsymbol{z} \in \boldsymbol{\mathcal{X}}$ *such that* $\int \boldsymbol{z} \in \text{ri } (\overline{\mathcal{X}} \cap \mathcal{A})$ *and, for almost all* $\theta \in \Theta$, $d(\boldsymbol{z}_\theta, \text{rbd } \mathcal{X}_\theta) > 3\rho^*$.

Proof. Take $\bar{\boldsymbol{x}}$ the nonatomic action profile in Assumption 6.5 and an arbitrary $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ such that $\int \boldsymbol{y} \in \text{ri } (\overline{\mathcal{X}} \cap \mathcal{A})$.

Let us denote $t \overset{\text{def}}{=} \frac{d(\int \boldsymbol{y}, \text{rbd } (\overline{\mathcal{X}} \cap \mathcal{A}))}{3M}$ and define the profile $\boldsymbol{z} \in \boldsymbol{\mathcal{X}}$ by $\boldsymbol{z} = \boldsymbol{y} - t(\boldsymbol{y} - \bar{\boldsymbol{x}})$.

Firstly, we have:

$$\|\int \boldsymbol{y} - \int \boldsymbol{z}\| = t\|\int \boldsymbol{y} - \int \bar{\boldsymbol{x}}\| \leqslant t2M \leqslant \tfrac{2}{3} d(\int \boldsymbol{y}, \text{rbd } (\overline{\mathcal{X}} \cap \mathcal{A})) \,,$$

hence $\int \boldsymbol{z} \in \text{ri } (\overline{\mathcal{X}} \cap \mathcal{A})$.

Besides, for any θ, $\boldsymbol{z}_\theta = \boldsymbol{y}_\theta - t(\boldsymbol{y}_\theta - \bar{\boldsymbol{x}}_\theta)$. Since $d(\bar{\boldsymbol{x}}_\theta, \text{rbd } \mathcal{X}_\theta) > \eta$, $\boldsymbol{y}_\theta \in \mathcal{X}_\theta$, and $\mathcal{X}_\theta$ is convex, we have:

$$d(\boldsymbol{z}_\theta, \text{rbd } \mathcal{X}_\theta) > \eta t = \frac{\eta}{3M} d(\int \boldsymbol{y}, \text{rbd } (\overline{\mathcal{X}} \cap \mathcal{A})) \,.$$

One concludes by defining $\rho^* \overset{\text{def}}{=} \frac{\eta}{9M} d(\int \boldsymbol{y}, \text{rbd } (\overline{\mathcal{X}} \cap \mathcal{A}))$. □

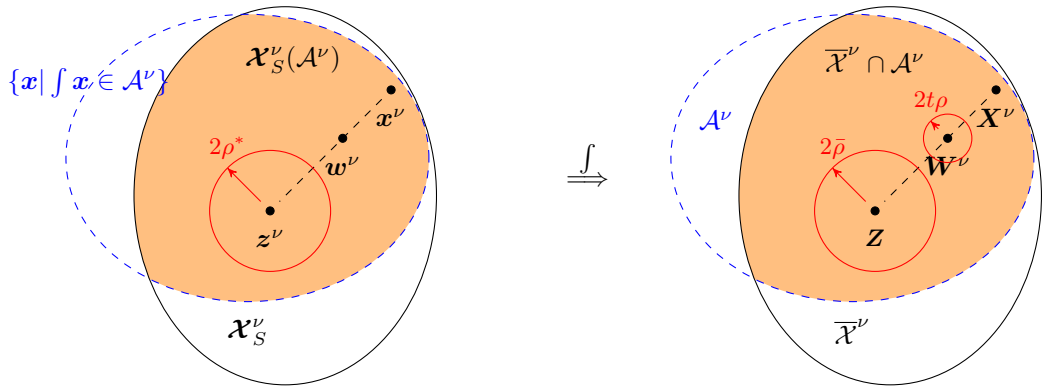


FIGURE 6.1: Illustration of the mapping $\int_\Theta(.)$ between $\boldsymbol{\mathcal{X}}^\nu(\mathcal{A}^\nu)$ and $\overline{\mathcal{X}} \cap \mathcal{A}$ used in Lemma 6.7

Notation Denote $\boldsymbol{Z} = \int \boldsymbol{z}$ where $\boldsymbol{z}$ is the one in Lemma 6.6 and define

$$\bar{\rho} \stackrel{\text{def}}{=} \frac{1}{3} d(\boldsymbol{Z}, \text{rbd } \overline{\mathcal{X}} \cap \mathcal{A}) > 0$$

and the parameter $\rho \stackrel{\text{def}}{=} \min(\rho^*, \bar{\rho})$, appearing in the bounds of Theorem 6.5.

The following lemma shows that the space of symmetric action profiles in the finite-type game *with aggregative constraint*, $\boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$, is converging to the space of action profiles in the nonatomic aggregative game *with aggregative constraint*, $\boldsymbol{\mathcal{X}}(\mathcal{A})$.

Lemma 6.7 (Convergence of $\boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$ to $\boldsymbol{\mathcal{X}}(\mathcal{A})$)**.** *Under Assumptions 6.1 and 6.5, let* $K_{\mathcal{A}} = \frac{R+1}{\rho}$. *Then, for all* $\nu \in \mathbb{N}^*$ *such that* $\max(\bar{\delta}^\nu, D^\nu) < \rho$,

i) for each $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$, $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}(\mathcal{A})) \leqslant 2K_{\mathcal{A}} \max(D^\nu, \bar{\delta}^\nu)$;

ii) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A})$, $d_2(\psi^\nu(\boldsymbol{x}), \boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)) \leqslant 2K_{\mathcal{A}} \max(D^\nu, \bar{\delta}^\nu)$.

Proof. i) Consider $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$ and $\boldsymbol{X}^\nu = \int \boldsymbol{x}^\nu$. Let $\boldsymbol{z}^\nu \stackrel{\text{def}}{=} \psi^\nu(\boldsymbol{z})$ where $\boldsymbol{z}$ is defined in Lemma 6.6. Since for each θ, $d(\boldsymbol{z}_\theta, \text{rbd } \mathcal{X}_\theta) > 3\rho > \bar{\delta}^\nu$, one has $\boldsymbol{z}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$ according to Lemma 6.4.iv).

Define $\boldsymbol{w}^\nu \stackrel{\text{def}}{=} \boldsymbol{x}^\nu + t(\boldsymbol{z}^\nu - \boldsymbol{x}^\nu)$ with $t \stackrel{\text{def}}{=} \frac{\max(D^\nu, \bar{\delta}^\nu)}{\rho} < 1$, then $\boldsymbol{w}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$ from convexity and:

- we have:

$$\begin{aligned} \|\boldsymbol{w}^\nu - \boldsymbol{x}^\nu\|_2 &= \max(D^\nu, \bar{\delta}^\nu) \frac{\|\boldsymbol{z}^\nu - \boldsymbol{x}^\nu\|_2}{\rho} \\ &\leqslant \max(D^\nu, \bar{\delta}^\nu) \frac{2(R+1)}{\rho} \\ &\leqslant 2K_{\mathcal{A}} \max(D^\nu, \bar{\delta}^\nu) ; \end{aligned}$$

- let us show that $\boldsymbol{w}^\nu \in \boldsymbol{\mathcal{X}}$: from Lemma 6.4.iii), it is sufficient to show that $d(\boldsymbol{w}_i^\nu, \text{rbd } \mathcal{X}_i^\nu) > \bar{\delta}_i^\nu$. For that, we show $d(\boldsymbol{z}_i^\nu, \text{rbd } \mathcal{X}_i^\nu) \geqslant 2\rho$ which implies:

$$d(\boldsymbol{w}_i^\nu, \text{rbd } \mathcal{X}_i^\nu) \geqslant 2t\rho > \bar{\delta}^\nu .$$

For any arbitrary $\boldsymbol{y}_i \in B_{2\rho}(\boldsymbol{z}_i^\nu) \cap \text{span } \mathcal{X}_i^\nu$, let $\boldsymbol{y}_\theta \stackrel{\text{def}}{=} \boldsymbol{z}_\theta + \boldsymbol{y}_i - \boldsymbol{z}_i^\nu$ for $\theta \in \Theta_i^\nu$. Then, $\boldsymbol{y}_\theta \in B_{2\rho}(\boldsymbol{z}_\theta) \cap \text{span } \mathcal{X}_\theta$ as $\text{span } \mathcal{X}_\theta = \text{span } \mathcal{X}_i^\nu$, and:

$$d(\boldsymbol{y}_\theta, \text{rbd } \mathcal{X}_\theta) \geqslant d(\boldsymbol{z}_\theta, \text{rbd } \mathcal{X}_\theta) - \|\boldsymbol{z}_\theta - \boldsymbol{y}_\theta\| > 3\rho - 2\rho = \rho > \bar{\delta}^\nu .$$

Hence, from Lemma 6.4.iv), one has $\boldsymbol{y}_\theta \in \mathcal{X}_i^\nu$ and, from convexity, $\boldsymbol{y}_i = \frac{1}{\mu_i^\nu} \int_{\Theta_i} \boldsymbol{y}_\theta \in \mathcal{X}_i^\nu$ which concludes;

• let us show that $\boldsymbol{W}^\nu = \int \boldsymbol{w}^\nu \in \mathcal{A}$: from Lemma 6.5.iii), it is sufficient to show that $\boldsymbol{W}^\nu \in \operatorname{ri} \mathcal{A}^\nu$ and $d(\boldsymbol{W}^\nu, \operatorname{rbd} \mathcal{A}^\nu) > D^\nu$. First, since $\max(\bar{\delta}^\nu, D^\nu) < \bar{\rho}$, one has:

$$d(\boldsymbol{Z}, \operatorname{rbd}(\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu)) \geqslant 3\bar{\rho} - \max(\bar{\delta}^\nu, D^\nu) \geqslant 2\bar{\rho}\,.$$

The linear mapping $\boldsymbol{x} \mapsto \int \boldsymbol{x}$ maps the segment linking $\boldsymbol{x}^\nu$ and $\boldsymbol{z}^\nu$ in $\boldsymbol{\mathcal{X}}^\nu(\mathcal{A}^\nu)$ to a segment linking $\boldsymbol{X}^\nu$ and $\boldsymbol{Z} = \boldsymbol{Z}^\nu$ in $\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu$. Hence, by the definition of $\boldsymbol{w}^\nu$, $B_{2t\rho}(\boldsymbol{W}^\nu) \cap \operatorname{span}(\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu) \subset \overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu$, because each point in $B_{2t\rho}(\boldsymbol{W}^\nu) \cap \operatorname{span}(\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu)$ is on the segment linking $\boldsymbol{X}^\nu$ and some point in $B_{2\rho}(\boldsymbol{Z}) \cap \operatorname{span}(\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu) \subset \overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu$. Finally we have:

$$d(\boldsymbol{W}^\nu, \operatorname{rbd}(\overline{\mathcal{X}}^\nu \cap \mathcal{A}^\nu)) \geqslant 2t\rho = 2\max(D^\nu, \bar{\delta}^\nu) > D^\nu\,,$$

which concludes.

ii) The proof is similar and omitted.

□

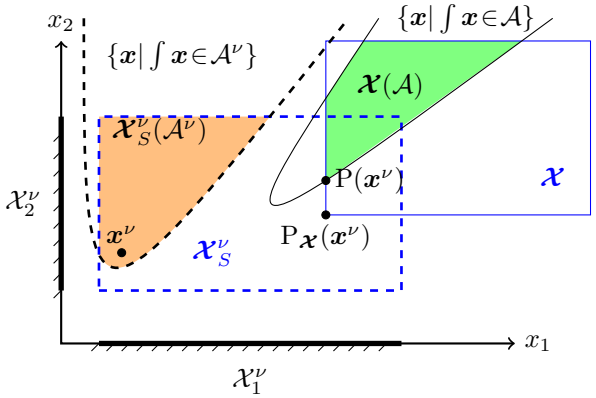


FIGURE 6.2: Difference between projections on $\boldsymbol{\mathcal{X}}$ and on $\boldsymbol{\mathcal{X}}(\mathcal{A})$. (Since it is impossible to draw the graph of a L^2 action-profile space with a continuum of players, we illustrate the idea with two players.)

Remark 6.9 (Difference between unilateral projections of actions and collective projection of the action profile)**.** *Lemma 6.7 shows that*

$$d_2(\boldsymbol{x}^\nu, \mathrm{P}(\boldsymbol{x}^\nu)) \leqslant 2K_{\mathcal{A}} \max(D^\nu, \bar{\delta}^\nu) \quad \text{and} \quad d_2(\psi^\nu(\boldsymbol{x}), \mathrm{P}^\nu(\psi^\nu(\boldsymbol{x}))) \leqslant 2K_{\mathcal{A}} \max(D^\nu, \bar{\delta}^\nu),$$

with P *and* P^ν *the projection functions onto* $\boldsymbol{\mathcal{X}}(\mathcal{A})$ *and* $\boldsymbol{\mathcal{X}}^\nu(\mathcal{A}^\nu)$*. Lemma 6.7 is of first importance for our proof of Theorem 6.5: without it, we only have the convergence of individual, i.e. unilateral action spaces in the sequence of approximating games, to the unilateral action spaces in the nonatomic aggregative game, as shown in Lemma 6.4. Without coupling constraints, this should be sufficient in the proof of the convergence of a sequence of SVWE. However, in the presence of coupling aggregative constraints, this convergence of unilateral action spaces is not enough. Given a profile in* $\boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$*, unilateral projection of each type i player's action onto the corresponding nonatomic player's action space in* $G(\mathcal{A})$*, i.e. from* $\mathcal{X}_i^\nu$ *to* $\mathcal{X}_\theta$*, cannot guarantee that the resulting action profile is in* $\boldsymbol{\mathcal{X}}(\mathcal{A})$*, and vice versa. Lemma 6.7 shows that, for each action profile* $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu(\mathcal{A}^\nu)$*, its projection on the space of nonatomic action profiles in* $\boldsymbol{\mathcal{X}}(\mathcal{A})$ *is close to* $\boldsymbol{x}^\nu$*, and vice versa.*

We are finally ready to prove Theorem 6.5.

Proof of Theorem 6.5. Fix $\nu \in \mathbb{N}^*$, define $\hat{\boldsymbol{z}}^\nu \overset{\text{def}}{=} \mathrm{P}(\boldsymbol{x}^\nu) \in L^2([0,1], \mathbb{R}^T)$. Then $\hat{\boldsymbol{z}}^\nu \in \boldsymbol{\mathcal{X}}(\mathcal{A})$ is an action profile in nonatomic aggregative game $G(\mathcal{A})$. By the definition of VWE (Definition 6.4), we have $\int_0^1 \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta),\ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta^\nu \rangle \,\mathrm{d}\theta \leqslant 0$.

Secondly, by the definition of SVWE, we have $\int_0^1 \langle \mathbf{h}_{\boldsymbol{x}^\nu}(\theta),\ \boldsymbol{x}_\theta^\nu - \boldsymbol{z}_\theta^\nu\rangle \,\mathrm{d}\theta \leqslant 0$ for any $\boldsymbol{z}^\nu \in \boldsymbol{\mathcal{X}}^\nu(A)$, where $\mathbf{h}_{\boldsymbol{x}^\nu}(\theta) = \nabla_1 f_\theta^\nu(\boldsymbol{x}_\theta^\nu, \boldsymbol{X}^\nu) = \nabla_1 f_i^\nu(\boldsymbol{x}_i^\nu, \boldsymbol{X}^\nu) \stackrel{\text{def}}{=} \mathbf{h}_{\boldsymbol{x}^\nu}(i)$, for each $\theta \in \Theta_i^\nu$ and each type $i \in \mathcal{I}^\nu$.

For all $i \in \mathcal{I}^\nu$ and $\theta \in \Theta_i^\nu$, we also have $\|\mathbf{h}_{\boldsymbol{x}^\nu} - \mathbf{g}_{\boldsymbol{x}^\nu}\|_2 \leqslant d_i^\nu$ by the definition of d_i^ν (cf. (6.13)).

Thirdly, $\|\boldsymbol{x}^\nu - \hat{\boldsymbol{z}}^\nu\|_2 \leqslant 2K_{\mathcal{A}} \max(D^\nu, \overline{\delta}^\nu)$ by Lemma 6.7.

With these two results, while noticing that $\boldsymbol{x}_\theta^\nu \leqslant R + \overline{\delta}^\nu$ for all θ by Lemma 6.3, one has:

$$\begin{aligned}
&\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{g}_{\boldsymbol{x}^\nu}(\theta),\ \boldsymbol{x}_\theta^* - \boldsymbol{x}_\theta^\nu\rangle \,\mathrm{d}\theta \\
&= \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta),\ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta^\nu\rangle \,\mathrm{d}\theta + \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta),\ \hat{\boldsymbol{z}}_\theta^\nu - \boldsymbol{x}_\theta^\nu\rangle \,\mathrm{d}\theta \\
&\quad + \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^\nu}(\theta) - \mathbf{h}_{\boldsymbol{x}^\nu}(\theta),\ \boldsymbol{x}_\theta^\nu - \boldsymbol{x}_\theta^*\rangle \,\mathrm{d}\theta + \int_\Theta \langle \mathbf{h}_{\boldsymbol{x}^\nu}(\theta),\ \boldsymbol{x}_\theta^\nu - \boldsymbol{x}_\theta^*\rangle \,\mathrm{d}\theta \\
&\leqslant 0 + \|\mathbf{g}_{\boldsymbol{x}^*}\|_2 \,\|\hat{\boldsymbol{z}} - \boldsymbol{x}^\nu\|_2 + \|\mathbf{g}_{\boldsymbol{x}^\nu} - \mathbf{h}_{\boldsymbol{x}^\nu}\|_2 \,\|\boldsymbol{x}^\nu - \boldsymbol{x}^*\|_2 + J^\nu \\
&\leqslant 2L_{\mathbf{f}}\, K_{\mathcal{A}} \max(D^\nu, \overline{\delta}^\nu) + (2M + \overline{\delta}^\nu)\overline{d}^\nu + J^\nu
\end{aligned} \tag{6.18}$$

where $J^\nu \stackrel{\text{def}}{=} \int_\Theta \langle \mathbf{h}_{\boldsymbol{x}^\nu}(\theta),\ \boldsymbol{x}_\theta^\nu - \boldsymbol{x}_\theta^*\rangle \,\mathrm{d}\theta = \sum_{i\in\mathcal{I}^\nu} \int_{\Theta_i^\nu} \langle \mathbf{h}_{\boldsymbol{x}^\nu}(i),\ \boldsymbol{x}_\theta^\nu - \boldsymbol{x}_\theta^*\rangle \,\mathrm{d}\theta$.

Next, for the VWE $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(\mathcal{A})$, let $\boldsymbol{y}^{*\nu} = \psi(\boldsymbol{x}^*) \in L^2([0,1], \mathcal{M})$ and $\boldsymbol{z}^{*\nu} \stackrel{\text{def}}{=} \mathrm{P}^\nu(\boldsymbol{y}^{*\nu}) \in \boldsymbol{\mathcal{X}}^\nu(\mathcal{A}^\nu)$, such that we have:

$$\begin{aligned}
J^\nu &= \sum_{i\in\mathcal{I}^\nu} \Big\langle \mathbf{h}_{\boldsymbol{x}^\nu}(i), \int_{\Theta_i^\nu} \boldsymbol{x}_\theta^\nu - \boldsymbol{x}_\theta^* \,\mathrm{d}\theta \Big\rangle \\
&= \sum_{i\in\mathcal{I}^\nu} \big\langle \mathbf{h}_{\boldsymbol{x}^\nu}(i), \mu_i^\nu(\boldsymbol{x}_i^\nu - \boldsymbol{y}^{*\nu}{}_i^\nu)\big\rangle \\
&= \sum_{i\in\mathcal{I}^\nu} \big\langle \mathbf{h}_{\boldsymbol{x}^\nu}(i), \mu_i^\nu(\boldsymbol{x}_i^\nu - \boldsymbol{z}^{*\nu}{}_i^\nu)\big\rangle + \sum_{i\in\mathcal{I}^\nu} \big\langle \mathbf{h}_{\boldsymbol{x}^\nu}(i), \mu_i^\nu(\boldsymbol{z}^{*\nu}{}_i^\nu - \boldsymbol{y}^{*\nu}{}_i^\nu)\big\rangle \\
&\leqslant 0 + (L_{\mathbf{f}} + \overline{d}^\nu)\left\|\boldsymbol{z}^{*\nu} - \boldsymbol{y}^{*\nu}\right\|_2 \leqslant (L_{\mathbf{f}} + \overline{d}^\nu) 2K_{\mathcal{A}} \max(D^\nu, \overline{\delta}^\nu)\,,
\end{aligned} \tag{6.19}$$

from the definition of SVWE for $\boldsymbol{x}^\nu$ and Proposition 6.2, the definition of $\overline{d}^\nu$ and Lemma 6.7.i).

Let us summarize by combining (6.18) and (6.19), and considering ν large enough such that $\overline{d}^\nu, \overline{\delta}^\nu \leqslant 1$:

$$\begin{aligned}
&\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\boldsymbol{x}^\nu}(\theta), \boldsymbol{x}_\theta^* - \boldsymbol{x}_\theta^\nu\rangle \,\mathrm{d}\theta \leqslant \Omega^\nu \\
&\qquad \text{with } \Omega^\nu \stackrel{\text{def}}{=} (4L_{\mathbf{f}} + 1)K_{\mathcal{A}} \max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu.
\end{aligned} \tag{6.20}$$

Last, using the monotonicity definitions (Definition 6.3):

- if G is strongly monotone with modulus α, then $\alpha \|\boldsymbol{x}^\nu - \boldsymbol{x}^*\|_2^2 \leqslant \Omega^\nu$;
- if G is aggregatively strongly monotone with modulus β, then $\beta\|\boldsymbol{X}^\nu - \boldsymbol{X}^*\|^2 \leqslant \Omega^\nu$,

leading to the results announced in Theorem 6.5. □

Remark 6.10. *The strong monotonicity of the nonatomic aggregative game G, either with respect to action profile or with respect to aggregate-action profile, is essential in this result. In contrast to finite-player games (cf. [Wan12a]), strict monotonicity is not enough to obtain such results using the same techniques. Indeed, since $L^2([0,1], \mathcal{M}^T)$ is only weakly compact, one cannot ensure that $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\boldsymbol{x}^\nu}(\theta), \boldsymbol{x}_\theta^* - \boldsymbol{x}_\theta^\nu\rangle \,\mathrm{d}\theta$ tends to $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\hat{\boldsymbol{z}}}(\theta), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta\rangle \,\mathrm{d}\theta$ in (6.20), where $\hat{\boldsymbol{z}}$ is an accumulation point of $(\boldsymbol{x}^\nu)_\nu$ in the weak topology.*

6.3.3 Construction of a sequence of finite-type approximating games

As seen in our previous results, a nonatomic player θ is characterized by two elements: her action set $\mathcal{X}_\theta$, and her gradient $\nabla_1 f_\theta$ defined from $\mathcal{M}^2$ to $\mathbb{R}^T$: $(\boldsymbol{x}, \boldsymbol{Y}) \mapsto \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})$.

Note that it is the gradient of the cost function $\nabla_1 f_\theta$, instead of the cost function f_θ itself, that characterizes a nonatomic player's type. For example, two players θ and ξ with $\mathcal{X}_\theta = \mathcal{X}_\xi$ and $f_\theta(\boldsymbol{x}, \boldsymbol{Y}) \equiv f_\xi(\boldsymbol{x}, \boldsymbol{Y}) + C$ where C is a strictly positive constant can be seen as identical in their behavior.

This section presents the construction of a sequence of finite-type approximating games for a given nonatomic aggregative game G in two particular cases: 1) the player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ is piecewise continuous (cf. Definition 6.7) and, 2) $\{\mathcal{X}_\theta, \theta \in \Theta\}$ and $\{f_\theta, \theta \in \Theta\}$ are respectively polytopes and functions parameterized by a finite number of real parameters. Those two cases are fairly general. As illustrated in Section 6.4, they emerge naturally when the nonatomic game comes from the modeling of a population described by parametric probability distributions, the main motivation for considering infinite nonatomic games.

Case 1: Piecewise Continuous Characteristics – Uniform Splitting

Definition 6.7 (Continuity of nonatomic player characteristic profile)**.** The player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ in nonatomic aggregative game G is *continuous* at $\theta \in \Theta$ if, for all $\varepsilon > 0$, there exists $\eta > 0$ such that: for each $\theta' \in \Theta$

$$|\theta - \theta'| \leqslant \eta \;\Rightarrow \begin{cases} d_H(\mathcal{X}_\theta, \mathcal{X}_{\theta'}) \leqslant \varepsilon \\ \sup_{(\boldsymbol{x},\boldsymbol{Y}) \in \mathcal{M}\times\mathcal{M}} \|\nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y}) - \nabla_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\| \leqslant \varepsilon \,. \end{cases} \tag{6.21}$$

If (6.21) is true for all θ and θ' on an interval $\Theta' \subset \Theta$, then the player characteristic profile is *uniformly continuous* on Θ'.

Assume that the player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ of nonatomic aggregative game G is piecewise continuous, with a finite number $(K+1)$ of discontinuity points $\sigma_0 = 0 \leqslant \sigma_1 < \sigma_2 < \cdots < \sigma_{K-1} \leqslant \sigma_K = 1$, and that it is uniformly continuous on (σ_k, σ_{k+1}), for each $k \in \{0, \ldots, K-1\}$. For $\nu \in \mathbb{N}^*$, define an ordered set of I_ν cutting points by

$$\{v_i^\nu, i = 0, \ldots, I^\nu\} := \left\{\tfrac{k}{\nu}\right\}_{0\leqslant k\leqslant \nu} \cup \{\sigma_k\}_{1\leqslant k\leqslant K} \,,$$

and the corresponding partition $(\Theta_i^\nu)_{i\in\mathcal{I}^\nu}$ of Θ by:

$$\Theta_i^\nu = [v_{i-1}^\nu, v_i^\nu) \text{ for } i \in \{1, \ldots, I^\nu - 1\} \,; \quad \Theta_{I^\nu}^\nu = [v_{I_\nu - 1}^\nu, 1].$$

Hence, $\mu_i^\nu = v_i^\nu - v_{i-1}^\nu$. Denote $\bar{v}_i^\nu = \frac{v_{i-1}^\nu + v_i^\nu}{2}$.

Proposition 6.4. *Let Assumptions 6.1 to 6.4 hold, and assume that* $\{\text{span } \mathcal{X}_\theta\}_{\theta\in\Theta}$ *has a finite number of elements. For* $\nu \in \mathbb{N}^*$*, consider the finite-type game* $G^\nu(\mathcal{A}^\nu)$ *with aggregative constraint* $\mathcal{A}^\nu \overset{\text{def}}{=} \mathcal{A}$*, set of types* $\mathcal{I}^\nu \overset{\text{def}}{=} \{1 \ldots I^\nu\}$*, where for each type* $i \in \mathcal{I}^\nu$*:*

$$\mathcal{X}_i^\nu \overset{\text{def}}{=} \mathcal{X}_{\bar{v}_i^\nu} \;\; \textit{and} \;\; f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) \overset{\text{def}}{=} f_{\bar{v}_i^\nu}\Big(\boldsymbol{x}, \boldsymbol{Y}\Big), \;\; \forall (\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M} \times \mathcal{M}.$$

Then $\big(G^\nu(\mathcal{A})\big)_\nu = \big(\mathcal{I}^\nu, \boldsymbol{\mathcal{X}}^\nu, \mathcal{A}, (f_i^\nu)_{i\in\mathcal{I}^\nu}\big)_\nu$ *is a sequence of finite-type approximating games of nonatomic aggregative game* $G(\mathcal{A})$*.*

Proof. Let us show the three points required by Definition 6.6 as follows.

i) Given an arbitrary $\varepsilon > 0$, there is a common modulus of uniform continuity η such that (6.21) holds for all the intervals (σ_k, σ_{k+1}). For ν large enough, one has, for each $i \in \mathcal{I}^\nu$, $\mu_i^\nu < \eta$ so that for all $\theta \in \Theta_i^\nu$, $|\bar{v}_i^\nu - \theta| < \eta$; hence $d_H\big(\mathcal{X}_\theta, \mathcal{X}_i^\nu\big) = d_H(\mathcal{X}_\theta, \mathcal{X}_{\bar{v}_i^\nu}) < \varepsilon$.

ii) According to the continuity property, for all $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$:

$$\big\|\nabla_1 f_i^\nu\,(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y}) \;-\; \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\big\| = \Big\|\nabla_1 f_{\bar{v}_i^\nu}\Big(\boldsymbol{x}, \boldsymbol{Y}\Big) \;-\; \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\Big\| < \varepsilon.$$

To be rigorous, one would we need to ensure span $\mathcal{X}_\theta$ to be the same for all $\theta \in \Theta_i^\nu$: if not, one can further divide Θ_i^ν into a finite number of groups so that players in each group have the same span $\mathcal{X}_\theta$. This is possible because $\{\text{span } \mathcal{X}_\theta\}_\theta$ is finite.

iii) By definition, $D^\nu = 0$. □

Case 2: Finite-dimensions Parameterized Characteristics – Meshgrid Approximation

Assume that the nonatomic aggregative game G satisfy two conditions:

(i) The feasible action sets are K-dimensional polytopes: there exists a constant real-valued $K \times T$ matrix $\boldsymbol{P}$, and a bounded mapping $\boldsymbol{b} : \Theta \to \mathbb{R}^K$, such that for any θ, $\mathcal{X}_\theta = \{\boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{P}\boldsymbol{x} \leqslant \boldsymbol{b}_\theta\}$, which is a nonempty, bounded, closed and convex polytope in $\mathbb{R}^T$.

(ii) There is a bounded mapping $\boldsymbol{s} : \Theta \to \mathbb{R}^l$ such that for any $\theta \in \Theta, f_\theta(\cdot,\cdot) = f(\cdot,\cdot;\boldsymbol{s}_\theta)$. Furthermore, for all $(\boldsymbol{x},\boldsymbol{Y}) \in \mathcal{M}^2$, $\nabla_1 f(\boldsymbol{x},\boldsymbol{Y};\cdot)$ is Lipschitz-continuous in $\boldsymbol{s}$ and with a Lipschitz constant L_3, independent of $\boldsymbol{x}$ and $\boldsymbol{Y}$.

Let us consider:

$$\underline{b}_k \overset{\text{def}}{=} \min_\theta b_{\theta,k},\ \ \overline{b}_k \overset{\text{def}}{=} \max_\theta b_{\theta,k}, \forall k \in \{1 \dots K\}$$
$$\text{and } \ \underline{s}_k \overset{\text{def}}{=} \min_\theta s_{\theta,k}\ ,\ \ \overline{s}_k \overset{\text{def}}{=} \max_\theta s_{\theta,k}, \forall k \in \{1 \dots l\}\ .$$

The characteristics of player θ are parameterized by a point $(\boldsymbol{b}_\theta, \boldsymbol{s}_\theta)$ in

$$C \overset{\text{def}}{=} \textstyle\prod_{k=1}^K [\underline{b}_k, \overline{b}_k] \times \prod_{k=1}^l [\underline{s}_k, \overline{s}_k]\ ,$$

a compact subset of $\mathbb{R}^{K+l}$.

Fix $\nu \in \mathbb{N}^*$, consider a uniform partition of the compact set C , obtained by dividing each dimension of this compact set into ν equal parts. Hence, the partition is composed of $I^\nu \overset{\text{def}}{=} \nu^{K+l}$ equal-sized subsets of C. The cutting points of the partition are $\underline{b}_{k,n_k} \overset{\text{def}}{=} \underline{b}_k + \frac{n_k}{\nu}(\overline{b}_k - \underline{b}_k)$ for $k \in \{1,\dots,K\}$, and $\underline{s}_{k,n_k} \overset{\text{def}}{=} \underline{s}_k + \frac{n_k}{\nu}(\overline{s}_k - \underline{s}_k)$ for $k \in \{1,\dots,l\}$, with $n_k \in \{0,\dots,\nu\}$. Let the set of *vectorial* indices, indexing the partition, be denoted by:

$$\Gamma^\nu \overset{\text{def}}{=} \{\boldsymbol{n} = (n_k)_{k=1}^{K+l} \in \mathbb{N}^{K+l} \,|\, n_k \in \{1,\dots,\nu\}\}\ .$$

Define the corresponding partition of the interval Θ: $\Theta = \dot{\bigcup}_{\boldsymbol{n}\in\Gamma^\nu} \Theta_{\boldsymbol{n}}^\nu$, where:

$$\Theta_{\boldsymbol{n}}^\nu \overset{\text{def}}{=} \Big\{\theta \in \Theta : b_{\theta,k} \in [\underline{b}_{k,n_k-1}, \underline{b}_{k,n_k}) \text{ for } 1 \leqslant k \leqslant K;\ s_{\theta,k} \in [\underline{s}_{k,n_k-1}, \underline{s}_{k,n_k}) \text{ for } 1 \leqslant k \leqslant l\Big\}.$$

To be rigorous, when $\underline{b}_{k,n_k} = \overline{b}_k$ or $\underline{s}_{k,n_k} = \overline{s}_k$, the parameter interval is closed at the right.

Finally, define the set of players $\mathcal{I}^\nu$ as the elements $\boldsymbol{n}$ in Γ^ν such that $\mu(\Theta_{\boldsymbol{n}}^\nu) > 0$.

Proposition 6.5. *For $\nu \in \mathbb{N}^*$, let the nonatomic finite-type game $G^\nu(\mathcal{A}^\nu)$ with an aggregative constraint $\mathcal{A}^\nu \overset{\text{def}}{=} \mathcal{A}$, set of types $\mathcal{I}^\nu \overset{\text{def}}{=} \{\boldsymbol{n} \in \Gamma^\nu : \mu(\Theta_{\boldsymbol{n}}^\nu) > 0\}$ and, for each type $\boldsymbol{n} \in \mathcal{I}^\nu$,*

$$\mathcal{X}_{\boldsymbol{n}}^\nu \overset{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{R}^T | \boldsymbol{P}\boldsymbol{x} \leqslant \textstyle\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta\, d\theta\}\ ,$$
$$f_{\boldsymbol{n}}^\nu(\boldsymbol{x},\boldsymbol{Y}) \overset{\text{def}}{=} \mu_{\boldsymbol{n}}^\nu f\big(\tfrac{1}{\mu_{\boldsymbol{n}}^\nu}\boldsymbol{x}, \boldsymbol{Y}; \tfrac{1}{\mu_{\boldsymbol{n}}^\nu}\textstyle\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{s}_\theta\, \mathrm{d}\theta\big), \quad \forall (\boldsymbol{x},\boldsymbol{Y}) \in \mu_i^\nu \mathcal{M} \times \mathcal{M}.$$

Then, under Assumptions 6.1 to 6.4, $(G^\nu(\mathcal{A}))_\nu = \big(\mathcal{I}^\nu, \boldsymbol{\mathcal{X}}^\nu, \mathcal{A}, (f_i^\nu)_{i\in\mathcal{I}^\nu}\big)_\nu$ is a sequence of finite-type approximating games of the nonatomic aggregative game $G(\mathcal{A})$ with an aggregative constraint.

Proof. Let us show the three properties required by Definition 6.6 as follows.

i) For each $\boldsymbol{n} \in \mathcal{I}^\nu$, $\mathcal{X}_{\boldsymbol{n}}^\nu = \Big\{\boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{P}\boldsymbol{x} \leqslant \frac{1}{\mu_{\boldsymbol{n}}^\nu}\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta\ \mathrm{d}\theta\Big\}$. Then, from Theorem 5.3 in Section 5.F, there is a constant C_0 such that, for each $\theta' \in \Theta_{\boldsymbol{n}}^\nu$, we have:

$$d_H\left(\mathcal{X}_{\theta'}, \mathcal{X}_{\boldsymbol{n}}^\nu\right) \leqslant C_0 \left\|\boldsymbol{b}_{\theta'} - \frac{1}{\mu_{\boldsymbol{n}}^\nu}\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta\, \mathrm{d}\theta\right\| \leqslant \frac{C_0}{\nu}\left\|\overline{\boldsymbol{b}} - \underline{\boldsymbol{b}}\right\|\ .$$

Hence, $\overline{\delta}^\nu$ tends to 0. Note that in this case, the assumption of $\{\text{span } \mathcal{X}_\theta\}_{\theta\in\Theta}$ being finite is naturally satisfied.

ii) For each $\boldsymbol{n} \in \mathcal{I}^\nu$ and each $\theta' \in \Theta^\nu_{\boldsymbol{n}}$, for all $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$, one has:

$$\begin{aligned}\|\nabla_1 f^\nu_{\boldsymbol{n}}(\boldsymbol{x}, \boldsymbol{Y}) \ - \ \nabla_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\| &= \left\|\nabla_1 f\left(\boldsymbol{x}, \boldsymbol{Y}; \tfrac{1}{\mu^\nu_{\boldsymbol{n}}} \textstyle\int_{\Theta^\nu_{\boldsymbol{n}}} \boldsymbol{s}_\theta \,\mathrm{d}\theta\right) \ - \ \nabla_1 f(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{s}_{\theta'})\right\| \\ &\leqslant L_3 \| \frac{1}{\mu^\nu_{\boldsymbol{n}}} \int_{\Theta^\nu_{\boldsymbol{n}}} \boldsymbol{s}_\theta \,\mathrm{d}\theta - \boldsymbol{s}_{\theta'}\| \leqslant \frac{L_3}{\nu} \|\overline{\boldsymbol{s}} - \underline{\boldsymbol{s}}\| \ ,\end{aligned}$$

by the Lipschitz continuity of $\nabla_1 f(\boldsymbol{x}, \boldsymbol{Y}; \cdot)$. Hence, $\overline{d}^\nu$ tends to 0.

iii) By definition, we have $D^\nu = 0$. □

Remark 6.11. *One can obtain similar results as Proposition 6.5 by considering the characteristics $(\boldsymbol{b}_\theta, \boldsymbol{s}_\theta)$ for an aribtrary $\theta \in \Theta^\nu_{\boldsymbol{n}}$ for each $\boldsymbol{n}$, instead of the average values $(\int_{\Theta^\nu_{\boldsymbol{n}}} \boldsymbol{b}_\theta \,\mathrm{d}\theta, \int_{\Theta^\nu_{\boldsymbol{n}}} \boldsymbol{s}_\theta \,\mathrm{d}\theta)$.*

Remark 6.12. *By construction, in both sequences above, the compacity and convexity of the feasibility sets $(\mathcal{X}_i)_i$ and the convexity and continuity of cost functions $(f_i)_i$ are naturally inherited from the properties assumed on $(\mathcal{X}_\theta)_\theta$ and $(f_\theta)_\theta$. This should often be the case when building a sequence of approximating games from a nonatomic game with an infinity of types.*

6.4 Illustration on a Smart Grid Example

In this section the results are derived on a simple example for illustration, in the framework stated in introduction. In this example, we will be able to compute explicitly the aggregate equilibrium of the infinite-type nonatomic game.

We suppose that the energy operator has access to the probability distribution of the amount of flexible energy in the $N = 30$ millions French households: let us assume that this distribution is uniform on $[0, E_{\max}]$ with $E_{\max} = 20$kWh (kiloWatthour), that is $\phi_E(E) = \frac{1}{E_{\max}}$ for $E \in [0, E_{\max}]$.

Then the quantile function (or inverse cumulative distribution function), scaled by the population size, is given by $E_\theta = F_E^{-1}(\theta) = \theta E_{\max} N$, for each $\theta \in \Theta = [0, 1]$. In this case, the action set mapping $\mathcal{X}_\cdot : \Theta \rightrightarrows \mathbb{R}^2$ is given by:

$$\forall \theta \in \Theta, \ \mathcal{X}_\theta = \left\{\boldsymbol{x}_\theta = (x_{\theta,O}, x_{\theta,P}) \in \mathbb{R}^2_+ \mid x_{\theta,O} + x_{\theta,P} = E_\theta\right\} \ , \tag{6.22}$$

which gives an infinity of different action sets. Let us consider, as said in the introduction, that there are two prices:

$$c_O(\boldsymbol{X}) = \frac{a_O}{N} X_O \text{ and } c_P(\boldsymbol{X}) = \frac{a_P}{N} X_P$$

for off peak and on peak periods, with $a_P > a_O$, that depend only on the aggregate energy on off peak period X_O (resp. on peak period X_P), or rather on the average energy that consumers ask on these periods. Thus, the cost function of each player θ is given, as in the example of public products game given by (6.6):

$$\forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta, \ f_\theta(\boldsymbol{x}_\theta) = x_{\theta,O} \times c_O(\boldsymbol{X}) + x_{\theta,P} \times c_P(\boldsymbol{X}) \ = \langle \boldsymbol{x}_\theta, \boldsymbol{c}(\boldsymbol{X})\rangle,$$

where $\boldsymbol{c} = (c_0, c_P)$. Hence, all players have the same cost function: the infinite number of types is only due to the infinite number of different action sets.

Owing to Proposition 6.1, the nonatomic game G obtained is aggregatively strongly monotone with modulus $\beta = \frac{a_O}{N}$. However, the game is not strongly monotone.

It turns out that on this toy example, one can directly compute the aggregate profile of the VWE, as the IDVI of Definition 6.4 asks to find $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ such that:

$$\begin{aligned} &\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta \geqslant 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \\ \iff &\int_\Theta \langle \boldsymbol{c}(\boldsymbol{X}^*), \boldsymbol{x}_\theta - \boldsymbol{x}^*_\theta \rangle \,\mathrm{d}\theta \geqslant 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \\ \iff &\langle \boldsymbol{c}(\boldsymbol{X}^*), \boldsymbol{X} - \boldsymbol{X}^* \rangle \geqslant 0, \quad \forall \boldsymbol{X} \in \overline{\mathcal{X}} \,. \end{aligned} \tag{6.23}$$

This simplification holds because, for each θ, $\mathbf{g}_{\boldsymbol{x}^*}(\theta)$ depends only on the aggregate $\boldsymbol{X}^*$ (which would not be the case for general nonlinear cost functions). As a result the VI obtained is of finite dimension.

In this example, the aggregate action set $\overline{\mathcal{X}}$ can also be characterized easily, although this would not be the case for arbitrary sets $(\mathcal{X}_\theta)_\theta$. In fact, as the aggregate flexible energy available is $E_{\mathrm{tot}} \stackrel{\mathrm{def}}{=} \int_\Theta E_\theta \,\mathrm{d}\theta = \frac{1}{2} N E_{\max}$, we obtain:

$$\overline{\mathcal{X}} = \left\{ (X_O, X_P) \in \mathbb{R}^2_+ \mid X_O + X_P = E_{\mathrm{tot}} \right\} . \tag{6.24}$$

Indeed, if $X_O + X_P = E_{\mathrm{tot}}$, then taking $\boldsymbol{x}_\theta = \boldsymbol{X} \frac{E_\theta}{E_{\mathrm{tot}}}$, we have $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ and $\int_\Theta \boldsymbol{x}_\theta \,\mathrm{d}\theta = \boldsymbol{X}$. The converse inclusion is clear. Consequently, we obtain from (6.23) that $\boldsymbol{X}^*$ is the solution of the quadratic program:

$$\begin{aligned} &\min_{\boldsymbol{X}} \tfrac{a_O}{N} \times \tfrac{1}{2} X_O^2 + \tfrac{a_P}{N} \times \tfrac{1}{2} X_P^2 \\ &X_O + X_P = E_{\mathrm{tot}} \\ &0 \leqslant X_O, X_P \end{aligned}$$

that is: $\boldsymbol{X}^* = (X^*_O, X^*_P) = (\frac{a_P}{a_O + a_P} E_{\mathrm{tot}}, \frac{a_O}{a_O + a_P} E_{\mathrm{tot}})$.

Now, let us define a sequence of finite-type approximating games G^ν to approximate G, with for each $\nu \in \mathbb{N}^*$, $I^\nu = \nu$. Let us drop the index ν for simplicity in the remaining. Let us split up the population uniformly with $\Theta^\nu_i = [\frac{i-1}{I}, \frac{i}{I}]$, for each $i \in \mathcal{I} = \{1, \dots, I\}$.

Because of the linearity of $\theta \mapsto E_\theta$, considering the uniform approximation detailed in Section 6.3.3 case 1, one will obtain directly $\boldsymbol{X}^*$. For the example, let us rather consider the approximating games defined with, for each $i \in \mathcal{I}$:

$$\mathcal{X}_i \stackrel{\mathrm{def}}{=} \{ \boldsymbol{x}_i \in \mathbb{R}^2_+ \mid x_{i,O} + x_{i,P} = E_i \stackrel{\mathrm{def}}{=} \tfrac{i}{I} N E_{\max} \} . \tag{6.25}$$

Besides, we naturally take $f_i \stackrel{\mathrm{def}}{=} f_\theta$ for each i (as the cost function is the same for each player).

One can observe that we get for each i, $\delta_i = \frac{N E_{\max}}{I} = \frac{2E_{\mathrm{tot}}}{I} \to 0$, and of course $d_i = 0$.

On the other hand, computing the aggregate approximate equilibrium, similarly to (6.23), one obtains:

$$\boldsymbol{X}^I = \left(\tfrac{a_P}{a_O + a_P} E_{\mathrm{tot}} (1 + \tfrac{1}{I}), \tfrac{a_O}{a_O + a_P} E_{\mathrm{tot}} (1 + \tfrac{1}{I}) \right) = (1 + \tfrac{1}{I}) \boldsymbol{X}^* \,,$$

and thus we have:

$$\|\boldsymbol{X}^I - \boldsymbol{X}^*\| = \frac{\|\boldsymbol{X}^*\|}{I} = \frac{\sqrt{a_O^2 + a_P^2}}{a_O + a_P} E_{\mathrm{tot}} \times \frac{1}{I}. \tag{6.26}$$

However, from Theorem 6.5, as we can compute $L_{\mathbf{f}} = \max_{\boldsymbol{X} \in \overline{\mathcal{X}}} \|\boldsymbol{c}(\boldsymbol{X})\| = \frac{a_P}{N} E_{\mathrm{tot}}$, we obtain the more conservative convergence bound (we can replace $(4L_{\mathbf{f}} + 1)$ by $2L_{\mathbf{f}}$ since $D^\nu = \overline{d}^\nu = 0$):

$$\begin{aligned} &\|\boldsymbol{X}^I - \boldsymbol{X}^*\|^2 \leqslant \frac{1}{\beta} 2 L_{\mathbf{f}} \overline{\delta}^I = \frac{N}{a_O} 2 \frac{a_P}{N} E_{\mathrm{tot}} \times \frac{2E_{\mathrm{tot}}}{I} \\ \iff &\|\boldsymbol{X}^I - \boldsymbol{X}^*\| \leqslant 2E_{\mathrm{tot}} \sqrt{\frac{a_P}{a_O}} \times \frac{1}{\sqrt{I}} \,. \end{aligned}$$

6.5 Conclusion

Theorem 6.5 provides a precise theoretical result for the use of symmetric finite-dimensional (variational) Wardrop equilibria (S(V)WE) as an approximation of the (V)WE in a strongly monotone or aggregatively strongly monotone nonatomic aggregative game with an infinity of players types, with or without aggregative constraints. There are numerous research themes related to this result and our topic in general.

First, one needs to find efficient algorithms for the computation of finite dimensional variational inequalities arising as the characterization of SVWE. An extensive literature exists in this regard but our particular case of aggregative game with aggregative constraints may lead to special methods or improvements on existing results [Gra17].

Then, the extension of evolutionary dynamics for population games and related algorithms to nonatomic games with infinitely many classes of players can be non trivial. A recent work [Had17] proposes online learning methods for population games with heterogeneous convex action sets. The presence of aggregate constraints adds two additional difficulties for considering evolutionary dynamics in population games, as those dynamics are based on unilateral adaptations from players. On the one hand, in the presence of coupling constraints, unilateral deviations by players may well lead to an action profile violating the coupling constraint. On the other hand, a feasible deviation in the action profile cannot always be decomposed into unilateral deviations of players.

Last, our results are limited to monotone games and the convergence result is limited to strongly monotone games. The study of nonatomic aggregative games that are not monotone needs probably other approaches. Indeed, even for population games where there are only finitely many types of players, there exist much fewer results for games that are not linear, potential or monotone. The question of whether or not it is possible to obtain similar convergence results as those stated in Theorem 6.5 without monotonicity assumptions constitutes an interesting path for future work.

Appendix

6.A Extension of results to the subdifferentiable case

In this section, we explain briefly how our results extend to the case of convex nonsmooth cost functions, considering subdifferential instead of gradients of convex costs.

The essence of the proofs are roughly the same as in the smooth case, but considering subdifferentials requires some additional technical arguments. The full proofs can be found in [JW18a]. The authors decided to formulate the results in the smooth case so that the key arguments and ideas appear clearly.

Recall that the *subdifferential*, i.e. set of *subgradients* of a convex function f at $\boldsymbol{x} \in \mathbb{R}^T$ in its domain C, which is a convex set in $\mathbb{R}^T$, is denoted by $\partial f(\boldsymbol{x})$. Recall that $\mathbf{g} \in \mathbb{R}^T$ is a subgradient of f at $\boldsymbol{x}$, denoted $\mathbf{g} \in \partial f(\boldsymbol{x})$, iff for all $\boldsymbol{z} \in C$, $f(\boldsymbol{z}) \geqslant f(\boldsymbol{x}) + \langle \mathbf{g}, \boldsymbol{z} - \boldsymbol{x} \rangle$.

One has to consider the correspondence $\mathcal{H} : L^2([0,1], \mathcal{M}) \rightrightarrows L^2([0,1], \mathbb{R}^T)$ of the subdifferentials, associating to each profile $\boldsymbol{x} \in L^2([0,1], \mathcal{M})$ and each player θ the set of subgradients of her cost functions:

$$\mathcal{H}(\boldsymbol{x}) \stackrel{\text{def}}{=} \{\mathbf{g} = (\mathbf{g}_\theta)_{\theta \in \Theta} \mid \mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x}),\ \forall a.e. \theta \in \Theta\}, \quad \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}). \tag{6.27}$$

In other words, $\mathcal{H}(\boldsymbol{x})$ is the collection of measurable (and integrable because of Assumption 6.3.iii adapted to assume uniform boundedness of $\partial_1 f_\theta$) selections of a subgradient for each $\boldsymbol{x}_\theta$. Most of the chapter can be interpreted in the nonsmooth framework by

- replacing $\mathbf{g}_{\boldsymbol{x}}$ by an element of $\mathcal{H}(\boldsymbol{x})$ in the equations;

- considering the Hausdorff distance d_H between subdifferentials instead of the Euclidean distance between two gradients (e.g. for d in Definition 6.6 of a sequence of finite-type approximating games);

• a direct implication is that we have to consider *generalized* variational inequalities (GVI), finite or infinite-dimensional, instead of VIs.

For instance, $\langle \mathbf{g}_{x^*}, x^* - x \rangle \leqslant 0, \forall x \in \mathcal{X}$ becomes $\exists \mathbf{g} \in \mathcal{H}(x^*),\ \langle \mathbf{g}, x^* - x \rangle \leqslant 0, \forall x \in \mathcal{X}$.

It is usefull to introduce the best-reply correspondence $Br : \overline{\mathcal{X}} \rightrightarrows \mathcal{X}$:

$$Br(Y) \stackrel{\text{def}}{=} \{x \in \mathcal{X} : x_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, Y), \forall \theta \in \Theta\}, \quad \forall\, Y \in \overline{\mathcal{X}}\,,$$

and, for $Y \in \overline{\mathcal{X}}$ and $x \in Br(Y)$, the correspondence $\mathcal{D}(x, Y)$: $\Theta \to \mathbb{R}^T$ defined by:

$$\mathcal{D}(x, Y)(\theta) \stackrel{\text{def}}{=} \{\mathbf{g}_\theta \in \partial_1 f_\theta(x_\theta, Y) \mid \langle \mathbf{g}_\theta, z_\theta - x_\theta \rangle \geqslant 0\,, \quad \forall z_\theta \in \mathcal{X}_\theta\} \ \ \forall \theta \in \Theta$$

which is nonempty (by first order conditions) and closed-valued. To get similar results as in the smooth case, we need to make the following additional assumption:

Assumption 6.6. *For all $Y \in \overline{\mathcal{X}}$ and all $x \in Br(Y)$, $\mathcal{D}(x, Y)$ is a measurable correspondence.*

One can show that $\mathcal{H}(.)$ and Br have nonempty values. Then, instead of Lemma 6.1, we use the compact-valued selection theorem [Aum76] to obtain for each $x \in Br(Y)$ the existence of a measurable mapping $\theta \mapsto \mathbf{g}_x(\theta)$ such that $\forall \theta \in \Theta, \mathbf{g}_x(\theta) \in \mathcal{D}(x, Y)(\theta)$.

We can then obtain a characterization of WE similar to Theorem 6.1, where the differential $\mathbf{g}_{x^*}$ is replaced by the existence of an element in $\mathcal{H}(x^*)$. The existence result in Theorem 6.2 is also valid in the subdifferentiable case.

The monotonicity of G is defined as in Definition 6.3, where the inequalities on $\mathbf{g}_x$ and $\mathbf{g}_y$ now have to hold for each pair of elements of the correspondences $(\mathbf{g}_x, \mathbf{g}_y) \in \mathcal{H}(x) \times \mathcal{H}(y)$. Properties characterizing monotonicity given in Proposition 6.1 follow with essentially the same proof, having in mind that for any θ, $\partial_1 f_\theta(x_\theta, Y) = \{c(Y) + \mathbf{g} : \mathbf{g} \in \partial(-u_\theta)(x_\theta)\}$.

In presence of coupling constraints, a VWE (Definition 6.4) is also defined by $x^* \in \mathcal{X}(\mathcal{A})$ and the existence of an element $\mathbf{g} \in \mathcal{H}(x^*)$ satisfying the infinite dimensional GVI (6.7).

The existence of a VWE (similar to Theorem 6.3) can also be obtained in the nonsmooth case, where continuity of the gradient is replaced by upper-hemicontinuity of the correspondence $\mathcal{H}$ and applying results of [DT96, Corollary 2.1]. The uniqueness conditions associated to monotonicity detailed in Theorem 6.4 follow as well with essentially the same proof as for the smooth case.

The main result, Theorem 6.5, is obtained for the nonsmooth case with the same bounds on the convergence rate.

6.B On the Relationship between Nash and Wardrop Equilibria

The objective of this chapter is to approximate the equilibrium of a nonatomic game with an infinity of players types, by considering approximating games with a finite number of players types.

A natural idea would also be to consider approximating games with a finite number of *players*, this number of players growing to infinity to approximate the nonatomic population game.

Indeed, we briefly explain in this appendix how we can obtain similar convergence results by adopting this approach considering finite-player atomic games. The approach is fully developed in [JW18a].

The main difference and difficulty under this approach is that the equilibrium concept to consider for finite-player games is no longer Wardrop Equilibrium, but Nash Equilibrium (NE). As the number of players is finite, an individual action x_i of a player i does have an impact on the aggregate action $X = \sum_i x_i$.

As a result, the modified cost function $\hat{f}_i : (x_i, X_{-i}) \mapsto f_i(x_i, X_{-i} + x_i)$ naturally appears, where the action of i is taken into account in the aggregate action.

This modified cost function, and the impact of individual actions in general, have to be considered both in the assumptions and in the definitions of the different concepts used in this chapter.

Nash equilibrium are naturally characterized by finite-dimensional variational inequalities under convexity hypotheses: in the atomic case, we need the additional following assumption:

Assumption 6.7. *For an atomic game $\mathcal{G}(\mathcal{A}) = (\mathcal{I}, (f_i)_i, (\mathcal{X}_i)_i, A)$ with a finite set of players $\mathcal{I}$ and cost functions $(f_i)_i$, the associated functions $\left(\hat{f}_i(., \boldsymbol{X})\right)_i$ are convex.*

Note that this convexity is not necessarily implied by the convexity of $f_i(., \boldsymbol{X})$. Under this additional assumption, we obtain a GVI (where G stands for *generalized* in the nonsmooth case, see Appendix 6.A) characterization of NE, similar to the one for SVWE given in Proposition 6.2, and an existence result:

Definition 6.8 (Variational Nash Equilibrium (VNE), [Har91])**.** A variational Nash equilibrium of atomic game $\mathcal{G}(\mathcal{A})$ is a solution to the following GVI problem:

$$\text{Find } \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A}) \text{ s.t. } \exists\, \mathbf{g} \in \hat{H}(\boldsymbol{x}) \text{ s.t. } \left\langle \mathbf{g}, \boldsymbol{x} - \boldsymbol{x} \right\rangle \geqslant 0,\ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\mathcal{A}). \tag{6.28}$$

where the subgradients correspondence $H : \boldsymbol{\mathcal{X}} \rightrightarrows \mathbb{R}^{IT}$ is given as:

$$\forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}},\ \hat{H}(\boldsymbol{x}) \stackrel{\text{def}}{=} \{(\mathbf{g}_i)_{i\in\mathcal{I}} \in \mathbb{R}^{IT} : \mathbf{g}_i \in \partial_1 \hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i}),\ \forall i \in \mathcal{I}\} = \prod_{i\in\mathcal{I}} \partial_1 \hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i})\ .$$

In particular, if $\overline{\mathcal{X}} \subset A$, a VNE is a NE .

Proposition 6.6 (Existence of VNE)**.** *Under Assumptions 6.1 and 6.3 (compacity and convexity) on $(\mathcal{X}_i)_i$ and Assumption 6.7, the atomic game $\mathcal{G}(\mathcal{A}) = (\mathcal{I}, (f_i)_i, (\mathcal{X}_i)_i, A)$ admits a VNE.*

To obtain a convergence result of the (V)NEs in a sequence of atomic games, we need some stronger properties than for the sequence of finite-type approximating games. In addition to Definition 6.6, we assume that:

1. the number of players tends to infinity: $I^\nu \underset{\nu\to\infty}{\longrightarrow} \infty$;
2. each player becomes infinitesimal: $\mu_i = \mu(\Theta_i) \underset{\nu\to\infty}{\longrightarrow} 0$;
3. in the gradient (or subdifferential) of a player, the impact of her own action on the aggregate profile vanishes along the sequence, by considering the additional parameter (given in the subdifferential case):

$$\lambda_i^\nu \stackrel{\text{def}}{=} \sup_{(\boldsymbol{x},\boldsymbol{Y})\in\mathcal{M}^2} \sup_{\mathbf{g}\in\partial_1 \hat{f}_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y} - \mu_i^\nu \boldsymbol{x})} d\left(\mathbf{g}, \partial_1 f_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y})\right) \underset{\nu\to\infty}{\longrightarrow} 0 \tag{6.29}$$

Then one obtains similar convergence result as Theorem 6.5, with the only difference being that:

- in the upper bound, $\overline{d}^\nu$ is replaced by $(\overline{d}^\nu + \overline{\lambda}^\nu)$, where $\overline{\lambda}^\nu \stackrel{\text{def}}{=} \max_i \lambda_i^\nu$),
- for the convergence in $L_2([0,1], \mathcal{M})$, one has to consider a projection of the (V)NE on $L_2([0,1], \mathcal{M})$, where each $\theta \in \Theta_i$ is associated to the action of $i \in \mathcal{I}$.

Note that this approach has also another interest, as the convergence theorem that we obtain in this case can be interpreted in the reverse way: under the right assumptions, in a sequence of atomic games converging (in the sense given by the definition of finite-type approximating games) to a nonatomic aggregative game, the sequence of Nash equilibria converge to an equilibrium (the SVWE) of this limit nonatomic game.

In some particular cases where the continuous SVWE can be computed explicitly as a function $\boldsymbol{x} : \Theta \to \mathcal{M}$ (for instance in the example derived in Section 6.4), this would give an approximation of a Nash equilibrium in the atomic (reality) aggregative game, as the finite number of atomic players is very large.

Part IV

Decentralized Energy Exchanges in a Peer to Peer Framework

The practical framework we considered in this Part IV differs slightly from the *demand response* framework presented in the previous parts of the thesis, although the reader will find some similarities in the mathematical tools and notions considered in the following Chapter 7 (Generalized Nash Equilibrium, Price of Anarchy, etc.).

Part II presented the analysis and tools related to the management of flexibilities in a decentralized context (local individual consumers optimizing a signal and interacting in a non-cooperative game). In this context of demand response addressed in Part II, we considered the point of view of a global operator or aggregator in charge of the electric system.

The point of view adopted in this Part IV is different. We go one step further in the decentralization paradigm, and consider the possibility for individual energy consumers (or *prosumers* i.e. consumers with a small local renewable production) to self-manage—without the coordination of a central operator—their local energy community and engage in local and direct trades with their neighbors or within a community of other consumers. This is what we referred to as *peer-to-peer electricity market*, as opposed to conventional electricity markets and the historical need for individual consumers to depend on an electricity provider to obtain energy.

In the following Chapter 7, we consider a network of prosumers involved in peer-to-peer energy exchanges, with differentiation price preferences on the trades with their neighbors, and we analyze two market mechanisms:

(i) a centralized market which will be used as a benchmark, where a central operator optimizes the flows (trades) between the nodes, local demand and flexibility activation to maximize the system overall social welfare;

(ii) a distributed peer-to-peer market design where prosumers in local energy communities optimize selfishly their trades, demand, and flexibility activation. We first characterize the solution of the peer-to-peer market as a Variational Equilibrium and prove that the set of Variational Equilibria coincides with the set of social welfare optimal solutions of market design **(i)**.

We give several results that help understanding the structure of the trades at an equilibrium or at the social welfare optimum. We characterize the impact of preferences on the network lines congestion and renewable energy surplus under both designs.

We provide a reduced example for which we give the set of all possible generalized equilibria, which enables to give an approximation of the price of anarchy. We provide a more realistic example which relies on the IEEE 14-bus network, for which we can simulate the trades under different preference prices. Our analysis shows in particular that the preferences have a large impact on the structure of the trades, but that one equilibrium (variational) is socially optimal.

Chapter 7 is based on the journal paper [LC+19b], accepted for publication.

Chapter 7

Peer-to-Peer Electricity Market Analysis: from Variational to Generalized Nash Equilibrium

This chapter is based on the paper [LC+19b], accepted for publication.

7.1 Introduction

New regulations are restructuring electricity markets in order to build the grid of the future. Instead of a centralized market design where all the operations are managed by a global central market operator [MVV18; Sch+13; Sto02], new *decentralized* models have emerged. These models involve local energy communities which can trade energy, either by the intermediate of a global market operator [LC18], or in a peer-to-peer setting [PS16; Sou+19]. Peer-to-peer energy trading allows flexible energy trades between peers, where, for instance, local prosumers exchange between them energy surplus from multiple small-scale distributed energy resources (DERs) [Liu+17; Lon+17].

Significant value is brought to the power system by coordinating local renewable energy source (RES)-based generators and DERs to satisfy the demand of local energy communities, since it decreases the need for investment in conventional generation and transmission networks. Also, thanks to the decreasing feed-in-tariffs, using RES-based generations on site (e.g., at household level, within the microgrid) is more attractive than feeding it into the grid, because of the difference between electricity selling and buying prices [Lon+17]. Peer-to-peer energy trading encourages the use of surplus energy within local energy communities, resulting in significant cost savings even for communities with moderate penetration of RES [Lon+17].

In practice, the radial structure of the distribution grid calls for hierarchical market designs, involving transmission and distribution network operators [LCMP19]. Nevertheless, various degrees of coordination can be envisaged: full coordination organized by a global market operator (transmission system operator), bilateral contract networks [MTM18], fully decentralized market designs allowing peer-to-peer energy trading between the prosumers in a distributed fashion [Mor+18; SBP18] or, still, within and between coalitions of prosumers, called community or hybrid peer-to-peer [MP18]. A community-based organization involves a community manager which organizes trades among the community and is in charge of the interactions with the rest of the market. A distributed market structure exists when the decentralized elements explicitly share, in a peer-to-peer fashion, local information, resulting in a system in which all the elements may not have access to the same level of information. This information asymmetry might create differences in valuations of the traded resource (e.g., price arbitrage) and result in market imperfections, implying that the prices associated with the bilateral trading of resource allocation between couples of agents do not coincide. This price gap can be interpreted as a bid-ask spread due to a lack of liquidity in the market [Ogg+12].

Energy exchange between production units and local demand of energy communities are formulated as a symmetric assignment problem. Its solution relies on two main streams

in the literature. The first stream deals with matching models which put in relation RES-based generators and consumers by the intermediate of a platform, with various consumers classes and different possible objective functions for the platform operator [Liu+17]. The second stream combines multi-agent modeling, as well as classical distributed optimization algorithms which are applied to solve the assignment problem in real-time [Mor+18; MM18; SBP18]. Auctions theory can be used, in addition, to schedule the DER commitment in day-ahead.

7.1.1 Matching Models for Peer-to-Peer Energy Trading

In the energy sector, peer-to-peer energy trading is a novel paradigm of power system operation. There, prosumers provide their own energy from solar panels, storage technologies, demand response mechanisms, and they exchange energy with one another in a distributed fashion. Zhang et al. provide in [Zha+17] an exhaustive list of projects and trails all around the world, which build on new innovative approaches for peer-to-peer energy trading. A large part of these projects rely on platforms, understood as two-sided markets, that match RES-based generators and consumers according to their preferences and locality aspects (e.g. *Piclo* in the UK, *TransActive Grid* in Brooklyn, US, *Vandebron* in the Netherlands, etc.). In the same vein, cloud-based virtual market places, which deal with excess generation within microgrids, are developed by *PeerEnergyCloud* and *Smart Watts* in Germany. Some other projects rely on local community building for investment sharing in batteries, solar PV panels, etc., in exchange for bill reduction or a certain level of autonomy with respect to the global grid (e.g. *Yeloha* and *Mosaic* in the US, *SonnenCommunity* in Germany, etc.). How other components of the platform's design can influence the nature and the preference of the prosumers involved is also studied in the literature. Typical elements of the platform's design are: the impact of pricing mechanism (e.g. setting one common market price versus individual prices per transaction set—for instance through auction design—or per class of prosumers), the platform's objective (e.g. maximizing the social welfare versus maximizing the platform's benefit), the influence of the platform's commission per transaction. For example, in [Ben+18], the authors study the impact of the price of the goods exchanged on the level of collaboration and also on the level of ownership among participants. In [FHW17], the impact of different platform's objective functions is analyzed considering a set of heterogeneous renters and owners. Dynamic pricing for operations of the platform based on supply and demand ratio of shared RES-based generation is investigated in [Liu+17]. Peer-to-peer organizations are also a way to enable small and flexible actors to enter markets by lowering the entrance barrier [EFL16].

Platform design constitutes an active area of research in the literature on two-sided markets [EFL16; FHW17]. Three needs are identified for platform deployment. Firstly, it should help buyers and sellers find each other, while taking into account the heterogeneity in their preferences. This requires the platform to find a trade-off between low entry cost and information retrieval from big, heterogeneous and dynamic information flows. Buyers' and sellers' search can be performed in a *centralized* (e.g. Amazon, Uber), effective *decentralized* (e.g. Airbnb, eBay), or even *fully distributed* (OpenBazaar, Arcade City) manner. Secondly, the platform must set prices that balance demand and supply, and ensure that prices are set competitively in a decentralized fashion. Finally, the platform ought to maintain trust in the market, relying on reputation and feedback mechanisms [FGV18]. Sometimes, supply might be insufficient so that subsidies need to be designed to encourage sharing on the platform [FHW17].

7.1.2 Distributed Optimization Approaches

Computational and communication bottlenecks have largely been alleviated by recent work on distributed and peer-to-peer optimization of large-scale optimal power flow [EAD16; Kra+14; PL14]. Mechanisms for the optimization of a common objective function by a decentralized system are known as decomposition-coordination methods [PC06]. In such methods, a centralized (large-scale) optimization problem is typically split into small-size local optimization problems whose outputs are coordinated dynamically by a central agent

(called "master") so that the overall objective of the system becomes aligned (after a certain number of iterations) with the (large-scale) centralized optimization problem outcome. Following this stream, a consensus-based Alternating Direction Method of Multipliers, or ADMM algorithm, is implemented in [MM18; SBP18; Sma] to approximate the optimal solution which maximizes the prosumers social welfare, in a peer-to-peer electricity market. Similar approaches relying on dual decomposition, which iteratively solves the problem in a distributed manner with limited information exchange, were implemented for energy trading between islanded microgrids in [GM14; MGD12]. Two main drawbacks of these algorithmic approaches are listed in [SBP18]: first, they do not take into account the strategic behaviors of the prosumers; second, they are computationally limited, which might constitute a blocking point when studying large-scale peer-to-peer networks. The latter issue is overcome in [Mor+18] with an improved consensus algorithm.

In addition, these distributed-optimization approaches enable incorporating heterogeneous energy preferences of individual prosumers in network management. In [MM18], the authors evaluate the added value of multi-class prosumer energy management in a distribution network that has a "green prosumer", a "philanthropic prosumer" and a "low-income household". Three energy classes are introduced to account for the prosumers' preferences: "green energy", "subsidized energy" and "grid energy". A platform agent is introduced to act as an auctioneer, allowing energy trading between the prosumers and the wholesale electricity market. The platform agent sets the price of each energy class in the distribution network. The tool of receding-horizon model predictive control is used to provide a real-time implementation. Consumer preferences are also introduced in [SBP18] in the form of product differentiation prices. They can either be pushed centrally as dynamic and specific tax payments, or be used to better describe the utility of the consumers who are willing to pay for certain characteristics of trades.

7.1.3 Privacy Issues

From the perspective of information and communication technology (ICT), a fully decentralized market design provides a robust framework since, if one node in a local market is attacked or in case of failures, the whole architecture should remain in place, while information could find other paths to circulate from one point to another, avoiding malicious nodes and corrupted paths.

From an algorithmic point of view, the implementation of a fully distributed market design might be challenging, since it has to deal with far more complex communication mechanisms than the centralized market design. Efficient communication will allow collaboration among prosumers, so that energy produced by one can be utilized by another in the network. Multiple peer-to-peer communication architectures exist in the literature, including structured, unstructured and hybrid ones. They are all based on common standards for the communication network operation, which are measured through latency, throughput, reliability and security [Jog+17]. In addition to the large size of the communication problem, privacy issues may also directly impact the market outcome. Indeed, if prosumers are allowed to keep some private information, then they might not have access to the same level of information, i.e. information asymmetry appears. Since the prosumers make decisions based on the information at their disposal, such asymmetry can introduce bias in the market outcome. To avoid or, at least, to limit bias introduced in the market outcome while guaranteeing the optimum of the social welfare, various algorithms that preserve local market agents' privacy have been discussed in the literature. For example, the algorithms can require the agents to update no more than their dual variables – e.g., local prices [EAD16; SBP18]. Of course, the efficiency of these algorithms depends on the level of privacy defined by the agents as well as which private information could be inferred from the released values.

7.1.4 Contributions

The peer-to-peer structure adopted in this work is different from the approaches involving decomposition-coordination methods. The latter approaches require for example to exchange Lagrangian multipliers updated at each iteration of the decentralized clearing

[MGD12; Mor+18; MP18; SBP18], that can be used by the coordinator to infer some information about the preferences of the peers. Such approaches have therefore two main drawbacks at the market level: first, for each market clearing, it requires in general a large number of iterations to reach an optimum—such latency in the clearing price computation might be difficult to allow from the point of view of market operators. Besides, it offers limited privacy guarantees as the market operator can infer private information from the peers under repeated interactions. In this chapter, we assume that there is no central authority coordinating the exchanges (in quantity, price and information) between the nodes. Within this framework, strategic communication mechanisms can appear, and nodes have the possibility to self-organize into coalitions or local energy communities, as reviewed in [Tus+18]. With such strategic behaviors, the equilibrium of the peer-to-peer market design might not coincide with the social welfare global optimum achieved with full coordination of the nodes by a "master" controlling all the information and decisions, as in [Wan+14] where the authors consider a noncooperative game involving storage units.

In this chapter, we first characterize the solution of a peer-to-peer electricity market as a Variational Equilibrium, assuming that all the agents have equal valuation of the price associated with the traded resource. We prove that the set of Variational Equilibria coincides with the set of social welfare optima. However, in a fully-distributed setting, it is very unlikely that each couple of agents coordinate on their valuations of the trading price. As a result, imperfections appear in the market, which we capture by considering Generalized Nash Equilibrium solutions as possible outcomes. We characterize analytically the impact of preferences on the network line congestion and energy surplus, both under centralized and peer-to-peer market designs. Our results are illustrated in two test cases (a three node network with arbitrage opportunity and the standard IEEE-14 bus network). We evaluate the loss of efficiency caused by peer-to-peer market imperfections in the three nodes network, with the Price of Anarchy as a performance measure. We also evaluate numerically the impact of the differentiation prices by computing the equilibria of our 14 nodes network under different price configurations.

For ease of reading, we also reference the link between the main results we obtain (summarized through propositions in the course of the text) below:

- Under *centralized market design*, we derive in Proposition 7.1 analytical expressions for the demand, flexibility activation and net import at the optimum, as linear functions of the nodal prices, at each node of the network.

- By substitution of these results at the optimum in the balancing equation, we observe that there might be energy surplus in the local energy community. We derive in Proposition 7.2 a necessary condition on technologies and RES generation to avoid energy surplus.

- This condition being not sufficient, we identify in Proposition 7.3 conditions on the nodal prices and preferences such that no congestion, and then no energy surplus, appears in the local distribution network. In Proposition 7.5, this result is extended by highlighting the link between the possible line congestions and the occurrence of strictly positive or negative cycles in the matrix of the differences of the trade preferences.

- In Proposition 7.4, we obtain analytical expressions of the nodal prices at the root node and at each node of the distribution grid.

- Under *decentralized market design*, assuming a complete market, we prove in Proposition 7.6 that the set of Variational Nash Equilibria (VNE), whose definition is recalled in Definition 7.2 coincides with the set of social welfare optima solutions of the centralized market clearing.

- However, there is no guarantee that there exists a market to determine the price system associated with the bilateral trade reciprocity constraints. In case it does not exist, the peer-to-peer market would become incomplete and the bilateral trade prices between any couple of nodes might diverge. We reformulate the Generalized Nash Equilibrium (GNE) problem (the notion of GNE being recalled in Definition 7.1) as an optimization problem

applying a parametrized variational inequality approach, enabling the computation of Generalized Nash Equilibria via a sampling method and a standard optimization algorithm, in Proposition 7.7.

- Making the parallel with the centralized market design results, we capture the impact of the capacity of the lines, preferences and structure of the matrix of the preference reciprocity gaps, on line congestion in Propositions 7.8, 7.9.

The chapter is organized as follows. In Section 7.2, we introduce the model of the generalized noncooperative game we consider in this work, and we give our main assumptions. In Section 7.3, the centralized market design **(i)** is formulated and its solutions characterized. We introduce the peer-to-peer market design **(ii)** in Section 7.4; its solutions are characterized in terms of VNE and GNE in the presence of market incompleteness. Congestion analysis and performance measure based on the Price of Anarchy are also introduced. These solutions concepts are then applied to two test cases in Section 7.5: a three node toy network and the IEEE 14-bus network.

7.2 Prosumers and Local Communities

In this section, we define the generic framework of agent (prosumer) interactions, and a stylized representation of the underlying (distribution) graph. We formulate the local supply and demand balancing constraint that holds in each node. To formalize the two market designs **(i)** and **(ii)**, we introduce the costs, utility functions, social welfare, private information and main assumptions on which our model relies.

7.2.1 Generic Framework

Let $\mathcal{N}$ be a set of N nodes, each of them representing an agent (prosumer), except the root node 0 which is assumed to contain only conventional generation. The root node belongs to the set $\mathcal{N}$. It can trade energy with any other node in $\mathcal{N}$. Under this assumption, the distribution network is a radial graph, with the root node being the interface between the local energy communities and the transmission network. Figure 7.1 illustrates such a graph structure.

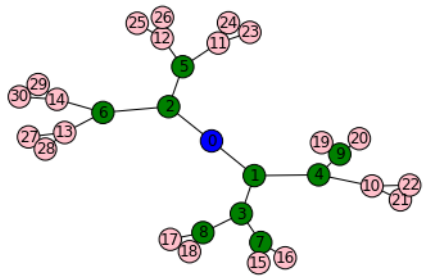


FIGURE 7.1: Example of a radial network. The root node at the interface of the distribution and transmission networks, can trade energy with any other node in the distribution network. In the distribution network, prosumer nodes organize in local energy communities, trading energy with neighbors inside their local community.

Let Ω_n be the set of neighbors of n, with the structure of a communication network (local energy community). It does not necessarily reflect the grid constraints. As usual, we assume that $n \in \Omega_n$, for all $n \in \mathcal{N}$. In particular, $\Omega_0 \stackrel{\text{def}}{=} \mathcal{N} \setminus \{0\}$.

In each node n, we introduce $\mathcal{X}_n \stackrel{\text{def}}{=} \{x_n \in \mathbb{R}_+ | \underline{x}_n \leqslant x_n \leqslant \overline{x}_n\}$ as agent n's demand set, with $\underline{x}_n$ and $\overline{x}_n$ being the lower and upper-bounds on demand capacity.

In parallel to the demand-side, we define the self-generation-side by letting $\mathcal{G}_n \stackrel{\text{def}}{=} \{G_n \in \mathbb{R}_+ | \underline{G}_n \leqslant G_n \leqslant \overline{G}_n\}$ be agent n's flexibility activation set, where $\underline{G}_n$ and $\overline{G}_n$ are the lower and upper-bounds on flexibility activation capacity.

The decision variables of each prosumer n are her demand x_n, flexibility activation G_n, and the quantity exchanged between n and m in the direction from m to n, q_{mn}, for all $m \in \Omega_n \setminus \{n\}$. If $q_{mn} \geqslant 0$, then n buys q_{mn} from m, otherwise ($q_{mn} < 0$) n sells $-q_{mn}$ to m. We impose an inequality on the trading reciprocity:

$$q_{mn} + q_{nm} \leqslant 0\,, \tag{7.1}$$

which means that, in the case where $q_{mn} > 0$, the quantity that n buys from m can not be larger that the quantity q_{nm} that m is willing to offer to n.

Remark 7.1. *In the present work, we model the trade reciprocity constraint as the inequality* (7.1)*. Other works, as [SBP18], consider a different model with an equality $q_{mn} = -q_{nm}$, meaning that the quantity proposed by agent n should be equal to the quantity the agent m wants. In our model, with* (7.1)*, those quantities do not necessarily correspond: n can be willing to offer more than the quantity wanted by m. If the inequality is strict (for instance, n has too much to offer), then part of her energy is produced in excess. Considering a model with an equality means that energy surplus is not allowed.*

Remark 7.2. *Considering an equality constraint instead of* (7.1) *might be intuitive and does not raise any problem when studying centralized solutions as in [SBP18], but this model becomes degenerated when studying GNEs, which is one of the main objective of this chapter. Indeed, a profile is a GNE if, by definition, it is optimal for each agent when considering the actions of the other agents fixed. Thus, if we impose an equality in* (7.1)*, any feasible solution $(\boldsymbol{q}_n)_n$ is a GNE as, for each player n, the quantities $(q_{mn})_n$ are fixed by the others. This degenerated situation does not appear when considering an inequality, as each agent n has a degree of freedom in her trade with other agents.*

The difference between the sum of imports and the sum of exports in node n is defined as the net import in that node: $Q_n \stackrel{\text{def}}{=} \sum_{m\in\Omega_n} q_{mn}$. Furthermore, each line is constrained in capacity. Let $\kappa_{nm} \in [0, +\infty[$ be the equivalent interconnection capacity between node n and node m, such that $q_{nm} \leqslant \kappa_{nm}$, $\kappa_{nm} = \kappa_{mn}$.

RES-based (solar PV panels) self-generation at each node n is modeled as a random variable ΔG_n. Its realization is exogenous to our model.

7.2.2 Local Supply and Demand Balancing

Local supply and demand equilibrium leads to the following equality in each node n in $\mathcal{N}$:

$$\begin{aligned} x_n &= G_n + \Delta G_n + \sum_{m\in\Omega_n} q_{mn}, \\ &= G_n + \Delta G_n + Q_n. \end{aligned} \tag{7.2}$$

Assuming perfect competition, a Market Operator (MO) maximizes the system social welfare, defined as the sum of the utilities of all the agents in the system, under a set of operational and *power flow* constraints, while checking that supply and demand balance each other at each node of the network. In nodal markets, allocative market efficiency can be achieved by setting (locational marginal) nodal price, λ_n, equal to the dual variable of the local supply and demand balancing equation [She+82].

In the present work, we consider an *innovative decentralized* market clearing, by comparison with the classical centralized approach, which is used for example in nodal markets. For that purpose, we introduce decentralization in agents' decision-making. This decentralization results firstly from the fact that demands, flexibility activation and trades are defined selfishly by each prosumer in the nodes; secondly from the fact that all the information regarding preferences and private information on target demands and RES-based generations is not available to all the nodes. The decentralized market clearing relies on a peer-to-peer market design, where each agent n computes the Lagrangian variable associated with her (local) supply and demand balancing equation, using the information at her disposal. Dual variables λ_n are kept private to agent n and used to compute her bilateral trading prices.

7.2.3 Cost and Usage Benefit Functions

Flexibility activation (production) cost in node n is modeled as a quadratic function of local activated flexibility, using three positive parameters a_n, b_n and d_n:

$$C_n(G_n) = \tfrac{1}{2} a_n G_n^2 + b_n G_n + d_n, \tag{7.3}$$

with $-\frac{b_n}{a_n} \geqslant \underline{G}_n$. We make the standard assumption that self-generation occurs at zero marginal cost.

The usage benefit perceived by agent n is modeled as a strictly concave function of node n demand [FHW17], using two positive parameters $\tilde{a}_n$, $\tilde{b}_n$ and a target demand defined exogenously by agent n:

$$U_n(x_n) = -\tilde{a}_n (x_n - x_n^\sharp)^2 + \tilde{b}_n. \tag{7.4}$$

The quantity $-U_n(.)$ can also be considered as the consumption cost of agent n [SBP18]. As $U_n(.)$ captures a usage benefit, which is interpreted as the comfort perceived by agent n, we impose that it always remains non-negative,that is:

$$\overline{x}_n - \sqrt{\tfrac{\tilde{b}_n}{\tilde{a}_n}} \leqslant x_n^\sharp \leqslant \underline{x}_n + \sqrt{\tfrac{\tilde{b}_n}{\tilde{a}_n}}\,.$$

The rational beneath this definition of usage benefit relates to the expected-utility theory [Rab00]: $U_n(x_n)$ represents the perceived comfort resulting from demand x_n satisfaction. The utility function is defined up to a positive affine transformation, and could be multiplied by a positive constant factor without changing the interpretation. The concavity of the function captures the absolute risk aversion of agent n. This is the most general class of utility functions that are often used because of their mathematical tractability. It is increasing for $x_n \leqslant x_n^\sharp$ (larger x_ns lead to higher usage benefits up to the maximum usage benefit), and decreasing for $x_n > x_n^\sharp$ (lower x_ns are better once the maximum usage benefit has been reached).

At the point of maximum usage benefit $x_n = x_n^\sharp$ gives $U_n(x_n^\sharp) = \tilde{b}_n$. We consider that usage benefit vanishes in case of zero demand, i.e., $U_n(0) = 0 \Leftrightarrow \tilde{a}_n = \frac{\tilde{b}_n}{(x_n^\sharp)^2}, \forall n \in \mathcal{N}$. This means that under the assumption that zero demand implies zero usage benefit, an explicit relation exists between the parameter $\tilde{a}_n$, the maximum usage benefit $\tilde{b}_n$, and the target demand $x_n^\sharp$.

In this work, we consider that prosumers have preferences on the possible trades with their neighbors. The preferences are modeled with (product) differentiation prices [SBP18]: each agent n has a positive price $c_{nm} > 0$ to buy energy to an agent m in her neighborhood Ω_n. The total trading cost function of agent n is denoted by:

$$\tilde{C}_n(\boldsymbol{q}_n) = \sum_{m \in \Omega_n, m \neq n} c_{nm} q_{mn}. \tag{7.5}$$

Parameters c_{nm} can model taxes to encourage/refrain the development of certain technologies (micro-CHPs, storage, solar panels) in some nodes. They can also capture agents' preferences to pay regarding certain characteristics of trades (RES-based generation, location of the prosumer, transport distance, size of the prosumer, etc.). If $q_{mn} > 0$ (i.e., n buys q_{mn} from m) then n has to pay the cost $c_{nm} q_{mn} > 0$. Thus, the higher c_{nm} is, the less interesting it is for n to buy energy from m but the more interesting it is for n to sell energy to m. On the other side, if $q_{mn} < 0$, then n sends the energy $-q_{mn}$ and receives the value $-c_{nm} q_{mn} > 0$ even if m does not accept all this energy (i.e. $q_{nm} + q_{mn} < 0$). In that case the energy surplus is bought by an aggregator and sold on the wholesale electricity market in exchange for a compensation intended for the prosumers with energy surpluses. This mechanism will be discussed in details in Section 7.3.

7.2.4 Utility Function and Social Welfare

Agent n's utility function is defined as the difference between the usage benefit resulting from the consumption of x_n energy unit and the sum of the flexibility activation and trading costs. Formally, it takes the form:

$$f_n(x_n, G_n, \boldsymbol{q}_n) = U_n(x_n) - C_n(G_n) - \tilde{C}_n(\boldsymbol{q}_n), \tag{7.6}$$

where $\boldsymbol{q}_n = (q_{mn})_{m\in\Omega_n, m\neq n}$.

We introduce the social welfare as the sum of the utility functions of all the agents in $\mathcal{N}$:

$$\mathrm{SW}(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q}) = \sum_{n\in\mathcal{N}} f_n(x_n, G_n, \boldsymbol{q}_n). \tag{7.7}$$

7.2.5 Private Information at the Nodes

There is private information at each node n that can be associated with:

- ΔG_n, local RES-based generation;
- $x_n^\sharp$, target demand;
- $C_n(.)$, flexibility activation cost function, more specifically parameters a_n, b_n, d_n;
- $U_n(.)$, usage benefit function, more specifically parameters $\tilde{a}_n$, $\tilde{b}_n$;
- $\tilde{C}_n(.)$, bilateral trade cost function, more specifically parameters $(c_{nm})_{m\in\mathcal{N}\setminus\{n\}}$.

In a centralized market design, all the private information is reported to the Market Operator (MO). This means that the local target demands $(x_n^\sharp)_{n\in\mathcal{N}}$ and RES-based generations $(\Delta G_n)_{n\in\mathcal{N}}$, are known by the MO. In contrast, in a peer-to-peer market design, $x_n^\sharp$ and ΔG_n are known only by agent n.

7.3 Centralized Market Design

The centralized market design is inspired from the existing pool-based markets. The global Market Operator (MO) maximizes the social welfare defined in Equation (7.7) under demand capacity constraints (7.8a) and flexibility activation capacity constraints (7.8b) in each node, capacity trading flow constraints for each couple of nodes (7.8c), trading reciprocity constraint (7.8d) and supply-demand balancing (7.8e) in each node:

$$\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{G},\boldsymbol{q}} \quad & \mathrm{SW}(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q}), & & \\
s.t. \quad & \underline{x}_n \leqslant x_n \leqslant \overline{x}_n, \forall n \in \mathcal{N}, & (\underline{\mu}_n, \overline{\mu}_n) & \quad (7.8a)\\
& \underline{G}_n \leqslant G_n \leqslant \overline{G}_n, \forall n \in \mathcal{N}, & (\underline{\nu}_n, \overline{\nu}_n) & \quad (7.8b)\\
& q_{mn} \leqslant \kappa_{mn}, \forall m \in \Omega_n, m \neq n, \forall n \in \mathcal{N}, & (\xi_{nm}) & \quad (7.8c)\\
& q_{mn} \leqslant -q_{nm}, \forall m \in \Omega_n, m > n, \forall n \in \mathcal{N}, & (\zeta_{nm}) & \quad (7.8d)\\
& x_n = G_n + \Delta G_n + Q_n, \forall n \in \mathcal{N}. & (\lambda_n) & \quad (7.8e)
\end{aligned}$$

Remark 7.3. *The constraint* (7.8d) *is indexed by $m > n$ so that the constraint is considered only once.*

Dual variables are denoted in blue font between brackets at the right of the corresponding constraints. Some of the dual variables can be interpreted as shadow prices, with classical interpretations in the energy economics literature. In the remainder, ξ_{nm} will be interpreted as the shadow price (congestion price) associated with capacity trading flow constraint (7.8c) between nodes n and m; ζ_{nm} will be understood as the bilateral trade price offered by n to m associated with the trading reciprocity constraint (7.8d); while λ_n is the nodal price

associated with the supply and demand balancing constraint in node n (7.8e), as discussed in Subsection 7.2.2.

The Social Welfare function is concave as the sum of concave functions defined on a convex feasibility set. Indeed, the feasibility set is obtained as Cartesian product of convex sets. We can compute the Lagrangian function associated with the standard constrained optimization problem of social welfare maximization under constraints (7.8a)-(7.8e):

$$\begin{aligned}\mathcal{L}(\boldsymbol{x},\boldsymbol{G},\boldsymbol{Q},\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\xi},\boldsymbol{\zeta},\boldsymbol{\lambda}) &= \sum_{n\in\mathcal{N}} \mathcal{L}_n(x_n, G_n, \boldsymbol{q}_n, \boldsymbol{\mu}_n, \boldsymbol{\nu}_n, \boldsymbol{\xi}_n, \boldsymbol{\zeta}_n, \lambda_n) \\ &= -\sum_{n\in\mathcal{N}} f_n(x_n, G_n, \boldsymbol{q}_n) + \sum_{n\in\mathcal{N}} \underline{\mu}_n(\underline{x}_n - x_n) \\ &+ \sum_{n\in\mathcal{N}} \overline{\mu}_n(x_n - \overline{x}_n) + \sum_{n\in\mathcal{N}} \underline{\nu}_n(\underline{G}_n - G_n) + \sum_{n\in\mathcal{N}} \overline{\nu}_n(G_n - \overline{G}_n) \\ &+ \sum_{n\in\mathcal{N}} \sum_{m\in\Omega_n, m\neq n} \xi_{nm}(q_{mn} - \kappa_{mn}) + \sum_{n\in\mathcal{N}} \sum_{m\in\Omega_n, m>n} \zeta_{nm}(q_{mn} + q_{nm}) \\ &+ \sum_{n\in\mathcal{N}} \lambda_n \Big(x_n - G_n - \Delta G_n - Q_n \Big).\end{aligned} \tag{7.9}$$

To determine the solution of the centralized market design optimization problem, we compute KKT conditions associated with Lagrangian function (7.9). Taking the derivative of the Lagrangian function (7.9) with respect to x_n, G_n, q_{mn}, for all n in $\mathcal{N}$ and all $m \in \Omega_n, m \neq n$, the stationarity conditions write down as follows:

$$\frac{\partial \mathcal{L}}{\partial x_n} = 0 \Leftrightarrow 2\tilde{a}_n(x_n - x_n^\sharp) - \underline{\mu}_n + \overline{\mu}_n + \lambda_n = 0\,, \quad \forall n \in \mathcal{N}\,, \tag{7.10a}$$

$$\frac{\partial \mathcal{L}}{\partial G_n} = 0 \Leftrightarrow a_n G_n + b_n - \underline{\nu}_n + \overline{\nu}_n - \lambda_n = 0\,, \quad \forall n \in \mathcal{N}\,, \tag{7.10b}$$

$$\frac{\partial \mathcal{L}}{\partial q_{mn}} = 0 \Leftrightarrow c_{nm} + \xi_{nm} + \zeta_{nm} - \lambda_n = 0, \quad \forall m \in \Omega_n, m \neq n, \forall n \in \mathcal{N}\,, \tag{7.10c}$$

where, for $m < n$, ζ_{nm} is defined as equal to ζ_{mn}.

From (7.10c), we infer that the nodal price at n can be expressed analytically as the sum of the node product differentiation prices regarding the other prosumers in her neighborhood, the congestion constraint dual variable from Equation (7.8c) and the bilateral trade prices:

$$\lambda_n = c_{nm} + \xi_{nm} + \zeta_{nm}, \quad \forall m \in \Omega_n, m \neq n, \quad \forall n \in \mathcal{N}\,. \tag{7.11}$$

The complementarity constraints take the following form:

$$0 \leqslant \underline{\mu}_n \perp x_n - \underline{x}_n \geqslant 0, \quad \forall n \in \mathcal{N}\,, \tag{7.12a}$$

$$0 \leqslant \overline{\mu}_n \perp \overline{x}_n - x_n \geqslant 0, \quad \forall n \in \mathcal{N}\,, \tag{7.12b}$$

$$0 \leqslant \underline{\nu}_n \perp G_n - \underline{G}_n \geqslant 0, \quad \forall n \in \mathcal{N}\,, \tag{7.12c}$$

$$0 \leqslant \overline{\nu}_n \perp \overline{G}_n - G_n \geqslant 0, \quad \forall n \in \mathcal{N}\,, \tag{7.12d}$$

$$0 \leqslant \xi_{nm} \perp \kappa_{mn} - q_{mn} \geqslant 0, \quad \forall m \in \Omega_n, m \neq n, \forall n \in \mathcal{N}\,, \tag{7.12e}$$

$$0 \leqslant \zeta_{nm} \perp -q_{mn} - q_{nm} \geqslant 0, \quad \forall m \in \Omega_n, m > n, \forall n \in \mathcal{N}\,. \tag{7.12f}$$

From (7.10c), we infer, for any couple of nodes $n \in \mathcal{N}, m \in \Omega_n, m > n$, that:

$$\zeta_{nm} = \lambda_n - c_{nm} - \xi_{nm} = \lambda_m - c_{mn} - \xi_{mn}\,, \tag{7.13}$$

Subtracting those two last members in (7.13), we infer that:

$$c_{nm} - c_{mn} + \xi_{nm} - \xi_{mn} = \lambda_n - \lambda_m, \forall m \in \Omega_n, m \neq n, \forall n \in \mathcal{N}\,. \tag{7.14}$$

From Equations (7.10a) and (7.10b), we infer that, at the optimum, for each node n:

$$x_n = x_n^\sharp - \frac{1}{2\tilde{a}_n}\Big(\lambda_n + (\overline{\mu}_n - \underline{\mu}_n)\Big)\,, \tag{7.15}$$

$$G_n = -\frac{b_n}{a_n} + \frac{1}{a_n}\Big(\lambda_n - (\overline{\nu}_n - \underline{\nu}_n)\Big)\,. \tag{7.16}$$

Substituting Equations (7.15) and (7.16) in the local demand and in the supply balance Equation (7.8e), we infer that the net import at node n can be expressed as a linear function of the nodal price:

$$Q_n = \left(x_n^\sharp - \frac{1}{2\tilde{a}_n}(\overline{\mu}_n - \underline{\mu}_n) + \frac{b_n}{a_n} + \frac{1}{a_n}(\overline{\nu}_n - \underline{\nu}_n)\right) - \left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right)\lambda_n - \Delta G_n\,. \tag{7.17}$$

The results are summarized in the following proposition.

Proposition 7.1. *In the quadratic model defined by equations (7.3-7.6), the optimal demands, flexibility activations and net imports at each node n can be expressed as linear functions of the nodal price at that node, given by Equations* (7.15), (7.16), *and* (7.17).

The total sum of the net imports at all nodes should be negative or null, i.e., $\sum_{n\in\mathcal{N}} Q_n \leqslant 0$. We observe from the supply-demand balancing (7.8e) that this is equivalent to $\sum_{n\in\mathcal{N}}(x_n - G_n) \leqslant \sum_{n\in\mathcal{N}} \Delta G_n$. A strict inequality would lead to a situation of energy surplus, i.e., the total energy generation is in excess compared to the total demand of the prosumers.

To deal with that energy surplus, we assume that a feed-in-tariff or feed-in-premium applies. The root node (node 0) who makes the link between the transmission and the distribution network could be a good candidate to manage the excess of generation. Indeed, she should be able to inject it in the transmission network. However, due to the radial structure of our network, all the distribution nodes are not directly connected to the root node. Relying on (7.11), this means that the bilateral trading prices between 0 and a node $n \in \mathcal{N} \setminus \{0\}$ cannot be the same for all the nodes in the distribution network because the trade price also depends on c_{0n} and ξ_{0n} which captures the congestion state of the path between 0 and n. As a result, node 0 cannot apply a feed-in-tariff in case of energy surplus. However, it might be possible to introduce another agent, such as an aggregator, having a very large demand and no generation capacity, that would be connected to all nodes of the distribution network. This aggregator would take care of the forecasting and bidding of the renewable generation and self-generation surpluses, while paying to prosumers the amount of energy they actually produced in excess at a price defined in advance (for example, the feed-in-tariff price or a premium). This compensation mechanism for the agents is similar to the *purchase obligations* or feed-in tariffs mechanism for renewable energy sources set in the European Union [FJ08].

Constraints on the technologies could also be applied at the prosumer level, to limit the RES-based generation and to choose large enough demand capacities. Note that the sizing of the prosumers' capacities and RES-based generation possible clipping strategies are out of the scope of this work. This result is formalized in the proposition below.

Proposition 7.2. *A necessary condition for no energy surplus is that there is at least one prosumer n in $\mathcal{N}$ whose capacities and RES-based generation are such that $\overline{x}_n - \underline{G}_n \geqslant \Delta G_n$.*

Proof. By combining (7.8a) and (7.8b), we obtain $\underline{x}_n - \overline{G}_n \leqslant x_n - G_n \leqslant \overline{x}_n - \underline{G}_n$. Subtracting ΔG_n in each part of the inequalities and applying (7.8e), we get $\underline{x}_n - \overline{G}_n - \Delta G_n \leqslant Q_n \leqslant \overline{x}_n - \underline{G}_n - \Delta G_n$. Then, $\overline{x}_n - \underline{G}_n - \Delta G_n < 0$ implies that $Q_n < 0$, i.e., there are more exports than imports from n. If $\overline{x}_n - \underline{G}_n - \Delta G_n < 0$, for all $n \in \mathcal{N}$ then, $\sum_{n\in\mathcal{N}} Q_n < 0$. No energy surplus is equivalent to $\sum_{n\in\mathcal{N}} Q_n = 0$. For this equality to hold, it is necessary that there exists at least one prosumer n in $\mathcal{N}$ such that $\overline{x}_n - \underline{G}_n \geqslant \Delta G_n$. □

In practice, this means that prosumers should size their capacities such that the difference between their upper-bound on demand capacity and lower-bound on flexibility activation capacity is larger than their RES-based generation. However, the previous proposition is a necessary condition.

The following proposition gives a sufficient condition on the locational marginal prices $(\lambda_n)_n$ for having no energy surplus at optimality. The latter condition ii) can be directly inferred by the complementarity condition (7.12f) and (7.10c).

Proposition 7.3. *Let $n_0 \in \mathcal{N}$ and a neighboring node $m_0 \in \Omega_{n_0}$. If there exists a node $m \in \mathcal{N}$ such that there exists a non congested path $(n_0, n_1, \dots, n_p = m)$ from n_0 to m with $\lambda_m > c_{n_0,m_0} - \sum_{i=0}^{p-1}(c_{n_i,n_{i+1}} - c_{n_{i+1},n_i})$, where λ_m is the optimal Lagrange multiplier (nodal price) defined in (7.8e), then there is no energy surplus at n_0 in the trade with m_0 (that is, $q_{n_0,m_0} + q_{m_0,n_0} = 0$). In particular:*

i) if all users have symmetric preferences $c_{nm} = c_{mn}$, there is no congestion and there exists $m \in \mathcal{N}$ with nodal price $\lambda_m > c_{n_0,m_0}$, then there is no energy surplus at n_0 in the trade with m_0;

ii) for $m = n_0$, if the nodal price realizes $\lambda_{n_0} > c_{n_0,m_0}$, then there is no energy surplus at n_0 in the trade with m_0.

Proof. Suppose on the contrary that there is some energy surplus at n_0: there exists some m_0 such that $q_{n_0,m_0} + q_{m_0,n_0} < 0$ and $q_{m_0,n_0} < 0$ (i.e. n_0 rejects energy). In the case where $G_m > \underline{G}_m$, Consider the infinitesimal transformation to the trades and production:

$$\begin{aligned} &q_{n_i,n_{i+1}} \leftarrow q_{n_i,n_{i+1}} + \varepsilon, \qquad q_{n_{i+1},n_i} \leftarrow q_{n_{i+1},n_i} - \varepsilon, \ \ \forall i \in \{0,\dots,p-1\}, \\ &q_{m_0,n_0} \leftarrow q_{m_0,n_0} + \varepsilon\,, \qquad G_m \leftarrow G_m - \varepsilon\,. \end{aligned} \tag{7.18}$$

Then, for ε small enough, all constraints are still satisfied and the variations in SW has the same sign as:

$$\lambda_m - c_{n_0,m_0} + \textstyle\sum_{i=0}^{p-1}(c_{n_i,n_{i+1}} - c_{n_{i+1},n_i}) > 0\,.$$

Hence, we can strictly increase SW, which contradicts the optimality. In the case where $G_m = \underline{G}_m$, then we necessarily have $x_m < \overline{D}_m$ (otherwise $\lambda_n = -2\tilde{a}_n(\overline{D}_n - x_n^\sharp) - \overline{\mu}_n < 0$ which is impossible from (7.10c)), and we can strictly increase x_m instead of decreasing G_m in (7.18), leading to the same contradiction. □

Remark 7.4. *From the previous proposition, we see that even if there is no excess in the renewable production, i.e. $\sum_n \Delta G_n < \sum_n x_n^\sharp$, we can still have some energy surplus if the trades preference prices $(c_{nm})_{n,m}$ are large enough.*

Hence, assuming no energy surplus, the total sum of the net imports in all nodes should vanish i.e. $\sum_{n\in\mathcal{N}} Q_n = 0$, which implies the following relation, by using (7.17):

$$\sum_{n\in\mathcal{N}}\left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right)\lambda_n = \sum_{n\in\mathcal{N}}\left(x_n^\sharp - \frac{1}{2\tilde{a}_n}(\overline{\mu}_n - \underline{\mu}_n) + \frac{b_n}{a_n} + \frac{1}{a_n}(\overline{\nu}_n - \underline{\nu}_n) - \Delta G_n\right). \tag{7.19}$$

From Equation (7.14), we infer that the nodal price at node n is a linear function of the nodal price at the root node, product differentiation and congestion prices with all the other nodes in $\mathcal{N}$:

$$\lambda_n = c_{n0} - c_{0n} + \xi_{n0} - \xi_{0n} + \lambda_0, \quad \forall n \in \Omega_0\,. \tag{7.20}$$

Substituting Equation (7.20) in Equation (7.19), we infer the closed form expression of the nodal price at the root node:

$$\begin{aligned} \lambda_0 \sum_{n\in\mathcal{N}}\left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right) = &\sum_{n\in\mathcal{N}}\left(x_n^\sharp - \frac{1}{2\tilde{a}_n}(\overline{\mu}_n - \underline{\mu}_n) + \frac{b_n}{a_n} + \frac{1}{a_n}(\overline{\nu}_n - \underline{\nu}_n) - \Delta G_n\right) \\ &- \sum_{n\in\Omega_0}\left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right)(c_{n0} - c_{0n} + \xi_{n0} - \xi_{0n})\,. \end{aligned} \tag{7.21}$$

From Equations (7.20) and (7.21), assuming that $(c_{n0})_n$, $(c_{0n})_n$, $(\xi_{n0})_n$, $(\xi_{0n})_n$ are known, the MO can iteratively compute all the $(\lambda_n)_{n\in\mathcal{N}}$. Note that $\underline{\boldsymbol{\mu}}, \overline{\boldsymbol{\mu}}$ and $\underline{\boldsymbol{\nu}}, \overline{\boldsymbol{\nu}}$ are determined by the MO when optimizing $\boldsymbol{x}$ and $\boldsymbol{G}$. Once computed by the MO, the nodal prices are announced to all the agents $n \in \mathcal{N}$. Then, to determine the optimal bilateral trading prices,

each agent n has to refer to Equation (7.13), which gives the bilateral trading prices as linear functions of the nodal price and congestion price. The results are summarized in the following proposition:

Proposition 7.4. *Assuming no energy surplus and knowing $(c_{n0})_n$, $(c_{0n})_n$, $(\xi_{n0})_n$, $(\xi_{0n})_n$, the MO computes the nodal price at the root node by Equation (7.21). The nodal prices in all the other nodes of the distribution network can be inferred from λ_0 according to Equation (7.20). Then, for each node $n \in \mathcal{N}$, bilateral trading prices can be computed for any node $m \in \Omega_n, n \neq m$ by Equation (7.13) provided congestion price $(\xi_{nm})_{m>n, m\in\Omega_n}$ is known*[1].

If all agents reveal their product differentiation prices $(c_{n0})_n$ to the MO and all the congestion prices $(\xi_{n0})_n$, $(\xi_{0n})_n$ in the lines involving the root node are known (or rationally anticipated), then the MO can compute all the nodal prices $(\lambda_n)_{n\in\mathcal{N}}$ from λ_0.

The result below shows a link between the market and the state of the distribution grid: Proposition 7.5 shows that the distribution grid lines become congested if there are "cycles" in the preferences as explained below.

Proposition 7.5. *Suppose that the matrix $\tilde{C} \overset{\text{def}}{=} (c_{nm} - c_{mn})_{nm}$ has a strictly negative cycle of length $k > 2$, i.e. there is a sequence of distinct indices $(n_i)_{1\leqslant i\leqslant k}$ such that $\sum_{1\leqslant i\leqslant k} \tilde{C}_{n_i,n_{i+1}} < 0$, where $n_{k+1} \overset{\text{def}}{=} n_1$. Then, at an optimal centralized solution, there is a trade opposed to the cycle made at full capacity, i.e. there exists $i \in \{1,\dots,k\}$ such that $q_{n_{i+1},n_i} = \kappa_{n_{i+1},n_i}$.*

Symmetrically, if there is a strictly positive cycle $(n_i)_{1\leqslant i\leqslant k}$ such that $\sum_{1\leqslant i\leqslant k} \tilde{C}_{n_i,n_{i+1}} > 0$, then at an optimal centralized solution, there is a trade in the direction of the cycle made at full capacity, i.e. there exists $i \in \{1,\dots,k\}$ such that $q_{n_i,n_{i+1}} = \kappa_{n_i,n_{i+1}}$.

Proof of Proposition 7.5. We prove the first part of the proposition as the second is symmetric.

Consider the trades $(q_{nm})_{nm}$ at an optimal solution and suppose on the contrary that there is $\epsilon > 0$ such that, for each $i \in \{1,\dots,k\}$, we have $q_{n_{i+1},n_i} \leqslant \kappa_{n_{i+1},n_i} - \epsilon$.

Then consider the same solution with trades $(\tilde{q}_{nm})_{nm}$ defined as follows: for each $i \in \{1,\dots,k\}$, let $\tilde{q}_{n_{i+1},n_i} \overset{\text{def}}{=} q_{n_{i+1},n_i} + \epsilon$ and $\tilde{q}_{n_i,n_{i+1}} \overset{\text{def}}{=} q_{n_i,n_{i+1}} - \epsilon$, while $\tilde{q}_{nm} = q_{nm}$ otherwise. Then all constraints are still feasible because, for each i, $\sum_{m\neq n_i} q_{m,n_i} = Q_n - \epsilon + \epsilon = Q_n$. Besides, by definition of $\tilde{q}$, we still have $\tilde{q}_{mn} = -\tilde{q}_{nm}$ for any $m > n$. Moreover, if we denote by SW the social welfare of the previous solution $(q_{nm})_{nm}$, the social welfare of this new solution is:

$$\begin{aligned}
\widetilde{\text{SW}} &= \text{SW} + \sum_n \sum_{m\neq n} c_{nm}(q_{mn} - \tilde{q}_{mn}) \\
&= \text{SW} + \sum_{1\leqslant i\leqslant k} \Big(c_{n_i,n_{i+1}}(q_{n_{i+1},n_i} - \tilde{q}_{n_{i+1},n_i}) + c_{n_i,n_{i-1}}(q_{n_{i-1},n_i} - \tilde{q}_{n_{i-1},n_i}) \Big) \\
&= \text{SW} + \sum_{1\leqslant i\leqslant k} \epsilon \left(c_{n_i,n_{i-1}} - c_{n_i,n_{i+1}} \right) = \text{SW} - \epsilon \sum_{1\leqslant i\leqslant k} \tilde{C}_{n_i,n_{i+1}} > \text{SW} \,,
\end{aligned}$$

which contradicts the fact that SW is maximal. □

Remark 7.5. *The property stated by Proposition 7.5 shows that the lines become congested if there is a strictly positive or negative cycle in the matrix $\tilde{C}$. In practice, a central MO should try to avoid such an outcome, since the congested lines are unavailable in case of unplanned real need (outages, peak demand). The existence of a positive cycle in $\tilde{C}$ means that there is an "arbitrage" opportunity in the network. In other words, one can strictly increase the social welfare by doing an exchange of energy quantities. We can make the assumption that this kind of opportunities do not exist in practice, since they should vanish quickly in a liquid market.*

From the point of view from mechanism design, we might also prevent this kind of cycling behavior by adding a transaction fee (e.g. $\tau \times |q_{mn}|$ with $\tau > 0$) on the trades, regardless they are positive or negative.

[1] Two assumptions can be made on the determination of the congestion prices: first, they are determined exogenously while checking the complementarity constraint (7.12e); second, they are determined through a market for (distribution) capacity line transmission. This second assumption enables the MO to complete the market. It will be discussed later in the chapter.

Section 7.5.1 shows an example where there is a cycling trade that is purely due to arbitrage opportunities because of the preferences.

7.4 Peer-to-Peer Market Design

The centralized market design is used, in this section, as a benchmark against which we test the performance of a fully distributed approach relying on peer-to-peer energy trading. We first start by defining in Subsection 7.4.1 the solution concepts of GNE and VNE, that we will use to analyze the outcome of the fully distributed market design. Then, various results are introduced to characterize the relations between these sets of solutions. Congestion issues and performance measures are discussed in Subsection 7.4.2.

7.4.1 Generalized Nash Equilibrium and Variational Equilibrium

In the peer-to-peer setting, each agent $n \in \mathcal{N}$ determines, by herself, her demand, flexibility activation and bilateral trades with other agents in her local energy community under constraints on demand, flexibility activation and transmission capacity so as to maximize her utility. A trade between two agents in a local energy community supposes that these two have decided on a certain quantity to be sent from one side and received by the other side. Therefore, there must be an "agreement" or trade constraint between each pair of agents in a local community, which couples their respective decisions. As a result, although the utility of a prosumer depends only on her own decisions, some of these decisions, such as the quantity she agrees to trade with all the other prosumers in her neighborhood, have an impact on the set of feasible actions of her neighbors. In the same way, her feasible actions are determined by the actions of her neighbors.

Formally, each agent in node $n \in \mathcal{N}$ solves the following optimization problem:

$$\max_{x_n, G_n, (q_{mn})_{m\in\Omega_n, m\neq n}} \quad f_n\Big(x_n, G_n, \boldsymbol{q}_n\Big), \tag{7.22a}$$

$$s.t. \quad \underline{x}_n \leqslant x_n \leqslant \overline{x}_n, \qquad (\underline{\mu}_n, \overline{\mu}_n) \tag{7.22b}$$

$$\underline{G}_n \leqslant G_n \leqslant \overline{G}_n, \qquad (\underline{\nu}_n, \overline{\nu}_n) \tag{7.22c}$$

$$q_{mn} \leqslant \kappa_{mn}, \forall m \in \Omega_n, m \neq n, \qquad (\xi_{nm}) \tag{7.22d}$$

$$q_{mn} \leqslant -q_{nm}, \forall m \in \Omega_n, m \neq n, \qquad (\zeta_{nm}) \tag{7.22e}$$

$$x_n = G_n + \Delta G_n + Q_n, \qquad (\lambda_n) \tag{7.22f}$$

where $\boldsymbol{q}_n = (q_{mn})_{m\in\Omega_n}$ are the trading decisions of agent n.

Hence, the peer-to-peer setting leads to N optimization problems, one for each agent $n \in \mathcal{N}$, with *individual constraints* on demand (7.22b), flexibility activation (7.22c), trade capacity (7.22d), supply and demand balancing (7.22f); as well as *coupling constraints* (7.22e) that ensure the reciprocity of the trades.

The Lagrangian function associated with optimization problem (7.22a) under constraints (7.22b)-(7.22f), writes down as $\mathcal{L}_n$ defined in equation (7.9).

For each agent n, the first order stationarity conditions are the same as (7.10a)-(7.10c), and the complementarity constraints are the same as (7.12a)-(7.12f), except that (7.12f) is indexed by all (m, n) with $m \neq n$ and that ζ_{nm} is not necessarily equal to ζ_{mn}. Let this condition system be denoted by KKT_n for each $n \in \mathcal{N}$.

As the problem given by (7.22) is convex, KKT_n are necessary and sufficient conditions for a vector $(x_n, G_n, \boldsymbol{q}_n)$ to be an optimal solution of (7.22).

Remark 7.6. *In Equation* (7.22), f_n *depends on the variables of player n only, and not on the variables of the other players. A consequence is that the social welfare function is decomposable:* $\mathrm{SW}(D, G, \boldsymbol{q}) = \sum_n f_n\big(x_n, G_n, \boldsymbol{q}_n\big)$. *Therefore, without the existence of the coupling transaction constraint* (7.22e), *the minimization of* SW *is equivalent to the minimization of each individual objective function* f_n. *We will see that this equivalence between social optimizer and equilibria also happens for the so-called Variational Equilibria.*

The notion of Generalized Nash Equilibrium (GNE) [Har91], generalizing Nash Equilibrium in the presence of coupling constraints has been commonly adopted:

Definition 7.1 (Generalized Nash Equilibrium [FK07])**.** A Generalized Nash Equilibrium of the game defined by the maximization problems (7.22) with coupling constraints, is a vector $(x_n, G_n, \boldsymbol{q}_n)_n$ that solves the maximization problems (7.22) or, equivalently, a vector $(x_n, G_n, \boldsymbol{q}_n)_n$ such that $(x_n, G_n, \boldsymbol{q}_n)$ solves the system KKT_n for each n.

The constraint (7.22e), $q_{mn} \leqslant -q_{nm}$, written both in the problem of n and in that of $m \neq n$ leads to the same inequality, but is associated to the multiplier ζ_{nm} in the problem of n and to ζ_{mn} in the problem of m. In this chapter, we consider two scenarios for the allocation of the resources represented in these coupling constraints:

Scenario (i) A market allocates the resources associated with (7.22e) through a single price system, therefore leading to the determination of one price for each constraint: $\zeta_{nm} = \zeta_{mn}$.

Scenario (ii) There does not exist any market to determine the price system associated with (7.22e). Hence, two prosumers n, m might attribute different evaluations of the same transaction $q_{mn} \leqslant -q_{nm}$ or, equivalently, the same dual variables to the trade constraint (7.22e) between n and m. This *can* lead to different prices $\zeta_{nm} \neq \zeta_{mn}$ for agents n and m.

The two scenarios have implications on the market organization. Let us discuss them one after another.

Scenario (i) corresponds to a complete market, where the common resources are shared in an efficient way. It suggests that all constraints are traded at a single price, which reflects the common valuation of each product from all agents. The associated solution concept is that of Variational Nash Equilibrium (VNE) [Har91] defined below, a refinement of GNE, where we ask for more symmetry: the Lagrangian multipliers associated to a constraint shared by several players have to be equal from one player to another. Note that a natural way to complete the market would be to introduce a market for (distribution) capacity line transmission, enabling the determination of congestion prices $(\xi_{nm})_{n,m}$. A similar idea was proposed by Oggioni et al. in [Ogg+12] at the transmission level for a subproblem of market coupling.

Definition 7.2 (Variational Nash Equilibrium (VNE) [FK07])**.** A Variational Nash Equilibrium (VNE) of the game defined by (7.22) is a solution $(x_n, G_n, \boldsymbol{q}_n)_n$ that solves the maximization problems (7.22) or, equivalently, a vector $(x_n, G_n, \boldsymbol{q}_n)_n$ such that $(x_n, G_n, \boldsymbol{q}_n)$ solves the system KKT_n for each n and, in addition, such that the Lagrangian multipliers associated to the coupling constraints (7.22e) are equal, i.e.:

$$\zeta_{nm} = \zeta_{mn},\ \forall n \in \mathcal{N}, \forall m \in \Omega_n, m \neq n\ . \tag{7.23}$$

The term "variational" refers to the variational inequality problem associated to such an equilibrium: indeed, if we define the set of admissible solutions as:

$$\mathcal{R} \stackrel{\text{def}}{=} \{\boldsymbol{x} = (x_n, G_n, \boldsymbol{q}_n)_n \,|(7.22\text{b}) - (7.22\text{f}) \text{ hold for each } n \in \mathcal{N}\}\ . \tag{7.24}$$

then $\hat{\boldsymbol{x}} \in \mathcal{R}$ is a VNE if, and only if, it is a solution of (cf. [FK07]):

$$\langle \sum_n \nabla f_n(\hat{\boldsymbol{x}}_n),\ \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \leqslant 0,\ \forall \boldsymbol{x} \in \mathcal{R}\ . \tag{7.25}$$

A remarkable fact is that VNEs exist under mild conditions [Har91; Ros65], even if the additional equality conditions on the multipliers seems restrictive.

We can observe, following Remark 7.6, that VNEs are defined by exactly the same KKT system than the social welfare maximizer (or equivalently as the solution of the same variational inequality (7.25)). Therefore, the following result is immediate:

Proposition 7.6. *The set of VNEs (such that $\zeta_{nm} = \zeta_{mn}$ for all $n \in \mathcal{N}$ and all $m \neq n \in \Omega_n$) coincides with the set of social welfare optima.*

Scenario (ii) corresponds to the case of partial price coordination or a completely missing market for some products. Agents with different willingness to pay for a certain resource face a price gap due to the lack of arbitrage opportunities that prevent price convergence. This imperfect coordination among agents relates to the notion of GNE, where nothing prevents the multipliers ζ_{nm} and ζ_{mn} to be different.

Remark 7.7. *A particular class of GNE is called restricted GNE [Fuk11]. It assumes that the dual variables of the shared constraint* (7.22e) *belongs to a non empty cone of* $\mathbb{R}^{N(N-1)}$.

A particular class of restricted GNE, introduced by Rosen [Ros65], are the normalized equilibria. There, the dual variables of the shared constraint (7.22e) *are equal up to a constant (endogenous) factor* r_n *that depends on prosumer* n*, but not on constraints. Mathematically, it means* $r_n\zeta_{nm} = r_m\zeta_{mn}$*, for all* $n \in \mathcal{N}$ *and all* $m \in \Omega_n, m \neq n$.

From KKT_n, we see that, as in the centralized case, $\lambda_n = \zeta_{nm} + c_{nm} + \xi_{nm}$, i.e., the per-unit nodal price at n is the sum of the transaction price, the preference price and the congestion price, all for getting one unit from m to n, for each neighbor of m. Besides,

$$\zeta_{nm} = \lambda_n - c_{nm} - \xi_{nm}, \forall m \in \Omega_n, m \neq n \,, \tag{7.26}$$

which gives the transaction price for agent n or, in other words, her evaluation of the trade q_{mn}.

In order to derive some results on GNE and simplify notations, let us introduce the coefficient r_n as:

$$\zeta_{0n} r_n = \zeta_{n0}, \quad \forall n \in \mathcal{N} \,. \tag{7.27}$$

Remark 7.8. *We interpret this situation as one where there is an imperfect market for determining the bilateral trade prices obtained as dual variables of the shared constraint* (7.22e)*. Between any couple of prosumer nodes, bilateral trade prices do tend to equalize (i.e.,* r_n *is close to* 1 *for any* $n \in \Omega_0$ *— meaning that the GNE approaches the VNE), but there remains a gap due to insufficient liquidity or differences in the price bids for the asked quantity [Ogg+12]. To some extent,* r_n *can be interpreted as a measure of the efficiency loss introduced by the GNE in comparison with the VNE.*

Using Equation (7.26) for the node 0 and an arbitrary node $n \in \Omega_0$ and for an arbitrary node $n \in \Omega_0$ and the node 0, and summing up both relations, we get:

$$\lambda_n = r_n\lambda_0 + \left(c_{n0} - r_n c_{0n}\right) + \left(\xi_{n0} - r_n\xi_{0n}\right), \ \forall n \in \Omega_0. \tag{7.28}$$

Similarly to the centralized market design case, since the total sum of the net imports in all nodes should vanish under no RES-based generation surplus, i.e., $\sum_{n\in\mathcal{N}} Q_n = 0$, we infer the closed form expression of the nodal price at the root node, similar to the centralized case:

$$\begin{aligned}\lambda_0 \sum_{n\in\mathcal{N}} \left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right) r_n = \sum_{n\in\mathcal{N}} \left(x_n^\sharp - \frac{1}{2\tilde{a}_n}(\overline{\mu}_n - \underline{\mu}_n) + \frac{b_n}{a_n} + \frac{1}{a_n}(\overline{\nu}_n - \underline{\nu}_n) - \Delta G_n\right) \\ - \sum_{n\in\Omega_0} \left(\frac{1}{2\tilde{a}_n} + \frac{1}{a_n}\right) \Big[\left(c_{n0} - c_{0n}r_n\right) + \left(\xi_{n0} - r_n\xi_{0n}\right)\Big].\end{aligned} \tag{7.29}$$

Let us denote $\mathrm{SOL}^{\mathrm{GNEP}}$ the set of GNEs of the peer-to-peer non-cooperative game. GNEs are not unique in general. It is relevant to study how efficient are those different outcomes in comparison to the VNE outcome (where the bilateral trades would be settled down by a MO).

Although there exist several standard methods to compute numerically a VNE in a generalized game (e.g. with variational inequalities methods [FP07]), it is in general harder to compute numerically other GNEs or even the complete set of GNEs.

A possible method to evaluate the set of GNEs is to apply the parameterized variational inequality approach [NTF11; Ogg+12] which enables to characterize each GNE as the solution of an optimization problem. Results based on a similar approach were also presented by Gabriel et al. [Gab+12] through an extensive study of Nash-Cournot and other energy

market models (some integrating market clearing conditions) that use mixed complementarity problems. However, peer-to-peer market design was not considered in this book, and we would like to highlight in this chapter how such results also apply in fully distributed markets. In our specific case, this leads to the optimization problem $\mathcal{P}_{\omega}^{\text{GNE}}$, parameterized by the coefficients $\omega_{nm} > 0$ corresponding to an additional value for user n for her trading constraint with m:

$$\left(\mathcal{P}_{\omega}^{\text{GNE}}\right) \qquad \max_{\boldsymbol{x},\boldsymbol{G},\boldsymbol{q}} \sum_{n\in\mathcal{N}} \left[f_n(x_n, G_n, \boldsymbol{q}_n) - \sum_{m\in\Omega_n, m\neq n} \omega_{nm} q_{mn} \right], \tag{7.30a}$$

$$s.t.\ \ \underline{x}_n \leqslant x_n \leqslant \overline{x}_n, \forall n \in \mathcal{N}, \qquad (\underline{\mu}_n, \overline{\mu}_n) \tag{7.30b}$$

$$\underline{G}_n \leqslant G_n \leqslant \overline{G}_n, \forall n \in \mathcal{N}, \qquad (\underline{\nu}_n, \overline{\nu}_n) \tag{7.30c}$$

$$q_{nm} \leqslant \kappa_{nm} \qquad (\xi_{nm}) \tag{7.30d}$$

$$q_{nm} + q_{mn} \leqslant 0, \forall m \in \Omega_n, m > n, \forall n \in \mathcal{N}. \qquad (\zeta_{nm}) \tag{7.30e}$$

$$x_n = G_n + \Delta G_n + Q_n, \forall n \in \mathcal{N}, \qquad (\lambda_n)\ . \tag{7.30f}$$

From [NTF11, Cor 3.1] and [NTF11, Thm. 3.3], we can make a link between the set of GNEs and the solutions of problem (7.30), as given in the following proposition:

Proposition 7.7. *(i) All GNEs can be found from problem* (7.30)*, that is:*

$$SOL^{GNEP} \subset \bigcup_{(\omega_{nm})\in\mathbb{R}_+^{*N(N-1)}} SOL\left(\mathcal{P}_{\omega}^{\text{GNE}}\right)\ ;$$

(ii) reciprocally, if $(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{\zeta})$ is a solution of $\mathcal{P}_{\omega}^{\text{GNE}}$ (where $\boldsymbol{\zeta}$ are multipliers associated to (7.30e)*), then*

$$(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q}, \boldsymbol{\zeta}) \text{ is a GNE } \iff \omega_{nm}(q_{nm} + q_{mn}) = 0,\ \forall n \neq m\ , \tag{7.31}$$

and in that case the multipliers associated to (7.22e) *in the GNE problem are defined by $\hat{\zeta}_{nm} = \zeta_{nm} + \omega_{nm}$.*

Proof. For (i), writing the KKT conditions verified by a solution $(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q})$ of the GNE problem (7.22) with Lagrangian multipliers $(\hat{\zeta}_{nm})_{n\neq m}$, it is easy to verify that $(\boldsymbol{x}, \boldsymbol{G}, \boldsymbol{q})$ verifies the KKT conditions of (7.30) $\mathcal{P}_{\hat{\zeta}}^{\text{GNE}}$, where the parameters are taken to $\boldsymbol{\omega} \overset{\text{def}}{=} \hat{\boldsymbol{\zeta}}$.

For (ii), we use the fact that problem (7.30) has linearly independent constraints, and apply [NTF11, Thm. 3.3] directly. □

Proposition 7.7 gives us a characterization of GNEs which enables their computation via a sampling method on $\boldsymbol{\omega}$ and the optimization of parameterized problems (7.30) (see Section 7.5).

7.4.2 Dealing with Congestion

Let us first explicit the following fact on congested lines:

Lemma 7.1. *For any couple of nodes $n \in \mathcal{N}, m \in \Omega_n, m \neq n$, such that $\kappa_{nm} > 0$, $\kappa_{mn} > 0$, $q_{nm} = \kappa_{nm}$ and $q_{mn} = \kappa_{mn}$ cannot hold simultaneously.*

The proof is direct from the capacity and transaction constraints. Then, we obtain the following sufficient condition for a line to be saturated:

Proposition 7.8. *Suppose $\xi_{n0} = \xi_{0n} = 0, \forall n \in \Omega_0$, i.e., there are large line capacities from and to node 0, $c_{n0} = c_{m0}$, i.e., the nodes have the same preferences for node 0, and the node 0 has the same preferences for any node, i.e., $c_{0n} = c_{0m}$. For any couple of nodes $n \in \mathcal{N}, m \in \Omega_n, m \neq n$, asymmetric preferences (such as $c_{mn} > c_{nm}$ or $c_{mn} < c_{nm}$) imply that the node with the smaller preference for the other saturates the line.*

Proof. For any $n \in \mathcal{N}, m \in \Omega_n, m \neq n$, applying Equation (7.14) for the three couples of nodes: $(n,0)$, $(0,m)$, (m,n), we obtain:

$$\begin{aligned} c_{n0} - c_{0n} + \xi_{n0} - \xi_{0n} &= \lambda_n - \lambda_0, \\ c_{0m} - c_{m0} + \xi_{0m} - \xi_{m0} &= \lambda_0 - \lambda_m, \\ c_{mn} - c_{nm} + \xi_{mn} - \xi_{nm} &= \lambda_m - \lambda_n. \end{aligned}$$

Summing up the three equations, we get:

$$\xi_{nm} - \xi_{mn} = (c_{0m} - c_{0n}) + (c_{n0} - c_{m0}) + (\xi_{n0} - \xi_{0n}) + (\xi_{0m} - \xi_{m0}) + (c_{mn} - c_{nm}).$$

Under the assumptions of the proposition, the equation can be simplified to give:

$$\xi_{nm} - \xi_{mn} = c_{mn} - c_{nm}.$$

Then, two cases arise depending on the order of (n,m) preferences:

i) if $c_{mn} > c_{nm}$ (meaning that m wants(buys) to sell to n more(less) than n wants to sell(buy) to m), $\xi_{nm} - \xi_{mn} > 0$, which implies from Lemma 7.1 that $\xi_{nm} > 0$. Then, for the complementarity constraint (7.12e) to hold we need to have $q_{nm} = \kappa_{nm}$, i.e., m saturates the line from m to n;

ii) if $c_{nm} > c_{mn}$ (meaning that n wants(buys) to sell to m more(less) than m wants to sell(buy) to n), $\xi_{nm} - \xi_{mn} < 0$, which implies from Lemma 7.1 that $\xi_{mn} > 0$. Then, for the complementarity constraint (7.12e) to hold we need to have $q_{mn} = \kappa_{mn}$, i.e., n saturates the line from n to m.

□

The following proposition gives a sufficient condition for the distribution grid lines to become congested along a cycle, analog to Proposition 7.5. The proof is similar and is omitted.

Proposition 7.9. *Suppose that there is a sequence of distinct indices $(n_i)_{1 \leqslant i \leqslant k}$ such that $\tilde{C}_{n_i,n_{i+1}} - \tilde{C}_{n_i,n_{i-1}} < 0$ for all $i = 1,\dots,k$, where $n_{k+1} \overset{\text{def}}{=} n_1$. Then, at an equilibrium, there is a trade opposed to the cycle made at full capacity i.e. there exists $i \in \{1,\dots,k\}$ such that $q_{n_{i+1},n_i} = \kappa_{n_{i+1},n_i}$.*

Remark 7.9. *Classically, the Price of Anarchy (PoA) is introduced as a performance measure to assess the performance of the peer-to-peer market design by comparison to the centralized market design. The PoA is defined as the ratio of the social welfare evaluated in the social welfare optimum to the social welfare evaluated in the worst GNE in the set SOL^{GNEP}. Formally, it is defined as follows:*

$$\text{PoA} \overset{\text{def}}{=} \frac{\max_{\boldsymbol{x},\mathbf{G},\mathbf{q}} \text{SW}(\boldsymbol{x},\mathbf{G},\mathbf{q})}{\min_{\boldsymbol{x},\mathbf{G},\mathbf{q} \in SOL^{GNEP}} \text{SW}(\boldsymbol{x},\mathbf{G},\mathbf{q})} \geqslant 1. \tag{7.32}$$

From Proposition 7.6, a VNE maximizes the social welfare SW*, because a VNE coincides with the optimum of the centralized social welfare optimization problem. However, the GNE set might contain equilibria that do not coincide with the (social welfare) optimum solution of the centralized optimization problem, and the* PoA *can be strictly greater than one.*

7.5 Test Cases

7.5.1 A Three Nodes Network with Arbitrage Opportunity

In this section, we illustrate our results and the equilibria on a toy model with only three nodes, indexed by $\mathcal{N} = \{0,1,2\}$, with parameters detailed in Figure 7.2.

The root node 0 has only conventional generation ($\Delta G_0 = 0$) with cost $(a_0, b_0) = (4, 30)$ and $(\underline{G}, \overline{G}) = (0, 10)$. Nodes 1 and 2 are prosumers with RES-based generators ($\overline{G}_n = \underline{G}_n = 0$ and $\Delta G_n > 0$ for $n \in \{1,2\}$). Each node is a consumer (with $(\underline{D}, \overline{D}) = (0, 10)$) and generator

(RES or conventional), therefore producing energy that can be consumed locally to meet demand x_n and exported to the other nodes to meet the unsatisfied demand.

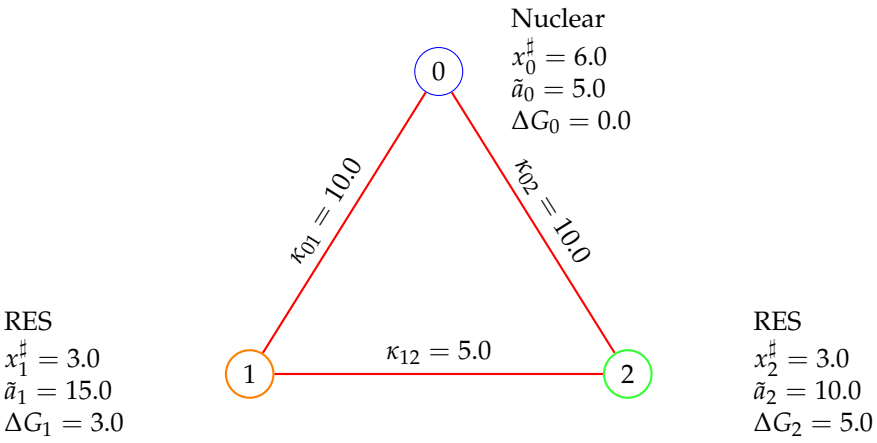


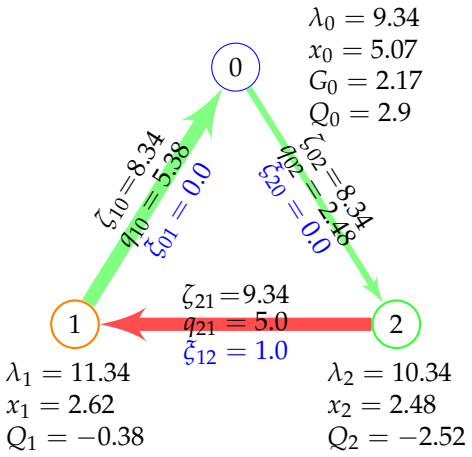
FIGURE 7.2: Three node network example.

Regarding the preferences $(c_{nm})_{nm}$, nodes 1 and 2 both prefer to buy local and to RES-based generators. Node 0 is assumed to be indifferent between buying energy from node 1 or node 2. Capacities are also defined larger from the source node 0 ($\kappa_{0n} = 10$) than between the prosumers nodes ($\kappa_{nm} = 5$).

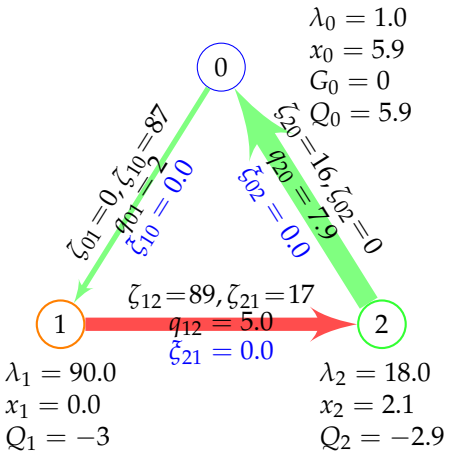
c_{nm}	0	1	2
0	–	1.0	1.0
1	3.0	–	1.0
2	2.0	1.0	–

$c_{nm} - c_{mn}$	0	1	2
0	–	-2.0	-1.0
1	2.0	–	0.0
2	1.0	0.0	–

TABLE 7.1: Price differentiation parameters and matrix of differences.



(A) Centralized solution (SW = 378.3)



(B) One GNE (SW = 255.5)

FIGURE 7.3: Comparison of the optimal centralized solution (a) and a GNE solution with low social welfare (b).

In Figure 7.3 (a), we illustrate the optimal solution of the centralized market design problem in which the global MO maximizes the social welfare under operational and power-flow constraints (7.8a)-(7.8e).

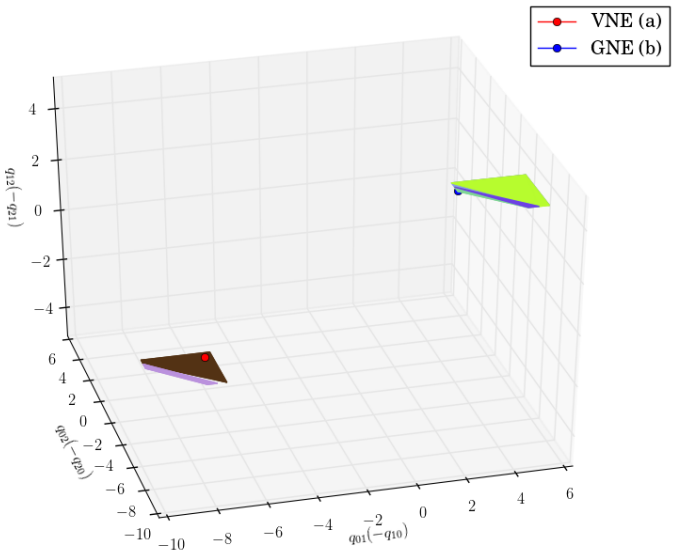


FIGURE 7.4: All existing GNEs in q-space. *The set of GNEs is given as two connected components, corresponding to the edge (1,2) saturated in one way and the other.*

We remark on this figure that the trade from node 1 to node 2 is at full capacity, which is explained by Proposition 7.5. Indeed, we see from Table 7.1 that there is a "cycle" in preferences $\tilde{C}_{01} + \tilde{C}_{12} + \tilde{C}_{20} = -1$ which explains why we obtain $q_{10} = \kappa_{10}$ and $q_{21} = \kappa_{21}$ in the centralized solution (Figure 7.3).

On the contrary, we remark that, in the GNE solution depicted in Figure 7.3b, the same edge is congested in the reverse way: Proposition 7.5 only applies in the case of a centralized solution.

In the example above, the cycle comes from the fact that it is easier for node 2 to buy from 0 than node 1 to buy from 0: thus, the social welfare can be increased if 1 buys from 2 who buys from 0. Changing the parameters to $c_{10} = c_{20} = 3$ removes the cycle in the optimal solution of $(q_{nm})_{nm}$.

In Figure 7.4, we show the different GNEs existing for this reduced problem in the three-dimensional space of transactions. As one can see on this figure, an interesting property is that, for any GNE, the edge from node 1 to node 2 is saturated in one way or the other.

Evaluating the GNE with the lowest social welfare is difficult because this task does not correspond to a convex problem (in particular, the SW is a concave function). However, the GNE depicted in Figure 7.3b is the worst GNE that we found with the sampling method given by Proposition 7.7, using a sampling $(\omega_{nm})_{n>m} \in \{0, \dots, 100\}^3$. Therefore, we can have the following bound on the PoA:

$$\mathrm{PoA} = \frac{\max_{\boldsymbol{x},\mathbf{G},\mathbf{q}} \mathrm{SW}(\boldsymbol{x},\mathbf{G},\mathbf{q})}{\min_{\boldsymbol{x},\mathbf{G},\mathbf{q}\in \mathrm{SOL}^{\mathrm{GNEP}}} \mathrm{SW}(\boldsymbol{x},\mathbf{G},\mathbf{q})} \geqslant \frac{378.3}{255.5} \simeq 1.48\,, \tag{7.33}$$

which means that, in the peer-to-peer market, in the presence of market imperfections, the resulting social welfare can be more than 50% smaller than the optimal social welfare (or, the VE obtained in the absence of market imperfections).

7.5.2 IEEE 14-bus Network

Let us consider the IEEE 14-bus network system as introduced in [Sou+19]. Each bus of the network corresponds to a prosumer in our model as described on Figure 7.5. The buses 3, 4, 5 and 9 to 14 contain only consumers without any production. Nodes 2 and 3 are prosumers node (consumption and RES production) and also contain thermal production plants. The

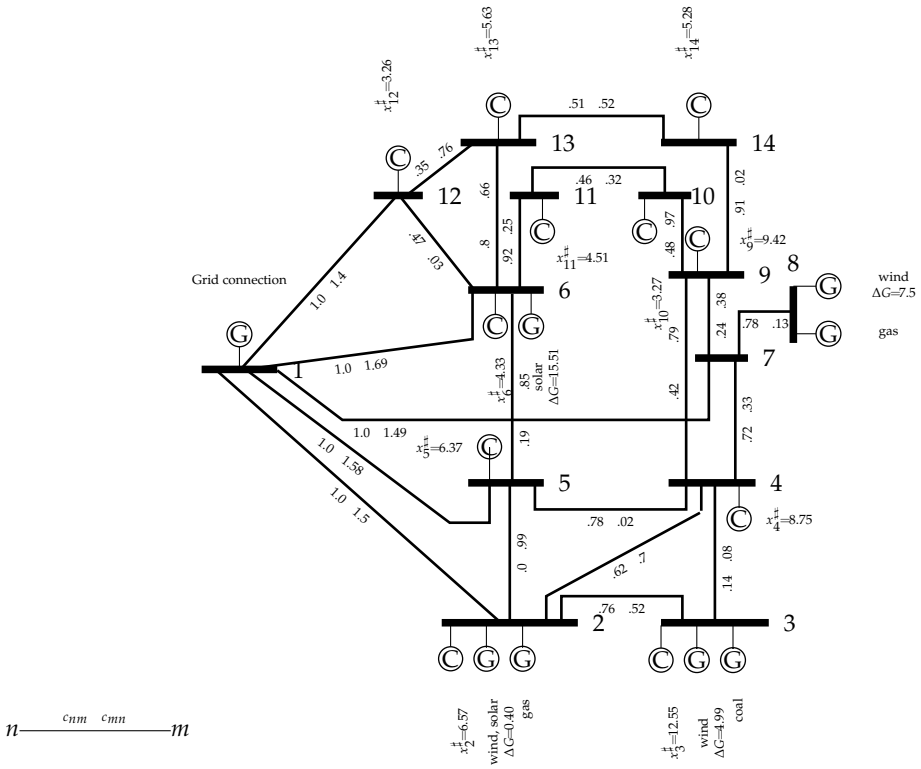


FIGURE 7.5: IEEE 14-bus network system

bus 6 is a prosumer with only intermittent solar energy production. Last, the bus 8 contains only production, renewable and thermal.

The bus 1 corresponding to the grid connection is also able to provide power to the busses linked to it.

Each pair of busses is able to trade with its neighboring busses, up to the capacity of the edge linking the pair of busses.

For simplicity, we compute the trades and optimal productions and consumptions for a particular unique time period. The renewable energy productions $(\Delta G_n)_n$ and the objective consumptions $(D_n^\star)_n$ for this time period are provided in Figure 7.5. Note that in this particular example, we have the inequality:

$$28.39 = \sum_{n\in\mathcal{N}} \Delta G_n < \sum_{n\in\mathcal{N}} D_n^\star = 69.94 \text{ [GWh]} ,$$

which explains partly why we do not have any energy surplus in the solutions depicted on Figure 7.6.

For the trade differentiation prices $(c_{nm})_{n,m}$, we consider four cases:

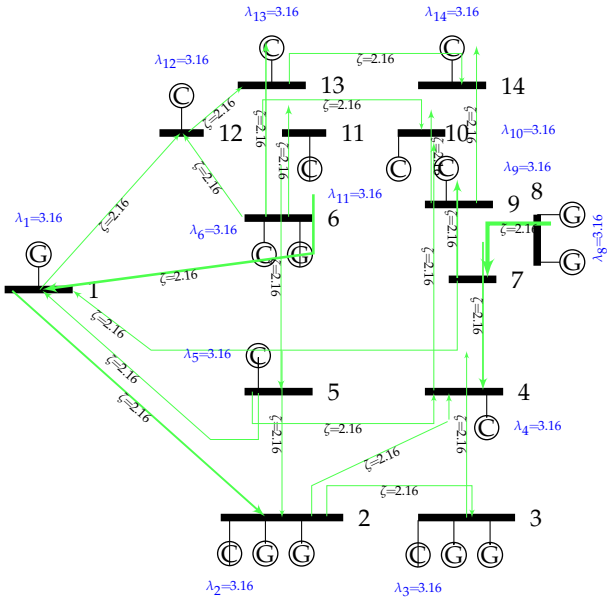
(a) uniform prices: $c_{nm} = 1$ for each n and m, so that we ensure that there does not exist any cycle in the matrix of price differences as described in Proposition 7.5;

(b) heterogeneous prices: for $n \neq 1$ and $m \neq 1$, c_{nm} is chosen uniformly in $[0, 1]$. We assume that agents have a preference for local trades so the price with the grid connection bus c_{n1} is larger and chosen uniformly in $[1, 2]$. The grid connection bus has no preferences so that $c_{1n} = 1$ for each n neighboring bus 1.

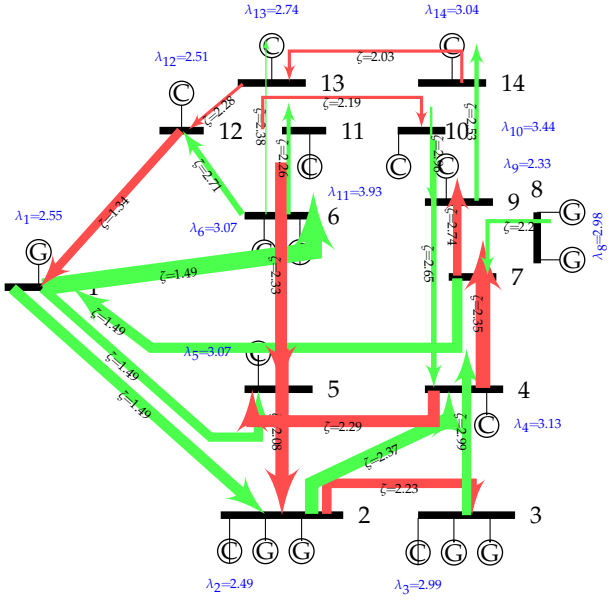
(c) symmetric prices: $(c_{nm})_{nm}$ random and symmetric (for $n < m$, c_{nm} is taken as in (b)).

(d) preferences for local trades with uniform prices: $(c_{nm})_{nm} = 1$ if $m \neq 1$ and $c_{n1} = 3$.

For each of theses cases, we compute the centralized solution (also corresponding to the VNE). The solutions are illustrated in Figure 7.5: directions of trades are represented by arrows, the width of each arrow is proportional to the quantity traded. Trades made at full capacity ($q_{nm} = \kappa_{nm}$) are represented by red arrows, while unsaturated trades are represented by green arrows. We observe that cases (c) and (d) give the same trade solutions $(q_{nm})_{nm}$ at VNE as case (a).



(A) with $(c_{nm})_{nm} = (1)_{nm}$ uniform, SW = 395.28 M$



(B) with $(c_{nm})_{nm}$ random, SW = 560.51 M$

FIGURE 7.6: Trades [$/MWh] at the VNE of the IEEE 14-bus network with homogeneous differentiation prices (left) and heterogeneous differentiation prices (right). *With heterogeneous prices, the quantities traded are larger, and some links become congested. In the homogenous case, marginal trade prices $(\zeta_{nm})_{n,m}$ are all equal. In the case of heterogeneous prices $(c_{nm})_{nm}$, marginal prices $(\zeta_{nm})_{n,m}$ are also heterogeneous.*

We see in Figure 7.6 that the differentiation prices $(c_{nm})_{nm}$ modify completely the solution. We observe that the quantities traded in case (b) are much larger. While some edges are almost unused in case (a) and no edge is congested, ten of the twenty-two edges become

congested in case (b) with heterogeneous prices. This effect can be explained by Propositions 7.5 and 7.8.

Also, we observe that marginal prices $(\zeta_{nm})_{n,m}$ are all equal to 2.16 \$/MWh in case (a), while they are heterogeneous in the case (b). In case (a), the equality is explained by both the absence of congestion ($\xi_{nm} = 0$) and the equality of $(c_{nm})_{nm}$ among users (absence of preferences).

As opposed to the reduced example with three nodes given in Section 7.5.1, it was not possible to compute a GNE different from the VNE for this 14 nodes network. The approach of Nabetani et al. [NTF11] that we used for the three node network is not possible here because of the dimension: to search for another GNE, we have to look on a space of dimension 22, e.g., the number of lines in the network. This observation also calls for the development of algorithms not based on brute force approach, enabling an efficient approximation of the GNEs. This could be the topic of future research.

7.6 Conclusion

We considered two market designs for a network of prosumers with differentiation price preferences: **(i)** a centralized market design where a global operator optimizes the trades to maximize the overall system social welfare, and **(ii)** a fully distributed peer-to-peer market design where prosumers in local energy communities optimize selfishly the trades, demand, and flexibility activation. We discussed the solution of the peer-to-peer market as a Variational Nash Equilibrium and prove that the set of Variational Equilibria coincides with the set of social welfare optima solutions of market design **(i)**. We also discussed the fact that other solutions of the peer-to-peer market **(ii)** may exist as Generalized Nash Equilibria of the problem. We characterized formally the impact of preferences on the network line congestion and energy surplus under both designs. Our results were illustrated in two test cases (a three nodes network and the IEEE 14-bus network). In the three nodes model, we provided a numerical bound on the Price of Anarchy capturing the maximal loss of efficiency caused by market imperfections, obtained at a generalized Nash equilibrium solution. Based on these performance analysis and numerical results, we conclude that peer-to-peer market design gives rise to similar performance than the classical centralized market design, provided market imperfections (resulting from the lack of coordination, insufficient market liquidity, information asymmetry resulting from privacy) can be corrected, and constitutes a relevant evolution for power system operation as it promises more robustness and resilience. Indeed, as the information and decisions are not optimized by a central single entity, in case of failure or if one node is attacked, the power system can still rely on the other nodes. Besides, as all prosumers are involved, they have the ability to adapt their actions to the state of grid.

Several extensions could be considered for further work. First, we could formally include the external aggregator in the study of the underlying economic system. In particular, instead of a premium or feed-in tariff—the sustainability of such mechanisms being questionable—we could consider an external aggregator, seen as a strategic agent, who charges the consumers for their energy surplus as done in [LC+19a; Rob+17]. Another possible extension of the proposed model would be to add taxes on the trades (either a constant tax or a quadratic term), designed by the market operator and in order to regulate or optimize the trades. Last, a different point of view than the one adopted in this chapter would be to consider the framework of cooperative games, in order to study the possibility of prosumers to form stable coalitions in which agents would share their locally produced energy and engage in trades with other coalitions [WH16].

Conclusion and Perspectives

This thesis addresses different problems in the context of decentralized electric systems, through a game-theoretic and an optimization point of view.

In this work we obtained several theoretical results (theorems of existence and uniqueness of *equilibria* in splittable congestion games, bounds on efficiency through the *price of anarchy*, approximation results of the equilibria in large games or games with an infinity of types), algorithmic methods (privacy-preserving disaggregation algorithm, distributed computation of equilibria) and convergence results (convergence of *Best Response*, rate of convergence of alternate projections) as well as modeling results (game-theoretic demand response, peer-to-peer energy exchanges) and real data-based numerical results.

All these results can help electricity operators, as well as other actors of the grid, in better understanding and assessing the potential of electricity consumption flexibilities, anticipating the effects and outcomes of incentive consumer-based mechanisms, and help in the practical implementation and the design of methods for an effective, efficient and decentralized optimization of these flexibilities.

The decentralization of electric systems and the development of the smart grid call for various considerations, and for many problems to be addressed by mathematicians and computer scientists. This thesis focused on the issues related to the aggregation and distributed optimization of consumption flexibilities, and several directions for further research on an extended framework could be investigated.

First, the study of the optimization of price incentives (or of the *billing mechanism*) given to consumers for demand response, would be a natural extension to Part II, where we analyzed two particular billing mechanisms. This issue could be addressed through a model of *bilevel optimization*, a mathematical framework that has not been considered in this thesis, see e.g. [AMDC17].

In the works presented here, we argued that obtaining the equilibrium profiles for players (i.e. consumers) in a *Demand Response* game, could be achieved automatically through electronic consumption schedulers. A different point of view, in line with Chapter 4 but not developed in this thesis, is that players will adapt and learn their profiles through the repetition of the game. To address these learning aspects, we could consider the framework of *dynamic games* or *repeated games* [MSZ15], instead of the "one-shot" games considered in this thesis.

Another natural extension of Part I and Part II we can think of, is to integrate in the model the interactions of the aggregator (in charge of the flexibilities of a set of consumers) with the electricity market or the other actors of the system. From a mathematical point of view, this could also give rise to multi-level optimization considerations, or to "multi-level" games, for instance *Stackelberg games*, see e.g. [Mah+13], [YH15].

Also, the works presented in this thesis adopt a deterministic point of view. As it is often the case when modeling practical situations, one can think of various sources of *stochasticity*. This is especially true in the context of decentralized electric systems: the intermittent generation of renewable energy sources, the prices on a volatile and complex electricity market and the unpredictable or unexpected actions of electricity consumers are all reasonable sources of stochasticity to take into account. Considering stochasticity would make the analysis conducted here much more complex, but a possible framework to study these aspects are *stochastic games*.

Another possible game framework to study demand response would be to consider a *mean field game* model [GLL11], [DGG19a]. This would be a direct extension to Chapter 6

where we considered the limit case of an infinite number of players $N \to \infty$, and would enable to integrate some stochastic aspects (consumption, renewable production) in the model. Mean field games found already some applications in the smart grid context, e.g. [Cou+12].

Besides, the work conducted in this thesis also calls for larger modeling considerations on decentralized electric systems, enlarging the scope of the optimization of consumption flexibilities.

In the questions to be addressed, an important issue for the future of the smart grid is the impact of the new electricity usages (local renewables, flexibilities, electric vehicles) on the physical network, in particular at the distribution level. One aspect of the value of consumption flexibilities is to limit the investments on the network needed to cope with these new usages. Network aspects are promptly discussed in Chapter 7, but considering a realistic power flow model [PL16], [KDS16] in our analysis would be a meaningful extension.

Let us point out that, except in Chapter 1, we considered in this thesis a *noncooperative* game-theoretic setting. An interesting direction of research would be to consider the framework of *cooperative games*, in order to study the possibility of consumers to form coalitions. This seems particularly relevant in the context of energy exchanges between consumers (Chapter 7) and in the context of self shared local production of electricity or shared storage [Cha+18].

Last, as already discussed in the introduction of this thesis, decentralized electric systems involve multiple actors. Not only consumers (or *prosumers*), but also a multiplicity of electricity producers and electricity suppliers competing and interacting with each others. The study and modeling of these competition aspects in a game theoretic environment would also constitute an interesting extension.

Aside from these considerations on modeling aspects of decentralized electric systems, this thesis led to several theoretical and algorithmic contributions in the fields of game theory (in particular congestion games), combinatorial and continuous optimization. These contributions also highlighted theoretical questions that remain open and which also constitute interesting avenues for further research.

First, in Chapter 1, we generalized our privacy-preserving decomposition algorithm to the case of arbitrary polyhedral constraints (Section 1.4). Additional results on the speed of convergence of the algorithm in this case would be interesting. In particular, an open question concerns the number of *cuts* added to the master problem: we did not prove that this number is finite as in the transportation case. In addition, even in this latter case, the upper bound on the number of cuts added remains exponential. In practice however, the number of constraints to consider remains small: a thinner, polynomial, upper bound on the number of cuts would constitute an interesting result.

Then, in Chapter 3, we obtained some convergence results on the Best Response (BR) algorithm in specific cases. The extension of these convergence results to a more general framework is still open, although it is observed numerically for a larger set of prices than affine functions. It is conjectured that this convergence occurs for monotone congestion games. This result would constitute an extension of [HS06] which shows that the *continuous Best Response* converges in zero-sum games. Another avenue for further work in this chapter is the improvement on our bound on the PoA: the bound we obtained for congestion games with affine price functions could be improved to be tighter to the numerical results, and generalized to a larger set of functions.

Another interesting direction for further research, from Chapter 5, is the development of algorithmic methods to compute a generalized equilibrium for a congestion game in a generic form and in the subdifferentiable case. The difficulty is that, due to the presence of coupling constraints, and to the subdifferentiability of the price functions, the standard approaches to solve generalized monotone variational inequalities cannot be directly applied to obtain an efficient and distributed method. However, one may need to exploit the specific structure of the operator, as a combination of monotone subgradients, appearing in the variational characterization of a generalized equilibrium, in order to obtain ad-hoc results, rather than focusing on general results for monotone generalized Variational Inequalities.

Last, Chapter 6 also highlighted the need for efficient algorithms for solving finite dimensional variational inequalities arising as the characterization of symmetric Variational

Wardrop equilibria: some specific methods could be obtained for the case of aggregative games with aggregative constraints.

Besides, the extension of evolutionary dynamics for population games and related algorithms to nonatomic games with infinitely many classes of players is nontrivial but will constitute interesting results. Aggregate constraints is a difficulty for considering evolutionary dynamics in population games, as those dynamics are generally based on unilateral adaptations from players. Another possible research direction in the spirit of Chapter 6 is to consider the possibility of an infinity of heterogeneous players in the framework of Mean Field Games, and obtain appropriate methods for the computation of equilibria.

Last, the results in this chapter are limited to monotone games and the convergence result is limited to strongly monotone games. The extension of our results for non monotone games remains an open question. However, this question is difficult as, even for population games with finitely many types of players, there exist much fewer results on equilibria for games that are not linear, potential or monotone.

Appendix A

Acronyms and Abbreviations

BR Best Response

BRD Best Response Dynamics

DER Distributed Energy Resource

DR Demand Response

DSM Demand Side Management

ECS Electronic Consumption Scheduler

EV Electric Vehicle

GNE Generalized Nash Equilibrium

KKT Karush-Kuhn-Tucker (conditions of optimality)

NE Nash Equilibrium

PoA Price of Anarchy

SC Social Cost

SW Social Welfare

VI Variational Inequality

VNE Variational Nash Equilibrium

WE Wardrop Equilibrium

Appendix B

Introduction (Français)

B.1 Système Électrique : du Centralisé vers le Décentralisé

Les innovations majeures et les nouvelles régulations apparues ces dernières années ont transformé les systèmes électriques de par le monde.

Grâce aux avancées technologiques et à des projets novateurs—l'installation massive de compteurs électriques intelligents [AY15] et les technologies de communication associées, l'intégration des véhicules électriques au réseau [TRY16], accompagnée des projets de recharge intelligente et des concepts *vehicle-to-grid*—de nouvelles perspectives se sont ouvertes pour un système électrique moderne.

Par ailleurs, de nouvelles régulations—ouverture à la concurrence, nouveaux marchés de l'électricité, possibilité de production renouvelable locale—émergeant des diverses institutions et législateurs en charge du système électrique (gouvernements, commission européenne de l'énergie, autorités de la concurrence) ont posé les bases d'un nouveau système électrique décentralisé [AO16]. En France, l'ouverture à la concurrence pour la production d'électricité et la fourniture de gros a débuté en 2000, après un monopole historique et une gestion centralisée par un unique acteur, EDF. L'ouverture à la concurrence pour la fourniture résidentielle a suivi, plus tard, en 2014. Depuis, le nombre de fournisseurs alternatifs a augmenté chaque année, pour atteindre plus de trente en 2019. Le nombre d'installations de production renouvelable a explosé, notamment du fait des faibles coûts des panneaux photovoltaïques et des politiques publiques incitatives. En France en 2019, plus de 36500 sites de productions photovoltaïques sont référencés par RTE [Rte].

Les changements et les différentes innovations sont par ailleurs motivés par des objectifs ambitieux de développement durable et de réduction d'émissions de gaz à effet de serre. Par exemple, l'Union Européenne s'est engagée à atteindre un objectif de réduction de 40% de ses émissions de gaz à effet de serre d'ici 2030 [Com19].

Parallèlement aux multiples actifs de productions d'énergie verte à petite échelle, différentes ressources énergétiques distribuées (DER, *Distributed Energy Resources*) comme les actifs de *stockage distribué* ou les *agrégateurs de flexibilités*, nouveaux acteurs du système électrique, viennent compléter ce paysage décentralisé.

Les agrégateurs de flexibilité [GKS13] agissent comme intermédiaires entre les consommateurs (résidentiels ou commerciaux) et l'opérateur du système électrique, possiblement via un marché dédié. Leur rôle est d'*agréger* un grand nombre de consommations potentiellement flexibles offertes par les consommateurs d'électricité (provenant d'usages flexibles tels que la charge de véhicule électrique, les équipements électroménagers connectés, l'air conditionné, les ballons d'eau chaude, etc, voir Figure B.1). Ces flexibilités, individuellement négligeables, peuvent, une fois agrégées, être valorisées comme un levier d'action sur la demande électrique globale via le marché ou directement comme un service offert à l'opérateur du système (voir Figure B.2).

Du point de vue de l'optimisation du système électrique et de l'équilibre production / demande, ces perspectives changent la façon de voir la demande électrique. Celle-ci était considérée comme un paramètre fixé dans le paradigme précédent, mais doit être considérée dans ce nouveau contexte comme une variable sur laquelle l'opérateur peut exercer un contrôle, bien que ce contrôle soit partiel et décentralisé.

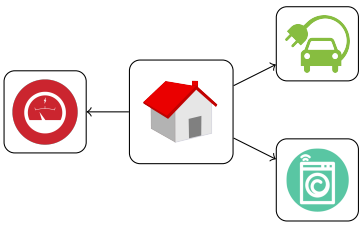


FIGURE B.1: Un consommateur flexible, équipé d'un compteur intelligent et d'un planificateur de consommation relié à un ou plusieurs usages électriques flexibles (véhicule électrique, air conditionné, machine à laver, etc.)

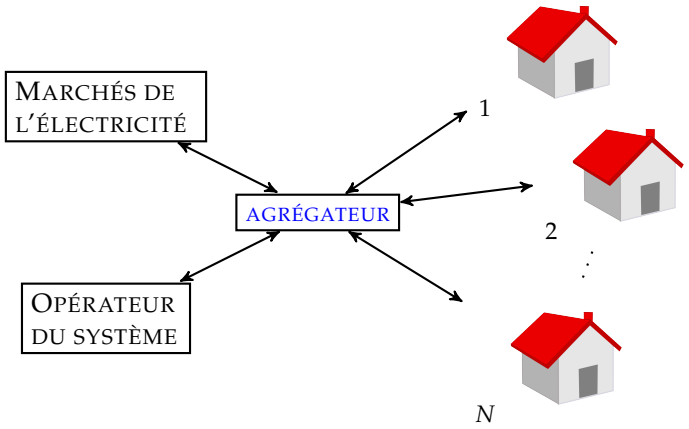


FIGURE B.2: Un agrégateur est en charge des flexibilités de nombreux consommateurs d'électricité individuels, et valorise ensuite ces flexibilités agrégées avec l'opérateur du système et sur les marchés de l'électricité.

Dans ce contexte, le concept de *Demand Response* (DR, qui pourrait être traduit par *réponse de la demande*) réfère à l'ensemble des techniques utilisées pour exploiter les flexibilités de consommations électriques, afin d'assurer l'équilibre production / demande ou des services auxiliaires [Sia14; Saa+12].

L'exploitation des flexibilités de consommation est considérée comme un levier majeur pour atteindre les objectifs de réductions d'émissions de carbone et augmenter de manière significative la part des énergies renouvelables dans la production, mais également pour limiter les investissements sur de nouvelles infrastructures [All11] et assurer la stabilité et la résilience du réseau électrique [Sia14].

Dans cette thèse, nous nous intéressons aux problèmes soulevés par la gestion des flexibilités de consommation, et développons une approche basée sur l'optimisation et la théorie des jeux.

Ce sujet a attiré l'attention de plusieurs communautés scientifiques, dans les domaines des systèmes électriques, des systèmes d'information, de l'optimisation et du contrôle, de la théorie des jeux et de l'informatique : le lecteur pourra se référer à [Den+15], [Sia14] ou [VZV15] pour des revues des travaux sur le sujet.

Fondamentalement, trois aspects rendent ce problème difficile. Tout d'abord, comme nous modélisons les flexibilités de chaque consommateur électrique individuel, cela implique un très grand nombre de variables. Ensuite, les informations concernant les contraintes des consommateurs ne sont pas complètement connues de l'opérateur central ou de l'agrégateur, et restent au niveau local du consommateur. En effet, un opérateur ne saurait avoir accès à l'ensemble de ces informations, d'une part du fait de la quantité et de la variabilité des données que cela représente, et d'autre part pour des raisons de confidentialité : les consommateurs sont en général réticents à fournir des informations à caractère privé, comme les heures de présence à leur foyer. Enfin, le paradigme décentralisé suppose qu'un

opérateur central possède seulement un contrôle partiel et indirect sur les profils de consommation de consommateurs flexibles, par exemple par le biais d'incitations ou de signaux qui peuvent être reçus localement par chaque consommateur, par un compteur communicant relié à un dispositif de planification des consommations (Figure B.1).

Le cadre conceptuel que nous considérons, en particulier pour les chapitres I et II, est le suivant : un opérateur central, par exemple l'opérateur d'un micro-réseau (*microgrid*) possédant des actifs de production d'électricité (moyens de production, contrats de fournitures, etc.) ou un agrégateur de flexibilités en interaction sur les marchés de l'électricité ou avec d'autres acteurs (voir Figure B.2), est en charge des flexibilités d'un ensemble de consommateurs. Dans un contexte de *smart grid*, chaque consommateur est équipé d'un planificateur de consommation électrique (ECS, *Energy Consumption Scheduler*), intégré à un compteur intelligent. Cet ECS est connecté à l'opérateur central (via le réseau électrique ou un réseau de communication ad-hoc), duquel il peut recevoir mais auquel il peut aussi transmettre des signaux. Il est également relié localement à divers appareils électriques flexibles (véhicule électrique, air conditionné, etc., voir Figure B.1), et dispose d'une petite capacité de calcul permettant l'implémentation locale d'algorithmes pour planifier la consommation de ces appareils.

B.2 Cadre Mathématique : de l'Optimisation aux Jeux

La transition d'un paradigme centralisé vers le paradigme décentralisé, tel que décrit dans la section précédente, requiert de remplacer les outils d'optimisation utilisés classiquement par des méthodes de théorie des jeux. Dans cette section, nous passons en revue les principales notions mathématiques utilisées dans ce contexte.

B.2.1 Optimisation Distribuée pour un Problème Centralisé

Adoptons le point de vue d'un opérateur central faisant face à une fonction de coût $\boldsymbol{p}, \boldsymbol{x} \mapsto f(\boldsymbol{p}, \boldsymbol{x})$, qui dépend de variables $\boldsymbol{p} \in \mathcal{P}$ qu'il contrôle directement, mais, possède également une dépendance en d'autres variables $\boldsymbol{x} \stackrel{\text{def}}{=} (\boldsymbol{x}_n)_{n\in\mathcal{N}}$, où $\mathcal{N}$ dénote un ensemble d'acteurs distincts de l'opérateur, par exemple un ensemble de consommateurs électriques flexibles. Chaque variable individuelle $\boldsymbol{x}_n$ (correspondant à l'agent n), est également soumise à des contraintes locales $\boldsymbol{x}_n \in \mathcal{X}_n$. Si les variables individuelles $(\boldsymbol{x}_n)_n$ sont contrôlées par l' opérateur , ce dernier fait face au problème d'optimisation centralisé suivant :

$$\min_{\boldsymbol{p},\boldsymbol{x}} f(\boldsymbol{p}, \boldsymbol{x}) \tag{B.1a}$$

$$\text{s.c. } \boldsymbol{p} \in \mathcal{P} \tag{B.1b}$$

$$\boldsymbol{x}_n \in \mathcal{X}_n, \ \forall n \in \mathcal{N}, \tag{B.1c}$$

$$\boldsymbol{h}(\boldsymbol{p}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \leqslant 0, \tag{B.1d}$$

où (B.1d) représente un couplage des variables de l'opérateur avec les variables des autres agents. Ce couplage peut être simple (par exemple si $\boldsymbol{p}$ représente le profil agrégé de l' ensemble des consommateurs $\mathcal{N}$, i.e. $\boldsymbol{p} = \sum_n \boldsymbol{x}_n$) mais aussi très complexe (par exemple pour modéliser les contraintes de flux du réseau électrique).

Une approche standard en optimisation pour résoudre les problèmes de type (1) d'une manière distribuée est de s'appuyer sur des techniques de décomposition [Coh78], [BT89], [BV04], en particulier sur les méthodes de décomposition lagrangienne [PC06], [XJB04], telles que les méthodes de *sous-gradients dual* [Ber99, Chapter 6] ou *ADMM* [GM75]. Ces approches ont notamment été appliquées dans le contexte des systèmes électriques décentralisés et de DR [Sag12], [Mol+17], [CBK17], [Shi+14], [LCL11], [Den+15]. Le principal inconvénient de ces méthodes est qu'elles s'appuient sur une hypothèse de convexité du problème (B.1), ce qui est assez restrictif pour les cas d'applications : par exemple, les contraintes des actifs de productions d'électricité [CA06], ou encore les contraintes de réseau, sont non-convexes. Dans le chapitre 1, nous proposons une méthode originale permettant la

résolution de problèmes de type (B.1) d'une manière décentralisée, et pour laquelle la convexité de la fonction f ou de l'ensemble $\mathcal{P}$ ne sont pas nécessaires. Notre méthode calcule une solution du problème ci-dessus sans que l'opérateur ne soit informé des contraintes et profils individuels des agents, garantissant ainsi la confidentialité des utilisateurs.

B.2.2 Modèles de Jeux pour la Gestion d'Énergie

Le problème centralisé (B.1) ne prend pas en compte le fait que les agents puissent être *stratégiques*, au sens que ces agents choisissent des actions (l'agent n contrôle la variable $\boldsymbol{x}_n$) et que ces actions influent sur des objectifs individuels (coût, facture d'énergie ou confort) qui peuvent être en compétition les uns les autres, ou en compétition avec l'objectif f de l'opérateur. Afin de modéliser ces aspects stratégiques, une large communauté de chercheurs a modélisé le problème de *Demand Response* dans le cadre de la théorie des jeux, e.g. [MR+10], [CK14], [Cha+14], [Saa+12], [Che+14], [Atz+13], [LCL11], [Bah+13].

Un cadre théorique différent—non sans liens avec les jeux—qui pourrait être considéré pour aborder le problème (B.1) et prendre en compte des objectifs individuels pour les agents, est l'*optimisation bi-niveaux* [CMS07]. Cependant, afin d'éviter la complexité de la programmation bi-niveaux, nous nous intéressons dans cette thèse à des problèmes et des modèles *à un seul niveau* dans les cadres de l'optimisation distribuée et de la théorie des jeux.

Un *jeu* [FT91] est une situation où un ensemble $\mathcal{N}$ d'*agents stratégiques* ou *joueurs* (par exemple, des consommateurs flexibles) sont en interaction. Chacun des joueurs $n \in \mathcal{N}$ choisit des actions dans un ensemble admissible, dénoté par $\mathcal{X}_n$ (e.g. l'ensemble des profils de consommation électrique réalisables).

L'interaction entre les joueurs est modélisée par une fonction coût individuelle, pour chaque joueur $n \in \mathcal{N}$, $(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \mapsto f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \in \mathbb{R}$, qui dépend non seulement de sa propre action $\boldsymbol{x}_n$ mais aussi des actions des autres joueurs $\boldsymbol{x}_{-n} \stackrel{\text{def}}{=} (\boldsymbol{x}_m)_{m \neq n}$.

Dans le contexte DR, nous pouvons modéliser les interactions entre consommateurs dans ce cadre, avec f_n représentant une *facture d'énergie* que le consommateur n souhaite minimiser. La dépendance de f_n aux profils de consommation des autres agents $\boldsymbol{x}_{-n}$ résulte du fait que l'opérateur central souhaite éviter la synchronisation des consommations. En effet, la demande résultante de l'ensemble des consommateurs détermine les coûts d' approvisionnement : plus la demande est importante sur un créneau temporel, plus le prix d'approvisionnement et de distribution d'électricité sera élevé, du fait des coûts de production marginaux croissants et des phénomènes de congestion.

Dans ce cadre, nous pouvons écrire un problème d'optimisation pour chaque agent n, étant données les actions des autres joueurs $\boldsymbol{x}_{-n}$:

$$\min_{\boldsymbol{x}_n} f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \tag{B.2a}$$

$$\text{s.c.} \quad \boldsymbol{x}_n \in \mathcal{X}_n \, . \tag{B.2b}$$

Dans le cadre des jeux non coopératifs, une notion fondamentale est donnée par le concept d'*équilibre de Nash* (NE, *Nash Equilibrium*) [Nas50]. Cette notion a été retenue comme le concept central de *solution* en théorie des jeux, avec des applications très diverses [Nis+07, Sec.1.3.3].

Une situation d'équilibre de Nash correspond à des *profils d'actions* $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_n)_{n \in \mathcal{N}}$ tels que chaque joueur $n \in \mathcal{N}$, si il considère l'action des autres joueurs $\hat{\boldsymbol{x}}_{-n}$ comme fixe, n'a pas intérêt à dévier de son action actuelle $\boldsymbol{x}_n$ (le joueur résout le problème de minimisation (B.2)). On a donc :

$$\forall n \in \mathcal{N}, \forall \boldsymbol{x}_n \in \mathcal{X}_n, \;\; f_n(\hat{\boldsymbol{x}}_n, \hat{\boldsymbol{x}}_{-n}) \leqslant f_n(\boldsymbol{x}_n, \hat{\boldsymbol{x}}_{-n}) \, .$$

L'équilibre de Nash capture la notion de solution stable, à partir de laquelle aucun joueur ne peut décroître sa fonction de coût en changeant d'action. Lorsque les joueurs sont dans une telle situation, il est dans l'intérêt de chacun d'eux de conserver l'action choisie.

B.2.3 Jeux de Congestion

Dans le contexte de DR, la classe des *jeux de congestion*, introduite par Rosenthal [Ros73b], s'avère particulièrement importante. Dans ces jeux, chaque joueur choisit un sous-ensemble d'un ensemble partagé de ressources $\mathcal{T}$. Chaque ressource $t \in \mathcal{T}$ possède un coût $c_t(\cdot)$, fonction croissante du nombre de joueurs sélectionnant cette ressource.

Dans le cas où l'action de chaque joueur n est réduite à la *sélection* de ressources, c'est-à-dire donnée par un vecteur $\boldsymbol{x}_n \in \{0,1\}^{\mathcal{T}}$, nous obtenons une classe de jeux nommée *jeux de congestion non-fractionnables* (*unsplittable* en anglais). Lorsque chaque joueur n décide d'une charge sur chacune des ressources de $\mathcal{T}$, c'est-à-dire d'un profil $\boldsymbol{x}_n \in \mathbb{R}_+^{\mathcal{T}}$, on parle d'un *jeu de congestion fractionnable* (*splittable*) [ORS93], [Wan12b, Sec. 1.3.3]. Cette extension a trouvé de premières applications dans les problèmes de routage sur des réseaux.

Ce cadre des jeux de congestion fractionnables est adapté à notre contexte. Si l'on considère que chaque agent choisit un profil de consommation $\boldsymbol{x}_n = (x_{n,t})_{t\in\mathcal{T}} \in \mathcal{X}_n$ sur un ensemble de créneaux temporels $\mathcal{T} = \{1, \dots, T\}$, et que sa fonction coût f_n correspond à une facture d'énergie avec, pour chaque période $t \in \mathcal{T}$, un prix unitaire de l'électricité $X_t \mapsto c_t(X_t)$, fonction de la consommation agrégée $X_t = \sum_{n\in\mathcal{N}} x_{n,t}$ à cette période t, on obtient une fonction de coût de type congestion donnée par :

$$f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) \stackrel{\text{def}}{=} \sum_{t\in\mathcal{T}} x_{n,t} c_t\Big(\sum_{m\in\mathcal{N}} x_{m,t}\Big). \tag{B.3}$$

Ce cadre correspond bien à un jeu de congestion fractionnable, ou jeu de routage, décrit par [ORS93], sur un réseau de T arcs parallèles représentant les T créneaux temporels, et où chaque fonction coût $c_t(.)$ correspond à une fonction de latence sur chaque arc, comme illustré en Figure 3.

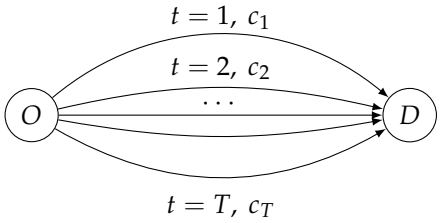


FIGURE B.3: Un jeu de congestion pour la consommation d'énergie sur T périodes de temps peut s'interpréter comme un jeu de routage sur un réseau de T arcs parallèles.

B.2.4 Efficacité des Équilibres

L'évaluation de l'*efficacité* d'un équilibre, ou plus généralement d'une issue d'un jeu, peut suivre plusieurs critères selon la situation : par exemple, dans notre cadre, un opérateur du système électrique s'intéresserait à l'évaluation des coûts totaux de production, ou bien des émissions de CO2 associées aux profils de consommation. Par ailleurs, un critère standard, largement adopté dans la littérature des jeux de congestion, est le *coût social*, notion qui réfère simplement à la somme des coûts de tous les joueurs, c'est à dire $\mathrm{SC}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} f_n(\boldsymbol{x})$.

Basée sur ce critère, une mesure quantitative de l'*efficacité* des équilibres d'un jeu est donnée par la notion de *Prix de l'Anarchie* (PoA, *Price of Anarchy*) [KP99], donné par le rapport du coût social maximal atteignable à un équilibre (pire équilibre) et du coût social optimal pouvant être atteint (cette situation optimale serait celle donnée par un opérateur central optimisant le coût social et choisissant les actions de tous les joueurs), c'est-à-dire :

$$\mathrm{PoA} = \frac{\max_{\hat{\boldsymbol{x}}\in\boldsymbol{\mathcal{X}}^{\mathrm{NE}}} \mathrm{SC}(\hat{\boldsymbol{x}})}{\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \mathrm{SC}(\boldsymbol{x})},$$

où $\boldsymbol{\mathcal{X}}^{\mathrm{NE}}$ désigne l'ensemble des NEs possibles du jeu. Un équilibre correspond à une situation localement stable mais à des choix individualistes et il n'y donc pas de raisons *a priori* qu'une telle situation minimise le coût social. Ainsi le PoA est généralement supérieur à 1. Différents travaux dans la littérature des jeux de congestion ont réussi à prouver des bornes sur le PoA dans certains cas particuliers : par exemple [CBCS16] et [CB+17] ont montré récemment que, dans les jeux de congestion sur réseaux, le PoA convergeait vers 1 dans la limite de demandes infiniment grandes. Dans la littérature appliquée au contexte *smart grid*, le PoA a aussi été adopté comme référence pour mesurer l'efficacité d'un système basé sur un jeu (voir par exemple [CBK17], [Zhu+12], [CK14], [Saa+12], [NAC14], [HRD14]).

Comme annoncé précédemment, la théorie des jeux offre un cadre naturel pour aborder les problèmes d'optimisation de façon distribuée et décentralisée [MRP13], [MS15] : un opérateur central, en charge de plusieurs agents (par exemple des consommateurs d'électricité), choisit les paramètres (incitations, tarifs, signaux) d'un jeu auquel les agents prennent part.

Dans ce contexte, un PoA faible caractérise l'efficience des incitations établies par l' opérateur. Pour ces considérations de type *mechanism design*, introduisons une autre classe importante de jeux, appelés *jeux de potentiel* (*potential games*) [MS96] : un jeu est un *jeu de potentiel exact* si il existe une fonction "potentiel" $\Phi : (\mathcal{X}_n)_{n\in\mathcal{N}} \to \mathbb{R}$ que tous les joueurs ont intérêt à minimiser, ou plus exactement, qui vérifie :

$$\forall n \in \mathcal{N},\ \forall \boldsymbol{x} \in (\mathcal{X}_m)_{m\in\mathcal{N}},\ \forall x'_n \in \mathcal{X}_n,\ f_n(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) - f_n(\boldsymbol{x}'_n, \boldsymbol{x}_{-n}) = \Phi(\boldsymbol{x}_n, \boldsymbol{x}_{-n}) - \Phi(\boldsymbol{x}'_n, \boldsymbol{x}_{-n})\ .$$

En particulier, si un opérateur central établit un jeu qui possède un potentiel [LM13], les NEs de ce jeu correspondent à des minima locaux de la fonction Φ. Si ce potentiel Φ est convexe, les minima d'autre part seront facile à calculer. Plusieurs travaux, dans le contexte DR, ont suivi une approche basée sur un jeu de potentiel, voir par exemple [Tus+18], [Wu+11], [BW15], [YH16].

B.3 Organisation du Manuscrit

Cette thèse est organisée en quatre parties. Les travaux et modèles sont présentés dans un ordre croissant de décentralisation et d'autonomie pour les consommateurs.

• La partie I aborde le problème de la gestion des flexibilités de consommation dans le cadre d'optimisation distribuée. Le point de vue adopté dans cette partie est celui d'un opérateur central en charge de ressources décentralisées (flexibilités) pour de nombreux agents, et souhaitant optimiser ces ressources distribuées tout en respectant la confidentialité des données des agents. Dans le chapitre 1, nous commençons par définir le cadre général d'un *problème distribué d'allocation de ressources*—les ressources correspondant dans notre cas particulier à l'énergie allouée à chaque période de temps—auquel fait face l'opérateur central en charge des agents. Chaque agent possède également des contraintes individuelles sur les profils qui lui sont admissibles, et il souhaite garder ces contraintes *confidentielles*. Nous proposons une méthode qui calcule une solution optimale de ce problème et qui garantit la confidentialité des agents, et nous prouvons la validité et l'efficacité (nombre d'itérations de l'algorithme) de cette méthode.

• La partie II considère le problème de la gestion des flexibilités de consommation *(Demand Response*), mais dans le cadre de la théorie des jeux. Dans les chapitres 2 et 3, les jeux sont davantage utilisés comme un outil pour l'optimisation décentralisée : un *mécanisme de prix*, défini par un opérateur, correspond à un signal de prix envoyé aux consommateurs et à des incitations à atteindre un *équilibre* dans le jeu correspondant. Nous comparons deux mécanismes de prix dans le chapitre 2 puis, dans le chapitre 3, nous nous focalisons sur un mécanisme de prix "par heure"—correspondant à un jeu de congestion atomique fractionnable—et sur les aspects de calcul d'équilibres dans ce jeu. Dans le chapitre 4, nous modélisons la possibilité pour les consommateurs d'avoir des objectifs locaux et potentiellement divergents : en plus de la "facture" donnée par l'opérateur, chaque consommateur possède un profil de consommation préférentiel.

• La partie III prouve des résultats théoriques sur l'approximation d'équilibres dans des jeux de congestion avec un très grand nombre de joueurs. Ceci peut s'appliquer au cas où les joueurs représentent des consommateurs électriques individuels, puisque l'on obtient ainsi un jeu avec des centaines de milliers de joueurs. Dans le chapitre 5, nous considérons un jeu de congestion avec un grand nombre de joueurs et la présence de *contraintes couplantes* entre ces joueurs, et nous montrons qu'un équilibre de Nash (NE) de ce jeu peut être approximé par un *équilibre de Wardrop* d'un jeu *nonatomique* (avec un continuum de joueurs), où les joueurs similaires sont regroupés en populations homogènes. Dans le chapitre 6, nous nous plaçons dans le cadre des *jeux non atomiques agrégatifs généralisés* avec une infinité de types de joueurs. Cette situation modélise par exemple les interactions d'une population de consommateurs, dont les caractéristiques (besoin en énergie, disponibilités temporelles) sont décrites par une fonction de distribution paramétrique. Nous montrons que dans ce jeu, un *équilibre de Wardrop variationnel*, une notion d'équilibre que nous définissons, peut être approximé par un équilibre de Wardrop d'un jeu de population de plus petite dimension.

• La partie IV va plus loin en termes de décentralisation, en considérant une situation sans opérateur central et où les consommateurs échangent de l'énergie directement dans des transactions *pair-à-pair*. Dans le chapitre 7, nous proposons un modèle original basé sur un *jeu généralisé*. Chaque joueur représente un consommateur flexible possédant une source de production renouvelable locale et qui, en plus d'un objectif local, possède des préférences sur les transactions qu'il peut engager avec ses pairs. Nous étudions les équilibres généralisés de ce jeu et, en particulier, nous montrons que les *équilibres variationnels* correspondent aux *optima sociaux* du jeu.

B.4 Contributions des Chapitres

Ce travail de thèse apporte des contributions de différentes natures : résultats théoriques et algorithmiques, expérimentations numériques et aspects de modélisation.

En effet, cette thèse introduit et analyse différents modèles et outils dans le cadre de l'optimisation d'un système électrique décentralisé, décrit en Section B.1, tout en apportant plusieurs résultats théoriques dans les domaines de l'optimisation décentralisée et de la théorie des jeux, en particulier dans le cadre des jeux de congestion fractionnables. Les principales contributions sont données ci-dessous, suivant leur ordre d'apparition :

Chapitre 1. Nous proposons un algorithme original (Algorithme 1.4) permettant de calculer un profil optimal agrégé d'allocation de ressources, solution d'un problème d' optimisation possiblement non convexe, et correspondant à l'*agrégation* de profils admissibles pour un ensemble d'agents ayant chacun des contraintes individuelles. Cet algorithme est distribué et préserve la confidentialité, au sens que les calculs sont effectués sans que les contraintes et profils individuels des agents ne soient révélés à un tiers, ni un autre agent, ni l'opérateur central. En pratique, cet algorithme peut être utilisé par exemple pour optimiser un profil de consommation électrique agrégé sur un ensemble de consommateurs flexibles, en assurant la confidentialité des informations de chaque consommateur. L'application de cette méthode dans ce contexte a fait l'objet d'un dépôt de brevet par EDF [Jac+18b].

Notre algorithme repose sur la méthode des projections alternées [VN50], [GPR67] sur des *polytopes de transport*, et un résultat clef pour l'algorithme est prouvé dans ce chapitre 1 : le théorème 1.3 montre que, lorsque les projections alternées ne convergent pas vers un même point, nous pouvons construire une coupe polyhédrale, correspondant à une coupe d'Hoffman [Hof60] dans le problème de flot associé à ce polytope de transport.

Enfin, le théorème 1.4, dont la preuve repose sur des arguments de théorie spectrale des graphes [CG97], donne une borne explicite de la vitesse de convergence des projections alternées sur des polytopes de transport. Ce résultat montre que la complexité temporelle de la méthode proposée dans ce chapitre évolue de façon linéaire avec le nombre d'agents considéré, et que cette méthode peut donc être utilisée à large échelle.

Chapitre 2. Nous considérons deux mécanismes de *Demand Response*, l'un basé sur un jeu de potentiel (dénoté mécanisme *Daily Proportional*, [MR+10]) et l'autre basé sur un jeu de

congestion fractionnable (dénoté mécanisme *Hourly Proportional*, [Bah+13]), et nous étudions les équilibres de Nash de ces jeux, en particulier l'existence et l'unicité de ces équilibres. Le théorème 2.1 donne une condition suffisante sur les dérivées des fonctions prix (correspondant aux fonctions de *latence* dans un jeu de congestion) pour l'unicité d'un équilibre. Ce résultat repose sur la condition standard de stricte monotonie [Ros65], explicitée en utilisant un résultat d'algèbre linéaire (perturbation de spectre matriciel). Dans le même chapitre, le théorème 2.2 établit une borne supérieure sur le prix de l'anarchie (défini en Section B.2) pour cette classe de jeu, obtenue par la technique de (λ, μ)-*régularité locale* introduite par [RS15].

Chapitre 3. Ce chapitre se focalise sur les aspects algorithmiques des jeux de congestion. Tout d'abord, le théorème 3.2 étend le résultat [ORS93] d'unicité d'un NE à un contexte plus général. Nous donnons différents résultats sur la convergence de l'algorithme de *meilleure réponse* (BR, *Best Response*, voir [GM91], [DG16], [DGG19b]), dans des cadres particuliers des jeux de congestion fractionnables. Le théorème 3.4 prouve la convergence de l'algorithme de BR cyclique, dans le cas potentiel, tandis que le théorème 3.2 prouve la convergence géométrique d'une version aléatoire de BR dans des cas spécifiques. Ces résultats montrent la convergence de l'algorithme BR dans un cadre différent de celui considéré par Mertzios [Mer08]. Nous étudions également une méthode de *gradient projeté*, et démontrons (théorème 3.5) que, sous une hypothèse de forte monotonie, cette méthode converge à une vitesse géométrique vers l'unique équilibre du jeu. En s'appuyant sur le même résultat de perturbation spectrale utilisé au chapitre précédent, la proposition 3.1 donne une condition suffisante sur les fonctions prix $c_t(.)$ (voir Section B.2) pour que la forte monotonie soit respectée. Enfin, dans ce chapitre, nous proposons une procédure de *demand response "online"* avec horizons glissants (Algorithme 3.4), en *commande prédictive* [Wu+11], afin de prendre en compte la mise à jour de prévisions dans un environnement stochastique. Le théorème 3.6 prouve que les profils de consommations calculés par cette procédure correspondent à l'équilibre de Nash, sous l'hypothèse de prévisions parfaites ;

Chapitre 4. Nous étendons le modèle de jeu de flexibilités de consommations électriques introduit au chapitre 2 en considérant des profils de consommation préférentiels pour les consommateurs. Dans le cadre simplifié de deux créneaux temporels, nous établissons plusieurs résultats théoriques sur l'impact de ces préférences sur l'équilibre du jeu, dans les deux mécanismes introduits au chapitre 2. Nous comparons le prix de l'anarchie dans les deux jeux correspondants ainsi que le *prix de l'efficacité*, un concept similaire au PoA que nous introduisons pour mesurer l'efficacité des équilibres du point de vue des coûts de l'opérateur du système. En particulier, le théorème 4.4 montre que les coûts du système pour le mécanisme *hourly proportional* sont toujours inférieurs à ceux obtenus dans le mécanisme *daily proportional*. Ce chapitre présente également des résultats numériques basés sur des données de consommation réelles, provenant de la base de données *PecanStreet* [Pec].

Chapitre 5. Nous considérons dans ce chapitre le cadre des jeux de congestion fractionnables avec *contraintes couplantes* [Har91], et avec un très grand nombre de joueurs, hétérogènes par leurs contraintes individuelles ou leur fonction objectif. En utilisant le cadre des inégalités variationnelles [FP07], nous montrons en théorème 5.1 que les équilibres de Wardrop Variationnels (VWE) fournissent une bonne approximation des équilibres de Nash Variationnels (VNE), et en théorème 5.2 que, en regroupant les joueurs similaires en populations homogènes, nous définissons un jeu approchant du jeu initial, et dont les équilibres sont proches du jeu initial. Ces résultats étendent le travail [Pac+18], où les auteurs montrent que les VNEs sont proches des VWEs dans les jeux agrégatifs atomiques avec un grand nombre de joueurs : le chapitre 5 montre que, de plus, nous pouvons considérer un petit nombre de populations dans le jeu approchant, et que le VWE obtenu reste proche d'un VNE du jeu initial. Ces résultats d'approximations peuvent être utilisés pour calculer rapidement un équilibre, puisque le problème caractérisant les équilibres dans le jeu approchant est de dimension bien inférieure au problème initial.

Chapitre 6. Nous introduisons une nouvelle famille de jeux, les jeux agrégatifs non atomiques avec une *infinité de types de joueurs* (Definition 6.1), où le *type* d'un joueur est défini par son ensemble d'actions admissibles et sa fonction objectif. Nous étendons ainsi la famille

de jeux considérée par [MZ97]. Ces jeux émergent naturellement lorsque l'on considère une population caractérisée par une fonction de distribution paramétrique. En présence de contraintes couplantes, nous définissons la notion d'*équilibre de Wardrop Variationnel* (VWE) pour un jeu nonatomique avec une infinité de types (Definition 6.4), caractérisé par une inégalité variationnelle de dimension infinie. Le théorème 6.5 prouve le résultat d'approximation suivant : un VWE dans un jeu nonatomique peut être approximé par un équilibre de Wardrop d'un jeu avec un nombre fini de populations homogènes, ce dernier équilibre étant symétrique au sens que les joueurs au sein d'une même population jouent la même action. Ce résultat présente également un fort intérêt numériquement, puisqu'il permet d'approcher l'équilibre du jeu initial (de dimension infinie) en restant en petite dimension.

Chapitre 7. Nous introduisons un modèle original d'échanges d'énergie *pair-à-pair* (*peer-to-peer*) au sein d'une communauté de consommateurs-producteurs d'électricité (*prosumers*). Le modèle obtenu est un *jeu généralisé* [HP90], puisque les contraintes de réciprocité des échanges définissent des contraintes couplantes entre chaque consommateur et ses voisins. Ce modèle de jeu permet de prendre en compte le caractère stratégique des interactions entre consommateurs, contrairement aux approches de décomposition et optimisation distribuée qui ont été considérées pour ce problème [SBP18]. Un des résultats obtenus dans ce chapitre est le suivant: la proposition 7.6 montre que les équilibres variationnels du jeu formulé correspondent à des optima du *coût social* (voir Section B.2). Nous étudions numériquement l'efficacité des équilibres, à travers le PoA, en calculant les *équilibres généralisés* du jeu sur différents cas tests, et nous comparons les flux d'énergie obtenus aux équilibres avec les flux obtenus lors de la solution centralisée, correspondant à l'optimum social.

Chacun des sept chapitres présentés dans cette thèse est basé soit sur un article accepté et présenté à une conférence avec comité de lecture, soit sur un article de journal, publié ou bien soumis pour publication:

- le chapitre 1 est basé sur l'article [Jac+19b] soumis pour publication à un journal, et sur l'article [Jac+19a] présenté à *2019 IEEE 58th Conference on Decision and Control* ;
- le chapitre 2 est basé sur l'article [Jac+17b] présenté à *2017 IEEE Conference on Innovative Smart Grid Technologies (ISGT)* ;
- le chapitre 3 est basé sur l'article de journal [Jac+19c], publié dans *IEEE Transactions on Smart Grid* ;
- le chapitre 4 est basé sur l'article [Jac+17a] présenté à *2017 IEEE Conference on Smart Grid Communications* ;
- le chapitre 5 est basé sur l'article [Jac+18a] soumis pour publication ;
- le chapitre 6 est basé sur l'article [JW19] soumis pour publication, ainsi que sur la pré-publication [JW18a] et l'article de conférence [JW18b], présenté à *2018 IEEE 57th Conference on Decision and Control* ;
- le chapitre 7 est basé sur l'article [LC+19b], accepté pour publication à *European Journal of Operational Research*.

Bibliography

[AB06] Charalambos D. Aliprantis and Kim Border. *Infinite Dimensional Analysis – A Hitchhiker's Guide*. 3rd ed. Springer-Verlag Berlin Heidelberg, 2006.

[AKH02] Eitan Altman, Hisao Kameda, and Yoshihisa Hosokawa. "Nash equilibria in load balancing in distributed computer systems". In: *International Game Theory Review* 4.02 (2002), pp. 91–100.

[All11] Hunt Allcott. "Rethinking real-time electricity pricing". In: *Resource and energy economics* 33.4 (2011), pp. 820–842.

[Alt+02] Eitan Altman, Tamer Basar, Tania Jimenez, and Nahum Shimkin. "Competitive routing in networks with polynomial costs". In: *IEEE Transactions on Automatic Control* 47.1 (2002), pp. 92–96.

[Alt+06] Eitan Altman, Thomas Boulogne, Rachid El-Azouzi, Tania Jiménez, and Laura Wynter. "A survey on networking games in telecommunications". In: *Computers & Operations Research* 33.2 (2006), pp. 286–311.

[ALT18] Miguel F Anjos, Andrea Lodi, and Mathieu Tanneau. "A decentralized framework for the optimal coordination of distributed energy resources". In: *IEEE Transactions on Power Systems* 34.1 (2018), pp. 349–359.

[AMDC17] Miguel Asensio, Gregorio Munoz-Delgado, and Javier Contreras. "Bi-level approach to distribution network and renewable energy expansion planning considering demand response". In: *IEEE Transactions on Power Systems* 32.6 (2017), pp. 4298–4309.

[AN02] Hesham K Alfares and Mohammad Nazeeruddin. "Electric load forecasting: literature survey and classification of methods". In: *International Journal of Systems Science* 33.1 (2002), pp. 23–34.

[AN79] Yash P Aneja and Kunhiraman PK Nair. "Bicriteria transportation problem". In: *Management Science* 25.1 (1979), pp. 73–78.

[AO16] Emmanuel Asane-Otoo. "Competition policies and environmental quality: Empirical analysis of the electricity sector in OECD countries". In: *Energy Policy* 95 (2016), pp. 212–223.

[AR08] Heiner Ackermann and Heiko Röglin. "On the convergence time of the best response dynamics in player-specific congestion games". In: *arXiv preprint arXiv:0805.1130* (2008).

[AT09] Alfred Auslender and Marc Teboulle. "Projected subgradient methods with non-Euclidean distances for non-differentiable convex minimization and variational inequalities". In: *Mathematical Programming* 120.1 (2009), pp. 27–48.

[Atz+13] Italo Atzeni, Luis G Ordóñez, Gesualdo Scutari, Daniel P Palomar, and Javier Rodríguez Fonollosa. "Demand-side management via distributed energy generation and storage optimization". In: *IEEE Transactions on Smart Grid* 4.2 (2013), pp. 866–876.

[Aum65] Robert J. Aumann. "Integrals of set-valued functions". In: *Journal of Mathematical Analysis and Applications* 12.1 (1965), pp. 1–12.

[Aum69] Robert J. Aumann. "Measurable utility and the measurable choice problem". In: *La Décision*. Paris. Colloque Internationaux du C. N. R. S., 1969, pp. 15–26.

[Aum76] Robert J. Aumann. "An elementary proof that integration preserves uppersemicontinuity". In: *Journal of Mathematical Economics* 3.1 (1976), pp. 15–18.

[AY15] Damminda Alahakoon and Xinghuo Yu. "Smart electricity meter data intelligence for future energy systems: A survey". In: *IEEE Transactions on Industrial Informatics* 12.1 (2015), pp. 425–436.

[Bah+13] Zahra Baharlouei, Massoud Hashemi, Hamed Narimani, and Hamed Mohsenian-Rad. "Achieving optimality and fairness in autonomous demand response: Benchmarks and billing mechanisms". In: *IEEE Transactions on Smart Grid* 4.2 (2013), pp. 968–975.

[Ban+14] Rupak Baniya, Mikko Maksimainen, Seppo Sierla, Cheng Pang, Chen-Wei Yang, and Valeriy Vyatkin. "Smart indoor lighting control: Power, illuminance, and colour quality". In: *Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium on*. IEEE. 2014, pp. 1745–1750.

[Bat87] Robert G Batson. "Combinatorial behavior of extreme points of perturbed polyhedra". In: *Journal of mathematical analysis and applications* 127.1 (1987), pp. 130–139.

[BB93] Heinz H Bauschke and Jonathan M Borwein. "On the convergence of Von Neumann's alternating projection algorithm for two sets". In: *Set-Valued Analysis* 1.2 (1993), pp. 185–212.

[BC11] Heinz H. Bauschke and Patrick Louis Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. 2nd ed. Springer International Publishing, 2011.

[BCW15] Heinz H Bauschke, Jiawei Chen, and Xianfu Wang. "A Bregman projection method for approximating fixed points of quasi-Bregman nonexpansive mappings". In: *Applicable Analysis* 94.1 (2015), pp. 75–84.

[Bea+16] Olivier Beaude, Samson Lasaulce, Martin Hennebel, and Ibrahim Mohand-Kaci. "Reducing the Impact of EV Charging Operations on the Distribution Network". In: *IEEE Transactions on Smart Grid* 7.6 (2016), pp. 2666–2679.

[Ben+18] Saif Benjaafar, Guangwen Kong, Xiang Li, and Costas Courcoubetis. "Peer-to-peer product sharing: Implications for ownership, usage, and social welfare in the sharing economy". In: *Management Science* 65.2 (2018), pp. 477–493.

[Ben62] Jacques F Benders. "Partitioning procedures for solving mixed variables programming problems". In: *Numerische mathematik* 4.1 (1962), pp. 238–252.

[Ber99] Dimitri P Bertsekas. *Nonlinear programming*. Athena Scientific, 1999.

[BH14] Zahra Baharlouei and Massoud Hashemi. "Efficiency-fairness trade-off in privacy-preserving autonomous demand side management". In: *IEEE Transactions on Smart Grid* 5.2 (2014), pp. 799–808.

[Bha+09] Umang Bhaskar, Lisa Fleischer, Darrell Hoy, and Chien-Chung Huang. "Equilibria of Atomic Flow Games are not Unique". In: *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*. 2009, pp. 748–757.

[BLY14] Jonathan M Borwein, Guoyin Li, and Liangjin Yao. "Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets". In: *SIAM Journal of Optimization* 24.1 (2014), pp. 498–527.

[BMW56] M. Beckmann, C. B. McGuire, and C. B. Winston. *Studies in the Economics of Transportation*. Yale University Press, 1956.

[BNH17] Zahra Baharlouei, Hamed Narimani, and Massoud Hashemi. "On the Convergence Properties of Autonomous Demand Side Management Algorithms". In: *IEEE Transactions on Smart Grid* (2017).

[BNMR12] Zahra Baharlouei, Hamed Narimani, and Hamed Mohsenian-Rad. "Tackling co-existence and fairness challenges in autonomous Demand Side Management". In: *IEEE Global Communications Conference (GLOBECOM)*. IEEE. 2012, pp. 3159–3164.

[Bol72] Ethan D Bolker. "Transportation polytopes". In: *Journal of Combinatorial Theory, Series B* 13.3 (1972), pp. 251–262.

[BPS13] Olivier Brun, Balakrishna J Prabhu, and Tatiana Seregina. "On the convergence of the best-response algorithm in routing games". In: *Proc.of the 7th Int. Conf. on Performance Evaluation Methodologies and Tools*. ICST. 2013, pp. 136–144.

[Bro+00] Luce Brotcorne, Martine Labbé, Patrice Marcotte, and Gilles Savard. "A Bilevel Model and Solution Algorithm for a Freight Tariff-Setting Problem". In: *Transportation Science* 34.3 (2000), pp. 289–302.

[Bro+01] Luce Brotcorne, Martine Labbé, Patrice Marcotte, and Gilles Savard. "A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network". In: *Transportation Science* 35.4 (2001), pp. 345–358.

[Bru84] Peter Brucker. "An O(n) algorithm for quadratic knapsack problems". In: *Operations Research Letters* 3.3 (1984), pp. 163–166.

[BT13] Amir Beck and Luba Tetruashvili. "On the convergence of block coordinate descent type methods". In: *SIAM journal on Optimization* 23.4 (2013), pp. 2037–2060.

[BT89] Dimitri P Bertsekas and John N Tsitsiklis. *Parallel and distributed computation: numerical methods*. Vol. 23. Prentice hall Englewood Cliffs, NJ, 1989.

[BTN01] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[BW15] Shahab Bahrami and Vincent WS Wong. "A potential game framework for charging PHEVs in smart grid". In: *Communications, Computers and Signal Processing (PACRIM), 2015 IEEE Pacific Rim Conference on*. IEEE. 2015, pp. 28–33.

[BWH17] Shahab Bahrami, Vincent WS Wong, and Jianwei Huang. "An online learning algorithm for demand response in smart grid". In: *IEEE Transactions on Smart Grid* (2017).

[CA06] Miguel Carrión and José M Arroyo. "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem". In: *IEEE Transactions on Power Systems* 21.3 (2006), pp. 1371 –1378.

[CB+17] Riccardo Colini-Baldeschi, Roberto Cominetti, Panayotis Mertikopoulos, and Marco Scarsini. "On the asymptotic behavior of the price of anarchy: Is selfish routing bad in highly congested networks". In: *ArXiv e-prints* (2017).

[CBCS16] Riccardo Colini-Baldeschi, Roberto Cominetti, and Marco Scarsini. "On the price of anarchy of highly congested nonatomic network games". In: *International Symposium on Algorithmic Game Theory*. Springer. 2016, pp. 117–128.

[CBK17] Pratyush Chakraborty, Enrique Baeyens, and Pramod P Khargonekar. "Distributed control of flexible demand using proportional allocation mechanism in a smart grid: Game theoretic interaction and price of anarchy". In: *Sustainable Energy, Grids and Networks* 12 (2017), pp. 30–39.

[CG97] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. 92. American Mathematical Soc., 1997.

[Cha+14] Bo Chai, Jiming Chen, Zaiyue Yang, and Yan Zhang. "Demand response management with multiple utility companies: A two-level game approach". In: *IEEE Transactions on Smart Grid* 5.2 (2014), pp. 722–731.

[Cha+18] Pratyush Chakraborty, Enrique Baeyens, Kameshwar Poolla, Pramod P Khargonekar, and Pravin Varaiya. "Sharing storage in a smart grid: A coalitional game approach". In: *IEEE Transactions on Smart Grid* (2018).

[Che+10] Lijun Chen, Na Li, Steven H Low, and John C Doyle. "Two market models for demand response in power networks". In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. 2010, pp. 397–402.

[Che+14] He Chen, Yonghui Li, Raymond HY Louie, and Branka Vucetic. "Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach". In: *IEEE Transactions on Smart Grid* 5.4 (2014), pp. 1744–1754.

[CK05] George Christodoulou and Elias Koutsoupias. "The price of anarchy of finite congestion games". In: *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. ACM. 2005, pp. 67–73.

[CK14] Pratyush Chakraborty and Pramod P Khargonekar. "A demand response game and its robust price of anarchy". In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE. 2014, pp. 644–649.

[CKS11] Chen Chen, Shalinee Kishore, and Lawrence V Snyder. "An innovative RTP-based residential power scheduling scheme for smart grids". In: *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE. 2011, pp. 5956–5959.

[Cla71] Edward H Clarke. "Multipart pricing of public goods". In: *Public choice* 11.1 (1971), pp. 17–33.

[CM87] Paul H Calamai and Jorge J Moré. "Projected gradient methods for linearly constrained problems". In: *Mathematical Programming* 39.1 (1987), pp. 93–116.

[CMS03] Jean-Philippe Côté, Patrice Marcotte, and Gilles Savard. "A bilevel modelling approach to pricing and fare optimisation in the airline industry". In: *Journal of Revenue and Pricing Management* 2.1 (2003), pp. 23–36.

[CMS07] Benoît Colson, Patrice Marcotte, and Gilles Savard. "An overview of bilevel optimization". In: *Annals of Operations Research* 153.1 (2007), pp. 235–256.

[Coh78] Guy Cohen. "Optimization by decomposition and coordination: A unified approach". In: *IEEE Transactions on automatic control* 23.2 (1978), pp. 222–232.

[Coh88] G Cohen. "Auxiliary problem principle extended to variational inequalities". In: *Journal of Optimization Theory and Applications* 59.2 (1988), pp. 325–333.

[Com19] European Commission. *Clean energy for all Europeans*. 2019.

[Con16] Laurent Condat. "Fast projection onto the simplex and the l1 ball". In: *Mathematical Programming* 158.1-2 (2016), pp. 575–585.

[Coo+09] William J Cook, WH Cunningham, WR Pulleyblank, and A Schrijver. *Combinatorial optimization*. Springer, 2009.

[Cor94] Luis C. Corchón. "Comparative statics for aggregative games the strong concavity case". In: *Mathematical Social Sciences* 28.3 (1994), pp. 151–165.

[Cos] *Coserv Electricity Distributor*. 2017. URL: http://www.coserv.com/Customer-Service/Electric-Rates-And-Tariff.

[Cou+12] Romain Couillet, Samir Medina Perlaza, Hamidou Tembine, and Mérouane Debbah. "A mean field game analysis of electric vehicles in the smart grid". In: *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*. IEEE. 2012, pp. 79–84.

[CP09] Guilherme Carmona and Konrad Podczeck. "On the existence of pure-strategy equilibria in large games". In: *Journal of Economic Theory* 144.3 (2009), pp. 1300–1319.

[CP82] D. Chan and J. S. Pang. "The Generalized Quasi-Variational Inequality Problem". In: *Mathematics of Operations Research* 7.2 (1982), pp. 211–222.

[CPP02] E. Cavazzuti, M. Pappalardo, and M. Passacantando. "Nash Equilibria, Variational Inequalities, and Dynamical Systems". In: *Journal of Optimization Theory and Applications* 114.3 (2002), pp. 491–506.

[CSSM04] José R. Correa, Andreas S. Schulz, and Nicolás E. Stier-Moses. "Selfish Routing in Capacitated Networks". In: *Mathematics of Operations Research* 29.4 (2004), pp. 961–976.

[CY11] Yunmei Chen and Xiaojing Ye. "Projection onto a simplex". In: *arXiv preprint arXiv:1101.6081* (2011).

[CZ84] Guy Cohen and Dao Li Zhu. "Decomposition and coordination methods in large scale optimization problems: The nondifferentiable case and the use of augmented Lagrangians". In: *Adv. in Large Scale Systems 1* (1984), pp. 203–266.

[Den+15] Ruilong Deng, Zaiyue Yang, Mo-Yuen Chow, and Jiming Chen. "A survey on demand response in smart grids: Mathematical models and approaches". In: *IEEE Transactions on Industrial Informatics* 11.3 (2015), pp. 570–582.

[Den+17] Wei Deng, Ming-Jun Lai, Zhimin Peng, and Wotao Yin. "Parallel multi-block ADMM with O (1/k) convergence". In: *Journal of Scientific Computing* 71.2 (2017), pp. 712–736.

[DG16] Stéphane Durand and Bruno Gaujal. "Complexity and optimality of the best response algorithm in random potential games". In: *International Symposium on Algorithmic Game Theory*. Springer. 2016, pp. 40–51.

[DGG19a] Josu Doncel, Nicolas Gast, and Bruno Gaujal. "Discrete Mean Field Games: Existence of Equilibria and Convergence". In: *Journal of Dynamics & Games* (2019), pp. 1–19.

[DGG19b] Stéphane Durand, Federica Garin, and Bruno Gaujal. "Distributed best response dynamics with high playing rates in potential games". In: *Performance Evaluation* 129 (2019), pp. 40–59.

[DT96] Xie Ping Ding and Enayet Tarafdar. "Monotone generalized variational inequalities and generalized complementarity problems". In: *Journal of Optimization Theory and Applications* 88.1 (1996), pp. 107–122.

[EAD16] Jonas Engels, Hamada Almasalma, and Geert Deconinck. "A distributed gossip-based voltage control algorithm for peer-to-peer microgrids". In: *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE. 2016, pp. 370–375.

[EFL16] Liran Einav, Chiara Farronato, and Jonathan Levin. "Peer-to-peer markets". In: *Annual Review of Economics* 8 (2016), pp. 615–635.

[Eli+13] Jocelyne Elias, Fabio Martignon, Lin Chen, and Eitan Altman. "Joint operator pricing and network selection game in cognitive radio networks: Equilibrium, system dynamics and price of anarchy". In: *IEEE Transactions on Vehicular Technology* 62.9 (2013), pp. 4576–4589.

[FA01] Murat Fahrioglu and Fernando L Alvarado. "Using utility information to calibrate customer demand management behavior models". In: *IEEE transactions on power systems* 16.2 (2001), pp. 317–322.

[FGV18] Magda Foti, Dimitrios Greasidis, and Manolis Vavalis. "Viability Analysis of a Decentralized Energy Market Based on Blockchain". In: *2018 15th International Conference on the European Energy Market (EEM)*. IEEE. 2018, pp. 1–5.

[FHW17] Zhixuan Fang, Longbo Huang, and Adam Wierman. "Prices and subsidies in the sharing economy". In: *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee. 2017, pp. 53–62.

[FJ08] Doerte Fouquet and Thomas B Johansson. "European renewable energy policy at crossroads—Focus on electricity support mechanisms". In: *Energy policy* 36.11 (2008), pp. 4079–4092.

[FK07] Francisco Facchinei and Christian Kanzow. "Generalized Nash equilibrium problems". In: *4OR: A Quarterly Journal of Operations Research* 5.3 (2007), pp. 173–210.

[FK10] Francisco Facchinei and Christian Kanzow. "Generalized Nash Equilibrium Problems". In: *Annals of Operations Research* 175.1 (2010), pp. 177–211.

[FP07] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[FPT04] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. "The complexity of pure Nash equilibria". In: *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*. ACM. 2004, pp. 604–612.

[Fri+94] Terry L. Friesz, David Bernstein, Nihal J. Mehta, Roger L. Tobin, and Saiid Ganjalizadeh. "Day-To-Day Dynamic Network Disequilibria and Idealized Traveler Information Systems". In: *Operations Research* 42.6 (1994), pp. 1120–1136.

[FT91] Drew Fudenberg and Jean Tirole. *Game theory, 1991*. Cambridge, Massachusetts, 1991, p. 393.

[Fuk11] Masao Fukushima. "Restricted generalized Nash equilibria and controlled penalty algorithm". In: *Computational Management Science* 8.3 (2011), pp. 201–218.

[Fuk86] Masao Fukushima. "A relaxed projection method for variational inequalities". In: *Mathematical Programming* 35.1 (1986), pp. 58–70.

[Gab+12] Steven A Gabriel, Antonio J Conejo, J David Fuller, Benjamin F Hobbs, and Carlos Ruiz. *Complementarity modeling in energy markets*. Vol. 180. Springer Science & Business Media, 2012.

[GJW82] Michael R Garey, David Johnson, and Hans Witsenhausen. "The complexity of the generalized Lloyd-max problem (corresp.)" In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 255–256.

[GKS13] Lazaros Gkatzikis, Iordanis Koutsopoulos, and Theodoros Salonidis. "The role of aggregators in smart grid demand response markets". In: *IEEE Journal on Selected Areas in Communications* 31.7 (2013), pp. 1247–1257.

[GLL11] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. "Mean field games and applications". In: *Paris-Princeton lectures on mathematical finance 2010*. Springer, 2011, pp. 205–266.

[GM14] David Gregoratti and Javier Matamoros. "Distributed energy trading: The multiple -microgrid case". In: *IEEE Transactions on industrial Electronics* 62.4 (2014), pp. 2551–2559.

[GM75] Roland Glowinski and A Marroco. "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires". In: *ESAIM* 9.R2 (1975), pp. 41–76.

[GM91] Itzhak Gilboa and Akihiko Matsui. "Social stability and equilibrium". In: *Journal of the Econometric Society* (1991), pp. 859–867.

[GPR67] L.G. Gubin, Boris T. Polyak, and E.V. Raik. "The method of projections for finding the common point of convex sets". In: *USSR Comput. Math. & Math. Phys.* 7.6 (1967), pp. 1 –24. ISSN: 0041-5553.

[Gra17] Sergio Grammatico. "Dynamic Control of Agents Playing Aggregative Games With Coupling Constraints". In: *IEEE Transactions on Automatic Control* 62.9 (2017), pp. 4537–4548.

[Had17] Saeed Hadikhanloo. "Learning in anonymous nonatomic games with applications to first-order mean field games". In: *arXiv preprint arXiv: 1704.00378* (2017).

[HAF05] Bernardo A Huberman, Eytan Adar, and Leslie R Fine. "Valuating privacy". In: *IEEE security & privacy* 3.5 (2005), pp. 22–25.

[Har91] Patrick T Harker. "Generalized Nash games and quasi-variational inequalities". In: *European journal of Operational research* 54.1 (1991), pp. 81–94.

[HB16] Josef Hofbauer and Peter Bednarik. "Discretized best-response dynamics for the rock-paper-scissors game". In: *Journal of Dynamics and Games* 4.1 (2016), pp. 75–86.

[HJ12] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[HKL80] R Helgason, J Kennington, and H Lall. "A polynomially bounded algorithm for a singly constrained quadratic program". In: *Mathematical Programming* 18.1 (1980), pp. 338–343.

[HM85] Alain Haurie and Patrice Marcotte. "On the relationship between Nash — Cournot and Wardrop equilibria". In: *Networks* 15.3 (1985), pp. 295–308.

[Hof60] Alan J Hoffman. "Some recent applications of the theory of linear inequalities to extremal combinatorial analysis". In: *Proc. of Symposia on Applied Mathematics*. 1960, pp. 113–127.

[Hon+17] Mingyi Hong, Xiangfeng Wang, Meisam Razaviyayn, and Zhi-Quan Luo. "Iteration complexity analysis of block coordinate descent methods". In: *Mathematical Programming* 163.1-2 (2017), pp. 85–114.

[HP90] Patrick T Harker and Jong-Shi Pang. "Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications". In: *Mathematical Programming* 48.1-3 (1990), pp. 161–220.

[HRD14] Qingqing Huang, Mardavij Roozbehani, and Munther A Dahleh. "Efficiency-risk tradeoffs in electricity markets with dynamic demand response". In: *IEEE Transactions on Smart Grid* 6.1 (2014), pp. 279–290.

[HS06] Josef Hofbauer and Sylvain Sorin. "Best response dynamics for continuous zero-sum games". In: *Discrete and Continuous Dynamical Systems Series B* 6.1 (2006), p. 215.

[HS09] Josef Hofbauer and William H Sandholm. "Stable games and their dynamics". In: *Journal of Economic theory* 144.4 (2009), pp. 1665–1693.

[HS66] Philip Hartman and Guido Stampacchia. "On some non-linear elliptic differential-functional equations". In: *Acta Mathematica* 115 (1966), pp. 271–310.

[HS98] Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[IA09] Ali Ipakchi and Farrokh Albuyeh. "Grid of the future". In: *IEEE power and energy magazine* 7.2 (2009), pp. 52–62.

[JMT05] Ramesh Johari, Shie Mannor, and John N Tsitsiklis. "Efficiency loss in a network resource allocation game: the case of elastic supply". In: *IEEE Transactions on Automatic Control* 50.11 (2005), pp. 1712–1724.

[Jog+17] Olamide Jogunola, Augustine Ikpehai, Kelvin Anoh, Bamidele Adebisi, Mohammad Hammoudeh, Haris Gacanin, and Georgina Harris. "Comparative analysis of P2P architectures for energy trading and sharing". In: *Energies* 11.1 (2017), p. 62.

[JPW06] Geetha Jagannathan, Krishnan Pillaipakkamnatt, and Rebecca N Wright. "A new privacy-preserving distributed k-clustering algorithm". In: *Proc. of the 2006 SIAM Int. Conf. on Data Mining*. SIAM. 2006, pp. 494–498.

[JT06] Ramesh Johari and John N Tsitsiklis. "A scalable network resource allocation mechanism with bounded efficiency loss". In: *IEEE Journal on Selected Areas in Communications* 24.5 (2006), pp. 992–999.

[Kat+08] Farid Katiraei, Reza Iravani, Nikos Hatziargyriou, and Aris Dimeas. "Microgrids management". In: *IEEE power and energy magazine* 6.3 (2008).

[KDS16] Burak Kocuk, Santanu S Dey, and X Andy Sun. "Strong socp relaxations for the optimal power flow problem". In: *Operations Research* (2016).

[KP99] Elias Koutsoupias and Christos Papadimitriou. "Worst-case equilibria". In: *Annual Symposium on Theoretical Aspects of Computer Science*. Springer. 1999, pp. 404–413.

[Kra+14] Matt Kraning, Eric Chu, Javad Lavaei, Stephen Boyd, et al. "Dynamic network energy management via proximal message passing". In: *Foundations and Trends® in Optimization* 1.2 (2014), pp. 73–126.

[KS12] Ankur A. Kulkarni and Uday V. Shanbhag. "On the variational equilibrium as a refinement of the generalized Nash equilibrium". In: *Automatica* 48.1 (2012), pp. 45–55.

[KW18] Max Klimm and Philipp Warode. "Parametrized Nash Equilibria in Atomic Splittable Congestion Games via Weighted Block Laplacians". In: *arXiv preprint arXiv:1811.08354* (2018).

[LC+19a] Hélène Le Cadre, Bernardo Pagnoncelli, Tito Homem-De-Mello, and Olivier Beaude. "Designing coalition-based fair and stable pricing mechanisms under private information on consumers' reservation prices". In: *European Journal of Operational Research* 272.1 (2019), pp. 270–291.

[LC18] Hélène Le Cadre. "On the efficiency of local electricity markets under decentralized and centralized designs: a multi-leader Stackelberg game analysis". In: *Central European Journal of Operations Research* (2018), pp. 1–32.

[LCL11] Na Li, Lijun Chen, and Steven H Low. "Optimal demand response based on utility maximization in power networks". In: *Power and Energy Society General Meeting*. IEEE. 2011, pp. 1–8.

[LCMP19] Hélène Le Cadre, Ilyès Mezghani, and Anthony Papavasiliou. "A game-theoretic analysis of transmission-distribution system operator coordination". In: *European Journal of Operational Research* 274.1 (2019), pp. 317–339.

[Liu+17] Nian Liu, Xinghuo Yu, Cheng Wang, Chaojie Li, Li Ma, and Jinyong Lei. "Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers". In: *IEEE Transactions on Power Systems* 32.5 (2017), pp. 3569–3583.

[LLC95] K. K. Lai, Kokin Lam, and W. K. Chan. "Shipping Container Logistics and Allocation". In: *Journal of the Operational Research Society* 46.6 (1995), pp. 687–697.

[Llo82] Stuart Lloyd. "Least squares quantization in PCM". In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 129–137.

[LM13] Na Li and Jason R Marden. "Designing games for distributed optimization". In: *IEEE Journal of Selected Topics in Signal Processing* 7.2 (2013), pp. 230–242.

[LMS98] Martine Labbé, Patrice Marcotte, and Gilles Savard. "A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing". In: *Management Science* 44.12-part-1 (1998), pp. 1608–1622.

[LNN95] Claude Lemaréchal, Arkadii Nemirovskii, and Yurii Nesterov. "New variants of bundle methods". In: *Mathematical Programming* 69.1-3 (1995), pp. 111–147.

[Lon+17] Chao Long, Jianzhong Wu, Chenghua Zhang, Lee Thomas, Meng Cheng, and Nick Jenkins. "Peer-to-peer energy trading in a community microgrid". In: *2017 IEEE Power & Energy Society General Meeting*. IEEE. 2017, pp. 1–5.

[LP99] Torbj´orn Larsson and Michael Patriksson. "Side constrained traffic equilibrium models – analysis, computation and applications". In: *Transportation Research Part B: Methodological* 33.4 (1999), pp. 233–264.

[Ma+82] Perng-Yi Richard Ma et al. "A task allocation model for distributed computing systems". In: *IEEE Transactions on Computers* 100.1 (1982), pp. 41–47.

[Mah+13] Sabita Maharjan, Quanyan Zhu, Yan Zhang, Stein Gjessing, and Tamer Basar. "Dependable demand response management in the smart grid: A Stackelberg game approach". In: *IEEE Transactions on Smart Grid* 4.1 (2013), pp. 120–132.

[MC84] Andreu Mas-Colell. "On a theorem of Schmeidler". In: *Journal of Mathematical Economics* 13.3 (1984), pp. 201–206.

[Mer08] George B Mertzios. "Fast convergence of routing games with splittable flows". In: *Proc. 2nd Int. Conf. Th. Math. Foundations Comp. Sci. (TMFCS)*. 2008, pp. 28–33.

[MGD12] Javier Matamoros, David Gregoratti, and Mischa Dohler. "Microgrids energy trading in islanding mode". In: *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*. IEEE. 2012, pp. 49–54.

[Mic86] Christian Michelot. "A finite algorithm for finding the projection of a point onto the canonical simplex of $\mathbb{R}^n$". In: *Journal of Optimization Theory and Applications* 50.1 (1986), pp. 195–200.

[Mil00] Igal Milchtaich. "Generic uniqueness of equilibrium in large crowding games". In: *Mathematics of Operations Research* 25.3 (2000), pp. 349–364.

[MM18] Thomas Morstyn and Malcolm McCulloch. "Multi-class energy management for peer-to-peer energy trading driven by prosumer preferences". In: *IEEE Transactions on Power Systems* (2018).

[MNS04] Patrice Marcotte, Sang Nguyen, and Alexandre Schoeb. "A Strategic Flow Model of Traffic Assignment in Static Capacitated Networks". In: *Operations Research* 52.2 (2004), pp. 191–212.

[Mol+17] Daniel K Molzahn, Florian Dörfler, Henrik Sandberg, Steven H Low, Sambuddha Chakrabarti, Ross Baldick, and Javad Lavaei. "A survey of distributed optimization and control algorithms for electric power systems". In: *IEEE Transactions on Smart Grid* 8.6 (2017), pp. 2941–2962.

[Mor+18] Fabio Moret, T Baroche, E Sorin, and P Pinson. "Negotiation algorithms for peer-to-peer electricity markets: Computational properties". In: *2018 Power Systems Computation Conference (PSCC)*. IEEE. 2018, pp. 1–7.

[MP07] Patrice Marcotte and Michael Patriksson. "Traffic Equilibrium". In: *Transportation*. Ed. by Cynthia Barnhart and Gilbert Laporte. Vol. 14. Handbooks in Operations Research and Management Science. Elsevier, 2007, pp. 623–713.

[MP18] Fabio Moret and Pierre Pinson. "Energy collectives: a community and fairness based approach to future electricity markets". In: *IEEE Transactions on Power Systems* (2018).

[MR+10] Amir-Hamed Mohsenian-Rad, Vincent WS Wong, Juri Jatskevich, Robert Schober, and Alberto Leon-Garcia. "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid". In: *IEEE Transactions on Smart Grid* 1 (2010), pp. 320–331.

[MRLG10] Amir-Hamed Mohsenian-Rad and Alberto Leon-Garcia. "Optimal residential load control with price prediction in real-time electricity pricing environments". In: *IEEE Transactions on Smart Grid* 1.2 (2010), pp. 120–133.

[MRP13] Jason R Marden, Shalom D Ruben, and Lucy Y Pao. "A model-free approach to wind farm control using game theoretic methods". In: *IEEE Transactions on Control Systems Technology* 21.4 (2013), pp. 1207–1214.

[MS01] Hervé Moulin and Scott Shenker. "Strategyproof sharing of submodular costs: budget balance versus efficiency". In: *Economic Theory* 18.3 (2001), pp. 511–533.

[MS15] Jason R Marden and Jeff S Shamma. "Game theory and distributed control". In: *Handbook of game theory with economic applications*. Vol. 4. Elsevier, 2015, pp. 861–899.

[MS82] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.

[MS96] Dov Monderer and Lloyd S Shapley. "Potential games". In: *Games and Economic Behavior* 14.1 (1996), pp. 124–143.

[MSZ15] Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. *Repeated games*. Vol. 55. Cambridge University Press, 2015.

[MTM18] Thomas Morstyn, Alexander Teytelboym, and Malcolm D McCulloch. "Bilateral contract networks for peer-to-peer energy trading". In: *IEEE Transactions on Smart Grid* 10.2 (2018), pp. 2026–2035.

[Mun57] James Munkres. "Algorithms for the assignment and transportation problems". In: *SIAM Journal of Applied Mathematics* 5.1 (1957), pp. 32–38.

[MVV18] Mehdi Madani and Mathieu Van Vyve. "Revisiting minimum profit conditions in uniform price day-ahead electricity auctions". In: *European Journal of Operational Research* 266.3 (2018), pp. 1072–1085.

[MZ97] Patrice Marcotte and Dao Li Zhu. "Equilibria with infinitely many differentiated classes of customers". In: (1997).

[Mül+17] Fabian L Müller, Jácint Szabó, Olle Sundström, and John Lygeros. "Aggregation and disaggregation of energetic flexibility from distributed energy resources". In: *IEEE Transactions on Smart Grid* (2017).

[NAC14] Ehsan Nekouei, Tansu Alpcan, and Deb Chattopadhyay. "Game-theoretic frameworks for demand response in electricity markets". In: *IEEE Transactions on Smart Grid* 6.2 (2014), pp. 748–758.

[Nas50] John F Nash. "Equilibrium points in n-person games". In: *Proc. Nat. Acad. Sci. USA* 36.1 (1950), pp. 48–49.

[Nis+07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Vol. 1. Cambridge University Press Cambridge, 2007.

[NJJ14] Robert Nishihara, Stefanie Jegelka, and Michael I Jordan. "On the convergence rate of decomposable submodular function minimization". In: *NIPS*. 2014, pp. 640–648.

[NL14] Duong Tung Nguyen and Long Bao Le. "Joint optimization of electric vehicle and home energy scheduling considering user comfort preference". In: *IEEE Transactions on Smart Grid* 5.1 (2014), pp. 188–199.

[NTF11] Koichi Nabetani, Paul Tseng, and Masao Fukushima. "Parametrized variational inequality approaches to generalized Nash equilibrium problems with shared constraints". In: *Computational Optimization and Applications* 48.3 (2011), pp. 423–452.

[Ogg+12] Giorgia Oggioni, Yves Smeers, Elisabetta Allevi, and Siegfried Schaible. "A generalized Nash equilibrium model of market coupling in the European power system". In: *Networks and Spatial Economics* 12.4 (2012), pp. 503–560.

[ORS93] Ariel Orda, Raphael Rom, and Nahum Shimkin. "Competitive routing in multiuser communication networks". In: *EEE/ACM Transactions on Networking* 1.5 (1993), pp. 510–521.

[Pac+18] Dario Paccagnan, Basilio Gentile, Francesca Parise, Maryam Kamgarpour, and John Lygeros. "Nash and Wardrop equilibria in aggregative games with coupling constraints". In: *IEEE Transactions on Automatic Control* 64.4 (2018), pp. 1373–1388.

[PC06] Daniel Pérez Palomar and Mung Chiang. "A tutorial on decomposition methods for network utility maximization". In: *IEEE Journal on Selected Areas in Communications* 24.8 (2006), pp. 1439–1451.

[PD11] Peter Palensky and Dietmar Dietrich. "Demand side management: Demand response, intelligent energy systems, and smart loads". In: *IEEE Transactions on Industrial Informatics* 7.3 (2011), pp. 381–388.

[Pec] *Pecan Street Inc. Dataport*. 2016. URL: https://dataport.pecanstreet.org/data.

[PGL17] Francesca Parise, Basilio Gentile, and John Lygeros. "A distributed algorithm for average aggregative games with coupling constraints". In: *arXiv preprint arXiv:1706.04634* (2017).

[PL14] Qiuyu Peng and Steven H Low. "Distributed algorithm for optimal power flow on a radial network". In: *53rd IEEE Conference on decision and control*. IEEE. 2014, pp. 167–172.

[PL16] Qiuyu Peng and Steven H Low. "Distributed optimal power flow algorithm for radial networks, I: Balanced single phase case". In: *IEEE Transactions on Smart Grid* 9.1 (2016), pp. 111–121.

[PS16] Yael Parag and Benjamin K Sovacool. "Electricity market design for the prosumer era". In: *Nature energy* 1.4 (2016), p. 16032.

[Rab00] Matthew Rabin. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem Econometrica". In: *September* 68 (2000), p. 5.

[Rat92] Kali P Rath. "A direct proof of the existence of pure strategy equilibria in games with a continuum of players". In: *Economic Theory* 2.3 (1992), pp. 427–433.

[Rob+17] Valentin Robu, Meritxell Vinyals, Alex Rogers, and Nicholas R Jennings. "Efficient buyer groups with prediction-of-use electricity tariffs". In: *IEEE Transactions on Smart Grid* 9.5 (2017), pp. 4468–4479.

[Ros65] J Ben Rosen. "Existence and uniqueness of equilibrium points for concave n-person games". In: *Journal of the Econometric Society* (1965), pp. 520–534.

[Ros73a] Robert W. Rosenthal. "A class of games possessing pure-strategy Nash equilibria". In: *International Journal of Game Theory* 2.1 (1973), pp. 65–67.

[Ros73b] Robert W. Rosenthal. "The network equilibrium problem in integers". In: *Networks* 3.1 (1973), pp. 53–59.

[Rou06] Tim Roughgarden. "Potential functions and the inefficiency of equilibria". In: *Proceedings of the International Congress of Mathematicians (ICM)*. Vol. 3. 2006, pp. 1071–1094.

[Rou15] Tim Roughgarden. "Intrinsic Robustness of the Price of Anarchy". In: *J. ACM* 62.5 (Nov. 2015), 32:1–32:42. ISSN: 0004-5411.

[Rou16] Tim Roughgarden. *Twenty Lectures on Algorithmic Game Theory*. Cambridge University Press, 2016.

[RS07] Oran Richman and Nahum Shimkin. "Topological Uniqueness of the Nash Equilibrium for Selfish Routing with Atomic Users". In: *Mathematics of Operations Research* 32.1 (2007), pp. 215–232.

[RS15] Tim Roughgarden and Florian Schoppmann. "Local smoothness and the price of anarchy in splittable congestion games". In: *Journal of Economic theory* 156 (2015), pp. 317–342.

[Rte] *RTE (French TSO) database*. URL: https://opendata.reseaux-energies.fr/.

[RV11] Abdur Rais and Ana Viana. "Operations research in healthcare: a survey". In: *International Transactions in Operational Research* 18.1 (2011), pp. 1–31.

[RW09] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*. Vol. 317. Springer Science & Business Media, 2009.

[Saa+12] Walid Saad, Zhu Han, H Vincent Poor, and Tamer Basar. "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications". In: *IEEE Signal Processing Magazine* 29.5 (2012), pp. 86–105.

[Sag12] Claudia Sagastizábal. "Divide to conquer: decomposition methods for energy optimization". In: *Mathematical programming* 134.1 (2012), pp. 187–222.

[Sam+12] Pedram Samadi, Hamed Mohsenian-Rad, Robert Schober, and Vincent WS Wong. "Advanced demand side management for the future smart grid using mechanism design". In: *IEEE Transactions on Smart Grid* 3.3 (2012), pp. 1170–1180.

[San01] William H. Sandholm. "Potential games with continuous player sets". In: *Journal of Economic Theory* 97.1 (2001), pp. 81–108.

[San11] William H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2011, pp. xiv+589.

[SBP18] Etienne Sorin, Lucien Bobo, and Pierre Pinson. "Consensus-based approach to peer-to-peer electricity markets with product differentiation". In: *IEEE Transactions on Power Systems* 34.2 (2018), pp. 994–1004.

[Sch+13] Fred C Schweppe, Michael C Caramanis, Richard D Tabors, and Roger E Bohn. *Spot pricing of electricity*. Springer Science & Business Media, 2013.

[Sch73] David Schmeidler. "Equilibrium points of nonatomic games". In: *Journal of statistical Physics* 7.4 (1973), pp. 295–300.

[Scu+12] Gesualdo Scutari, Daniel P Palomar, Francisco Facchinei, and Jong-Shi Pang. "Monotone games for cognitive radio systems". In: *Distributed Decision Making and Control*. Springer, 2012, pp. 83–112.

[Sha53] Lloyd S Shapley. "A value for n-person games". In: *Contributions to the Theory of Games* 2.28 (1953), pp. 307–317.

[She+82] Hanif D Sherali, Allen L Soyster, Frederic H Murphy, and Suvrajeet Sen. "Linear programming based analysis of marginal cost pricing in electric utility capacity expansion". In: *European Journal of Operational Research* 11.4 (1982), pp. 349–360.

[Shi+14] Wenbo Shi, Na Li, Xiaorong Xie, Chi-Cheng Chu, and Rajit Gadh. "Optimal residential demand response in distribution networks". In: *IEEE Journal on Selected Areas in Communications* 32.7 (2014), pp. 1441–1450.

[Sia14] Pierluigi Siano. "Demand response and smart grids - A survey". In: *Renewable and Sustainable Energy Reviews* 30 (2014), pp. 461–478.

[Sma] *P2P Energy Trading: Market design and optimization, P2P-SmarTest H2020 project*. URL: http://www.p2psmartest-h2020.eu/.

[SMC06] Kibeom Seong, Mehdi Mohseni, and John M Cioffi. "Optimal resource allocation for OFDMA downlink systems". In: *Information Theory, 2006 IEEE Int. Sym.* IEEE. 2006, pp. 1394–1398.

[Smi84a] Michael J. Smith. "A descent algorithm for solving monotone variational inequalities and monotone complementarity problems". In: *Journal of Optimization Theory and Applications* 44.3 (1984), pp. 485–496.

[Smi84b] Michael J. Smith. "The stability of a dynamic model of traffic assignment – an application of a method of Lyapunov". In: *Transportation Science* 18.3 (1984), pp. 245–252.

[Sou+19] Tiago Sousa, Tiago Soares, Pierre Pinson, Fabio Moret, Thomas Baroche, and Etienne Sorin. "Peer-to-peer and community-based markets: A comprehensive review". In: *Renewable and Sustainable Energy Reviews* 104 (2019), pp. 367–378.

[Sto02] Steven Stoft. "Power system economics". In: *Journal of Energy Literature* 8 (2002), pp. 94–99.

[SW15] Sylvain Sorin and Cheng Wan. "Finite composite games: Equilibria and dynamics". In: *arXiv preprint arXiv:1503.07935* (2015).

[TJ78] Peter D. Taylor and Leo B. Jonker. "Evolutionary stable strategies and game dynamics". In: *Mathematical Biosciences* 40 (1978), pp. 145–156.

[TK18] Tatiana Tatarenko and Maryam Kamgarpour. "Learning Generalized Nash Equilibria in a Class of Convex Games". In: *IEEE Transactions on Automatic Control* (2018).

[TRY16] Kang Miao Tan, Vigna K Ramachandaramurthy, and Jia Ying 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.

[Tus+18] Wayes Tushar, Chau Yuen, Hamed Mohsenian-Rad, Tapan Saha, H Vincent Poor, and Kristin L Wood. "Transforming energy networks via peer to peer energy trading: Potential of game theoretic approaches". In: *arXiv preprint arXiv:1804.00962* (2018).

[UO30] George E. Uhlenbeck and Leonard S. Ornstein. "On the Theory of the Brownian Motion". In: *Phys. Rev.* 36 (5 Sept. 1930), pp. 823–841.

[VN50] John Von Neumann. *Functional operators: Measures and integrals*. Vol. 1. Princeton University Press, 1950.

[VZV15] John S Vardakas, Nizar Zorba, and Christos V Verikoukis. "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms". In: *IEEE Communications Surveys and Tutorials* 17.1 (2015), pp. 152–178.

[Wan+14] Yunpeng Wang, Walid Saad, Zhu Han, H Vincent Poor, and Tamer Başar. "A game-theoretic approach to energy trading in the smart grid". In: *IEEE Transactions on Smart Grid* 5.3 (2014), pp. 1439–1450.

[Wan12a] Cheng Wan. "Coalitions in nonatomic network congestion games". In: *Mathematics of Operations Research* 37.4 (2012), pp. 654–669.

[Wan12b] Cheng Wan. "Contributions to Evolutionary and Congestion Game Theory". PhD thesis. Université Pierre et Marie Curie - Paris VI, 2012.

[War52] John Glen Wardrop. "Some theoretical aspects of road traffic research". In: *Proc. of the Inst. of Civil Eng., Part II, 1*. 1952, pp. 325–378.

[WCP13] Weiran Wang and Miguel A Carreira-Perpinán. "Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application". In: *arXiv preprint arXiv:1309.1541* (2013).

[WH16] Hao Wang and Jianwei Huang. "Cooperative planning of renewable generations for interconnected microgrids". In: *IEEE Transactions on Smart Grid* 7.5 (2016), pp. 2486–2496.

[WL15] Weiran Wang and Canyi Lu. "Projection onto the capped simplex". In: *arXiv preprint arXiv:1503.01002* (2015).

[WMRH12] Chenye Wu, Hamed Mohsenian-Rad, and Jianwei Huang. "Wind power integration via aggregator-consumer coordination: A game theoretic approach". In: *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*. IEEE. 2012, pp. 1–6.

[WP17] Zhanle Wang and Raman Paranjape. "Optimal residential demand response for multiple heterogeneous homes with real-time price prediction in a multiagent framework". In: *IEEE Transactions on Smart Grid* 8.3 (2017), pp. 1173–1184.

[Wu+11] Chenye Wu, Hamed Mohsenian-Rad, Jianwei Huang, and Amy Yuexuan Wang. "Demand side management for wind power integration in microgrid using dynamic potential game theory". In: *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. IEEE. 2011, pp. 1199 –1204.

[XB06] Lin Xiao and Stephen Boyd. "Optimal scaling of a gradient method for distributed resource allocation". In: *Journal of Optimization Theory and Applications* 129.3 (2006), pp. 469–488.

[XJB04] Lin Xiao, Mikael Johansson, and Stephen P Boyd. "Simultaneous routing and resource allocation via dual decomposition". In: *IEEE Transactions on Communications* 52.7 (2004), pp. 1136–1144.

[Yao86] Andrew C. Yao. "How to generate and exchange secrets". In: *27th Annual Symp. Found. of Comp. Sci. (SFCS)*. 1986, pp. 162–167.

[YH15] Mengmeng Yu and Seung Ho Hong. "A real-time demand-response algorithm for smart grids: A stackelberg game approach". In: *IEEE Transactions on Smart Grid* 7.2 (2015), pp. 879–888.

[YH16] Maojiao Ye and Guoqiang Hu. "Game design and analysis for price-based demand response: An aggregate game approach". In: *IEEE transactions on cybernetics* 47.3 (2016), pp. 720–730.

[YM13] Naouar Yaagoubi and Hussein T Mouftah. "A comfort based game theoretic approach for load management in the smart grid". In: *Green Technologies Conference, 2013 IEEE*. IEEE. 2013, pp. 35–41.

[YN17] Hao Yu and Michael J Neely. "A Simple Parallel Algorithm with an O(1/t) Convergence Rate for General Convex Programs". In: *SIAM Journal of Optimization* 27.2 (2017), pp. 759–783.

[YP17] Peng Yi and Lacra Pavel. "A distributed primal-dual algorithm for computation of generalized Nash equilibria with shared affine coupling constraints via operator splitting methods". In: *arXiv preprint arXiv 1703. 05388* (2017).

[YP18] Peng Yi and Lacra Pavel. "Asynchronous distributed algorithm for seeking generalized Nash equilibria". In: *arXiv preprint arXiv: 1801. 02967* (2018).

[Zha+17] Chenghua Zhang, Jianzhong Wu, Chao Long, and Meng Cheng. "Review of existing peer-to-peer energy trading projects". In: *Energy Procedia* 105 (2017), pp. 2563–2568.

[Zho+11] Renxin Zhong, Agachai Sumalee, Terry L. Friesz, and William H.K. Lam. "Dynamic user equilibrium with side constraints for a traffic network: Theoretical development and numerical solution algorithm". In: *Transportation Research Part B: Methodological* 45.7 (2011), pp. 1035–1061.

[ZHS10] Mohammad Zulhasnine, Changcheng Huang, and Anand Srinivasan. "Efficient resource allocation for device-to-device communication underlaying LTE network". In: *WiMob, 2010 IEEE 6th Int. Conference*. IEEE. 2010, pp. 368–375.

[Zhu+12] Quanyan Zhu, Jiangmeng Zhang, Peter W Sauer, Alejandro Dominguez-Garcia, and Tamer Başar. "A game-theoretic framework for control of distributed renewable - based energy resources in smart grids". In: *2012 American Control Conference (ACC)*. IEEE. 2012, pp. 3623–3628.

[Zie+08] Anthony Ziegelmeyer, Frédéric Koessler, Kene Boun My, and Laurent Denant-Boèmont. "Road traffic congestion and public information: an experimental investigation". In: *Journal of Transport Economics and Policy* 42.1 (2008), pp. 43–82.

[ZM93] Daoli Zhu and Patrice Marcotte. "Modified descent methods for solving the monotone variational inequality problem." In: *Operations Research Letters* 14.2 (1993), pp. 111–120.

[ZM94] Daoli Zhu and Patrice Marcotte. "An extended descent framework for variational inequalities". In: *Journal of Optimization Theory and Applications* 80.2 (1994), pp. 349–366.

[ZN97] D. Zhang and A. Nagurney. "Formulation, stability, and computation of traffic network equilibria as projected dynamical systems". In: *Journal of Optimization Theory and Applications* 93.2 (1997), pp. 417–444.

[Zoh+12] Ahmed Zoha, Alexander Gluhak, Muhammad Ali Imran, and Sutharshan Rajasegarar. "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey". In: *Sensors* 12.12 (2012), pp. 16838–16866.

List of publications, proceedings and preprints

[Jac+17a] Paulin Jacquot, Olivier Beaude, Stéphane Gaubert, and Nadia Oudjane. "Demand Response in the Smart Grid: the Impact of Consumers Temporal Preferences". In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. 2017. arXiv: 1711.11304.

[Jac+17b] Paulin Jacquot, Olivier Beaude, Stéphane Gaubert, and Nadia Oudjane. "Demand Side Management in the Smart Grid: an Efficiency and Fairness Tradeoff". In: *IEEE/PES 8thInnovative Smart Grid Technologies Europe (ISGT)*. IEEE. 2017. arXiv: 1711.11129.

[Jac+18a] Paulin Jacquot, Cheng Wan, Olivier Beaude, and Nadia Oudjane. "Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players". In: *arXiv preprint* (2018). arXiv: 1810.01436.

[Jac+18b] Paulin Jacquot, Nadia Oudjane, Olivier Beaude, Pascal Benchimol, and Stéphane Gaubert. "Procédé de gestion décentralisée de consommation électrique non-intrusif". French Patent FR1872553. EDF and Inria. 2018. filed to INPI on 7 Dec. 2018.

[Jac+19a] Paulin Jacquot, Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, and Nadia Oudjane. "A Privacy-preserving Disaggregation Algorithm for Non-intrusive Management of Flexible Energy". In: *IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019. arXiv: 1903.03053.

[Jac+19b] Paulin Jacquot, Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, and Nadia Oudjane. "A Privacy-preserving Method to optimize distributed resource allocation". In: *arXiv preprint* (2019). arXiv: 1908.03080.

[Jac+19c] Paulin Jacquot, Olivier Beaude, Stéphane Gaubert, and Nadia Oudjane. "Analysis and Implementation of an Hourly Billing Mechanism for Demand Response Management". In: *IEEE Transactions on Smart Grid* 10.4 (July 2019), pp. 4265–4278. ISSN: 1949-3053. arXiv: 1712.08622.

[JW18a] Paulin Jacquot and Cheng Wan. "Nonsmooth Aggregative Games with Coupling Constraints and Infinitely Many Classes of Players". In: *arXiv preprint* (2018). arXiv: 1806.06230.

[JW18b] Paulin Jacquot and Cheng Wan. "Routing Game on Parallel Networks: the Convergence of Atomic to Nonatomic". In: *IEEE 57th Conference on Decision and Control (CDC)* (2018). arXiv: 1804.03081.

[JW19] Paulin Jacquot and Cheng Wan. "Nonatomic Aggregative Games with Infinitely Many Types". In: *arXiv preprint* (2019). arXiv: 1906.01986.

[LC+19b] Hélène Le Cadre, Paulin Jacquot, Cheng Wan, and Clémence Alasseur. "Peer-to-Peer Electricity Market Analysis: From Variational to Generalized Nash Equilibrium". In: *European Journal of Operational Research* (2019). arXiv: 1812.02301.

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Méthodes de Théorie des jeux et d'Optimisation appliquées aux Systèmes Électriques Décentralisés

Mots clés : jeux de congestion, optimisation décentralisée, smart grid, gestion de flexibilités

Résumé : Dans le contexte de transition vers un système électrique décentralisé et intelligent, nous abordons le problème de la gestion des flexibilités de consommation électriques. Nous développons différentes méthodes basées sur l'optimisation distribuée et la théorie des jeux. Nous commençons par adopter le point de vue d'un opérateur central en charge de la gestion des flexibilités de plusieurs agents. Nous présentons un algorithme distribué permettant le calcul des profils de consommations des agents optimaux pour l'opérateur. Cet algorithme garantit la confidentialité des agents : les contraintes individuelles, ainsi que le profil individuel de consommation de chaque agent, ne sont jamais révélés à l'opérateur ni aux autres agents. Ensuite, nous adoptons dans un second modèle une vision plus décentralisée et considérons un cadre de théorie des jeux pour la gestion des flexibilités de consommation. Cette approche nous permet en particulier de modéliser les comportements stratégiques des consommateurs. Dans ce cadre, une classe de jeux adéquate est donnée par les jeux de congestion atomiques fractionnables. Nous obtenons plusieurs résultats théoriques concernant les équilibres de Nash dans cette classe de jeux, et nous quantifions l'efficacité de ces équilibres en établissant des bornes supérieures sur le prix de l'anarchie. Nous traitons la question du calcul décentralisé des équilibres de Nash dans ce contexte en étudiant les conditions et les vitesses de convergence des algorithmes de meilleure réponse et de gradient projeté. En pratique un opérateur peut faire face à un très grand nombre de joueurs, et calculer les équilibres d'un jeu de congestion dans ce cas est difficile. Afin de traiter ce problème, nous établissons des résultats sur l'approximation d'un équilibre dans les jeux de congestion et jeux agrégatifs avec un très grand nombre de joueurs et en présence de contraintes couplantes. Ces résultats, obtenus dans le cadre des inégalités variationnelles et sous certaines hypothèses de monotonie, peuvent être utilisés pour calculer un équilibre approché comme solution d'un problème de petite dimension. Toujours dans la perspective de modéliser un très grand nombre d'agents, nous considérons des jeux de congestion nonatomiques avec contraintes couplantes et avec une infinité de joueurs hétérogènes : ce type de jeux apparaît lorsque les caractéristiques d'une population sont décrites par une fonction de distribution paramétrique. Sous certaines hypothèses de monotonie, nous prouvons que les équilibres de Wardrop de ces jeux, définis comme solutions d'une inégalité variationnelle de dimension infinie, peuvent être approchés par des équilibres de Wardrop symétriques de jeux annexes, solutions d'inégalités variationnelles de petite dimension. Enfin, nous considérons un modèle de jeu pour l'étude d'échanges d'électricité pair-à-pair au sein d'une communauté de consommateurs possédant des actifs de production électrique renouvelable. Nous étudions les équilibres généralisés du jeu obtenu, qui caractérisent les échanges possibles d'énergie et les consommations individuelles. Nous comparons ces équilibres avec la solution centralisée minimisant le coût social, et nous évaluons l'efficacité des équilibres via la notion de prix de l'anarchie.

Title: Game theory and Optimization Methods for Decentralized Electric Systems

Keywords: congestion games, decentralized optimization, smart grid, demand response

Abstract: In the context of smart grid and in the transition to decentralized electric systems, we address the problem of the management of distributed electric consumption flexibilities. We develop different methods based on distributed optimization and game theory approaches. We start by adopting the point of view of a centralized operator in charge of the management of flexibilities for several agents. We provide a distributed and privacy-preserving algorithm to compute consumption profiles for agents that are optimal for the operator. In the proposed method, the individual constraints as well as the individual consumption profile of each agent are never revealed to the operator or the other agents. Then, in a second model, we adopt a more decentralized vision and consider a game theoretic framework for the management of consumption flexibilities. This approach enables, in particular, to take into account the strategic behavior of consumers. Individual objectives are determined by dynamic billing mechanisms, which is motivated by the modeling of congestion effects occurring on time periods receiving a high electricity load from consumers. A relevant class of games in this framework is given by atomic splittable congestion games. We obtain several theoretical results on Nash equilibria for this class of games, and we quantify the efficiency of those equilibria by providing bounds on the price of anarchy. We address the question of the decentralized computation of equilibria in this context by studying the conditions and rates of convergence of the best response and projected gradients algorithms. In practice an operator may deal with a very large number of players, and evaluating the equilibria in a congestion game in this case will be difficult. To address this issue, we give approximation results on the equilibria in congestion and aggregative games with a very large number of players, in the presence of coupling constraints. These results, obtained in the framework of variational inequalities and under some monotonicity conditions, can be used to compute an approximate equilibrium, solution of a small dimension problem. In line with the idea of modeling large populations, we consider nonatomic congestion games with coupling constraints, with an infinity of heterogeneous players: these games arise when the characteristics of a population are described by a parametric density function. Under monotonicity hypotheses, we prove that Wardrop equilibria of such games, given as solutions of an infinite dimensional variational inequality, can be approximated by symmetric Wardrop equilibria of auxiliary games, solutions of low dimension variational inequalities. Again, those results can be the basis of tractable methods to compute an approximate Wardrop equilibrium in a nonatomic infinite-type congestion game. Last, we consider a game model for the study of decentralized peer-to-peer energy exchanges between a community of consumers with renewable production sources. We study the generalized equilibria in this game, which characterize the possible energy trades and associated individual consumptions. We compare the equilibria with the centralized solution minimizing the social cost, and evaluate the efficiency of equilibria through the price of anarchy.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France