INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAX084

Thèse de doctorat

ÉCOLE POLYTECHNIQUE

IP PARIS

Jeux de Stackelberg, tarification optimale et application aux marchés de l'électricité

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

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques Appliquées

Thèse présentée et soutenue à Palaiseau, le 24 Octobre 2023, par

Quentin Jacquet

Composition du Jury :

Nizar Touzi Professeur, NYU Tandon School of Engineering	Président
Luce Brotcorne Directrice de recherche, INRIA Lille Nord-Europe	Rapporteuse
Claudia Sagastizábal Professeure, IMECC-Unicamp et CEMEAI	Rapporteuse
Guillaume Carlier Professeur, Université Paris Dauphine	Examinateur
Ivana Ljubić Professeure, ESSEC Business School	Examinatrice
Stéphane Gaubert Directeur de recherche, INRIA et Ecole polytechnique (CMAP)	Directeur de thèse
Wim van Ackooij Chercheur expert, EDF R&D (Département Osiris, groupe R36)	Co-directeur de thèse
Clémence Alasseur Chercheuse experte, EDF R&D (Département Osiris, groupe R33)	Co-directrice de thèse

CMAP Inria EDF

Contents

***

Remerciements

***

Je tiens tout d'abord à remercier vivement Luce Brotcorne et Claudia Sagastizábal d'avoir consacré leur temps et leur grande expertise à l'élaboration des rapports de cette thèse. Je souhaite également remercier sincèrement Guillaume Carlier, Ivana Ljubić et Nizar Touzi pour avoir pris part à mon jury de thèse. C'est pour moi un réel plaisir et une chance de pouvoir soutenir ma thèse en présence d'un tel jury.

Je tiens ensuite à remercier grandement mes encadrants de thèse Wim van Ackooij, Clémence Alasseur et Stéphane Gaubert pour leur indéfectible soutient, tant scientifique qu'humain, durant cette longue et belle aventure qu'est la thèse. Cher Stéphane, tu n'auras eu de cesse de me partager avec passion ton savoir. Après trois ans passés à tes côtés, je ne sais toujours pas si celui-ci admet une limite. Semaine après semaine, tu as pris sur ton temps pour creuser avec moi, au tableau, chacune des pistes que j'envisageai, et apporter à tout instant ta vision experte. Mais au-delà de l'expertise scientifique, j'ai toujours pu compter sur ton soutien moral et ta bienveillance pour m'aider à avancer au mieux dans la bonne direction. Sache que je mesure pleinement la chance de t'avoir eu en tant que directeur de thèse. Clémence, ma maman de thèse, toi par qui tout a commencé. Je te remercie de m'avoir fait confiance dès le début en me proposant ce sujet de thèse. Tu as été une encadrante et une collègue hors pair : d'abord par ton sourire et ta gentillesse, qui ont fait souffler sur ces trois années un vent de fraîcheur et de légèreté, mais aussi par ta connaissance des problématiques métiers : c'est par ta compréhension fine des besoins opérationnels que de nouvelles problématiques ont émergé pour donner lieu à de nouveaux axes de recherches. Tu mérites amplement l'inscription sur ta tasse ! Wim, je te remercie d'avoir accepté de continuer l'aventure avec moi, toi qui été déjà de la partie pour mon stage de fin d'études. Tu as été à l'écoute durant ces trois années, me laissant une grande liberté dans le choix des problématiques et des outils mathématiques, et je savais qu'en cas de question, la porte de ton bureau m'était toujours grande ouverte.

Je tiens aussi à remercier également Roxana Dumitrescu. Ce fût un réel plaisir de travailler et de discuter avec toi, que ce soit en France, à Londres (merci encore pour l'invitation) ou en conférence. J'espère que notre collaboration sur les jeux appliqués aux marchés de l'énergie se poursuivra et sera tout aussi fructueuse qu'elle l'a été pendant ces derniers mois. Je remercie également Riadh Zorgati pour avoir alimenter sans cesse mes réflexions avec de nouvelles problématiques et pour avoir collaborer avec moi sur une partie de mes recherches. Depuis des années, tu valorises le sujet de la tarification au sein d'EDF, et c'est grâce à toi que l'idée de la thèse a émergée.

Je tiens ensuite à remercier chaleureusement toutes les personnes qui m'ont aidé à Osiris pendant ces trois années (et demie). Tout d'abord, merci à Sandie Balaguer de m'avoir fait confiance pour le stage de fin d'étude, ce fût un grand plaisir de travailler sous ton aile pendant six mois. Sache que tu n'es pas étrangère à ma volonté d'avoir voulu rester à EDF pour trois ans de plus. Merci également à Julien Loron pour avoir été à l'écoute durant ces trois années et pour avoir contribué depuis le premier jour à me sentir intégré dans l'équipe. Je remercie enfin Sylvie Malézieux, Christelle Roger, Christelle Boyer et Audrey Cathala, qui m'auront permis de franchir à peu près avec succès les méandres administratifs.

Je remercie également mes collègues de R36 pour tous les moments de convivialité passés à leurs côtés, qui ont su, malgré un démarrage de thèse en période de Covid, garder une super ambiance au sein du groupe. Merci d'abord à mon co-bureau Olivier Juan, qui entre deux réunions trouve toujours un moment pour discuter (de division par 0 entre autres). Merci à Andréa pour les discussions de jardinage, de packages Nyan Cat, et bien-sûr pour avoir exporté le désormais sacro-saint rendez-vous du 12. Merci à Rodolphe et Cécile pour leur bonne humeur, même en cas de défaite à la coinche. Merci à Guilhem, compagnon de route depuis le MPRO et éco-pirate à ses heures perdues. Merci enfin à Nadia, Alex, Pascale, Joseph, Raphaël, Jean-Armand, Gaëtan, Claire, Orane, Sandrine, Jean-Yves et Antoine. Je n'oublie pas les anciens du groupe, en particulier Fatine, Rebecca et les deux fervents défenseurs des territoires oubliés, à savoir les deux massicois Saad et Maxime. Je remercie également les personnes du groupe R33 avec qui j'ai pu échanger tout au long de ces trois ans, au premier rang desquels Cheng, qui restera à mes yeux toujours un peu une R36. Également, je remercie Paulin Jacquot, que je copie par le nom et par le choix de mes encadrements, mais pas par le choix de groupe (désolé Clémence ...). Aussi, je tiens à remercier Claire Del Gatto et Yohann Moreau du département SEQUOIA pour avoir contribué à l'élaboration d'un cas d'étude réaliste et pour l'intérêt porté à mes travaux de recherches. Bien évidemment, je me dois de remercier vivement tous les thésards de R36. Adrien, tu as fini cette épreuve avec brio il y a quelque mois, mais tu as été pendant ces trois ans un véritable compagnon de route, et un entraîneur sportif de grande qualité. Je remercie Alexandre avec qui c'est toujours un réel plaisir de discuter et à qui je laisse la place du plus vieux thésard pour quelques mois. Je remercie enfin l'excellent chef-pâtissier Luca et les non-moins brillants Paul, Bianca, Naomi et Thibaud.

Je souhaite aussi remercier sincèrement les personnes du CMAP et de l'INRIA qui m'ont accompagné pendant ces trois années. Je remercie plus particulièrement les membres de l'équipe Tropical pour leur gentillesse et leur écoute : Marianne Akian, Xavier Allamigeon, Yang Qi, Quentin Canu, Antoine Béreau, Shanqing Liu et Nicolas Vandame. Je n'oublie pas Hanadi Dib sans qui j'aurai été totalement dépassé dans les procédures administratives.

Je remercie également mes professeurs de l'ENSTA Paris et du MPRO, qui m'ont donné goût à la recherche. Je tiens tout particulièrement à remercier Sourour Elloumi, toujours attentive au bien-être et au devenir de ses (anciens) élèves. Aussi, je remercie toutes les personnes que j'ai rencontré durant les conférences. En particulier, merci aux organisateurs et aux participants du séminaire à Dagsthul sur le bi-niveau, fort en interactions scientifiques, ainsi que les participants aux conférences SIAM OP23 et FM23 avec qui j'ai pu discuter et enrichir ma culture scientifique.

Enfin, je tiens à remercier mes proches pour le soutien à tout épreuve depuis trois ans, et bien plus. Je commencerai par mes amis de longue date Sansay, David et Paul qui m'ont témoigné au fil des ans leur amitiés et leur soutien. Je remercie également Lucas Trogno avec qui je prends toujours un grand plaisir à perdre au tennis. Je remercie aussi ma filleule de promo et amie Zoé Verchère, avec qui je partage la passion du badminton. Enfin, je remercie tout particulièrement les "chameaux" : Julien Nespoulous dont l'amitié s'est forgée dans l'adversité des parties de billard endiablées avec Louis Declerck, Antoine Crépin à qui je n'ai toujours pas mis un but suuuur Fifa, mon moulin à parole préféré Jean Jodeau, le photographe Jean-Loup Hannebicq, l'exilé Shubo Cheng, et Arthur Poulain, l'homme qui gagne des jambons à la belote.

Je remercie également mes partenaires de badminton, devenus au fil du temps de vrais amis: de manière non-exhaustive, Mathieu, Maxime, Luce, Clémentine, Antoine, Pierre (x2), Hugo et Marion. De même, je remercie Jean-Maurice et toute l'équipe d'A-Systems pour les soirées passées à jouer au billard.

Je remercie enfin mes parents, mon frère et ma grand-mère qui m'ont toujours accompagné au fil des années, et à qui je dois énormément. Mille merci d'avoir su m'épauler et me guider en me témoignant à chaque instant votre inaltérable soutien. Je remercie aussi chaleureusement ma belle-famille qui m'a accueilli les bras ouverts dès le premier jour. Pour finir, Clotilde, toi qui m'accompagnes depuis maintenant 6 ans, c'est sans doute à toi que je dois le plus, merci pour ton soutien, ta patience, ta compréhension, et même ta contribution à travers tes compétences en dérivées et ton œil expert pour repérer les imperfections dans les slides. Simplement, merci d'être là.

Chapter 1

Introduction

***

1.1 Electricity markets: description and current context

1.1.1 The delicate balancing act of demand satisfaction

The electricity network operates at a predetermined frequency (50Hz in Europe), and only a very small and short deviation is allowed. When the power generation does not meet the demand, the lack of electricity in the grid is compensated by a frequency reduction. This can be the starting point of a cascade failure leading to *black-outs*: some producers are forced to unplug their generators from the network, as they are not able to produce at such a low frequency, so that the imbalance increases and the frequency reduces even more. To prevent this, the network regulators monitor carefully this equilibrium, by both anticipating the demand and reacting swiftly in case of peak consumption (as an example, the last huge *black-out* in Europe dates back to 2003 [Ber04]). The last winter was particularly scrutinized in France, where an important part of the nuclear power plants was unavailable due to maintenance. In particular, the government and the French regulator (RTE) have deployed an "electricity weather" forecast system, called Ecowatt [1], in order to alert the population when a consumption peak is expected in the following days. This system relies on good citizenship and individual responsibility to reduce consumption and avoid imbalances.

From electricity generation to end-user fulfillment, the electricity is exchanged through two markets: the *wholesale market* and the *retail market*, see Figure 1.1 for a schematic illustration. We describe hereafter the modus operandi of the two markets (actors, quantity exchanged, time window).

[1] `https://www.monecowatt.fr/`

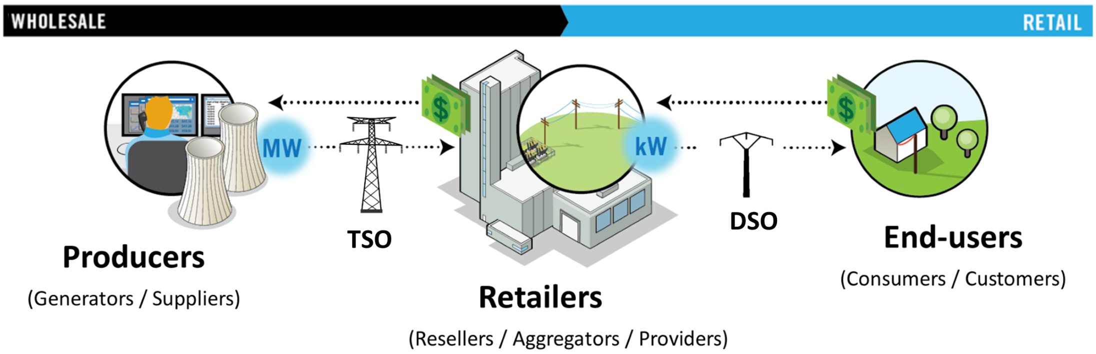


Figure 1.1: Wholesale and retail markets (Source: PJM)

Equivalent terminologies are used in the sequel, and are clarified in the figure.

1.1.2 Wholesale market

Several actors participate in the wholesale market[2]:

- *Producers*: Companies exploiting big power plants (nuclear power plants, wind farms, hydroelectric power dams, etc.) for electricity production purposes. In France, 95% of the production is achieved by EDF and Engie[3]. The distribution of French production on each resource is depicted on Figure 1.2.
- *Retailers*: Agents having a portfolio of end-users and trading electricity by aggregating the consumption of their customers. Like suppliers, they are also responsible for paying the aggregated imbalance of their portfolio. In France, as in Europe, there are several huge retailers (EDF, Engie, TotalEnergies) and some smaller actors (Eni, Enercoop, ekWateur, etc.).
- *Market regulator*: Entity ensuring that the markets correctly function, for the benefit of end consumers and in line with energy policy objectives. In France, this role is assumed by the CRE (Commision de Régulation de l'Energie).

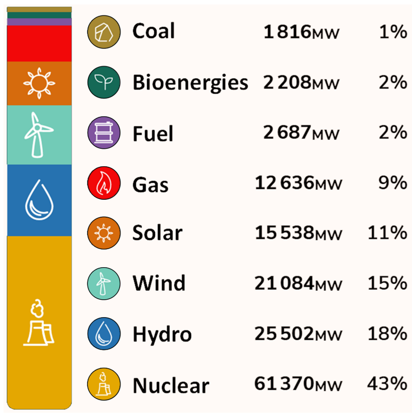


Figure 1.2: French installed power plants capacities (Source: RTE, 12/2022)

- *Market exchange*: Entity which operates the market platform and eases the energy exchanges between actors. In France, this role is assumed by EPEX spot for the day-ahead market, and by EEX for the future market.

The quantities exchanged on this market are huge, and the power circulating in the electric grid are typically in MW. This high-voltage transmission grid is maintained by a *Transmission System Operator* (TSO). The TSO is in charge of balancing the market: it measures the imbalance and take corrective measures if needed. The French TSO is RTE.

The exchange of energy in the wholesale market is divided in several sub-markets, corresponding to several time horizons:

[2]The description is inspired by the following web-pages: `https://www.incite-itn.eu/blog/introduction-to-electricity-markets-its-balancing-mechanism-and-the-role-of-renewable-sources/` and `https://learn.pjm.com/electricity-basics/market-for-electricity.aspx`

[3]`https://selectra.info/energie/guides/comprendre/electricite/production`

(i) *Future market*: Contracts between suppliers and retailers are signed **months** in advance. These contracts constitute a safe option in terms of revenue (fixed price), but it is also the most risky in terms of electricity consumption/production scheduling (high uncertainty).

(ii) *Day-ahead market*: Electricity is sold **one day before** the exchange of electricity occurs in the grid. This is a *spot* market: the supply and demand are aggregated to find the equilibrium: once all the bids are received, the market clearing price (*spot price*) represents the value where supply and demand meet. At the end, all the supply and demand bids that are equal or below the market clearing price are approved. This mechanism favors the "cheapest" (lowest marginal cost) power plants: the cheapest resource will "clear" the market first (participates to the electricity generation), followed by the next cheapest option and so forth until demand is met, see Figure 1.3. The clearing price is then interpreted as the price of the most expensive resource that contributes to the energy generation.

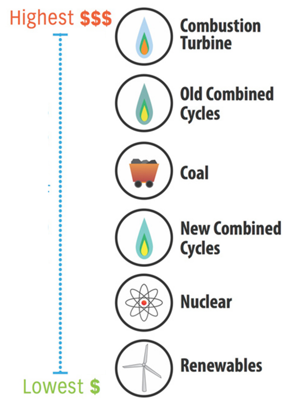


Figure 1.3: Sorted resources (Source: PJM)

(iii) *Intraday market*: Producers and retailers can correct **within the day** their past transaction to adjust more precisely the supply to the demand.

(iv) *Imbalance market*: This market is responsible for the **real time** adjustment in order to keep the frequency of the network as close as possible to the reference frequency.

The current energy crisis[4]

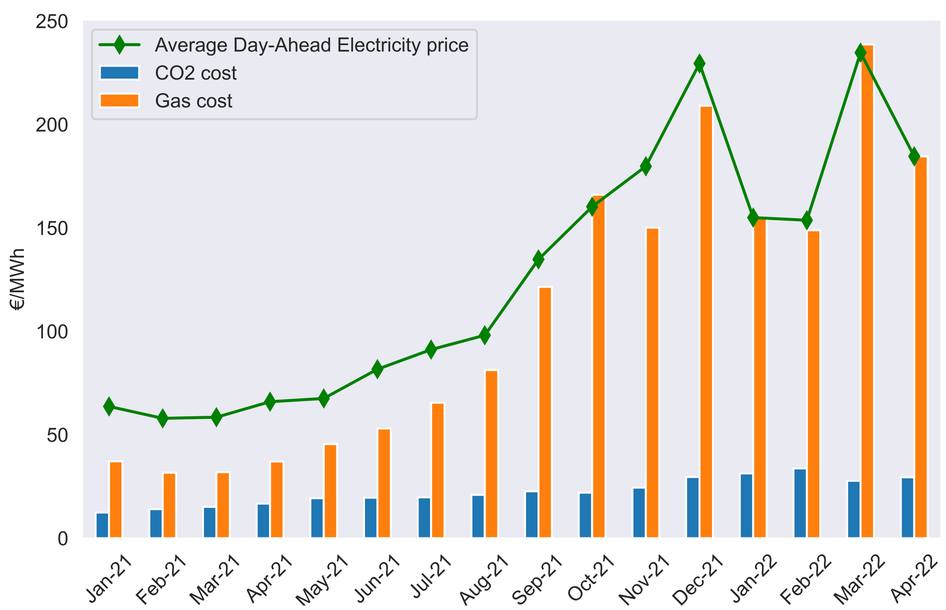


Figure 1.4: Correlation between Gas cost, CO2 cost and electricity prices in Europe (Source: eurelectric.org)

The causes of the electricity crisis which we have been observing since Fall 2021 are various:

[4] `https://www.epexspot.com/en/energycrisis`

(i) The first reason is the huge gas price spike which appends at the end of the covid period. As explained previously, the spot price is determined by the most expensive power resource used in the electricity mix, which often corresponds to a gas turbine. Figure 1.4 shows the strong correlation between the evolution of day-ahead price and gas cost in Europe.

(ii) The second reason is the low generation capacity of many resources: low wind, lower-than-usual gas storage, low hydro-reservoirs levels and low nuclear production since 2022.

(iii) The third reason is the increase of CO_2 prices, driven by shift to -55% emission reduction target for 2030 (Package "Fit for 55", European Council[5]), which further drove up the costs of conventional power plants.

1.1.3 Retail Market

A wide variety of offers...

After electricity is bought in the wholesale market by the resellers, it can be sold to end-users (the population) in the retail market. These contracts are often *affine* in the consumption: they are composed of a fixed price (subscription and installation) plus a variable price (in €/kWh). In France, some contracts has several variable portions, corresponding to several periods (Peak/Off-peak contracts).

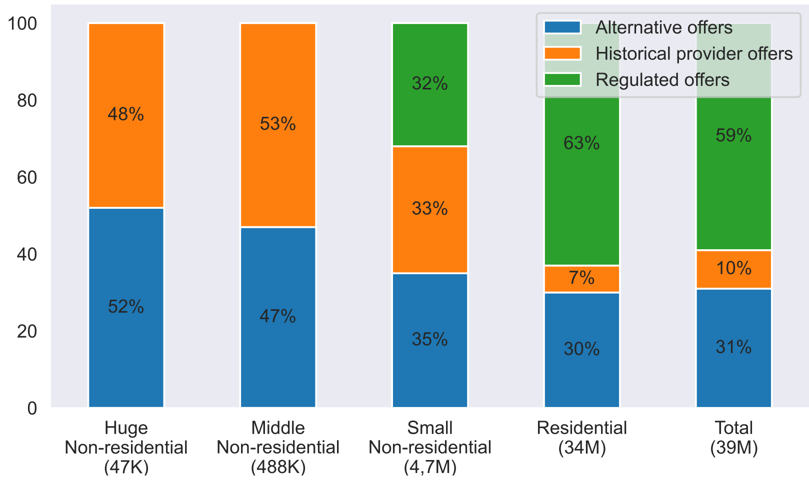


Figure 1.5: Distribution of French consumers in 2022 (Source: CRE)

Many consumers have options for purchasing electricity (this is the case in France for individual consumers since 2007). In fact, they can choose among numerous competitive providers (not only the historical one) to find the contract that best fits their needs. When it comes to regulated offers (for instance, "Tarif Bleu" in France), the providers resell the electricity at prices determined by government regulators. The quantities exchanged in the retail market are of a lower order than in the wholesale market (the maximum power of end-users are in general around few tens of kW). Electricity reaches final consumers through the low-voltage grid, operated by a *Distribution System Operator (DSO)*, responsible for the correct transportation of electricity to end-users. In France, the DSO is Enedis. Figure 1.5 depicts the current distribution of consumers, distinguishing the contracts from the historical provider Électricité de France (EDF) to the new providers.

[5] `https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/`

...and a myriad of customers

A specificity of retail electricity markets is the asymmetry between the number of sellers (electricity providers) and buyers (end-users). In fact, the set of end-users corresponds to the households of/in a state. In this market, each retailer aims at attracting a portion of the population in their portfolio, by designing a competitive menu of offers/contracts.

The customers choice is strongly determined by the electricity invoice: a fully rational end-user will select the cheapest contract for their needs. Classically, given a contract, the invoice of a customer is proportional to their consumption, influenced by several factors:

- *Heating and cooling devices*: a major factor is the type of heating. In France, many households (around 35%[6]) use electricity as heating source.
- *Composition of the household*: the consumption is heavily dependent on the composition of the household: as an example, both the period and the quantity will differ between a retired person and a family with a child.
- *Geography:* In some countries, especially large countries, the weather can substantially vary according to the region. For instance in France, the consumption needs and habits significantly differ from north to south.

To help customers to choose the right offer among the vast jungle of contracts, price comparison engines have been deployed, either from companies or from state agencies (see Figure 1.6 for the French case). These tools basically ask customers to provide the three factors mentioned above, and display the annualized bill for each available offer in the retail market.

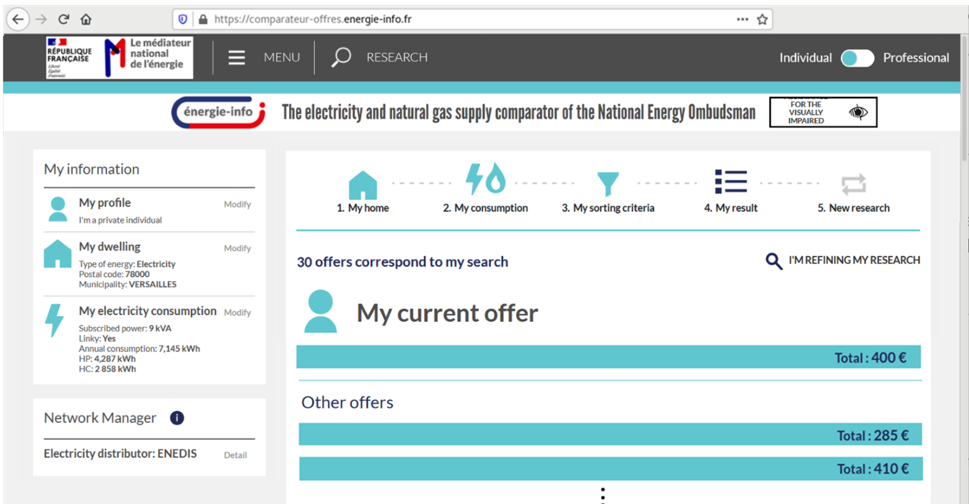
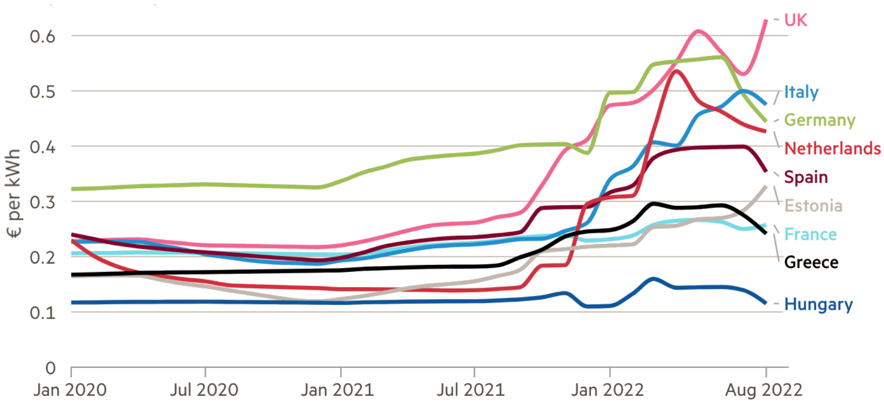


Figure 1.6: Example of price comparison engine (where name of contracts were hidden).

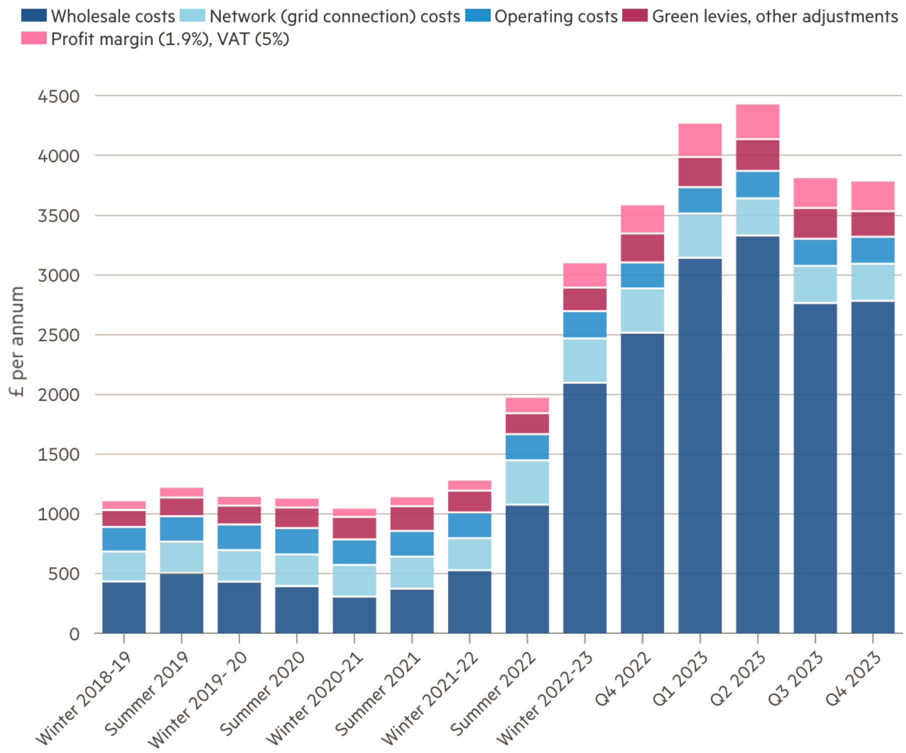
The current energy crisis

The spot prices spike in the wholesale market translates as a cost explosion for resellers in the retail market. Figure 1.7a highlights that every European country was impacted by this crisis. Nonetheless, the intensity of the household electricity price increase was not uniform: in such countries (like in France), governments have initiated tariff shields to cap the prices and

[6] https://www.voltalis.com/comprendre-electricite/les-types-de-chauffage-preferes-des-foyers-francais-1772



(a) Average household price in Europe (Source: energy.eu).



(b) Typical household invoices in UK (consumption of 2900 kWh for electricity and 12000kWh for gas).

Figure 1.7: Consequence of energy crisis on the retail market.

prevent from a too huge increase. The most striking hike appended in the U.K. where the annual electricity invoice has been multiplied by a factor 3 between 2020 and 2023, see Figure 1.7b.

In addition to soaring prices and the resulting consequences for end-users, many suppliers have been forced to file for bankruptcy. In France, three alternative suppliers (E.Leclerc, Planète Oui, Cdiscount) have already stopped their activities and in the U.K., the seventh largest electricity supplier (Bulb Energy, 5% market share, 1.7 million consumers) also had to cease operations. In these cases, customers are most of the time automatically redirected to the historical provider.

1.2 Arising challenges

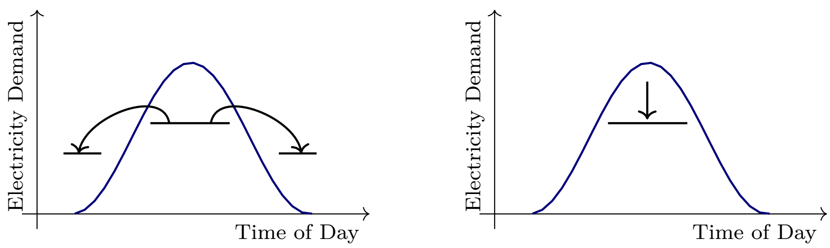
Complex behavior of users

Several (non-monetary) components can influence the decision of end users. Among them, a growing part of the consumers carefully look at the origin of the electricity, i.e., which resource was used for the production of the electricity. As an example, for a decade, there has been a rapid emergence of "green" offers, guaranteeing that electricity is obtained from renewable resources [HR14; MZ16; AF19].

In addition, three behavioral characteristics are often studied as they strongly impact the customers choice:

(i) *Partial/Bounded rationality*: due to possible paucity of information about their consumption characteristics, or unawareness about price updates, customers do not always select the contract that an omniscient agent would have picked. This bounded rationality behavior, i.e., an uncertainty of some kind in the decision is now incorporated into a growing number of models [ES17; RH18; RDP19], often implying customers choices of probabilistic nature (e.g. logit choice models [Tra09]), more difficult to handle into optimization problems [BLS23].

(ii) *Price elasticity*: the energy crisis has highlighted a non-negligible flexibility in the end-users consumption, induced by high price fluctuation (here soaring prices). They have restricted their consumption to focus only on essential usages. Many studies tackle this phenomenon by estimating its intensity [And+97; Lij07; ACR20; NYK20]. This important adaptability is the foundation of Demand Response [AE08], aiming at modifying the consumers load curves by sending to the latter price incentives in order to lift down the peak consumption ("peak clipping" and "load shifting"), see Figure 1.8.



(a) Load shifting (b) Peak clipping

Figure 1.8: Two types of Demand Response mechanisms

(iii) *Inertia in the response*: as shown in Figure 1.6, price comparison engines often compare the current offer of the end-user with the alternatives available in the market. There is indeed an asymmetry between the already subscribed contract with the rest of the market since customers are facing *switching costs* which favors an inertia in the choice dynamics: end-users will only change their offer for another one if there exists a substantial difference between the two contracts which compensates switching costs. These costs have various origins: as illustrated in Figure 1.9, it can be monetary costs (exit fees, equipment costs), but can be psychological (brand loyalty, time and effort spent to make the switch of offer). We refer to [Stu02; Sal22; DW23] for estimations of switching costs in electricity retail markets. Empirical studies analyzes the impact of these switching costs on both the customers decision and on the pricing strategies [FH08; HP10; Abd17; Mys+18], and more recently some works integrate the time dependence (inertia due to switching costs)

Figure 1.9: Example of switching costs (Source: sketchbubble.com)

into optimization models [PE17; Hua22], recovering in particular the relationship between brand loyalty and promotions.

Readability of the market

In recent years, offers have abounded: both the number of resellers and the number of contracts per provider have increased (for instance, from 18 in 2008 to more than 100 retailers nowadays in France[7]). As a consequence, consumers suffer from a reduced readability of the market and are sometimes confused in this vast jungle of offers. In the meantime, having a larger *menu* of offers enables a better fit to customers preferences and habits, and so a potential reinforcement of the attractiveness. A natural question then arises for each provider: what is the optimal trade-off between these two opposite trends ? This question is somehow less regarded in the literature though it is an important concern for resellers concerns.

Incentive mechanisms for consumption reduction purposes

In a context of sobriety, due to both energy crisis (see previous sections) and ecological considerations, retailers aim at incentivizing their customers for consumption reduction, not only in peak period (as for demand response) but in average. These incentives can take several forms. First, they can be based on customers civism and consciousness: some applications show to each end-user how far their consumption is from the average consumption of similar customers (this is the case in France for most of the providers, such as Engie and EDF). Another approach is based on monetary reward, most of the time offering discounts if the customer was able to lower their consumption compared to past years ("SimplyEnergy"[8], "Plüm énergie"[9] or "OhmConnect"[10]). Economical studies looking at the impact of these monetary incentives on customers can be found in the literature (e.g., [MT12; Pra+18; Rus+23]).

[7] Source: Commission de Régulation de l'Energie (CRE)
[8] `https://www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward`
[9] `https://plum.fr/cagnotte/`
[10] `https://www.ohmconnect.com/`

1.3 Pricing models as Stackelberg games in the literature

Game theory is a mathematical domain which analyzes the interaction between rational agents (called *players*). The first contribution may be the duopolist problem formulated by Cournot in 1838 [Cou38], where the notion of equilibrium is tackled, i.e., a situation where each player does not want to deviate from their current situation. Mathematical games are then theorized by von Neumann and Morgenstern in 1944 [VM47]. In the fifties, Nash [Nas53] formalized the notion of equilibrium, very well-known and used today. In this thesis, we focus on a specific class of games: Stackelberg games [Sta52]. These 2-players problems reflect the asymmetry of the players' roles: the leader first decides, then the follower replies (sequential games). This class of games is particularly well-adapted to pricing problems in which a price-setting entity aims at optimizing its revenue. The customers decision, then, always appends after the establishment of the new pricing strategy.

We focus here on two frameworks dealing with Stackelberg games, and with numerous application to pricing management: bilevel optimization, and principal-agent models. For a gentle introduction on these two areas, we refer to preliminary chapter 3. We present hereafter an overview of the existing contributions in the pricing literature.

Bilevel pricing

Bilevel problems constitute an important part of the mathematical literature of pricing problems. They consist in problems with two players, where a first player (leader) proposes prices to customers (followers) who maximize their own utility functions and take into account the prices settled by the leader. We refer to Labbé and Violin [LV15] for a detailed presentation of the concepts and solution models on bilevel pricing, as to the recent survey of Dempe [Dem20] for an extensive review of the bilevel literature, with a particular focus on pricing problems.

Early works in network pricing. Labbé et al. [LMS98] first proposed a taxation model applied to highway pricing, later extended by Brotcorne et al. [Bro+06]. Larsson and Patriksson [LP98] studied the question of the traffic congestion through bilevel programming. In freight management, Brotcorne et al. [Bro+00] proposed a model and algorithms to determine optimal tariffs. These models are applied to networks, therefore the followers problem inherits the graph structure induced by the network topology. For instance, given the leader decision, the follower problem is a shortest path problem. Dewez et al. [Dew+08] took advantage of the structure to propose new valid inequalities to tighten the relaxations, as with path-based formulations.

Envy-free product pricing. The envy-free product pricing problem (sketched in Chapter 3) was first analyzed by Guruswami et al. [Gur+05] from a complexity angle, showing the APX-Hardness of the problem. Then, several reformulations of the bilevel problem have been proposed by Shioda, Tunçel and Myklebust [STM11] and Fernandes et al. [Fer+13]. Heilporn [Hei+10] and Fernandes et al. [Fer+16] studied links between the latter product pricing problem and network pricing models, showing in particular that valid inequalities from one model can be used in the other one (such as triangle inequality).

Applications on electricity pricing. For a decade, many pricing problems arose in energy management. With the liberalization of retail electricity markets, Leyffer and Mundson [LM10] tackled the issue of the optimal tariff-setting faced by a retailer in a competitive environment, leading to an Equilibrium Problem with Equilibrium constraints (EPEC). Luna et al. [Lun+20] then proposed a primal-dual regularization technique to overcome the ill-posedness of the problem arising from the non-uniqueness of lower decision. Besides, many contributions have been done in Demand Response: Zugno et al. [Zug+13] might be pioneers on this topics. Afşar et al. [Afş+16] looked at the optimal tariff in order to incentive consumers to clip the peak consumption. Also, Kovacs [Kov16] proposed an alternative time-variant tariff for demand response. Alekseeva et al. [Ale+19] studied peak shifting strategies for cost reduction purposes. Aussel et al. [Aus+20] enriched the model by adding one level of optimization (suppliers, aggregators, end-users) and developed alternative tie-breaking rules to optimistic and pessimistic formulation. Abate, Riccardo and Ruiz [ARR21] introduced uncertainty into the consumers utility functions to take into account elasticity in the demand. Grimm et al. [Gri+21] compared time-of-use (TOU), critical-peak-pricing (CPP) and real-time-pricing tariff (RTP) in a numerical study made on realistic use-cases. Kozanidis and Kostarelou [KK23] proposed a bilevel model to deal with the optimal price-bidding of energy producers in day-ahead electricity markets.

Works	Actors	(i) Time horizon (ii) Stochasticity	Lower level nature	(i) Approach (ii) Type of solution
[LM10]	Multi-L. Common-F.	(i) Static (ii) Deterministic	Continuous	(i) Nonlinear models (NCP formulations) (ii) Stationary
[Zug+13]	Single-L. Multi-F.	(i) Discrete time (ii) Discrete scenarios	Linear	(i) KKT reformulation (ii) Global
[Afş+16]	Single-L. Multi-F.	(i) Discrete time (ii) Deterministic	Linear	(i) KKT reformulation (ii) Global
[Kov16]	Single-L. Multi-F.	(i) Discrete time (ii) Deterministic	Linear	(i) Heuristic
[Ale+19]	Single-L. Multi-F.	(i) Discrete time (ii) Deterministic	Linear	(i) KKT reformulation (ii) Global
[Lun+20]	Multi-L. Common-F.	(i) Static (ii) Deterministic	Continuous	(i) Nonlinear models (Dual regularization) (ii) Stationary
[Aus+20]	Single-L. Multi-F.	(i) Discrete time (ii) Deterministic	Linear	(i) KKT reformulation (ii) Local
[ARR21]	Single-L. Multi-F.	(i) Discrete time (ii) Discrete scenarios	Linear	(i) KKT reformulation (ii) Local
[Gri+21]	Single-L. Single-F.	(i) Discrete time (ii) Deterministic	Linear	(i) S.-D. reformulation (ii) Global
[KK23]	Single-L. Single-F.	(i) Discrete time (ii) Deterministic	Integer	(i) Opt.-V. reformulation (ii) Global

Table 1.1: Non-exhaustive classification of bilevel problems applied to electricity pricing. "S.-D." stands for "Strong-Duality" and "Opt.-V." stands for "Optimal-Value", "L." for "Leader" and "F." for "Follower".

Tropical angle. An emerging direction is the viewpoint adopted by Baldwin and Klemperer [BK19] and Tran and Yu [TY19], where the customer decision is analyzed using tropical geometry. This geometry uses the max-plus algebra, see e.g. [MS21]. The utility maximization problem of the follower, then, consists in a selection of kind "one out of N possibilities", and can be viewed as the evaluation of a tropical polynomial. The agents response map is in this context a polyhedral complex where each cell corresponds to a subset of leader's decisions (prices) that induce a common customers decision. When the problem involves several independent followers, the "overlay" of polyhedral complexes give a full visual understanding of the pricing problem. We refer to [Eyt18] for a tropical approach of bilevel pricing applied to mobile telecommunication networks.

Principal-Agent models for energy management

Principal-Agent models belongs to the theory of incentives, where the first approach dates back to the works of Barnard [Bar38]. This theory was first designed in order to improve the productivity of workers, in the same spirit as the Taylorism. As in bilevel pricing, the Principal (a.k.a. the leader) designs a contract taking into account the rational behavior (so-called *moral hazard* in this field) of the agent (a.k.a. the follower), who aims at maximizing its own utility.

In the last decade, contributions about Principal-Agent models applied to energy management have been numerous. Carmona, Fehr and Hinz [CFH09] introduced an optimal stochastic control problem for carbon pricing. Féron, Tankov and Tinsi [FTT20] studied the trading process in intra-day electricity markets using Principal-Agent models. Alasseur, Farhat and Saguan [AFS20] analyzed through risk sharing mechanism new levers to favor capacity investments. About Demand response, Aïd, Possamï and Touzi [APT22] exhibited optimal incentives for reducing both the average electricity consumption and its variance. In [Éli+20], Élie et al. proved that a closed-form expression of optimal contracts can be found in the case of linear energy valuation and when the contracts are indexed by aggregated consumption. The question of electricity pricing in the retail market has been addressed by Alasseur et al. in [Ala+20] where elasticity of the demand is described through a Constant Relative Risk Aversion measure (CRRA). Recent application emerged, as renewable markets for example. Aïd, Kemper and Touzi [AKT23] introduced a Principal-Agent framework in order to model investments in renewable energy. Also, Shrivats, Firoozi and Jaimungal [SFJ21] proposed a principal-agent mean-field game applied to the market of Renewable Energy Certificates.

1.4 Objectives and contents of the thesis

In this PhD thesis, we analyze pricing problems which arise in energy management, and more precisely at the interface between retailers and consumers. First, we focus on the pricing of offers for the highly competitive retail market, where the crux of the matter is to fairly construct a menu of offers, each of them selectable by any customer, but designed to be attractive for a targeted portion of the population, the most profitable one for the electricity provider. Second, we study incentive mechanisms which aim at reducing the electricity consumption of a portfolio of end-users by constructing optimal rewards that favor the most sobriety-compliant customers. Moreover, in most of electricity pricing problems, a critical data that needs to be estimated is the typical (based on past data) consumption and invoice of customers. We look here at refined estimations to this purpose. This PhD dissertation introduces modeling, theoretical and algorithmic contributions to answer these issues. We now give more details on the content of this dissertation:

- ◇ We first establish a bilevel framework to tackle the question of the optimal pricing of a menu of offers, a problem faced by a company whose aim is to stay competitive by designing its tariff menu in reaction to a given market context. This work extends the models developed in standard approaches for product pricing. This framework is of a combinatorial nature, as customer decision consists in a discrete choice among the range of offers. In Chapter 4, we introduce a new customer behavior – seen as a regularization of purely deterministic choice models – where the decision of each clusterized end-user is now expressed as a discrete probability measure on the alternatives, concentrated on the most advisable contracts for the latter. This new model is motivated by the huge number of consumers, viewing each cluster of end-users (consumers with similar characteristics) as a representative end-user of an infinite-size homogeneous subpopulation. In Chapter 5, we enrich the customers behavior by inserting into the lower problem switching costs, acting as an elastic restoring force that keeps the customers attached to their current offer. This inertial effect dives the problem into a dynamical control framework: the model now corresponds to the control of a Markovian decision process where the follower transition probabilities (from a contract to another) are determined by the leader decision at each time step. We study in Chapter 6 another point of interest for operational services: the influence of the size of the menu, i.e., the number of contracts to design. This question is tackled by a two step-strategy: we first look at the mean-field limit case – in which a continuum of offers is designed – and then we search for the best finite-size approximated menu that maximizes the leader objective. This quantization problem enables us to determine a loss function which maps the limit number of contracts with the relative revenue loss induced by this constraint.

◇ We then gather contributions of different nature. In Chapter 7, we first consider the question of the energy sobriety by focusing on a Principal-Agent framework, in which a company aims at proposing a new type of contract which encourages the customers to make some efforts in view of reducing their consumption. This contract provides a variable reward in addition to the standard linear pricing (invoice proportional to the consumption) which is based on the rank of the end-user within the set of similar users. This ranking game scheme introduces a competition (mean-field equilibrium) between consumers, driving them to lower their mean power consumption. In Chapter 8, motivated by the wide number of applications where invoice estimations are required, we focus on concentration inequalities and their embedding into optimization models. In particular, we look at Bennett-type inequalities and show that a convex reformulation can be obtained, leading to tractable chance-constrained programs, as the one coming from the electricity invoice estimations context. Finally, we study in Chapter 9 a family of nonconvex approximations of optimization problems under sparsity requirements. These sparse optimization problems naturally appear when it comes to the design of a menu of offers, which is indeed the core of Chapter 6. We tackle the issue under a more general angle, considering a general problem under cardinality constraints, and show that tractable approximated constraints, based on Rényi entropies, lead to an acceptable solution, i.e., with a limited violation of the original cardinality requirement.

1.5 Contributions of the chapters

We now give more details on the contributions of each chapter of this PhD thesis:

1. Chapter 4 states the problem faced by an electricity provider who aims at designing a menu of offers intended for a wide number of end-users. A complete description of the model would lead to a Multi-Leader-Common-Follower Game, see [LM10], where a Nash equilibrium should be found between the providers. Here, we study a static version where a provider optimizes the offers given a strategy (prices) of the other actors of the market (competitive providers and regulated offers), leading to a Single-Leader-Single-Follower problem, see Figure 1.10. This can be interpreted as an instantaneous reaction/adjustment of the studied provider to a current market situation. We first introduce deterministic models based on the bilevel (Envy-free unit-demand, see e.g., [Gur+05]) product pricing framework, and derive MILP formulations coming from single-level reformulations. We also recall Logit models [GMS15], viewed here as probabilistic regularization of customer behavior, and analyze their asymptotic behavior. The main concern is about a both fair and tractable model, and we develop to this purpose a quadratic regularization of the lower level. We exhibit a polyhedral structure of the customers response (Theorem 4.3.1), and show that when the regularization parameter – interpreted as the rationality intensity

of end-users – tends to infinity (fully rational), deterministic response is recovered, i.e., the polyhedral complex converges to the cell description developed by Baldwin and Klemperer [BK19] in deterministic settings (Theorem 4.3.2). An extensive numerical study is achieved on a realistic instance, showing the benefit of this new model (Figure 4.9) and the absolute necessity of smooth customers behavior to avoid too optimistic prices (non profitable in practice), see Figure 4.10.

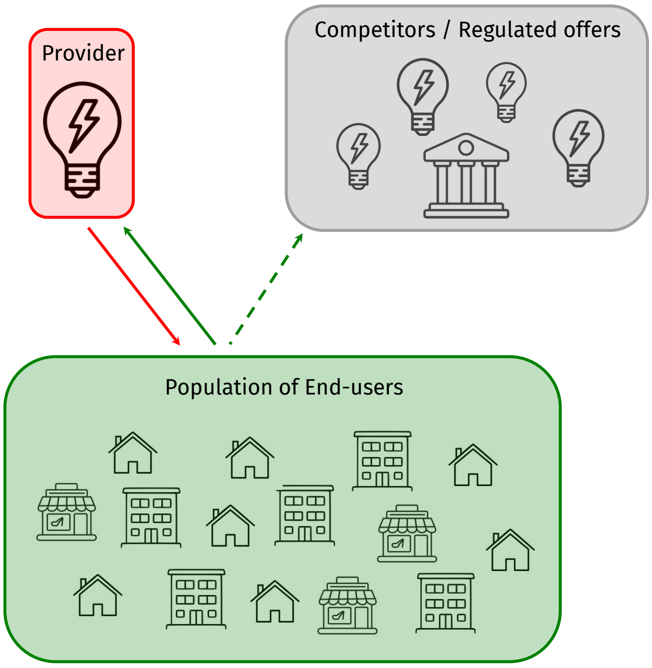


Figure 1.10: Bilevel problem faced by a provider willing to optimize its menu of offers in a competitive market

We look at the provider reaction to given prices of the (static) competition. The provider first updates the prices (red arrow), then each end-user in the heterogeneous population chooses an offer – either from the provider (solid green arrow) or from the competitors (dashed green arrow).

2. In Chapter 5, we incorporate into the previous static bilevel model (Chapter 4) the influence of switching costs. The resulting inertia in the customers behavior translates as a past dependence of the contract choice. We model the problem as a Markov Decision Process where each transition problem between two discrete time steps corresponds to an instance of the static bilevel model, see Figure 1.11. We formalize the problem as an optimal control problem aiming at maximizing the average long-term reward (ergodic control, see e.g. [ABG11]), and show that this can be equivalently solved by an ergodic eigenvalue problem (Proposition 5.2.1). We prove using a contraction argument on the dynamics that the eigenvalue problem has a regular solution (Theorem 5.2.2), from which the optimal leader strategy can be obtained. We introduce a Policy Iteration algorithm [Put94; Gau96] adapted to the control of decomposable population, recomputing on the fly the transitions between states to drastically reduce the needed storage space, while keeping fast computational cost in comparison with Relative-Value-Iteration algorithms (RVI), see Table 5.2. We then analyze the impact of the switching cost intensity on the pricing policy. We observe that above a threshold, periodic promotions are applied, providing a strictly higher reward to the leader than constant-price strategies (Figure 5.5). We finally showcase the extent of this behavior in an example, proving the superiority of cycling policies over static

ones, see Section 5.7.

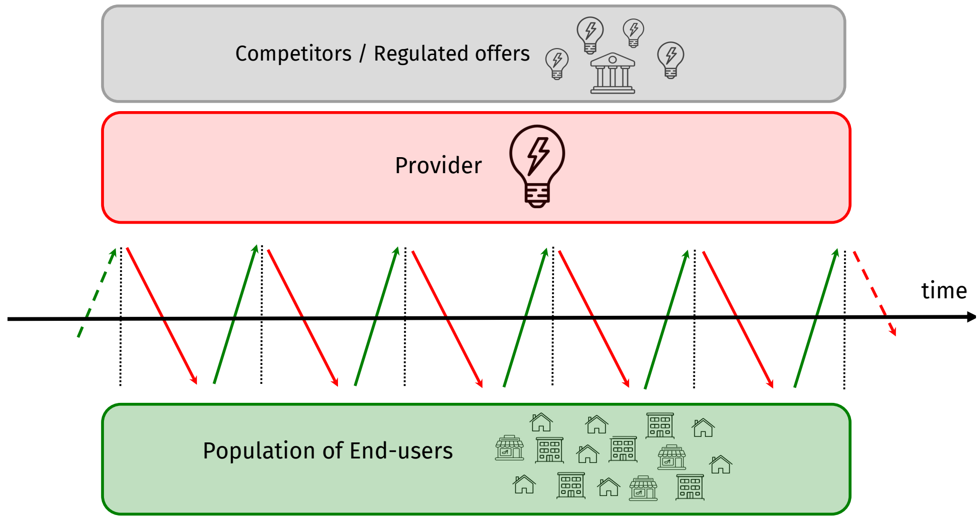


Figure 1.11: Iterated bilevel pricing problem as a Markov Decision Process.

At each time step, the provider is able to update the prices, and each end-user of the heterogeneous population can either keep the current contract, or change for a new one (from the same provider or from the static competitors). In the latter case, she will suffer switching costs.

3. Chapter 6 explores another important factor in the design of tariff menus, that is the question of the optimal number of offers (size of the menu) that a retailer should propose to the population. To this aim, we first relax the condition on the size of the menu to obtain as a limit case a problem where an affine contract can be especially designed for each end-user (infinite size of menu). We show that this relaxed problem can be reformulated as a generalized monopolist problem à la Rochet-Choné [RC98] (Theorem 6.5.1), which is convex when the customers portfolio is fixed (e.g. when full-participation constraint is imposed, see (6.3)). Given the optimal menu (potentially of infinite size) of the relaxed problem, we approximate the latter by a menu of finite prescribed number of contracts using a pruning procedure (Algorithm 6), initially designed for control problems. This pruning scheme iteratively reduces the size of the menu by removing the most redundant remaining contract, i.e., the one which induces the least approximation error when it is removed. We adapt this new reverse greedy approach for several criteria (e.g. L_∞ or L_1 norm), and show that a partial recomputation of the solution at each iteration is always sufficient (Propositions 6.3.1 and 6.3.2). We numerically observe that this local update leads to huge computation time gains, see Figure 6.3. We then exploit the structure of the problem to interpret the decision of the customers as a Bregmann Voronoi diagram [BNN10] – see Proposition 6.4.1 – and derive from quantization theory [Pag15] a Lloyd's procedure that aims at finding the best L_1 approximation of the infinite-size menu by a finite one, showing the efficiency and robustness of the new pruning methods.

4. Chapter 7 introduces a novel incentives scheme, applied to energy sobriety concerns. In this scheme, the retailer (Principal) – motivated by regulation agencies/governments – contracts with a continuum of customers by proposing to them a monetary reward based on the individual ranking of each customer within the population, see Figure 1.12. The consumers are then encouraged by this financial compensation to be ranked within the best energy savers in order to receive the highest reward. Here, the competition between

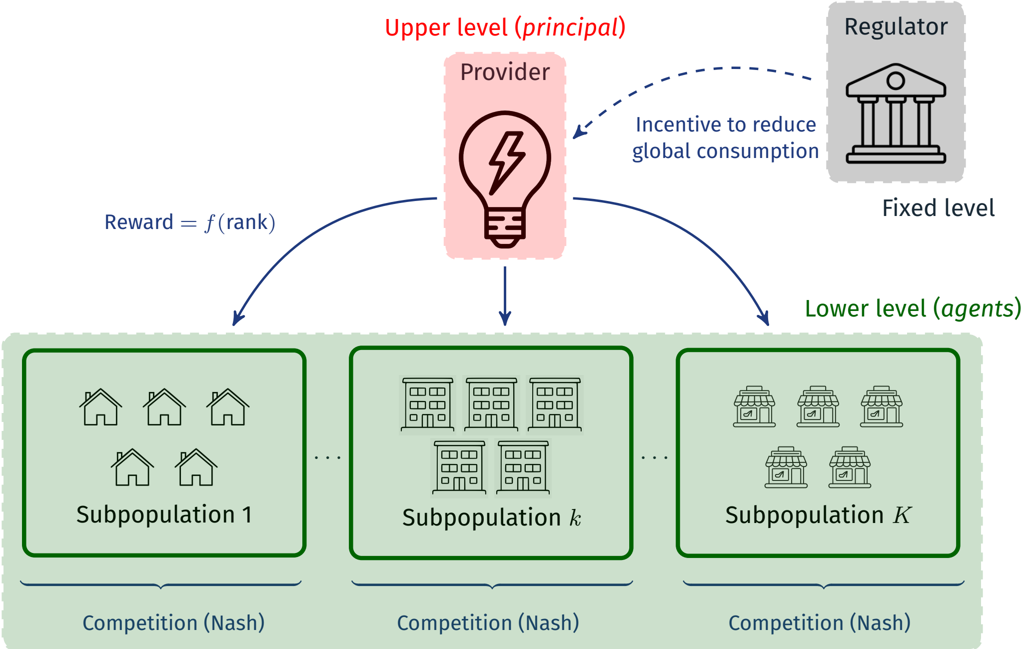


Figure 1.12: Principal-Agent relationship between an energy supplier and a field of agents, the latter competing with each other to obtain the best reward possible by reducing as possible their energy consumption.

Motivated by governing energy reduction policies, the provider designs an incentive reward function based on the ranking of the agents so that the relative best energy saver will receive the higher reward. Given the incentive function, a Nash equilibrium has then to be found for each subpopulation of homogeneous end-users.

end-users, triggered by a given reward function, is modeled as a mean-field game, and we prove that it has a unique Nash equilibrium, which can be analytically determined (Theorem 7.2.3). This characterization holds in particular for purely rank-based rewards, widely studied in [BZ21]. Then, incorporating this closed-form formula of the equilibrium in the retailer problem, we show that when the price-elasticity of the agents is uniform, the optimal reward can be obtained through the resolution of a fixed-point equation on the targeted optimal mean consumption. To this purpose, we exploit the optimality conditions of an equivalent convex reformulation of the problem in the distribution space, see Theorem 7.2.4. The latter characterization enables a fine understanding of the reward function, establishing sufficient conditions for the development of such incentive schemes. When a reward must be offered to a decomposable heterogeneous population with non-uniform price elasticity, a numerical algorithm (Algorithm 10) is proposed, and simulations on realistic examples shows that the approach is likely to produce important consumption reductions while ensuring a mean end-users satisfaction greater than the traditional (non-incentivized) case, see Figure 7.5.

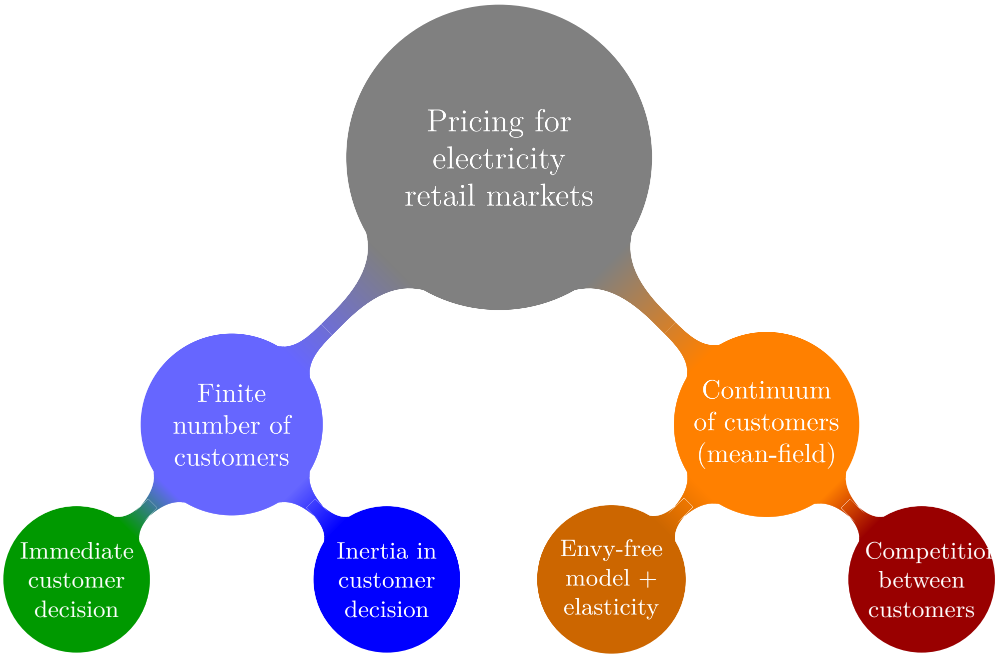
5. In Chapter 8, we study Bennett-type concentration inequalities in view of integrating them into chance-constrained programs [Pré95]. First, we introduce a double bisection algorithm enabling to compute confidence bounds (Algorithm 12). This can be applied for estimating invoices when consumption load curves are only known through their moments. Then, we study tight conservative approximations of chance-constrained programs with information on means and variances (Proposition 8.4.1 and Proposition 8.4.2). We

show that these approximations can be embedded into convex optimization programs as in [NS07], and are tractable in many use cases. We first focus on a canonical discrete problem, that is the integer knapsack problem, and show that a cutting-planes description of the convex Bennett-type estimation (added iteratively in the Branch-and-Bound) leads to more profitable robust solutions while keeping a reasonable computational time (Table 8.2). We also use these estimators in continuous optimization problems, by focusing on the robust support vector machines problem, and show on instances of the literature that the approach reduces miss-classification error (Section 8.4.2).

6. In Chapter 9, we study generic sparse optimization problems. Such problems naturally arise in pricing management. As an example, for the sake of readability, retailers often constrain their menu of offers to be of limited prescribed size, see Chapter 6. Here, we introduce a family of cardinality's lower bounds, involving Rényi's entropy [Rén+61]. Focusing on application to selection problems – where the optimization is performed on discrete probability measure space (simplex), we proved (Theorem 9.3.1) that the entropic bounds we developed can control the sparsity of the solution, recovering as limit case the exact ℓ_0-norm. In numerical results, we study the specific case of the Shannon entropy and its ability to impose sparsity in the solution. We show that the use of this entropic bound for portfolio selection problems ensures a good compromise between the control of the cardinality and optimization performance (Figure 9.6).

Schematic classification of the chapters

In Figure 1.13, we briefly outline the mathematical focuses of each chapter by keywords, and classify the chapters according to the main modeling differences. The concepts mentioned in the keywords are then introduced in the preliminary chapter (Chapter 3).



Chapter 4:
- Bilinear bilevel optimization
- Regularized lower levels
- Polyhedral complex analysis

Chapter 5:
- Ergodic control of populations
- Deterministic MDPs
- Convergence analysis (turnpike-like behaviors)

Chapter 6:
- Principal-Agent problem with adverse selection
- Quantization of menus
- Bregman Voronoï diagrams

Chapter 7:
- Second-best Principal-Agent problem
- Mean-field games
- Rank-based incentives

Figure 1.13: Mind map expressing the differences between the chapters in terms of modeling, as with keywords outlining the main tools used in each chapter.

Chapter 2

Introduction (en français)

***

2.1 Description des marchés de l'électricité et contexte actuel

2.1.1 L'équilibre du réseau: une tâche ardue

Le réseau électrique fonctionne à une fréquence prédéterminée (50 Hz en Europe), et seule une déviation très faible et temporaire est autorisée. Lorsque la production d'électricité ne répond pas à la demande, le manque d'électricité dans le réseau est compensé par une réduction de la fréquence. Cela peut être le point de départ d'une panne en cascade conduisant à des *black-out* : certains producteurs sont obligés de déconnecter du réseau certaines installations, car ils ne sont pas en mesure de produire à une fréquence aussi basse, ce qui augmente le déséquilibre et diminue encore la fréquence. Pour éviter cela, les régulateurs de réseau surveillent attentivement cet équilibre, à la fois en anticipant la demande et en réagissant rapidement en cas de pic de consommation (à titre d'exemple, le dernier grand black-out en Europe remonte à 2003 [Ber04]). L'hiver dernier a été particulièrement scruté en France, où une partie importante du parc nucléaire était indisponible pour cause de maintenance. Le gouvernement et le régulateur français (RTE) ont notamment déployé une "météo de l'électricité", baptisée Ecowatt [1], afin d'alerter la population en cas de pic de consommation prévu dans les jours suivants. Ce système repose alors sur le civisme et la responsabilité individuelle pour réduire la consommation et éviter les déséquilibres.

De la production d'électricité à la satisfaction de l'utilisateur final, l'électricité est échangée via deux marchés : le *marché de gros* et le *marché de détail*, voir Figure 2.1 pour une illustration schématique. Nous décrivons ci-après le modus operandi des deux marchés (acteurs, quantité échangée, fenêtre temporelle).

[1] `https://www.monecowatt.fr/`

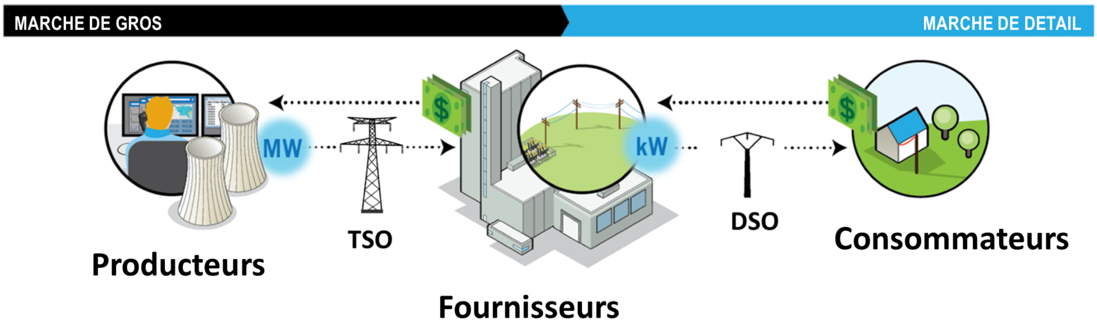


Figure 2.1: Marché de gros et de détail (Source: PJM)

2.1.2 Marché de gros

Plusieurs acteurs participent au marché de gros[2] :

- ◇ *Producteurs* : Entreprises exploitant de grandes centrales (centrales nucléaires, parcs éoliens, barrages hydroélectriques, etc.) à des fins de production d'électricité. En France, 95% de la production est réalisée par EDF et Engie[3]. La répartition de la production française sur chaque ressource est représentée en Fig. 2.2.

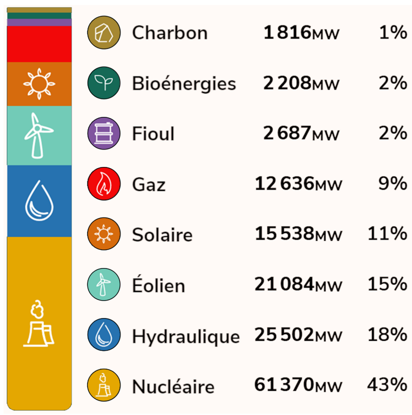


Figure 2.2: Capacités installées des centrales françaises (Source : RTE, 12/2022)

- ◇ *Fournisseurs* : Agents disposant d'un portefeuille d'utilisateurs finaux et négociant de l'électricité en agrégeant la consommation de leurs clients. Comme les producteurs, ils sont également responsables du paiement du déséquilibre global de leur portefeuille. En France, comme en Europe, il existe quelques grands distributeurs (EDF, Engie, TotalEnergies) et plusieurs petits acteurs (Eni, Enercoop, ekWateur, etc.).

- ◇ *Régulateur du marché* : entité veillant au bon fonctionnement des marchés, au bénéfice des consommateurs finaux et conformément aux objectifs de la politique énergétique. En France, ce rôle est assumé par la CRE (Commission de Régulation de l'Energie).

- ◇ *Marché d'échange* : Entité qui exploite la plate-forme de marché et facilite les échanges d'énergie entre les acteurs. En France, ce rôle est assumé par EPEX spot pour le marché $J-1$ et par EEX pour les marchés à termes.

Les quantités échangées sur ce marché sont importantes, et les puissances circulant dans le réseau électrique sont typiquement en MW. Ce réseau de transport à haute tension est entretenu par un *Gestionnaire de Réseau de Transport* (GRT). Le GRT est chargé d'équilibrer le marché : il mesure le déséquilibre et prend des mesures correctives si nécessaire. Le GRT français est RTE.

L'échange d'énergie sur le marché de gros est divisé en plusieurs sous-marchés, correspondant à plusieurs horizons temporels :

[2] La description est inspirée des pages Web suivantes : `https://www.incite-itn.eu/blog/introduction-to-electricity-markets-its-balancing-mecanism-and-the-role-of-renewable-sources/` et `https://learn.pjm.com/electricity-basics/market-for-electricity.aspx`

[3] `https://selectra.info/energie/guides/comprendre/electricite/production`

(i) *Marché à terme* : les contrats entre les fournisseurs et les détaillants sont signés des **mois** à l'avance. Ces contrats constituent une option sûre en termes de revenus (prix fixe), mais c'est aussi la plus risquée en termes de programmation de la consommation/production d'électricité (incertitude élevée).

(ii) *Marché J-1* : l'électricité est vendue **un jour avant** que l'échange d'électricité se produise dans le réseau. Il s'agit d'un marché *spot* : l'offre et la demande sont agrégées pour trouver l'équilibre : une fois toutes les offres reçues, le prix d'équilibre du marché (*prix spot*) représente la valeur où l'offre et la demande se rencontrent. À la fin, toutes les offres de production et de demande qui sont égales ou inférieures au prix d'équilibre du marché sont approuvées. Ce mécanisme favorise les centrales électriques "les moins chères" (coût marginal le plus bas) : la ressource la moins chère participera en premier à la production d'électricité, suivie de l'option suivante la moins chère et ainsi de suite jusqu'à ce que la demande soit satisfaite, voir Figure 2.3. Le prix d'équilibre est alors interprété comme le prix de la ressource la plus chère qui contribue à la production d'énergie.

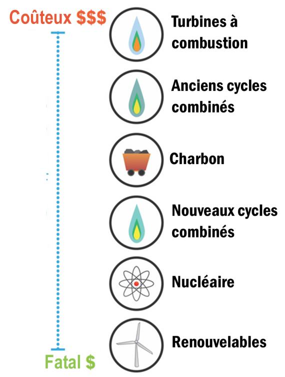


Figure 2.3: Ressources triées (Source : PJM)

(iii) *Marché infra-journalier* : les producteurs et les détaillants peuvent corriger **dans la journée** leur transaction passée pour ajuster plus précisément l'offre à la demande.

(iv) *Mécanismes d'ajustement* : ce marché est responsable de l'ajustement en **temps réel** afin de maintenir la fréquence du réseau aussi proche que possible de la fréquence de référence.

La crise énergétique actuelle[4]

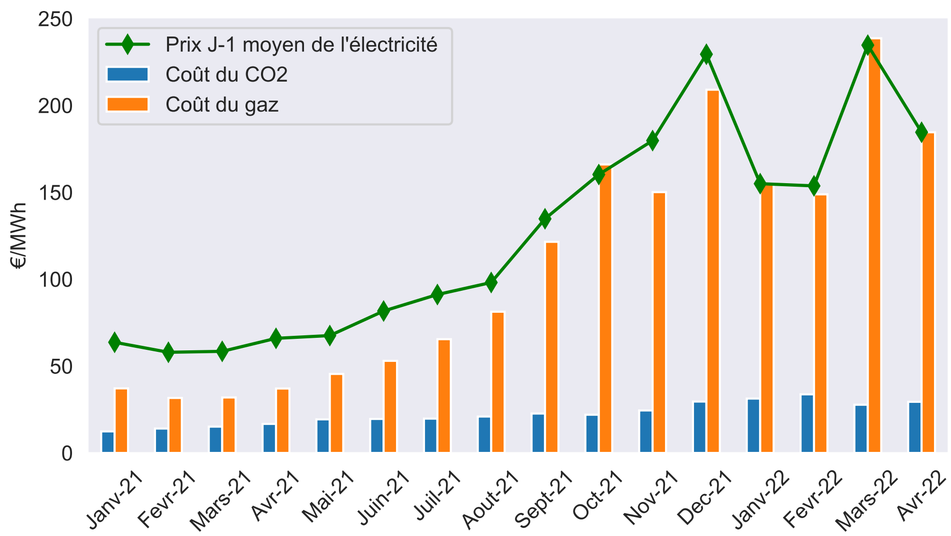


Figure 2.4: Corrélation entre coût du gaz, coût du CO2 et prix de l'électricité en Europe (Source : eurelectric.org)

Les causes de la crise électrique que nous observons depuis l'automne 2021 sont diverses :

[4] https://www.epexspot.com/en/energycrisis

(i) La première raison est l'énorme flambée des prix du gaz qui s'ajoute à la fin de la période Covid. Comme expliqué précédemment, le prix spot est déterminé par la ressource électrique la plus chère utilisée dans le mix électrique, qui correspond souvent à une turbine à gaz. Figure 2.4 montre la forte corrélation entre l'évolution du prix day-ahead et le coût du gaz en Europe.

(ii) La deuxième raison est la faible capacité de production de nombreuses ressources : faible éolien, stockage de gaz inférieur à la normale, faibles niveaux des réservoirs d'eau et faible production nucléaire à partir de la fin 2022.

(iii) La troisième raison est l'augmentation des prix du CO_2, entraînée par le passage à l'objectif de 55% de réduction des émissions pour 2030 (Package "Fit for 55", Conseil européen[5]), qui a encore fait grimper les coûts des centrales électriques conventionnelles.

2.1.3 Marché de détail

Une grande variété d'offres...

Une fois l'électricité achetée sur le marché de gros par les revendeurs, elle peut être vendue aux utilisateurs finaux (la population) sur le marché de détail. Ces contrats correspondent souvent à une fonction affine de la consommation : ils sont composés d'un prix fixe (abonnement et installation) et d'un prix variable (en €/kWh). En France, certains contrats comportent plusieurs parts variables, correspondant à plusieurs périodes (Contrats Heures Pleines/Heures Creuses par exemple).

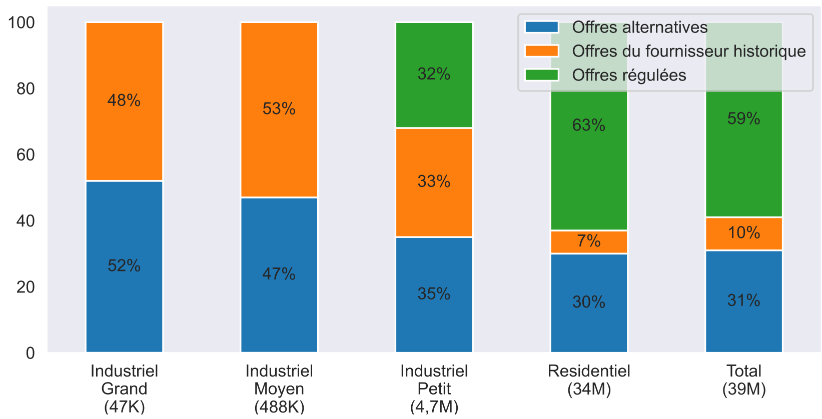


Figure 2.5: Répartition des consommateurs français en 2022 (Source : CRE)

Les consommateurs peuvent désormais choisir librement leur offre d'électricité (c'est le cas en France pour les particuliers depuis 2007). En effet, il existe aujourd'hui de nombreux fournisseurs alternatifs au fournisseur historique. Il existe tout de même des offres réglementées (par exemple, Tarif Bleu en France) où les fournisseurs revendent l'électricité à des prix déterminés par les régulateurs gouvernementaux. Les quantités échangées sur le marché de détail sont d'un ordre inférieur à celles du marché de gros (la puissance maximale des utilisateurs finaux est en général de l'ordre de quelques dizaines de kW). L'électricité parvient aux consommateurs finaux via le réseau basse tension, exploité par un *Gestionnaire du réseau de distribution (GRD)*, responsable de l'acheminement de l'électricité jusqu'aux utilisateurs finaux. En France, le GRD est

[5]`https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/`

Enedis. Figure 2.5 décrit la répartition actuelle des consommateurs, en distinguant les contrats du fournisseur historique (Électricité de France) aux nouveaux fournisseurs.

...et une myriade de clients

Une spécificité des marchés de détail de l'électricité est l'asymétrie entre le nombre de vendeurs (fournisseurs d'électricité) et d'acheteurs (utilisateurs finaux). En effet, les utilisateurs finaux correspondent à l'ensemble des ménages d'un pays. Sur ce marché, chaque fournisseur vise à attirer une partie de la population dans son portefeuille, en concevant un menu compétitif d'offres/contrats.

Le choix des clients est fortement déterminé par la facture d'électricité : un utilisateur final pleinement rationnel sélectionnera le contrat le moins cher pour ses besoins. Classiquement, étant donné un contrat, la facture d'un client est proportionnelle à sa consommation, influencée par plusieurs facteurs :

- *Appareils de chauffage et de climatisation* : un facteur important est le type de chauffage. En France, de nombreux foyers (environ 35%[6]) utilise l'électricité comme source de chauffage.
- *Composition du ménage* : la consommation dépend fortement de la composition du ménage : à titre d'exemple, la période et la quantité seront différentes entre une personne retraitée et une famille avec un enfant.
- *Géographie :* Dans certains pays, en particulier les pays étendus, la météo peut varier considérablement selon les régions. Par exemple en France, les besoins et les habitudes de consommation diffèrent sensiblement du nord au sud.

Pour aider les clients à choisir la bonne offre parmi l'ensemble des contrats du marché, des comparateurs de prix ont été déployés, soit par des entreprises, soit par des agences étatiques (voir Figure 2.6 pour le cas français). Ces outils demandent essentiellement aux clients de fournir les trois facteurs mentionnés ci-dessus et affichent la facture annualisée pour chaque offre disponible sur le marché de détail.

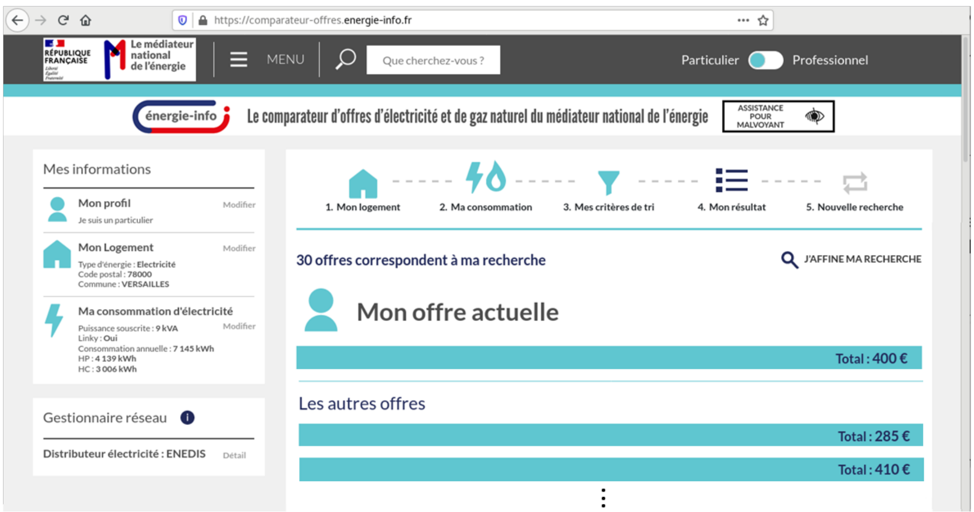
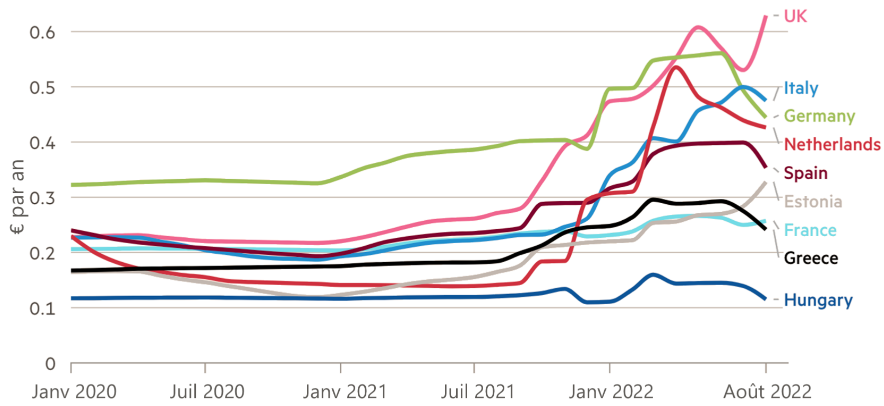


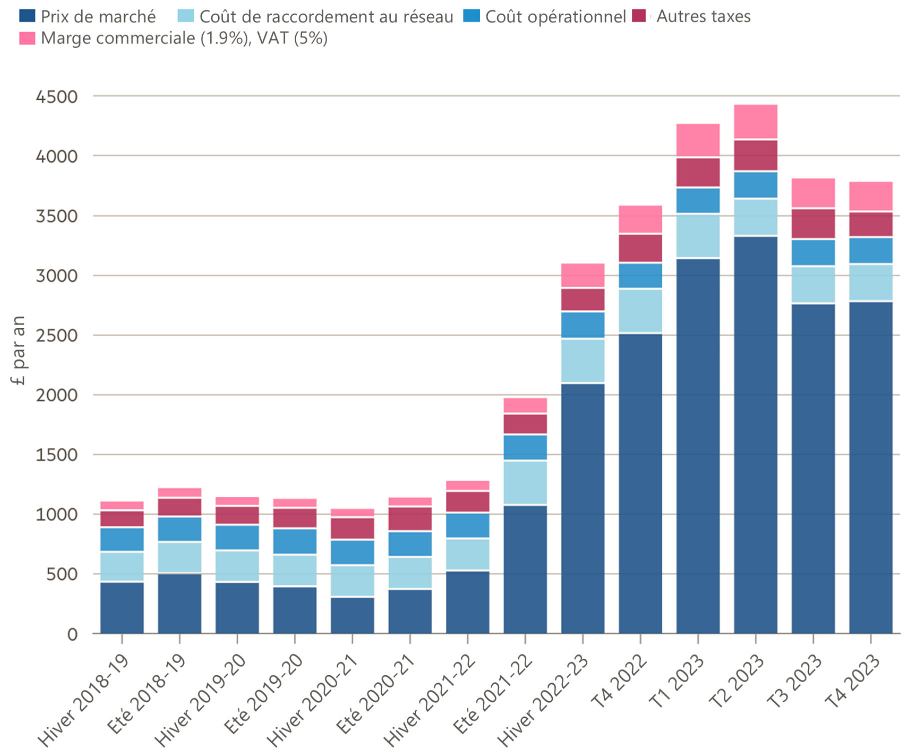
Figure 2.6: Exemple de comparateur de prix (où le nom des contrats a était masqué).

[6] https://www.voltalis.com/comprendre-electricite/les-types-de-chauffage-preferes-des-foyers-francais-1772

La crise énergétique actuelle



(a) Prix moyen par ménage en Europe (Source: energy.eu).



(b) Factures d'un ménage typique au Royaume-Uni (consommation de 2900 kWh pour l'électricité et 12000 kWh pour le gaz, source: Finantial Times et Cornwall Insight, Ofgem).

Figure 2.7: Conséquence de la crise énergétique sur le marché de détail.

La flambée des prix spot sur le marché de gros se traduit par une explosion des coûts pour les revendeurs sur le marché de détail. Figure 2.7a souligne que tous les pays européens ont été impactés par cette crise. Néanmoins, l'intensité de la hausse du prix de l'électricité pour les ménages n'a pas été uniforme : dans certains pays (comme en France), les gouvernements ont mis en place des boucliers tarifaires pour plafonner les prix et éviter une hausse trop importante. La hausse la plus frappante a eu lieu au Royaume-Uni où la facture annuelle d'électricité a été multipliée par un facteur 3 entre 2020 et 2023, voir Figure 2.7b.

Outre la flambée des prix et les conséquences qui en résultent pour les utilisateurs finaux, de nombreux fournisseurs ont été contraints de déposer le bilan. En France, trois fournisseurs alternatifs (E.Leclerc, Planète Oui, Cdiscount) ont déjà cessé leurs activités et au Royaume-Uni, le septième fournisseur d'électricité (Bulb Energy, 5% de part de marché, 1,7 million de consommateurs) a également dû cesser ses activités. Dans ces cas, les clients sont la plupart du temps redirigés vers le fournisseur historique.

2.2 Défis émergents

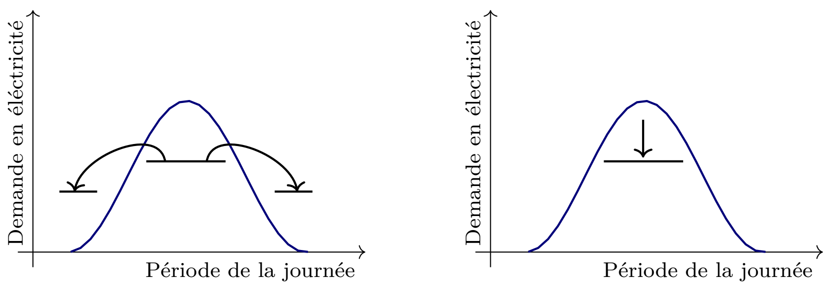
Comportement complexe des utilisateurs

Plusieurs composants (non monétaires) peuvent influencer la décision des utilisateurs finaux. Parmi eux, une partie croissante des consommateurs regarde attentivement l'origine de l'électricité, c'est-à-dire quelle ressource a été utilisée pour la production de l'électricité. A titre d'exemple, depuis une décennie, on assiste à l'émergence d'offres “vertes”, garantissant une électricité issue de ressources renouvelables [HR14; MZ16; AF19].

Par ailleurs, trois caractéristiques comportementales sont souvent étudiées car elles impactent fortement le choix des clients :

(i) *Rationalité partielle/limitée* : en raison du manque d'informations sur leurs caractéristiques de consommation ou de l'ignorance des mises à jour des prix, les clients ne sélectionnent pas toujours le contrat qu'un agent omniscient aurait choisi. Ce comportement sous rationalité limitée, c'est-à-dire avec une certaine incertitude dans la décision, est maintenant intégré dans un nombre croissant de modèles [ES17; RH18; RDP19], impliquant souvent des choix de clients de nature probabiliste (par exemple, les modèles de choix logit [Tra09]), plus difficile à gérer dans les problèmes d'optimisation [BLS23].

(ii) *Élasticité au prix* : la crise énergétique a mis en évidence une flexibilité non négligeable dans la consommation des utilisateurs finaux, induite par une forte fluctuation des prix (ici la flambée des prix). Ils ont restreint leur consommation pour se concentrer uniquement sur les usages essentiels. De nombreuses études abordent ce phénomène en estimant son intensité [And+97; Lij07; ACR20; NYK20]. Cette importante adaptabilité est à la base du “Demand Response”, voir par exemple [AE08]. Cette technique vise à modifier les courbes de charge des consommateurs en envoyant à ces derniers des incitations tarifaires, le but étant souvent de faire baisser le pic de consommation (“peak clipping” et “load shifting”), voir Figure 2.8.



(a) Load shifting (b) Peak clipping

Figure 2.8: Deux mécanismes de Demand response

(iii) *Inertie dans la réponse*: comme le montre Figure 2.6, les moteurs de comparaison de prix comparent souvent l'offre actuelle de l'utilisateur final avec les alternatives disponibles sur le marché. Il existe en effet une asymétrie entre le contrat déjà souscrit avec le reste du marché puisque les clients font face à des *coûts de changement* ce qui provoque une inertie dans la dynamique de choix : les utilisateurs finaux ne changeront leur offre pour une autre que s'il existe un différence substantielle entre les deux contrats qui compense les coûts de changement. Ces coûts ont diverses origines : comme l'illustre Figure 2.9, il peut s'agir de coûts monétaires (frais de sortie, coûts d'équipement), mais ils peuvent être psychologiques (fidélité à la marque, temps et efforts consacrés au changement d'offre). Nous nous référons à [Stu02; Sal22; DW23] pour les estimations des coûts de changement sur les marchés de détail de l'électricité. Des études empiriques analysent l'impact de ces coûts de changement

Figure 2.9: Exemple de coûts de changement (Source : sketchbubble.com)

à la fois sur la décision des clients et sur les stratégies de tarification [FH08; HP10; Abd17; Mys+18], et plus récemment certains travaux intègrent la dépendance temporelle (inertie due aux coûts de changement) dans des modèles d'optimisation [PE17; Hua22], montrant notamment le lien entre coûts de changement et promotions.

Lisibilité du marché

Ces dernières années, les offres se sont multipliées : tant le nombre de revendeurs que le nombre de contrats par fournisseur ont augmenté (par exemple de 18 en 2008 à plus de 100 revendeurs aujourd'hui en France[7]). En conséquence, les consommateurs souffrent d'une lisibilité réduite du marché et sont parfois désorientés dans cette vaste jungle d'offres. Mais dans le même temps, disposer d'un *menu* d'offres plus large permet de mieux s'adapter aux préférences et aux habitudes des clients, et donc de renforcer son attractivité. Une question naturelle se pose alors pour chaque fournisseur: quel est le compromis optimal entre ces deux tendances opposées ? Cette question est moins abordée dans la littérature alors qu'elle est une préoccupation majeure chez les fournisseurs.

Mécanismes incitatifs à des fins de réduction de la consommation

Dans un contexte de sobriété, dû à la fois à la crise énergétique (voir sections précédentes) et à des considérations écologiques, les distributeurs visent à inciter leurs clients à réduire leur consommation, non seulement en période de pointe (comme pour l'effacement) mais en moyenne.

[7]Source : Commission de Régulation de l'Energie (CRE)

Ces incitations peuvent prendre plusieurs formes. Tout d'abord, elles peuvent être basées sur le civisme et la conscience des clients : certaines applications montrent à chaque utilisateur final à quel point sa consommation est éloignée de la consommation moyenne des clients similaires à ce dernier (c'est le cas en France pour la plupart des fournisseurs, comme Engie et EDF). Une autre approche est basée sur la récompense monétaire, offrant la plupart du temps des remises si le client a pu réduire sa consommation par rapport aux années précédentes ("SimplyEnergy"[8], "Plüm énergie"[9] ou "OhmConnect"[10]). Des études économiques examinant l'impact de ces incitations monétaires sur les clients peuvent être trouvées dans la littérature (par exemple [MT12; Pra+18; Rus+23]).

2.3 Tarification et jeux de Stackelberg dans la littérature

La théorie des jeux est un domaine mathématique qui analyse l'interaction entre des agents rationnels (appelés *joueurs*). La première contribution est sans doute le problème du duopole formulé par Cournot en 1838 [Cou38], où la notion d'équilibre est abordée, c'est-à-dire une situation où chaque joueur ne veut pas dévier de sa situation actuelle. Les jeux mathématiques sont ensuite théorisés par von Neumann et Morgenstern en 1944 [VM47]. Dans les années 50, Nash [Nas53] formalise la notion d'équilibre, très connue et utilisée aujourd'hui. Dans cette thèse, nous nous concentrons sur une classe spécifique de jeux, à savoir les jeux de Stackelberg [Sta52]. Ces problèmes à 2 joueurs reflètent l'asymétrie des rôles des joueurs : le meneur décide d'abord, puis le suiveur répond (jeux séquentiels). Cette classe de jeux est particulièrement bien adaptée aux problèmes de tarification dans lesquels une entité fixant les prix vise à optimiser ses revenus. La décision du client intervient donc toujours après l'établissement de la nouvelle stratégie tarifaire.

Nous nous concentrons ici sur deux cadres traitant des jeux Stackelberg, et avec de nombreuses applications à la gestion des prix : l'optimisation bi-niveau et les modèles principal-agent. Pour une introduction sur ces deux domaines, nous renvoyons au chapitre préliminaire 3. Nous présentons ci-après un aperçu des travaux existants sur la tarification.

Tarification en optimisation bi-niveau

Les problèmes à deux niveaux constituent une partie importante de la littérature mathématique des problèmes de tarification. Il s'agit de problèmes à deux joueurs, où un premier joueur (meneur) propose des prix aux clients (suiveurs) qui maximisent leurs propres fonctions d'utilité et tiennent compte des prix fixés par le meneur. Nous renvoyons à Labbé et Violin [LV15] pour une présentation détaillée des concepts et des modèles de solutions sur la tarification bi-niveau, ainsi qu'aux travaux de Dempe [Dem20] pour une revue approfondie de la littérature bi-niveau.

Premiers travaux sur la tarification des réseaux. Labbé et al. [LMS98] ont d'abord proposé un modèle de taxation appliqué à la tarification routière, étendu ensuite par Brotcorne et al. [Bro+06]. Larsson et Patriksson [LP98] ont étudié la question de la congestion du trafic à travers la programmation à deux niveaux. En gestion du fret, Brotcorne et al. [Bro+00] ont proposé un modèle et des algorithmes pour déterminer les tarifs optimaux. Ces modèles sont appliqués à des réseaux, et le problème rencontré par les suiveurs hérite de la structure de graphe induite par la topologie du réseau. Par exemple, étant donné la décision du meneur, le problème du suiveur peut être un problème de plus court chemin. Dewez et al. [Dew+08] ont profité de la

[8] `https://www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward`
[9] `https://plum.fr/cagnotte/`
[10] `https://www.ohmconnect.com/`

structure pour proposer de nouvelles inégalités valides pour resserrer les relaxations, ainsi que de nouvelles formulations basées sur la description des chemins.

Tarification de biens. Le problème de la tarification de produits "sans convoitise", i.e., où chaque agent est toujours en mesure de choisir l'option qu'il désire le plus (voir Chapter 3) a été analysé pour la première fois par Guruswami et al. [Gur+05] sous l'angle de la complexité, montrant le caractère APX-Hard du problème. Par la suite, plusieurs reformulations du problème à deux niveaux ont été proposées par Shioda, Tunçel et Myklebust [STM11] et Fernandes et al. [Fer+13]. Heilporn [Hei+10] et Fernandes et al. [Fer+16] ont étudié les liens entre ce dernier problème de tarification des produits et les modèles de tarification rencontrés dans les réseaux, montrant en particulier que des inégalités valides d'un modèle peuvent être utilisées dans l'autre.

Applications dans le domaine électrique. Pendant une décennie, de nombreux problèmes de tarification se sont posés dans la gestion de l'énergie. Avec la libéralisation des marchés de détail de l'électricité, Leyffer et Mundson [LM10] ont abordé la question de la tarification optimale à laquelle doit faire face un détaillant dans un environnement concurrentiel, amenant à la résolution d'un problème d'équilibre sous contraintes d'équilibres (EPEC). Luna et al. [Lun+20] ont ensuite proposé une technique de régularisation primale-duale pour contourner le caractère mal posé du problème résultant de la non-unicité de la solution du niveau bas. De plus, de nombreuses contributions ont été faites en "Demand Response" : Zugno et al. [Zug+13] font figure de pionniers sur ce sujet. Afşar et al. [Afş+16] ont étudié le tarif optimal afin d'inciter les consommateurs à écrêter la pointe de consommation. Sur le même sujet, Kovacs [Kov16] a proposé un tarif alternatif dépendant du temps. Alekseeva et al. [Ale+19] ont étudié des stratégies pour déplacer le pic de consommation à des fins de réduction des coûts. Aussel et al. [Aus+20] ont enrichi le modèle en ajoutant un niveau d'optimisation (fournisseurs, agrégateurs, utilisateurs finaux) et ont développé des règles de partage alternatives en cas de choix équivalents, s'interprétant comme un intermédiaire entre la formulation optimiste et la formulation pessimiste. Abate, Riccardo et Ruiz [ARR21] ont introduit de l'incertitude dans les fonctions d'utilité des consommateurs pour tenir compte de l'élasticité de la demande. Grimm et al. [Gri+21] ont comparé le temps d'utilisation (TOU), la tarification de pointe critique (CPP) et le tarif de tarification en temps réel (RTP) dans une étude numérique réalisée sur des cas d'utilisation réalistes. Kozanidis et Kostarelou [KK23] ont proposé un modèle à deux niveaux pour gérer les offres de prix optimales des producteurs d'énergie sur les marchés de l'électricité J-1.

Géométrie tropicale. Une approche émergente est le point de vue adopté par Baldwin et Klemperer [BK19] et Tran et Yu [TY19], où la décision du client est analysée à l'aide de la géométrie tropicale. Cette géométrie utilise l'algèbre max-plus, voir par exemple [MS21]. Le problème de maximisation de l'utilité du suiveur consiste alors en une sélection de type "une parmi N possibilités", et peut être vu comme l'évaluation d'un polynôme tropical. La carte de réponse des agents est dans ce contexte un complexe polyédral où chaque cellule correspond à un sous-ensemble de décisions du meneur (prix) qui induisent une même décision client. Lorsque le problème implique plusieurs suiveurs indépendants, la "superposition" de complexes polyédriques donne une compréhension visuelle complète du problème de tarification. Nous renvoyons à [Eyt18] pour une approche tropicale de la tarification à deux niveaux appliquée aux réseaux de télécommunications mobiles.

Travaux	Acteurs	(i) Temporalité (ii) Incertitude	Nature du niveau bas	(i) Approche (ii) Type de solution
[LM10]	Multi-L. Common-F.	(i) Statique (ii) Déterministique	Continu	(i) Non-linéaire (formulation NCP) (ii) Stationnaire
[Zug+13]	Single-L. Multi-F.	(i) Temps discret (ii) Par scénarios	Linéaire	(i) Reformulation KKT (ii) Globale
[Afş+16]	Single-L. Multi-F.	(i) Temps discret (ii) Déterministique	Linéaire	(i) Reformulation KKT (ii) Globale
[Kov16]	Single-L. Multi-F.	(i) Temps discret (ii) Déterministique	Linéaire	(i) Heuristique
[Ale+19]	Single-L. Multi-F.	(i) Temps discret (ii) Déterministique	Linéaire	(i) Reformulation KKT (ii) Globale
[Lun+20]	Multi-L. Common-F.	(i) Statique (ii) Déterministique	Continu	(i) Non-linéaire (régularisation duale) (ii) Stationnaire
[Aus+20]	Single-L. Multi-F.	(i) Temps discret (ii) Déterministique	Linéaire	(i) Reformulation KKT (ii) Locale
[ARR21]	Single-L. Multi-F.	(i) Temps discret (ii) Par scénarios	Linéaire	(i) Reformulation KKT (ii) Locale
[Gri+21]	Single-L. Single-F.	(i) Temps discret (ii) Déterministique	Linéaire	(i) Reformulation S.-D. (ii) Globale
[KK23]	Single-L. Single-F.	(i) Temps discret (ii) Déterministique	Entier	(i) Reformulation Opt.-V. (ii) Globale

Table 2.1: Classification non-exhaustive des travaux en optimisation bi-niveau appliqués à la tarification électrique.
“S.-D.”: “Strong Duality”, “Opt.-V.”: “Optimal-Value”, “L.”: “Leader”, “F.”: “Follower”.

Modèles principal-agent pour la gestion de l’énergie

Les modèles Principal-Agent appartiennent à la théorie des incitations, dont la première approche remonte aux travaux de Barnard [Bar38]. Cette théorie a d’abord été conçue dans le but d’améliorer la productivité des travailleurs, dans le même esprit que le taylorisme. Comme dans la tarification à deux niveaux, le principal (alias le leader) conçoit un contrat en tenant compte du comportement rationnel (appelé *aléa moral* dans ce domaine) de l’agent (alias le suiveur), qui vise à maximiser sa propre fonction d’utilité.

Au cours de la dernière décennie, les contributions sur les modèles Principal-Agent appliqués à la gestion de l’énergie ont été nombreuses. Carmona, Fehr et Hinz [CFH09] ont introduit un problème de contrôle stochastique optimal pour la tarification du carbone. Féron, Tankov et Tinsi [FTT20] ont étudié le processus d’échange sur les marchés infrajournaliers de l’électricité à l’aide de modèles principal-agent. Alasseur, Farhat et Saguan [AFS20] ont analysé à travers le mécanisme de partage des risques de nouveaux leviers pour favoriser les investissements capacitaires. Concernant l’effacement de la demande, Aïd, Possamï et Touzi [APT22] ont présenté des incitations optimales pour réduire à la fois la consommation moyenne d’électricité et sa variance. Dans [Éli+20], Élie et al. ont prouvé qu’une expression analytique des contrats optimaux peut être trouvée dans le cas de la valorisation linéaire de l’énergie et lorsque les contrats sont indexés par la consommation agrégée. La question de la tarification de l’électricité sur le marché de détail a été abordée par Alasseur et al. in [Ala+20] où l’élasticité de la demande est décrite par une mesure relative d’aversion au risque (CRRA). Des applications récentes ont émergé, comme les marchés des énergies renouvelables par exemple. Aïd, Kemper et Touzi [AKT23] ont introduit un cadre Principal-Agent afin de modéliser les investissements dans les énergies renouvelables. Aussi, Shrivats, Firoozi et Jaimungal [SFJ21] ont proposé un jeu de champ moyen principal-agent appliqué au marché des certificats d’énergie renouvelable.

2.4 Objectifs et contenu de la thèse

Dans cette thèse, nous analysons les problèmes de tarification qui se posent dans la gestion de l'énergie, et plus précisément à l'interface entre les distributeurs et les consommateurs. Premièrement, nous nous concentrons sur la tarification des offres pour le marché de détail. Dans ce cadre, il s'agit de construire un menu d'offres le plus rentable possible pour le fournisseur d'électricité, où chaque offre peut être sélectionnée par n'importe quel client mais est conçue pour être attrayante pour une partie ciblée de la population. Dans un deuxième temps, nous étudions des mécanismes incitatifs qui visent à réduire la consommation d'électricité d'un portefeuille d'utilisateurs finaux en construisant des récompenses optimales favorisant les clients les plus sobres. De plus, dans la plupart des problèmes de tarification de l'électricité, une donnée critique qui doit être estimée est la consommation et la facture typiques (basées sur des données passées) des clients. Nous examinons ici des estimations fines dans ce but. Cette thèse contribue à répondre à ces différentes problématiques de par ses aspects novateurs tant d'un point de vue théorique, modélisation et algorithmique. Nous donnons maintenant plus de détails sur le contenu de cette thèse :

- ◇ Nous abordons d'abord à travers l'optimisation bi-niveau la question de la tarification optimale d'un menu d'offres, problème crucial pour une entité qui cherche à rester compétitive en concevant son menu tarifaire en réaction à un contexte de marché. Ce travail étend les modèles développés dans les approches standard pour la tarification de biens. Ce cadre est de nature combinatoire, la décision des clients consistant en un choix discret parmi la gamme d'offres. Dans le chapitre 4, nous introduisons un nouveau comportement client – vu comme une régularisation de modèles de choix purement déterministes – où la décision de chaque utilisateur final est maintenant exprimée comme une mesure de probabilité discrète sur les alternatives, concentrée sur les contrats les plus avantageux pour le consommateur. Ce nouveau modèle est motivé par le grand nombre de consommateurs, considérant chaque groupe d'utilisateurs finaux (consommateurs ayant des caractéristiques similaires) comme un utilisateur final représentatif d'une sous-population homogène de taille infinie. Dans le chapitre 5, nous enrichissons le comportement des clients en insérant dans le problème du suiveur des coûts de changement, agissant comme une force de rappel qui maintient les clients attachés à leur offre actuelle. Cet effet d'inertie plonge le problème dans un cadre de contrôle dynamique : le modèle correspond maintenant au contrôle d'un processus de décision markovien où les probabilités de transition du suiveur (d'un contrat à un autre) sont déterminées par la décision du meneur à chaque pas de temps. Nous étudions dans le chapitre 6 un autre point d'intérêt pour les opérationnels : l'influence de la taille du menu, c'est-à-dire le nombre de contrats à concevoir. Cette question est abordée par une stratégie en deux étapes : nous examinons d'abord le cas limite champ moyen - dans lequel un continuum d'offres est conçu - puis nous recherchons le meilleur menu de taille finie qui approche au mieux le cas limite. Ce problème de quantification permet de déterminer une fonction de perte qui met en correspondance le nombre de contrats utilisés avec la perte relative de revenus induite par cette restriction.

- ◇ Nous rassemblons ensuite des contributions de diverses natures. Dans le chapitre 7, nous considérons d'abord la question de la sobriété énergétique en nous focalisant sur un cadre Principal-Agent, dans lequel une entreprise vise à proposer un nouveau type de contrat qui incite les clients à faire des efforts en vue de réduire leur consommation. Ce contrat prévoit une rémunération variable en plus de la tarification linéaire standard (facture proportionnelle à la consommation) qui est basée sur le rang de l'utilisateur final dans l'ensemble des utilisateurs similaires. Ce jeu de classement introduit une compétition (équilibre de champ moyen) entre les consommateurs, les poussant à baisser leur consommation électrique moyenne. Dans le chapitre 8, motivés par le grand nombre d'applications nécessitant des estimations de factures, nous nous concentrons sur les inégalités de concentration et leur intégration dans le modèle d'optimisation. En particulier, nous nous intéressons aux inégalités de type Bennett et montrons qu'il est possible d'obtenir une reformulation convexe, conduisant à des programmes avec contraintes en probabilité traitables. Enfin, nous étudions dans le chapitre 9 une famille d'approximations non convexes de problèmes d'optimisation sous contrainte de cardinalité. Ces problèmes d'optimisation parcimonieuse apparaissent naturellement lors de la conception d'un menu d'offres, qui est au cœur du chapitre 6. Nous abordons ici la question de manière générale et montrons que l'approche non linéaire basée sur les entropies de Rényi est peu gourmande en temps tout en conduisant à une solution acceptable, c'est-à-dire avec une violation limitée de l'exigence de cardinalité d'origine.

2.5 Contributions des chapitres

Nous donnons maintenant plus de détails sur les apports de chaque chapitre:

1. Le chapitre 4 énonce la problématique rencontrée par un fournisseur d'électricité qui vise à concevoir un menu d'offres destinées à un grand nombre d'utilisateurs finaux. Une description complète du modèle conduirait à un Multi-Leader-Common-Follower Game, voir [LM10], où un équilibre de Nash devrait être trouvé entre les fournisseurs. Ici, nous étudions une version statique où un fournisseur optimise ses offres compte tenu de la stratégie (prix) des autres acteurs du marché (fournisseurs concurrents et offres régulées), conduisant à un problème Single-Leader-Single-Follower, voir Figure 2.10. Cela peut être interprété comme un ajustement immédiat d'un fournisseur à une situation de marché. Nous introduisons d'abord des modèles déterministes basés sur le cadre de tarification bi-niveau de biens (Envy-free unit-demand, voir par exemple [Gur+05]), et dérivons des formulations MILP issues d'une reformulation à un seul niveau. Nous rappelons également les modèles Logit [GMS15], vus ici comme une régularisation probabiliste du comportement des clients, et analysons leur comportement asymptotique. Le souci principal est d'avoir un modèle à la fois juste et traitable, et nous développons à cet effet une régularisation quadratique pour le problème du suiveur. Nous exposons une structure polyédrique de la réponse des clients (Théorème 4.3.1), et montrons que lorsque le paramètre de régularisation - interprété comme l'intensité de rationalité des utilisateurs finaux - tend vers l'infini (entièrement rationnel), la réponse déterministe est récupérée, c'est-à-dire que le complexe polyédrique converge vers la description cellulaire développée par Baldwin et Klemperer [BK19] dans des contextes déterministes (Théorème 4.3.2). Une étude numérique poussée est réalisée sur une instance réaliste, montrant l'intérêt de ce nouveau modèle (Figure 4.9) et la nécessité de modéliser un comportement lisse des clients pour éviter des prix trop optimistes (non réalistes en pratique), voir Figure 4.10.

2. Dans le chapitre 5, nous intégrons dans le précédent modèle (Chapitre 4) l'influence des coûts de changement. L'inertie qui en résulte dans le comportement des clients se traduit par une dépendance au choix du contrat précédent. Nous modélisons le problème comme un processus de décision de Markov où chaque problème de transition entre deux pas de temps discrets correspond à une instance du modèle statique à deux niveaux, voir Figure 2.11. Nous formalisons le problème comme un problème de contrôle optimal visant à maximiser la récompense moyenne sur un horizon de temps infini (contrôle ergodique, voir par exemple [ABG11]), et montrons que cela peut être résolu de manière équivalente par

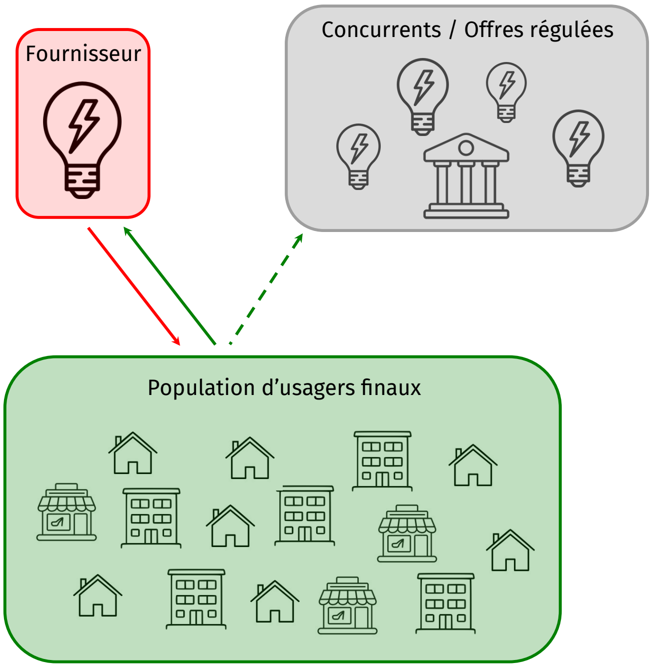


Figure 2.10: Problème de bi-niveau rencontré par un fournisseur désireux d'optimiser son menu d'offres dans un marché concurrentiel

Nous examinons la réaction du fournisseur aux prix donnés de la concurrence (statique). Le fournisseur met d'abord à jour ses prix (flèche rouge), puis chaque utilisateur final de la population hétérogène choisit son offre – appartenant soit au fournisseur (flèche verte pleine), soit à la concurrence (flèche en pointillé).

un problème aux valeurs propres (Corrollaire 5.2.1). Nous prouvons à l'aide d'un argument de contraction sur la dynamique que le problème aux valeur propres a une solution régulière (Théorème 5.2.2), à partir de laquelle la stratégie de leader optimale peut être déduite. Nous introduisons un algorithme d'itération sur les politiques [Put94; Gau96] adapté au contrôle de population décomposable, recalculant à la volée les transitions entre états pour réduire drastiquement l'espace de stockage nécessaire, tout en gardant un coût de calcul rapide par rapport aux algorithmes d'itération sur les valeurs relatives (RVI), voir Table 5.2. Nous analysons ensuite l'impact de l'intensité des coûts de changement sur la politique tarifaire. Nous observons qu'au-delà d'un seuil, des promotions périodiques sont appliquées, offrant une rémunération strictement supérieure au meneur par rapport aux stratégies à prix constants (Figure 5.5). Nous montrons enfin l'étendue de ce comportement dans un exemple, prouvant la supériorité des politiques cycliques sur les politiques à prix constant, voir Section 5.7.

3. Le chapitre 6 explore un autre facteur important dans la conception des menus tarifaires, à savoir la question du nombre optimal d'offres (taille du menu) qu'un fournisseur devrait proposer à la population. Dans ce but, nous assouplissons d'abord la condition sur la taille du menu pour obtenir comme cas limite un problème dans lequel un contrat peut être spécialement conçu pour chacun des utilisateurs finaux (menu de taille infinie). Nous montrons que ce problème relaxé peut être reformulé comme une généralisation du problème présenté par Rochet et Choné [RC98] (Théorème 6.5.1), qui est convexe lorsque le portefeuille de clients est fixé (c-à-d lorsque la souscription au contrat est imposée), voir (6.3)). Étant donné le menu optimal (potentiellement de taille infinie) du problème relaxé, nous

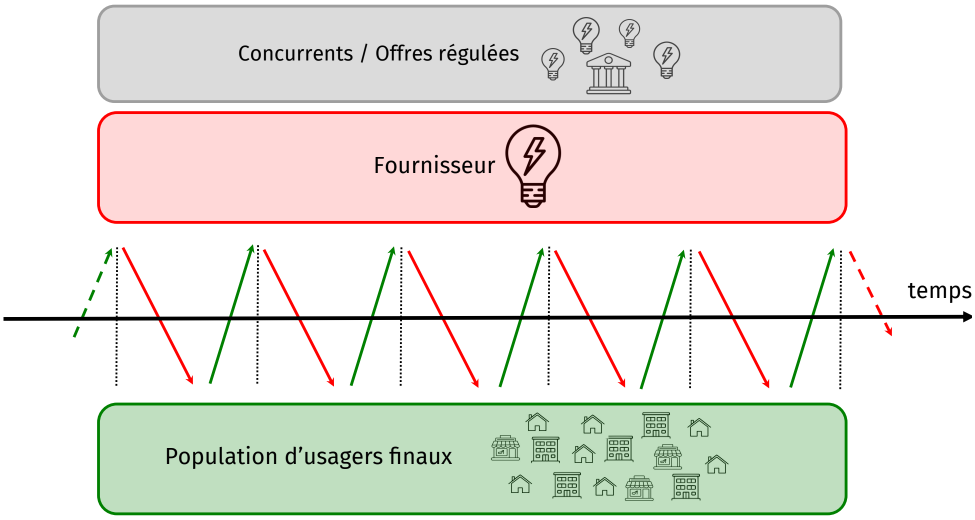


Figure 2.11: Problème itéré de tarification à deux niveaux en tant que processus de décision de Markov.

A chaque pas de temps, le fournisseur est en mesure de mettre à jour ses prix, et chaque utilisateur final de la population hétérogène peut soit conserver son contrat actuel, soit en changer pour un nouveau (du même fournisseur ou des concurrents statiques). Dans ce dernier cas, elle subira des frais de changement.

approchons ce dernier par un menu de nombre fini de contrats en utilisant une procédure d'élagage, initialement conçue pour les problèmes de contrôle, voir Algorithme 6. Ce schéma d'élagage réduit itérativement la taille du menu en supprimant le contrat restant le plus redondant, c'est-à-dire celui qui induit le moins d'erreur d'approximation lorsqu'il est supprimé. Nous adaptons cette nouvelle approche de type "glouton inversé" pour plusieurs critères (par exemple erreur en norme L_∞ ou L_1), et montrons qu'un recalcul partiel de la solution à chaque itération est toujours suffisant (Propositions 6.3.1 et 6.3.2). Nous observons numériquement que cette mise à jour locale entraîne d'énormes gains de temps de calcul, voir Figure 6.3. Nous exploitons ensuite la structure du problème pour interpréter la décision des clients comme un diagramme de Voronoi utilisant les distances de Bregmann [BNN10] – voir Proposition 6.4.1 – et dérivons de la théorie de la quantification [Pag15] une procédure de Lloyd qui vise à trouver la meilleure approximation L_1 du menu de taille infinie par un menu fini, montrant l'efficacité et la robustesse des nouvelles méthodes d'élagage.

4. Le chapitre 7 introduit un nouveau dispositif incitatif, appliqué aux préoccupations de sobriété énergétique. Dans ce schéma, le fournisseur (Principal) – motivé par les agences de régulation – contracte avec un continuum de clients homogènes en leur proposant une récompense monétaire basée sur le classement individuel de chaque client au sein de la population, voir Figure 2.12. Les consommateurs sont alors incités par cette compensation financière à se classer parmi les meilleurs économes en énergie afin de recevoir la plus haute récompense.

   Ici, la concurrence entre les utilisateurs finaux, initiée par une fonction de récompense donnée, est modélisée comme un jeu à champ moyen, et nous prouvons qu'elle admet un unique équilibre de Nash, qui peut être déterminé analytiquement (Théorème 7.2.3). Cette caractérisation vaut en particulier pour les récompenses purement basées sur le rang, large-

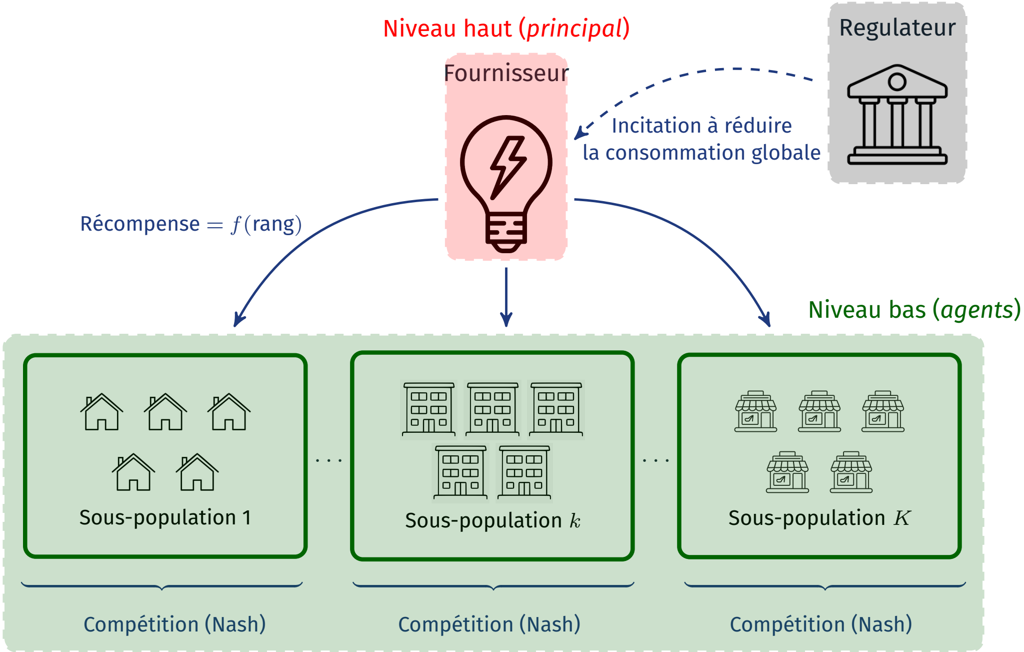


Figure 2.12: Relation Principal-Agent entre un fournisseur d'énergie et un ensemble d'agents, ces derniers rivalisant les uns avec les autres pour obtenir la meilleure récompense possible en réduisant au maximum leur consommation d'énergie.

Motivé par des politiques gouvernementales de réduction d'énergie, le fournisseur conçoit une fonction de récompense incitative basée sur le classement des agents afin que l'agent économisant le plus d'énergie reçoive la récompense la plus élevée. Compte tenu de la fonction d'incitation, un équilibre de Nash doit alors être trouvé pour chaque sous-population d'utilisateurs finaux homogènes.

ment étudiées dans [BZ21]. Ensuite, en incorporant cette formule analytique de l'équilibre dans le problème du fournisseur, nous montrons que lorsque l'élasticité de la demande induite par les prix est uniforme, la récompense optimale peut être obtenue par la résolution d'une équation de point fixe sur la consommation moyenne cible optimale. Pour cela, nous exploitons les conditions d'optimalité d'une reformulation convexe dans l'espace des distributions, équivalente au problème initial, voir Théorème 7.2.4. Cette dernière caractérisation permet une compréhension fine de la fonction de récompense, établissant des conditions suffisantes pour le développement de tels dispositifs incitatifs. Quand la population est hétérogène et avec une élasticité au prix non uniforme, un algorithme numérique (Algorithme 10) est proposé, et des simulations sur des exemples réalistes montrent que l'approche est susceptible de produire d'importantes réductions de consommation tout en garantissant une satisfaction moyenne des utilisateurs finaux supérieure au cas traditionnel (non incitatif), voir Figure 7.5.

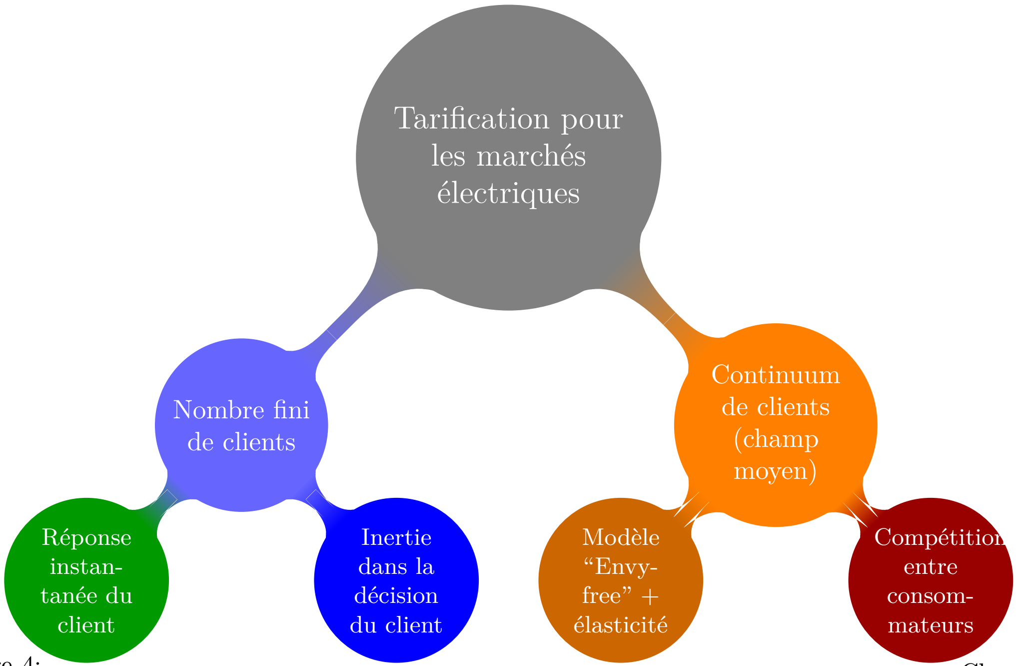
5. Dans le chapitre 8, nous étudions les inégalités de concentration de type Bennett en vue de les intégrer dans des programmes sous contraintes en probabilité [Pré95]. Premièrement, nous introduisons d'abord un algorithme à double dichotomie permettant de calculer des bornes de confiance (Algorithme 12). Ceci peut être appliqué pour l'estimation des factures lorsque les courbes de charge de consommation ne sont connues qu'à travers leurs moments. Ensuite, nous étudions des approximations distributionnellement robuste de programmes sous contraintes en probabilité avec des informations sur les moyennes et les variances (Corollaires 8.4.1 et 8.4.2). Nous montrons que ces approximations peuvent être intégrées

dans des programmes d'optimisation convexe comme dans [NS07], et sont traitables dans de nombreux cas d'utilisation. Nous nous intéressons d'abord à un problème discret de référence, à savoir le problème du sac à dos entier, et montrons qu'une description par plans sécants de l'estimation convexe de type Bennett (ajoutés itérativement dans le Branch-and-Bound) conduit à des solutions robustes plus rentables tout en gardant un temps de calcul raisonnable (Table 8.2). Nous utilisons également ces estimateurs dans des problèmes d'optimisation continue, en nous concentrant sur le problème des séparateurs à vaste marge robustes, et montrons sur des instances de la littérature que l'approche réduit l'erreur de classification (Section 8.4.2).

6. Dans le chapitre 9, nous étudions des problèmes génériques d'optimisation parcimonieux. De tels problèmes se posent naturellement dans la gestion des prix. Par exemple, pour des raisons de lisibilité, les détaillants contraignent souvent leur nombre d'offres à une taille prescrite, voir le chapitre 6. Ici, nous introduisons une famille de bornes inférieures pour cardinalité, faisant intervenir l'entropie de Rényi [Rén+61]. En nous concentrant sur l'application aux problèmes de sélection - où l'optimisation est effectuée sur l'espace des mesure de probabilité discrètes (simplexe), nous prouvons (Théorème 9.3.1) que les bornes entropiques que nous avons développées peuvent contrôler la parcimonie de la solution, récupérant comme cas limite la norme exacte ℓ_0. Dans les résultats numériques, nous étudions le cas particulier de l'entropie de Shannon et sa capacité à imposer de la parcimonie dans la solution. Nous montrons que l'utilisation de cette borne entropique pour problèmes de sélection de portefeuille assure un bon compromis entre le contrôle de la cardinalité et les performances d'optimisation (Figure 9.6).

Classification schématique des chapitres

En Figure 2.13, nous décrivons brièvement les objectifs mathématiques de chaque chapitre par mots-clés et classons les chapitres en fonction des principales différences de modélisation. Les concepts mentionnés dans les mots clés sont introduits dans le chapitre préliminaire (Chapitre 3).



Chapitre 4:

- ⋄ Optimisation bi-niveau bilinéaire
- ⋄ Problèmes bas régularisés
- ⋄ Analyse de complexes polyédraux.

Chapitre 5:

- ⋄ Contrôle ergodique de populations
- ⋄ MDPs déterministes
- ⋄ Analyse de convergence (propriétés "turnpike")

Chapitre 6:

- ⋄ Problème Principal-Agent avec selection adverse
- ⋄ Quantification de menus d'offres
- ⋄ Diagrammes de Voronoï

Chapitre 7:

- ⋄ Problème Principal-Agent avec aléa moral
- ⋄ Jeux à champ moyen
- ⋄ Incitations basées sur le rang

Figure 2.13: Schéma exprimant les différences entre les chapitres en termes de modélisation, ainsi que des mots-clés décrivant les principaux outils utilisés dans chaque chapitre.

Publications and Works related to this PhD thesis

***

[Jac+23a] Quadratic Regularization of Bilevel Pricing Problems and Application to Electricity Retail Markets.
Joint work with Wim van Ackooij, Clémence Alasseur and Stéphane Gaubert
European Journal of Operational Research, 2023, DOI: 10.1016/j.ejor.2023.05.006
See Chapter 4.

[Jac+22] Ergodic control of a heterogeneous population and application to electricity pricing.
Joint work with Wim van Ackooij, Clémence Alasseur and Stéphane Gaubert
2022 IEEE 61st Conference on Decision and Control (CDC), pp. 3617-3624
See Chapter 5.

[Jac+23b] A Quantization Procedure for Nonlinear Pricing with an Application to Electricity Markets.
Joint work with Wim van Ackooij, Clémence Alasseur and Stéphane Gaubert
Preprint arXiv:2303.17394 – To appear in the proceedings of *IEEE CDC 2023*
See Chapter 6.

[Ala+22] A Rank-Based Reward between a Principal and a Field of Agents: Application to Energy Savings.
Joint work with Clémence Alasseur, Erhan Bayraktar and Roxana Dumitrescu
Preprint arXiv:2209.03588 – Submitted to *Operations Research*
See Chapter 7.

[JZ23] Tight Bound for Sum of Heterogeneous Random Variables: Application to Chance Constrained Programming .
Joint work with Riadh Zorgati
Preprint arXiv:2211.12275 – Submitted to *Journal of Optimization Theory and Applications*
See Chapter 8.

Chapter 3

Prolegomena

***

In this preliminary chapter, we introduce three main theories – bilevel programming, Principal-Agent models and control of Markov processes – related to the study of Stackelberg games, and by which is underpinned the core of this PhD thesis. For each of the approaches, we give a simple example on the product pricing problem, which is the basic problem underlying all the following chapters.

3.1 Bilevel programming

The special structure of pricing problems can be generally cast into the *bilevel* framework, see [Bar13] and [Dem+15]. In this setting, a *leader* (here the company) aims at optimizing its own objective (*upper level*), integrating into the constraints the decision of the *follower* (here the consumers), viewed as the solution of an inner optimization problem (*lower level*). We denote by $x \in \mathbb{R}^M$ the set of variables controlled by the leader (upper variables), and by $y \in \mathbb{R}^N$ the set of variables controlled by the follower (lower variables). The two players have distinct objectives:

- $F : (\mathbb{R}^M \times \mathbb{R}^N) \to \mathbb{R}$ the *leader*'s objective function (modeling here a profit to maximize)
- $f : (\mathbb{R}^M \times \mathbb{R}^N) \to \mathbb{R}$ the *follower*'s objective function (modeling here a cost to minimize).

The strategy chosen by the leader is supposed to belong to a subset $\mathcal{X} \subseteq \mathbb{R}^M$ (if the upper constraints are linear, then $\mathcal{X}$ is a polyhedron). For a leader strategy $x \in \mathcal{X}$, the follower minimizes f by finding an optimal strategy $y^* \in \Psi(x)$, where the set of optimal lower decisions $\Psi(x)$ is defined as

$$\Psi(x) := \operatorname*{arg\,min}_{y \in \mathcal{Y}(x)} f(x, y) \quad . \tag{3.1}$$

In (3.1), the subset $\mathcal{Y}(x) \subseteq \mathbb{R}^N$ represents the feasible values of y for a given upper decision x^*, taking into account the constraints of the follower, and is called *feasible set mapping.* We consider here for sake of simplicity that $\mathcal{Y}(x)$ is a convex subset of the form:

$$\mathcal{Y}(x) := \{y \in \mathcal{Y} \mid g(x, y) \leqslant 0\} \quad , \tag{3.2}$$

where $g : \mathbb{R}^M \times \mathbb{R}^N \to \mathbb{R}^P$ and $y \mapsto g_k(x, y)$ is convex for any $1 \leqslant k \leqslant P$ and $x \in \mathbb{R}^M$.

The set of optimal strategies $\Psi(x)$ may not be reduced to a singleton, i.e, there may exist several distinct lower decisions which appear to the follower as equivalent. However, these decisions are not equivalent in general for the leader. To overcome this issue, we can develop an *optimistic* model where the follower acts cooperatively by choosing $y_o^*(x)$ such that $y_o^*(x) = \arg\max_{y \in \Psi(x)} F(x, y)$ or a *pessimistic* model where the follower realizes the worst case for the leader by choosing $y_p^*(x) = \arg\min_{y \in \Psi(x)} F(x, y)$.

Remark 3.1.1

The optimistic case is often easier to solve than pessimistic model, see e.g. [Dem02]. Wiesemann et al. [Wie+13] presented conditions guaranteeing the existence of optimal solutions for pessimistic bilevel problems, and developed iterative solution scheme adapted to this class of instances.

The general optimistic bilevel problem can be written

$$\max_{x \in \mathcal{X}, y^*} \left\{ F(x, y^*) \quad \text{s.t.} \quad y^* \in \Psi(x) = \mathop{\arg\min}_{y \in \mathcal{Y},\, g(x,y) \leqslant 0} f(x, y) \right\} \,. \tag{3.3}$$

Notation 3.1.1

In Stackelberg literature, the leader is denoted by “she” and the follower by “he”. We will adopt this notation for the sequel. Besides, in graphic depiction, the first player will be represented in red, and the second player in green, see Figure 1.10 to Figure 3.2.

Problems of the form Equation (3.3) are NP-Hard in general (even when the leader’s and the follower’s objective functions are linear), and if we deal with mixed-integer programs, bilevel problems can be Σ_2^P-Hard, see [Jer85]. However, several methods have been introduced in order to solve them. The major part of the literature aims at reformulating bilevel instances as *single-level* problems. We present hereafter two techniques:

Classical KKT reformulation. This first approach consists in replacing the lower level by the Karush-Kuhn-Tucker (KKT) optimality conditions when it is suitable : let us assume that

(i) $\forall x \in \mathbb{R}^M$, $y \mapsto f(x, y)$ is convex,

(ii) and for each $x \in \mathcal{X}$, there exists $\hat{y} \in \mathbb{R}^N$ such that $g(x, \hat{y}) < 0$ (Slater condition).

Then, the set of *global* solutions of (3.3) coincides with the one of

$$\begin{aligned} \max_{x \in \mathcal{X}, y \in \mathcal{Y}} \; & F(x, y) \\ \text{s.t. } & \partial_y f(x, y) + \lambda^\top \partial_y g(x, y) = 0 \\ & g(x, y) \leqslant 0 \\ & \lambda_k g_k(x, y) = 0, \ 1 \leqslant k \leqslant P \\ & \lambda \in \mathbb{R}^P_{\geqslant 0} \end{aligned} \tag{3.4}$$

Dempe shows that the equivalence between single-level and bilevel formulation doesn’t hold for local solutions: indeed, a local optimum of the single-level reformulation need not be a local optimal solution to the bilevel programming problem, see [DD12, Example 3.1]. Methods searching for stationary points require certain regularity [FL04b; Dus+17].

The general difficulty in the solving of (3.4) lies in the complementarity constraints $\lambda_k g_k(x, y) = 0$, as they reveal a combinatorial nature behind bilevel problems. Various mixed-integer methods have been used, and we refer to the recent survey of Kleinert et al. [Kle+21] for an overview on these techniques, among which Branch-and-Bound methods can be used to treat this difficulty for medium-size instances [FM81; MB90]. Alternatively, nonlinear methods are often used to smooth the complementarity constraints using penalization techniques, see e.g. [LM10; FL04a; Dus+20].

Optimal-value reformulation. We define the *optimal value function* of the follower $\varphi : \mathbb{R}^M \to \mathbb{R}$ as

$$\varphi(x) := \min_{y \in \mathcal{Y},\, g(x,y) \leqslant 0} f(x, y) \ . \tag{3.5}$$

The optimistic bilevel problem can, then, be recast as a one-level instance :

$$\begin{aligned} \max_{x \in \mathcal{X}, y \in \mathcal{Y}} \ & F(x, y) \\ \text{s.t. } & f(x, y) \leqslant \varphi(x) \\ & g(x, y) \leqslant 0 \ . \end{aligned} \tag{3.6}$$

This alternative single-level reformulation to the KKT reformulation was initiated by Outrata [Out90], and is used in branching algorithms [DK16; Fis+17]: in a schematic view, the constraint “$f(x, y) \leqslant \varphi(x)$” in Equation (3.6) is hard to handle (it is in general nonconvex, and even if it is convex, it can not be qualified). Therefore, these methods first ignore this constraint (*High-Point Relaxation* [Bar13]) and then iteratively add cuts to take into account the optimality of follower decision. Alternatively, in [AO22], the authors enhanced DC (Difference of Convex functions) algorithms by defining trial points as inexact solutions, reducing computational burden.

A very studied case is the *product pricing problem* [STM11; Dew+08; Fer+16], with applications to toll pricing [LMS98], freight pricing [Bro+00] or tickets selling [Hoh20] :

$$\begin{aligned} \max_{x, y^*} \ & \sum_{i=1}^{M} \sum_{j=1}^{N} x_j y^*_{ij} \\ \text{s.t. } & y^* \in \arg\min_{y} \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} (x_j - r_{ij}) y_{ij} \ \text{ s.t. } \ \sum_{j=1}^{N} y_{ij} \leqslant 1,\ y_{ij} \geqslant 0,\ \forall i, j \right\} \ . \end{aligned} \tag{3.7}$$

In (3.7), each customer $1 \leqslant i \leqslant M$ aims at choosing among $N (\ll M)$ products the one maximizing the quantity $r_{ij} - x_j$, which can be interpreted as the utility of product j (r_{ij} is a reservation utility). If none of the products $1 \leqslant j \leqslant N$ gives a positive utility for customer i, then the customer will not purchase any product and $y_{ij} = 0$ for all j. In this model, even if several customers are reacting to the strategy of the leader, each customer reacts independently as there is not any capacity constraint on the number of available products (*Envy-free* problems, see e.g [STH07]). Therefore, the agents can be merged into one agent making a global decision for each individual customer.

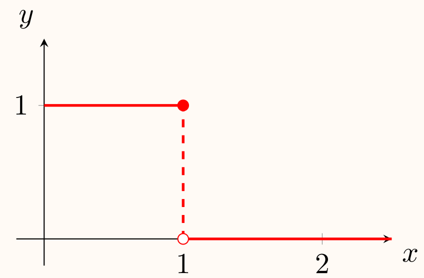
The formulation (3.7) has a specific structure: the two objective functions are bilinear. As a consequence, the lower problem is linear when the decision of the leader is taken. Moreover, the lower level is somehow simple, in the sense that we only require for the follower decision to belong to the simplex $\{y_{i,\cdot} \in \mathbb{R}^N_{\geqslant 0} \mid \sum_j y_{ij} \leqslant 1\}$. Depending on the application, unitary selling costs c_{ij} are sometimes added leading to an upper objective $\max_x \sum_{i=1}^{M} \sum_{j=1}^{N} (x_j - c_{ij}) y^*_{ij}$.

Example 3.1.1 (Product pricing problem)

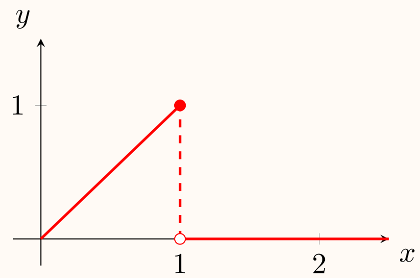
Let us consider the simplest case where there is only one customer ($M = 1$) and a single product to price ($N = 1$), with reservation utility $r = 1$. Then, the problem simplifies to

$$\max_{x\in\mathbb{R}} xy^* \text{ s.t. } y^* \in \arg\min_{0\leqslant y\leqslant 1}\{(x-1)y\} \tag{3.8}$$

In this toy example, the set of optimal lower solution $\Psi(x)$ is always a singleton, except when $x = r = 1$, which corresponds to the case where the utility of the product is exactly zero: at this price, every lower response is optimal, i.e., $\Psi(1) = [0, 1]$. An optimistic lower response would be $y = 1$ as it maximizes the leader objective for $x = 1$, whereas a pessimistic solution would be $y = 0$ to disadvantage the leader. A huge difference between the two approaches is that the pessimistic solution has in general no guarantee to be attained [Wie+13; Dem02]: in this example, the optimistic solution is $x = 1$, leading to a revenue for the leader of $F(1, 1) = 1$, whereas the pessimistic solution is not attained (the objective value is also 1, but only obtained at the limit case $x \to 1$).



(a) Optimal optimistic follower response



(b) Leader objective function whithin the optimal optimistic lower response

3.2 Reverse Stackelberg games: a bridge to the theory of incentives

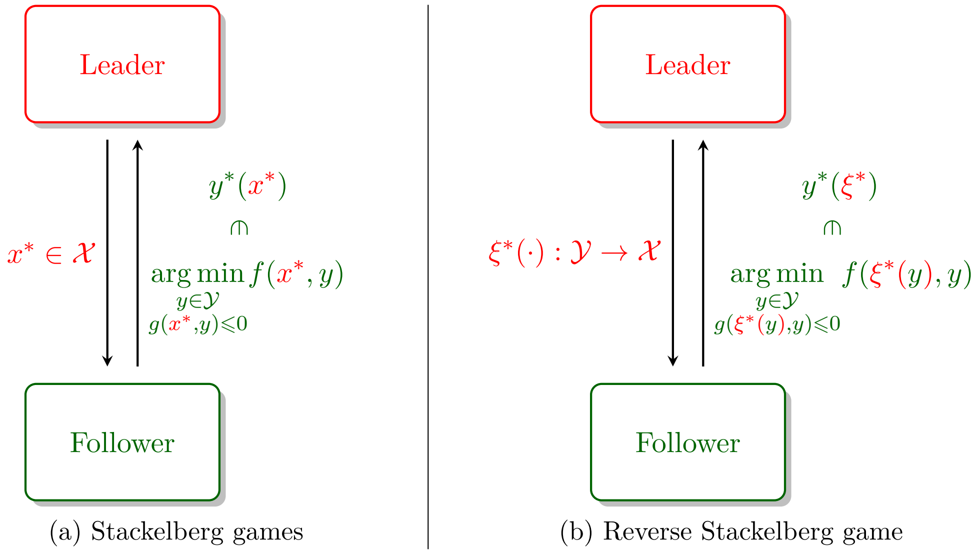
In the previous section, the follower decides his strategy by knowing the upper decision $x^* \in \mathcal{X}$. In the pricing context, this means that the follower reacts to a price signal. This is realistic in many situations, where the piece of information that the leader has agreed to provide is only a price. For example, in the retail electricity market, the offers are only based on price coefficients in € or €/kWh.

Let us imagine now that the leader aims at incentivising the follower to reach a target, for example a consumption reduction. To this purpose, it is more natural (and more powerful) to provide to the follower a *reward function* instead of a single price, describing for each level of follower effort (lower decision) the corresponding price/reward that can be obtained.

Figure 3.1 shows the difference between standard (without incentives) Stackelberg games and reverse Stackelberg games. Instead of a single price vector $x^* \in \mathcal{X}$, the leader now provides a function $\xi^* : \mathcal{Y} \to \mathcal{X}$. This function can be viewed as the mirror of the lower response map $y^*(\cdot)$, hence the named of *reverse* game. The general optimistic reverse Stackelberg game can be written as the following bilevel instance:

$$\max_{\xi(\cdot):\mathcal{Y}\to\mathcal{X},\, y^*} \left\{ F(\xi(y^*), y^*) \quad \text{s.t.} \quad y^* \in \Psi(x) = \mathop{\arg\min}_{y\in\mathcal{Y},\, g(\xi(y),y)\leqslant 0} f(\xi(y), y) \right\} \tag{3.9}$$

Concepts and basis framework can be found in [GSH12], and we refer to [BLS22] for recent



(a) Stackelberg games

(b) Reverse Stackelberg game

Figure 3.1: Incentived schemes

A reverse Stackelberg game is a Stackelberg game where we allow the leader decision to belong to an infinite-dimensional space.

advances on application to grid pricing.

Example 3.2.1

To illustrate that reverse Stackelberg games are, in some cases, a way to incentivise the follower to better act in the leader sense, we recall the example of [Ols09]: we consider the 1-dimensional unconstrained bilevel problem with

$$F(x, y) = -(y-5)^2 - x^2 \quad , f(x, y) \quad = x^2 + y^2 - xy \quad .$$

Then, the leader wants to attain the point $(x, y) = (0, 5)$ which gives a zero objective value. However, in the standard Stackelberg game, the follower response to $x = 0$ is $y = 0$. If now the leader provides to the follower the function $\xi^*(y) = 2y - 10$, then the follower response will be

$$y^*(\xi^*) \in \arg\min_{y} \left\{ (2y-10)^2 + y^2 - (2y-10)y \right\} = \{5\} \quad .$$

The value of the incentive is then $x^* = \xi^*(y^*) = 0$, leading to a zero objective value.

3.3 Principal-Agent problems

The theory of incentives (or contract theory) was initially introduced in Economics [LM09] and developed in the seventies in particular with the various works of Mirrlees (see e.g. [Mir76]). This theory analyses the relation between two asymmetric players, respectively named in this context *Principal* and *Agent* (equivalent of leader and follower in bilevel).

Compared with reverse Stackelberg games, Principal-Agent problems also focus on the optimal design of an incentive/contract, but often take into account uncertainty due to unobservable characteristics for the principal.

Presentation in discrete time. Let $(\Omega, \mathcal{F})$ be a measurable space, and Y be a real-valued random variable, representing the observable output process. In this context, the Principal sets up a contract $\xi : \mathbb{R} \to \mathbb{R}$ (Borel measurable map) at time 0 with the agent, and the agent is payed $\xi(Y)$ at time 1.

Given a contract, the agent chooses an effort $a \in \mathbb{R}$ at time 0, generating a probability measure $\mathbb{P}^a$ on $(\Omega, \mathcal{F})$ such that under $\mathbb{P}^a$, Y has a continuous distribution with density $f(\cdot, a)$ (the effort is the way for the agent to influence/control the stochastic output process Y). The agent has, then, a utility $V_0^A(\xi, a)$ for each contract ξ and effort a, that he wants to maximize:

$$V_0^A(\xi) = \sup_{a \in \mathbb{R}} V_0^A(\xi, a) \ \ .$$

For instance, a typical objective function considered in the literature (see e.g. the seminal paper of Holmström and Milgrom [HM91]) is defined through an exponential function called *Constant Absolute Risk Aversion* (CARA) :

$$V_0^A(\xi, a) = \mathbb{E}^{\mathbb{P}^a}\left[-e^{-\theta_A(\xi - c(a))}\right] \ \ ,$$

where c is the cost of effort for the agent and $\theta_A > 0$ is the risk aversion intensity.

In the same way, the Principal aims at maximizing a utility function $V_0^P(\xi, a)$. To this purpose, several models coexist, and can be divided into three main categories, depending on the available quantity of information:

(i) *Full information / First-best*: the principal is able both to design the contract and to control the action of the agent, then the provider optimization problem is

$$V_0^{P,FB} = \sup_{\xi, a} \left\{ V_0^P(\xi, a) \text{ s.t. } V_0^A(\xi, a) \geqslant R \right\} \ \ , \tag{3.10}$$

where the constraint (called *Individual rationality*, or IR in short) ensures that the utility of the follower is greater than a reservation utility. This class of problems are called *first-best* as they do not consider the optimal behavior of the agent, i.e., that a must belong to the set of maximizers of $a \mapsto V_0^A(\xi, a)$. This relaxed case is a standard single-level problem, simpler to solve. As an example, one can obtain using Borch' rule [Bor62] that if V_0^P is also define through a CARA utility, i.e., $V_0^P(\xi, a) := \mathbb{E}^{\mathbb{P}^a}\left[-e^{-\theta_P \xi}\right]$, then the optimal contract is affine: $\xi(Y) = c + \frac{\theta_P}{\theta_A + \theta_P} Y$, see e.g. [MR18].

(ii) *Moral-hazard / Second-best*: the principal is able to observe the output process Y, but cannot directly influence the action, optimally chosen by the agent. In this case, the principal problem is

$$\begin{aligned} V_0^{P,SB} &= \sup_{\xi, a^*} V_0^P(\xi, a^*) \\ &\text{s.t. } a^* \in \arg\max_{a \in \mathbb{R}} V_0^A(\xi, a) \\ &\qquad V_0^A(\xi, a^*) \geqslant R \end{aligned} \tag{3.11}$$

The additional constraint (called *incentive-compatibility* condition, or IC in short) expresses the optimality of the agent choice, the latter player maximizing his own utility function. Problem (3.11) is in fact the **stochastic analog** of the reverse Stackelberg game introduced in (3.9), and corresponds to a bilevel problem. The question of existence was initially tackled by Page [Pag87], and the first widespread method was the *First-order approximation* (FOA) where the IC condition is replaced by the weaker first-order condition $\partial_a V_0^A(\xi, a) = 0$, see [Rog85]. This idea ties up with the classical KKT transformation: in both cases, single-level reformulations are obtained exploiting the optimality conditions of the inner problem (note that the first-order approximation does not guarantee equivalence with the two-level problem, as pointed out by Mirrlees [Mir99]).

(iii) *Adverse selection / Third-best*: in this situation, the agent has a type (characteristic) which is unknown by the principal. This uncertainty is *ex ante* the design of the reward (whereas the moral hazard is uncertainty *ex post* the establishment of the contract), see e.g. [Sal05] for the general theory. In practice, in models under adverse selection (we also speak about models with *screening*), the distribution of types within the population is supposed to be known by the Principal.

Principal-Agent problems under adverse selection include in particular the so-called *monopolist* problem [MR78; GL84; RC98]:

$$\begin{array}{ll} \max\limits_{\xi\colon \mathcal{Y}\to\mathcal{X}} & \int_\Omega \left[\xi(y^*(\tau)) - C(\xi(y^*(\tau)))\right]\rho(\tau)d\tau \\ \text{s.t. } \forall\tau\in\mathcal{T} & y^*(\tau)\in\arg\max\limits_{y}\{\tau y - \xi(y)\} \\ & u(\tau) := \tau y^*(\tau) - \xi(y^*(\tau)) \geqslant R(\tau) \end{array} . \tag{3.12}$$

In (3.12), each customer of type $\tau \in \mathcal{T}$ aims at choosing among the set of goods $\mathcal{Y} \subseteq \mathbb{R}$ the one maximizing the quantity $\tau y - \xi(y)$, which can be interpreted as the utility of good y. The price of a good y is determined by the principal and corresponds to $\xi(y)$. The upper objective is then the maximization of the mean profit (difference between the price and the cost of the good). Here, contrary to the product pricing problem in (3.7), there is a continuum of types and a continuum of goods. Moreover, all the consumers choose a product of the principal, as we ensure that the utility of each customer is greater than the reservation utility (hence the name of monopolist problem). Equivalently, the contract menu offered by the principal can be written as a pair of functions $\tau\in\mathcal{T}\mapsto(\hat{\xi}(\tau),\hat{y}(\tau))$, see e.g.[CD17]:

$$\begin{array}{ll} \max\limits_{(\hat{\xi},\hat{y})\colon \mathcal{T}\to\mathcal{X}\times\mathcal{Y}} & \int_{\mathcal{T}} \left[\hat{\xi}(\tau) - C(\hat{\xi}(\tau))\right]\rho(\tau)d\tau \\ \text{s.t. } \forall\tau\in\mathcal{T} & \tau\hat{y}(\tau) - \hat{\xi}(\tau) \geqslant \tau\hat{y}(\tau') - \hat{\xi}(\tau'),\ \tau'\in\mathcal{T} \\ & u(\tau)\geqslant R(\tau) \end{array} . \tag{3.13}$$

The fist constraint, so-called *incentive-compatibility* condition, ensures that the optimal choice $y^*(\tau)$ for customers of type τ under incentive $\hat{\xi}$ effectively corresponds to the good $\hat{y}(\tau)$.

Remark 3.3.1

The choice of the customer can be decided in advance by the principal so that a customer of type τ will necessarily choose the good $\hat{y}(\tau)$. This is possible since we can dedicate one specific good for each type of customers. Therefore, in contrast with the bilevel product pricing problem (3.7), there is no longer any combinatorial issue, i.e., the choice of good j is not encoded anymore with a binary variable y_{ij} but is now replaced by the decision $\hat{y}(\tau)$.

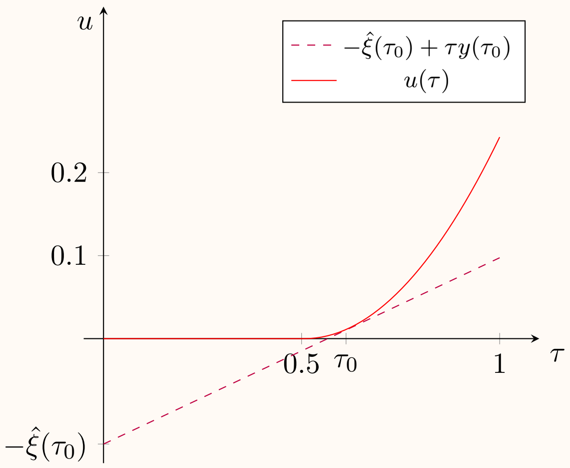
In [RC98], Rochet and Choné reformulate (3.12) as a variational problem under convexity constraints as follows:

$$\min_{u\colon\mathcal{T}\to\mathcal{Y}} \left\{ \int_{\mathcal{T}} \left[u - \tau\nabla u(\tau) - C(\nabla u(\tau))\right]\rho(\tau)d\tau \;\middle|\; \begin{array}{l} u \text{ convex} \\ u(\tau)\geqslant R(\tau) \end{array} \right\} . \tag{3.14}$$

From the optimal solution u, the pair $(\hat{\xi},\hat{y})$ can be recovered as $\hat{y} = \nabla u$ and $\hat{\xi}(\tau) = \tau\hat{y}(\tau) - u(\tau)$. Conditions for the existence of a solution to the latter problem have been tackled in [Car01] and Carlier and Zhang [CZ20] extended these results to partial participation (relaxing the constraint $u \geqslant R$). When a solution exists, the optimal (nonlinear) contract is a function that depends on the type of the consumer, revealed after the signing of the contract. Rochet and Choné [RC98] show that there is in general *bunching* in the solution, i.e., agents with different types are treated identically (same contract) in the optimal solution.

Example 3.3.1 (Product pricing as variational problem)

Let us consider a continuum of customers uniformly distributed according to their types $\tau \in [0,1]$. The reservation utility is set to 0 and the cost function to $C(y) = 1/2y^2$. We display below the optimal solution of the variational problem in this simple setting (this is done here using a discretization as in [EM09]).



(a) Optimal utility function for Problem (3.14).

We observe that the utility is constant and equal to 0 for $\tau \leqslant 0.5$. This means that every customer of type $\tau \leqslant 0.5$ chooses good $y = 0$ with price $\hat{\xi} = 0$. This illustrates that bunching appears on subsets where the utility function is affine. For $\tau > 0.5$, the good is different for each type and has an increasing price. We represent the tangent plane to u at $\tau = \tau_0$ showing that the price and the good can be recovered from the utility function.

Time-dependent models. Suppose now that the process Y is a time-dependent process, and that the action taken by the agent also evolves in time: for simplicity, let us suppose that the process Y_t is given by the following controlled state equation:

$$Y_t = \int_0^t a_s ds + \int_0^t \sigma \mathrm{d}W_s,\ t \in [0,T]\ ,$$

where the first integral is the drift, controlled by the action process a, and σ is the volatility (intensity of the Brownian motion). The principal-agent problem with the latter setting has been first studied by Sannikov [San08]. The standard steps that should be followed in order to solve the time-dependent moral-hazard principal agent problem can be summarized as follows (see e.g. [CPT17]):

1. Restrict the search space for the leader by finding a subset of contracts (with a fixed form) for which we are able to compute analytically the best response of the agent. These contracts are called *revealing* contracts.
2. Then, it has to be proved that an optimal contract belongs to the determined subset.
3. Finally, it remains to solve the (easier) problem defined on the subset of revealing contracts.

Multi-agents models with competition. In many applications, the Principal designs a contract not for a single follower but for a population of agents, distributed according to a given distribution. If each agent reacts independently (as in Envy-free problems), the model corresponds to a Principal-Agent instance under adverse-selection, see e.g. the monopolist problem (3.12). If now there is interaction/competition between the agents, then a Nash equilibrium (no agent can do better by unilaterally changing their strategy) has to be found between the agents *a posteriori* the signing of the contract. When the population is of large size, a mean-field assumption is often supposed, considering that the population is a continuum of agents so that the impact of each individual agent is negligible.

Given the contract, the decision of the population corresponds to a Nash equilibrium of a mean-field games, see e.g. [HMC06; HCM07; Pie06]. These mean-field games are studied for their own sake in various application fields (for instance in energy [ABM20] or in finance [GPY13]). The embedding of these games into Principal-Agent problems has been regarded more recently by Élie, Mastrolia and Possamaï [EMP19], and also Carmona and Wang [CW21]. Ranking games [BZ16; BZ21] can be viewed as a special case, where the contract depends on the ranking within the population, so that each agent competes to obtain the best ranking.

3.4 Markov Decision Processes for the control of populations

Notation 3.4.1

We bring to the reader's attention that we still denote by $x \in \mathcal{X}$ the leader/controller action to keep the same notation as in bilevel optimization theory, but this should not be confused with standard state space notation in control theory.

A Markov chain $(Y_t)_{t\geqslant 0}$ is a stochastic sequence of events/states, where the probability that a given event happens at the next time step only depends on the current state (and not on the sequence of all past events), see e.g., [Chu60]. In the discrete-time setting and in case of a finite *state space* $\mathcal{Y} := \{1,\dots,N\}$ of N possible states, the Markov process (Y_t) is fully determined by the initial state Y_0 and the *transition matrix* $P = (\mathbb{P}[Y_{t+1} = m \mid Y_t = n])_{1\leqslant n,m\leqslant N}$.

The associated *Markov Decision Process* (or shortly MDP) is then a control problem where a controller (this would correspond to the leader in bilevel or to the principal in Principal-Agent theory) is able to influence the transition matrix by some action/control $x \in \mathcal{X}$. We refer to [Put94], [Ber12] and [Fru19] for introduction and main results on MDPs. At each time step, if the process, or *environment* (it would correspond to the follower in bilevel optimization or to the agent in Principal-Agent theory) is in state $n \in \{1,\dots,N\}$ at time t and the controller chooses action x, then the process will randomly move in state $m \in \{1,\dots,N\}$ at time $t+1$ with probability $P(x)_{n,m}$, and gives to the controller an instantaneous reward $\theta(x)_m$. The transition matrix is then indexed by the action chosen by the controller. The MDP is therefore represented by a 4-tuple $\mathcal{M} = (\mathcal{Y}, \mathcal{X}, P(\cdot), \theta(\cdot))$, where

(i) $\mathcal{X}$ is the *action/control* space,
(ii) $P(x) \in \mathbb{R}^{N\times N}$ is the *transition matrix* associated with control $x \in \mathcal{X}$,
(iii) $\theta(x) \in \mathbb{R}^N$ is the *instantaneous reward* to be in a given state due to control $x \in \mathcal{X}$.

Given an initial *probability measure* $\mu_0 \in \mathcal{P}(\mathcal{Y}) = \Delta_N := \{\nu \in \mathbb{R}^N_{\geqslant 0} \mid \sum_{i=1}^N \nu_i = 1\}$ and a sequence of actions $(x_t)_{t\geqslant 0}$, the probability measure $\mu_t = (\mathbb{P}[Y_t = n])_{1\leqslant n\leqslant N}$ satisfies the deterministic transition equation:

$$\mu_{t+1} = \mu_t P(x_t) \ .$$

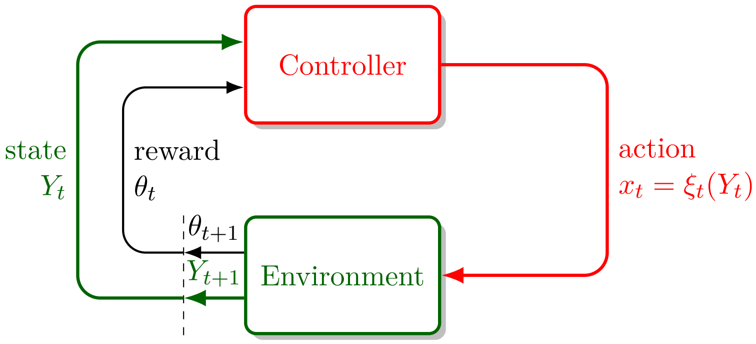


Figure 3.2: Schematic representation of Markovian Decision Process.

The dynamics is here represented as a loop, where at each iteration an action is chosen by the controller, influencing the random choice of the next state.

At each time step and given the current state, the controller chooses an action x_t according to a *decision rule* $\xi_t : \mathcal{Y} \to \mathcal{X}$. The collection of decision rules $\xi = \{\xi_t\}_t$ is then called a *policy*. *Stationary* policies $\xi = (\xi_0, \xi_0, \xi_0; \dots)$ are an important and well-studied specific case.

Three different controller objective functions are considered in the literature:

(i) *Finite-horizon gain*: $\sup_{\xi} \mathbb{E}\left[\sum_{t=1}^{T} \theta(x_t)_{Y_t} \,\middle|\, x_t = \xi_t(Y_t),\ Y_{t+1} \sim \mathbb{P}\left[\,\cdot \mid Y_t, x_t\right]\right]$.

(ii) *Discounted infinite-horizon gain*: $\sup_{\xi} \mathbb{E}\left[\sum_{t=1}^{\infty} \alpha^{t-1}\theta(x_t)_{Y_t} \,\middle|\, x_t = \xi_t(Y_t),\ Y_{t+1} \sim \mathbb{P}\left[\,\cdot \mid Y_t, x_t\right]\right]$.

(iii) *Average long-term gain*: $\sup_{\xi} \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} \theta(x_t)_{Y_t} \,\middle|\, x_t = \xi_t(Y_t),\ Y_{t+1} \sim \mathbb{P}\left[\,\cdot \mid Y_t, x_t\right]\right]$.

Finite-horizon problems can be solved by backward induction (dynamic programming) using a Bellman equation [Bel54], whereas discounted infinite-horizon problems and average long-term gain problems are typically solved by a fixed-point equation and respectively discounted and relative Value-Iteration Algorithm, see e.g. [Put94] for details on the finite action spaces and [Sch85] for compact action spaces.

Finite population and convergence to mean-field limit case

In the context of the control of a population, the environment corresponds to a finite set of I agents. When the agents are *indistinguishable* (same behavior, i.e., same transition kernel $P(\cdot)$), Gast and Gaujal [GG10] and more recently Motte and Pham [MP22] introduced so-called *I-agent Markov Decision Processes*, which are represented by a 5-tuple $(\mathcal{Y}, \mathcal{X}, P(\cdot), \theta(\cdot), I)$. As for standard MDPs, the controller chooses an action x, and each agent $i \in [I]$ is influenced by this action so that he moves from n_i to m_i with probability $P(x)_{n_i,m_i}$. The controller's reward is then $\frac{1}{I}\sum_{i\in[I]} \theta(x)_{n_i}$.

Remark 3.4.1

The I-agent MDP is equivalent to a standard MDP with state space $\mathcal{Y}^I$ and transition matrix $Q(u) = \operatorname{diag}(P(x), \dots, P(x)) \in \mathbb{R}^{N^I \times N^I}$.

As in [MP22], one can define the *lifted MDP* associated with $\mathcal{M}$ as the (deterministic) MDP $(\mathcal{P}(\mathcal{Y}), \mathcal{X}, T(\cdot), r(\cdot))$, where

- ⋄ $\mathcal{P}(\mathcal{Y}) = \Delta_N$ is the set of *probability measures* on $\mathcal{Y}$,
- ⋄ $T(x) := [\mu \in \Delta_N \mapsto \mu P(x)]$ is the *transition function* which gives the next distribution after executing action x,
- ⋄ $r(x) := [\mu \in \Delta_N \mapsto \langle \theta(x), \mu \rangle_N]$ is the *expected* instantaneous reward according to a given measure due to action x.

Gast and Gaujal [GG10] showed that for an infinite number of *indistinguishable* players ($I \to \infty$), the I-agent MDP coincides with the lifted MDP for finite-time horizon and discounted infinite horizon and provided convergence rates. Motte and Pham [MP22] extended this result to a broader class of MDPs. In [Bäu23], Bäuerle focused on the average-gain optimality criteria and shows that mean-field limit is ϵ-optimal for the discounted problem if the number of agents is large and the discount factor close to one.

Many mean-field control problems in demand-side management have been studied since the eighties: Malhame and Chong [MC85] introduces a control problem for a population of thermostatically controlled loads (TCL) and in [KM13], Kizilkale and Carmona circumvented the difficulty of large population by mean-field approximation. Le Floch, Can Kara and Moura [FKM18] studied the control problem of a large fleet of electrical vehicles needed to be charged. Cammardella et al. [Cam+19] considered a quadratic criterion and a Kullback-Leibler penalty in order to learn both the control and the transition kernel. In [Bré+19], electricity consumers are incentivized to shift a part of the peak consumption to the off-peak period by designing dynamic pricing mechanisms through the use of multi-armed bandit techniques. Finally, Moreno et al. [Mor+23] introduces a new mirror descent approach applied to demand-side management control problem.

Let us come back to the product pricing problem, constituting our leitmotiv in this preliminary chapter. In discrete choice theory (see e.g. [Tra09]), probabilities of choice are often supposed to follow Gumbell distributions (*logit* models): still considering as in (3.7) M consumers and N products, the probability that customer $1 \leqslant i \leqslant M$ chooses product $1 \leqslant l \leqslant N$ after having chosen product $1 \leqslant n \leqslant N$ is

$$[P_i(x)]_{n,l} = \mathbb{P}[u_{in}^l \geqslant u_{in}^k,\ \forall\, 0 \leqslant k \leqslant N] \ ,$$

where the utility for customer i to select product k after product n is

$$u_{in}^k = \beta^{-1}\epsilon_{ik} + \gamma \mathbb{1}_{l=n} + \begin{cases} r_{ik} - x_k & (k \geqslant 1) \\ 0 & (k = 0) \end{cases} \quad .$$

As in the deterministic case (3.7), each customer i focuses on product n maximizing the quantity $r_{in} - x_n$, but with an additional perturbation (of intensity β^{-1}) coming from a family of independent and identically distributed Gumbell noises $\{\epsilon_{ij}\}_{ij}$. Each customer still has the possibility not to choose one of the N products (and choose option $k = 0$). The correlation between two consecutive customer choices is represented by an extra term $\gamma \mathbb{1}_{l=n}$, modeling an additional utility to maintain the same choice from one time step to another. The parameter $\gamma \geqslant 0$ is then viewed as a switching cost that appears when it comes to making a change of offer/product, see e.g. [Dub+08; DHR09; DHR10]. This dynamic extension of the product pricing problem is tackled e.g. in [PE17]. In the logit setting, the transition matrix $P_i(x)$, coming from the optimal choices of customer i, has a closed-form formula:

$$[P_i(x)]_{n,l} = e^{\beta(r_{il} - x_l) + \gamma \mathbb{1}_{l=n}} \ / \ \left(1 + \sum_{k=1}^{N} e^{\beta(r_{ik} - x_k) + \gamma \mathbb{1}_{k=n}}\right) \quad .$$

Example 3.4.1 (Product pricing as a deterministic MDP)

As in Example 3.1.1, let us consider the case where $M = 1$, $N = 1$ and $r = 1$. Then, the finite-horizon problem simplifies to

$$\max_{x_1,\dots,x_T \in \mathcal{X}^T} \left\{ \sum_{t=1}^{T} x_t \mu_t \text{ s.t. } \mu_{t+1} = \frac{\mu_t}{1 + e^{-\gamma - \beta(1-x_t)}} + \frac{1-\mu_t}{1 + e^{\gamma - \beta(1-x_t)}} \right\}. \tag{3.15}$$

Compared with the (non-dynamic) case in (3.8), the lower decision is here of probabilistic nature and is unique (and explicit) in the logit setting, so that there is no ill-posedness issue. Given a price x, the deterministic (binary) response y^* can be viewed as the limit of μ_t when $\beta \to \infty$ and $\gamma \to 0$.

Despite its apparent simplicity, the problem (5.25) is already challenging: backward induction approaches imply to solve transition problems (in the computation of Bellman operator [Bel54]) which are here non concave (in x).

Chapter 4

Quadratic Regularization of Bilevel Pricing Problems and Application to Electricity Retail Markets

***

This chapter is based on the published paper [Jac+23a], to which we add an extended study of the deterministic and logit models (Section 4.2). We also provide in appendix (see 4.8.3) a presentation of the numerical tool that embedded the different resolution methods we present in the chapter.

Abstract. We consider the profit-maximization problem solved by an electricity retailer who aims at designing a menu of contracts. This is an extension of the unit-demand envy-free pricing problem: customers aim to choose a contract maximizing their utility based on a reservation bill and multiple price coefficients (attributes). A basic approach supposes that the customers have deterministic utilities; then, the response of each customer is highly sensitive to price since it concentrates on the best offer. A second classical approach is to consider logit model to add a probabilistic behavior in the customers' choices. To circumvent the intrinsic instability of the former and the resolution difficulties of the latter, we introduce a quadratically regularized model of customer's response, which leads to a quadratic program under complementarity constraints (QPCC). This allows to robustify the deterministic model, while keeping a strong geometrical structure. In particular, we show that the customer's response is governed by a polyhedral complex, in which every polyhedral cell determines a set of contracts which is effectively chosen. Moreover, the deterministic model is recovered as a limit case of the regularized one. We exploit these geometrical properties to develop a pivoting heuristic, which we compare with implicit or non-linear methods from bilevel programming, showing the effectiveness of the approach.

Throughout the chapter, the electricity retailer problem is our guideline, and we present a numerical study on this application case.

4.1 Introduction

4.1.1 Context

For a company, the question of determining the correct prices of its products is crucial: a compromise has to be found between having enough consumers buying products and setting prices that are sufficiently important to cover the production cost. Profit-maximization models have been extensively studied. They consist in maximizing the seller profit taking in account the customer behavior. The special structure of these problems can be generally cast into the bilevel framework, see [Bar13] and [Dem+15]. In this setting, a *leader* (here the company) aims at optimizing its own objective (*upper level*), taking into account the decision of the *follower* (here the consumers), obtained as the solution of an inner optimization problem (*lower level*). This 2-player problem is known in game theory as a Stackelberg game, see [Sta52], and reflects the asymmetry of the players' roles: the leader moves first, then the follower replies (sequential games). As detailed in [Kle+21], two classical approaches consist in reformulating the problem as a single-level one, either using strong-duality or the KKT conditions, to express the optimality of the lower decision and constrain the upper problem. Formulations based on KKT conditions lead to Mathematical Programs with Complementarity Constraints (MPCC), a class of optimization problems whose interest has been growing in recent years, and particularly in the energy sector, see [Afş+16; Ale+19; Aus+20; ARR21].

The *unit-demand envy-free* pricing problem is a specific case. We consider a finite number of customers (or *segments* of customers) who are supposed to buy precisely one product, among the ones maximizing their utility. Moreover, products are available in unlimited supply. [Gur+05] showed that this problem is APX-hard (even on a restricted class of instances). [STM11] developed Mixed-Integer Programming (MIP) formulations, along with valid cuts and heuristics. They also enhance the model to ensure that each customer faces a unique maximum utility. [Fer+16] compare several MIP formulations and reinforce them with new valid cuts. All these approaches are based on deterministic models of customer's response. By their deterministic nature, they lead to instability features: the customer's response is discontinuous, resulting in typical "sawtooth" shaped profit functions, see e.g. [LMS98; GMS15] or Figure 4.3 below.

There are situations in which revenue management data are uncertain, and as noted in [Tun08], it is desirable that "optimal or near-optimal prices delivered by the optimization techniques [be] robust under modest perturbations of the reservation prices of the potential customers and the competitors' prices". This question of uncertain data and/or uncertain decision of the followers is nowadays a central question in the bilevel community, see e.g., the recent survey [BLS23]. To overcome instability issues, one approach is to consider choice models of a probabilistic nature. Then, the value of the lower level objective determines the probability distribution of the customer's choice. The most studied case concerns the *logit* model, see [McF74; Tra09]. [LH11] suppose that the population is homogeneous, meaning that there is only one segment. They reformulate the problem as a concave maximization problem by a *market-share* transformation. [SK20] extend this approach to the case of multiple price attributes. Logit pricing models with multiple consumers segments have only been studied very recently: [Li+19] formalize the pricing problem under the Mixed Multinomial Logit (MMNL), and develop algorithms to find good solutions. [Hoh20] applies such models to the revenue management case study of the German long-distance railway network. However, logit-based models are in general hard to solve with guarantees of optimality, owing to their nonlinear and nonconvex nature.

4.1.2 Contribution

We consider a *multi-attribute* version of the *unit-demand envy-free pricing* problem that we model by a bilinear bilevel formulation. This applies in particular to the pricing of electricity offers, which is our driving case study.

Our main contribution is the development of a new model, based on a *quadratic regularization* of customer's response: it has the same benefits as the logit-based models in terms of realism and robustness, whereas its quadratic nature allows one to apply efficient algorithms based on polyhedral geometry.

First, we give a closed-form expression of the lower response and highlight its polyhedral structure (Theorem 4.3.1). This shows in particular that in the presence of near ties (contracts with similar utilities), customer's response distributes among the best contracts, rather than concentrating on a single one. The regularization parameter measures the "rationality" of the customer, in particular, the deterministic response is recovered as a limit case – with perfectly rational customers (Theorem 4.3.2). More precisely, we show that in the regularized model, the response is governed by a polyhedral complex, in which each open cell determines a set of contracts which are effectively chosen.

Besides, we show that this model has the same good theoretical properties as the logit model (stability) and provide metric estimates showing that the responses of the two models are close (Section 4.8.2). The main interest of quadratic regularization, then, lies in computational tractability. We show that the regularized bilevel model reduces to a convex Quadratic Program with Complementary Constraints (QPCC). Powerful methods based on mixed or semidefinite programming allow one to solve instances of significant size of QPCC with optimality guarantees, although problems of this kind are generally difficult. In fact, we show in Section 4.8.1 that solving the present quadratic model is APX-Hard, by reusing the transformation introduced for the deterministic case in [Gur+05]. We develop in Section 4.4 an efficient local search method, QSPC (Quadratic Search on the Price Complex), exploiting the polyhedral structure of the customer's response.

Finally, we consider realistic instances arising from French electricity markets, and analyze the optimal solution in both deterministic and quadratic cases. In particular, we look at the customers' distribution to illustrate the influence of a regularized lower level. A performance comparison between the proposed algorithm and other methods from the literature is also given in Section 4.5.

Our study is inspired by several works. We adopt the viewpoint of [GMS15] in that we consider the MMNL model [Li+19; Hoh20] as a regularized version of its deterministic analog [STM11; Fer+16]. They look at a related problem that studies the toll pricing optimization, and demonstrate, among other things, asymptotic convergence of the logit regularization to the deterministic model. Besides, [STH07] introduced several probabilistic choice models as alternatives to the logit approach, and developed convex mixed-integer formulations to solve them. In particular, they considered a model which depends on the surplus of the products, and we design a new customer's response that satisfies this assumption. By comparison with all these works, the main novelty is the introduction of the quadratic regularized model as a new probabilistic customer's response and the evidences that it has the same good features as the logit model, in terms of economic realism and robustness, while being computationally more tractable. [DB01] also investigate quadratic regularization on bilinear bilevel problems and develop bundle trust region algorithm to solve them. We differ from their work by specializing the lower level to be defined on the simplex, and by describing the customers' choices as a polyhedral complex. This interpretation is inspired by the study of [BK19], who showed that for deterministic models, agent's response can be represented by a polyhedral complex, a tropical hypersurface. This tropical complex is recovered as a limit case of the present polyhedral complex when the regularization term vanishes. Finally, to overcome the ill-posedness of bilevel problems (coming from the non-uniqueness of the followers response), primal-dual techniques have been introduced: [SSC17] directly regularized the KKT conditions of the lower problem whereas [Lun+20] inserted a regularization term into the objective of the dual lower problem. Both approximation schemes are proved to converge when the regularization term is driven to zero. In comparison, we perform here a primal regularization so that the (unique) follower decision is always primal feasible and corresponds to a modified customer behavior, interpreted as a probabilistic choice.

The chapter is organized as follows. In Section 4.2, we present the deterministic multi-attribute unit-demand envy-free pricing problem and establish basic properties of the model (optimality of integer low-level solutions, reformulation as a single level problem using the KKT conditions). For comparison, we also recall the definition of the logit-based model. In Section 4.3, we introduce the quadratically regularized model, in particular, we describe the geometric properties of customer's response, and provide a reformulation as a single level QPCC. In Section 4.4, we develop the local search method (QSPC), exploiting the polyhedral structure of customer's response. In Section 4.6, we provide a numerical analysis on instances from the electricity pricing problem.

4.2 Preliminaries

4.2.1 Notation

In the sequel, we denote by Δ_N the simplex of $\mathbb{R}^N$, and by $\|x\|_N$ the Euclidean norm associated with the canonical scalar product $\langle x, y\rangle_N$ on $\mathbb{R}^N$. For any polyhedron Q, $\mathrm{Vert}(Q)$ denotes the set of vertices of Q. Moreover, for any optimization problem (P), the value $v(P) \in \mathbb{R} \cup \{\pm\infty\}$ denotes its optimal value (that can be infinite if (P) is infeasible or unbounded).

4.2.2 Deterministic model

We suppose that a company has W different types of contracts and that a market study has distinguished beforehand S customer segments, each of them gathering consumers that have approximately the same behavior. Given a segment $s \in [S] := \{1, \dots, S\}$ and a product $w \in [W]$, the *reservation bill* R_{sw} is the maximum bill that customers of this segment are willing to pay on w. In the classical product pricing model, the items to sell are only characterized by a price (determined by the company) and each customer faces the same price. In our setting, we consider the *multi-attribute* case where the bill of each contract w is determined by a finite number $H > 1$ of variables (or attributes), denoted by x_w^h. For instance, in the French electricity market, the invoice of a customer depends on at least two variables, representing a fixed and a variable component, the former depending on the subscribed power of the customer and the latter depending on his electricity consumption, see [CRE04]. Moreover, in the peak/off-peak contract, the variable component distinguishes between the peak and off-peak consumption. Then, the invoice is determined by at least three variables. The following assumption captures such contracts.

Assumption 4.2.1. The bill $\theta_{sw}(x)$ paid by segment s for contract w is a linear form:

$$\theta_{sw}(x) := \langle E_{sw}, x_w\rangle_H \quad , \tag{4.1}$$

where $E_{sw} = (E_{sw}^h)_{h\in H} \in \mathbb{R}^H_{\geqslant 0}$. Besides, the price coefficients x_w^h are supposed to be in a non-empty polytope $X \subset \mathbb{R}^{W\times H}$.

In the electricity market context, E_{sw} represents the electricity consumption of the customers of segment s who choose the contract w. It depends on h (the period of the day) and on the contract type w. This is realistic, since the notion of peak and off-peak period can vary along the contracts, and since customers adapt their electricity consumption depending on their choice of contract. Note that, in this model, the consumption does not depend on the price. This (strong) simplification is justified by a high inelasticity of the electricity demand in the short run, see e.g., [Cse20]. Hence, this model constitutes a first-order model, and aims at focusing on the uncertainty of the decision, an active research field in bilevel programming [BLS23]. Here, we are not looking at long-term policies, but we focus on finding the best price policy to a given set of competitors' offers at a given time. The situation in which the bill $\theta_{sw}(x)$ is *affine* in the energy consumption, instead of being linear as in (4.1), reduces to the latter case by adding to the set H an extra element $h = 0$, with $E^0_{sw} = 1$ for all s, w. This is the case here, where (4.1) simultaneously takes into account the fixed part (contracted power) and the variable portion (depending on the consumption). We also make classical assumptions:

Assumption 4.2.2. (i) *Unit-Demand*: Each customer purchases exactly one contract.

(ii) *Envy-free*: There is no limitation on the number of customers able to purchase the same contract and so each customer chooses a contract maximizing his utility.

(iii) *No-purchase option*: Consumers have the option not to purchase any contract, or in a competitive environment, to choose a contract from a competitor.

The *utility* of segment s for contract w is the difference between the reservation bill and the invoice, i.e.,

$$U_{sw}(x) := R_{sw} - \theta_{sw}(x) \ .$$

The *disutility* is then the opposite of the utility. The no-purchase option corresponds to the fact that in competitive environment, customers can choose a contract among those proposed by competitors. We assume here that the competition is static, meaning that competitors do not react to the company prices. Therefore, competing contracts could be understood in our context as regulated alternatives (for instance, in the French electricity market, there are several such offers with prices determined by a regulation authority). More generally, the contracts from different static competitors can be aggregated in a unique contract of a virtual competitor, and the reservation bill R_{sw} consists here in the infimum of the bills proposed by the competition to segment s (there can be an additional term representing a given preference for the contract w). This utility is also called *surplus*, as it corresponds to the additional gain in terms of utility that a consumer can expect by choosing an offer from the leader, compared to the no-purchase option. The utility of the no-purchase option is therefore set to be 0.

Remark 4.2.1

We could also set the utility to be the opposite of the bill, i.e., $U_{sw}(x) = -\theta_{sw}(x)$ and the no-purchase utility to be R_{sw}, but as the utilities are defined up to an additive constant in choice models, the standard normalization is to set the no-purchase utility to 0.

Finally, when a segment s chooses a contract w, the company has to fulfill the service, implying a cost C_{sw}. In the case of an electricity retailer, it has to supply electricity and we suppose that they have the same structure as the bills θ i.e.,

$$\forall s \in [S],\, w \in [W], \quad C_{sw} = \left\langle E_{sw}, \check{C}_w \right\rangle_H$$

where $\check{C}_w = (\check{C}_w^h)_{h\in H} \in \mathbb{R}^H$. In this way, $\check{C}_w^h$ represents the *unitary production cost* at period h for the contract type w. Note that this cost depends on the contract: for instance, a "green electricity" contract may induce a higher production costs than a classical contract. A fixed cost (not proportional to the consumption) can be incorporated in this model by introducing the dummy period $h = 0$ with a unit virtual consumption $E_{sw}^0 = 1$, as explained above.

To model the customers behavior of segment s, we define the variables $y_s \in \mathbb{R}^W$ such that

$$\forall s \in [S],\ w \in [W], \quad y_{sw} = \begin{cases} 1 \text{ if segment } s \text{ chooses } w, \\ 0 \text{ otherwise.} \end{cases} \tag{4.2}$$

To make explicit the no-purchase option, we introduce a variable y_{s0} and denote the extended choice vector for segment s by $\bar{y}_s := (y_{s0}, y_s) \in \mathbb{R} \times \mathbb{R}^W = \mathbb{R}^{W+1}$. We shall think of an element $\bar{y}_s \in \Delta_{W+1}$ as a *relaxed* choice of segment s. When $\bar{y}_s$ is a vertex of Δ_{W+1}, $y_s = (y_{sw})_{w\in W}$ determines the behavior of segment s, according to (4.2). The no-purchase option corresponds to $y_{s0} = 1$. For a price strategy $x \in X$, the customers behavior is defined by the solution set mapping Ψ defined as

$$\Psi(x) := \operatorname*{arg\,min}_{\bar{y}' \in (\Delta_{W+1})^S} \left\{ \sum_{s\in[S]} \langle \theta_s(x) - R_s, y'_s \rangle_W \right\} . \tag{4.3}$$

Note that the scalar product that appears in the objective is on $\mathbb{R}^W$ since the no-purchase option induces a zero utility for any customer.

The *multi-attribute unit-demand envy-free pricing problem* can now be expressed as the following bilinear bilevel model

$$\max_{x\in X, \bar{y}} \left\{ F(x, \bar{y}) := \sum_{s\in[S]} \rho_s \langle \theta_s(x) - C_s, y_s \rangle_W \;\middle|\; (x, \bar{y}) \in \operatorname{gph} \Psi \right\} \tag{o-BP}$$

In the model, ρ_s stands for the weight of segment s in terms of company's profit. Note that there is an asymmetry in the two objective functions: the leader aims at maximizing quantities $\theta_{sw}(x) - C_{sw}$ while the follower aims at minimizing the disutility $\theta_{sw}(x) - R_{sw}$. The very special case $C = R$ would lead to a subclass of bilevel problems, known as zero-sum games, see e.g., [Was14]. The label (o-BP) refers to the *optimistic* nature of this bilevel problem: if the lower level problem has several optimal solutions, the upper level optimizer takes into account the most favorable of these optimal solutions, see e.g. [Dem+15].

Remark 4.2.2

Because all the segments react *independently*, we can aggregate all their actions under the same problem. Hence, the minimization in the lower level problem (4.3) is made over the Cartesian product of simplices. The vertices of each of these simplices represent the possible decisions of a given segment.

The following result justifies the minimization over relaxed choices in (o-BP).

Proposition 4.2.1

There exists an optimal solution of (o-BP) with integer lower values y.

Proof. We denote by $(x^*, \bar{y}^*)$ an optimal solution, which exists because $\operatorname{gph} \Psi$ is compact and non-empty (from Assumption 4.2.1). The argmin set $\Psi(x^*)$ is a face of the Cartesian product of simplices $(\Delta_{W+1})^S$, since it arises from the minimization of a linear objective on this product. So it is a non-empty integer polyhedron. Moreover, there exists an extreme point of $\Psi(x^*)$, denoted by $\hat{y}$, such that $F(x^*, \hat{y}) = F(x^*, \bar{y}^*)$ owing to the linearity in y of the upper objective. To conclude, $(x^*, \hat{y})$ is also an optimal solution and $\hat{y}$ is integer as extreme point of $\Psi(x^*)$. □

Problem (o-BP) is a very specific bilinear bilevel problem with a quite simple lower problem (minimization over the simplex, without integrity constraints). However, despite its apparent simplicity, this model is APX-hard since it includes as a special case the unit-demand envy-free pricing model, which was shown to be APX-hard, see [Gur+05].

The problem (o-BP) is a *profit-maximization* problem: in fact, we can define the optimistic leader profit function π^{opt} for a given price strategy x as

$$\pi^{opt}(x) := \sum_{s\in[S]} \rho_s \sum_{w\in[W]} (\theta_{sw}(x) - C_{sw}) y_{sw}^{opt}(x) \tag{4.4}$$

where $y_{sw}^{opt}(x)$ is the optimistic lower response (which is binary, see (4.2)). The problem (o-BP) is therefore the maximization of the function π^{opt} over X. The optimistic profit function π^{opt} is piecewise linear (the profit is linear for a given customers distribution y, and the possible customers distribution lies in a discrete set). However, π^{opt} is in general discontinuous at prices inducing *ties* (multiple minimum disutilities for a segment), see Fig. 4.3.

4.2.3 Tie-breaking rules

Proposition 4.2.2 (Degeneracy)
Let $(x^*, \bar{y}^*)$ be an optimal solution of (o-BP) and suppose that all the contracts are chosen by at least one segment (otherwise the contract is not useful). If $x^* \in \mathrm{Int}(X)$, then there are at least W segments that face ties i.e. for all $w \in [W]$, there exists $s \in [S]$ such that $y_{sw}^* = 1$ and

$$\theta_{sw}(x^*) - R_{sw} = \min\left\{0, \{\theta_{sw'}(x^*) - R_{sw'}\}_{(w'\neq w)}\right\} .$$

Proof. Suppose that

$$\exists w \in [W], y_{sw}^* = 1 \Rightarrow \theta_{sw}(x^*) - R_{sw} < \min\left\{0, \{\theta_{sw'}(x^*) - R_{sw'}\}_{(w'\neq w)}\right\} .$$

Then, one could increase the price of the contract w by a little amount – which is possible because $x^* \in \mathrm{Int}(X)$ – and keep the same customers response, contradicting the optimality of the leader's decision. □

Proposition 4.2.2 proves that solutions always contain ties between invoices for some customers. Therefore, the relevance of the optimistic hypothesis has to be discussed. To this end, we consider two other versions of the problem (o-BP). First let us consider the *pessimistic* version in which customers having ties are supposed to choose the worst invoice in terms of leader's profit:

$$\begin{aligned} \sup_{x\in X} \min_{\bar{y}} \quad & \sum_{s\in[S]} \rho_s \left\langle \theta_s(x) - C_s, y_s \right\rangle_W \\ \text{s. t.} \quad & \bar{y}_s \in \operatorname*{arg\,min}_{\bar{y}'_s \in \Delta_{W+1}} \left\langle \theta_s(x) - R_{sw}, y'_s \right\rangle_W , \forall s \end{aligned} \tag{p-BP}$$

Remark 4.2.3

The existence of a solution is not guaranteed. Dempe gives in [Dem02, Theorem 3.3] a pessimistic problem such that the optimum is not attained.

We also introduce an intermediary model between optimistic and pessimistic, called *uniform* model, defined as

$$\begin{aligned} \sup_{x \in X} \quad & \sum_{s \in [S]} \rho_s \langle \theta_s(x) - C_s, y_s \rangle_W \\ \text{s.t.} \quad & y_{sw} = \mathbb{1}_{w \in \Phi_s(x)} / |\Phi_s(x)|, \ \forall s, w \end{aligned} \tag{u-BP}$$

where $\Phi_s(x) := \arg\min \left\{0, \{\theta_{sw}(x) - R_{sw}\}_{w \in [W]}\right\}$ denotes the set of minimum disutilities among the $W+1$ possibilities. In this model, when a tie occurs, customers have no preference on the contracts and spread their choice on the different possibilities (of course, the lower response cannot be binary anymore). We then define the uniform leader profit function as

$$\pi^{uni}(x) := \sum_{s \in [S]} \sum_{w \in [W]} (\theta_{sw}(x) - C_{sw}) y^{uni}_{sw}(x) \tag{4.5}$$

where y^{uni} stands for the uniform lower response. The three responses (optimistic, pessimistic and uniform) only differ when a tie between two contracts appears. Using an idea of Gilbert, Marcotte and Savard [GMS15], the following proposition establishes the equality of the three versions in terms of optimal value, assuming a condition on the costs.

Assumption 4.2.3 (No-profit option). We say that the bilevel models (o-BP), (p-BP) and (u-BP) allow the *no-profit option* if we can set the price to be equal to $\check{C}$ i.e., $\check{C} \in X$.

In the electricity provider context, this condition is realistic: setting prices equal to $\check{C}$ yields a public service type policy (with no benefit), in which the company aims to cover exactly its costs.

Theorem 4.2.1 (Indifference to tie-breaking rule)

The inequalities $v(o\text{-}BP) \geqslant v(u\text{-}BP) \geqslant v(p\text{-}BP)$ always hold. Moreover, as soon as the models allow the no-profit option,

$$v(o\text{-}BP) = v(u\text{-}BP) = v(p\text{-}BP) \ .$$

Proof. The inequalities $v(o\text{-}BP) \geqslant v(u\text{-}BP) \geqslant v(p\text{-}BP)$ are immediate. Let $(x^*, \bar{y}^*)$ be an optimistic optimal solution. For any $\delta > 0$, we consider the perturbed price matrix x^δ defined by

$$\forall w \in [W],\ h \in [H],\ (x^\delta)^h_w = \frac{1}{1+\delta}\left((x^*)^h_w + \delta \check{C}^h_w\right) \ .$$

This new price matrix lies in the polytope X , since it a barycenter of $x^* \in X$ and $\check{C} \in X$. Suppose that for a given segment s, there is a tie between contract w_1 and w_2 i.e.,

$$\theta_{sw_1}(x^*) - R_{sw_1} = \theta_{sw_2}(x^*) - R_{sw_2} \ .$$

In this optimistic problem, the segment s chooses between w_1 and w_2 the contract that maximizes the profit of the leader i.e., the contract with the highest value $\theta_{sw}(x^*) - C_{sw}$. Without loss of generality, we suppose that it is w_1.

The new price policy x^δ is constructed in order to break the tie while keeping the same choice of contract: from the definition of x^δ, one can obtain that for any segment s and any contract w $\theta_{sw}(x^\delta) - \theta_{sw}(x^*) = -\frac{\delta}{1+\delta}(\theta_{sw}(x^*) - C_{sw})$. Then,

$$\begin{aligned}
&\left[\theta_{sw_1}(x^\delta) - R_{sw_1}\right] - \left[\theta_{sw_2}(x^\delta) - R_{sw_2}\right] \\
=&\left[\theta_{sw_1}(x^\delta) - \theta_{sw_1}(x^*)\right] - \left[\theta_{sw_2}(x^\delta) - \theta_{sw_2}(x^*)\right] \\
=&-\frac{\delta}{1+\delta}\left([\theta_{sw_1}(x^*) - C_{sw_1}] - [\theta_{sw_2}(x^*) - C_{sw_2}]\right) \leqslant 0 \ .
\end{aligned}$$

Note that the only possibility to conserve the tie between w_1 and w_2 with price strategy x^δ is when both contracts yield the same profit for the leader. Therefore, for any δ sufficiently close to 0, $(x^\delta, \bar{y}^*)$ is a pessimistic solution with objective $\frac{1}{1+\delta}v(o\text{-}BP)$. Hence, $v(p\text{-}BP) \geqslant \frac{1}{1+\delta}v(o\text{-}BP)$, leading to $v(p\text{-}BP) \geqslant v(o\text{-}BP)$ when $\delta \to 0^+$. □

Theorem 4.2.1 proves that under the Assumption 4.2.3 the tie-breaking rule at the lower level does not affect the optimal value. In consequence, considering the optimistic behavior does not introduce any bias.

4.2.4 Single-level reformulation

Classical KKT transformation. The most common way to express the optimality of the lower problem as a system of inequalities is to use the Karush-Kuhn-Tucker (KKT) conditions. Applying this idea to (o-BP) leads to the following formulation

$$\begin{aligned}
\max_{x\in X,\bar{y}} \quad & \sum_{s\in[S]} \rho_s\mu_s + \rho_s \langle R_s - C_s, y_s\rangle_W \\
\text{s.t.} \quad & 0 \leqslant y_{sw} \perp \theta_{sw}(x) - R_{sw} - \mu_s \geqslant 0, \forall s, w \\
& 0 \leqslant y_{s0} \perp \mu_s \leqslant 0, \forall s \\
& \bar{y}_s \in \Delta_{W+1}, \forall s
\end{aligned} \tag{o-KKT}$$

To numerically solve this formulation, we usually replace the complementarity constraints by Big-M constraints introducing new binary variables. Using Proposition 4.2.1, we provides a compact formulation in which the lower variables y_s are the only binary variables:

$$\begin{aligned}
\max_{x\in X,\bar{y}} \quad & \sum_{s\in[S]} \rho_s\mu_s + \rho_s \langle R_s - C_s, y_s\rangle_W \\
\text{s.t.} \quad & 0 \leqslant \theta_{sw}(x) - R_{sw} - \mu_s \leqslant M_{sw}(1 - y_{sw}), \forall s, w \\
& 0 \leqslant -\mu_s \leqslant M_{s0}(1 - y_{s0}), \forall s \\
& \bar{y}_s \in \mathrm{Vert}(\Delta_{W+1}), \forall s
\end{aligned} \tag{4.6}$$

Here, the set of vertices $\mathrm{Vert}(\Delta_{W+1})$ is known and is equal to $\left\{y \in \{0,1\}^{W+1} \mid \sum_{w=0}^{W} y_w = 1\right\}$. The Big-$M$ parameters $M_{sw} > 0$ must be chosen to be sufficiently large to prevent the elimination of any optimal solution, see [PM19; KS23]. This is in general as hard as solving the initial bilevel problem, see [Kle+20]. However, in the present case, owing to the boundedness of the pricing variables $x \in X$ and the structure of the constraints, we can explicitly find valid Big-M values. If $X \subseteq \prod_{1\leqslant w\leqslant W}[x_w^-, x_w^+]$, then it sufficies to take:

$$M_{sw} = \theta_{sw}(x^+) - R_{sw} + M_{s0}\,, \quad M_{s0} = \max\{0, \max_{1\leqslant w\leqslant W}\{R_{sw} - \theta_{sw}(x^-)\}$$

Remark 4.2.4

The formulations (*o-KKT*) and (4.6) generalize the (U) formulation introduced by [Fer+13] that applies in the single-attribute case: the variables μ_s express the disutilities of each segment s.

Using the strong-duality condition. We next present an alternative formulation exploiting strong-duality, following an idea of Kleinert et al. [Kle+21] and the references therein. It uses the dual of the lower problem, and expresses the equality of the primal and dual objectives.

Proposition 4.2.3

We can reformulate (*o-BP*) as

$$\begin{aligned} \max_{x \in X, \bar{y}} \quad & \sum_{s \in [S]} \rho_s \langle \theta_s(x) - C_s, y_s \rangle_W \\ \text{s.t.} \quad & \langle \theta_s(x) - R_s, y_s \rangle_W \leqslant \theta_{sw}(x) - R_{sw}, \forall s, w \\ & \langle \theta_s(x) - R_s, y_s \rangle_W \leqslant 0, \forall s \\ & \bar{y}_s \in \Delta_{W+1}, \forall s \end{aligned} \tag{o-SDC}$$

Proof. For a segment s, the dual of the lower problem (4.3) is expressed as

$$\max_{\mu_s \in \mathbb{R}} \left\{ \mu_s \;\middle|\; \begin{array}{l} \mu_s \leqslant \theta_{sw}(x) - R_{sw}, \forall w \\ \mu_s \leqslant 0 \end{array} \right\} .$$

Due to strong duality, the primal objective is lower than the dual one and observing that the dual variable can be eliminated gives us the result. □

Remark 4.2.5

In our special case, the approach by strong duality leads to the same formulation as the *Optimal Value Transformation*[Dem+15]. In fact, being lower than all possible values is here equivalent to being lower than the extreme points of the simplex, which is exactly what the strong-duality reformulation gives us.

The formulation (*o-SDC*) does not contain complementarity constraints but there is no free-lunch: the difficulty is recast in the bilinear terms $\theta_{sw}(x)y_{sw}$, $\forall s, w$. These terms are non-convex and they are often relaxed into lifting variables that have to verify the McCormick inequalities. Here, thanks to Proposition 4.2.1, we can suppose that the lower variables are binary, and thus the McCormick relaxation becomes exact. Formulation (*o-SDC*) is therefore equivalent to

$$\begin{aligned} \max_{x \in X, \bar{y}} \quad & \sum_{s \in [S]} \rho_s \sum_{w \in [W]} \nu_{sw} - C_{sw} y_{sw} \\ \text{s.t.} \quad & \textstyle\sum_{w' \in [W]} \nu_{sw'} - R_{sw'} y_{sw'} \leqslant \theta_{sw}(x) - R_{sw}, \forall s, w \\ & \textstyle\sum_{w' \in [W]} \nu_{sw'} - R_{sw'} y_{sw'} \leqslant 0, \forall s \\ & \nu_{sw} \leqslant \theta_{sw}(x), \forall s, w \\ & \nu_{sw} \leqslant R_{sw} y_{sw}, \forall s, w \\ & \nu_{sw} \geqslant \theta_{sw}(x) - \max_{x \in X}\{\theta_{sw}(x)\}(1 - y_{sw}), \forall s, w \\ & \bar{y}_s \in \mathrm{Vert}(\Delta_{W+1}), \; \nu_s \in \mathbb{R}^W_{\geqslant 0}, \forall s \end{aligned} \tag{4.7}$$

Once again, the compactness of X ensures that the Big-M value $\max_{x\in X}\{\theta_{sw}(x)\}$ is sufficiently large to keep the validity of the reformulation. Note that the lifting variable ν_{sw} represents the product $\theta_{sw}(x)y_{sw}$ which cannot exceed R_{sw}.

The MIP formulation of equation (4.7) extends the formulation HLMS, cited in [Fer+16] and initially developed by Heilporn et al [Hei+10], to the multi-attribute case.

4.2.5 Logit regularization

The formulation (o-BP) models customers reactions as deterministic behaviors. It relies on two assumptions:

(i) customers have perfect rational and deterministic behavior,

(ii) parameters such as reservation bills and costs are perfectly known.

Both assumptions can be discussed: not only real customers are not purely rational agents in that they can choose a contract that does not maximize the utility, but also a segment is the aggregation of quasi-similar customers, not strictly identical ones. Therefore in reality, when a segment faces two very close disutilities, customers of this segment are likely to spread themselves over the two possibilities. Besides, the reservation bills and costs are estimations obtained by analysis on the market but cannot be known exactly. Hence, assuming lower response to be binary as in the optimistic model can be quite unrealistic and may lead to an unachievable optimum. This can be avoided by *Logit* modeling which captures the probabilistic nature of customers' choice by adding a Gumbel uncertainty. There is a wide literature which uses this approach as choice models, see e.g. [Tra09] and the references therein.

Previously, consumers were supposed to choose a contract minimizing their deterministic disutility i.e., each segment $s \in [S]$ selects $w^* \in \{0 \dots W\}$ such as $V_{sw^*} = \min_{w\in\{0\dots W\}} V_{sw}$ where $V_{sw} := \theta_{sw}(x) - R_{sw}$ for all $w \in [W]$ and $V_{s0} := 0$. We now suppose that their disutilities are defined as

$$U_{sw} := \beta V_{sw} + \varepsilon_{sw},\ \forall s, w,$$

where $\{\varepsilon_{sw}\}_w$ is a family of Gumbel random variables, distributed identically and independently, and $\beta \geqslant 0$ is an inverse temperature in the sense of physics. The choice of Gumbel uncertainties is standard in discrete choice theory, and the main underlying assumption is not so much about the shape of the uncertainty but rather on the independence of the noises [Tra09, Chapter 3]. Here, we suppose that the utilities capture enough information so that the remaining part of the uncertainty behaves as a white noise.

> **Remark 4.2.6**
>
> In the sequel, we consider a common β across the segments, but all the results still apply for a differentiated value $\beta_s = d_s\beta$, where d_s is a given parameter. This corresponds to a rescaling of β, adapted to each segment.

As a consequence, the lower response is expressed as

$$y_{sw} = \mathbb{P}[U_{sw} \leqslant U_{sw'},\ \forall w' \neq w],\ \forall s, w\ . \tag{4.8}$$

Hence, a customer has a probability to choose a contract which is not the optimal one in terms of deterministic utility. The calculation of the probability y_{sw} arising in equation (4.8) is done in [Tra09] and it has an explicit form. Replacing the deterministic lower response by this expression of y_{sw} leads to the following *Mixed Multinomial Logit* model:

$$\begin{aligned} \max_{x\in X,y} \quad & \sum_{s\in[S]} \rho_s \langle \theta_s(x) - C_s, y_s \rangle_W \\ \text{s.t.} \quad & y_{sw} = \frac{e^{-\beta(\theta_{sw}(x)-R_{sw})}}{1+\sum_{w'\in[W]} e^{-\beta(\theta_{sw'}(x)-R_{sw'})}}, \ \forall s,w \end{aligned} \tag{β-BP}$$

Remark 4.2.7

The '1' in the denominator corresponds to the no-purchase option.

Equivalently, we recall here a standard reformulation of (β-BP) :

Proposition 4.2.4

Problem (β-BP) is equivalent to

$$\begin{aligned} \max_{x\in X,\bar{y}} \quad & \sum_{s\in[S]} \rho_s \langle \theta_s(x) - C_s, y_s \rangle_W \\ \text{s.t.} \quad & \bar{y}_s \in \operatorname*{arg\,min}_{\bar{y}'_s\in\Delta_{W+1}} \left\{ \langle \theta_s(x) - R_s, y'_s \rangle_W + \frac{1}{\beta} \langle \log(\bar{y}'_s), \bar{y}'_s \rangle_{W+1} \right\}, \ \forall s \end{aligned} \tag{4.9}$$

Proof. Given $V \in \mathbb{R}^W$, we study the problem: $\min_{\bar{y}\in\Delta_{W+1}} \left\{ \langle V, y \rangle_W + \beta^{-1} \langle \log(\bar{y}), \bar{y} \rangle_{W+1} \right\}$. First note that the positivity assumption is always satisfied at the optimum, since the function $y\log(y)$ acts a barrier. Looking at the KKT optimality conditions , we then obtain that there exists $\mu \in \mathbb{R}$ (dual variable of the constraint $\sum_w y_w = 1$) such that for any $w \leqslant W$, $0 = V_w + \frac{1}{\beta}(\log(y_w)+1) - \mu$. This implies that $y_w = \exp(\beta\mu - 1)\exp(-\beta V_w)$. As $\bar{y}$ must lie in the simplex, we recover the standard expression of the logit model. □

This model highlights that the logit expression is the optimum of a strictly convex minimization problem (the property was pointed out in [Fis80] and [GMS15]). The objective function is the deterministic function to which we add the entropic regularization term $\beta^{-1} \langle \log(\bar{y}'_s), \bar{y}'_s \rangle_{W+1}$, attracting the lower response to the center of the simplex Δ_{W+1}.

The model (β-BP) is intrinsically defined as a single-level problem since the lower response for any segment s is unique and analytically known. For a given price strategy x, we define the leader profit function $\pi^{log}(x;\beta)$ as

$$\pi^{log}(x;\beta) := \sum_{s\in[S]} \rho_s \sum_{w\in[W]} (\theta_{sw}(x) - C_{sw}) y^{log}_{sw}(x;\beta) \tag{4.10}$$

where y^{log} stands for the logit lower response. This objective function π^{log} is in general neither concave nor convex, see [Li+19].

We can think of the logit model as a regularization of the deterministic case, so it is of interest to look at the convergence for $\beta \to +\infty$, expecting that the regularized optimum converges to the deterministic optimum. Such a study has been done by Gilbert, Marcotte and Savard [GMS15] in the context of toll pricing. In our setting, we prove asymptotic results without requiring equality between the optimistic and pessimistic versions.

Proposition 4.2.5

For a price strategy $x \in X$, $\lim_{\beta\to+\infty} y^{log}(x;\beta) = y^{uni}(x)$ and so $\lim_{\beta\to+\infty} \pi^{log}(x;\beta) = \pi^{uni}(x)$. Moreover,

$$v(u\text{-}BP) \leqslant \liminf_{\beta\to+\infty} v(\beta\text{-}BP) \leqslant \limsup_{\beta\to+\infty} v(\beta\text{-}BP) \leqslant v(o\text{-}BP)$$

and the equalities occurs under the no-profit option.

Proof. ◇ Inequality $\limsup_{\beta\to+\infty} v(\beta\text{-}BP) \leqslant v(o\text{-}BP)$:

Let $(\beta_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ such that $\lim_{n\to+\infty}\beta_n = +\infty$ and $(z_n)_{n\in\mathbb{N}} \in (X\times\Delta_{W+1})^{\mathbb{N}}$ the sequence of $z_n = (x_n, \bar{y}_n)$ solutions of $(\beta_n\text{-}BP)$. By compactness of X and Δ_{W+1}, $(z_n)_{n\in\mathbb{N}}$ has accumulation points.

Let $(x^*, \bar{y}^*)$ be one of these points, we must show that $\bar{y}^* \in \Psi(x^*)$. We take $s \in [S]$ and two contracts w_1 et w_2 such that

$$\theta_{sw_2}(x^*) - R_{sw_2} = \gamma + \theta_{sw_1}(x^*) - R_{sw_1},\ \gamma > 0.$$

By definition of accumulation points, there exists a sub-sequence $(x_k, \bar{y}_k)_{k\in K}$ converging to $(x^*, \bar{y}^*)$ and by continuity of θ, $\exists k_1 \in \mathbb{N}, \forall k \geqslant k_1,\ \theta_{sw_2}(x_k) - R_{sw_2} \geqslant \theta_{sw_1}(x_k) - R_{sw_1} + \frac{1}{2}\gamma$.
Therefore, using the definition of logit probabilities, $\forall k \geqslant k_1,\ 0 \leqslant (\bar{y}_k)_{sw_2} \leqslant e^{-\frac{1}{2}\beta_k\gamma}$.
Because β_k goes to infinity, we can conclude that $(\bar{y}^*)_{sw_2} = 0$ and since only minimum disutilities have positive probabilities, $\bar{y}^* \in \Psi(x^*)$.
To conclude, all adherent points $(x^*, \bar{y}^*)$ of the sequence are feasible solution of $(o\text{-}BP)$.

◇ Inequality $v(u\text{-}BP) \leqslant \liminf_{\beta\to+\infty} v(\beta\text{-}BP)$:

Because we can't be sure that the supremum of the uniform problem $(u\text{-}BP)$ is attained, we take a sequence $(x_n, \bar{y}_n)_{n\in\mathbb{N}} \in (X\times\Delta_{W+1})^{\mathbb{N}}$ of solutions of the uniform model. We denoted by v_n their objective value which converges to $v(u\text{-}BP)$.
Let $\varepsilon > 0$ be given, from the convergence, $\exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ v_n \geqslant v(u\text{-}BP) - \varepsilon/2$.
Besides, from each x_n we construct $\tilde{y}_n(\beta)$ such that $\tilde{y}_n(\beta) = \text{logit}(x_n;\beta)$ and $\lim_{\beta\to+\infty}\tilde{y}_n(\beta) = \bar{y}_n$. For all β, $(x_n, \tilde{y}_n(\beta))$ is a valid solution for $(\beta\text{-}BP)$ and its objective funtion converges to v_n when $\beta \to +\infty$ i.e.,

$$\exists \beta_n \in \mathbb{R},\ \forall \beta \geqslant \beta_n,\ v(\beta\text{-}BP) \geqslant v_n - \varepsilon/2\ .$$

In particular, for all $\beta \geqslant \beta_{n_1}$, $v(\beta\text{-}BP) \geqslant v(u\text{-}BP) - \varepsilon$.

□

The last proposition confirms that $(\beta\text{-}BP)$ is a valid regularization in that it consists on a smooth approximation of $(o\text{-}BP)$ for sufficiently big β value. Nonetheless, we want to go further in the analysis by obtaining indications on the evolution of the optimal value as β grows. To do so, we study the simpler case where there is a unique segment (homogeneous population) and unconstrained prices. This leads to a pricing model under the standard Mixed Multinomial Logit (MNL) customer behavior:

$$v_\beta := \max_{\theta,y}\left\{\langle\theta - C, y\rangle_W \;\middle|\; y_w = \frac{e^{-\beta(\theta_w - R_w)}}{1+\sum_{w'\in[W]} e^{-\beta(\theta_{w'} - R_{w'})}},\ \forall w\right\}\ . \tag{4.11}$$

In [LH11], Li et al. deeply study the model defined in (4.11) and provide in particular a characterization of its optimal solution. We reuse this property in the next proposition to show an asymptotic result. To this end, we set $v_\infty := \max_w(R_w - C_w)$ and $\#v_\infty$ the cardinality of the latter argmax.

4. Quad. Regul.

Proposition 4.2.6 (Customers behavior)
For the standard MNL model defined in equation (4.11),

(i) $v_\beta \underset{\beta\to 0}{=} \frac{1}{\beta}\mathcal{W}_0\left(W/e\right) + o\left(\frac{1}{\beta}\right)$; where $\mathcal{W}_0$ denotes the Lambert function [Cor+96].

(ii) if $v_\infty > 0$ then $v_\beta \underset{\beta\to+\infty}{=} v_\infty - \frac{\ln(\beta v_\infty)}{\beta} + \frac{\ln(\# v_\infty)-1}{\beta} + o\left(\frac{1}{\beta}\right)$.

Proof. We know by [LH11, Theorem 2] that v_β satisfies

$$\beta v_\beta e^{\beta v_\beta} = \sum_{w\in[W]} e^{-1+\beta(R_w - C_w)} \ .$$

and so $\beta v_\beta = \mathcal{W}_0(f(\beta))$, where $f(\beta) = \sum_{w\in[W]} e^{-1+\beta(R_w-C_w)}$.

The result for the first item comes naturally.
For the second, because we suppose $v_\infty > 0$, we have $\lim_{\beta\to+\infty} f(\beta) = +\infty$.
An elementary calculation shows that

$$\ln(f(\beta)) \underset{\beta\to+\infty}{=} \beta v_\infty - 1 + \ln(\# v_\infty) + o(1) \ .$$

and it follows also $\ln\ln(f(\beta)) \underset{\beta\to+\infty}{=} \ln(\beta v_\infty) + o(1)$.
From the properties of the Lambert function,

$$\mathcal{W}_0(f(\beta)) \underset{\beta\to+\infty}{=} \ln(f(\beta)) - \ln\ln(f(\beta)) + O\left(\frac{\ln\ln(f(\beta))}{\ln(f(\beta))}\right) \ .$$

We therefore obtain that $\beta v_\beta \underset{\beta\to+\infty}{=} \beta v_\infty - \ln(\beta v_\infty) - 1 + \ln(\# v_\infty) + o(1)$. The result is obtained by dividing by β. □

In Figure 4.1, we draw the optimal value on a simple case, along with the first-order asymptotic expansions found in proposition 4.2.6. The result for small β values is quite intuitive: with customers randomly reacting, the company can impose very high price in such a way that there are always some consumers buying its products. Hence, the company's profit becomes infinite as $\beta \to 0$. The result for large β values is not so evident and can be interpreted as a *moral hazard*, see the numerical section for an interpretation in the general case.

4.3 Quadratic regularization

In the case of a homogeneous population and unconstrained prices, [Li+19] express the problem (β-BP) in terms of lower variables to obtain a concave maximization problem. If we add bounds on prices and consider multi-attribute utilities, [SK20] show another concave transformation that keeps tractability in the resolution. However, with heterogeneous segments as it is the case here, no tractable transformation is known, and only local optimum of (β-BP) can generally be found. This motivates us to look at a new convex penalization, replacing the entropy penalization term in (4.9) by a quadratic one.

$$\begin{aligned} &\max_{x\in X, y} \sum_{s\in[S]} \rho_s \langle \theta_s(x) - C_s, y_s\rangle_W \\ &\text{s. t. } \bar{y}_s \in \operatorname*{arg\,min}_{\bar{y}'_s \in \Delta_{W+1}} \left\{ \langle \theta_s(x) - R_s, y'_s\rangle_W + \frac{1}{\beta}\langle \bar{y}'_s - 1, \bar{y}'_s\rangle_{W+1} \right\}, \forall s \end{aligned} \qquad (q\beta\text{-}BP)$$

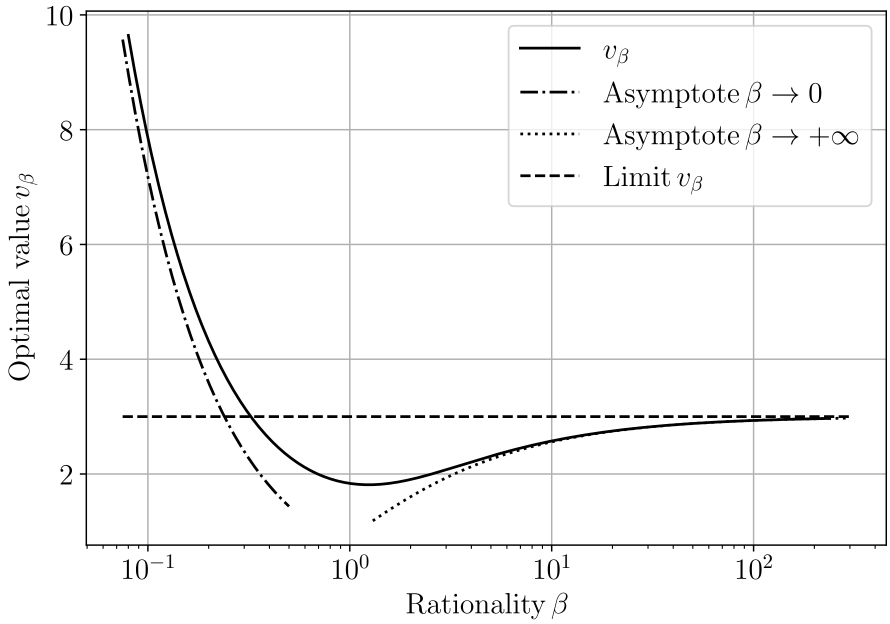


Figure 4.1: Optimal value of (4.11) according to β.

The solid line represents the optimal value of the example.
The asymptotic results found in 4.2.6 are drawn with dotted lines.

The quadratic term $\beta^{-1} \langle \bar{y} - 1, \bar{y} \rangle$ is chosen so that it vanishes at any vertex of the simplex Δ_{W+1}. The following result shows that two perhaps more intuitive quadratic terms lead to the same optimum.

Proposition 4.3.1
The two following penalizations are equivalent to the one in ($q\beta$-BP):

(i) $\frac{1}{\beta} \left\| \bar{y}_s - \frac{1}{W+1} \right\|^2_{W+1}$ (uniform law attractor),

(ii) $\frac{1}{\beta} \left\| \bar{y}_s \right\|^2_{W+1}$.

Proof. $\|\bar{y}_s - \alpha\|^2_{W+1} - \langle \bar{y}_s - 1, \bar{y}_s \rangle_{W+1} = (1 - 2\alpha) \left(\sum_{w=0}^{W} y_{sw} \right) + (W+1)\alpha^2 = (1-\alpha)^2 + W\alpha^2$. The two objective functions are equal up to a constant for valid lower responses, thus the argmins are the same. □

The first item suggests that our new penalization acts as an attractor to the uniform law whose intensity is inversely proportional to β. The bigger β is, the more customers will uniformly spread their choices on all the possibilities. This asymptotic behavior is therefore identical to the one of logit regularization. In Section 4.8.2, we provide metric estimates – along with illustrations – in order to compare the logit and quadratic regularizations. [DB01] have introduced such a quadratic regularization in order to avoid discontinuities that appears in the deterministic version (o-BP), and theoretically analyze the convergence of a bundle trust region algorithm specifically designed for this problem. Here, the second level is of a particular nature: we focus on lower problem defined on simplices, which allows us to interpret the customers' decision as a geometric object. In particular, for $W+1$ disutilities $V_{s0}, \ldots, V_{sW}$, the follower response of a given segment s can be written as

$$\arg\min_{\Delta_{W+1}} \left\{ \sum_{w=0}^{W} V_{sw} y_{sw} + \frac{1}{\beta} y_{sw}^2 \right\} = \arg\min_{\Delta_{W+1}} \left\| y_s - \left(-\frac{\beta}{2} V_s \right) \right\|_{W+1} = \operatorname*{Proj}_{\Delta_{W+1}} \left(-\frac{\beta}{2} V_s \right) . \tag{4.12}$$

Here again, the disutility V_{sw} of a segment s stands for a certain $\theta_{sw}(x) - R_{sw}$ in the problem ($q\beta$-BP). The response can be understood as a projection on the simplex of a specific vector whose intensity varies proportionally to β.

Remark 4.3.1

The projection on a closed convex set is Lipschitz of constant one in the Euclidean norm, a fortiori, it is continuous. Therefore, the quadratic lower response $y^{quad}(x;\beta)$, solution of the lower problem in ($q\beta$-BP), is a continuous function of the price variables x.

4.3.1 Lower Response and Leader's Profit

In the logit model, the lower response of a segment s is analytically known and is defined by the logit expression. To better understand the customer behavior, we aim to find an explicit calculation of the lower response for a segment s that faces disutilities $V_{s0}, \dots, V_{sW}$. We assume that these disutilities are *sorted in ascending order*. The lower response y that satisfies (4.12) is the solution of the KKT conditions expressed as:

$$\begin{aligned}
&V_{sw} + \frac{2}{\beta} y_{sw} - \lambda_{sw} - \mu_s = 0, && w \in \{0 \dots W\} \\
&0 \leqslant y_{sw} \perp \lambda_{sw} \geqslant 0, && w \in \{0 \dots W\} \\
&y_s \in \Delta_{W+1},\ \lambda_s \in \mathbb{R}^{W+1}_{\geqslant 0},\ \mu_s \in \mathbb{R}
\end{aligned} \tag{4.13}$$

These conditions are necessary and sufficient because we study a convex minimization problem where the Slater's condition holds. In the sequel, we analyze the KKT system (4.13) to characterize the customer's response.

Lemma 4.3.1 (Monotonicity)

If y satisfies (4.13), the sequence $(y_{sw})_{w=0..W}$ is decreasing for disutilities sorted in ascending order.

Proof. We consider $V_{sw_1} \leqslant V_{sw_2}$. If $y_{sw_2} = 0$, there is nothing to prove, the inequality $y_{sw_1} \geqslant y_{sw_2}$ is automatically satisfied. If however $y_{sw_2} > 0$, $\lambda_{sw_2} = 0$ by complementarity, and therefore $V_{sw_1} + \frac{2}{\beta} y_{sw_1} - \lambda_{sw_1} = V_{sw_2} + \frac{2}{\beta} y_{sw_2}$. Since $V_{sw_2} - V_{sw_1} \geqslant 0$ and $\lambda_{sw_1} \geqslant 0$, $\frac{2}{\beta}(y_{sw_1} - y_{sw_2}) \geqslant 0$. Therefore, in any case, $\forall w_1, w_2$, $V_{sw_1} \leqslant V_{sw_2} \Rightarrow y_{sw_1} \geqslant y_{sw_2}$. □

Proposition 4.3.2 (Lower response algorithm)

For any segment s, let the sequence $(c_{sw})_{w \in [W]}$ be

$$c_{sw} := \frac{1}{w} \left[\frac{2}{\beta} + \sum_{w'=0}^{w-1} V_{sw'} \right]$$

and let the index τ be defined as $\tau = \min \{ w \in [W], | \, V_{sw} \geqslant c_{sw} \}$. Then, the sequence (c_{sw}) verifies the following property:

$$V_{sw} < c_{s\tau} \text{ for } w < \tau\,;\ V_{sw} \geqslant c_{s\tau} \text{ for } w \geqslant \tau \tag{4.14}$$

Moreover, the solution (y_s, λ_s, μ_s) of (4.13) can be expressed as follows

(i) $y_{sw} = \frac{\beta}{2}\left[c_{s\tau} - V_{sw}\right]$ for $w < \tau$; $y_{sw} = 0$ for $w \geqslant \tau$,

(ii) $\lambda_{sw} = 0$ for $w < \tau$; $\lambda_{s\tau} = V_{s\tau} - c_{s\tau}$; $\lambda_{sw} = \lambda_{s,w-1} + V_{sw} - V_{s,w-1}$ for $w > \tau$,

(iii) $\mu_s = c_{s\tau}$.

The index τ is therefore the index from which the probability y becomes zero.

Proof. The first property on (c_{sw}) comes with the ascending sort of V_s and the definition of τ: $V_{s\tau} \geqslant c_{s\tau}$ and therefore $V_{sw} \geqslant c_{s\tau}$ for $w \geqslant \tau$. Besides, by minimality of τ, $V_{s,\tau-1} < c_{s,\tau-1}$. Using the definition of (c_{sw}), for all $w < \tau$, $V_{sw} \leqslant V_{s,\tau-1} = c_{s\tau} - \frac{\tau-1}{\tau}(c_{s,\tau-1} - V_{s,\tau-1}) < c_{s\tau}$.

Concerning the second part of the proposition, one can first remark that solution of (4.13) is unique since it is a projection on the simplex, see (4.12). The procedure returns a certain (y_s, λ_s, μ_s) which is feasible for (4.13): by construction, y is nonnegative, $\sum_{w=0}^{\tau} y_{sw} = 1$ and the complementarity constraints are satisfied. As the disutilities are sorted, $\lambda_{sw} \geqslant \lambda_{s\tau} \geqslant 0$ for any $w \geqslant \tau$. The solution we obtain is therefore the unique solution of (4.13). □

From the explicit calculation of the lower response, one can observe the following property

Corollary 4.3.1 (Soft threshold)

If y_s satisfies (4.13), the first disutility is chosen with probability 1 if and only if the difference between any other disutility and the one chosen is higher than $2/\beta$ i.e.,

$$y_{s0} = 1 \text{ and } \forall w > 0,\, y_{sw} = 0 \iff \forall w > 0,\, V_{sw} \geqslant V_{s0} + \frac{2}{\beta}\ . \tag{4.15}$$

Proof. From the last proposition, the condition $\forall w > 0$, $V_{sw} \geqslant V_{s0} + \frac{2}{\beta}$ is equivalent to $V_{s1} \geqslant V_{s0} + \frac{2}{\beta}$ which means that $\tau = 1$. □

Coming back to problem ($q\beta$-BP), we summarize the properties of the lower response in the following corollary

Corollary 4.3.2 (Lower response of ($q\beta$-BP))

For a price strategy x and a given β, the quadratic lower response $\left(y_{sw}^{quad}(x;\beta)\right)_{w=0\ldots W}$ for a segment s can be computed by the following algorithm:

1. Compute $V_{sw}(x) := \theta_{sw}(x) - R_{sw}$ for all $w \in [W]$ and $V_{s0} = 0$,
2. Reindex the disutilities so that they are sorted in the ascending order,
3. Calculate the solution y defined in Proposition 4.3.2,
4. The value $y_{sw}^{quad}(x;\beta)$ is the component of y that corresponds to the disutility V_s initially indexed by w.

As pointed out in equation (4.12), the lower response can be viewed as a projection on the simplex. Five algorithms to compute the projection are provided in [Con16]. The first one, applied to the projection $\mathrm{Proj}_{\Delta_{W+1}}\left(-\frac{\beta}{2}V_s\right)$, allows us to recover the response found in Proposition 4.3.2. Other algorithms are faster but do not contain such a clear interpretation that customers select disutilities with the lowest values.

The next proposition studies the impact of β on the solution, using majorization preorder:

Definition 4.3.1 (Majorization,[MOA11]). For a vector $a \in \mathbb{R}^d$, we denote by $a^\downarrow \in \mathbb{R}^d$ the vector with the same components, but sorted in descending order. Given $a, b \in \Delta_d$, we say that a majorizes b from below written $a \succ b$ iff

$$\sum_{i=1}^{k} a_i^\downarrow \geqslant \sum_{i=1}^{k} b_i^\downarrow \quad \text{for } k = 1, \dots, d \ .$$

Proposition 4.3.3 (Majorization ordering)
Let $u \in \mathbb{R}^n$. For any $0 \leqslant \alpha \leqslant \beta$, $\mathrm{Proj}_{\Delta_n}(\alpha u) \prec \mathrm{Proj}_{\Delta_n}(\beta u)$. As a consequence, the customer decision $y_s^{quad}(x;\beta)$ is increasing according to β for the *majorization preorder*, i.e., for $0 \leqslant \alpha \leqslant \beta$,

$$y_s^{quad}(x;\alpha) \prec y_s^{quad}(x;\beta) \ .$$

Proof. Using Proposition 4.3.2 (or [Con16], Algorithm 1]), if $y = \mathrm{Proj}_{\Delta_n}(u)$, then

$$y = \max\{u - c_\kappa(u^\downarrow), 0\} \ , \quad y^\downarrow = \max\{u^\downarrow - c_\kappa(u^\downarrow), 0\} \ ,$$

with $c_k(u) = \frac{1}{k}\left[\sum_{i=1}^{k} u_i - 1\right]$ and $\kappa = \max\left\{k \in [n] \ \mid \ c_k(u^\downarrow) \leqslant u_k\right\}$.

As we are looking to majorization property, we can suppose w.l.o.g. that u is ordered in a decreasing fashion, see Definition 4.3.1, so that the projection vector are also ordered. Let $\kappa_\alpha := \max\{k \in [n] \ \mid \ c_k(\alpha u) \leqslant \alpha u_k\}$ and $\kappa_\beta := \max\{k \in [n] \ \mid \ c_k(\beta u) \leqslant \beta u_k\}$. First note that $\kappa_\beta \leqslant \kappa_\alpha$ since $\kappa_\alpha = \max\{k \ \mid \ \frac{1}{k}\sum_{i=1}^{k} u_i - u_k \leqslant \frac{1}{k\alpha}\}$ (idem for κ_β). Let also $y = \mathrm{Proj}_\Delta(\alpha u)$ and $z = \mathrm{Proj}_\Delta(\beta u)$. Then, for $d \geqslant \kappa_\beta$,

$$\sum_{i=1}^{d} z_i = 1 \geqslant \sum_{i=1}^{d} y_i \ .$$

For $d < \kappa_\beta (\leqslant \kappa_\alpha)$,

$$\begin{aligned}
\frac{1}{d}\sum_{i=1}^{d} z_i - \frac{1}{d}\sum_{i=1}^{d} y_i &= (\beta - \alpha)\frac{1}{d}\sum_{i=1}^{d} u_i + c_{\kappa_\alpha}(\alpha u) - c_{\kappa_\beta}(\beta u) \\
&= \underbrace{(\beta - \alpha)\left[\frac{1}{d}\sum_{i=1}^{d} u_i - \frac{1}{\kappa_\beta}\sum_{i=1}^{\kappa_\beta} u_i\right]}_{\geqslant 0} + \underbrace{\frac{1}{\kappa_\alpha}\left[\sum_{i=1}^{\kappa_\alpha} \alpha u_i - 1\right]}_{:= c_{\kappa_\alpha}(\alpha u)} - \underbrace{\frac{1}{\kappa_\beta}\left[\sum_{i=1}^{\kappa_\beta} \alpha u_i - 1\right]}_{:= c_{\kappa_\beta}(\alpha u)} \\
&\geqslant \sum_{k=\kappa_\beta}^{\kappa_\alpha - 1} c_{k+1}(\alpha u) - c_k(\alpha u) = \sum_{k=\kappa_\beta}^{\kappa_\alpha - 1} \frac{1}{k}\left[\alpha u_{k+1} - c_{k+1}(\alpha u)\right] \geqslant 0
\end{aligned}$$

where the last inequality comes from the definition of $c_k(\alpha u)$ for $k \leqslant \kappa_\alpha$. □

Proposition 4.3.3 has a strong qualitative interpretation: the more the regularization parameter β is, the less "diversified" the choice will be. As an example, in Δ_n,

$$\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \prec \left(\frac{1}{n-1}, \dots, \frac{1}{n-1}, 0\right) \prec \cdots \prec \left(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0\right) \prec (1, 0, \dots, 0) \, .$$

Corollary 4.3.1 gives a quantitative information about when β is sufficiently large so that the customer response concentrates on a single contract. It shows a soft threshold effect at a finite rationality ($\beta < \infty$): we allow the variables y_{sw} to be fractional values, but they will concentrate on a unique contract per segment if the disutilities are sufficiently separated. This effect only occurs asymptotically ($\beta = \infty$) in the logit model. Corollary 4.3.1 also has an intuitive economic interpretation. In fact, one can link the estimation of the regularization intensity β with the minimal gap (in €) above which the decision coincides with the best deterministic one (probability one to choose the offer giving the highest utility). For example, the $2/\beta$ threshold in (4.15) reveals that a value $\beta = 0.2$ then corresponds to the minimal difference of 10€ to recover a binary decision.

For a given price strategy x, the leader profit function $\pi^{quad}(x;\beta)$ is then defined as

$$\pi^{quad}(x;\beta) := \sum_{s\in[S]} \rho_s \sum_{w\in[W]} (\theta_{sw}(x) - C_{sw}) y^{quad}_{sw}(x;\beta) \tag{4.16}$$

where $y^{quad}(x;\beta)$ is defined as explained in corollary 4.3.2. The problem (q-BP) is therefore the maximization of the function π^{quad} over X.

4.3.2 Price complex and convergence to the deterministic model

[BK19] have introduced a geometric approach to analyze the response of agents to prices, in a discrete choice model. They showed that the deterministic response is governed by a polyhedral complex: all prices in a given cell yield the same response. Here, we generalize this approach to continuous responses, since in our regularized model, responses do not concentrate anymore on a single contract. However, the closed-form formula we found for the lower response highlights the sparsity in terms of customer choices. In fact, in a feasible solution, only few contracts have positive probabilities to be chosen by a segment s (we call them *active contracts*). Now, all the prices in a given cell yield (different) responses encoded with the same "sparcity pattern" i.e., the responses share the same set of active contracts.

Definition 4.3.2.
1. A matrix $A \in \{0,1\}^{S\times(W+1)}$ is called a *pattern*. We denote by $|A_s|$ the number of positive coefficients in row s, and $|A| = \sum_{s\in[S]} |A_s|$ the total number of positive coefficients of A.
2. We denote by $X(A;\beta)$ the price strategies that have an active-contract set corresponding to the pattern A, i.e.,

$$X(A;\beta) := \left\{ x \in X \mid \mathbb{1}_{(y^{quad}_{sw}(x;\beta)>0)} = A_{sw},\ \forall s,w \right\} \ .$$

Thus, the set of active contracts stays unchanged on the set $X(A;\beta)$. We call this price region a *unique pattern region* (UPR).

The UPRs are not closed since we look at the prices that give a positive probability. Thus, we define $\overline{X}(A;\beta)$ to be the closure of the UPR $X(A;\beta)$.

Definition 4.3.3.
1. A pattern A is said to be *feasible* if $\overline{X}(A;\beta)$ is non-empty, and $\mathcal{A}^\beta \subseteq \{0,1\}^{S\times(W+1)}$ is then the set of feasible patterns.
2. A *pure* pattern A is a pattern containing only pure strategies i.e., each segment has a unique active contract ($|A| = S$). The other patterns are called *mixed* patterns ($|A| > S$).
3. A *price complex cell* is a non-empty set $P \subseteq X$ such that there exist $A^1,\dots,A^k \in \mathcal{A}^\beta$, with $k \geqslant 1$, satisfying $P = \bigcap_{1\leqslant i\leqslant k} \overline{X}(A^i;\beta)$.

4. The *price complex* is the collection of all price complex cells.

Proposition 4.3.4 (Characterization of price complex cells)
For any pattern $A \in \mathcal{A}^\beta$ and any $\beta > 0$, the UPR $X(A;\beta)$ is defined as $X(A;\beta) = \overline{X}^0(A;\beta) \cap X^1(A;\beta)$ where

$$\overline{X}^0(A;\beta) = \left\{ x \in X \,\middle|\, \begin{array}{l} \forall s, w, \text{ if } A_{sw} = 0, \\ \quad |A_s| V_{sw}(x) \geqslant 2\beta^{-1} + \sum\limits_{w' \mid A_{sw'}=1} V_{sw'}(x) \end{array} \right\}, \tag{4.17a}$$

$$X^1(A;\beta) = \left\{ x \in X \,\middle|\, \begin{array}{l} \forall s, w, \text{ if } A_{sw} = 1, \\ \quad |A_s| V_{sw}(x) < 2\beta^{-1} + \sum\limits_{w' \mid A_{sw'}=1} V_{sw'}(x) \end{array} \right\}. \tag{4.17b}$$

where $|A_s|$ corresponds to the number of active contracts for s and $V_{sw}(x)$ is defined as in Corollary 4.3.2. As a consequence, $\overline{X}(A;\beta) = \overline{X}^0(A;\beta) \cap \overline{X}^1(A;\beta)$ where $\overline{X}^1(A;\beta) := \mathrm{cl}\left(X^1(A;\beta)\right)$, obtained by weakening the inequalities (4.17b).

Proof. Given a pattern $A \in \mathcal{A}^\beta$ and a $\beta > 0$, we can assume w.l.o.g. that for any segment s the disutilites are sorted in ascending order so that the active contracts are the first $|A_s|$ ones. First, we consider a price strategy $x \in X(A;\beta)$. Using the notation of Proposition 4.3.2, $\tau = |A_s|$ and equation (4.14) gives us exactly that $x \in \overline{X}^0(A;\beta) \cap X^1(A;\beta)$. Reciprocally, we suppose that $x \in \overline{X}^0(A;\beta) \cap X^1(A;\beta)(4.17)$ is satisfied i.e., $V_{sw} < c_{s,|A_s|}$ for $w < |A_s|$ and $V_{sw} \geqslant c_{s,|A_s|}$ for $w \geqslant |A_s|$. Then, $\tau = |A_s|$ and $x \in X(A;\beta)$. $\square$

Theorem 4.3.1
The collection of price complex cells constitutes an $|X|$-dimensional polyhedral complex, and the $|X|$-cells are closures of UPRs i.e., $\overline{X}(A;\beta)$ for some pattern $A \in \mathcal{A}^\beta$.

Proof. It is clear that the collection of price complex cells covers the space X. Besides, from the definition of a cell, the intersection of two cells P and P' is again a price complex cell or is empty. Finally, Proposition 4.3.4 gives us a characterization of the cells with linear inequalities, therefore the intersection of P with another P' is then characterized by the same inequalities as P but with some of them saturated. Hence, the intersection is a common face of P and P'. $\square$

We now study the asymptotic behavior of the price complex ($\beta \to \infty$) and show how it embeds in the deterministic complex introduced in [BK19]. To this end, we first denote by $\overline{X}(A;\infty)$ the polytope defined by the same inequalities as in $\overline{X}(A;\beta)$ setting $\beta^{-1} = 0$ (idem for $\overline{X}^0$ and $\overline{X}^1$), and by $\mathcal{A}^\infty$ the set of patterns inducing a non-empty $\overline{X}(A;\infty)$. We next make use of the notion of Painlevé-Kuratowski limits of sets. We refer to [RW09, Chapter 4] for background on this notion, including the definition and properties of upper and lower limits. We first prove the following preliminary lemma:

Lemma 4.3.2
Consider two sequences of polyhedra P_β^+ and P_β^- defined as $P_\beta^\pm := \left\{x \in X : Ax \leqslant b \pm \beta^{-1} e\right\}$ (e is the all-ones vector), and the limit case $P := \{x \in X : Ax \leqslant b\}$. Then, $P_\beta^+ \xrightarrow[\beta\to\infty]{} P$

and $\lim_\beta P^-_\beta \subseteq P$. Moreover, if $\mathrm{Int}(P) \neq \varnothing$, $P^-_\beta \xrightarrow[\beta\to\infty]{} P$.

Proof. Throughout the proof, we consider a sequence (β_n) converging to ∞, and the notation $P^\pm_n$ has to be understood as $P^\pm_{\beta_n}$.

The two monotone sequences have a limit: $\lim_n P^+_n = \bigcap_n P^+_n$ and $\lim_n P^-_n = \bigcup_n P^-_n$, see [RW09, Exercise 4.3], it remains to prove that this limit coincides with P. Two first inclusions come with the definition of the sequences: $\lim_n P^-_n \subseteq P$ and $P \subseteq \lim_n P^+_n$.

Let us consider $x \notin P$. If $x \in X \backslash P$, then there exists a row i such that $A_i x = b_i + \epsilon$ where $\epsilon > 0$. Therefore, for $\beta_n \geqslant \epsilon^{-1}$, $x \notin P^+_n$. Otherwise, if $x \notin X$, x cannot be in any P^+_n. In any case, $x \notin P \Rightarrow x \notin \lim_n P^+_n$, and therefore $\lim_n P^+_n \subseteq P$.

We now assume that $\mathrm{Int}(P) \neq \varnothing$. For any given $x \in P$, let us define the sequence $x_n := \mathrm{Proj}_{P^-_n}(x)$. Since the $P^- \nearrow$, the distance $\|x_n - x\|$ is a decreasing sequence bounded from below by 0 and converges to a distance $d \geqslant 0$. Suppose now that $d > 0$, then for any unitary vector u, $x + du \notin P^-_n, n \in \mathbb{N}$. Besides, there exists $0 \leqslant d' \leqslant d$ and a unitary vector v such that $x + d'v \in \mathrm{Int}(P)$. Defining $y = x + d'v$, we obtain that $y \in \mathrm{Int}(P)$ and $y \notin P^-_n, n \in \mathbb{N}$. As it belongs to the interior of P, $Ay \leqslant b - \epsilon e, \epsilon > 0$ and for any $\beta_n \geqslant \epsilon^{-1}$, $y \in P^-_n$. This yields a contradiction: d must be equal to 0, and therefore $x_n \to x$. To conclude, for any $x \in P$, we can exhibit a sequence of points $x_n \in P^-_n$ converging to x, so $P \subseteq \lim_n P^-_n$. ☐

We are now ready to prove the following result:

Proposition 4.3.5 (Convergence)

For any pattern A, $\limsup_{\beta\to\infty} \overline{X}(A;\beta) \subseteq \overline{X}(A;\infty)$. Moreover, if $\mathrm{Int}\left(\overline{X}(A;\infty)\right) \neq \varnothing$,

$$\overline{X}(A;\beta) \xrightarrow[\beta]{} \overline{X}(A;\infty) \ .$$

Proof. Using Lemma 4.3.2, one can obtain the following inclusions:

$$\begin{aligned}\limsup_\beta \overline{X}(A;\beta) &= \limsup_\beta \left(\overline{X}^0(A;\beta) \cap \overline{X}^1(A;\beta)\right) \\ &\subseteq \lim_\beta \overline{X}^0(A;\beta) \cap \lim_\beta \overline{X}^1(A;\beta) \subseteq \overline{X}^0(A;\infty) \cap \overline{X}^1(A;\infty) \ .\end{aligned}$$

Moreover, if $\mathrm{Int}\left(\overline{X}(A;\infty)\right) \neq \varnothing$, then $\lim_\beta \overline{X}^0(A;\beta) = \overline{X}^0(A;\infty)$, see Lemma 4.3.2. Besides, $\overline{X}^0(A;\infty)$ and $\overline{X}^1(A;\infty)$ cannot be separated, and therefore $\overline{X}^0(A;\beta) \cap \overline{X}^1(A;\beta) \xrightarrow[\beta]{} \overline{X}^0(A;\infty) \cap \overline{X}^1(A;\infty)$, see [RW09, Theorem 4.32c]. ☐

Lemma 4.3.3

For any pattern $A \in \mathcal{A}^\infty$, the asymptotic cell $\overline{X}(A;\infty)$ can be equivalently defined by the following system

$$\begin{aligned} \forall s, w, w', \text{if } A_{sw} = A_{sw'} = 1, \quad & V_{sw}(x) = V_{sw'}(x), \\ \text{if } A_{sw} = 1 \text{ and } A_{sw'} = 0, \quad & V_{sw'}(x) \geqslant V_{sw}(x). \end{aligned} \tag{4.18}$$

Proof. We first define the mean active disutility for a segment s as $\tilde{V}_s = \frac{1}{|A_s|} \sum_{w' \mid A_{sw'}=1} V_{sw'}$. Then, we know by (4.17b) that for any active contract w, $V_{sw} - \tilde{V}_s \leqslant \frac{2}{\beta}$. Denoting by V^+ and V^- the extreme disutilities of active contracts, we obtain $0 \leqslant V^+ - V^- \leqslant \frac{2}{\beta}$. At the limit, active contracts share a same disutility, equal to $\tilde{V}_s$. Besides, we also know from (4.17a) that for any inactive contract w, $V_{sw} \geqslant \frac{2}{\beta} + \tilde{V}_s$. At the limit, any inactive contract has disutilities greater than the active contracts. $\square$

Lemma 4.3.4

For any mixed pattern A, there exist $k > 1$ pure patterns $A^1, \dots, A^k$ such that

$$\overline{X}(A;\infty) = \bigcap_{1 \leqslant i \leqslant k} \overline{X}(A^i;\infty) \ .$$

Proof. Suppose that for a given segment s, $|A_s| = k$, then we can construct patterns $A^1, \dots, A^k$ such that A^i is a copy of A where the row s is replaced by 1 on the ith active contract, and 0 everywhere else. From the characterization (4.18), we obtain that $\overline{X}(A;\infty) = \bigcap_{1 \leqslant i \leqslant k} \overline{X}(A^i;\infty)$. Each pattern A^i has pure strategy for segment s. If there still exist mixed strategies for other segment, we can start again the transformation until all the patterns are pure. $\square$

At the limit $\beta = \infty$, each mixed pattern is a face of some pure patterns. The pure patterns are therefore sufficient to describe any cell.

Theorem 4.3.2 (Asymptotic cells and UPRs)

Let $A^1, \dots, A^k$ be k pure patterns, then

$$x \in P = \bigcap_{1 \leqslant i \leqslant k} \overline{X}(A^i;\infty) \iff \{A^1, \dots, A^k\} \subseteq \Psi(x) \ .$$

where $\Psi(x)$ is the set of optimistic best responses, defined in (4.3). Moreover, for any pure pattern A,

$$\operatorname{Int}\left(\overline{X}(A;\infty)\right) = \{x \in X : \{A\} = \Psi(x)\} \ .$$

Proof. The equivalence is a direct consequence of the Lemma 4.3.3. The equality also arises from this lemma: the set $\{x \in X : \{A\} = \Psi(x)\}$ is characterized by (4.18) with strict inequalities. $\square$

Theorem 4.3.2 establishes a link with the approach of [BK19]: we generalize the price complex to relaxed choices and the definition we introduce in Definition 4.3.3 is equivalent to their definition in the specific case $\beta = \infty$. Moreover, Baldwin and Klemperer define *unique demand region* (UDR) where the set $\Psi(x)$ has a unique element, and Theorem 4.3.2 proves that any pure UPR converges to the corresponding UDR. To illustrate Proposition 4.3.4, Figure 4.2 shows the complex cells for a single customer making a choice among two contracts from the company and one from a competitor. The deterministic complex was depicted in [BK19, Figure 1] or in [Eyt18] for bilevel models, and Figure 4.2 illustrates the generalization of the price complex to relaxed choices: note that new types of full-dimensional cells, representing choices concentrated on several contracts, appear.

The logit profit function has no good convexity properties in our context of a heterogeneous population. Thanks to the properties of the lower response and the notion of polyhedral complex, we can prove that its quadratic analog is more structured:

Lemma 4.3.5

For $K \geqslant N$, the function $J : x \in \mathbb{R}^N \mapsto \sum_{i=1}^N x_i^2 - \frac{1}{K}\left(\sum_{i=1}^N x_i\right)^2$ is convex.

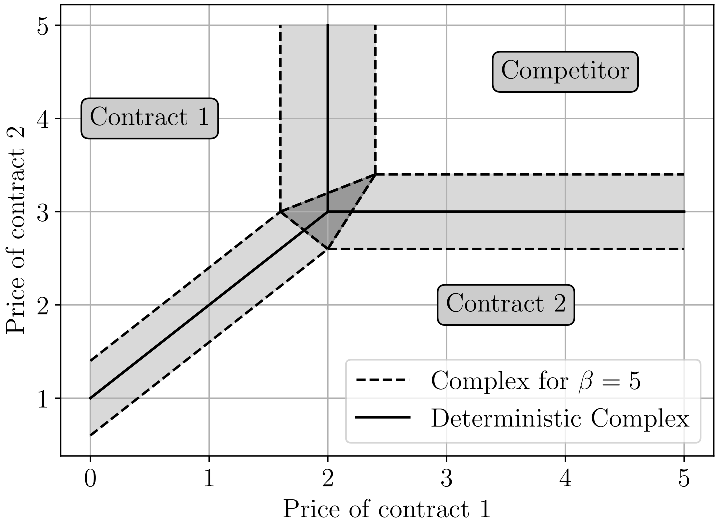


Figure 4.2: Price complex in a simple case.

For $\beta = \infty$ (deterministic case, solid line), the three cells correspond to the the choice of a unique contract (rectangles indicate the choice). For $\beta < \infty$, each line "splits" to create intermediate cells (mixed strategies). Pure strategies correspond to white zones, strategies mixing two contracts correspond to light gray zones and the strategy mixing all contracts corresponds to the dark gray zone.

Proof. The Hessian H of the function J is $H_{ij} = -2/K$ for $i \neq j$ and $H_{ii} = 2 - 2/K$. Using the Gershgorin circle theorem, any eigen value λ_i has to verify $|\lambda_i - (2 - 2/K)| \leqslant \sum_{j \neq i} 2/K$. Therefore, $\lambda_i \geqslant 2 - 2N/K$ and we deduce that all eigen values of H are nonnegative. $\square$

Theorem 4.3.3 (Profit decomposition)
The quadratic leader profit function $\pi^{quad}(x; \beta)$ is continuous. Moreover, the problem ($q\beta$-BP) is equivalent to the following problem

$$\max_{A \in \mathcal{A}^\beta} \left\{ \varphi(A; \beta) := \max_{x \in \overline{X}(A;\beta)} \pi^{quad}(x; \beta) \right\} \tag{4.19}$$

where $\pi^{quad}(x; \beta)$ is concave on each price complex cell $\overline{X}(A; \beta)$, defined in Proposition 4.3.4.

Proof. The continuity of the lower response suffices to ensure the continuity of π^{quad}. The difficulty lies in the concave foundation. Because the profit function is a sum over the segments, we may assume that there is only one segment s. Let us consider a feasible pattern $A \in \mathcal{A}^\beta$. On the cell $\overline{X}(A; \beta)$ associated with this pattern, the profit function is expressed as

$$J_s^A(x) := \sum_{w \in [W] \,|\, A_{sw} = 1} (\theta_{sw}(x) - C_{sw}) y_{sw}^{quad}(x; \beta) \ .$$

To keep compact notation, we define $\mathcal{W}_s^A := \{w \in [W] \,|\, A_{sw} = 1\}$, and $V_{sw} := \theta_{sw}(x) - R_{sw}$ for

$w \in [W]$ and $V_{s0} = 0$. Using corollary 4.3.2, we can rewrite J^A as

$$\begin{aligned}
J_s^A(x) &= \frac{\beta}{2} \sum_{w \in \mathcal{W}_s^A} (V_{sw} + R_{sw} - C_{sw})(c_{s,|A_s|} - V_{sw}) \\
&= \frac{\beta}{2} \sum_{w \in \mathcal{W}_s^A} (R_{sw} - C_{sw})(c_{s,|A_s|} - V_{sw}) - \frac{\beta}{2} \left[\sum_{w \in \mathcal{W}_s^A} V_{sw}^2 - c_{s,|A_s|} \sum_{w \in \mathcal{W}_s^A} V_{sw} \right] \\
&= L - \frac{\beta}{2} \left[\sum_{w \in \mathcal{W}_s^A} V_{sw}^2 - \frac{1}{|A_s|} \left(\sum_{w \in \mathcal{W}_s^A} V_{sw} \right)^2 \right]
\end{aligned}$$

where $L = \frac{1}{|A_s|} \sum_{w \in \mathcal{W}_s^A} V_{sw} + \frac{\beta}{2} \sum_{w \in \mathcal{W}_s^A} (R_{sw} - C_{sw})(c_{s,|A_s|} - V_{sw})$ denotes the linear part. The set $\mathcal{W}_s^A$ has a cardinality of $|A_s|$ or $|A_s| - 1$ depending on if the no-purchase option appears in the first $|A_s|$ disutilities. Therefore, by Lemma 4.3.5, J_s^A is concave in V_s, and thus is concave in x since the functions θ are linear. Finally, exploring $\overline{X}(A;\beta), A \in \mathcal{A}^\beta$ is sufficient to cover the whole space X. □

Theorem 4.3.3 paves the way to enumerative scheme resolutions: it shows that the problem can be polynomially solved on each cells of the polyhedral complex, and if all the cells are explored it gives a global optimum. Nonetheless, it could be very cumbersome (especially for low β values).

4.3.3 QPCC Reformulation

As in the deterministic case, the model can be recast into a single-level program with complementarity constraints using the KKT conditions. Moreover, we are able to replace the bilinear terms using manipulations on the constraints:

Theorem 4.3.4

The problem ($q\beta$-BP) is equivalent to the following concave QPCC problem

$$\begin{aligned}
\max_{x \in X, \mu \in \mathbb{R}^S, \bar{y}} \quad & \sum_{s \in [S]} \rho_s \mu_s + \rho_s \langle R_s - C_s, y_s \rangle_W - 2\beta^{-1} \rho_s \|\bar{y}_s\|_{W+1}^2 \\
\text{s. t.} \quad & 0 \leqslant y_{sw} \perp \theta_{sw}(x) - R_{sw} + 2\beta^{-1} y_{sw} - \mu_s \geqslant 0, \ \forall s, w \qquad (q\beta\text{-}QPCC) \\
& 0 \leqslant y_{s0} \perp 2\beta^{-1} \bar{y}_s - \mu_s \geqslant 0, \ \forall s \\
& \bar{y}_s \in \Delta_{W+1}, \ \forall s
\end{aligned}$$

Proof. The KKT optimality condition have been detailed in (4.13). One can remark that the variable λ can be removed to obtain the KKT system of ($q\beta$-$QPCC$). We then reformulate the objective by using the constraints: for a given $s \in [S]$,

$$\begin{aligned}
\langle \theta_s(x), y_s \rangle_W &= \langle \mu_s e_W + R_s, y_s \rangle_W - 2\beta^{-1} \|y_s\|_W^2 \\
&= \mu_s - \mu_s y_{s0} + \langle R_s, y_s \rangle_W - 2\beta^{-1} \|y_s\|_W^2 \quad .
\end{aligned}$$

Finally, the objective in ($q\beta$-$QPCC$) is obtained using the complementarity constraint on the no-purchase option: $\mu_s y_{s0} = 2\beta^{-1} y_{s0}^2$. □

As in the deterministic case, we can replace the complementarity constraints in ($q\beta$-BP) by Big-M constraints to obtain a mixed-integer quadratic problem (MIQP). However, the introduction of binary variables is unavoidable since the existence of a solution with integer lower response is no longer true:

$$\begin{aligned}
\max_{x\in X,\mu\in\mathbb{R}^S,\bar{y},z} \quad & \sum_{s\in[S]} \rho_s\mu_s + \rho_s \langle R_s - C_s, y_s\rangle_W - 2\beta^{-1}\rho_s \|\bar{y}_s\|^2_{W+1} \\
\text{s.t.} \quad & 0 \leqslant \theta_{sw}(x) - R_{sw} + 2\beta^{-1}y_{sw} - \mu_s \leqslant M_{sw}(1-z_{sw}),\ \forall s, w \\
& 0 \leqslant 2\beta^{-1}\bar{y}_s - \mu_s \leqslant M_{s0}(1-z_{s0}),\ \forall s \\
& y \leqslant z \\
& \bar{y}_s \in \Delta_{W+1},\ z\in\{0,1\}^{W+1},\ \forall s
\end{aligned} \tag{4.20}$$

QPCC problems have been recently studied, using conic relaxations – [Den+17; ZX19] – or logical Benders – [BMP13; Jar+20]. In the latter, they introduce the notion of *complementarity piece* defined by a valuation of the binary vector z. The complementarity pieces of ($q\beta$-$QPCC$) coincide with the cells $\overline{X}(A;\beta)$ of the price complex (4.17): admissible valuations of z define feasible patterns, and vice versa.

4.3.4 Comparison with the logit model

Quadratic and logit regularizations share a parameter β, interpreted as a rationality parameter. It will be convenient to replace the regularization parameter β in the quadratic model by $\beta' = \beta e/4$, leaving the value β in the logit model. In fact, the minimum of $\frac{1}{\beta}y(y-1)$ is $-\frac{1}{4\beta}$ whereas the minimum of $\frac{1}{\beta}y\log(y)$ is $-\frac{1}{e\beta}$, and so this choice of β equalizes the minimal intensity of the regularization term. To have a better intuition on the differences and similarities between the logit and quadratic regularization, we study a simple case where there is one single-attribute contract and five customers. We provide in Fig. 4.3 the leader profit as a function of the contract price for multiple configurations (the optimistic version, the quadratic version and the logit version for two values of β). The behavior of the deterministic and logit profit have already

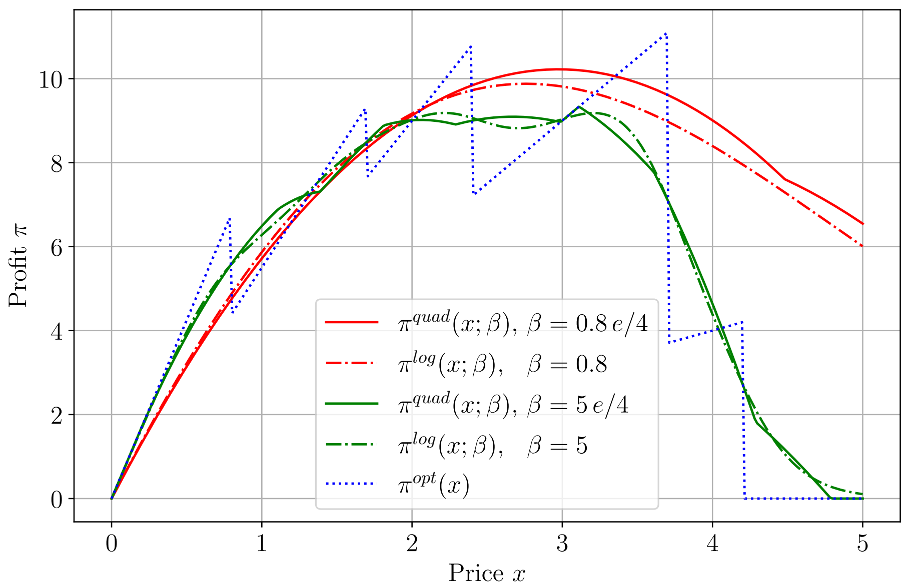


Figure 4.3: Comparison of profit functions π^{opt}, π^{log} and π^{quad}

been compared in another context in [GMS15]. We now include the quadratic model in this comparison. The following properties of profit functions can be identified:

- The deterministic profit is piecewise linear but contains discontinuities that arise when two contracts share the same minimal disutilities (here between the only contract and the no-purchase option). The optimal profit is always attained at such a frontier price, leading to an instability: for this specific case, the optimal deterministic profit is higher than 11 and is achieved for $x = 3.7$. Nevertheless, a price of $x = 3.71$ induces a profit lower than 4.
- The logit regularization smooths the deterministic profit function while maintaining its global shape for β large enough. Nonetheless, the function is non-convex and we can observe for $\beta = 5$ two local maxima.
- The quadratic regularization and its logit analog share the same behavior: in fact, the shape is very similar for both values of β. The difference lies in the structure of the quadratic model: the profit function is piecewise concave, see Theorem 4.3.3.

4.3.5 Comparison with the primal-dual regularization of [SSC17]

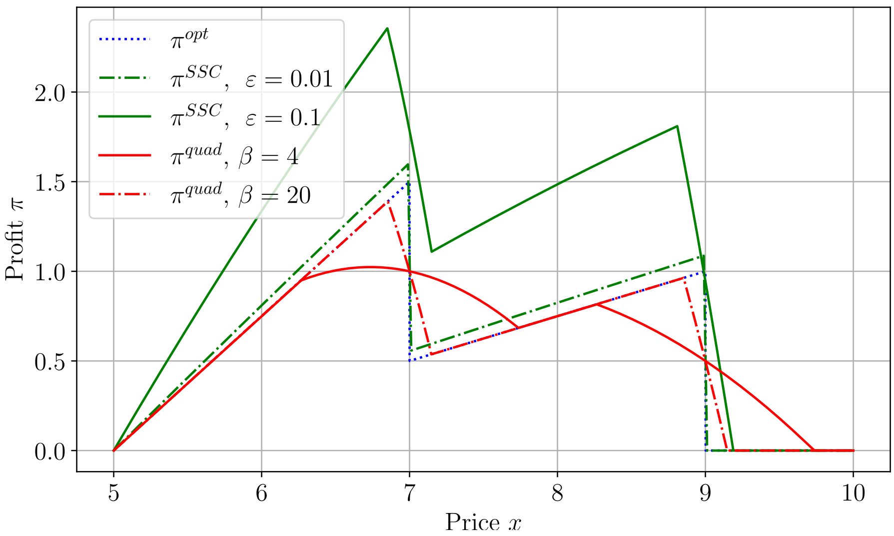


Figure 4.4: Comparison between the quadratic regularization

Our primal quadratic regularization (π^{quad}) is drawn for $\beta = 4$ and $\beta = 20$ and the primal-dual quadratic regularization of [SSC17] (π^{SSC}) is drawn for $\varepsilon = 0.1$ and $\varepsilon = 0.01$.

Sun, Su and Chen [SSC17] have already used a quadratic regularization for the multi-product pricing problem. As we do in this chapter, they provide a closed-form formula and look at the leader profit function. The main difference between the two quadratic regularizations lies in the fact that our version is only a primal regularization whereas the regularization of [SSC17] is of a primal-dual nature. This leads to distinct regularized solutions – in particular, the ones that we obtain remain feasible solutions of the deterministic model. Figure 4.4 illustrates the profit functions obtained with the two versions, reusing Example 2 provided in [SSC17].

4.4 Local Search by Pivoting on the Price Complex

In the previous sections, we established geometrical properties of the quadratic regularization. In particular, Theorem 4.3.4 provides a direct formulation which allows us to find a global optimum via MIQP techniques. However, such methods are workable only up to a limited instance size, above which a good optimality gap cannot be obtained in reasonable time. Therefore, it is of interest to develop a local search method taking advantage of the structure highlighted in Theorem 4.3.3: finding the optimal solution is no more than finding the cell of the polyhedral complex containing this solution. Indeed, computing the optimum on a given cell reduces to a (simple) quadratic program. Given a cell, a neighbor cell can be obtained by reversing one of the inequalities (4.17), however, computing all the neighbor cells is computationally expensive (there is a large number of inequalities (4.17) and moreover some of them are redundant). Hence, we introduce a narrow neighborhood which selects specific neighbors obtained by reversing the inequalities associated to contracts near the active/inactive frontier, as these yields good candidates in the search for better solutions, see Algorithm 1.

Algorithm 1 `exploreGoodNeighbors`

Require: $A, x_A, \varphi_A = \pi^{quad}(x_A;\beta)$ ▷ x_A optimum on the initial pattern A
$A^*, x^*, \varphi^* \leftarrow A, x_A, \varphi_A$ ▷ A^* will be the best neighboring pattern
for $s = 1 \dots S$ **do**
  $A^- \leftarrow A$, $A^+ \leftarrow A$
  $w^- \leftarrow \max_{0\leqslant w \leqslant W} \{V_{sw} \mid A_{sw} = 1\}$ ▷ worst active contract
  $w^+ \leftarrow \min_{0\leqslant w \leqslant W} \{V_{sw} \mid A_{sw} = 0\}$ ▷ best nonactive contract
  $A^-_{s,w^-}, A^+_{s,w^+} \leftarrow 0, 1$ ▷ new patterns
  for $\dagger \in \{-,+\}$ **do**
    $x^\dagger, \varphi^\dagger \leftarrow$ solution of $\max_{x\in\overline{X}(A^\dagger;\beta)} \pi^{quad}(x;\beta)$
    if $\varphi^\dagger \geqslant \varphi^*$ **then**
      $A^*, x^*, \varphi^* \leftarrow A^\dagger, x^\dagger, \varphi^\dagger$ ▷ update best pattern
    end if
  end for
end for
return A^*, x^*, φ^*

More precisely, for a given feasible pattern A and for any segment s, we select two inequalities on which we will pivot:

(i) the inequality (4.17b) of index (s, w^-) where w^- is the active contract with the greatest disutility for s (i.e., with the lowest positive probability y_{sw}),

(ii) and the inequality (4.17a) of index (s, w^+) where w^+ is the non-active contract with the lowest disutility.

By pivoting, we mean that, starting from this feasible pattern A, we consider a new cell, in which all the inequalities in (4.17a) and (4.17b) stay unchanged, except the two ones of indices (s, w^-) and (s, w^+) that are reversed. This leads to a new pattern.

Although this strategy does not explore the whole neighborhood (it only changes $2S$ inequalities among WS), pivoting on the selected inequalities is likely to produce relevant new cells. We will consider two methods for exploring these cells:

1. computing $\varphi(\,\cdot\,;\beta)$ for each of the $2S$ neighboring patterns by solving $2S$ quadratic programs, see (4.19), and returning the best pattern A' with its value $\varphi(A';\beta)$ (it could be the initial pattern if no improvement was made),

2. or solving the MIQP (4.20) where the only unfixed binary variables z are the $2S$ variables indexed by the selected inequalities (the other variables z are equal to the current pattern values) and returning the pattern A' obtained by the solver with its value $\varphi(A';\beta)$.

The second option is computationally more expensive (as it relies on a MIQP) but it explores a wider neighborhood, since several of the S-groups of 2 inequalities can be reversed in a single step.

Iterating the procedure `exploreGoodNeighbors` (Algorithm 1) produces a local search, which always terminates because the number of patterns is finite and we continue only if we found a better pattern than the previous one.

> **Remark 4.4.1**
> In Algorithm 1, the exploration runs along segments, but it could also be made in the reversed order (loop on the contracts and selection of the worst active / best nonactive segments). It appears in the numerical tests that the latter option is less efficient.

The local search ends up with a local optimum in the sense that there is no neighbor (achievable by `exploreGoodNeighbors`) that produces a better solution. Then, to improve this solution, we need to consider a larger neighborhood. This is the object of the procedure `MIQP_restart`, described in Algorithm 2, in which we construct a small MIQP, fixing binary variables, except for the following ones:

(i) γ^S segments: for such a segment s, the variables $z_{s,w}$ in (4.20) become free for all $w \in 0 \dots W$; in other words, the whole row s in the pattern may be changed;

(ii) γ^W contracts: for such a contract w, the variables $z_{s,w}$ in (4.20) become free for all $s \in [S]$; in other words, the whole column w in the pattern may be changed;

(iii) every variable $z_{s'w'}$ with $s' \neq s$ and $w' \neq w$ is made free with probability $\sigma \in [0, 1]$.

This restart procedure uses a pattern as input and ends either with this pattern or a better one if the MIQP has found such a pattern.

Algorithm 2 `MIQP_restart`

Require: A ▷ initial pattern
1: Select γ^S segments, γ^W contracts and coefficients (s, w) with probability σ
2: Constrain z to be equal to A, except for the chosen segments, contracts and coefficients
3: $A^*, x^*, \varphi^* \leftarrow$ optimum of (4.20) with the additional constraints on z.
4: **return** A^*, x^*, φ^*

The complete heuristic (Algorithm 3), which we call *Quadratic Search on Price Complex* (`QSPC`), alternates between the local search and the restart phase until no progress is made, i.e., several iterations do not have produced any improvement.

We provide a graphical illustration on Figure 4.5 showing the path (in terms of cells) achieved by the algorithm through the iterations, and the final cell with the corresponding customers response.

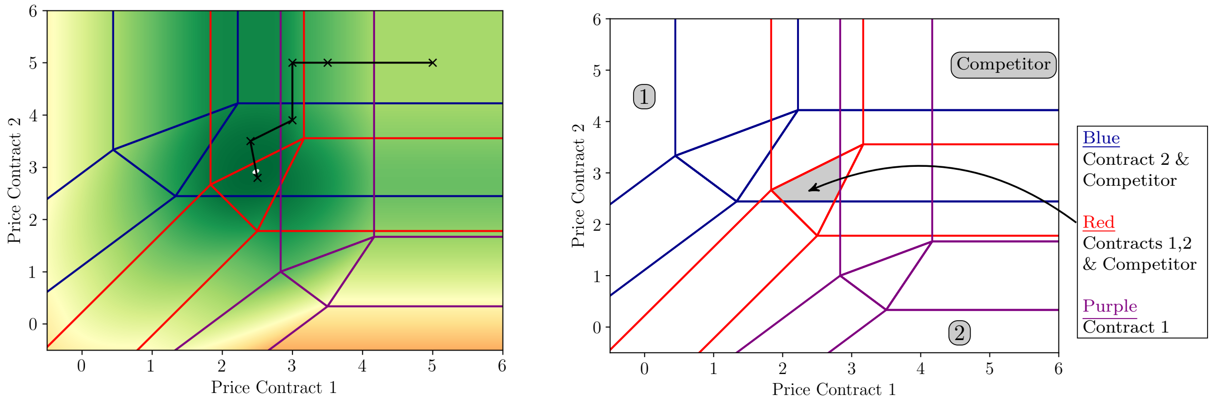
4.5 Performance analysis of the proposed method

The pre and post processing algorithms are implemented in `Python 3.7`, whereas the optimization methods are implemented in `C++` for numerical efficiency. Besides, we use `Cplex v12.10` [IBM09] as a MIQP solver and the tests are performed on a laptop `Intel Core i7 @2.20GHz` $\times$ 12. We ran `Cplex` on 4 threads.

Algorithm 3 Quadratic Search on Price Complex (`QSPC`)

Require: $A, x_A, \varphi_A = \pi^{quad}(x_A; \beta), r_{max}$ ▷ x_A optimum on the initial pattern A

$r \leftarrow 0$
$A^*, x^*, \varphi^* \leftarrow A, x_A, \varphi_A$ ▷ A^* will be the best pattern found
while $r < r_{max}$ **do** ▷ $r = \#$ restarts without improvement
  if $r = 0$ **then**
    $opt_{loc} \leftarrow$ false
    while opt_{loc} is false **do** ▷ Until a local optimum is found
      $A', x_{A'}, \varphi_{A'} \leftarrow$ `exploreGoodNeighbors`(A^*, x^*, φ^*)
      $opt_{loc} \leftarrow (A' = A^*)$
      $A^*, x^*, \varphi^* \leftarrow A', x_{A'}, \varphi_{A'}$
    end while
  end if
  $A', x_{A'}, \varphi_{A'} \leftarrow$ `MIQP_restart`(A^*)
  if $A' = A^*$ **then**
    $r \leftarrow r + 1$
  else
    $A^*, x^*, \varphi^*, r \leftarrow A', x_{A'}, \varphi_{A'}, 0$ ▷ Update best pattern
  end if
end while
return A^*, x^*, φ^*



(a) The algorithm starts at the upper right cell (every segment chose the competitor), and evolves through the cells to finally stop at a local maximum (here it is also the global one).

(b) For each customer, we show on the right the active contracts (chosen with positive probability).

Figure 4.5: Graphical illustration of the algorithm with $S = 3$ customers (blue, red and purple) and $W = 2$ contracts of $H = 1$ price attribute.

4.5.1 Comparison with implicit method

Another way to solve the model ($q\beta$-BP) is from the profit-maximization point of view, considering directly the nonsmooth problem

$$\max_{x \in X} \pi^{quad}(x; \beta) \tag{4.21}$$

where the function $\pi^{quad}(\,\cdot\,; \beta)$ is defined in (4.16). Taking advantage of the lower response uniqueness to end in a nonsmooth problem – where lower variables are functions of the upper ones – constitutes the basis of implicit methods for bilevel problems, see [KLM20].

Implicit methods require an oracle able to evaluate the objective function for any given point. Therefore, the explicit calculation of the lower response given by corollary 4.3.2 turns out to be essential in order to design the oracle. Powerful algorithms are already available, and we focus on *Covariance matrix adaptation evolution strategy* (`CMA-ES`, [Han06; Han+10]). In our problem the search space X has a reasonable dimension $(W \times H)$. Therefore, we can expect `CMA-ES` to find good solutions. For the numerical tests, we used an existing library available in `C++`[1].

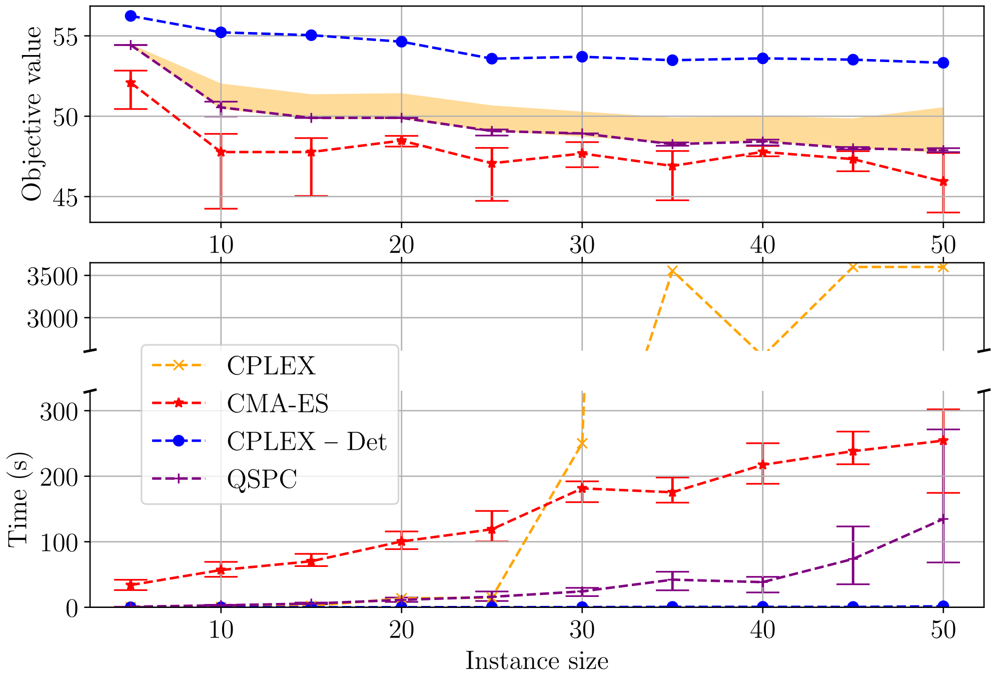


Figure 4.6: Numerical results with `CPLEX`, `CMA-ES`, `QSPC`.

The upper graph shows the objective value and the lower graph shows the resolution time for a segments number S varying between 5 to 50 and a β fixed to 0.5. For heuristic methods, five tries have been done and vertical lines indicate the least and the greatest value. The final gap obtained with the quadratic method is represented with a yellow zone (between the best solution and the best upper bound). For comparison, results of the deterministic model (`CPLEX - Det`) are given.

	`CMA-ES`	`CPLEX`	`QSPC`	`CPLEX - Det`
Problem	$(q\beta\text{-}BP)$			$(o\text{-}BP)$
Method	Section 4.5.1	Eq.(4.20)	Algorithm 3	$(o\text{-}KKT)$
Parameters	$\sigma = 0.005$ $\lambda = 1000$	MIP Gap : 3% Max time: 3600s	$\sigma = 0.05$ $\gamma^S = \gamma^W = 1$	MIP Gap: 1% Max time: 3600s

Table 4.1: Methods used in the numerical tests

Figure 4.6 shows the performances of methods listed in Table 4.1. The numerical tests highlight the combinatorial explosion induced by the direct resolution of the quadratic model with `CPLEX` for a finite β. The critical size seems to be around 30 segments on our data set. In contrast, the deterministic value is very fast to obtain up to 50 segments. This emphasizes the need of heuristics to rapidly obtain good solutions of the quadratic model.

The method `CMA-ES` is rather suitable for very large instances. In fact, the algorithm explores

[1] `https://github.com/CMA-ES/libcmaes`

the domain X which does not depend on the number of segments S, and the time to compute the lower response (by Proposition 4.3.2) linearly increases in S. The overall resolution time of CMA-ES has therefore an affine growth in the number of segments. Besides, the best solution found by CMA-ES seems to edge closer to optimum as the size grows. Increasing the number of segments dwindles the weight of each one in the objective, that tends to smooth the profit function and, as a consequence, facilitates CMA-ES in the resolution.

The great power of QSPC is to systematically find very good solutions (no large variance of the optimal value), even for large instances. Of course, this is only possible because we exploit the special geometry of our problem (as opposed to a generic algorithm like CMA-ES). Concerning resolution time, QSPC is also faster. However, QSPC becomes computationally more expensive as the number of segments increases, since it involves the restart phase (solution of a MIQP problem).

Finally, this numerical study gives us an *a posteriori* way to know how many segments are needed to accurately represent the population. After 30 segments the objective value seems to reach a plateau: using more segments does not seem to add useful information (at least in terms of optimal value).

4.5.2 Comparison with NLP solvers

Non-linear programming (NLP) constitutes a third alternative – with implicit methods and combinatorial methods – in the resolution of complementarity problems. Solvers have been designed/adapted to deal with these reformulations, see [KLM20] for a recent practical survey. For the numerical tests, we focus on two solvers:

(i) KNITRO [BNW06], which is a powerful commercial solver, able to recognize if the problem contains complementarity constraints to reformulate them as a non-linear inequalities,

(ii) filterMPEC [FL04a], which is an extension a Sequential Quadratic Programming (SQP) solver designed to solve MPECs. The theoretical material is described in [Ley06]. Note that we keep the scalar product form (compl_frm = 1) in all the runs.

Both solvers are available through the platform NEOS [CMM98].

Figure 4.7 compares the results obtained by KNITRO and filterMPEC with our heuristic. We still display the value returned by CPLEX to bound the optimality gap. The whole graph is computed with instances that slightly differ from the ones on Figure 4.6: the polytope X only contains the bounds on prices and not any other constraint. In fact, the solution returned by NLP methods violates the constraints by an ϵ and if the polytope X were more complicated than a box, it would require a finer post-processing to reconstruct a valid price vector x that exactly respects the inequalities/equalities of X.

The two NLP solvers are very fast to return a solution, either KNITRO or filterMPEQ, even if the time cannot be considered as a uniform indicator since the calculations were achieved on NEOS servers whereas QSPC was run on a personal computer. On these instances, QSPC always returns better solutions. In fact, only Clarke-Stationary points can be ensured by NLP solvers, see [KLM20] and the references therein. Of the two solvers, KNITRO seems to be the fastest, but we run it on 4 threads whereas filterMPEC uses a SQP algorithm which is difficult to parallelize.

4.6 Application to Electricity Pricing

4.6.1 Instance definition

In the numerical tests, we consider an electricity pricing problem: a power retailer has $W = 4$ different contracts that need to be optimized, each one depending on $H = 3$ coefficients

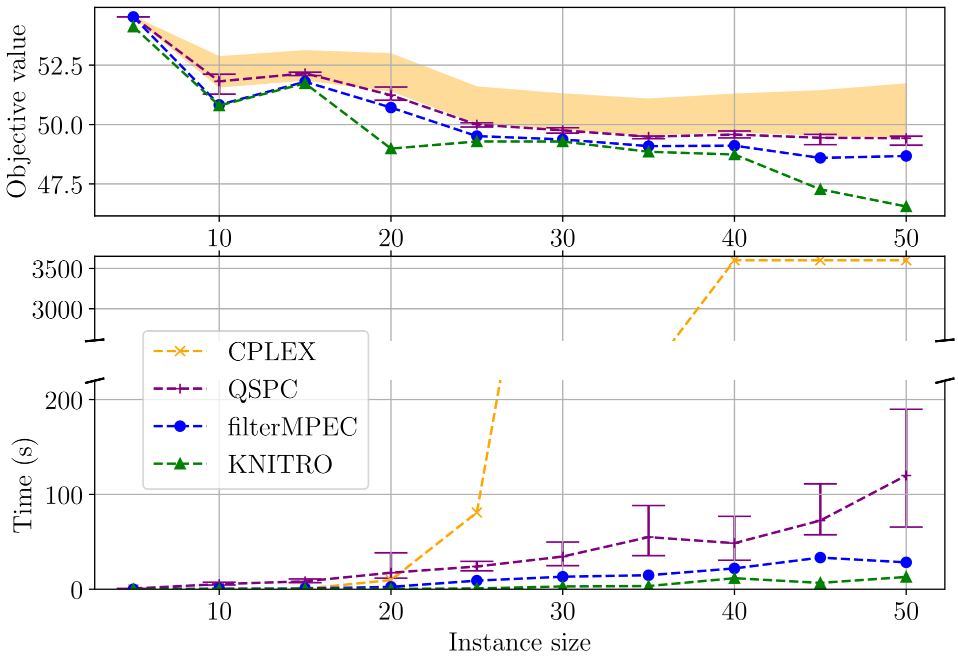


Figure 4.7: Comparison with NLP methods.

The upper graph shows the objective value and the lower graph shows the resolution time for a segments number S varying between 5 to 50 and a β fixed to 0.5. For heuristic methods, five tries have been done and vertical lines indicate the least and the greatest value. The final gap obtained with the quadratic method is represented with a yellow zone (between the best solution and the best upper bound).

(peak/off-peak/fixed part)[2]. These contracts mimic the most common type of contracts existing in the French power markets, and are listed in Table 4.2. To evaluate the costs C_{sw}, we use the methodology from the French regulator which consists in summing the different costs such as electricity production cost, taxes, transport and distribution network charges or commercial margin, see e.g. [CRE21, Figure 1]. The costs of electricity production are evaluated as the average of historical market prices to represent that the retailer buys the energy for its customers on power exchanges over the whole year. Costs are therefore not reflecting the hourly variability of electricity market prices but this is coherent with our approach which is not a dynamic time pricing but a fixed one.

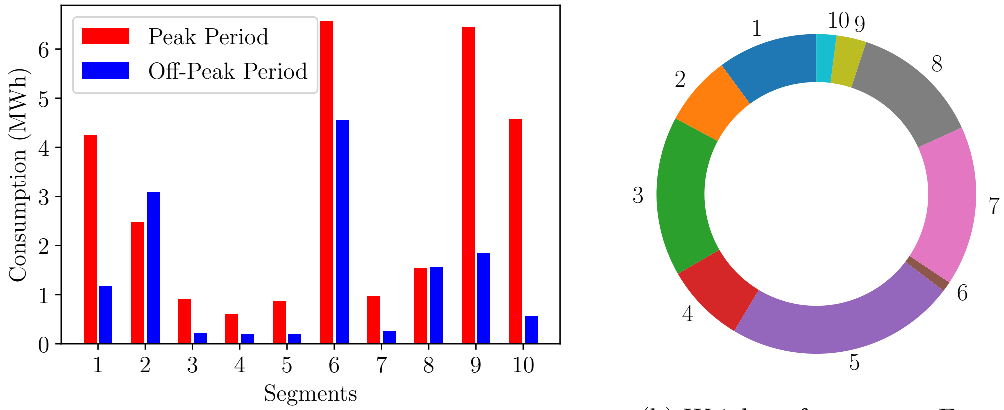
1 2	Base Peak/Off peak	Standard	Low cost offers (digital-only customer services)
3 4	Base Peak/Off peak	Green[3]	Higher costs, but preferred by some segments (higher reservation bill)

Table 4.2: Contracts used in the instances

Each offer has a base load version (no price difference between peak and off-peak periods) and a version with different prices at peak and off-peak periods, making a total of 4 contracts.

Concerning the customers, a thousand load curves (obtained by the `SMACH` simulator of EDF, see [Hur+15]) represent various power consumption profiles and mimic the entire French population, taking into account different household compositions, locations, and electrical equip-

[2]Here, we call "peak period" the interval 8am – 8pm. The other twelve hours defines the "off-peak" period.
[3]This type of contract provides power generated from renewable source such as on-shore wind and the retailer has to provide guarantees of origin which induces additional costs.



(a) Nominal consumption of segments, over one year. For each segment, the consumption is separated into the Peak period and the Off-peak period.

(b) Weights of segments. For each segment, the size of the section corresponds to the proportion of users in this segment.

Figure 4.8: Clustering for 10 segments.

ments. To construct our set of instances, we used the k-means algorithm to obtain S clusters (segments), where $S = 10$. In this way, customers that have similar consumption profile and contract preferences are aggregated in the same cluster. Figure 4.8 displays the nominal consumption after the clustering process, i.e., the aggregated year-based consumption that a typical customer of the segment is expected to consume when he faces a constant price. Segments 6 and 9 correspond to consumers with high electricity consumption and typically have individual houses with full-electric equipments and especially electrical heating. By contrast, segments 3, 4 and 5 are low energy consumers which are small households without electrical heating and cooking. Segments 1, 9 and 10 have a highly differentiated peak/off-peak profile compared with the others which consume in a more regular way. For peak/off-peak contracts, we suppose that each customer can shift a part of the consumption from peak period to off-peak period (*load shifting*). Here, we suppose that 15% of the nominal peak consumption can be shifted to off-peak periods. Moreover, we suppose that the green preference is cast into three categories: highly / mediumly / lowly eco-friendly. This corresponds to an additional utility of 4% / 2% / 0% of their bill computed with regulated prices[4]. For instance, segments 3, 4 and 5 have similar nominal consumptions (see Figure 4.8) but different green preferences (see Figure 4.9). In this study, we consider 6 competitors' offers, defined with real prices that can be found in the French market. These offers are depicted in Table 4.3.

Competitors	1	2	3	4	5	6
Peak (€/kWh)	0.174	0.1819	0.1840	0.19	0.166	0.23
Off peak (€/kWh)			0.147	0.155		0.135
Fixed portion (€)	136	136	144	144	148	141

Table 4.3: Competitors prices. Contract 2 and 4 are green contracts

<table>
<tr><td>Contract</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Peak (€/kWh)</td><td rowspan="2">0.166</td><td rowspan="2">0.1819</td><td>0.1768</td><td>0.2215</td></tr>
<tr><td>Off peak (€/kWh)</td><td>0.1607</td><td>0.1391</td></tr>
<tr><td>Fixed portion (€)</td><td>148</td><td>136</td><td>136.29</td><td>120</td></tr>
</table>

(a) Optimal prices with deterministic setting

<table>
<tr><td>Contract</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Peak (€/kWh)</td><td rowspan="2">0.1693</td><td rowspan="2">0.1834</td><td>0.1863</td><td>0.1895</td></tr>
<tr><td>Off peak (€/kWh)</td><td>0.1491</td><td>0.1626</td></tr>
<tr><td>Fixed portion (€)</td><td>133.7</td><td>129.29</td><td>122.95</td><td>128.19</td></tr>
</table>

(b) Optimal prices with quadratic regularization of intensity $\beta = 0.2$

Table 4.4: Optimal prices

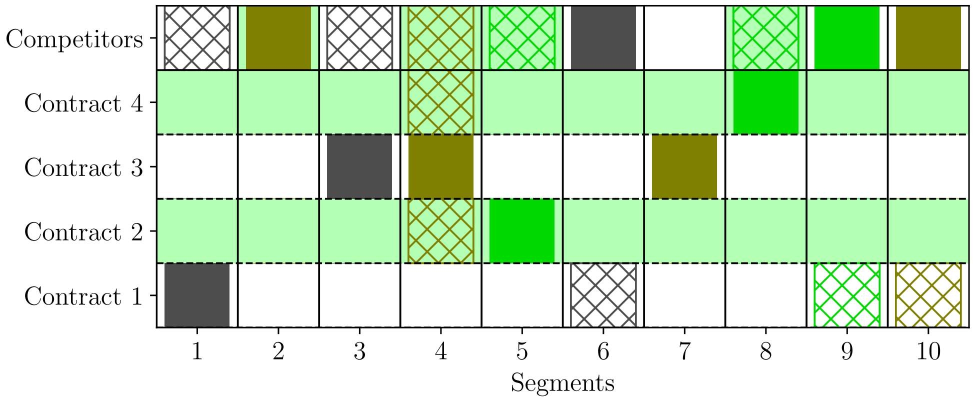
4.6.2 Numerical analysis

Table 4.4a shows the optimal prices for the deterministic case, and Figure 4.9a shows the corresponding distribution over the contracts. As previously explained on a theoretical example (Figure 4.3), the deterministic model adjusts the prices so that many customers face two contracts with equal utilities. This can be viewed in Figure 4.9a where the hatched bars represent the ties in the choice. In Table 4.4a, we have the extreme case where the second retailer's contract contends exactly the same coefficients as the second competitors' offer. We also noticed that the segments who naturally favor green energy (segments 5, 8 and 9) chose a green contract and that some segments are attributed to the competitors (such as segments 2, 6, 8 and 9) as they must be too costly for the retailer.

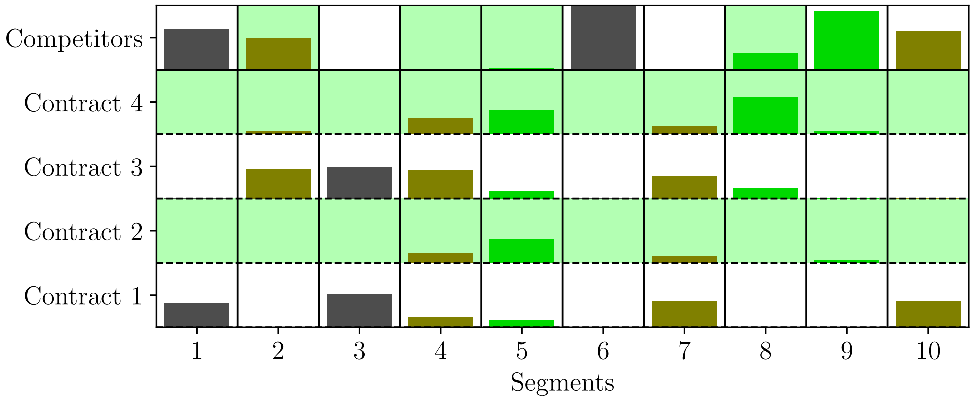
We also display the results for the regularized case $\beta = 0.2$. In this example, this choice of β appears to be close from the worst case from a retailer's point of view (see Figure 4.10) and is, in a sense, robust to any choice of β value. We observe that the optimal price grid (Table 4.4b) is somehow different from the optimistic one. Every contract has a lower fixed part in the regularized case compared with the deterministic case, but the variable portions can be either lower or greater. Concerning the customers' distribution along the contracts, we observe that the choices are globally preserved in the sense that every deterministic decision stays privileged in the regularized case. Let's notice that high consumption segments (segments 6 and 9) and highly differentiated peak/off-peak (segments 1, 9 and 10) are for a great part not favored by the retailer and let to competitors for a high proportion. On the other side, our retailer manages to attract green segments. In order to compare with logit approach, we also show the logit customers' distribution computed using the optimal quadratic prices. This distribution is very similar to the quadratic case, see Figure 4.9b and Figure 4.9c. The main difference lies in the small probabilities: the logit choice is slightly more spread on the different contracts, but the probabilities stay highly comparable. We refer to Section 4.8.2 where we provide a more detailed comparison between quadratic and logit approaches and develop metric estimates to quantify the deviation between the two models.

To analyze the impact of the regularization, Figure 4.10 draws the profit function (retailer objective) as a function of the regularization intensity. About the logit and quadratic model, the result for small β values is quite intuitive: with customers randomly reacting, the company can impose very high prices since there will always be some consumers taking its contracts. Hence, the company's profit becomes infinite as $\beta \to 0$. For the company, having deterministic customers is more beneficial since the price can be adjusted to perfectly fit the population

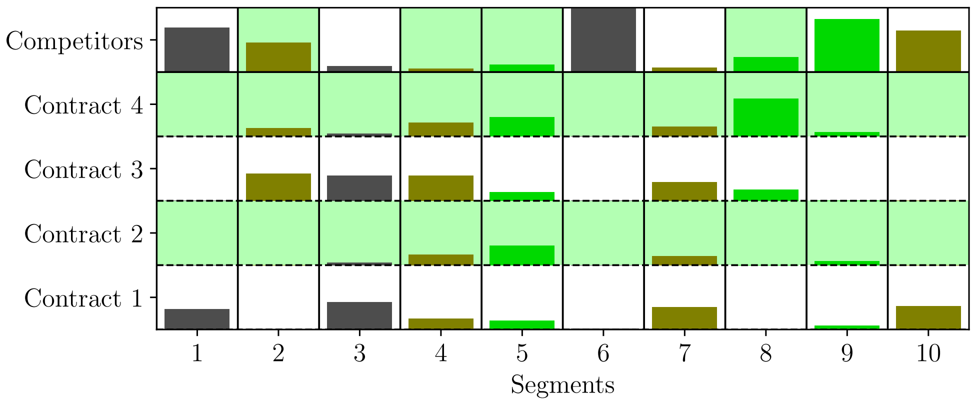
[4] The instances are not intended to fully depict the reality of the market, but they are already enough rich to deliver some useful insights on the effectiveness of the model.



(a) Optimal customers' distribution with deterministic setting.
A hatched bar means that the segment had the same utility as the chosen contract, but favors the retailer by choosing the one with the highest profit value (it could be a competitors' offer).



(b) Optimal customers' distribution with quadratic regularization of intensity $\beta = 0.2$.
The size of the bar defines the probability of choices, i.e., a bar taking a fourth of the rectangle height represents a choice probability of 25%.



(c) Customers' logit distribution with intensity $\beta = 0.8/e$ for prices of Table 4.4b.

Figure 4.9: Optimal customers' distributions.
Green contracts are displayed with a green-filled rectangle. Decisions of highly (resp. mediumly / lowly) eco-friendly clusters are displayed with green (resp. brown / gray) bars. The six offers of Table 4.3 are summed up into the first line, where only the best competitors' offer is displayed.

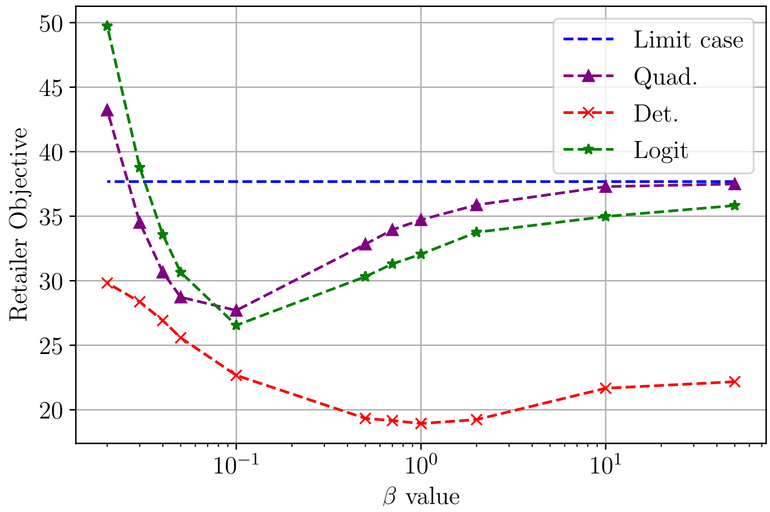


Figure 4.10: Optimal value as a function of the rationality parameter β.
We display the results for the model under logit response (Logit) and for the model under quadratic response (Quad.). In addition, we display the objective value obtained by applying the optimal prices of the deterministic model (Table 4.4a), assuming a quadratic response of the customers (Det.).

behavior. This can be interpreted as the result of moral hazard: the randomness in the followers decision negatively impacts the leader revenue.

We also display the objective function value obtained by fixing the price to the optimal prices found in the deterministic setting, i.e., supposing $\beta = \infty$, but recalculating the response and the optimal objective value in the uncertain context (finite value β). We see that the objective value is far below the optimal quadratic solution (indeed, around 40% of revenue are lost for large values of β). This highlights that the deterministic solution is unstable, and not robust to uncertainty, see comments in Section 4.3.4. It is then necessary to consider regularized consumers behavior to obtain a reliable menu of offers.

Remark 4.6.1

For completeness, one can find in Section 4.5 a numerical comparison of the `QSPC` solver with other methods. This study is performed on various instance sizes and compares the proposed method with direct resolution, MPEC solvers (via nonlinear reformulations), and on-the-shelf heuristics. In particular, we could solve instances of substantial size (10 contracts, 50 segments) in a reasonable time with a MIP gap tolerance of 3%.

4.7 Conclusion

We explored an extension of the unit-demand envy-free pricing problem, in which the customer invoice is determined by multiple price coefficients. We first analyzed a bilevel programming model, assuming a fully deterministic behavior of customers (every customer takes only one contract, maximizing her utility). This is inspired by known models in the case of a unique price coefficient. Such bilevel problems reduce to mixed linear programming, allowing one to solve instances of intermediate size to optimality. However, the assumption of deterministic behavior is not realistic, at least for the class of electricity pricing problems that motivate this work. So, we developed a new, alternative model, based on a quadratic regularization, which combines tractability and realism. We demonstrated that the lower response map of this quadratic model is characterized by a polyhedral complex, and using this geometrical property, we designed a heuristic which showed its efficiency in terms of optimality and time on our data set. We finally analyze the behaviors of the three models (deterministic, logit and quadratic) on a use case and highlight once again the need of a (tractable) probabilistic choice model to avoid unrealistic solutions.

Several extensions may be considered to further improve the realism. In particular, throughout the chapter, competitors are supposed not to adjust their prices to the strategy of the company (static competition). Relaxing this assumption would imply to consider a Nash equilibrium between leaders (multi-leader-common-follower games). In particular, this has been studied by [LM10], where an application to the electricity market is also the main motivation. Nonetheless, even considering the deterministic case (perfect knowledge and purely rational decision), only stationary points (not necessarily local solutions) can be numerically found in general, and for relatively small instance size. Besides, we also suppose that end-customers immediately react to the prices (no switching cost). Modeling such features would lead to dynamic games, increasing a lot the computational time, and making the above numerical study intractable.

4.8 Appendices

4.8.1 Complexity

[Gur+05] proved that the deterministic model is APX-hard (see [Pas09] for a description of this class). Using this result, we prove that the quadratic case is also APX-hard:

Proposition 4.8.1

The problem ($q\beta$-BP) is APX-hard, even in the single-attribute setting and without price constraints.

Proof. Reusing the same polynomial transformation (and the same notations) as in [Gur+05], we claim the existence of a sufficiently large parameter β ($\beta \geqslant 8(n+m)$) such that the quadratic optimal value is not far from the deterministic one i.e., $|\, v(q\beta\text{-}BP) - v(o\text{-}BP)| \leqslant 1/4$.
First, it can be noticed that the optimal prices cannot be any values: for any product,

- if the price is in $]\frac{2}{\beta}, 1-\frac{2}{\beta}[$, then customers having a null reservation bill for the contract will have no chance to purchase it and customers having reservation bill of 1 or 2 will purchase it with probability 1. So the company has more interest in setting the price at $1-\frac{2}{\beta}$.
- With the same logic, if the price is in $]1+\frac{2}{\beta}, 2-\frac{2}{\beta}[$, then the company has more interest in setting the price at $2-\frac{2}{\beta}$.
- If the price is less than $\frac{2}{\beta}$, the profit made by the company with this contract is less than $1/4$, so setting the price to $1-\frac{2}{\beta}$ is more beneficial.
- Finally, a price greater than $2+\frac{2}{\beta}$ does not make any profit.

For an optimal solution, the price values can only be in $[1-\frac{2}{\beta}, 1+\frac{2}{\beta}] \cup [2-\frac{2}{\beta}, 2+\frac{2}{\beta}]$. Taking the optimal quadratic prices and rounding them to obtain a price vector of values 1 or 2 provides a price vector for the deterministic problem with a value closed to the quadratic optimum i.e., $v(o\text{-}BP) \leqslant v(\beta\text{-}BP) - \frac{2}{\beta}(n+m)$.

For the converse, taking the optimal deterministic solution (we know that the prices can only be 1 or 2) and subtracting $\frac{2}{\beta}$ to each price gives a quadratic solution with objective value closed to the deterministic optimum i.e., $v(\beta\text{-}BP) \leqslant v(o\text{-}BP) - \frac{2}{\beta}(n+m)$.

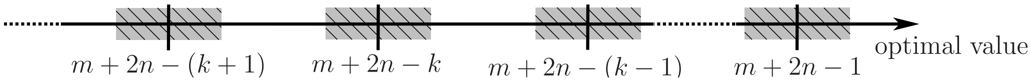


Figure 4.11: Representation of the objective value for the transformation of Guruswami et al.

Deterministic optimum is integer and the quadratic one lies in a small interval centered on it (hashed zones).

Computing the quadratic optimum for $\beta \geqslant 8(n+m)$ and rounding it gives us the deterministic optimum. Thus, the quadratic case is at least as hard as the deterministic case, which was proved to be APX-hard. $\square$

Remark 4.8.1

The structure of this specific instance allows us to exhibit a threshold from which the quadratic model is a sufficiently good approximation for the deterministic model. In a more general case, even if we have established the convergence of the quadratic model to the deterministic one, we are not able to provide such a threshold.

4.8.2 Metric estimates to compare logit and quadratic regularization

Proposition 4.8.2

Consider a segment s facing $W+1$ disutilities $V_{s0}, \dots, V_{sW}$ sorted in ascending order. For a given $\beta > 0$, we denote by $(y_{sw}^{quad})_w$ the quadratic response (computed with a parameter $\beta' = \beta e/4$) and by $(y_{sw}^{log})_w$ its logit analog (computed with β). Then,

$$\text{If } y_{sw}^{quad} = 0, \text{ then } y_{sw}^{log} \leqslant \gamma_w := \left(1 + w e^{\frac{8}{we}}\right)^{-1} \quad (\leqslant 1/9) \tag{4.22}$$

Conversely,

$$\text{If } y_{sw}^{log} \leqslant \eta_w^W := \left(W + 1 + w(e^{\frac{8}{e}} - 1)\right)^{-1}, \text{ then } y_{sw}^{quad} = 0 \tag{4.23}$$

Proof. Suppose that $y_{sw}^{quad} = 0$, then from Proposition 4.3.2, $V_{sw} \geqslant c_{sw} = \frac{1}{w}\left[\frac{8}{e\beta} + \sum_{k=0}^{w-1} V_{sw}\right]$ and thus

$$\exp\left(\frac{8}{we} - \beta V_{sw}\right) \leqslant \exp\left(-\frac{1}{w}\sum_{k=0}^{w-1} \beta V_{sk}\right) \leqslant \frac{1}{w}\sum_{k=0}^{w-1} e^{-\beta V_{sk}} \ ,$$

where the latter inequality is obtained by convexity of the exponential. We then deduce that $\gamma_w^{-1} e^{-\beta V_{sw}} \leqslant \sum_{k=0}^{w} e^{-\beta V_{sk}}$. Using the logit expression gives us the desired result.

Suppose that $y_{sw}^{log} \leqslant \eta$ for a given η. We exploit the ascending sort on V in the logit expression to obtain

$$\eta \geqslant \frac{e^{-\beta V_{sw}}}{\sum_{k=0}^{w-1} e^{-\beta V_{sk}} + \sum_{k=w}^{W} e^{-\beta V_{sk}}} \geqslant \frac{e^{-\beta V_{sw}}}{\sum_{k=0}^{w-1} e^{-\beta V_{s0}} + \sum_{k=w}^{W} e^{-\beta V_{sw}}} \ .$$

Continuing the simplifications, $\eta^{-1} \leqslant w e^{-\beta(V_{s0} - V_{sw})} + (W - w + 1)$ and therefore

$$V_{sw} \geqslant V_{s0} + \frac{1}{\beta}\log\left(\frac{\eta^{-1} - (W - w + 1)}{w}\right) \ .$$

Finally, taking $\eta = \eta_w^W$ implies that $V_{sw} \geqslant V_{s0} + \frac{8}{e\beta}$, insuring that $y_{sw}^{quad} = 0$. $\square$

The technical Proposition 4.8.2 shows that there is a common convergence speed to the deterministic behavior: in fact, for any value of β, if we have no "quadratic chance" to choose a contract w then we have a very little logit probability to choose w. The converse applies but it depends on the total number of contracts; in the logit version, the probability depends on the whole set of contracts whereas the quadratic version does not care of the contracts that have a very large disutility. It is important to note that the bounds γ_w and η_w^W in (4.22) and (4.23) are valid for any value of β.

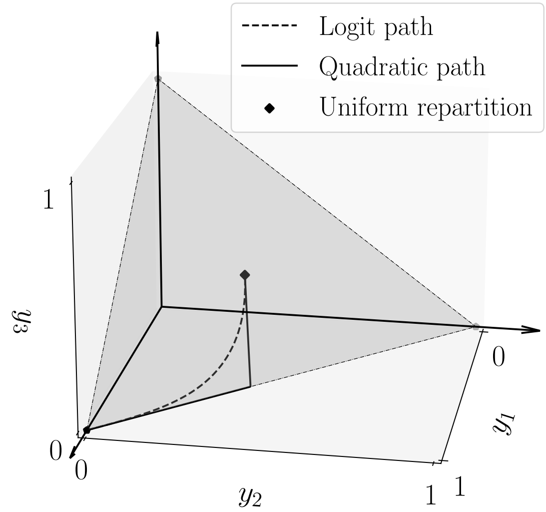


Figure 4.12: Logit and quadratic path on the simplex, as functions of β

Figure 4.12 illustrates Proposition 4.8.2 and shows the logit and quadratic paths for a disutility vector $V = (0, \frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}})$. The trajectory shares the same start point (the simplex center for $\beta = 0$) and the same end point (the vertex $y = (1, 0, 0)$ for $\beta \to +\infty$). However, for the rest of the path the trajectories slightly deviate: we observe the sparsity effect of the projection operator in the behavior of the quadratic path whereas the logit trajectory always lies in the interior of the simplex.

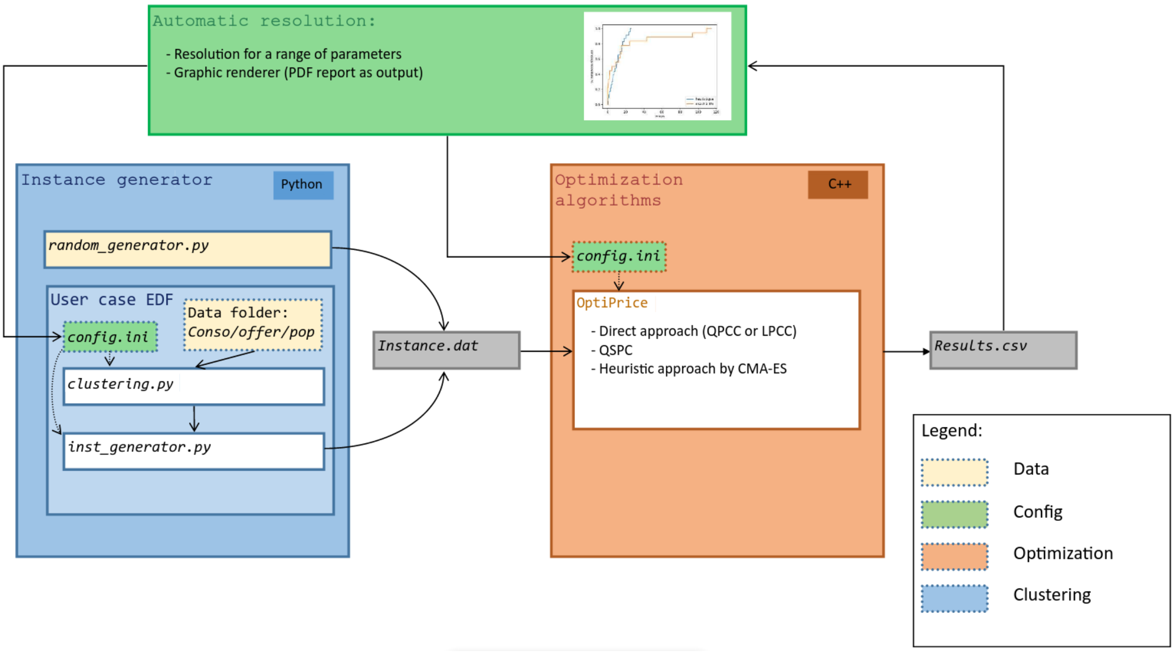
4.8.3 Description of the code

The different models (optimistic, logit and quadratic) have been implemented into a digital mock-up, thought to be generic and user-friendly thanks to a graphic interface, see Figure 4.13.

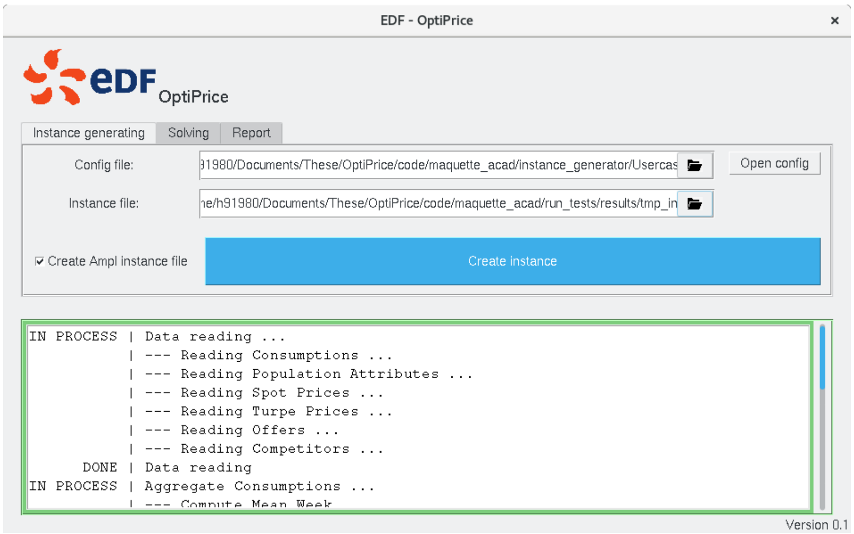
A first step, more data oriented, reads several excel files (mostly information on consumers load curves and costs). After a pre-processing, a k-means clustering is performed to obtain a prescribed number of segments (parameter to give in a configuration file). At the end of this step, an .dat file is created, corresponding to the mathematical description of the instance (with the same notation as in the previous sections), see Figure 4.14. The generation of instance files makes possible the use of not only realistic instances (obtained by the preprocessing step previously described) but also academic instances that can be randomly generated.

Then, the optimization is performed. The different optimization models are encoded in C++ (using the IloCplex library for QPCC and LPCC models, after a mixed-integer reformulation). The solving of the logit model is also possible (using the black-box solver CMA-ES). The configuration file allows to run a single method or several methods/models at once. When the computation is performed, a solution file is written, so that the performances can be analyzed.

The third step on the tool is a post-processing phase, in which the solution files are read and a pdf report is automatically generated, including visualization of the solutions (as obtained in Figure 4.9).



(a) Flow chart of the code



(b) Screenshot of the graphic interface

Figure 4.13: Flow chart and screenshot of the graphic interface

.readme

```
 Instance Format:
------------------

Row 1:
S W H X
where S nb of segments, W nb of contracts, H, nb of periods, X nb of equalities/inequalities
describing X

Row 2:
(rho_s)_{s in [S]}
where rho_s is the weight of segment s in terms of profit for the company

Row 3:
(beta_s)_{s in [S]}
where beta_s is the rationality of segment s.  This parameter is ignored for optimistic resolution

Row 4:
(alpha_w)_{w in [W]}
where alpha_w is the minimum fraction of the population that has to chose the contract

Rows 5 to 4+S:
(E^h_sw)_{w in [W], h in [H]}
where E^h_sw is the comsumption of s at period h if he chooses w.  Note that the 2-dimension indices
(w,h) are linearized such that k = H*(w-1) + h

Rows 5+S to 4+2*S:
(C_sw)_{w in [W]}
where C_sw is the cost to provide the service to s if he chooses w

Rows 5+2*S to 4+3*S:
(R_sw)_{w in [W]}
where R_sw is the reservation price of s for contract w

Rows 5+3*S to 4+3*S+X:
w E/G b A_h
where w is the concerned contract,
E/G equals 0 if sum_h (A_h*x_wh) = b and E/G equals 1 if sum_h (A_h*x_wh) >= b
```

Figure 4.14: Explaination of the instance format (in the “.readme” of the tool)

Chapter 5

Ergodic control of a heterogeneous population and application to electricity pricing

***

This chapter is based on the proceedings paper [Jac+22], to which we add the study of a particular example (Section 5.7) showing that the solution to the ergodic eigenproblem is in general not unique. We also add a section (7.5) that first extends the results in the presence of noise in the dynamics, and then relates these results with weak-KAM theory. In particular, we show for controllable systems that the turnpike property can be viewed as a particular case of the convergence to the Aubry set.

Abstract. We consider a control problem for a heterogeneous population composed of customers able to switch at any time between different contracts, depending not only on the tariff conditions but also on the characteristics of each individual. A provider aims to maximize an average gain per time unit, supposing that the population is of infinite size. This leads to an ergodic control problem for a "mean-field" MDP in which the state space is a product of simplices, and the population evolves according to controlled linear dynamics. By exploiting contraction properties of the dynamics in Hilbert's projective metric, we show that the ergodic eigenproblem admits a solution. This allows us to obtain optimal strategies, and to quantify the gap between steady-state strategies and optimal ones. We illustrate this approach on examples from electricity pricing, and show in particular that the optimal policies may be cyclic –alternating between discount and profit taking stages.

5.1 Introduction

5.1.1 Motivation and Context

Most OECD[1] members have engaged a reform of their retail electricity markets. Historical providers are now facing competition with new entrants. Opening up markets to competition aims to improve their efficiency and to lower the prices for consumers, proposing a wider choice of offers.

In theory, consumers are often supposed to be fully rational, and their reactions to price to be instantaneous. However, many studies highlight that switching costs and limited awareness conjointly lead to inertia in retail electricity market, which hinders efficient choices, see [HMP17; NMS19; DW19]. Inertia in imperfect markets impacts the decision of the providers and modifies their pricing strategies. Then, what is the optimal tariff strategy for a company? In general, two opposing forces arise: a *harvesting motive* and a *incentive motive*. Either the company favors immediate rewards by taking advantage of the static market power, either the firm proposes attractive offers to increase its market share and secure greater harvest in the future [Cab09]. Studies also tend to show the importance of promotions in the pricing behavior of firms, see [HP10; AR12]. In particular, empirical analyses show how the depth and frequency of promotions are linked with the level of inertia.

5.1.2 Contributions

We consider a population of customers, that have different types (consumption profiles). Each customer chooses between several energy contracts, taking into account the price offers of a provider, who aims at optimizing a mean reward per time unit. This is represented by an ergodic control problem, in which the state –the population– belongs to a product of simplices. We suppose that the population evolves according to the Fokker-Planck equation of a controlled Markov chain. In this work, we directly study the "mean-field" model where the population is supposed to be of infinite size. This choice is motivated by our application where the population is in fact the whole set of French households (around 30 million), leading to untractable model without such mean-field hypothesis. Our first main result, Theorem 5.2.2, shows that the ergodic eigenproblem does admit a solution. This entails that the value of the ergodic control problem is independent of the initial state, and this also allows us to determine optimal stationary strategies. Theorem 5.2.2 requires a primitivity assumption on the semigroup of transition matrices; it applies in particular to positive transition matrices, such as the ones arising from logit based models. The proof relies on contraction properties of the dynamics in Hilbert's projective metric, which allow us to establish compactness estimates which guarantee the existence of a solution.

We then study stationary pricing strategies. Owing to the contraction properties of the dynamics, these are such that the population distribution converges to a stationary state. Then, we refine a result from [Fly79], providing a bound on the loss of optimality arising from the restriction to stationary pricing strategy. We define a family of Lagrangian functions, whose duality gap provides an explicit bound on the optimality loss, see Proposition 5.3.1. In particular, a zero duality gap guarantees that stationary pricing policies are optimal.

Finally, we apply these results to a problem of electricity pricing, inspired by a real case study (French contracts). An essential feature of this model is to take into account the *inertia* of customers, i.e., their tendency to keep their current contract even if it is not the best offer. This is represented by a logit-based stochastic transition model with switching costs. Theorem 5.4.1 provides a closed-form formula for the stationary distribution. We present numerical tests on

[1] https://www.oecd.org/about/document/ratification-oecd-convention.htm

examples of dimension 2 and 4. These reveal the emergence of optimal cyclic policies for large switching costs, recovering the empirical notion of "promotions" of [DHR09] and [PE17].

5.1.3 Related works

As mentioned above, several studies brought to light complex phenomena that emerge when considering pricing on imperfect markets with inertia. However, this dynamic pricing problem has been theoretically studied only recently: Pavlidis and Ellickson [PE17] focus on the discounted infinite horizon pricing problem, and numerically solved it in small dimension. They directly suppose a continuum of customers in each segment of the (heterogeneous) population, leading to a "mean-field" system. In the context of discounted horizon, and in absence of common-noise, the derivation of this model as a limit of a large finite population is achieved in [GG10]. In particular, Gast and Gaujal provide guarantees on the speed of convergence of order $1\sqrt{N}$. Motte and Pham [MP22] generalize the results in the presence of common-noise. In [Bäu23], Bauerle focuses on a different criteria: the average long-term reward. This criterion has been widely studied in control processes, but much less in the mean-field context. Biswas studied mean-field games in discrete time, and proved that, under particular conditions, the optimum is characterized by an ergodic eigenproblem [Bis15].

In contrast, the ergodic eigenproblem studied here is of a deterministic nature, more degenerate than its stochastic analogue studied in the context of average cost Markov Decision Processes. In particular, the Doeblin-type conditions generally used in this setting to obtain the existence of an eigenvector [Kur89; HL96] do not apply. In fact, we end up with a special case of the "max-plus" or "tropical" infinite dimensional spectral problem [KM97; AGW09], or of the eigenproblem studied in discrete weak-KAM and Aubry Mather theory [Fat08; GT11]. Basic spectral theory results require the Bellman operator to be compact. This holds under demanding "controllability" conditions (see e.g. [KM97, Theorem 3.6]), not satisfied in our setting. Extenstions of these results rely on quasi-compactness techniques [MN02; AGN11], which also do not apply to our problem. In comparison with [Bäu23], the equi-Lipschitz property of the value function is not assumed a priori. Instead, we exploit here the contraction properties of the dynamics to obtain the existence of the eigenvector. This is partly inspired by a previous work of Calvez, Gabriel and Gaubert [CGG14], in which contraction techniques in Hilbert metric were applied to a different problem (growth maximization). Also, [CGG14] deals with a PDE rather than discrete setting. Our result should also be compared with [Bis15, Th. 3.1], in which different conditions, based on geometric ergodicity are used to guarantee the existence of an eigenvector; these conditions do not apply to our case, in fact, they entail that the eigenvector is unique up to an additive constant, and this is generally not true in our model. In fact, we provide an explicit counter example showing that the eigenvector may not be unique, see Example 5.2.1, and this entails that our existence result cannot be obtained using geometric ergodicity arguments.

The ergodic eigenproblem, in the deterministic "0-player case", has been studied under the name of cohomological equation in the field of dynamical systems. The existence of a regular solution is generally a difficult question, a series of results going back to the work of Livšic [Liv72], show that a Hölder continuous eigenvector does exists if the payment function is Hölder continuous, and if the dynamics is given by an Anosov diffeomorphism. The latter conditions requires the tangent bundle of the state space to split in two components, on which the dynamics is either uniformly expanding or uniformly contracting. Here, we establish a "one player" version, but requiring a uniform contraction assumption.

[2]the Markov Chain induced by any deterministic stationary policy consists of a single recurrent class plus a –possibly empty– set of transient states (i.e., there exists a subset of states that are visited infinitely often with probability 1 independently of the starting state)

	Time	Transitions	Assumption
[Sch85]	discrete	stochastic	unichain[2]
[Bis15]	discrete	stochastic	Doeblin / minorization[3]
[MN02]	discrete	deterministic	quasi-compactness
[Fat08]	continuous	deterministic	controlability[4]
[Zav12]	discrete	deterministic	controlability
[CGG14]	continuous	deterministic	contraction of the dynamics

Table 5.1: Non-exhaustive comparison of approaches to prove the existence of a solution to the ergodic eigenproblem.

This chapter is organized as follows. In Section 5.2, we first define the model and prove the results on the ergodic eigenproblem. We study steady-states and their optimality in Section 5.3, and illustrate the electricity application in Section 5.4.

5.2 Ergodic control

5.2.1 Notations

We denote by Δ_n the simplex of $\mathbb{R}^n$, and by $\langle \cdot, \cdot \rangle_n$ the scalar product on $\mathbb{R}^n$. We denote by $\operatorname{sp}(f) := \max_{x\in E} f(x) - \min_{x \in E} f(x)$ the span of the function $f : E \to \mathbb{R}$. We say that a matrix P is positive, and we write $P \gg 0$, if all the coefficients of P are positive. The set of convex functions with finite real values on a space K is denoted by $\operatorname{Vex} K$, and the convex hull of a set K is denoted by $\operatorname{vex} K$. Moreover, the set of Lipschitz function on E is denoted by $\operatorname{Lip}(E)$, and the relative interior of a set E is denoted by $\operatorname{relint}(E)$.

The *Hilbert projective metric* d_H on $\mathbb{R}^n_{>0}$ is defined as $d_H(u,v) = \max_{1\leqslant i,j \leqslant n} \log(\frac{u_i}{v_i}\frac{v_j}{u_j})$. see [LN09]. It is such that $d_H(u,v) = 0$ iff the vectors u and v are proportional, hence, the name "projective". For a set $E \subseteq \mathbb{R}^n_{>0}$, we denote by $\operatorname{Diam}_H(E) := \max_{u,v\in E} d_H(u,v)$ the diameter of the set E, and for a matrix $P \in \mathbb{R}^{n\times n}$ we denote by $\operatorname{Diam}_H(P) := \max_{1\leqslant i,j\leqslant n} d_H(P_i, P_j)$ the *diameter* of P, where P_i denotes the ith row of P. This can be seen to coincide with the diameter, in Hilbert's projective metric, of the image of the set $\mathbb{R}^n_{>0}$ by the transpose matrix of P.

Finally, for a sequence $(a_t)_{t\geqslant 1}$, we respectively denote by $a_{s:t}$, and $a_{:t}$ the subsequences $(a_\tau)_{s\leqslant\tau\leqslant t}$ and $(a_\tau)_{1\leqslant\tau\leqslant t}$.

5.2.2 Model

We consider a large population model composed of K clusters of indistinguishable individuals. Each cluster $k \in [K] := \{1,\dots,K\}$ represents a proportion ρ_k of the overall population, and is supposed to react independently from the other clusters.

Let $\mathcal{X}$ and $\mathcal{A}$ be respectively the state and action spaces. We suppose in the sequel that $\mathcal{X}$ is finite and w.l.o.g. $\mathcal{X} = \{1,2,\dots,N\}$. We suppose also that $\mathcal{A}$ is a compact set (in Section 5.4, we will consider a subspace of $\mathbb{R}^N$).

For any time $t \geqslant 0$ and any cluster k, we denote by $\mu_t^k \in \Delta_N$ the distribution of the population of cluster k over $[N]$.

[3] for all state s, action a and measurable subset B of the state space, $P(B|x,a) \geqslant \epsilon\mu(B)$

[4] for every pair of states (s,s'), there exists an action a making s' accessible from s

At every time $t \geqslant 1$, a controller chooses an action $a_t \in \mathcal{A}$. She obtains a reward $r : \mathcal{A} \times \Delta_N^K \to \mathbb{R}$ defined as

$$r : (a_t, \mu_t) \mapsto \sum_{k \in [K]} \rho_k \left\langle \theta^k(a_t), \mu_t^k \right\rangle_N \quad , \tag{5.1}$$

where $\theta^{kn}(a)$ is the unitary reward for the controller coming from an individual of cluster k in state n after executing action a.

We suppose that the dynamics of the system are deterministic, linear, with a Markov transition matrix. We then denote by $P^k(a)$ the transition matrix for cluster k such that

$$\mu_t^k = \mu_{t-1}^k P^k(a_t) \quad . \tag{5.2}$$

The (deterministic) semi-flow ϕ of the state μ is then defined by

$$\phi_t(a_{:t}, \mu_0) := \mu_t \quad .$$

We also denote by Π the set of policies. Then, for a given policy $\pi = \{\pi_t\}_{t \geqslant 1}$, the action taken by the controller at time t is $a_t = \pi_t(\mu_t)$.

In the sequel, the following assumptions will be used:

(A1) The transition matrix $P^k(\cdot)$ is a continuous function of the action for any k.

(A2) There exists $L \in \mathbb{N}$ such that for any sequence of actions $a_{:L} \in \mathcal{A}^L$ and cluster k, $\prod_{l \in [L]} P^k(a_i) \gg 0$.

Recall that in Perron-Frobenius theory, a nonnegative matrix M is said to be *primitive* if there is an index l such that $M^l \gg 0$, see [BP94, Ch. 2]. Assumption (A2) holds in particular under the following elementary condition:

(A2') For any action $a \in \mathcal{A}$, $P(a) \gg 0$.

(A3) There exists M_r such that, $|\theta^{kn}(a)| \leqslant M_r$ for every $k \in [K]$, $n \in [N]$ and $a \in \mathcal{A}$.

Condition (A2) has appeared in [Gau96] in the context of semigroup theory, it can be checked algorithmically by reduction to a problem of decision for finite semigroups, see Rk. 3.8, *ibid.* Observe that (A3) is very reasonable in practice.

We equip the product of simplices Δ_N^K with the norm $\|\mu\| := \sum_{k=1}^K \|\mu^k\|_1$. It follows from (A3) that for any action a, the total reward function $\mu \mapsto r(a, \mu)$ is a M_r-Lipschitz real-valued function from $(\Delta_N^K, \|\cdot\|)$ to $(\mathbb{R}, |\cdot|)$.

Remark 5.2.1

Assumption (A2') is not a Doeblin minorization condition (as in [Bis15] for instance):this would suppose that there would exist $\epsilon > 0$ and a measure $f(\cdot)$ such that for all subsets $D \subseteq \Delta_N^K$,

$$\inf_{\mu \in \Delta_N^K} \inf_{a \in \mathcal{A}} \mathbb{P}\left[\mu_t \in D \mid \mu_{t-1} = \mu, a_t = a\right] \geqslant \epsilon f(D) \quad .$$

Here, this condition does not hold since the transitions are deterministic: $\mathbb{P}\left[\mu_t = \nu \mid \mu_t = \mu, a_t = a\right] = \mathbb{1}_{(\nu = \mu P(a))}$. This constitutes the most degenerate (and difficult) case to handle.

5.2.3 Optimality criteria

We suppose that the controller aims to maximize her average long-term reward, i.e.,

$$g^*(\mu_0) = \sup_{\pi\in\Pi} \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} r(\pi_t(\mu_t), \mu_t) \ . \tag{5.3}$$

Starting from μ_0, the population distribution will evolve in Δ_N^K according to a policy $\pi \in \Pi$. Nonetheless, with the assumptions we made, we next show that the dynamics effectively evolves on a particular subset.

Let $Q_L^k(a_{:L}) := \prod_{l\in[L]} P^k(a_l)$ be the transition matrix over L time steps, and $\mathcal{D}_L$ be defined as $\mathcal{D}_L = \bigtimes_{k\in[K]} \mathcal{D}_L^k$ where

$$\mathcal{D}_L^k = \operatorname{vex}\left(\{\mu^k Q_L^k(a_{:L}) \mid a_{:L} \in \mathcal{A}^L, \mu^k \in \Delta_N\}\right) \ . \tag{5.4}$$

Lemma 5.2.1

Let (A1)-(A2) hold. Then $\mathcal{D}_L$ is a compact set included in the relative interior of Δ_N^K. Moreover, for $t \geqslant L$, $\mu_t \in \mathcal{D}_L$ for any policy $\pi \in \Pi$.

Proof. The set $\{\mu^k Q_L^k(a_{:L}) \mid (a_{:L}, \mu^k) \in \mathcal{A}^L \times \Delta_N\}$ is compact, since $(a_{:L}, \mu^k) \mapsto \mu^k Q_L^k(a_{:L})$ is continuous and Δ_N and $\mathcal{A}$ are both compact. Therefore, $\mathcal{D}_L$ is compact as it is the convex hull of a compact set in finite dimension. Then, the positiveness of Q_L^k implies that $\mathcal{D}_L^k \subset \operatorname{relint}(\Delta_N)$. Moreover, by property of the semiflow, $\phi_t(a_{:t}, \mu_0) = \phi_L\left(a_{t-L+1:t}, \phi_{t-L}(a_{:t-L}, \mu_0)\right) \in \mathcal{D}_L$. $\square$

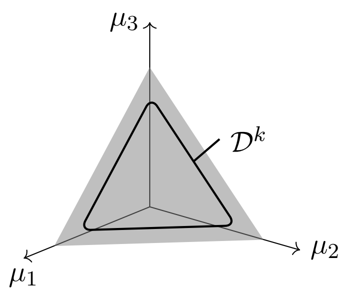


Figure 5.1: Effective domain $\mathcal{D}^k$ for Δ_3.

We recall that the relative interior of the simplex, equipped with Hilbert's projective metric, is a complete metric space, on which the Hilbert's metric topology is the same as the Euclidean topology. Hence, under (A1) and (A2), $(\mathcal{D}_L, d_H)$ is a complete metric space. We also recall *Birkhoff theorem*, which shows that every matrix $Q \gg 0$ is a contraction in Hilbert's projective metric, i.e.,

$$\forall \mu, \nu \in (\mathbb{R}_{>0}^N),\ d_H(\mu Q, \nu Q) \leqslant \kappa(Q) d_H(\mu, \nu) \ , \tag{5.5}$$

where

$$\kappa(Q) := \tanh\left(\operatorname{Diam}_H(Q)\,/\,4\right) < 1 \ ,$$

see [LN09, Appendix A]. This property applies to the transition matrix $P^k(a)$ under (A2'), or to Q_L^k under (A2).

Theorem 5.2.1

Suppose that Assumptions (A1) and (A2) hold. Then, the gain g^* does not depend on the initial population distribution.

Proof. Suppose that (A2') hold. Let μ_0 and ν_0 two initial distributions such that $g^*(\mu_0) < g^*(\nu_0)$. The optimal gain $g^*(\nu_0)$ is achieved by an infinite sequence of action $a_1^*, a_2^*, \ldots$. Then, by applying the same sequence of action starting from μ_0, we obtain

$$\begin{aligned}\left|\frac{1}{T}\sum_{t=1}^{T} r\left(a_t, \mu_t\right) - r\left(a_t, \nu_t\right)\right| &\leqslant \frac{1}{T} M_r^{\Phi} \sum_{t=1}^{T} d_H(\mu_t, \nu_t) \\ &\leqslant \frac{1}{T} M_r^{\Phi} \sum_{t=1}^{T} \kappa^t d_H(\mu_0, \nu_0) \\ &\leqslant \frac{1}{T}\frac{\kappa}{1-\kappa} d_H(\mu_0, \nu_0) \ .\end{aligned}$$

The property (5.5) is used with $\kappa = \max_{a \in \mathcal{A}} \kappa(P(a))$. We therefore constructed a policy that gives an average gain of $g^*(\nu_0)$ starting from μ_0.

The generalization for Section 5.2.2 is achieved by looking at a larger time subdivision where the transition between two steps is given by $Q_L \gg 0$. □

In the sequel, the average gain will be simply written g^*.

5.2.4 Ergodic eigenproblem

For any real-valued function $v : \Delta_N^K \to \mathbb{R}$, the Bellman operator $\mathcal{B}$ is defined as

$$\mathcal{B}\, v\,(\mu) = \max_{a \in \mathcal{A}} \{ r(a, \mu) + v(\mu P(a)) \} \ .$$

A first observation is that $\mu \mapsto (\mathcal{B}\, v)(\mu)$ is convex for any real-valued convex function v. Indeed, the transition is linear in μ, as well as the reward; therefore, for any $a \in \mathcal{A}$, the expression under the maximum is convex in μ, and since the maximization preserves the convexity, the observation is established. For a feedback policy π, we also define $\mathcal{B}^\pi$ the Kolmogorov operator such that $\mathcal{B}^\pi\, v\,(\mu) = r(\pi(\mu), \mu) + v(\mu P(\pi(\mu)))$.

Existence of a solution

The ergodic control problem for a Markov decision process with Bellman operator $\mathcal{B}$, on a compact state space $\mathcal{X}$, is classically studied by means of the ergodic eigenproblem

$$g 1_{\mathcal{X}} + h = \mathcal{B} h \ , \tag{5.6}$$

in which h is a bounded function on the state space, called the bias or potential, and g is a real constant. If the ergodic eigenproblem is solvable, then, g yields the optimal mean payoff per time unit, and it is independent of the initial state. Moreover, an optimal policy can be obtained by selecting maximizing actions in the expression of $\mathcal{B}h$. When the state and action spaces are finite, the ergodic eigenproblem is well understood, in particular, a solution does exist if every policy yields a unichain transition matrix (i.e., a matrix with a unique final class), see e.g. [Put94]. In the case of in infinite state space, the existence of a solution to the ergodic eigenproblem is a more difficult question [KM97; Fat08; MN02; AGN11]. This is especially the case for *deterministic* Markov decision processes, owing to the lack of regularizing effect of stochastic transitions. Here, we exploit the contraction properties of the dynamics, with respect to Hilbert's projective metric, together with the vanishing discount approach, to show the existence. First let us show a preliminary result on metrics comparison:

Lemma 5.2.2
Let $\mathcal{D} \subset \mathrm{relint}(\Delta_n)$, $n \in \mathbb{N}$ and $x, y \in \mathcal{D}$. Then,

$$n \|x - y\|_\infty \leqslant d_H(x, y) \Upsilon(\mathrm{Diam}_H(\mathcal{D})) \tag{5.7}$$

where $\Upsilon(d) = \frac{1}{d} e^d (e^d - 1)$.

Proof. We use the results in [AGN15]: Lemma 2.3 shows that for any vectors $u, x, y \in \mathcal{D}$ such that there exist $a, b > 0$ satisfying $ax \leqslant u \leqslant bx$ and $ay \leqslant u \leqslant by$, we have the following inequality:

$$\|x - y\|_u \leqslant \left(e^{d_T(x,y)} - 1\right) e^{\max(d_T(x,u), d_T(y,u))} \ ,$$

where d_T denotes the Thompson distance, and $\|z\|_u = \inf\{a > 0 \mid -au \leqslant z \leqslant au\}$. In particular, by choosing $u = (1/n, \dots 1/n)$ as the center of the simplex, $\|\cdot\|_u = n \|\cdot\|_\infty$. Moreover, $d_T(\cdot, \cdot) \leqslant d_H(\cdot, \cdot)$ on $\mathrm{relint}(\Delta_N^K)$, see [AGN15, Eq. 2.4]. Therefore,

$$\begin{aligned} n \|x - y\|_\infty &\leqslant \left(e^{d_H(x,y)} - 1\right) e^{\max(d_H(x,u), d_H(y,u))} \\ &\leqslant \left(e^{d_H(x,y)} - 1\right) e^{\mathrm{Diam}_H(\mathcal{D})} \ . \end{aligned}$$

We easily conclude using the fact that $f : x \mapsto e^x - 1$ is a convex function, and so for all $0 \leqslant x \leqslant \bar{x}$, $f(x) \leqslant x \frac{e^{\bar{x}} - 1}{\bar{x}}$.

□

Let us define the optimal infinite horizon discounted objective V_α^*, defined as

$$V_\alpha^*(\mu_0) = \sup_{\pi \in \Pi} \sum_{t \geqslant 1} \alpha^{t-1} r(\pi_t(\mu_t), \mu_t) \ , \tag{5.8}$$

where α is the discount factor and μ_0 is the initial distribution. As a consequence of Lemma 5.2.2, we obtain that the value functions of the discounted problems constitute an equi-Lipschitz family:

Lemma 5.2.3 (Equi-Lipschitz property)
Assume that (A1)-(A3) hold. Then, $(V_\alpha^*)_{\alpha \in (0,1)}$ is $\left(\frac{\kappa M_r^{\mathcal{D}}}{1-\kappa}\right)$-equi-Lipschitz on $\mathcal{D}_L$ for the Hilbert metric, i.e.,

$$\forall \mu_0, \nu_0 \in \mathcal{D}_L, \ |V_\alpha^*(\mu_0) - V_\alpha^*(\nu_0)| \leqslant \frac{\kappa M_r^{\mathcal{D}}}{1 - \kappa} d_H(\mu^0, \nu^0) \ .$$

Proof. We first make the proof under the stronger assumption (A2'), and then deduce the general result.

Let a be the sequence of actions derived from an ϵ-optimal policy π and initial condition $\mu_0 \in \mathcal{D}_1$. Then, for $\nu_0 \in \mathcal{D}_1$

$$V_\alpha^*(\mu_0) - V_\alpha^*(\nu_0) \leqslant \sum_{t \geqslant 1} \alpha^{t-1} \begin{bmatrix} r\left(a_t, \phi_t(a_{:t}, \mu_0)\right) \\ - r\left(a_t, \phi_t(a_{:t}, \nu_0)\right) \end{bmatrix} + \epsilon \ .$$

The total reward is (NM_r)-Lipschitz for the infinite norm. Therefore, using Lemma 5.2.2, $\mu \mapsto r(a,\mu)$ is Lipschitz of constant $M_r^{\mathcal{D}} := \frac{1}{K} M_r \Upsilon(\mathrm{Diam}_H(\mathcal{D}_1))$ for the Hilbert metric. Hence,

$$V_\alpha^*(\mu_0) - V_\alpha^*(\nu_0) \leqslant M_r^{\mathcal{D}} \sum_{t \geqslant 1} \alpha^{t-1} d_H(\phi_t(a_{:t},\mu_0), \phi_t(a_{:t},\nu_0)) \\ + \epsilon \ .$$

From the Birkhoff theorem, one can derive that $d_H(\mu P(a), \nu P(a)) \leqslant \kappa\, d_H(\mu,\nu)$ for $\mu,\nu \in \mathcal{D}_1$, $a \in \mathcal{A}$, where $\kappa = \max_{a\in\mathcal{A}} \kappa(P(a)) < 1$. As a consequence, $d_H(\phi_t(a_{:t},\mu_0), \phi_t(a_{:t},\nu_0)) \leqslant \kappa^t d_H(\mu_0,\nu_0)$ and

$$\begin{aligned} V_\alpha^*(\mu_0) - V_\alpha^*(\nu_0) &\leqslant M_r^{\mathcal{D}} \sum_{t\geqslant 1} \alpha^{t-1}\kappa^t d_H(\mu^0,\nu^0) + \epsilon \\ &\leqslant \frac{\kappa M_r^{\mathcal{D}}}{1-\alpha\kappa} d_H(\mu^0,\nu^0) + \epsilon \leqslant \frac{\kappa M_r^{\mathcal{D}}}{1-\kappa} d_H(\mu^0,\nu^0) + \epsilon \ . \end{aligned}$$

The value function V_α^* is therefore $\left(\frac{\kappa M_r^{\mathcal{D}}}{1-\kappa}\right)$-equi-Lipschitz for the Hilbert metric.

To deduce the general result with (A2), we define

- $\widetilde{\mathcal{A}} := \mathcal{A}^L$, $\tilde{\alpha} := \alpha^L$,
- $\tilde{\phi}_\tau(\tilde{a}_{:\tau},\mu_0) := \mu_0 \prod_{1\leqslant t\leqslant \tau} Q(\tilde{a}_t)$,
- $\tilde{r}(a_{:L},\mu) := \sum_{l\in[L]} \alpha^{l-1} r(a_l, \phi_l(a_{:l},\mu))$,
- and $\tilde{\mathcal{B}} : V \mapsto \max_{\tilde{a}\in\widetilde{\mathcal{A}}} \{\tilde{r}(\tilde{a},\mu) + V(\nu) \mid \nu = \mu Q_L(\tilde{a})\}$.

and observe that

$$V_\alpha^*(\mu_0) = \sum_{\tau\geqslant 1} \tilde{\alpha}^{\tau-1} \tilde{r}(\tilde{a}_\tau, \tilde{\phi}_\tau(\tilde{a}_{:\tau},\mu_0)) \ .$$

We have rescaled the time (τ instead of t) so that the transition matrix between time τ and time $\tau+1$ is $Q_L(\tilde{a}_\tau)$. One τ-time step corresponds to L t-time steps. □

We are now able to prove the main result:

Theorem 5.2.2 (Existence of a solution)
Assume that (A1)-(A3) hold. Then, the ergodic eigenproblem

$$g\, 1_{\mathcal{D}_L} + h = \mathcal{B}\, h \tag{5.9}$$

admits a solution $h^* \in \mathrm{Lip}(\mathcal{D}_L) \cap \mathrm{Vex}(\mathcal{D}_L)$ and $g^* \in \mathbb{R}$.

Proof. Let us define a reference distribution $\overline{\mu} \in \Delta_N^K$, $g_\alpha^* = (1-\alpha)V_\alpha^*(\overline{\mu})$ and $h_\alpha^* = V_\alpha^* - V_\alpha^*(\overline{\mu})\, 1_{\mathcal{D}_L}$. Then, as V_α^* is equi-Lipschitz on $\mathcal{D}_L$ (Lemma 5.2.3), h_α^* is equi-bounded and equi-Lipschitz on $\mathcal{D}_L$ (in particular equi-continuous). By the Arzelà-Ascoli theorem, $h_\alpha^* \to h^* \in \mathcal{C}^0(\mathcal{D}_L)$.

Finally, from the discounted reward approach, we get $\mathcal{B}(\alpha V_\alpha^*) = V_\alpha^*$, therefore

$$\frac{g_\alpha^*}{1-\alpha}\, 1_{\mathcal{D}_L} + h_\alpha^* = \mathcal{B}\left(\frac{\alpha g_\alpha^*}{1-\alpha}\, 1_{\mathcal{D}_L} + \alpha h_\alpha^*\right) \ .$$

By the additive homogeneity property of the Bellman function, $g_\alpha^*\, 1_{\mathcal{D}_L} + h_\alpha^* = \mathcal{B}(\alpha h_\alpha^*)$. The fixed-point equation (5.29) is then obtained by continuity of the Bellman operator $\mathcal{B}$.

To conclude, h^* is convex since V_α^* is convex and the pointwise convergence preserves the convexity.

□

5. Switching costs

Proposition 5.2.1

For any solution (g^*, h^*) of (5.29), g^* satisfies (5.3), and a maximizer $a^*(\cdot) \in \arg\max \mathcal{B}\, h^*$ defines an optimal *stationary* policy for the average gain problem.

Proof. Let $\pi \in \Pi$ be a policy. By definition, for every t, $\mathcal{B}^{\pi_t}\, h^* \leqslant \mathcal{B}\, h^* = h^* + g^*\, 1_{\mathcal{D}_L}$. Therefore, iterating the Kolmogorov operator, we obtain

$$(\mathcal{B}^{\pi_1} \circ \ldots \circ \mathcal{B}^{\pi_t}) h^* \leqslant h^* + t g^*\, 1_{\mathcal{D}_L} \quad .$$

Let $\underline{h}^* := \min_{\mu \in \mathcal{D}_L} h^*(\mu)$ be the minimum of h^*. Then, $0_{\mathcal{D}_L} \leqslant h - \underline{h}^*\, 1_{\mathcal{D}_L}$, and so $(\mathcal{B}^{\pi_1} \circ \ldots \circ \mathcal{B}^{\pi_t})(0_{\mathcal{D}_L}) \leqslant h^* + (t g^* - \underline{h}^*)\, 1_{\mathcal{D}_L}$. Finally,

$$\liminf_{t \to \infty} \frac{1}{t} (\mathcal{B}^{\pi_1} \circ \ldots \circ \mathcal{B}^{\pi_t})(0_{\mathcal{D}_L})(\mu_0) \leqslant g^* \quad .$$

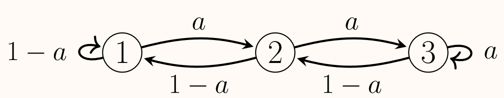
Any strategy has an average reward lower than g^*. As we have proved that the bias function h^* is continuous on $\mathcal{D}_L$, a maximizer $a^*(\mu)$ can be found for any state μ, and so playing the strategy $a^*(\mu)$ achieves the best possible average gain g^*. $\square$

In particular, the constant g^* in (5.29) is unique, and it coincides with the optimal average long-term reward, for all choices of the initial state μ_0. However, even if the payoff g^* is unique, the bias function h^* is not (and so neither is the optimal policy).

Non-uniqueness of the solution

In contrast with results where geometric ergodicity is assumed to guarantee the existence (and the uniqueness) of the eigenvector, see e.g. [Bis15; HL96], the uniqueness of the latter is generally not true. Indeed, we can construct instances where there exist several "attractor" states, and where a family of strategies can be found so that each of them secures the optimal mean payoff. In these instances, the eigenvector does exist but is not unique. To illustrate this fact, we introduce in Example 5.2.1 a model satisfying (A1) to (A3). Note that taking the same dynamics as in the example but without node 2 can also lead to a non-unique solution of the eigenroblem as long as we allow for a more general form of reward $r(a, \mu)$. Here, we aim at fitting exactly with our application case by considering that the reward function satisfies (5.1).

Example 5.2.1 (Non-uniqueness of the eigenvector)



Let us consider the dynamical system described by the following transition matrix:

$$P(a) = \begin{bmatrix} 1-a & a & 0 \\ 1-a & 0 & a \\ 0 & 1-a & a \end{bmatrix} ,$$

where the a is supposed to belong to the action space $\mathcal{A}$, which is of the form

$$\mathcal{A} = [a_0, a_1] \ , \quad 0 < a_0 < 1/2 \ \text{ and } \ a_1 = 1 - a_0 \quad .$$

We consider the following unitary reward $\theta(\cdot)$:

$$\theta(a)_n = \begin{cases} 1-a \text{ if } n=1, \\ 0 \quad \text{if } n=2, \\ a \quad \text{if } n=3, \end{cases}$$

The reward function is then $r(a,\mu) = (1-a)\mu_1 + a\mu_3$ for any $\mu \in \Delta_3$, and

$$r(a,\mu P(a)) = (1-a)^2(1-\mu_3) + a^2(1-\mu_1) \ .$$

In the sequel, we work in the sub-simplex $\Delta_3^{\leqslant} := \{(x,y) \in \mathbb{R}_+^2 \mid x+y \leqslant 1\}$, considering that μ_2 can be reconstructed as $\mu_2 = 1-\mu_1-\mu_3$.

The associated ergodic eigen problem. For any real-valued function $v : \Delta_3^{\leqslant} \to \mathbb{R}$, let us define the Bellman operator $\mathcal{B}$ as

$$\mathcal{B}\, v(\mu_1,\mu_3) = \max_{a\in\mathcal{A}} \left\{ \begin{matrix} (1-a)^2(1-\mu_3) + a^2(1-\mu_1) \\ +\, v\left((1-a)(1-\mu_3), a(1-\mu_1)\right) \end{matrix} \right\} \ .$$

In Example 5.2.1, the transition $a \in \mathcal{A} \mapsto P(a)$ is linear. Moreover, the transition matrix over two time steps is then

$$(P(a))^2 = \begin{bmatrix} (1-a)^2 + a(1-a) & a(1-a) & a^2 \\ (1-a)^2 & 2a(1-a) & a^2 \\ (1-a)^2 & a(1-a) & a^2 + a(1-a) \end{bmatrix}$$

and has positive coefficients. Therefore, the transition matrix $P(a)$ satisfies the primitivity assumption (A2) for all $a \in \mathcal{A}$. Using Theorem 5.2.2, the ergodic eigenproblem

$$g 1_{\mathcal{D}_1} + h = \mathcal{B}\, h \tag{5.10}$$

admits a solution $h^* \in \mathrm{Lip}(\mathcal{D}_1) \cap \mathrm{Vex}(\mathcal{D}_1)$ and $g^* \in \mathbb{R}$, where $\mathcal{D}_1$ is defined one can construct the effective domain $\mathcal{D}_1$ as in (5.4). As the quantity in the maximum is convex, for any convex function $v : \mathcal{D}_1 \to \mathbb{R}$, the maximum value in $\mathcal{B}\, v$ is obtained for $a = a_0$ or $a = a_1$. Therefore, in the sequel, we restrict wlog the state space to be $\mathcal{A} = \{a_0, a_1\}$.

Steady states. Let $k \in \{0,1\}$. The equilibrium distribution $\hat{\mu}$ achieved by a constant decision a_k is given by the equation $\hat{\mu} P(a_k) = \hat{\mu}$, which has a unique solution:

$$\hat{\mu}_1^k = \frac{(1-a_k)^2}{1-a_k(1-a_k)}, \ \hat{\mu}_3^k = \frac{a_k^2}{1-a_k(1-a_k)} \ . \tag{5.11}$$

Bias function for the Kolmogorov operator. Let us define $\mathcal{B}^k$ the Kolmogorov operator associated to the constant strategy $\pi : \mu \mapsto a_k$, i.e.,

$$\mathcal{B}^k\, v(\mu_1,\mu_3) = (1-a_k)^2(1-\mu_3) + a_k^2(1-\mu_1) + v\left((1-a_k)(1-\mu_3), a_k(1-\mu_1)\right) \ .$$

Then, the linear function $h^k(\mu_1,\mu_3) = \alpha^k \mu_1 + \beta^k \mu_3$ and the gain g^k are solutions of

$$h^k(\mu_1,\mu_3) + g^k = \mathcal{B}^k\, h^k(\mu_1,\mu_3), \ (\mu_1,\mu_3) \in \mathcal{D}_1 \tag{5.12}$$

5. Switching costs

if and only g^k, α^k and β^k satisfy the following system

$$\begin{cases} g^k = (1-a_k)^2 + a_k^2 + (1-a_k)\alpha^k + a_k\beta^k \\ \alpha^k = -a_k^2 - a_k\beta^k \\ \beta^k = -(1-a_k)^2 - (1-a_k)\alpha^k \end{cases},$$

where the unique solution of the latter system is given by

$$g^k = \frac{a_k^3 + (1-a_k)^3}{1-a_k(1-a_k)},\ \alpha^k = \frac{a_k(1-a_k)^2 - a_k^2}{1-a_k(1-a_k)},\ \beta^k = \frac{(1-a_k)a_k^2 - (1-a_k)^2}{1-a_k(1-a_k)} \ . \tag{5.13}$$

Note that $g^0 = g^1$ since $a_0 + a_1 = 1$, and we simply denoted it by g^*.

Solution for the ergodic eigen problem. We now exhibit a family of eigenvectors where each of them constitutes a solution to the ergodic eigenproblem associated with Example 5.2.1:

Theorem 5.2.3 (Non-uniqueness of the eigenvector)
Let $v^k : \mathcal{D} \to \mathbb{R}$, $k \in \{0,1\}$, be defined as

$$v^k(\mu_1,\mu_3) = \hat{h}^{k0}(\mu_1,\mu_3) \vee \hat{h}^{k1}(\mu_1,\mu_3) \vee \hat{h}^{k2}(\mu_1,\mu_3) \ , \tag{5.14}$$

with

- $\hat{h}^{ij}(\cdot,\cdot) := h^j(\cdot,\cdot) - h^j(\hat{\mu}_1^i,\hat{\mu}_3^i),\ i,j \in \{0,1\}$,
- $\hat{h}^{k2}(\cdot,\cdot) := \mathcal{B}^{1-k}\,\hat{h}^{kk}(\cdot,\cdot) - g^*$.

Then, for any $\lambda \in [0,1]$, the couple (v^λ, g^*) is solution the ergodic eigenproblem (5.10) – corresponding to Example 5.2.1 – with g^* defined in (5.13) and

$$v^\lambda(\mu_1,\mu_3) := \left(v^0(\mu_1,\mu_3) - \tfrac{\lambda}{1-\lambda}\right) \vee \left(v^1(\mu_1,\mu_3) - \tfrac{1-\lambda}{\lambda}\right) \ .$$

Proof. As first observation, the couple (h^k, g^*), solution of (5.12), is not solution of the ergodic eigenproblem (5.10). Therefore, let us try to construct a solution as a mixture of h^0 and h^1. To this purpose, let us define the function the function $u^0 : \mathcal{D}_1 \to \mathbb{R}$ as

$$u^0(\mu_1,\mu_3) = \hat{h}^{00}(\mu_1,\mu_3) \vee \hat{h}^{01}(\mu_1,\mu_3) \ .$$

For $(\mu_1,\mu_3) \in \mathcal{D}_1$, the value of $\mathcal{B}\,u^0(\mu_1,\mu_3)$ is given by the maximum of 4 quantities:

(i) $\mathcal{B}^0\,\hat{h}^{00}(\mu_1,\mu_3) = \hat{h}^{00}(\mu_1,\mu_3) + g^*$,

(ii) $\mathcal{B}^0\,\hat{h}^{01}(\mu_1,\mu_3)$,

(iii) $\mathcal{B}^1\,\hat{h}^{00}(\mu_1,\mu_3)$,

(iv) $\mathcal{B}^1\,\hat{h}^{01}(\mu_1,\mu_3) = \hat{h}^{01}(\mu_1,\mu_3) + g^*$,

The equality in (i) and (iv) comes from the fact that $\hat{h}^{00}$ and $\hat{h}^{01}$ are solutions for (5.12). Besides, we can prove using basic algebra that

$$\mathcal{B}^0\,\hat{h}^{01}(\mu_1,\mu_3) - \mathcal{B}^1\,\hat{h}^{00}(\mu_1,\mu_3) = g^*(\hat{\mu}_3^0 - \hat{\mu}_1^0) \leqslant 0 \ .$$

Therefore, the maximum is obtained either with (i), (iii) or (iv). We consider now the function

$$v^0(\mu_1, \mu_3) = \hat{h}^{00}(\mu_1, \mu_3) \vee \hat{h}^{01}(\mu_1, \mu_3) \vee \hat{h}^{02}(\mu_1, \mu_3) \ , \tag{5.15}$$

with $\hat{h}^{02}(\mu_1, \mu_3) := \mathcal{B}^1 \hat{h}^{00}(\mu_1, \mu_3) - g^*$. By construction, $v^0 = \mathcal{B}\, u^0 - g^*$. Moreover, one can show that $\mathcal{B}\, \hat{h}^{02}(\mu_1, \mu_3) - g^* \leqslant v^0(\mu_1, \mu_3)$. Therefore,

$$\forall (\mu_1, \mu_3) \in \mathcal{D}_1, \quad \mathcal{B}\, v^0(\mu_1, \mu_3) = \mathcal{B}\, u^0(\mu_1, \mu_3) = v^0(\mu_1, \mu_3) - g^* \ .$$

As a conclusion, (v^0, g^*) is a solution of (5.10).

By symmetry of the problem, we can construct the function $v^1(\mu_1, \mu_3) = v^0(\mu_3, \mu_1)$, and (v^1, g^*) is a different solution of (5.10). Finally, each max-plus combination of v^0 and v^1 also constitutes a solution of the ergodic eigenproblem. □

We display in Figure 5.2 the eigenvector v^0, $v^{1/2}$ and v^1, obtained numerically (using the RVI procedure, see Algorithm 4), as with the eigenvector v^0, obtained theoretically (see above).

Remark 5.2.2

We continue the analysis of Example 5.2.1 in Section 7.5, where the convergence of the dynamical systems under optimal policies is studied through the weak KAM angle, showing in particular that the projected Aubry set is reduced to two points, corresponding to the states defined in (5.11). From this result, we will prove that any eigenvector is a tropical linear combination of the two basis eigenvectors v^0 and v^1, showing that the family $\{v^\lambda\}_{\lambda \in [0,1]}$ describes all the eigenvectors that are solutions of the ergodic eigenproblem.

5.3 Steady-state optimality

5.3.1 Definition

The solution of dynamic programming problems, including the ergodic eigenproblem (5.29), is subject to the "curse of dimensionality" . Therefore, it is of interest to investigate cases in which the dynamic problem reduces to a static one. In fact, in some cases the optimal stationary policy may be a simple policy that attracts the system to a steady-state ("get there, stay there" – [Fly79]). We next formalize this property:

Definition 5.3.1. Let $\mathcal{S} = \{(a, \mu) \in \mathcal{A} \times \Delta_N^K \,|\, \mu = \mu P(a)\}$ be the action-space domain of stationary probabilities. Then, $\mu \in \Delta_N^K$ is a *steady-state* if there exists $a \in \mathcal{A}$ such that $(a, \mu) \in \mathcal{S}$.

If (A2) holds, then for any cluster k and any price $a \in \mathcal{A}$, the Markov chain induced by the transition matrix $P^k(a)$ has a unique stationary distribution. We denote by $\overline{\mu}(\cdot) : \mathcal{A} \mapsto \Delta_N^K$ the mapping sending an action to the stationary distribution it induces.

Definition 5.3.2. The *optimal steady-state gain* $\overline{g}$ is defined as

$$\overline{g} := \max_{(a,\mu) \in \mathcal{S}} r(a, \mu) \ . \tag{5.16}$$

If (A2) holds, (5.16) is in general a static nonconvex maximization problem over the actions. Nonetheless, we can expect to solve it efficiently in the case where $\overline{\mu}(\cdot)$ is analytically known, see e.g. Section 5.4. Maximizers $\overline{a}$ are called *optimal steady-state price*, they correspond to a steady-state distribution $\overline{\mu}(\overline{a})$.

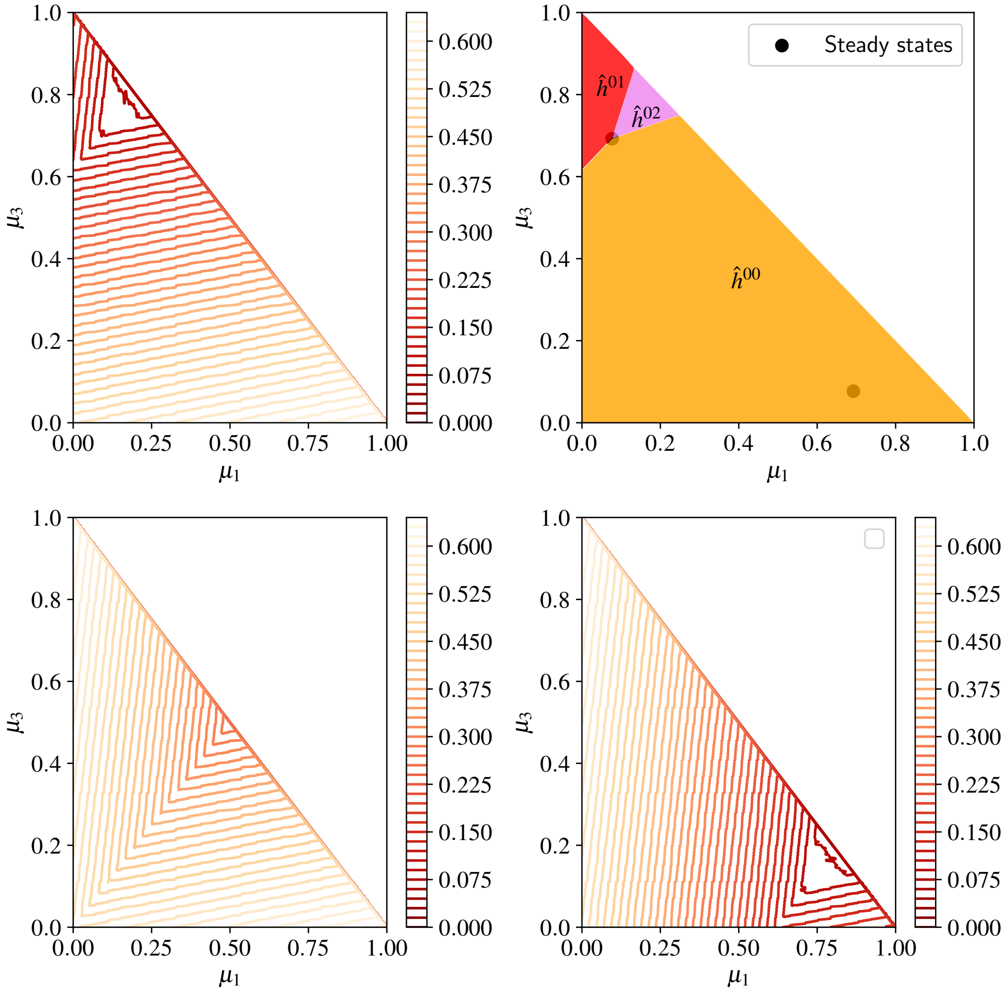


Figure 5.2: Eigenvectors for Example 5.2.1 with $\mathcal{A} = [0.25, 0.75]$.
On the upper left, we display the eigenvector v^0 (obtained by the RVI algorithm). On the upper right, we display the theoretical v^0 found in (5.15) as with the two steady states to which each of optimal strategies converges. On the lower left, we display the eigenvector $v^{1/2}$ and on the lower right, we display the eigenvector v^1.

5.3.2 Optimality gap

In this section we introduce a class of Lagrangian functions designed so that each dual problem turns out to be an upper bound of g^*. This extends the result of [Fly79] involving usual Lagrangian functions. We use here a more general Lagrangian, depending on the choice of a non-linear function φ. This leads to much tighter bounds, and allows us to prove the optimality of a steady-state strategy whenever a zero duality gap is obtained. Let Φ be defined as

$$\Phi = \{\varphi : \Delta_N^K \to \Delta_N^K \text{ injective and bounded}\} \ .$$

For a given function $\varphi \in \Phi$, we define the Lagrangian function $\mathcal{L}^{(\varphi)} : (\mathcal{A}, \Delta_N^K, \mathbb{R}^{KN}) \to \mathbb{R}$ by

$$\mathcal{L}^{(\varphi)}(a, \mu, \lambda) := r(a, \mu P(a)) + \langle \lambda, \varphi\left(\mu P(a)\right) - \varphi(\mu) \rangle_{KN} \ .$$

As a direct consequence of the injectivity of φ, we obtain that for any given $\varphi \in \Phi$,

$$\overline{g} = \max_{(a,\mu) \in \mathcal{A} \times \Delta_N^K} \inf_{\lambda \in \mathbb{R}^{KN}} \mathcal{L}^{(\varphi)}(a, \mu, \lambda) \ .$$

We also define the dual problem $g^{(\varphi)}$ as

$$g^{(\varphi)} := \inf_{\lambda \in \mathbb{R}^{KN}} \max_{(a,\mu)\in\mathcal{A}\times\Delta_N^K} \mathcal{L}^{(\varphi)}(a,\mu,\lambda) \ . \tag{5.17}$$

Proposition 5.3.1
With (g^*, h^*) solution of (5.29) and $\overline{g}$ defined in (5.16),

$$\overline{g} \leqslant g^* \leqslant g^{(\varphi)}, \ \forall \varphi \in \Phi \ .$$

Proof. The proof extends the arguments in [Fly79, Remark 5.1] to nonlinear functions $\varphi \in \Phi$.

First, from the geometrical convergence the dynamic (see Theorem 5.2.1), the valid strategy consisting in executing action $\overline{a}$ each period of time induces an average reward of $\overline{g}$, regardless the initial distribution. Therefore, $\overline{g} \leqslant g^*$.

Then, for $\epsilon > 0$, there exists λ^ϵ such that for any $(a, \mu) \in \mathcal{A} \times \Delta_N^k$,

$$r(a, \mu P(a)) + \langle \lambda^\epsilon, \varphi(\mu P(a)) - \varphi(\mu) \rangle_{KN} \leqslant g^{(\varphi)} + \epsilon \ .$$

We construct a sequence of decision $a_1, \ldots, a_T$ leading to distribution $\mu_1, \ldots, \mu_T$. Then, at each period t,

$$r(a_t, \mu_t) + \langle \lambda^\epsilon, \varphi(\mu_t) - \varphi(\mu_{t-1}) \rangle_{KN} \leqslant g^{(\varphi)} + \epsilon \ .$$

Therefore, we take the mean over $t = 1, \ldots, T$ to recover the average reward criteria:

$$\frac{1}{T}\sum_{t=1}^{T} r(a_t, \mu_t) + \frac{1}{T}\langle \lambda^\epsilon, \varphi(\mu_T) - \varphi(\mu_0) \rangle_{KN} \leqslant g^{(\varphi)} + \epsilon \ .$$

The second term converges to zero when $T \to \infty$ as we suppose that φ is bounded on the simplex. So,

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} r(a_t, \mu_t) \leqslant g^{(\varphi)} + \epsilon \ .$$

The latter inequality is valid for any $\epsilon > 0$, and any sequence of action $(a_t)_{t\in\mathbb{N}}$, so $g^* \leqslant g^{(\varphi)}$. $\square$

We define the duality gap $\delta_{\mathcal{L}^{(\varphi)}}$ as

$$\delta_{\mathcal{L}^{(\varphi)}} := g^{(\varphi)} - \overline{g}.$$

As an immediate consequence of Proposition 5.3.1, if there exists $\varphi \in \Phi$ such that $\delta_{\mathcal{L}^{(\varphi)}} = 0$, then $g^* = \overline{g}$, and the dynamic program 5.3 reduces to the static optimization program (5.16). Depending on the problem parameters, the duality gap may, or may not, vanish, see Figure 5.6.

5.4 Application to electricity pricing

We suppose that an electricity provider has N-1 different types of offers and that a study has distinguished beforehand K customer segments, assuming that customers of a given segment have approximately the same behavior. Given a segment k and an offer $n \in [N\text{-}1]$, the *reservation price* R^{kn} is the maximum price that customers of this segment are willing to spend on n, and E^{kn} is the (fixed) quantity a customer of segment k will purchase if he chooses n. The *utility* for these customers is linear and is defined as

$$U^{kn}(a) := R^{kn} - E^{kn} a^n \ .$$

where a^n is the price for one unit of product n. The action space is then a compact subset of $\mathbb{R}^{N\text{-}1}$.

To model the competition between the provider and the other providers of the market, consumers have an alternative option (state of index N). We suppose that this alternative offer is fixed over time (for example a regulated contract). Then, under this assumption, it can be modelized w.l.o.g. by a null utility for each cluster ($U^{kN} = 0$).

If a customer of segment k chooses the contract $n < N$ at price a^n, then the provider receives $E^{kn}a^n$ from the electricity consumption of the customer and has an induced cost of C^{kn}. Note that the cost should depend on the quantity E^{kn}, but as it is supposed to be a parameter, we omit this dependency. The (linear) reward for the provider is then

$$\theta^{kn}(a) = E^{kn}a^n - C^{kn},\ n < N, \quad \theta^{kN} = 0\ .$$

We suppose that the transition probability follows a logit response, see e.g. [PE17]:

$$[P^k(a)]_{n,m} = \frac{e^{\beta[U^{km}(a)+\gamma^{kn}\,1_{m=n}]}}{\sum_{l\in[N]} e^{\beta[U^{kl}(a)+\gamma^{kn}\,1_{l=n}]}}\ , \tag{5.18}$$

where the parameter γ^{kn} is the cost for segment k to switch from contract n to another one, and β is the intensity of the choice (it can represent a "rationality parameter"). One can easily check that (A1)-(A3) are satisfied.

In the no-switching-cost case ($\gamma = 0$), we say that the customers response is *instantaneous*, and corresponds to the classical logit distribution, see e.g. [Tra09]:

$$\mu_L^{kn} = e^{\beta U^{kn}(a)} \Big/ \sum_{l\in[N]} e^{\beta U^{kl}(a)}\ . \tag{5.19}$$

Steady-states. The application scope of the transition model we defined in (5.18) is broader than electricity pricing. For this specific kernel, we derive a closed-form expression for the stationary distributions, fully characterized by the instantaneous response:

Theorem 5.4.1

Given a constant action a, the distribution μ_t^k converges to $\overline{\mu}^k(a)$, defined as

$$\overline{\mu}^{kn}(a) = \frac{\eta^{kn}(a)\mu_L^{kn}(a)}{\sum_{l\in[N]}\eta^{kl}(a)\mu_L^{kl}(a)}\ . \tag{5.20}$$

where $\eta^{kn}(a) := 1 + \left[e^{\beta\gamma^{kn}} - 1\right]\mu_L^{kn}(a)$, and μ_L is defined in (5.19).

Proof. In the proof, we forget the dependence on k and a. The stationary probability is defined as $\forall m \in [N]$, $\mu^m\,[1 - P^{mm}] = \sum_{n\neq m}\mu^n P^{nm}$. We can then replace by the definition of the probabilities (5.18) to obtain

$$\mu^m\left[\frac{\sum_{l\neq m} e^{\beta U^m}}{\sum_l e^{\beta[U^l + 1_{l=m}\,\gamma^m]}}\right] = \sum_{n\neq m}\mu^n\left[\frac{e^{\beta U^m}}{\sum_l e^{\beta[U^l+1_{l=n}\,\gamma^n]}}\right]\ .$$

Defining $\tilde{\mu}^n := \frac{\mu^n}{\sum_l e^{\beta[U^l+1_{l=n}\,\gamma^n]}}$, we obtain

$$\forall m \in [N],\ \tilde{\mu}^m \sum_{l\neq m} e^{\beta U^l} = e^{\beta U^m}\sum_{l\neq m}\tilde{\mu}^l\ .$$

The solution $\tilde{\mu}^n := \lambda e^{\beta U^n}, n \in [N]$ is then a valid solution, and the constant λ is chosen so that $\sum_{l\in[N]} \mu^l = 1$:

$$\begin{aligned} \overline{\mu}^{kn}(a) &= \lambda e^{\beta U^{kn}(a)} \sum_{m\in[N]} e^{\beta[U^{kn}(a)+1_{m=n}\gamma^{kn}]} \\ \lambda^{-1} &= \sum_{n\in[N]} e^{\beta U^{kn}(a)} \sum_{m\in[N]} e^{\beta[U^{km}(a)+1_{m=n}\gamma^{kn}]} \end{aligned} \tag{5.21}$$

Finally, $\eta^{kn} = \sum_l e^{\beta[U^{kl}+1_{l=n}\gamma^{kn}]}/\sum_l e^{\beta U^{kl}}$. We recover the definition of $\overline{\mu}$ (5.21). □

As a consequence, the optimal steady-state can be found by solving

$$\overline{g} = \max_{a\in\mathcal{A}} r(a, \overline{\mu}(a)) \ . \tag{5.22}$$

Problem (5.22) has no guarantee to be convex. However, it is a box-constrained smooth optimization problem which can be much more efficiently solved (at least up to local maximum) than the original time-dependent problem.

In addition, if we suppose that $\gamma^{kn} = \gamma^k > 0$ for all n, then for any $a \in \mathcal{A}$, we get the two following properties as immediate consequence of Theorem 5.4.1:

- $\lim_{\gamma^k\to 0} \overline{\mu}^k(a) = \mu_L^k(a)$,
- $(\overline{\mu}^{kn})$ and (μ_L^{kn}) are sorted in the same order.

We now aim to compare the steady-state $\overline{\mu}^k(a)$ with the logit distribution $\mu_L^k(a)$ using the majorization theory:

Definition 5.4.1 (Majorization,[MOA11]). For a vector $a \in \mathbb{R}^d$, we denote by $a^{\downarrow} \in \mathbb{R}^d$ the vector with the same components, but sorted in descending order. Given $a, b \in \Delta_d$, we say that a majorizes b from below written $a \succ b$ iff

$$\sum_{i=1}^k a_i^{\downarrow} \geqslant \sum_{i=1}^k b_i^{\downarrow} \quad \text{for } k = 1, \dots, d \ .$$

Proposition 5.4.1 (Majorization property of the steady-state)
Let $k \in [K]$ and $a \in \mathcal{A}$ be given. Suppose that $\gamma^{kn} = \gamma^k > 0$ for all $n \in [N]$, then the stationary distribution majorizes the instantaneous logit response i.e.,

$$\overline{\mu}^k(a) \succ \mu_L^k(a) \ . \tag{5.23}$$

Proof. Let us suppose that we reorder the probabilities (and the η) such that they are sorted in the decreasing order.

$$\begin{aligned} \left(\sum_{m=1}^n \overline{\mu}^m\right)^{-1} &= \frac{\sum_{l=1}^n \eta^l \mu_L^l + \sum_{l=n+1}^N \eta^l \mu_L^l}{\sum_{m=1}^n \eta^m \mu_L^m} \\ &= 1 + \frac{\sum_{l=n+1}^N \eta^l \mu_L^l}{\sum_{m=1}^n \eta^m \mu_L^m} \\ &\leqslant 1 + \frac{\sum_{l=n+1}^N \mu_L^l}{\sum_{m=1}^n \mu_L^m} = \left(\sum_{m=1}^n \mu_L^m\right)^{-1} \ . \end{aligned}$$

The inequality comes from the sorting of η, and the last equality from $\sum \mu_L = 1$. □

Proposition 5.4.1 establishes a qualitative feature of this model: if the price is kept constant over time, then, in the model with inertia, the stationnary distribution of population *majorized* the one obtained in the corresponding logit-model without inertia. Recalling that the majorization order expresses a form of dispersion, this means that inertia increases the concentration of the population on its favorite offers.

Lemma 5.4.1

Let us consider a and b in Δ_d. If $a \succ b$, then for all i, $a_i \leqslant db_i$.

Proof. $a_i^{\downarrow} \leqslant \sum_{j=i}^{d} a_j^{\downarrow} = 1 - \sum_{j=1}^{i-1} a_j^{\downarrow} \leqslant 1 - \sum_{j=1}^{i-1} b_j^{\downarrow} = \sum_{j=i}^{d} b_j^{\downarrow} \leqslant (d-i+1)b_i^{\downarrow} \leqslant db_i^{\downarrow}$. □

Proposition 5.4.2 (Boundedness of the steady-state gain)

Even with $\mathcal{A} = \mathbb{R}^{N-1}$, the optimal steady-state gain $\overline{g}$ is bounded independently of γ.

Proof. Suppose that the optimal steady-state gain is attained for an action a, then

$$
\begin{aligned}
\overline{g} = \sum_{k\in[K]} \rho_k \sum_{n\in[N]} (E^{kn}a^n - C^{kn})\overline{\mu}^{kn}(a) &\leqslant \max_{k,n}(R^{kn} - C^{kn}) + \sum_{k\in[K]} \rho_k \sum_{n\in[N]} (E^{kn}a^n - R^{kn})\overline{\mu}^{kn}(a) \\
&\leqslant \max_{k,n}(R^{kn} - C^{kn}) + \sum_{k\in[K]:U^{kn}(a)<0} \rho_k \left\langle -U^k(a), \overline{\mu}^k(a) \right\rangle_N \\
&\leqslant \max_{k,n}(R^{kn} - C^{kn}) + N \sum_{k\in[K]:U^{kn}(a)<0} \rho_k \left\langle -U^k(a), \mu_L^k(a) \right\rangle_N \\
&\leqslant \max_{k,n}(R^{kn} - C^{kn}) + \frac{N}{\beta e} \ .
\end{aligned}
$$

The third inequality comes from Lemma 5.4.1. For the fourth one, since the logit expression contains a no-purchase option, $\mu_L^{kn} \leqslant \frac{1}{1+e^{-\beta U^{kn}(a)}}$. To conclude, it remains to see that $1 + e^{\beta z} \geqslant (\beta e)z$ for all z. □

Proposition 5.4.2 proves that the optimal steady-state gain cannot diverge to infinity when the inertia grows. This qualitative result is no longer true for the optimal strategy (that may be a periodic sequence of actions instead of a single constant one), see Section 5.7.1.

5.5 Numerical resolution

5.5.1 Relative Value Iteration with Krasnoselskii-Mann damping

Relative Value Iteration (RVI) has been extensively studied to solve unichain finite-state MDP [Put94; Ber98]. Simplicial state-spaces appear in particular in the definition of *belief state* for partially observable MDP [Hau00]. For such continuous state-spaces, a discretization must be done as a prerequisite to RVI algorithm. Here, we define a regular grid Σ of the simplex Δ_N^K, and $\mathcal{B}^\Sigma$ the Bellman Operator with a linear point approximation on the grid Σ, achieved by a Freudenthal triangulation [Lov91].

With this simple framework, we have the following property:

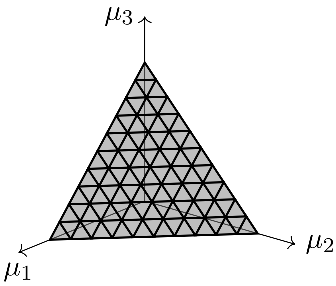


Figure 5.3: Freudenthal triangulation of Δ_3

Proposition 5.5.1 ([Hau00], Thm 12)
For any $v \in \mathrm{Vex}(\Delta_N^K)$,

$$\mathcal{B}\, v \leqslant \mathcal{B}^{\Sigma}\, v \quad .$$

As the bias function $\hat{h}$ is convex at each iteration, the solution returned by Algorithm 4 provides a gain which is an upper bound of the optimal gain g^*.

Algorithm 4 RVI with Mann-type iterates

Require: Grid Σ, Bellman operator $\mathcal{B}^{\Sigma}$, initial function $\hat{h}_0$
1: Initialize $\hat{h} = \hat{h}_0$, $\hat{h}'(\mu) = \mathcal{B}^{\Sigma}\, \hat{h}$
2: **while** $\mathrm{sp}(\hat{h}' - \hat{h}) > \epsilon$ **do**
3: $\quad \hat{h} \leftarrow (\, \hat{h}' - \max\{\hat{h}'\}e + \hat{h}\,)/2$
4: $\quad \hat{h}'(\hat{\mu}) \leftarrow (\mathcal{B}^{\Sigma}\, \hat{h})(\hat{\mu})$ for all $\hat{\mu} \in \Sigma$
5: **end while**
6: $\hat{g} \leftarrow (\, \max(\hat{h}' - \hat{h}) + \min(\hat{h}' - \hat{h})\,)/2$
7: **return** $\hat{g}, \hat{h}$

In Algorithm 4, we use, following [GS20], a mixture of the classical relative value iteration algorithm [Put94] with a *Krasnoselskii-Mann* damping. As detailed in [GS20] (Th. 9 and Coro 13), it follows from a theorem of Ishikawa that the sequence of bias function $\hat{h}$ does converge, and it follows from a theorem of Baillon and Bruck that $\hat{g}$ provides an ϵ approximation of the optimal average cost g^* after $O(1/\epsilon^2)$ iterations.

5.5.2 Howard algorithm with on-the-fly transition generation

With value-iteration algorithms, the second main class of iterative methods to solve MDPs are *policy iteration algorithms*, initiated by Howard (see e.g [DF68] of [Put94]). We describe here the method adapted to "decomposable" state spaces. Domain decomposition has already been exploited by Festa [Fes18] in a different setting to obtain parallel Howard's algorithm, as well as memory space gain.

Let $\Lambda = (\hat{\mu}_i)_{i \in [M]}$ be a *local* semi-Lagrangian discretization of the simplex Δ_N of size $M := |\Lambda|$. We refer the grid to be *local*, since the discretization is done for the probability space of one sub-population and not on the *global* probability space Δ_N^K. The *global* discretization is then

$$\Sigma = (\hat{\mu}_{\vec{i}_1}, \ldots, \hat{\mu}_{\vec{i}_K})_{\vec{i} \in [M]^K} \quad .$$

We define the local transition operator $T^{\Lambda,k} : (i,a) \in [M] \times \mathcal{A} \mapsto \arg\min_{j \in [M]} \|\hat{\mu}_i P^k(a) - \hat{\mu}_j\|_\infty$. For each $k \in [K]$, this operator can be computed in a preprocessing step, and stored in $O(M \times$

$|\mathcal{A}|$). Contrary to the RVI algorithm where a Freudenthal triangulation is performed during the computation of $\mathcal{B}^\Sigma$, Figure 5.4 illustrates that the transition operator is approximated by finding the closest discretization point (in the L_∞-norm) to the true next state.

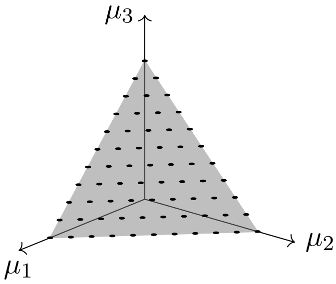


Figure 5.4: Semi-Lagrangian discretization of Δ_3

The *global* transition can then be obtained *on-the-fly*, i.e., for any action $a \in \mathcal{A}$ and global index $\vec{i} \in [M]^K$, $T^\Sigma(\vec{i}, a)$ can be recomputed whenever it is required in the algorithm knowing the sub-transition $T^{\Lambda,k}(\vec{i}_k, a)$ for all $k \in \mathbb{N}$:

$$T^\Sigma : (\vec{i}, a) \in [M]^K \times \mathcal{A} \mapsto (T^{\Lambda,k}(\vec{i}_k, a))_{k\in[K]} \ . \tag{5.24}$$

Remark 5.5.1

A complete storage of T^Σ would lead to a memory occupation in $O(M^K \times |\mathcal{A}|)$, whereas the storage of every $T^{\Lambda,k}$ is $O(K \times M \times |\mathcal{A}|)$.

Algorithm 5 shows the Howard algorithm with on-the-fly transition generation. It consists in alternating a policy evaluation step with a policy improvement step. We implemented a parallelized version of this algorithm[5] by adapting the code of [Coc+98][6], initially intended for computing spectral elements in max-plus algebra. The algorithm is known to have experimentally a superlinear convergence which, in finite action-space setting, is reached in finitely many steps, see e.g.[Put94]. Despite the decomposable transition, all the subpopulations $k \in [K]$ are linked together through a common policy. In the implementation, both the policy $\hat{d}$ and the bias function $\hat{h}$ depend on the global state $\vec{i} \in \Sigma$. Therefore, the memory needed to run the algorithm is still exponential in the number of segments – in $O(M^K)$ – but would have been worst with stored global transition T^Σ – in $O(M^K \times |\mathcal{A}|)$ – the action space being very large in general.

5.6 Numerical results

The numerical results were obtained on a laptop i7-1065G7 CPU@1.30GHz. We solved the problem up to dimension 4 (2 provider offers, 2 clusters) with high precision ($\delta_\mu = 50$ points for each dimension, 1.6 million discretization points, precision $\epsilon = 10^{-5}$) in 7 hours for Value-Iteration algorithm and in 70 seconds for the Howard algorithm adapted to decomposable state-spaces (both methods parallelized on 8 threads), see Table 5.2. The Policy Iteration algorithm adapted to decomposable state-spaces appears as the faster algorithm while occupying a reasonable memory space.

[5] Available at `https://gitlab.com/these_tarif/ergodic_inertia`
[6] Available at `http://www.cmap.polytechnique.fr/~gaubert/HOWARD2.html`

Algorithm 5 Howard Algorithm with on-the-fly transition generation

Require: Local grid Λ, family of local transitions $(T^{\Lambda,k})_{k\in[K]}$, initial decision rule $\hat{d}'$
1: **do**
2: $\quad \hat{d} \leftarrow \hat{d}'$
3: $\quad \hat{g}, \hat{h}$ solution of $\begin{cases} \hat{g} + \hat{h}_{\vec{i}} = r(\hat{d}_{\vec{i}}, \hat{\mu}_{\vec{i}}) + \hat{h}_{\vec{j}},\ \vec{i} \in \Sigma \\ \vec{j} = T^{\Sigma}(\vec{i}, \hat{d}_{\vec{i}}) \end{cases}$ ▷ Policy Evaluation
4: $\quad \hat{d}'_{\vec{i}} \leftarrow \arg\min_{a\in\mathcal{A}} \{r(a, \hat{\mu}_{\vec{i}}) + \hat{h}_{\vec{j}} \mid \vec{j} = T^{\Sigma}(\vec{i}, a)\}$ for all $\vec{i} \in \Sigma$ ▷ Policy Improvement
5: **while** $\hat{d}' \neq \hat{d}$
6: **return** $\hat{g}, \hat{d}$

Instance	(node, arcs)	RVI	PI [Coc+98]	This work
$K = 1, N = 1$ $\delta_\mu = 1/2000$	$(2e3, 2.5e6)$	70s 0.8Mo	1s 30Mo	0.2s 9Mo
$K = 2, N = 2$ $\delta_\mu = 1/50$	$(7.4e5, 6.9e8)$	7h 15Mo	390s 13Go	70s 103Mo

Table 5.2: Comparison RVI / Howard

We provides running times that include the graph building step, which is a very costly operation for high dimensional graph in the standard Policy-Iteration algorithm. Every method ran on 8 threads.

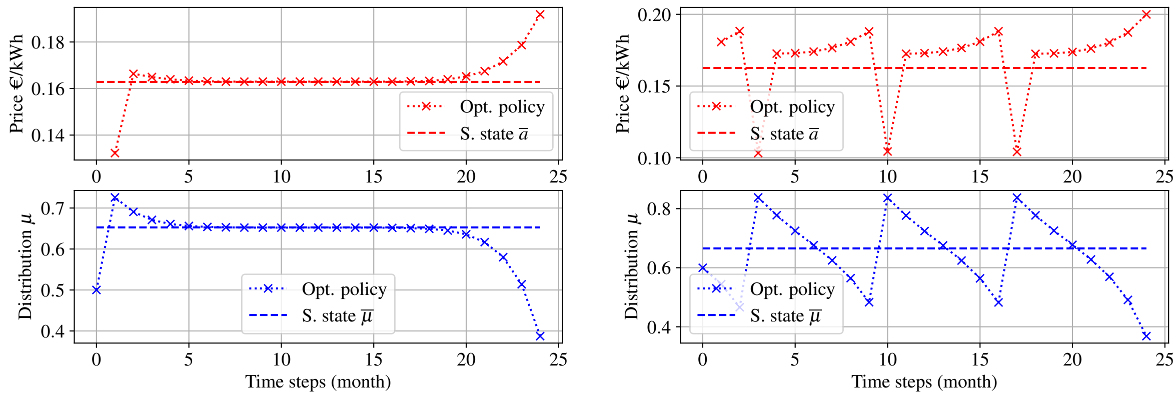
In order to visualize qualitative results, we focus on the minimal non-trivial example (1 offer and 1 cluster). Note that the conclusions we draw from this example remain valid for the case 2 offers / 2 clusters. We use data of realistic orders of magnitude: we consider a population that checks monthly the market offers and consumes $E = 500$kWh each month. The provider competes with a regulated offer of 0.17€/kWh (inducing a reservation price of 85€), and has a cost of 0.13€/kWh. We suppose that the prices are freely chosen by the provider in the range 0.08-0.22€/kWh. The intensity parameter β is fixed to 0.1. In a quadratic setting (see previous chapter), this means that a customer will choose with probability one a contract as long as its utility is higher of $2/\beta = 20$€in comparison to other offers.

Numerical experiments in Fig. 5.6-5.5 emphasize the role of the switching cost. There exists a threshold – around $\gamma = 22$ in Fig. 5.6 – above which the steady-state policy becomes dominated by a cyclic strategy, where a period of promotion is periodically applied to recover a sufficient market share (period of 7 time steps on this example, see Fig. 5.5b and Fig. 5.5d). Below this threshold, the optimal policy has an attractor point which is exactly the best steady-state price, see Fig. 5.5c. The finite horizon policy is therefore a "turnpike" like strategy [Dam+14]: we rapidly converge to the steady-state and diverge at the end of the horizon, see Fig. 5.5a. Fig. 5.6 highlights that the adding of a convex function φ strengthens the upper bound, so that the optimality of the steady-state strategy is guaranteed up to γ around 19.

5.7 Study of the minimal non-trivial model

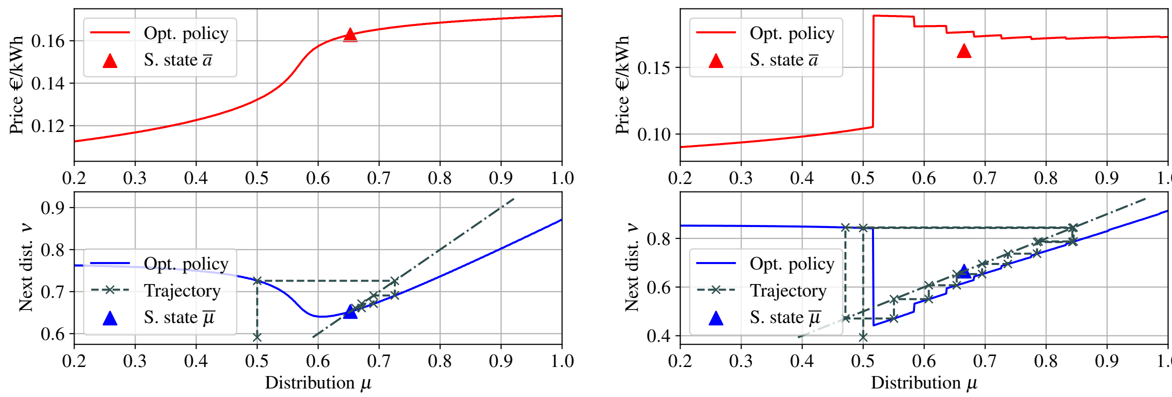
Let us study the simple (yet non-trivial) case where the company has 1 contract ($N = 1$) and the population is homogeneous ($K = 1$). Numerical results have been shown in previous section.

In this setting, the probability μ to choose the retailer contract lies in the segment $[0, 1]$. For



(a) Optimal finite horizon trajectory (provider action and customer distribution) for *low* switching cost.

(b) Optimal finite horizon trajectory (provider action and customer distribution) for *high* switching cost.



(c) Optimal decision for the long-run average reward (provider action and next customer distribution) for *low* switching cost. Graphical iteration is drawn in dotted lines.

(d) Optimal decision for the long-run average reward (provider action and next customer distribution) for *high* switching cost. Graphical iteration is drawn in dotted lines.

Figure 5.5: Numerical results for both the finite horizon and long-term average reward criteria. *Low* (resp. *high*) switching cost stands for $\gamma = 20$ (resp. 25).

the finite-horizon setting, the toy model is therefore defined as

$$v_T^*(\mu_0) = \max_{a_1,\dots,a_T \in \mathcal{A}^T} \left\{ \begin{array}{l} \displaystyle\sum_{t=1}^{T} (a_t - C)\mu_t \\ \text{s.t. } \begin{bmatrix} \mu_t & 1-\mu_t \end{bmatrix} = \begin{bmatrix} \mu_{t-1} & 1-\mu_{t-1} \end{bmatrix} \begin{bmatrix} \frac{e^{\beta\gamma}e^{-\beta(a_t-R)}}{1+e^{\beta\gamma}e^{-\beta(a_t-R)}} & \frac{1}{1+e^{\beta\gamma}e^{-\beta(a_t-R)}} \\ \frac{e^{-\beta(a_t-R)}}{e^{\beta\gamma}+e^{-\beta(a_t-R)}} & \frac{e^{\beta\gamma}}{e^{\beta\gamma}+e^{-\beta(a_t-R)}} \end{bmatrix} \end{array} \right\}. \tag{5.25}$$

In the sequel, the data is $C = 2$, $R = 3$, $\beta = 3$, $T = 45$.

5.7.1 Cycling strategies

Figure 5.5 suggests a threshold (in terms of switching cost intensity) that separates the decision behavior into two different regimes: the convergence to a steady state for low switching costs intensity and the convergence to periodic strategies above the threshold (see Figure 5.5b). Therefore, in order to better understand this cycling behavior, we define the set of periodic strategies in the one-dimensional case as follows:

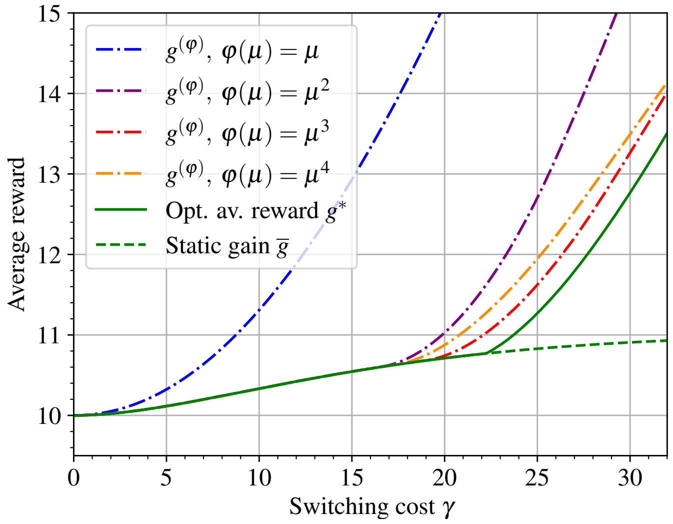


Figure 5.6: Optimal gain g^* for a range of switching costs, along with lower bound $\overline{g}$ and upper bounds $g^{(\varphi)}$, $\varphi(\cdot) = (\cdot)^{1,2,3,4}$.

Definition 5.7.1. A τ-*cycle* is a cycling strategy of τ time steps, defined by the customer response $(\mu_0, \ldots, \mu_{\tau-1}, \mu_\tau)$, with the cycling condition $\mu_0 = \mu_\tau$. For any τ-cycle l, we denote by

(i) $\overline{\mu}[l] = \frac{1}{\tau}\sum_{t=1}^{\tau} \mu_t$ and $V[l] = \frac{1}{\tau}\sum_{t=1}^{\tau}(\mu_t - \overline{\mu})^2$ the mean and the variance of the customer distribution over the cycle,

(ii) $g[l] = \frac{1}{\tau}\sum_{t=1}^{\tau}(a_t - C)\mu_t$ the gain (mean profit over the cycle).

Proposition 5.7.1

Let $\gamma > 0$, knowing μ_{t-1} and μ_t in $[0, 1]$, there exists a unique a_t verifying the constraint in (5.25), defined as

$$\hat{a}_t := e^{-\beta(a_t - R)} = \frac{2\mu_t - \kappa_t + \sqrt{(2\mu_t - \kappa_t)^2 + 4\hat{\gamma}^2\mu_t(1-\mu_t)}}{2\hat{\gamma}(1-\mu_t)} \tag{5.26}$$

where $\hat{\gamma} = e^{\beta\gamma}$ and $\kappa_t = 1 + (\hat{\gamma}^2 - 1)(\mu_{t-1} - \mu_t)$.

Proof. From (5.25), one obtains the following equation:

$$\mu_t = \left[\frac{\hat{\gamma}\hat{a}_t}{1 + \hat{\gamma}\hat{a}_t} - \frac{\hat{a}_t}{\hat{\gamma} + \hat{a}_t}\right]\mu_{t-1} + \frac{\hat{a}_t}{\hat{\gamma} + \hat{a}_t} \ ,$$

that can be equivalently written as a second-order equation: $0 = \hat{a}_t^2\,[\hat{\gamma}(\mu_t - 1)] + \hat{a}_t\,[2\mu_t - \kappa_t] + [\hat{\gamma}\mu_t]$ of discriminant $\Delta = (2\mu_t - \kappa_t)^2 + 4\hat{\gamma}^2\mu_t(1-\mu_t) \geqslant 0$. ☐

Corollary 5.7.1

As a special case of Proposition 5.7.1,

(i) if $\gamma = 0$, $\hat{a}_t = \frac{\mu_t}{1-\mu_t}$,

(ii) the steady-state policy that converges to $\mu \in]0,1[$ is obtained by fixing the price to

$$\hat{a} = \frac{2\mu - 1 + \sqrt{(2\mu-1)^2 + 4\hat{\gamma}^2\mu(1-\mu)}}{2\hat{\gamma}(1-\mu)} \ . \tag{5.27}$$

Proof. Items (i) and (ii) are obtained with $\kappa_t = 1$, either with $\hat{\gamma} = 1$ or $\mu_{t-1} = \mu_t$. □

Proposition 5.7.1 gives an explicit expression of the (unique) action that allows a transition between state μ_{t-1} and μ_t. The uniqueness can be extended to transitions $\mu_t = \mu_{t-1}P(a)$ in higher dimension, but the explicit characterization of the action is not straightforward. We now want to compare the gain over a τ-cycle l and the steady-state gain $\overline{g}$. A first result is readily obtained in absence of switching costs, i.e., $\gamma = 0$, showing that constant-price policies are in this case optimal:

Proposition 5.7.2 (Gain without switching cost)
Suppose that $\gamma = 0$, then, the optimal steady-state policy induces a gain greater than the one achieved by any τ-cycle l of at least $\frac{V[l]}{\beta}$, i.e.,

$$g[l] \leqslant \overline{g} - \frac{V[l]}{\beta} \ .$$

As a consequence, the optimal cycle corresponds to a constant-price policy.

Proof. Using Corollary 5.7.1, $a_t = R - \frac{1}{\beta}\log\left(\frac{\mu_t}{1-\mu_t}\right)$, and the mean profit of a τ-cycle l is

$$g[l] = (R-C)\overline{\mu}[l] - \frac{1}{\beta\tau}\sum_{t=1}^{\tau}\mu_t \log\left(\frac{\mu_t}{1-\mu_t}\right) \ .$$

The function $\mu \mapsto \mu\log\left(\frac{\mu}{1-\mu}\right)$ is strongly convex of modulus 1. Therefore, using Jensen's inequality for strongly convex function, see e.g. [MN10], we obtain that

$$g[l] \leqslant (R-C)\overline{\mu}[l] - \frac{1}{\beta}\overline{\mu}[l]\log\left(\frac{\overline{\mu}[l]}{1-\overline{\mu}[l]}\right) - \frac{V[l]}{\beta} \leqslant \overline{g} - \frac{V[l]}{\beta} \ .$$

□

Let us specialize the τ-cycles to a particular sub-class:

Definition 5.7.2. A (s,S,τ)-*cycle* is a specific τ-cycle, in which $\mu_t = S + \frac{s-S}{\tau}t, t \leqslant \tau$.

Proposition 5.7.3
Let us consider a (s,S,τ)-cycle l. Then,

$$g[l] \geqslant \frac{s(\tau-1)-S}{\tau}\gamma + O(1) \text{ as } \gamma \to \infty \ .$$

As a consequence, there exists a threshold $\Gamma > 0$ such that for any $\gamma \geqslant \Gamma$, the optimal steady-state policy is dominated by a (s,S,τ)-cycle.

Proof. Recalling that $\sqrt{a^2+b} \leqslant |a| + \frac{b}{2|a|}$, we have $\sqrt{(2\mu-\kappa)^2 + 4\hat{\gamma}^2\mu(1-\mu)} \leqslant |2\mu-\kappa| + \dfrac{\hat{\gamma}^2\mu(1-\mu)}{|2\mu-\kappa|}$. We first look at a period $1 \leqslant t < \tau$ where $\mu_{t-1} - \mu_t = \frac{S-s}{\tau}$. As we suppose that $\gamma \to \infty$, $\hat{\gamma} \geqslant \sqrt{1+\frac{\tau}{S-s}}$ and so $\kappa \geqslant 2\mu$. Therefore,

$$\hat{a} \leqslant \frac{\hat{\gamma}\mu}{1 + (\hat{\gamma}^2-1)\frac{S-s}{\tau} - 2\mu}$$

and

$$a \geqslant R + \frac{1}{\beta}\log\left(\frac{(\hat{\gamma}^2-1)(S-s)-\tau}{\hat{\gamma}\tau}\right) \simeq \frac{1}{\beta}\log(\hat{\gamma}) + O(1) = \gamma + O(1) \ .$$

If now we look at the last period $t = \tau$. Then, as we suppose that $\gamma \to \infty$, $\hat{\gamma} \geqslant \sqrt{1+\frac{1}{S-s}}$. Therefore,

$$\hat{a} \leqslant \frac{1+(\hat{\gamma}^2-1)(S-s)}{\hat{\gamma}(1-y)} + \frac{\hat{\gamma}}{(\hat{\gamma}^2-1)(S-s)-1}$$

and

$$a \geqslant R - \frac{1}{\beta}\log\left(\frac{1+(\hat{\gamma}^2-1)(S-s)}{\hat{\gamma}(1-y)} + \frac{\hat{\gamma}}{(\hat{\gamma}^2-1)(S-s)-1}\right) \simeq -\gamma + O(1) \ .$$

The mean profit is finally bounded by below : $g[l] \geqslant \frac{1}{\tau}\left[(\tau-1)s - S\right]\gamma + O(1)$. To conclude, any (s,S,τ)-cycle satisfying $\tau \geqslant 1 + \frac{S}{s}$ induces a mean profit that diverges with respect to γ. In the meantime, the steady-state optimum is bounded, see Proposition 5.4.2, and so dominated for sufficiently large switching cost γ.

□

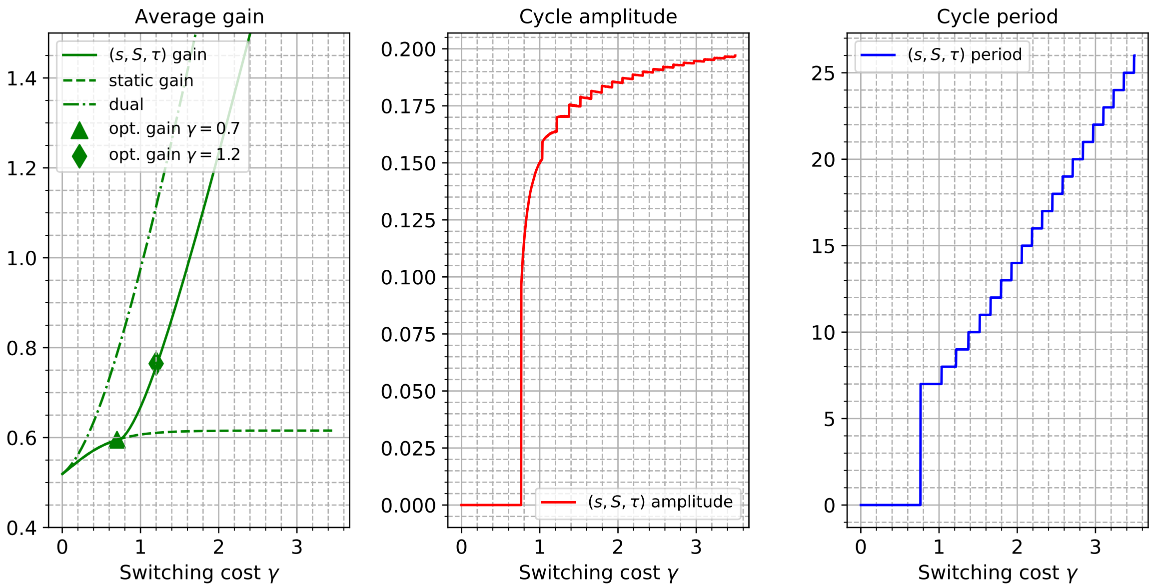


Figure 5.7: Evolution of the optimal (s,S,τ)-cycle for a range of switching cost values

The left (resp. middle, right) panel shows the gain (resp. cycle amplitude, cycle period) of the optimal (s,S,τ)-cycle. The steady-state gain $\overline{g}$ is displayed for comparison, as well as the optimal gain obtained in Figure 5.5. A kink appears at $\gamma \simeq 0.762$, indicating the separation of the cycling behavior from the steady-state behavior.

In Figure 5.7, we compute the optimal (s, S, τ)-cycle, by iterating over the possible values of s, S, and τ for each given value of γ. Before the kink, the optimal cycle is in reality the constant-price strategy (cycle of amplitude 0), and after this point, there exists cycle of positive amplitude that outperforms the steady-state strategy. The results found in Figure 5.5 for a broader class of cycles are consistent, and the (s, S, τ)-cycles are good approximations of the optimal policy.

5.8 Extensions and Perspectives

5.8.1 Existence of a solution to the ergodic eigenproblem with common-noise

Here, we aim at extending the existence result proved in Theorem 5.2.2 to a stochastic framework, where both the transition matrices and the instantaneous rewards are impacted by an hazard, interpreted here as a common-noise in the dynamics, see [Bäu23]. Let Ω be a compact Lebesgue measurable subset of $\mathbb{R}^S$, $S \in \mathbb{N}$, anf $\mathbb{F}$ its natural filtration. For any action $a \in \mathcal{A}$ and state $\mu \in \Delta_N^K$, we denote by $\mathbb{P}_{a,\mu}$ the state and action-dependent probability measure on $(\Omega, \mathbb{F})$ along with its corresponding density function $f_{a,\mu}$. For any realization $\omega \in \Omega$, we now denote by $r(a, \mu, \omega)$ (resp. $P(a, \omega)$) the instantaneous reward (resp. the transition matrix) for $\mu \in \Delta_N^K$ and $a \in \mathcal{A}$.

For any real-valued function $v : \Delta_N^K \to \mathbb{R}$ and discount factor $\alpha \in]0, 1]$, the Bellman operator $\mathcal{B}_\alpha$ is now defined as

$$\mathcal{B}_\alpha\, v(\mu) = \max_{a\in\mathcal{A}} \int_\Omega \left[r(a, \mu, \omega) + \alpha v(\mu P(a, \omega))\right] f_{a,\mu}(\omega) d\omega \ , \tag{5.28}$$

where $\int_\Omega f_{a,\mu}(\omega)d\omega = 1$. For $\alpha = 1$, we simply write $\mathcal{B} = \mathcal{B}_1$. For $\alpha < 1$, we denote by v_α the solution of $\mathcal{B}_\alpha\, v = v$, which can be obtained as the limit of the sequence $(v^j_\alpha)_{j\in\mathbb{N}}$ where $v^{j+1}_\alpha = \mathcal{B}_\alpha\, v^j_\alpha$ and $v^j_\alpha \equiv 0$. This result derives from the contraction property of the Bellman operator $\mathcal{B}_\alpha$ for $\alpha < 1$.

As in the deterministic case, we make the following assumptions:

(B1) The transition $(a, \omega) \mapsto P^k(a, \omega)$ is a continuous function for any k.

(B2) For any action $a \in \mathcal{A}$, cluster k and $\omega \in \Omega$, $P^k(a, \omega) \gg 0$.

(B3) There exists M_r such that, $|\theta^{kn}(a, \omega)| \leqslant M_r$ for every $k \in [K]$, $n \in [N]$, $a \in \mathcal{A}$ and $\omega \in \Omega$.

Let $\mathcal{D}$ be defined as $\mathcal{D} = \bigtimes_{k\in[K]} \mathcal{D}^k$ where

$$\mathcal{D}^k = \text{vex}\left(\{\mu^k P^k(a, \omega) \mid a \in \mathcal{A}, \mu^k \in \Delta_N, \omega \in \Omega\}\right) \ .$$

Lemma 5.8.1

Let (B1)-(B2) hold. Then $\mathcal{D}$ is a compact set included in the relative interior of Δ_N^K. Moreover, for $t \geqslant 1$, $\mu_t \in \mathcal{D}$ for any policy $\pi \in \Pi$.

Proof. The set $\{\mu^k P^k(a, \omega) \mid (a, \mu^k, \omega) \in \mathcal{A} \times \Delta_N \times \Omega\}$ is compact, since $(a, \mu^k, \omega) \mapsto \mu^k P^k(a, \omega)$ is continuous and Δ_N, $\mathcal{A}$ and Ω are compact. Therefore, $\mathcal{D}$ is compact as it is the convex hull of a compact set in finite dimension. Then, the positiveness of P^k implies that $\mathcal{D}^k \subset \text{relint}(\Delta_N)$. Moreover, by property of the semiflow, $\phi_t(a_{:t}, \mu_0) \in \mathcal{D}$. ☐

Under (B1) and (B2), $(\mathcal{D}, d_H)$ is then a complete metric space

$$\forall w, a \in \Omega \times \mathcal{A},\ \forall \mu, \nu \in \mathcal{D},\ d_H(\mu P(a,\omega), \nu P(a,\omega)) \leqslant \kappa(P(a,\omega)) d_H(\mu,\nu) \ ,$$

where $\kappa(P) := \tanh\left(\mathrm{Diam}_H(P)\,/\,4\right) < 1$. We are now able to derive the stochastic analog of Lemma 5.2.3, where the equi-Lipschitz of the value function is proved.

Lemma 5.8.2 (Equi-Lipschitz property for state-independent noise)
Assume that (B1)-(B3) hold, and that $f_{a,\mu}(\omega) = f_a(\omega)$ for all $\mu \in \mathcal{D}$. Then, v_α is equi-Lipschitz, i.e.,

$$\forall \alpha < 1,\ |v_\alpha(\mu) - v_\alpha(\nu)| \leqslant M d_H(\mu,\nu) \ ,$$

with $M := \frac{M_r^{\mathcal{D}}}{1-\kappa}$, $\kappa = \sup_{a\in\mathcal{A},\omega\in\Omega} \kappa(P(a,\omega))$ and $M_r^{\mathcal{D}}$ defined in Lemma 5.2.2.

Proof. We denote by $(v_\alpha^j)_{j\in\mathbb{N}}$ the sequence defined as $v_\alpha^{j+1} = \mathcal{B}_\alpha\, v_\alpha^j$ and $v_\alpha^0 \equiv 0$. Let us assume that for a given $j \in \mathbb{N}$, v_α^j is M_α^j-Lipschitz w.r.t the Hilbert metric , i.e.,

$$\left|v_\alpha^j(\mu) - v_\alpha^j(\nu)\right| \leqslant M_\alpha^j d_H(\mu,\nu) \ .$$

Then, for $\mu, \nu \in \mathcal{D} \subset \Delta_K^N$, we have:

$$\begin{aligned}
\left|\mathcal{B}_\alpha\, v_\alpha^j(\mu) - \mathcal{B}_\alpha\, v_\alpha^j(\nu)\right| &\leqslant \int_\Omega |f_{a,\mu}(\omega) r(a,\mu,\omega) - f_{a,\nu}(\omega) r(a,\nu,\omega)|\, d\omega \\
&\qquad + \alpha \int_\Omega \left|f_{a,\mu}(\omega) v_\alpha^j(\mu P(a,\omega)) - f_{a,\nu}(\omega) v_\alpha^j(\nu P(a,\omega))\right| d\omega \\
&\leqslant \int_\Omega |f_{a,\mu}(\omega) - f_{a,\nu}(\omega)|\, r(a,\mu,\omega) d\omega \\
&\qquad + \int_\Omega |r(a,\mu,\omega) - r(a,\nu,\omega)|\, f_{a,\nu}(\omega) d\omega \\
&\qquad + \alpha \int_\Omega |f_{a,\mu}(\omega) - f_{a,\nu}(\omega)|\, v_\alpha^j(\mu P(a,\omega)) d\omega \\
&\qquad + \alpha \int_\Omega \left|v_\alpha^j(\mu P(a,\omega)) - v_\alpha^j(\nu P(a,\omega))\right| f_{a,\nu}(\omega) d\omega \\
&\leqslant \left[\left(\mathrm{Lip}_H\, \bar{f}_a\right)\left(\|r\|_\infty + \alpha \|v_\alpha^j\|_\infty\right) + \alpha\kappa M_\alpha^j + M_r^{\mathcal{D}}\right] d_H(\mu,\nu) \ ,
\end{aligned}$$

with $\bar{f}_a : \mu \in \mathcal{D} \mapsto \int_\Omega f_{a,\mu}(\omega) d\omega$. Therefore, v_α^{j+1} is M_α^{j+1}-Lipschitz with

$$M_\alpha^{j+1} = \left(\mathrm{Lip}_H\, \bar{f}_a\right)\left(\|r\|_\infty + \alpha\|v_\alpha^j\|_\infty\right) + \alpha\kappa M_\alpha^j + M_r^{\mathcal{D}} \ .$$

By assumption, $\bar{f}_a$ is a constant function, so $\mathrm{Lip}_H\, \bar{f}_a = 0$, and $M_\alpha^{j+1} = \alpha\kappa M_\alpha^j + M_r^{\mathcal{D}} \leqslant \kappa M_\alpha^j + M_r^{\mathcal{D}}$. Therefore, for all $j \in \mathbb{N}$, $M_\alpha^j \leqslant M := \frac{M_r^{\mathcal{D}}}{1-\kappa}$, which is independent of j. So, at the limit, v_α is M-equi-Lipschitz w.r.t. the Hilbert pseudo-metric. □

Remark 5.8.1

In the proof, if the hazard depends on the state, i.e., $\bar{f}_\alpha$ is not constant, then there is no guarantee that M_α^{j+1} is bounded. In fact, there is no reason that $\|v_\alpha^j\|_\infty$ be uniformly bounded (when $\alpha \to 1$, it goes to infinity in general).

As in the deterministic setting, the existence of a solution to the ergodic eigenproblem can be derived from the equi-Lipschitz property of the value function, by applying the vanishing discount approach. Following the same steps as in the proof of Theorem 5.2.2, we obtain:

Theorem 5.8.1 (Existence of a solution for state-independent noise)
Assume that (B1)-(B3) hold. Then, the ergodic eigenproblem

$$g\,1_{\mathcal{D}} + h = \mathcal{B}\,h \tag{5.29}$$

admits a solution $h^* \in \mathrm{Lip}(\mathcal{D}) \cap \mathrm{Vex}(\mathcal{D})$ and $g^* \in \mathbb{R}$.

In Theorem 5.8.1, by restricting ourselves to state-independent noises, we can still apply the vanishing discount approach, as in the deterministic case Theorem 5.2.2. In [Bäu23], Bäuerle also used this approach in the stochastic case, but here we do not assume a priori the equiboundedness of the optimal discounted objective functions V^*_α. Instead, using a contraction argument on the dynamics, we obtained that $(V^*_\alpha)_{\alpha\in(0,1)}$ is equi-Lipschitz (see Lemma 5.8.2). In particular, it entails that any optimal eigenvector h^* is Lipschitz (and not only upper semicontinuous).

5.8.2 Generalization of turnpike properties via weak-KAM theory

In the numerical results (see Section 6.6 and Section 5.7), we observed that the dynamics does not always converge to a steady-state, but converges for high switching costs to periodic strategies. In this section, we aim at generalizing the turnpike properties (convergence to a single state) to cycling policies. To this purpose, we exhibit connections between turnpike results in control theory and weak-KAM theory [Fat97; FS04; FS05; Fat08; Dav+16]. Note that the latter link was mentioned in [Tré22] where Trélat noticed that the existence of a storage function, often assumed in control theory to force the dynamics to be *strictly dissipative* at the optimal static point (and so to obtain turnpike results, e.g. in [Grü+21]), is related to the weak KAM theory.

Control problem. We first briefly generalize the results of previous sections to a slightly broader setting: the set of states $\mathcal{X}$ is a separable and *compact* space for the topology of a metric d; and the set of actions/controls $\mathcal{A}$ is a *compact* topological space. Let $\mathcal{C}(\mathcal{X},\mathbb{R})$ be the set of continuous function from $\mathcal{X}$ to $\mathbb{R}$. We focus on the following Bellman operator:

$$\forall h \in \mathcal{C}(\mathcal{X},\mathbb{R}), \quad \mathcal{B}\,h : x \in \mathcal{X} \mapsto \sup_{a\in\mathcal{A}} \left\{ r(x,u) + h\left(\Gamma(x,a)\right)\right\}, \tag{5.30}$$

where the following assumptions hold:

Assumption 5.8.1. (i) The dynamics is described by a family of self-maps $(\Gamma(\cdot,a))_{a\in\mathcal{A}}$ of $\mathcal{X}$ that are uniformly contracting, meaning that there exists a constant $\kappa < 1$ such that $\mathrm{d}\left(\Gamma(x,a),\Gamma(y,a)\right) \leqslant \kappa\, \mathrm{d}(x,y)$ for all $x,y \in \mathcal{X}$.

(ii) The *stage reward* function $r : \mathcal{X}\times\mathcal{A}$ is jointly continuous and the maps $r(\cdot,a)$ are M-Lipschitz from $(\mathcal{X},\mathrm{d})$ to $\mathbb{R}$ equipped with the usual distance.

Weak KAM analog of the control problem. Aubry-Mather theory [MF94] and weak KAM (Kolmogorov-Arnold-Moser) theory [Fat97; FS04; FS05; Fat08; Dav+16] aim at studying orbits minimizing the action, in relation with Hamiltonian dynamical systems. A discrete version of the weak KAM theory was then introduced by Zavidovique [Zav23], from which we extracted the following definitions.

Let $c : \mathcal{X}^2 \to \overline{\mathbb{R}}$ be the stage cost function indexed by pairs of states:

$$c(x,y) := \inf_{a\in\mathcal{A}} \left\{ -r(x,a) \mid \Gamma(x,a) = y \right\} \quad . \tag{5.31}$$

The stage cost $c(x, y)$ is then the opposite of the maximum reward over the actions that can be obtained when the dynamics go from state x to state y. Let T_c^+ be the positive Lax-Oleinick semi-group, defined as

$$T_c^+ \, h(x) := \sup_{y \in \mathcal{X}} \{h(y) - c(x, y)\} \ . \tag{5.32}$$

With the definition of the stage cost function (in (5.31)), we easily get that $\mathcal{B} = T_c^+$. Therefore, we can re-interpret the existence of a solution for the ergodic eigen problem, stated in Theorem 5.2.2, as the existence of a positive weak KAM solution:

Theorem 5.8.2 (Existence of positive weak KAM solution, case of a contracting dynamics) Assume that Assumption 5.8.1 hold. Then, the problem

$$T_c^+ h = h + g \tag{5.33}$$

admits a solution $h^* \in \mathrm{Lip}(\mathcal{X}) \cap \mathrm{Vex}(\mathcal{X})$ and $g^* \in \mathbb{R}$. Moreover, any sequence $(x_n)_{n \in \mathbb{N}}$ satisfying $x_{n+1} \in \arg\max T_c^+ h^*(x_n)$ for $n \in \mathbb{N}$ minimizes the average stage cost:

$$g^* = \inf_{(x_n)_{n \in \mathbb{N}}} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} c(x_n, x_{n+1}) \ . \tag{5.34}$$

Theorem 5.8.2 is no more than a re-writing of Theorem 5.2.2 in the weak-KAM framework. This result relies on the contraction property satisfied by the dynamics of the underlying MDP. In weak KAM theory, another convenient assumption which guarantees the solvability of the ergodic eigenproblem is the following:

Assumption 5.8.2 (Controllability)**.** The cost function $c(\cdot, \cdot)$ is uniformly bounded and jointly continuous on $\mathcal{X}$.

This controllability assumption was in particular considered by Kolokoltsov and Maslov [KM97] in order to show the existence of a solution. It supposes for the underlying dynamics that for each pair of states $(x, y) \in \mathcal{X}^2$ there exist a control $a \in \mathcal{A}$ so that y is attainable from x via action a. Under this assumption, Zavidovique [Zav12] proved that a positive weak KAM solution to (5.35) exists:

Theorem 5.8.3 (Existence of positive weak KAM solution, case of controllable system) Assume that 5.8.2 hold. Then, the problem

$$T_c^+ h = h + g \tag{5.35}$$

admits a solution $h^* \in \mathrm{Vex}(\mathcal{X})$ and $g^* \in \mathbb{R}$. Moreover, any sequence $(x_n)_{n \in \mathbb{N}}$ satisfying $x_{n+1} \in \arg\max T_c^+ h^*(x_n)$ for $n \in \mathbb{N}$ minimizes the average stage cost:

$$\lambda^* = \inf_{(x_n)_{n \in \mathbb{N}}} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} c(x_n, x_{n+1}) \ . \tag{5.36}$$

Note that extensions of this existence result have been proved using quasi-compactness techniques, see e.g. [MN02]. Also, in [AL10], the authors relax the controllability assumption by finding conditions to recover the continuity of the optimal control cost function $c(x, y)$ on finite-cost region. Here, the controllability assumption is not satisfied: due to the inertia in the response, the customers cannot switch from an offer to another too rapidly. This means that only a restricted region around the previous state can be explored by the set of feasible actions. Nevertheless, the connection between turnpike properties and weak-KAM solutions appears to be simpler under a controllability assumption, and so, as a first approach, we assume that Assumption 5.8.2 holds in the rest of this chapter.

We now define the Aubry set through the use of critical subsolutions:

Definition 5.8.1 (Critical subsolutions, [Zav12]). Given $C \in \mathbb{R}$, a function $h : \mathcal{X} \to \mathbb{R}$ is called a *C-subsolution* if

$$\forall (x,y) \in \mathcal{X}^2,\ h(y) - h(x) \leqslant c(x,y) + C \ .$$

The set of the C-subsolutions is denoted by $\mathcal{S}_C$, and the set of critical subsolutions $\mathcal{S}_{g^*}$ is denoted by $\mathcal{S} = \mathcal{S}_{g^*}$.

Note that a weak KAM solution u belongs to $\mathcal{S}$: $\forall (x,y) \in \mathcal{X}^2, h(y) - c(x,y) \leqslant T_c^+ h(x) = h(x) + g^*$.

Definition 5.8.2 (Aubry set, [Zav12]). Let $h \in \mathcal{S}$ be a critical subsolution. The *Aubry set of* h, $\widetilde{\mathbb{A}}_h \in \mathcal{X}^{\mathbb{N}}$, is defined as

$$\widetilde{\mathbb{A}}_h = \left\{ (x_n)_{n\in\mathbb{N}} \mid \forall n < p, h(x_p) - h(x_n) = \sum_{k=n}^{p-1} c(x_k, x_{k+1}) + (p-n) g^* \right\} \ .$$

The Aubry set $\widetilde{\mathbb{A}}$ is then the intersection over all the critical subsolutions, i.e., $\widetilde{\mathbb{A}} = \cap_{h\in\mathcal{S}} \widetilde{\mathbb{A}}_h$. Finally, the projected Aubry set $\mathbb{A}$ refers to the projection of the Aubry set on the first component, and is given by

$$\mathbb{A} = \left\{ x_0 \mid (x_n)_{n\in\mathbb{Z}} \in \widetilde{\mathbb{A}} \right\} \subseteq (\mathcal{X})^{\mathbb{N}} \ .$$

The projected Aubry set can be understood as the collection of states where an optimal strategy can go through infinitely-many times. In particular, from the definition of the Aubry set, a τ-cycle $(x_n)_{n\in\mathbb{N}}$ (as introduced in Definition 5.7.1), where $x_{i+\tau} = x_i$ for all $i \in \mathbb{N}$, belongs to the Aubry set if $\sum_{i=1}^{\tau} c(x_k, x_{k+1}) = -\tau g^*$, which means that the sequence produces an optimal average long-term reward. Therefore, Aubry sets are able to capture the "optimal support" of the dynamics.

Proposition-Definition 5.8.1 (Mañé potential). For every $(x,y) \in \mathcal{X}^2$, we define $c_n(x,y)$ the shortest-path of length n from x to y as

$$c_n(x,y) := \inf_{(x_1,\dots,x_{n-1})\in\mathcal{X}^{n-1}} \{c(x,x_1) + c(x_1,x_2) + \ldots + c(x_{n-1},y)\} \ .$$

We define the *Mañé potential* as $c^*(x,y) := \sup_{h\in\mathcal{S}} h(y) - h(x)$. As $\mathcal{X}$ is compact, it can be equivalently defined as

$$c^*(x,y) = \inf_{k\geqslant 1} c_k(x,y) + kg^* \ . \tag{5.37}$$

The potential function verifies the following properties:

(i) (Triangular inequality) For $x,y,z \in \mathcal{X}$, $c^*(x,y) + c^*(y,z) \geqslant c^*(x,z)$.

5. Switching costs

(ii) For all $y \in \mathcal{X}$, the function $h^y = -c^*(\cdot, y)$ is a critical sub-solution.

(iii) A point $y \in \mathcal{X}$ belongs to the projected Aubry set $\mathbb{A}$ if and only if the function h^y is a positive weak KAM solution.

Proof. (i) $\sup_{h\in\mathcal{S}}\{h(y)-h(x)\}+\sup_{h'\in\mathcal{S}}\{h'(z)-h'(y)\} \geqslant \sup_{h\in\mathcal{S}}\{h(z)-h(y)+h(y)-h(x)\} = c^*(x,z)$.

(ii) By triangular inequality, $c^*(x,z) - c^*(y,z) \leqslant c^*(x,y)$. Besides, $c^*(x,y) \leqslant c(x,y) + g^*$ by definition of the potential. Therefore, $h^z(y) - h^z(x) \leqslant c(x,y) + g^*$.

□

Mañé potentials encode the optimal "way" to go from x to y: $c^*(x,y)$ is computed using the least-cost path $c_k(x,y)$ such that $\frac{c_k(x,y)}{k}$ is minimized. A positive (resp. negative) potential means that the best average gain between x and y is lower (resp. greater) than the optimal one.

The link between the max-plus spectral problem and the weak KAM theory was initially studied in [AGW09], where the construction of a minimal Martin space coincides with the projected Aubry set when the state space is compact.

Strict subsolutions and strict dissipativity condition. Zavidovique showed that a subsolution can always be constructed so that the intersection of all the Aubry set $\mathbb{A}_h$, $h \in \mathcal{S}$, can be reduced to the Aubry set of this peculiar subsolution:

Theorem 5.8.4 (Strict subsolution, [Zav23], Theorem 1.4.1)

There exists a subsolution $h_0 \in \mathcal{S} \cap \mathcal{C}(\mathcal{X}, \mathbb{R})$ such that if the equality $h_0(y) - h_0(x) = c(x,y) + g^*$ holds for some $(x,y) \in \mathcal{X}^2$, then

$$\forall h \in \mathcal{S}, \quad h(y) - h(x) = c(x,y) + g^* \ .$$

In the proof of Theorem 5.8.4, Zavidovique defines h_0 as the mean of a dense sequence of subsolutions, which is a non constructive definition. Here, we make an explicit construction of the strict subsolution by defining h_0 as a mean only on a subset of subsolutions. Let us first recall a preliminary result:

Lemma 5.8.3

There exists a strictly positive measure on $\mathcal{X}$.

Proof. Let us recall that we assume that $\mathcal{X}$ is separable. Consider any countable subset $\{x_n \mid n \in \mathbb{N}\}$, dense in $\mathcal{X}$. We can define $\mu = \sum_n 2^{-n}\delta_{x_n}$. As every open set contains some x_n, the measure is strictly positive on $\mathcal{X}$. □

Theorem 5.8.5 (Explicit construction of a strict subsolution)

Let $h_0 : x \in \mathcal{X} \mapsto \int_{\mathcal{X}} h^z(x) d\mu(z)$, where μ is an arbitrary strictly positive measure, and h^z is defined in Definition 5.8.1. If the equality $h_0(y) - h_0(x) = c(x,y) + g^*$ holds for some $(x,y) \in \mathcal{X}^2$, then

$$\forall h \in \mathcal{S}, \quad h(y) - h(x) = c(x,y) + g^* \ .$$

Proof. Suppose that $h_0(y) - h_0(x) = c(x,y) + g^*$ holds for some $(x,y) \in \mathcal{X}^2$, and let $\beta(z) = c(x,y) + g^* - h^z(y) - h^z(x)$. Then, by definition, $\beta(z) \geqslant 0$ for all z, and

$$\int_{\mathcal{X}} \beta(z) d\mu(z) = 0 \ .$$

As μ is a strictly positive measure, $\beta(z) = 0$ for all $z \in \mathcal{X}$ and so $h^z(y) - h^z(x) = c(x,y) + g^*$ for all $z \in \mathcal{X}$. Using [AGW09, Theorem 6.2], we know that h^z are the exteme generators of $\mathcal{S}$ (as $\mathcal{X}$ is compact, the minimal Martin space is a subset of $\{h^z\}_{z\in\mathcal{X}}$) and thus, the strict positivity of the quantity $h^z(y) - h^z(x) = c(x,y) + g^*$ for any $z \in \mathcal{X}$ induces the strict positivity of $h(y) - h(x) = c(x,y) + g^*$ for any $h \in \mathcal{S}$. □

The 2-Aubry set is then defined as

$$\widehat{\mathbb{A}} = \left\{(x,y) \in \mathcal{X}^2 \mid h_0(y) - h_0(x) = c(x,y) + g^*\right\}$$

and is the restriction of the Aubry set to sequences of length 2. The solution h_0 is therefore *strict* on the set $\mathcal{X}^2 \setminus \widehat{\mathbb{A}}$, meaning that

$$h_0(y) - h_0(x) < c(x,y) + g^* \text{ for all } (x,y) \notin \widehat{\mathbb{A}} \ . \tag{5.38}$$

This viewpoint of the Aubry set (as the set of states where h_0 is not strict) establishes a link with optimal control problems: a condition of *strict dissipativity* is often required to obtain turnpike properties (convergence to a steady-state). In [Grü+21], denoting by $s(x,u) := -r(x,u) + g^*$ the so-called *supply rate*, the strict dissipativity condition is then expressed as follows:

Assumption 5.8.3 (Strict dissipativity condition). Let $x_e \in \mathcal{X}$ an equilibrium. Then the system is strictly x-dissipative at x_e if it exists $h \in \mathcal{X} \to \mathbb{R}$ and $\alpha \in \mathcal{K}_\infty$ such that

$$\begin{aligned} &\text{(control setting)} \;\; s(x,a) + h(x) - h(\Gamma(x,a)) \geqslant \alpha(\|x - x_e\|),\; x \in \mathcal{X}, a \in \mathcal{A} \\ &\text{(weak-KAM setting)} \;\; h(y) - h(x) + \alpha(\|x - x_e\|) \leqslant c(x,y) + g^*,\; x,y \in \mathcal{X} \end{aligned} \ .$$

Here, $\mathcal{K}_\infty$ is the set of continuous, zero at zero, unbounded and strictly increasing function. The function $\check{c}(x,u) := s(x,u) + h(x) - h(\Gamma(x,u))$ is called the *rotated stage cost* in [Grü+21], where Grüne et al. proved in particular that if Assumption 5.8.3 holds, then the dynamics converges to the steady-state x_e (turnpike property). In the weak-KAM setting, this condition implies that the projected Aubry set is reduced to the singleton $\{x_e\}$:

Proposition 5.8.1

If Assumption 5.8.3 holds, then $\widetilde{\mathbb{A}} = \{(x_n)_{n\in\mathbb{N}}\}$ where $x_n = x_e$ for all $n \in \mathbb{N}$.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of states. Then, summing the strict dissipativity conditions for each pair (x_n, x_{n+1}), we obtain:

$$h(x_p) - h(x_n) + \sum_{k=n}^{p-1} \alpha(\|x_k - x_e\|) \leqslant \sum_{k=n}^{p-1} c(x_k, x_{k+1}) + (p-n)g^* \ .$$

By definition, the function h in the strict dissipativity assumption is a subsolution, and therefore, the sequence $(x_n)_{n\in\mathbb{N}}$ belongs to $\widetilde{\mathbb{A}}_h$ if $\sum_{k=n}^{p-1} \alpha(\|x_k - x_e\|) = 0$, which implies that $x_n = x_e$ for all $n \in \mathbb{N}$. Finally, $\widetilde{\mathbb{A}} = \widetilde{\mathbb{A}}_h$ since the constant sequence $(x_e, x_e, \ldots)$ belongs to any $\widetilde{\mathbb{A}}_{h'}$ for h' subsolution. □

Convergence to the Aubry set. As we previously seen in numerical results, the strict dissipativity condition is a strong assumption, supposing the dynamics converges to a single state x_e (that should be known in advance). Here, we relax this dissipativity assumption and exploit the existence of the peculiar strict subsolution h_0 (recall that this is only possible thanks to controllability assumption):

Theorem 5.8.6 (Turnpike theorem, case of controllable system)
Let h^* be a positive weak KAM solution, and $x_0 \in \mathcal{X}$. We denote by $\pi^*(\cdot) \in \arg\max T_c^+ h^*$ an optimal stationary policy and $\{x_i^*\}$ the sequence of states generated by the policy π^*. Then, all the accumulation points of the sequence $\{x_i\}$ belong to the projected Aubry set $\mathbb{A}$.

Proof. The operator T^+ is nonexpensive therefore for any h_1 and h_2 in $\mathcal{B}(\mathcal{X}, \mathbb{R})$,

$$\left\|(T^+)^N(h_1) - (T^+)^N(h_2)\right\|_\infty \leqslant \|h_1 - h_2\|_\infty$$

In particular for $h_1 = 0$ and $h_2 = h^*$, we obtain that

$$(T^+)^N(0)(x_0) - h^*(x_0) \geqslant \|h^*\|_\infty$$

and $(T^+)^N(0)(x_0) = -\sum_{k=0}^{N-1} c(x_k^*, x_{k+1}^*)$ (optimal trajectory from x_0). Therefore,

$$\sum_{k=0}^{N-1} c(x_k^*, x_{k+1}^*) \leqslant -\|h^*\|_\infty - h^*(x_0).$$

Let $\check{c}(x,y) := c(x,y) + g + h_0(x) - h_0(y) \geqslant 0$ the rotated stage cost associated with the peculiar subsolution h_0. Then,

$$0 \leqslant \sum_{k=0}^{N-1} \check{c}(x_k^*, x_{k+1}^*) \leqslant g + h_0(x_0) - h_0(x_N^*) - \|h^*\|_\infty - h^*(x_0) \leqslant M \quad , \tag{5.39}$$

where M is a constant independent of N (the positiveness of the sum comes from the definition of a subsolution).

The 2-state space $\mathcal{X}^2$ is compact, therefore the sequence $\{(x_k^*, x_{k+1}^*)\}_{k\in\mathbb{N}}$ admits accumulation points. Let $\overline{z} \in \mathcal{X}^2$ be such a point, and $\{z_k\}_{k\in\mathbb{N}} \in (\mathcal{X}\times\mathcal{X})^{\mathbb{N}}$ be a subsequence converging to $\overline{z}$. Suppose that $z \notin \widehat{\mathbb{A}}$. Then, for all $\epsilon > 0$, there exists index k_ϵ such that $\forall k \geqslant k_\epsilon$, $\|z_k - \overline{z}\| \leqslant \epsilon$. Moreover, by continuity of the subsolution and $c(\cdot,\cdot)$,

$$\exists (\epsilon_c, \eta) \in \mathbb{R}_+^* \times \mathbb{R}_+^*,\ \forall z' \in \mathcal{X}^2, \quad \|z' - \overline{z}\| \leqslant \epsilon_c \Rightarrow \check{c}(z') \geqslant \eta \ .$$

So, for all $k \geqslant k_{\epsilon_c}$, $\check{c}(z_k) \leqslant \eta$, and $\sum_{k=k_{\epsilon_c}}^{N-1} \check{c}(z_k) \geqslant (N - k_{\epsilon_c})\eta \xrightarrow[N\to\infty]{} \infty$, which is a contradiction with (5.39). Therefore, all the accumulation points of the sequence $\{(x_k^*, x_{k+1}^*)\}_{k\in\mathbb{N}}$ belongs to the 2-Aubry set $\widehat{\mathbb{A}}$. □

Remark 5.8.2
If the Aubry set is reduced to a singleton, then we recover the notion of "turnpike": the dynamics converges to a *steady-state*. Note that the strict solution h_0 can be explicitly computed – see Theorem 5.8.5 and (5.37), and so the Aubry set too. Therefore, we are able to numerically say when the optimal strategy describes a periodic orbit or simply converges to a steady-state.

5.8.3 Continuation of Example 5.2.1

We now complete the description of the optimal strategies in Example 5.2.1 by exhibiting the 2-Aubry set:

Theorem 5.8.7 (Description of the Aubry set)
Let us consider the dynamical system described in Example 5.2.1. Then, the 2-Aubry set $\widehat{\mathbb{A}} \subseteq \left(\Delta_3^{\leqslant}\right)^2$ only contains the two steady-states defined in (5.11), i.e.,

$$(\mu, \nu) \in \widehat{\mathbb{A}} \iff (\mu, a) \in \{(\hat{\mu}^k, a_k)\}_{k\in\{0,1\}} \text{ and } \nu = \mu P(a) \enspace .$$

The projected Aubry set is then $\mathbb{A} = \{\hat{\mu}^0, \hat{\mu}^1\}$.

Proof. By Definition 5.8.2, we directly get that $(\hat{\mu}^k, a_k)$, $k \in \{0, 1\}$, belongs to the 2-Aubry set, as $c(\hat{\mu}^k, \hat{\mu}^k) = -g^*$.

Conversely, let us define the cells $\mathcal{C}^{ij} \in \Delta_3^{\leqslant}$, $i, j \in \{0, 1\}$, as

$$\mathcal{C}^{ijk} = \left\{(\mu_1, \mu_3) \in \Delta_3^{\leqslant} \mid \hat{h}^{ij}(\mu_1, \mu_3) \geqslant \hat{h}^{ik}(\mu_1, \mu_3)\right\} \enspace ,$$

where $\hat{h}^{ij}$ is defined in (5.12). Let us also define $w^{ij} = \hat{h}^{ii} \vee \hat{h}^{ij}$. Then, w^{ij}, $i, j \in \{0, 1\}$, are critical sub-solutions and

(i) $\mathcal{B}^1\, w^{00}(\mu_1, \mu_3) = w^{00}(\mu_1, \mu_3) + g^* \iff (\mu_1, \mu_3) \in \mathbb{A}^{00} := \mathcal{C}^{002} \cap \mathcal{C}^{020}$,

(ii) $\mathcal{B}^1\, w^{01}(\mu_1, \mu_3) = w^{01}(\mu_1, \mu_3) + g^* \iff (\mu_1, \mu_3) \in \mathbb{A}^{01} := \mathcal{C}^{010} \cap \mathcal{C}^{012}$,

(iii) $\mathcal{B}^1\, w^{02}(\mu_1, \mu_3) = w^{02}(\mu_1, \mu_3) + g^* \iff (\mu_1, \mu_3) \in \mathbb{A}^{02} := (\mathcal{C}^{020} \cap \mathcal{C}^{021}) \cup (\mathcal{C}^{010} \cap \mathcal{C}^{001})$.

Therefore, as the $\widehat{\mathbb{A}} = \bigcap_{h\in\mathcal{S}} \widehat{\mathbb{A}}_h$, see Definition 5.8.2, if $(\mu, \mu P(a_0)) \in \widehat{\mathbb{A}}$, then $\mu \in \mathbb{A}^{00} \cap \mathbb{A}^{01} \cap \mathbb{A}^{02} = \{\hat{\mu}^0\}$. By symmetry, if $(\mu, \mu P(a_1)) \in \widehat{\mathbb{A}}$, then $\mu \in \mathbb{A}^{11} \cap \mathbb{A}^{10} \cap \mathbb{A}^{12} = \{\hat{\mu}^1\}$. ☐

The complete description of the Aubry set in Theorem 5.8.7 shows that there exists only two "attractors", which are the steady-states.

Characterization of the weak KAM solutions. It is shown in [AGW09, Theorem 8.1] that the functions h^z where z belongs to the projected Aubry sets are precisely the extremal generators of the set of weak-KAM solutions. Recall that the set of weak-KAM solutions is a tropical linear space, meaning that it is stable by taking the pointwise maximum of functions, and by translating a function by a constant. An element of a tropical linear space is extreme if it cannot be written as the pointwise maximum of two elements of this space which are both different from it. Theorem 8.1 of [AGW09] also states that any weak-KAM solution h of (5.10) can be written as

$$h(x) = \sup_{z\in\mathbb{A}}\{h^z(x) - \zeta(z)\} \enspace , \tag{5.40}$$

where $\zeta : \mathcal{X} \to \mathbb{R}$. In our case, this means that the functions $\{v^\lambda\}_{\lambda\in[0,1]}$ describes the whole set of solutions (up to a constant). In particular, v^k reads as the Mañé potential in $\hat{\mu}^k$ (up to a constant), i.e., $v^k(\mu) = -c^*(\mu, \hat{\mu}^k)$.

Note that the representation of the solutions as a supremum in (5.40) provides a discrete time analogue of the representation of weak-KAM solutions established by Fathi in the continuous time setting, see Theorem 8.6.1 in [Fat08].

Remark 5.8.3 (Explicit strict subsolution)

The proof of Theorem 5.8.7 allows us to derive a strict subsolution: let h_0 be defined as

$$h_0 = \sum_{i=0}^{1} \sum_{j=1}^{3} \alpha_{ij} w^{ij} \ ,$$

where $w^{ij} = \hat{h}^{ii} \vee \hat{h}^{ij}$ and $\sum_{k,j} \alpha_{ij} = 1$. Then, h_0 is strict on $(\Delta_3^{\leqslant}) \setminus \mathbb{A}$.

In this case, it is not necessary to make the combination of all the subsolutions h^z, $z \in \Delta_3^{\leqslant}$, but only the combination of 6 subsolutions. As a consequence, we can here state an analog result to Theorem 5.8.6, ensuring that any optimal policy converges to one of the steady states $\hat{\mu}^k$ and proving that no cycling behavior (such as periodic promotions like in Section 6.6) can happen.

5.9 Conclusion

We developed an ergodic control model to represent the evolution of a large population of customers, able to actualize their choices at any time. Using qualitative properties of the population dynamics (contraction in Hilbert's projective metric), we showed the existence of a solution to the ergodic eigenproblem, which we applied to a problem of electricity pricing. A numerical study reveals the existence of optimal cyclic promotion mechanisms, that have already been observed in economics. We also quantified the suboptimality of constant-price strategy in terms of a specific duality gap.

We extended the results in the presence of noise, and show that contraction arguments still lead to existence of solution for the ergodic eigenproblem. We also analyzed the problem through the weak KAM angle. In particular, we focused on the description of the Aubry set, to which the dynamics converge under any optimal policy. We provided an example that shows the non-uniqueness of the optimal bias function, and computed the projected Aubry set for this example, showing that the latter is not reduced to a single state (in contrast with situations where turnpikes property holds).

The present model has connections with partially observable MDPs, in which the state space is also a simplex. We plan to explore such connections in future work. Besides, the convergence of the solution of the discretized ergodic equation (associated to the grid Σ) to the continuous solution will also be studied.

Chapter 6

A Quantization Procedure for Nonlinear Pricing with an Application to Electricity Market

***

This chapter is based on the proceedings paper [Jac+23b], to which we add the reinterpretation of the customer choice as a Bregman Voronoï diagram (Section 6.4) and the comparison of the proposed pruning method with the Lloyd's procedure for Bregman criterion.

Abstract. We consider a revenue maximization model, in which a company aims at designing a menu of contracts, given a population of customers. A standard approach consists in constructing an incentive-compatible continuum of contracts, i.e., a menu composed of an infinite number of contracts, where each contract is especially adapted to an infinitesimal customer, taking his type into account. Nonetheless, in many applications, the company is constrained to offering a limited number of contracts. We show that this question reduces to an optimal quantization problem, similar to the *pruning problem* that appeared in the max-plus based numerical methods in optimal control. We develop a new quantization algorithm, which, given an initial menu of contracts, iteratively prunes the less important contracts, to construct an implementable menu of the desired cardinality, while minimizing the revenue loss. We apply this algorithm to solve a pricing problem with price-elastic demand, originating from the electricity retail market. Numerical results show an improved performance by comparison with earlier pruning algorithms.

6.1 Introduction

6.1.1 Motivation from electricity markets

Electricity retail markets are now open to competition in most countries, and providers are free to design a *menu* of offers/contracts in addition to regulated alternatives (fixed prices), so that each consumer can select among the vast jungle of offers the one which maximizes his utility. In this chapter, the choice of a contract is based on the minimization of the invoice (rational choice theory, see e.g. [Sco00]), and we suppose that each customer can adjust his consumption to the electricity prices (*price elasticity*). This phenomenon is highlighted by the current spike in energy prices: consumers are likely to make huge consumption reduction efforts in order to save money.

A key problem for electricity providers is to design an optimal menu of offers, maximizing their revenue, under a restriction on the "size" of the menu (number of contracts). In fact, from an optimization point of view, proposing more contracts increases the revenue, as it allows one to adjust the menu to the individual preferences of the different types of customers. However, in practice, it is essential to restrict the number of contracts, in order to make the commercial offer more visible to agents, easier to understand, and also to keep an implementable menu for the company.

6.1.2 The optimal nonlinear pricing problem

We consider more generally the revenue maximization problem faced by a seller, called *principal*, or *leader* in the setting of Stackelberg games [SC73]. This problem has been addressed by the theory of mechanism design [BKS15] through the question of nonlinear pricing. The so-called *monopolist problem* is among the most studied ones: in this approach, the population is represented as a continuum of buyers (called *agents* or *followers*), and a contract can be specifically designed for each agent (continuum menu). In the seminal paper [RC98], Rochet and Choné study the monopolist problem by introducing a dual approach. In some specific cases (linear-quadratic setting and specific agents distribution), analytic solutions can be found in one [MR78] or many dimensions [Arm96], via reformulation as welfare maximization using virtual valuation technique. Extending the framework of Rochet and Choné to decomposable variational problem under convexity requirement, Carlier [CD17] addresses the question of the existence and uniqueness of a solution, and proposes an iterative algorithm. In the specific case $\mathbb{R}^2$, Mirebeau [Mir16] introduces a more efficient method using an adaptive mesh based on stencils. The infinite-size menu is therefore characterized by a value-function satisfying the incentive-compatibility conditions as with the full-participation condition, the latter supposing that contracting with the whole population is optimal. Bergemann, Yeh and Zhang recently considered the question of the optimal quantization of a menu [BYZ21].

6.1.3 Contributions

Our main contribution is the development of new *quantization* algorithms which, given the infinite-size menu, aim at finding the best n-contracts approximation that maximizes the revenue. This 2-step strategy bypasses the combinatorial difficulty tackled in bilevel pricing – see e.g. [LMS98; BK19] – where formulations directly embed customer choices over the n contracts, becoming rapidly untractable for large size of menu. We show that the quantization problem is equivalent to the *pruning problem*, which arose, following McEneaney [McE07], in the development of the max-plus based curse-of-dimensionality attenuation methods in numerical optimal control, see [MDG08; GMQ11; GQS14], and [MD15] for an application. In these methods, the

6. Nb of offers

value function of an optimal control problem is represented as a supremum of "basis functions", and one looks for a sparse representation – with a prescribed number of basis functions. In the present application, the basis functions are linear functions, representing contracts. We develop a *greedy descent* algorithm which iteratively removes the less "important" contracts. We consider different importance measures, taking into account the L_1 and L_∞ approximation errors previously considered in the study of the pruning problem, and also a specific measure of the loss of revenue, see Algorithms 6 and 7. An essential feature of these algorithms is the low incremental cost per iteration, with an update rule requiring only *local* computations – in a "small neighborhood" of the active set of a basis function. To do so, we exploit discrete geometry techniques, by associating to a basis decomposition a polyhedral complex, which is updated dynamically.

To apply this algorithm to the optimal design of a menu in the electricity retail market, we generalize the framework of [CD17] to allow for a nondecomposable (still convex) cost. Indeed, the revenue of the provider depends on the supply cost, supposed to be an increasing function of the *global* consumption, see e.g. ([Ack+18; Ale+19]). In this extended setting, we prove the existence and uniqueness of the solution for the infinite-size menu (Theorem 6.2.1). The solving of this problem is then tackled by a direct method (discretization of the variational problem). We also take into account the elastic behavior of customers, who adapt their consumption according to prices, even on an annual scale (we focus here on year-based contracts). We show that, after a change of variables, the addition of a uniform elasticity actually reduces to the previous model (Theorem 6.5.1). Numerical tests, on a realistic instance (arising from the French electricity market), illustrate the efficiency of our approach both in terms of revenue gain and of computational time, see Figure 6.3. Our algorithm also allows one to estimate the minimal admissible number of contracts, given a target of acceptable revenue loss by comparison with the infinite-size case.

6.1.4 Related works

In the nonlinear pricing context, the restriction to a finite number of offers has been regarded only recently. In [BYZ21], the authors analyze the loss of revenue induced by this restriction, exhibiting upper bounds of order $1/n^{2/d}$, where d is the dimension and n the maximal number of contracts. A similar asymptotic error rate arose in a different setting of quantization theory, see e.g. [GMQ11]. Moreover, in the linear-quadratic setting of [BYZ21], the extreme distributions realizing the worst revenue loss satisfies separability conditions à la Armstrong [Arm96], leading to an explicit expression for the optimal quantization. We do not satisfy these requirements here, as we tackle a broader class of variational problem, hence the need of efficient methods to solve the pricing problems with a finite number of contracts. In [EM09], a discretization is obtained by writing the utility function as a supremum of finitely many affine functions, and so the solution they obtain can be viewed as a n-contracts menu. However, the scheme also discretizes the population (with the same size as the contracts). In the present application, this is not desirable, since the size of the population has to be much larger that the size of the menu.

The present algorithms should be compared with the pruning methods to compute a sparse representation of a function as maximum of a prescribed number of basis functions. The pruning problem was shown in [GMQ11] to be a continuous version of the facility location problem, a hard combinatorial optimization problem. The pruning algorithms developed in [MDG08; GMQ11] rely on a notion of importance metric, measuring the contribution of each basis function to the approximation error. A basic algorithm in [MDG08; GMQ11] performs a single pass which keeps only the n basis functions with the highest importance metric, the latter being evaluated either by solving a convex programming problem or in approximate way, after a discretization of the state space. A *greedy ascent* algorithm is also implemented in [GMQ11], adding incrementally functions by decreasing order of importance. In contrast, the present algorithm does not require a discretization of the state space. Moreover, the use of fast (local) updates of the importance measure allows us to perform a greedy descent starting from the complete family of basis functions, and removing at each stage the less important one. This leads to improved performances on our application case.

The chapter is organized as follows: in Section 7.2, we define the nonlinear pricing problem, adapted to our application case, and encompassing the monopolist framework. In Section 6.3, we approximate the continuum menu by a finite set of contracts, and present refined pruning algorithms with local update. Then, in Section 6.5, we specify the problem encountered in electricity markets, and show how it boils down to the general case of Section 7.2. Finally, we numerically study the effectiveness of our approach in Section 6.6.

6.2 Nonlinear Pricing with Coupling Costs

6.2.1 Notation

For two vectors x and y of $\mathbb{R}^d$, we denote by $\langle x, y\rangle$ the scalar product and $x \odot y$ the entrywise product. Moreover, for a discrete set S, we denote by $|S|$ the cardinality of S.

6.2.2 Generalized monopolist problem

Let us consider a heterogeneous population, where each agent in the population is defined by a d-dimensional vector of characteristics $x \in X$. We suppose that $X \subset \mathbb{R}^d_{>0}$ is a compact polyhedral domain. An agent of type x will derive a utility $\langle x, \alpha \odot q_k\rangle - p_k$ from consuming a good k with quality $q_k \in \mathbb{R}^d_{>0}$ and price $p_k \in \mathbb{R}_{>0}$. The vector $\alpha \in (\mathbb{R}^*)^d$ is an exogeneous data, viewed as a varying perception of the quality. The agents are distributed according to ρ satisfying $\int_X \rho(x)\mathrm{d}x = 1$.

Let us consider a *monopolist* (principal) who designs a contract menu represented by a pair of functions $x \mapsto (p(x), q(x)) \in P \times Q$. For each agent x, these functions indicate respectively the price and the quality that the agent is supposed to prefer. Here, P and Q are compact subsets of $\mathbb{R}_{>0}$ and $\mathbb{R}^d_{>0}$. To ensure that the contract $(p(x), q(x))$ really satisfies agent of type x (i.e., is optimal for him), an additional constraint on the shape of the function, called *incentive-compatibility* condition is required: denoting by $u(x) := \langle x, \alpha \odot q(x)\rangle - p(x)$ the utility function for the menu designed by the monopolist,

$$u(y) - u(x) \geqslant \langle y - x, \alpha \odot q(x)\rangle \,, \ \forall x, y \in X \ . \tag{6.1}$$

Let U_x be the set of admissible values of u for type x:

$$U_x := \{\langle x, \alpha \odot q\rangle - p \mid (p, q) \in P \times Q\} \ .$$

Each set U_x is compact by compactness of P and Q.

> **Proposition 6.2.1** ([Roc87])
> Let $q(\cdot)$ be defined on X, with values in Q. There exists a function $p : X \to P$ such that $u(\cdot)$ satisfies (6.1) if and only if
>
> (i) $u(x) \in U_x$, for $x \in X$,
>
> (ii) u is convex on X,

(iii) $\nabla u(x) = \alpha \odot q(x)$ for a.e. $x \in X$.

The aim of the monopolist is then to maximize a revenue function, defined as

$$J(u,q) := \int_X L(x,u(x),q(x))\mathrm{d}x - C\left(\int_X M(x,q(x))\mathrm{d}x\right) , \tag{6.2}$$

In (6.2), the cost function C takes as input data aggregated on the whole domain X. Such coupling cost naturally appears in some applications, for instance in electricity retail market, see Section 6.5.

Assumption 6.2.1. The integrand L is linear in u and q. Moreover, the integrand M is strictly convex in q, and C is increasing and strictly convex.

In addition to the incentive-compatibility condition, the utility must be greater than a reservation utility:

$$u(x) \geqslant R(x) \ . \tag{6.3}$$

The problem solved by the monopolist is then

$$\max_{u,q} \left\{ J(u,q) \ \middle| \begin{array}{l} u,q \text{ satisfy } (6.1),(6.3) \\ (u(x),q(x)) \in U_x \times Q \text{ for } x \in X \end{array} \right\} \tag{6.4}$$

Theorem 6.2.1

Under Assumption 6.2.1, Problem (6.4) has a unique optimal solution.

Proof. Using Proposition 6.2.1, for any solution, $\nabla u(x) = \alpha \odot q(x)$ for a.e. $x \in X$. We then directly study the existence and uniqueness in $(u, \nabla u)$.

Existence. Let $H^1(X)$ be the Sobolev space associated with X. We define

$$\mathcal{K} = \left\{ u \ \middle| \begin{array}{l} u \text{ convex and } u \geqslant R \\ u(x) \in U_x,\ \alpha^{-1} \odot \nabla u(x) \in Q \ ,\ \forall x \in X \end{array} \right\} .$$

The set $\mathcal{K}$ is a closed, convex and bounded subset of $H^1(X)$ (it is bounded since X is bounded and $\|u\|_{L_\infty}$ and $\|\nabla u\|_{L_\infty}$ are bounded too; it is convex since R is convex).

Besides, J is concave (Assumption 6.2.1). Morover, as $Q \subset \mathbb{R}^d_{>0}$, there exist $a, b > 0$ such that for any $x \in X$, $a \leqslant \|\nabla u(x)\| \leqslant b$. Therefore, there exists $c \in \mathbb{R}_+$ such that $|J(u, \nabla u)| \leqslant c$, and as a consequence, J is continuous on $\mathcal{K}$, see [ET99, Chapter 1, Proposition 2.5].

Using the fact that $H^1(X)$ is reflexive and [ET99, Chapter 2, Proposition 1.2], Problem (6.4) admits at least one solution.

Uniqueness. (Same arguments as in [RC98]) Let now consider two distinct solutions u_1 and u_2. Then, if $\nabla u_1 \neq \nabla u_2$ on a measurable subset, any function $tu_1 + (1-t)u_2$ is valid and gives a strictly better solution than u_1 and u_2 (due to strict convexity of the cost function $u \mapsto C(\int_X M(x, \nabla u(x))\mathrm{d}x)$ and linearity of L). Therefore, $u_1 - u_2$ is a constant function. By linearity of L, the objective value obtained with u_1 and u_2 differ by the same constant. This contradicts the optimality of the two solutions u_1 and u_2. □

This result should be compared with [Car01], where the (decomposable) criteria is defined by an integrand that must satisfy coercivity condition, which entails that a minimizing sequence (u_n) must be bounded in the $W^{1,1}$ Sobolev norm. Here, J is not necessarily coercive. Instead, the compactness argument directly comes with assumptions on P and Q.

6.2.3 Discretized solution for the infinite-size case

As an extension of the monopolist problem, Problem (6.4) can be solved to optimality through a discretization scheme. In [EM09], the authors proved the convergence of the discretized problem to the continuous one, which can be extended to nondecomposable cost. Efficient numerical methods have been proposed in [CD17] and [Mir16].

Definition 6.2.1 (Maximization/minimization diagram [BNN10]). Let $S \subseteq \{1, n\}$ and $\{\hat{u}_i\}_{i\in S}$ be a set of $|S|$ continuous functions defined on $X \subset \mathbb{R}^d$.

(i) We call *upper envelope* (resp. *lower envelope*) of the functions $\{\hat{u}_i\}_{i\in S}$ the function $x \mapsto \max_{i\in S} \hat{u}_i(x)$ (resp. $x \mapsto \min_{i\in S} \hat{u}_i(x)$).

(ii) We call *maximization diagram* (resp. *minimization diagram*) of $\{\hat{u}_i\}_{i\in S}$ the subdivision of X into cells such that, on each cell, $\arg\max_i \hat{u}_i$ (resp. $\arg\min_i \hat{u}_i$) is fixed.

Let us define a regular grid Σ of X. Each of the methods provides a solution $\{(\hat{p}_i, \hat{q}_i)\}_{i\in\Sigma}$, inducing a convex utility function $\hat{u}_\Sigma$ that can be represented as the supremum of affine functions, with the notation:

$$\hat{u}_S(x) = \max_{i\in S} \hat{u}_i(x),\ S \subseteq \Sigma \ , \tag{6.5}$$

where $\hat{u}_i : x \in \mathbb{R}^d \mapsto \langle \alpha \odot \hat{q}_i, x\rangle - \hat{p}_i$. Therefore, the discretized solution $\hat{u}_\Sigma$ is the upper envelope of the family $\{\hat{u}_i\}_{i\in\Sigma}$ and induces a maximization diagram. In the context of max-plus methods [McE07; AGL08], the functions $\hat{u}_i$ are called *basis functions* and can be more general than affine functions, but we focus here on this specific case, as this naturally appears in the model (affine contracts).

6.3 Pruning procedures

6.3.1 Pruning method for max-plus basis decomposition

Let us now suppose that the monopolist has a maximal number of n contracts he can design. Given the discretized infinite-size solution u_Σ, the question can be recast as the following combinatorial problem:

$$\min_{S\subseteq\Sigma} \{d(\hat{u}_S, \hat{u}_\Sigma) \text{ s.t. } |S| \leqslant n\} \ , \tag{6.6}$$

where the function $d(\cdot)$ can be either

(i) the L_∞ norm $d_\infty(u, v) = \|u - v\|_{L_\infty(X)}$,

(ii) the L_1 norm $d_1(u, v) = \|u - v\|_{L_1(X)}$,

(iii) and the J-based criterion $d_J(u, v) = J(v, \alpha^{-1} \odot \nabla v) - J(u, \alpha^{-1} \odot \nabla u)$.

The third case corresponds to the maximization of the function J, where $\alpha^{-1} \odot q = \nabla u$ thanks to Proposition 6.2.1.

Theorem 6.3.1 ([GMQ11])
Let $X \subseteq \mathbb{R}^d$ and $v : X \to \mathbb{R}$ strongly convex of class $\mathcal{C}^2$. Then, both L_1 and L_∞ approximation errors are $\Omega\left(\frac{1}{n^{2/d}}\right)$ as $n \to \infty$.

Theorem 6.3.1 exhibits an error rate identical to the complexity bound proved in [BYZ21] in a different setting.

We define the *importance metric* of basis function i as

$$\nu(S,i) = d(\hat{u}_{S\backslash\{i\}}, \hat{u}_S) \ . \tag{6.7}$$

This corresponds to an incremental version of the criteria (6.6). For the L_∞ and L_1 case, if $\nu(S,i) = 0$, then the i-th basis function does not contribute to the max-sum. Otherwise, if $\nu(S,i) > 0$, then it expresses the maximal difference between the shape of $\hat{u}_S$ with and without $\hat{u}_i$, depending on the criterion. For the criterion d_J, it expresses the loss of revenue for the principal when contract i is removed.

6.3.2 Specific case: minimizing L_∞ error

For an L_∞ approximation error, the importance metric (6.7) can be computed by solving a linear program, see [GMQ11]:

$$\begin{aligned} \max_{x\in X,\,\nu} \quad & \nu \\ \text{s.t} \quad & \forall j \in S\backslash\{i\}, \quad \hat{u}_i(x) - \hat{u}_j(x) \geqslant \nu \quad (\lambda_{ij}) \end{aligned} \tag{P_i^S}$$

In (P_i^S), we denote by $(\lambda_{ij})_j$ the dual variable associated with each constraint. The set of saturated constraints is then characterized by the positive variables λ_{ij}.

Algorithm 6 Pruning for L_∞ importance metric

```
Require: n                                    ▷ Desired number of contracts
  S ← Σ                                       ▷ Indices of kept contracts
  I ← Σ                                       ▷ Indices of problems to re-compute
  for t = 1 : |Σ| − n do
      for i ∈ I do
          ν_i, λ_i ← solution of (P_i^S)
          J_i ← {j ∈ S\{i} | λ_ij > 0}
      end for
      r ← arg min_{i∈S} ν_i                   ▷ Contract to remove
      S ← S\{r}
      I ← {i ∈ S | r ∈ J_i}
  end for
  return S
```

Algorithm 6 describes a *greedy descent* procedure: we start from the complete set of contracts S, and iteratively remove the less important contract exploiting a fast local update of the importance metric. Compared with [GMQ11], the importance metric is computed *exactly*, i.e., without discretization of the space X. Moreover, we take advantage of the linearity of the basis functions $\hat{u}_i$ to exploit the optimal dual variables λ_{ij} in the linear program (P_i^S):

Proposition 6.3.1 (Local update)
Let λ_{ij} be the optimal dual variables in (P_i^S) for a contract $i \in S$. Then, the importance metric of i stays *unchanged* when we remove a contract $j \in S$ s.t. $\lambda_{ij} = 0$, i.e., $\nu(S\backslash\{j\}, i) = \nu(S,i)$.

Proposition 6.3.1 ensures the correctness of Algorithm 6, where we only re-compute at each iteration the values ν_i for a very small subset of Σ. This leads to a huge gain in computation time, see Section 6.6.

6.3.3 L_1 and J-based approximation error

Contrary to the L_∞ case, the computation exploits the geometric structure. Indeed, the representation of the function $\hat{u}_S$ as a maximum of basis functions $\hat{u}_j, j \in S$ induces a maximization diagram, see Section 6.2.3. Each cell is denoted by $\mathcal{C}_i$ and corresponds to the types $x \in X$ such that $\hat{u}_S(x) = \hat{u}_i(x)$. Removing a basis function $\hat{u}_i$ from the supremum $\hat{u}_S = \sup_{j \in S} \hat{u}_j$ yields a local modification of the latter supremum, concentrated on a neighborhood of the cell $\mathcal{C}_i$. Hence, we will need to compute at each iteration the neighbors of each contract cell $\mathcal{C}_i$ with $i \in S$. This idea may be compared with the notion of Delaunay triangulation associated to a Voronoï diagram [For95]. During the algorithm, we keep in memory two sets: J_i represents the neighboring cells of cell i, and V_i is the vertex representation of cell i. Two routines are used for both the L_1 and J-based criterion:

- ◇ VREP(S, i) returns the V-representation (representation by vertices) of the polyhedral cell $\mathcal{C}_i$ induced by contract i for a given set S, taking as input the H-representation (representation by half-spaces) $\{x \in X \mid \hat{u}_i(x) \geqslant \hat{u}_j(x), \ \forall j \in S\}$ of the cell i. This is done using the revised *reverse search* algorithm implemented in the library `lrs`, see [Avi00].
- ◇ UPDATENEIGHBORS$((V_S)_{i \in I})$ updates the neighbors of each cell $i \in I$ knowing the vertex representation.

Proposition 6.3.2 (Local update)
The importance metric of a contract $i \in S$ stays *unchanged* when we remove a contract j which is not in the neighborhood of i, i.e., $\nu(S \backslash \{j\}, i) = \nu(S, i)$ for $j \in S \backslash J_i$.

Proposition 6.3.2 ensures the correctness of Algo. 7, where we only re-compute vertex representations for a small subset of contracts (corresponding to the neighboring cells of the lastly removed contract, see line 8 of the algorithm). This local update is illustrated in Figure 6.1. The update of the importance metric in line 11 differs between the L_1 and J-based cases, and is described in Algos. 8a–8b. In Algo. 8a, the integral that appears in the computation of ν_i can be evaluated analytically using Green's formula, as it integrates a linear form over a polytope, see Appendix 6.8.1. In Algo. 8b, δ_L can be computed in the same way. For M_0 and δ_M, this generally involves the integration of the function $x \mapsto M(x, \hat{q}_i)$. In the present application, this function is linear, and so the direct integration is possible, see (6.17)–(6.18).

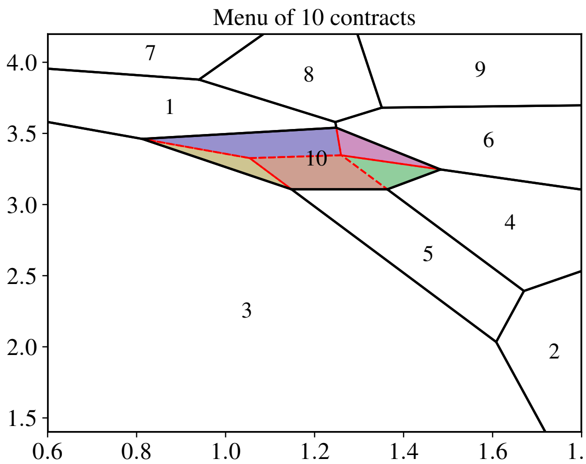


Figure 6.1: Evaluation of contract by dividing into subregions ($d = 2$)
The green polyhedron corresponds to $F_{4,-10} \cap V_{10}$.

Algorithm 7 Pruning with local update (for L_1 and J-based)

Require: n ▷ Desired number of contracts
for $i \in \Sigma$ **do**
  $V_i \leftarrow \textsc{Vrep}(\Sigma, i)$ ▷ Vertex representation
end for
$S \leftarrow \Sigma$ ▷ Indices of kept contracts
$I \leftarrow \Sigma$ ▷ Indices of problems to re-compute
for $t = 1 : |\Sigma| - n$ **do**
  $(J_i)_{i \in I} \leftarrow \textsc{updateNeighbors}((V_i)_{i \in I})$
  for $i \in I$, $j \in J_i$ **do**
    $F_{j,-i} \leftarrow \textsc{Vrep}(S \backslash \{i\}, j)$ ▷ Future cells
  end for
  $\nu \leftarrow \textsc{updateImpMetric}(I, (V_i)_{i \in S}, (F_{j,-i})_{j \in J_i, i \in S})$
  $r \leftarrow \arg\min_{i \in S} \nu_i$ ▷ Contract to remove
  $S \leftarrow S \backslash \{r\}$
  for $j \in J_r$ **do**
    $V_j \leftarrow F_{j,-r}$ ▷ Update vertex representation
  end for
  $I \leftarrow J_r$
end for
return S

Algorithm 8a $\textsc{updateImpMetric}$ (L_1 error)

Require: I, $(V_i)_{i \in S}$, $(F_{j,-i})_{i \in I, j \in J_i}$
for $i \in I$ **do** ▷ Update metric on cells
  $\nu_i \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} (\hat{u}_i(x) - \hat{u}_j(x)) \mathrm{d}x$
end for
return ν

Algorithm 8b $\textsc{updateImpMetric}$ (J-based error)

Require: I, $(V_i)_{i \in S}$, $(F_{j,-i})_{i \in I, j \in J_i}$
$M_0 \leftarrow \sum_{i \in S} \iint_{V_i} M(x, \hat{q}_i) \mathrm{d}x$
for $i \in S$ **do** ▷ Update metric on cells
  $$\delta_L \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} L(x, \hat{u}_i(x), \hat{q}_i) - L(x, \hat{u}_j(x), \hat{q}_j) \mathrm{d}x$$
  $\delta_M \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} M(x, \hat{q}_j) - M(x, \hat{q}_i) \mathrm{d}x$
  $\nu_i \leftarrow \delta_L - C(M_0) + C(M_0 + \delta_M)$
end for

Proposition 6.3.3 (Critical steps)

Let m be the maximum number of neighbors of a polyhedral cell during the execution of the algorithm (for all t and i, $|J_i| \leqslant m$). Then,

- The number of linear programs (P_i^S) solved in Algo. 6 is in $O(m|\Sigma|)$,
- The number of computations of a vertex representation of a polyhedral cell (calls to $\textsc{Vrep}(S, i)$ in Algo. 7 / reverse search) is in $O(m^2|\Sigma|)$.

By comparison with Proposition 6.3.3, a naïve implementation (full recomputation of the importance metric at each step) of the two algorithms would respectively lead to a number of critical steps in $O(|\Sigma|^2)$ and $O(m|\Sigma|^2)$. Indeed, each linear program P_i^S can be solved in polynomial time (by an interior point method), and reverse search has an incremental running time of $O(|\Sigma|d)$ per vertex if the input is nondegenerate, see [Avi00].

6.4 Links with k-means clustering with Bregman divergence

Let us first introduce the notion of Bregman divergence and Bregman Voronoï diagram:

Definition 6.4.1 (Bregman Divergence [BNN10]). We define Bregman divergence $D_u : X \times X \to \mathbb{R}_+$ with respect to a convex diferentiable function u as

$$D_u(x,y) = u(x) - u(y) - \langle x - y, \nabla u(y)\rangle \tag{6.8}$$

Definition 6.4.2 (Bregman Voronoï diagram [BNN10]). Let $\mathcal{S} = \{\mu_1,\dots,\mu_n\}$ be a set of n points of X. We call *(first-type) Bregman Voronoï diagram* of $\mathcal{S}$, denoted by $\mathrm{vor}_u(\mathcal{S})$, the minimization diagram of the family $\hat{d}_i(x) := D_u(x,\mu_i)$, $i \in [n]$:

$$\mathrm{vor}_u(\mu_i) := \{x \in X \mid D_u(x,\mu_i) \leqslant D_u(x,\mu_j),\ \forall j \in [n]\} \ . \tag{6.9}$$

The point μ_i, associated with the Voronoï cell $\mathcal{C}_i = \mathrm{vor}_u(\mu_i)$, is called a *site*.

In the definitions, we do not require strict convexity of the function u, as it is generally the case in the definition of Bregman divergence, as we will study Bregman Voronoï diagrams of piecewise linear convex function. This means that we allow the Bregman pseudo-distance between two distinct points to be zero.

Proposition 6.4.1 (Interpretation as Voronoï diagram)
Let $\mathcal{S} = \{\mu_1,\dots,\mu_n\}$ be a set of n points of X. We define the family of functions $\hat{u}_i$ as the supporting hyperplanes of u at μ_i, i.e.,

$$\hat{u}_i(x) = u(\mu_i) + \langle x - \mu_i, \nabla u(\mu_i)\rangle \ .$$

Then, the maximization diagram of $\{\hat{u}_i\}_{1\leqslant i\leqslant n}$ and the Bregman Voronoï diagram of $\mathcal{S}$ coincides.

Proof.

$$\begin{aligned}
D_u(x,\mu_i) \leqslant D_u(x,\mu_j) &\iff u(x) - u(\mu_i) - \langle x-\mu_i, \nabla u(\mu_i)\rangle \leqslant u(x) - u(\mu_j) - \langle x-\mu_j, \nabla u(\mu_j)\rangle \\
&\iff u(\mu_i) + \langle x-\mu_i, \nabla u(\mu_i)\rangle \geqslant u(\mu_j) + \langle x-\mu_j, \nabla u(\mu_j)\rangle \\
&\iff \hat{u}_i(x) \geqslant \hat{u}_j(x)
\end{aligned}$$

□

Proposition 6.4.1 shows that the maximization diagram induced by a quantized n-contracts solution – obtained by pruning procedure from the complete solution $\hat{u}_\Sigma$ – is in fact a Bregman Voronoï diagram associated to $\hat{u}_\Sigma$ (the polyhedral complex displayed in Figure 6.1 is therefore an example of such a diagram).

In the sequel, let u be a convex function on a convex set $X \subset \mathbb{R}^d$, such that u is differentiable almost everywhere. We associate to X the p.d.f. ρ satisfying $\int_X \rho(x)\mathrm{d}x$. We also use the notation

$$\rho_{|\mathcal{C}}(x) = \rho(x) / \int_{\mathcal{C}} \rho(y)\mathrm{d}y \ .$$

Clustering. We denote by $L_u(\mathcal{S})$ the loss of optimality induced by a set of representatives $\mathcal{S} = \{\mu_1, \ldots, \mu_n\}$:

$$\begin{aligned} L_u(\mathcal{S}) &= \sum_{i=1}^{n} \int_{\mathrm{vor}_u(\mu_i)} D_u(x, \mu_i)\rho(x)\mathrm{d}x \\ &= \int_X \min_{1 \leqslant i \leqslant n} D_u(x, \mu_i)\rho(x)\mathrm{d}x \\ &= \int_X (u(x) - \max_{1 \leqslant i \leqslant n} \hat{u}_i(x))\rho(x)\mathrm{d}x \end{aligned}$$

If ρ is the uniform distribution over X, then $L_u(\mathcal{S})$ is the L_1-error between $u(\cdot)$ and the upper envelope of $\{\hat{u}_i\}_{1 \leqslant i \leqslant n}$. Algorithm 9 – initially introduced in [Ban+05] – adapts the Lloyd's procedure [Llo82] for clustering with Bregman divergence.

Algorithm 9 Bregman Hard Clustering – Lloyd procedure

Require: number of cluster n, initial centroids $\{\mu_i^{(0)}\}_{1 \leqslant i \leqslant n}$
$t \leftarrow 0$
do
  $\mathcal{C}_i^{(t)} \leftarrow \{x \in X \mid D_u(x, \mu_i^{(t)}) \leqslant D_u(x, \mu_j^{(t)}),\ \forall j \in [n]\}$ for all $i \in [n]$ ▷ Assignment step
  $\mu_i^{(t+1)} = \int_{\mathcal{C}_i^{(t)}} x \rho_{|\mathcal{C}_i^{(t)}}(x)\mathrm{d}x$ ▷ Centroid estimation step
  $t \leftarrow t + 1$
while there exist $i \in [n]$ such that $\mu_i^{(t)} \neq \mu_i^{(t-1)}$
return $\{\mu_i^{(t)}\}_{1 \leqslant i \leqslant n}$

At each iteration, the algorithm achieves two steps : (i) an assignment step constructing the cells of the Voronoï diagram, and (ii) a re-estimation step computing the coordinates of each centroïd. This iterative algorithm is proved to monotonically decrease the loss function (and therefore to monotonically decrease the L_1-error between the function u and its upper envelope $\{\hat{u}_i\}_{1 \leqslant i \leqslant n}$).

Proposition 6.4.2 ([Ban+05], Proposition 2)

Algorithm 9 produces a sequence of centroïds $\{\mathcal{S}^{(t)} := \{\mu_1^{(t)}, \ldots, \mu_n^{(t)}\}\}_{t \geqslant 0}$ such that loss sequence $\{L_u(\mathcal{S}^{(t)})\}_{t \geqslant 0}$ is decreasing.

The latter algorithm exploits the equality between the geometrical barycentre of the cell and the definition of the centroïd as the minimizer of the Bregman divergence integrated on the cell, see the following proposition:

Proposition 6.4.3 ([Ban+05], Proposition 1)

Let $\mathcal{C} \subseteq X$. Then, the centroïd of $\mathcal{C}$ coincides with the (unique) minimizer of the Bregman information:

$$\int_{\mathcal{C}} x \rho_{|\mathcal{C}}(x)\mathrm{d}x = \arg\min_{\mu \in \mathcal{C}} \int_{\mathcal{C}} D_u(x, \mu)\rho_{|\mathcal{C}}(x)\mathrm{d}x \ . \tag{6.10}$$

The terminal partition obtained by Algorithm 9 is locally optimal, i.e., the loss function cannot be reduced by any perturbation of the centroïds positioning. We refer to [LB16, Theorem 1] for asymptotic quantization rate, previously shown in another context in [GMQ11].

Algorithm 9 provides an alternative method to the pruning procedure to find the best n-contracts approximation of $\hat{u}_\Sigma$ that minimizes the L_1-error, see the next section for numerical comparison.

6.5 Application to Electricity Markets

6.5.1 Price elasticity

Let us consider a provider holding several contracts, each of them defined by a fixed price component $p \in \mathbb{R}$ (in €), and d variable price components $z \in \mathbb{R}^d$ (in €/kWh). In France, the contracts often take into account $d = 2$ time periods, with different prices for Peak / Off-peak consumptions. Moreover, the price coefficients (p, z) of each contract are supposed to belong to a non-empty polytope $P \times Z \subset \mathbb{R}^{d+1}$:

Assumption 6.5.1. Let p^-, p^+ be in $\mathbb{R}_{>0}$ and z^-, z^+ be in $\mathbb{R}^d_{>0}$. Then, $P = [p^-, p^+]$, and the polytope Z is of the following form:

$$Z := \left\{ z^- \leqslant z \leqslant z^+ \mid z_{i_1} \leqslant \kappa_{i_1,i_2} z_{i_2} \text{ for } i_1 \leqslant_{\mathcal{P}} i_2 \right\} \ ,$$

where $\mathcal{P}$ is a *partially ordered set* (poset) of $\{1, \dots, d\}$, and $\leqslant_{\mathcal{P}}$ the ordering relation, and $\kappa_{i_1,i_2} > 0$. When $\kappa \equiv 1$, $z^- \equiv 0$ and $z^+ \equiv 1$, Z is known as an *order polytope* [Sta86].

Assumption 6.5.1 is natural for the electricity pricing problem: the price can be freely determined within a box (bounds), as long as some inequalities between peak price coefficients and off-peak price coefficients are fulfilled.

We suppose that each agent in the (infinite-size) population is characterized by a *reference* consumption vector $\check{x} \in X \subset \mathbb{R}^d_{>0}$. Here, supposing a continuum of agents is justified since we consider in the application case the population of a whole country. We suppose that the consumption is *elastic* to prices, i.e., a consumer can deviate from its reference consumption $\check{x}$. In addition, we suppose that electricity elasticity can be captured into a utility-based framework, see e.g. [Sam+12] for the properties that the utility must satisfy. Here, we focus on isoelastic utilities:

Assumption 6.5.2 (Isoelastic utility function)**.** For a reference consumption $\check{x}$, the utility of consuming an amount of energy $x \in \mathbb{R}^d_{\geqslant 0}$ is depicted through a *Constant Relative Risk Aversion* (CRRA,[Pin12; Ala+20]) or isoelastic utility:

$$\mathcal{U}_{\check{x}} : x \in \mathbb{R}^d_{\geqslant 0} \mapsto \frac{1}{\eta} \sum_{i=1}^{d} \beta_{\check{x}i} (x_i)^\eta, \ \eta \in (-\infty, 0) \cup (0, 1] \ . \tag{6.11}$$

The coefficient η is called the *risk aversion* coefficient.

In this context, this elasticity measure depicts the easiness of a customer to adopt another energy source to fulfill his needs. In [Ala+20], the authors model the electric elasticity by this kind of utility function, and separate the case $\eta < 0$ and $\eta \in (0, 1]$. The first regime ($\eta < 0$) will model a household consumption: the satisfaction coming from consuming energy saturates to a maximum utility, and a zero consumption is prohibited. In contrast, the second regime ($\eta \in (0, 1]$) will represent the high flexibility of the industrial sector, which can adapt more easily its consumption according to price. We refer to [NC20] and references therein for empirical studies on the intensity of the elasticity coefficient η.

For a contract defined by price coefficients $(p, z) \in \mathbb{R} \times \mathbb{R}^d$, a consumer $\check{x}$ will optimize his consumption in order to maximize the *welfare function*, obtained by subtracting the electricity cost to (6.11):

$$\mathcal{U}^*_{\check{x}} : (p, z) \in \mathbb{R} \times \mathbb{R}^d \mapsto \max_{x \in \mathbb{R}_{\geqslant 0}{}^d} \{ \mathcal{U}_{\check{x}}(x) - \langle x, z \rangle \} - p \ . \tag{6.12}$$

6. Nb of offers

We denote by $\mathcal{U}^*_{\check{x}}$ the welfare function as the maximization term in (6.12) corresponds to a Fenchel-Legendre transform up to a change of sign. As a consequence, $\mathcal{U}^*_{\check{x}}$ is convex and nonincreasing. We now make the following assumption to fix the value of β:

Assumption 6.5.3. The reference consumption $\check{x} \in \mathbb{R}^d$ is obtained for reference prices $\check{p} \in \mathbb{R}$ and $\check{z} \in \mathbb{R}^d$.

Under Assumption 6.5.3, the optimal consumption of customer $\check{x}$ on period $i \in \{1, \dots, d\}$, denoted $\mathcal{E}_{\check{x}i}$, is given by

$$\mathcal{E}_{\check{x}i}(z) = \check{x}_i \left(z_i/\check{z}_i\right)^{\frac{-1}{1-\eta}} \geqslant 0 \ , \tag{6.13}$$

and the welfare function is given by

$$\mathcal{U}^*_{\check{x}}(p, z) = \left(\frac{1}{\eta} - 1\right) \sum_{i=1}^{d} \check{x}_h \check{z}_h \left(z_h/\check{z}_h\right)^{\frac{-\eta}{1-\eta}} - p \ . \tag{6.14}$$

Equations (6.13) and (6.14) are obtained from the first order optimality condition (zero derivative) for (6.12) ($\beta_{\check{x}i} = \check{z}_h \left(\check{x}_h\right)^{1-\eta}$).

6.5.2 Infinite-size menu of offers

In this section, we relax the assumption of a finite number of contracts, by supposing that the provider is able to define *as many offers as* consumers. Therefore, the infinite-size menu of offers can be represented by two functions $p : X \to \mathbb{R}$ and $z : X \to \mathbb{R}^d$, representing respectively the fixed price component and the variable price components. Let us define the (weighted) invoice of a consumer as

$$\mathcal{L}_{\check{x}} : (p, z) \in \mathbb{R} \times \mathbb{R}^d \mapsto (p + \langle \mathcal{E}_{\check{x}}(z), z\rangle)\rho(\check{x}) \ , \tag{6.15}$$

where $\rho(\check{x}) \geqslant 0$ represents the density of customers with reference consumption $\check{x}$. The provider's revenue maximization problem is then

$$\max_{p,z} \ \mathcal{J}^1(p, z) - \mathcal{J}^2(z) \tag{6.16a}$$

$$\text{s.t. } \mathcal{U}^*_x(p(x), z(x)) \geqslant \mathcal{U}^*_x(p(y), z(y)), \ \forall x, y \in X \tag{6.16b}$$

$$\mathcal{U}^*_x(p(x), z(x)) \geqslant R(x), \ \forall x \in X \tag{6.16c}$$

$$p(x) \in P, \ z(x) \in Z \tag{6.16d}$$

where $\mathcal{J}^1(p, z) = \int_X \mathcal{L}_x(p(x), z(x))\mathrm{d}x$ and $\mathcal{J}^2(z) = C\left(\int_X \sum_{i=1}^d \mathcal{E}_{xi}(z(x))\rho(x)\mathrm{d}x\right)$.

Equations (6.16b) and (6.16c) are respectively the *incentive-compatibility condition* and *participation constraint*. Taking C as a strictly convex increasing function of the global consumption is often considered in the literature. In particular, this cost function is often modeled as a piecewise linear function, see e.g. [Ale+19], or as a quadratic function, see e.g. [Ack+18]. In fact, the marginal cost to supply electricity is not constant and increases with the consumption. The convexity of the reservation utility is also a classical assumption, as this reservation utility should be a supremum over the utilities of alternative offers (each of them being a linear function of the reference consumption).

Let us make the following change of variables:

$$q_i := \left(z_i/\check{z}_i\right)^{\frac{-\eta}{1-\eta}} \ .$$

Then, the consumption on period $i \in \{1, \dots, d\}$ is a convex function of q_i, expressed as $\mathfrak{E}_{\check{x}i}(q_i) = \check{x}_i [q_i]^{\frac{1}{\eta}}$, and both the utility and the weighted invoice now read as linear functions of p and q: defining $\alpha = (\eta^{-1} - 1)\check{z}$,

$$\begin{aligned} u(x) &:= \langle x, \alpha \odot q(x) \rangle - p(x) \ , \\ L(x, u(x), q(x)) &:= \left(\tfrac{1}{\eta} \langle x, \check{z} \odot q(x) \rangle - u(x) \right) \rho(x) \ , \end{aligned} \tag{6.17}$$

Theorem 6.5.1

Under Assumption 6.5.1, the provider's revenue maximization problem (6.16) is equivalent to a monopolist problem of the form (6.4) with

$$M(x, q(x)) := \rho(x) \sum_{i=1}^{d} x_i [q_i(x)]^{\frac{1}{\eta}} \tag{6.18}$$

and, if $\eta < 0$,

$$Q = \left\{ q \in \mathbb{R}^d \;\middle|\; \begin{aligned} & \left(z^- / \check{z}\right)^{\frac{-\eta}{1-\eta}} \leqslant q \leqslant \left(z^+ / \check{z}\right)^{\frac{-\eta}{1-\eta}} \\ & q_{i_1} \leqslant \left(\kappa_{i_1, i_2} \frac{\check{z}_{i_2}}{\check{z}_{i_1}}\right)^{\frac{-\eta}{1-\eta}} q_{i_2} \text{ for } i_1 \leqslant_{\mathcal{P}} i_2 \end{aligned} \right\},$$

otherwise,

$$Q = \left\{ q \in \mathbb{R}^d \;\middle|\; \begin{aligned} & \left(z^+ / \check{z}\right)^{\frac{-\eta}{1-\eta}} \leqslant q \leqslant \left(z^- / \check{z}\right)^{\frac{-\eta}{1-\eta}} \\ & q_{i_1} \geqslant \left(\kappa_{i_1, i_2} \frac{\check{z}_{i_2}}{\check{z}_{i_1}}\right)^{\frac{-\eta}{1-\eta}} q_{i_2} \text{ for } i_1 \leqslant_{\mathcal{P}} i_2 \end{aligned} \right\}$$

Proof. Owing to assumption on the set Q and the strict monotonicity of $z \mapsto z^{\frac{-\eta}{1-\eta}}$ (increasing for $\eta < 0$ and decreasing for $\eta > 0$), one can explicitly derive the form of Q. The rest of the formulation is immediate. □

6.6 Numerical results

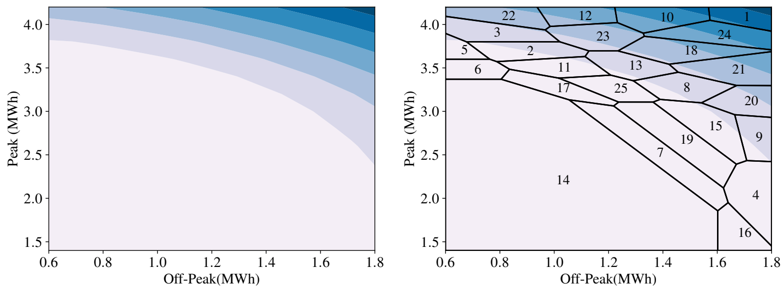
6.6.1 Instance

The numerical results were obtained on a laptop i7- 1065G7 CPU@1.30GHz. We provide in Table 6.1 the values of the parameters used in the application. In particular, we consider reference prices $(\hat{p}, \hat{z})$ corresponding to French regulated prices, and reference consumption spread around the mean French consumption per household ($\mathcal{E}_{\text{moy}} = 4$MWh). The cost function is taken as a quadratic function, scaled so that the marginal cost $C'(\mathcal{E}_{\text{moy}}) = 0.08$€/kWh. In comparison, the production cost is estimated in France around 0.05€/kWh for nuclear plants[1] and up to 0.09€/kWh for wind energy[2].

We display in Figure 6.2 the infinite-size menu and the quantized solution for two different sizes of menu (25 contracts and 10 contracts). In each cell $\mathcal{C}_i$, the contract i brings to customers of reference consumption $x \in \mathcal{C}_i$ the maximal utility given the quantized menu, i.e., $\hat{u}_S(x) = \hat{u}_i(x)$ for $x \in \mathcal{C}_i$. We observe that there is a region/cell (light gray region) where the monopolist reproduces the alternative option (of utility R). On the other side, for high consumption (peak or off-peak), the monopolist manages to design contracts that provide strictly higher utility than the regulated offer, and at the same time, procure to the monopolist a higher revenue.

[1] CRE (2022), *Délibération n° 2022-45*

[2] ADEME (2016), *Coûts des énergies renouvelables en France*



(a) Infinite-size menu (b) Menu of 25 contracts (c) Menu of 10 contracts

Figure 6.2: L_1-norm pruning for the electricity market case.
The normalized utility $u - R$ is depicted with colormap (light gray corresponds to the zero value and blue to high value).

η	-0.1
$\check{p}$	140€
$\check{z}$	(0,174,019)€/kWh
$C(\cdot)$	$0.01(\cdot)^2$
(p^-, p^+)	$(0, 500)$€
(q_1^-, q_1^+)	$(0.05, 0.5)$€/kWh
(q_2^-, q_2^+)	$(0.05, 0.5)$€/kWh
ρ	Uniform($[0.6, 1.8] \times [1.4, 4.2]$)
$R(\cdot)$	linear function (one regulated contract)

Table 6.1: Instance used in the numerical results.

6.6.2 Comparison of pruning objectives

In the upper graph of Figure 6.3, the three pruning objectives studied in the chapter (L_∞, L_1 and J-based) are compared with the 1-step approach of [MDG08; GMQ11]. The approach consists in sorting the importance metrics for all $i \in \Sigma$, and directly taking the n contracts with highest importance metric (here we consider the J-based importance metric). We display the relative objective loss, defined as 1 - J_t/J_{ref}, where J_t is the objective for a menu of size t and J_{ref} the objective obtained with the infinite-size menu. Note that removing a contract can induce a violation of the full-participation constraint ($u \geqslant R$). Therefore, in order to recover a feasible solution at each iteration, we lift up the solution with the simple rule $u \leftarrow u + \max\{\max_{x\in X}\{R(x) - u(x)\}, 0\}$.

On this example, the pruning procedure of Algo. 7 (greedy descent) leads to a significant loss reduction, whatever the criterion, compared with the 1-step approach. As expected, we observe that the J-based pruning has the smallest relative loss in the objective, as we minimize the error at each iteration of the process. In contrast, the L_∞-norm does not capture sufficiently well the behavior of the objective function J, and has larger objective loss, even for a large number of contracts.

We also depicted the cumulated time along the iterations in the lower graph of Figure 6.3 (we do not display the time for the 1-step procedure, as it is very fast, in less than 0.5s). For comparison, we add the cumulative time of a "naïve" J-based pruning, recomputing at each iteration the importance metric of each cell (global update). On this example, we observe that the computational time is already reduced by a factor almost 3 (this factor would be greater in higher dimension, as the neighborhood would be larger). As expected, the L_∞ criterion is the fastest, owing to the fast local update rule exploiting the sparsity of optimal Lagrange multipliers (Algorithm 6), and the J-based and L_1-norm criteria have similar computational time, as they use the same algorithmic architecture, see Algo. 7. In terms of loss minimization, the J-based pruning shows a loss of revenue reduced by a factor of around 2 by comparison with other methods. This approach allows us to determine the minimum number of contracts given an admissible revenue loss: e.g., Figure 6.3 shows that, with a J-based quantization, a menu of 10 contracts suffices to limit the revenue loss to 4%.

We finally compare the performance of the Lloyd algorithm with the pruning method with L_1-norm. We observe that the result of quite comparable. In particular, the pruning process and the Lloyd algorithm end up in this example with the same solution for a very restricted number of contracts. The difference lies in the computational time: the Lloyd algorithm is more adapted for a single run whereas the pruning procedure hightly benefit from the warm-start strategy and is particularly efficient when it comes to test a range of menu sizes.

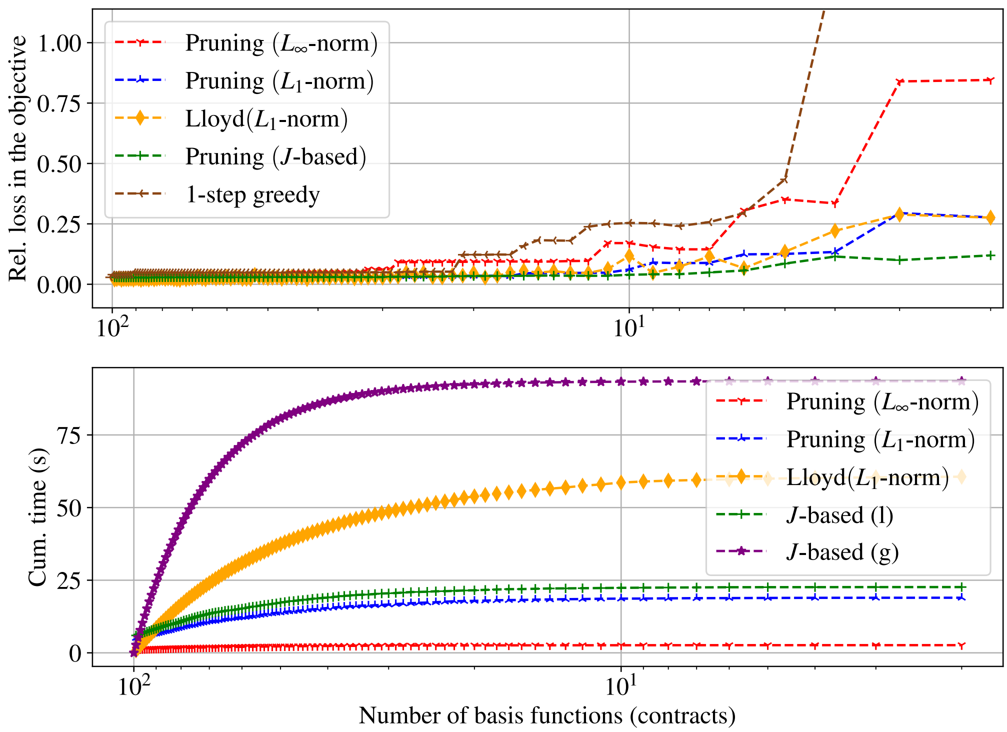


Figure 6.3: Comparison of error bounds for the three types of pruning objective, and for the LLoyd algorithm. The upper graph shows the loss of optimality induced by a reduced number of contracts. The lower graph depicts cumulative time along the iterations. (g) stands for global update while (l) stands for local update.

6.7 Conclusion

We have addressed a nonlinear pricing problem incorporating coupling costs. This arises in electricity markets, where supply costs depend on the global consumption. We have developed a quantization procedure, allowing to maximize the revenue of a provider, given a cardinality constraint on the set of contracts. This relies on refined pruning procedures, inspired by the max-plus basis methods in numerical optimal control. In particular, we exploited the local nature of the pruning process, in order to reduce the computational time. Thus, this leads to a new class of applications for methods originally developed in optimal control, and this also improves the complexity of a key ingredient of these methods.

A strong parallel with vector quantization can be made, see e.g. [Pag15]. In this context, a different quantization problem is addressed by Lloyd's procedures, *ibid.*. Whether these ideas can be adapted to the quantization of the maximum of affine functions with revenue criterion is left for further work.

6.8 Appendix

6.8.1 Green's formula on 2D-polytope

Proposition 6.8.1

Let P a 2D-polytope describes by its vertices $(x_i, y_i) \in \mathbb{R}^2$ (counter-clockwise ordered). Then for any $a, b, c \in \mathbb{R}$,

$$\iint_P (ax + by + c)dxdy = \sum_{i=1}^{N} \begin{bmatrix} \oint_{y_i}^{y_{i+1}} b(q_i + \frac{1}{\tau_i}y)ydy \\ - \oint_{x_i}^{x_{i+1}} (ax + c)(p_i + \tau_i x)dx \end{bmatrix}, \tag{6.19}$$

with $\tau_i = \frac{y_{i+1}-y_i}{x_{i+1}-x_i}$, $p_i := y_i - \tau_i x_i$ and $q_i := x_i - \frac{1}{\tau}y_i$.

Proof. The application of the Green formula gives :

$$\iint_P (ax + by + c)dxdy = \oint_{\mathcal{C}_P} (bxy)dy - (ax + c)ydx \ ,$$

where $\mathcal{C}_P$ is the contour of the polytope P. We then decompose on each edges, and use the change of variable $x = q + y/\tau$ in the first integral and $y = x + \tau x$ in the second one. □

Chapter 7

A Rank-Based Reward between a Principal and a Field of Agents: Application to Energy Savings

***

This chapter is based on the submitted paper [Ala+22].

Abstract. In this chapter, we consider the problem of a Principal aiming at designing a reward function for a population of heterogeneous agents. We construct an incentive based on the ranking of the agents, so that a competition among the latter is initiated. We place ourselves in the limit setting of mean-field type interactions and prove the existence and uniqueness of the equilibrium distribution for a given reward, for which we can find an explicit representation. Focusing first on the homogeneous setting, we characterize the optimal reward function using a convex reformulation of the problem and provide an interpretation of its behaviour. We then show that this characterization still holds for a sub-class of heterogeneous populations. For the general case, we propose a convergent numerical method which fully exploits the characterization of the mean-field equilibrium. We develop a case study related to the French market of Energy Saving Certificates based on the use of realistic data, which shows that the ranking system allows to achieve the sobriety target imposed by the European commission.

7.1 Introduction

7.1.1 Motivation

In Europe, energy retailers have incentives to generate energy consumption savings at the scale of their customer portfolio. For example in France, since 2006, power retailers – called *Obligés*

– have a target of a certain amount of Energy Saving Certificates[1] to hold at a predetermined future date (usually 3 or 4 years). If they fail to obtain this number of certificates, then they face financial penalties. Certificates can be acquired either by certifying energy savings at the customer or by buying certificates on the market. If a retailer holds more certificates than its target at the end of the period, the surplus can be sold on the Energy Saving Certificates market. The pluri-annual energy savings goal is determined by the government, and is function of the cumulative discounted amount of energy saved (thanks to thermal renovation for instance)[2]. Similar mechanisms – called *White certificates* – have been implemented in several countries in Europe (Great Britain, Italy or Denmark).

There is evidence from behavioral economics that energy consumption reductions can be motivated by providing a financial reward and/or information on social norms or comparison to customers, see e.g. see [AT14] or [DM15]. Especially, in [DM15], the authors find that social norms reduce consumption by around 6% (0.2 standard deviations). Secondly, they obtain that large financial rewards for targeted consumption reductions work very well in reducing consumption, with a 8% reduction (0.35 standard deviations) in energy consumption. For recent years, electricity providers are aware of this lever to make energy savings, and contracts offering bonus/rewards in compensation of reduction efforts appear, see e.g. the offers of "SimplyEnergy"[3], "Plüm énergie"[4] or "OhmConnect"[5]. The interest of this kind of solutions is reinforced in the current situation of gas and power shortage where many countries intend to diminish their global energy consumption[6].

7.1.2 Contributions

In this chapter, we design a monetary reward based on the *rank* of each consumer. In our context, the rank measures the reduction effort of a consumer compared with the rest of the population (a rank $r \in [0, 1]$ indicates that the consumer is among the r percent of the population with the highest consumption reduction). This new mechanism initiates a competition between similar consumers to be the best energy saver and unites the incentive potential of rankings with a financial reward.

We suppose that the interaction between the consumers is of mean-field type, i.e., the number of consumers is infinite. This choice is motivated by our application, where the game is played across a country (for e.g. around 30 millions of households in France). Given the reward, the problem reduces to a mean-field game. Our first main result is to characterize the (unique) mean-field Nash equilibrium of this game for rewards that linearly depend on the terminal consumption (Theorem 7.2.3).

We then study the Principal-Agent relation (Stackelberg game) between the provider and the population of consumers. We introduce the bi-level problem solved by the retailer, aiming at maximizing over reward functions the profit made on the whole time period, taking into account the consumption distribution achieved at the equilibrium. Our second main result is to derive a semi-explicit formula of the optimal reward in the homogeneous setting (Theorem 7.2.4), which follows by solving of a fixed-point equation. This relies on a convex reformulation of the problem, obtained by transforming the latter into an optimization over equilibrium distributions, and by expressing the sufficient optimality conditions for the reformulated problem. We show that the unique optimal reward can be approximated by a bounded function, where the sub-optimality

[1] https://www.powernext.com/french-energy-saving-certificates
[2] https://www.ecologie.gouv.fr/dispositif-des-certificats-deconomies-denergie
[3] https://www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward
[4] https://plum.fr/cagnotte/
[5] https://www.ohmconnect.com/
[6] https://www.politico.eu/article/eu-countries-save-energy-winter/

of the latter is controlled and converges to zero for sufficiently large bounds (Corollary 7.2.2). In the general setting (heterogeneous population), we show that under uniform price elasticity and uniform relative volatility, the problem reduces to the previous case (Proposition 7.2.2). For the more general case, the reformulation in the distributions space does not apply, and we introduce a numerical algorithm (Algorithm 10) to optimize the shape of the reward. This black-box optimization procedure relies on a fast evaluation of the retailer objective function at each iteration, which is done by exploiting the characterization of the mean-field Nash equilibrium.

We then apply our approach to the French market of Energy Saving Certificates using realistic data (Section 7.4). We show that the numerical procedure exhibits a fast convergence, and successfully finds the optimal reward in the homogeneous setting, and provides significant consumption reduction in the general setting, while maintaining the satisfaction (utility) of the consumers. We also simulate some trajectories of consumers by using the reward found in the mean-field context to highlight the energy reduction capacity of this mechanism. In particular, we show that the ranking system allows to achieve the sobriety target imposed by the European commission.

Finally, we consider several extensions suitable to our context. First, we show that, for the class of reward functions considered here, the addition of common-noise in the consumption process only shifts the equilibrium distribution by a (random) constant. Besides, we focus on time-dependent costs of effort for the agents, reflecting the collective awareness of agents on the energy reduction's necessity. We are also able to provide some invariance results, which show that the use of more sophisticated reward (a function that jointly depends on the rank and the consumption of the agent) is, at the equilibrium, equivalent to a reward that belongs to the class of purely rank-based rewards.

7.1.3 Related Works

Given the reward function provided by the retailer, the competition between agents is modeled by a mean-field game. These games have been introduced simultaneously by Lasry and Lions [LL06a; LL06b; LL07] and Huang, Caines and Malhame [HMC06; HCM07]. They refer to the study of differential games involving a large number of indistinguishable agents which interact through their empirical distribution. By looking at the limit case where a continuum of agents is involved, each of them asymptotically negligible, mean-field games provide efficient ways to compute approximations of Nash equilibria for stochastic games with large number of players (games which are otherwise rarely tractable). Among various techniques, the problem is often solved by a fixed-point method involving both a Hamilton-Jacobi-Bellman equation – characterizing the agents best response to a given population distribution – and a Fokker-Planck equation. Existence and uniqueness of a mean-field equilibrium are then analyzed through this system of coupled partial differential equations, see e.g. [Car+15; BCS17].

The design of a reward/incentive by the retailer is then modeled as a *Principal-Agent* problem, see e.g. the works of Sannikov [San08] and Capponi, Cvitanić and Yolcu [CCY12] in continuous-time settings. In such problems, the Principal (retailer) aims at designing a monetary reward that is offered to the agent, depending on the quantity of work achieved by the latter. In energy management, Aïd, Possamaï and Touzi introduced an incentive mechanism to control both the average consumption and the volatility of the agents consumption. The additional difficulty in our context is the presence of a continuum of agents, and the interaction between them which is expressed in terms of a mean-field game. Such extensions of the Principal-Agent problem have been considered by Elie, Mastrolia and Possamaï [EMP19] – where an explicit contract has been found for a specific class of dynamics (encompassing the linear-quadratic setting) – and by Carmona and Wang [CW21] – focusing on the linear-quadratic setting and finite-state spaces. Shrivats, Firoozi and Jaimungal [SFJ21] introduce a Principal-Agent formulation to study the interaction between a regulator and a field of providers in the market of Renewable Energy Certificate (REC).

Our study is inspired by several works. We focus on rank-based interactions, previously introduced in [BZ16], where results of existence and uniqueness of the mean-field Nash equilibrium are provided for a general class of rewards. Extensions to principal-agent problem are then studied in [BCZ19; BZ21], deriving explicit expressions of optimal contract for several principal's objectives (profit/effort/rank-performance maximization/distribution target). In comparison to these works, we provide new theoretical results for non purely rank-based reward in the case of a homogeneous population and general convex cost functions, and extends the latter to a sub-class of heterogeneous population, while keeping explicit characterizations of equilibria and optimal rewards. Finding such explicit expressions is rare in the literature, and is only possible by imposing a specific dynamics (as in [EMP19] and [CW21]). Another additional difficulty which arises from the application is to take into account the diversity of the agents: here, we consider that the overall population is clustered into a finite number of (infinite-size) independent sub-populations. This heterogeneous context (in absence of uniform elasticity) increases further the difficulty – both on analytic and numerical aspects – but is necessary on the application side for realism purposes, see e.g. [SFJ21; SFJ22] for applications of mean-field games to REC markets. In [Cam+21], Campbell et al. introduce deep learning algorithms to solve principal-agent mean field games under heterogeneity of agent types. Here, we propose an alternative method, which takes advantage of the specific structure of the problem (explicit solution of the underlying mean-field game and common rank-based reward across the sub-populations) to lower the numerical complexity and derive efficient computational methods.

The rest of the chapter is organized as follows: in Section 7.2, we first define the model and characterize the equilibrium for the mean-field game between the agents. In Section 7.3, we propose a numerical approach to solve the problem in the heterogeneous setting, for which the convex reformulation seems not extendable. In Section 7.4, we apply the results to the French market of Energy Savings Certificates, and finally in Section 7.5, we tackle some extensions that naturally arise in the context of the application. *The proofs of the main results are given in the appendix.*

7.2 Model

7.2.1 Notation and Assumptions

In the sequel, we denote by $\mathcal{P}(\mathbb{R})$ the set of distributions defined on $\mathbb{R}$ and by $\mathcal{P}^+(\mathbb{R})$ the set of distributions having strictly positive density. Moreover, for any $\mu \in \mathcal{P}(\mathbb{R})$, F_μ refers to the cumulative distribution function (cdf) of μ, and when it exits, f_μ (resp. q_μ) refers to the probability density function (pdf) (resp. the quantile function) of μ. Moreover, we write $X \sim \mu$ when X is distributed according to $\mu \in \mathcal{P}(\mathbb{R})$. The normal distribution centered in m with standard deviation σ is denoted by $\mathcal{N}(m, \sigma)$ and its pdf is denoted by $x \mapsto \varphi(x; m, \sigma)$.

Let us successively introduce the different players involved in the Stackelberg game:

Consumers. We consider a *heterogeneous* population of consumers, and we suppose that a clustering algorithm can be applied as a preprocessing step in order to split the population into K sub-populations (or clusters), each of them composed of similar customers. Each cluster $k \in [K] := \{1, \ldots, K\}$ represents a proportion ρ_k of the overall population and corresponds to a given class of customers, categorized for example according to their usages, their heating system or the household composition. Here, we directly tackle mean-field interactions between the agents:

Assumption 7.2.1. We assume that each sub-population is composed of an infinite number of indistinguishable agents, represented by a single consumer (*representative* agent).

Energy consumption. Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space, which supports a family of K independent Brownian motions $\{W_k\}_{1 \leqslant k \leqslant K}$, and $\mathbb{A}$ be the set of progressively measurable processes a satisfying the integrability condition $\mathbb{E} \int_0^T |a(s)|ds < \infty$. For a given control $a_k \in \mathbb{A}$, we denote by $X_k^a(t)$ the forecasted energy consumption of an agent from cluster k (typically an household), forecast made at time $t < T$ for the time period $[0, T]$. The controlled process X_k^a is described through the following stochastic differential equation:

$$\begin{cases} dX_k^a(t) = a_k(t)dt + \sigma_k dW_k(t), \\ X_k(0) = x_k^{\text{nom}}. \end{cases} \tag{7.1}$$

Here, we consider an arithmetic Brownian motion in the dynamics, expressing the uncertainty in the electricity needs. The use of such arithmetic noise (specific to Ornstein-Uhlenbeck processes) has been showed to be relevant for load modeling, see e.g. [RSM16]. Aïd, Possamaï and Touzi considered in [APT22] the multidimensional version of this dynamics, to the same purpose of representing the electricity consumption. The process a_k in Equation (7.1) is then viewed as the consumer's effort to reduce his electricity consumption. Without any effort, customers are expected to have a *nominal* consumption of x_k^{nom}, and we define by f_k^{nom} the p.d.f. of $X_k^a(T)$ under a zero effort (a_k is a constant process equals to 0):

$$f_k^{\text{nom}}(x) := \varphi\left(x\,;\, x_k^{\text{nom}}, \sigma_k\sqrt{T}\right) \quad . \tag{7.2}$$

Note that we do not explicitly impose bounds on the process X_k – typically non-negativity assumption – but this will be naturally enforced by the cost of effort and the volatility parameter σ_k so that the probability of negative consumption will be negligible.

Retailer. In this model, an electricity provider, incentivised by a regulation agency, aims at designing a reward function based on the *terminal ranking* of the agents in order to lower the global consumption of the customers: considering that the terminal consumption of the agents in the kth population, i.e. $X_k^a(T)$, is distributed according to μ_k, the ranking r of a player consuming the quantity x, is measured by the fraction of agents consuming less than x, i.e., $r = F_\mu(x)$, where F_μ denotes the cumulative distribution function on μ (so that the worst performer/the highest consumption has rank one and the top performer has rank 0).

A reward function in our context is then a continuous real-valued function $\mathbb{R} \times [0, 1] \ni (x, r) \mapsto R(x, r)$ that depends both on the terminal consumption x and the terminal ranking r. We consider only rewards that are non-increasing in both arguments, to favor low ranks. For any $\mu \in \mathcal{P}(\mathbb{R})$, we write $R_\mu(x) = R(x, F_\mu(x))$ and when $R(x, r)$ is independent of x, we say that the reward is *purely rank-based.* In the sequel, we will consider the following decomposition assumption:

Assumption 7.2.2. Each sub-population $k \in [k]$ receives a reward R_k has the form

$$R_k\,(x, r) = B_k(r) - px \quad , \tag{7.3}$$

where $p \in \mathbb{R}$ and $B_k \in \mathcal{B}$ with $\mathcal{B}$ the set of purely rank-based (decreasing) functions. We then call R the *total reward* and its rank-dependent part B_k the *additional reward* (financial "bonus" for the consumer).

In the energy context, the second member "$-px$" represents the classic invoice of the consumer, where p is the price to consume one unit of energy (e.g. in €/kWh). Here, this simple pricing strategy can be viewed as a regulated price (as this is the case in France for example[7]). The invoice is embedded in the reward function since it acts as a natural incentive to reduce the consumption. The first member B_k is then the additional financial reward offered to consumers based on their terminal ranking.

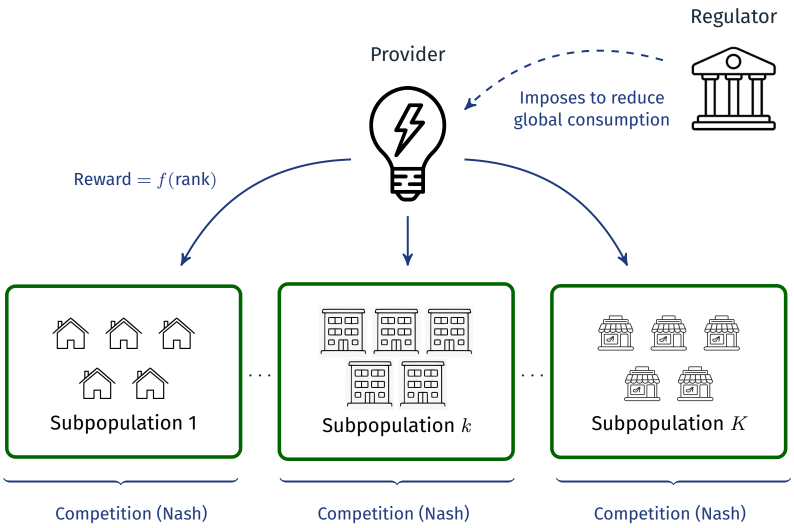


Figure 7.1: Relation between the Principal (provider) and the sub-populations composed of an infinite number of Agents (consumers).

In the modeling of energy consumption, a common-noise is often added (it can represent for example the outdoor temperature). However, we show that the insertion of such a noise only shifts the consumption distribution (by a random constant). This result was already mentioned for translation invariant functions (such as purely rank-based rewards), and we extend in Section 7.5 this property to the more general class of reward functions satisfying Assumption 7.2.2.

Assumption 7.2.3 (Fair reward mechanism)**.**

(i) Each cluster is *independent*: the rank of an agent of cluster $k \in [K]$ is only determined by the distribution of the cluster k.

(ii) The same *unitary* bonus is proposed to each cluster, i.e., $B_k(r) = x_k^{\text{nom}}\beta(r)$ for all $k \in [K]$.

Assumption 7.2.3 imposes that the sub-populations evolve separately, but are linked through a common reward function. This assumption is taken for the sake of a fair reward mechanism: on one hand, consumers only compete with similar agents, i.e., with agents having the same characteristics (type of heating, household composition, ...) and on the other hand, the shape of the reward should be identical for each the sub-population to prevent from favoring one cluster compared to another. The function β is then the unitary bonus received by every customer (in €/kWh).

Figure 7.1 outlines the Principal-Agent relation between the retailer and the field of consumers. We then first focus on the competition among the agents before studying the principal problem.

[7]"Tarif réglementé de vente" (TRV)

7.2.2 Mean-field game between agents

In all this section, let us fix a cluster $k \in [K]$, as there is no interaction between clusters. We suppose here that the reward $R_k(x, r)$ is given.

An agent of k is able to produce an effort a_k to reduce its consumption, but has to pay as a counter-part the quadratic cost $c_k\, a_k^2(t)$ with $c_k > 0$ a given positive constant. The convexity of the effort cost is natural in the context of our application. In particular, this cost either corresponds to the purchase of new equipment that is more efficient than the older one (new heating installation, isolation, ...) or corresponds to a change in the consumption pattern (sobriety). In the latter case, the convexity illustrates that small efforts (as for e.g. switching off the light when leaving a room) are easy to make while large consumption reduction (as for e.g. reducing heating or air conditioning) are more demanding. It is also possible to consider a more general convex cost, which is in non-quadratic form, since it would still lead to a tractable agent problem. However, quadratic costs are often considered in order to obtain explicit expression of the optimum, see e.g. [APT22] in the electricity context. In exchange of the effort, the consumer receives the reward $R_k(x, r)$, depending on his rank $r = F_{\mu_k}(x)$ within the sub-population, where μ_k is the k-sub-population distribution. His objective is then:

$$V_k(R_k, \mu_k) := \sup_{a\in\mathbb{A}} \mathbb{E}\left[R_{k,\mu_k}(X_k^a(T)) - \int_0^T c_k a_k^2(t)dt\right] \quad . \tag{P^{cons}}$$

The quantity $V_k(R_k, \mu_k)$ represents the *optimal expected utility* of an agent of class k, for a given provider's reward and population distribution.

We present below some results which will be used throughout the chapter. The first theorem gives the explicit solution of the agent's best response to a population distribution $\tilde{\mu}_k$:

Theorem 7.2.1 (Characterization of the best response)

Given a bounded total reward function R_k satisfying Assumption 7.2.2 and $\tilde{\mu}_k \in \mathcal{P}(\mathbb{R})$, let

$$\gamma_k(\tilde{\mu}) = \int_{\mathbb{R}} f_k^{\text{nom}}(x) \exp\left(\frac{R_{k,\tilde{\mu}}(x)}{2c_k\sigma_k^2}\right) dx \quad (< \infty) \quad . \tag{7.4}$$

Then, the optimal terminal distribution μ_k^* of a player from cluster k admits a pdf defined as

$$f_{\mu_k^*}(x) = \frac{1}{\gamma(\tilde{\mu}_k)} f_k^{\text{nom}}(x) \exp\left(\frac{R_{k,\tilde{\mu}_k}(x)}{2c_k\sigma_k^2}\right) \quad , \tag{7.5}$$

and the optimal value is then $V_k(R_k, \tilde{\mu}_k) = 2c_k\sigma_k^2 \ln \gamma_k(\tilde{\mu}_k)$.

The above result corresponds to [BZ21, Proposition 2.1] and is obtained using the Schrödinger bridge approach, see [CGP15] for connections with optimal transport theory. The consumption process X_k under the optimal effort then satisfies the equation

$$dX_k(t) = a_k(t, X_k(t); \mu_k^*)dt + \sigma_k dW_k(t),$$

where the optimal effort $a_k(\cdot, \cdot; \mu_k^*)$ is defined as $a_k(t, x, \mu) = \sigma_k^2 \partial_x \ln u_k(t, x, \mu)$ where

$$u_k(t, x, \mu) = \mathbb{E}\left[\exp\left(\frac{1}{2c_k\sigma_k^2} R_{k,\mu}(x + \sigma_k\sqrt{T-t}Z\right)\right], \quad Z \sim \mathcal{N}(0, 1) \quad . \tag{7.6}$$

We now introduce the notion of mean-field Nash equilibrium.

Definition 7.2.1 (Mean-field Nash equilibrium)**.** We say that $\mu_k \in \mathcal{P}(\mathbb{R})$ is an *equilibrium* (terminal distribution) if it is a fixed-point of the mapping $\Phi_k : \tilde{\mu}_k \mapsto \mu_k^*$, with μ_k^* given by the solution of the equation (7.5).

The existence of such an equilibrium has been proved in the general setting using Schauder's fixed point theorem (see [BZ16]). We give below a characterization of this equilibrium distribution, as well as an explicit expression for purely rank-based rewards:

Theorem 7.2.2 (Characterization of the equilibrium distribution)

Given a bounded total reward function $R_k : \mathbb{R} \times [0,1] \to \mathbb{R}$, the distribution $\mu_k \in \mathcal{P}(\mathbb{R})$ is an equilibrium terminal distribution for cluster k if and only if its quantile function q_{μ_k} satisfies

$$N\left(\frac{q_{\mu_k}(r) - x_k^{\mathrm{nom}}}{\sigma_k\sqrt{T}}\right) = \frac{\int_0^r \exp\left(-\frac{R_{k,\mu_k}(q_{\mu_k}(z))}{2c_k\sigma_k^2}\right) dz}{\int_0^1 \exp\left(-\frac{R_{k,\mu_k}(q_{\mu_k}(z))}{2c_k\sigma_k^2}\right) dz}\ , \tag{7.7}$$

where N is the standard normal c.d.f. In the specific case of a purely rank-based reward, we obtain that the equilibrium ν_k is unique and the quantile is given by

$$q_{\nu_k}(r) = x_k^{\mathrm{nom}} + \sigma_k\sqrt{T}N^{-1}\left(\frac{\int_0^r \exp\left(-\frac{B_k(z)}{2c_k\sigma_k^2}\right) dz}{\int_0^1 \exp\left(-\frac{B_k(z)}{2c_k\sigma_k^2}\right) dz}\right)\ . \tag{7.8}$$

The mean consumption at the equilibrium is then $m_{\mu_k} = \int_0^1 q_{\mu_k}(r)dr$.

The above result is provided in [BZ21, Theorem 3.2], and below we extend the explicit characterization to the more general case of reward maps R, which not only depend on the rank, but also have a linear dependence on x.

Theorem 7.2.3 (Explicit characterization for non purely rank-based rewards)

Suppose the reward is of the form defined in Assumption 7.2.2. Then, the equilibrium μ_k is unique, and it satisfies

$$q_{\mu_k}(r) = q_{\nu_k}(r) - \frac{pT}{2c_k}\ , \tag{7.9}$$

where ν_k is the (unique) equilibrium distribution for the specific case $p = 0$ (purely rank-based reward), defined in (7.8).

Theorem 7.2.3 shows that the addition of a linear part in the consumption acts as a shift on the probability density function. We emphasize that our uniqueness result of the equilibrium μ generalizes the one established in [BZ21], the latter being obtained under the additional assumptions that the map $r \mapsto R_k(x,r)$ is convex and $r \mapsto \partial_x R_k(x,r)$ is non decreasing. Instead, we assume a linear dependence on the consumption for the reward, but no convexity requirement is made on its purely rank-based component B.

Corollary 7.2.1 (Equilibrium without additional reward)

For $R_k(x,r) = -px$, the equilibrium follows the normal distribution $\mathcal{N}\left(x_k^{\mathrm{pi}}, \sigma_k\sqrt{T}\right)$, where $x_k^{\mathrm{pi}} = x_k^{\mathrm{nom}} - \frac{pT}{2c_k}$ is the consumption under the natural incentive associated with the price

p. Moreover, the optimal consumer's utility is

$$V_k(R,\mu_k) = V_k^{\mathrm{pi}} := -px_k^{\mathrm{pi}} - \frac{p^2T}{4c_k} \ . \tag{7.10}$$

Proof. For $B_k \equiv 0$, Eq. (7.8) gives us $q_{\nu_k}(r) = x_k^{\mathrm{nom}} + \sigma_k\sqrt{T}N^{-1}(r)$, therefore $\nu_k \sim \mathcal{N}(x^{\mathrm{nom}}, \sigma_k\sqrt{T})$. We then obtain by Theorem 7.2.3 the definition of the equilibrium μ_k. Finally, using Lemma 7.7.1, we get

$$2c_k\sigma_k^2 \ln \gamma_k(\tilde{\mu}_k) = \ln\left(\int_{\mathbb{R}} f_k^{\mathrm{nom}}(x)\exp\left(\tfrac{-px}{2c_k\sigma_k^2}\right)dx\right) = -px_k^{\mathrm{nom}} + \frac{p^2T}{4c_k} \ .$$

□

Corollary 7.2.1 shows that the price of electricity constitutes a natural incentive, as the consumer already makes an effort to reduce his consumption from x^{nom} to x^{pi}. However, it induces a disutility for consumers ($V^{\mathrm{pi}} \leqslant 0$). An increase of the price would lead to a supplementary consumption reduction but would decrease further the utility of the agents, and is therefore a non-desirable energy saving strategy.

7.2.3 The Principal's problem

In this section, we suppose that Assumption 7.2.2 is satisfied. Therefore, the equilibrium distribution is unique and is defined by (7.9). For a mean-field equilibrium $(\mu_k)_{k\in[K]}$, the mean consumption of the overall population is then $m_\mu = \sum_{k\in[K]} \rho_k m_{\mu_k}$.

For a given k, we denote by ϵ_k the mapping which associates to the total reward function the corresponding equilibrium distribution, i.e. $\epsilon_k(R_k) = \mu_k$, where μ_k satisfies (7.9). The problem of the retailer can then be written as

$$\pi^* := \max_{\beta\in\mathcal{B}} \left\{ pm_\mu - \kappa(m_\mu) - \sum_{k\in[K]} \rho_k x_k^{\mathrm{nom}} \int_0^1 \beta(r)dr \;\middle|\; \begin{array}{l} R_k(x,r) = x_k^{\mathrm{nom}}\beta(r) - px \\ \mu_k = \epsilon_k(R_k) \\ V_k(R_k,\mu_k) \geqslant V_k^{\mathrm{pi}} + \tau x_k^{\mathrm{nom}} \end{array} \right\} \tag{P^{ret}}$$

where $\kappa(\cdot)$ denotes the mean selling cost function and m_μ is the mean consumption at the equilibrium μ. The optimal objective π^* then corresponds to the profit per agent (mean over the population) made on the interval $[0,T]$ (in €). The inequality constraint on the utility ensures that consumers "play the game", as it procures a strictly better utility than without additional reward. Classically, $\tau = 0$, meaning that the effort achieved by consumers in order to save energy is compensated (in mean) by the reward offered by the retailer. Observe that with $\tau = 0$, some agents may have a negative reward, which is not always desirable. Therefore, for practical issue and acceptability, we allow for a positive *tau* to take into account switching costs that appear when it comes to subscribing to a reward mechanism, see e.g. [MMV23].

Assumption 7.2.4. The function $\kappa : \mathbb{R} \to \mathbb{R}$ is increasing, convex and differentiable. Moreover, $\kappa'(0) < p < \kappa'(x^{\mathrm{pi}})$.

Assumption 7.2.4 is natural in the context of our application. In practice, the selling cost function is defined as $\kappa : m \mapsto c_p(m) + s(m)$, where

- $s(\cdot)$ denotes the penalty imposed by the regulator to favor a reduction in consumption,
- and $c_p(\cdot)$ denotes the cost function, induced by the production of energy.

We assume here that the marginal cost $\kappa'(\cdot)$ is lower than the marginal price p at 0 – meaning that it is always profitable to sell a positive quantity of energy – and conversely we assume that the marginal cost $\kappa'(\cdot)$ is greater than the marginal price p at x^{pi} – meaning that it is not profitable to sell more electricity with the additional reward than without. The penalty function s is increasing and convex, since the regulator aims at encouraging consumption reduction by strongly penalizing huge consumption levels. Moreover, the retailer's aggregated cost function is often considered as increasing and convex, due to a decreasing return to scale, see e.g. [Ale+19; ABM20]: the mechanism of day-ahead markets favors the "cheapest" (lowest marginal cost) power plants as the cheapest resource will participate to the electricity generation first, followed by the second cheapest option, and so on, until the demand is satisfied. In the case of non-convex aggregated cost, the convex hull of the aggregated cost function is often considered, see e.g. [Sch+16].

In the case of a homogeneous population and *linear dependence of the objective function with respect to the equilibrium distribution*, the results are obtained in [BZ21]. We extend them here to the more general case of *convex nonlinear dependencies.*

Homogeneous population

We consider in this section the specific case where there is a unique cluster of customers (homogeneous population). Therefore, we omit the dependence in k. Using Lemma 7.7.2 (given in the Appendix), Problem (P^{ret}) can be reformulated as a constrained maximization problem on the distribution space:

Proposition 7.2.1
Let us consider the following minimization problem

$$\begin{aligned} \min_{\mu\in\mathcal{P}^+(\mathbb{R})} \quad & \kappa\left(\int_{\mathbb{R}} y f_\mu(y)dy\right) + 2c\sigma^2 \int_{\mathbb{R}} \ln\left(\frac{f_\mu(y)}{f^{\text{nom}}(y)}\right) f_\mu(y)dy \\ \text{s.t.} \quad & \int_{\mathbb{R}} f_\mu(y)dy = 1 \\ & y \mapsto \ln\left(\frac{f_\mu(y)}{f^{\text{nom}}(y)}\right) + \frac{p}{2c\sigma^2} y \quad \text{decreasing} \end{aligned} \quad . \tag{7.11}$$

Then, the reward $B_{\mu^*} \in \mathcal{B}$, constructed from an optimal distribution $\mu^* \in \mathcal{P}^+(\mathbb{R})$ of (7.11) as

$$B_{\mu^*}(r) = V^{\text{pi}} + \tau x^{\text{nom}} + 2c\sigma^2 \ln\left(\frac{f_{\mu^*}(q_{\mu^*}(r))}{f^{\text{nom}}(q_{\mu^*}(r))}\right) + pq_{\mu^*}(r) \tag{7.12}$$

is optimal for problem (P^{ret}).

Proof. From Lemma 7.7.2, B_μ defined in (7.12) is the reward that achieves a given equilibrium distribution μ with the lowest cost while satisfying the utility condition in (P^{ret}) (since $V(R,\mu) = (1+\tau)V^{\text{pi}}$ for any attainable equilibrium μ and $R(x,r) = B_\mu(r) - px$). The objective function is then rewritten as a function of the pdf f_μ using the expression of the reward. □

We now relax (7.11) by ignoring the decreasingness of the additional reward in (7.11):

$$\min_{f:\mathbb{R}\to\mathbb{R}} \left\{ \kappa\left(\int_{\mathbb{R}} y f(y)dy\right) + 2c\sigma^2 \int_{\mathbb{R}} \ln\left(\frac{f(y)}{f^{\text{nom}}(y)}\right) f(y)dy \,\middle|\, \int_{\mathbb{R}} f(y)dy = 1 \text{ and } f(x) \geqslant 0, x\in\mathbb{R} \right\} . \tag{$\widetilde{P}^{\text{ret}}$}$$

The discussion about the relation between the initial problem (7.11) and the relaxed one ($\widetilde{P}^{\text{ret}}$) is provided further. The optimal solution of this relaxed problem is then characterized by the following lemma:

Lemma 7.2.1 (Characterization of the optimal distribution for the relaxed problem)
Let Assumption 7.2.4 holds. Then, ($\widetilde{P}^{\text{ret}}$) defines a convex problem. Moreover, if μ^* admits a density f_{μ^*} which minimizes ($\widetilde{P}^{\text{ret}}$), then it satisfies the following optimality conditions: for μ^*-almost every x in $\mathbb{R}$,

$$f_{\mu^*}(x) = \frac{1}{\alpha(\mu^*)} f^{\text{nom}}(x) \exp\left(-x\frac{\kappa'(m_{\mu^*})}{2c\sigma^2}\right) \tag{7.13}$$

where

$$\alpha(\mu) = \int_{\mathbb{R}} f^{\text{nom}}(y) \exp\left(-y\frac{\kappa'(m_{\mu})}{2c\sigma^2}\right) dy \ .$$

Proof. The convexity of the objective functional with respect to f comes from the convexity of κ (see Assumption 7.2.4) and the convexity of $x \mapsto x\ln(x)$. The first-order conditions for ($\widetilde{P}^{\text{ret}}$) are detailed in Section 7.7. Furthermore, they are sufficient for this convex problem, see e.g. [LBD22, Theorem 3.3]. □

In contrast with [BZ21], the optimal distribution is not explicit anymore due to the general function $\kappa(\cdot)$. Instead, the optimal distribution is implicitly known through the fixed-point equation (7.13). We simplify this condition in the following theorem to end up with one-dimension fixed-point equation on the mean consumption.

Theorem 7.2.4
Let Assumption 7.2.4 holds, and let $\delta : \mathbb{R} \to \mathbb{R}$ be a function given by

$$\delta(m) = p - \kappa'(m) \ .$$

Then, the distribution $\mu^* = \mathcal{N}(m^*, \sigma\sqrt{T})$, where m^* satisfies the fixed-point equation

$$m - x^{\text{pi}} = \frac{T}{2c}\delta(m) \ , \tag{7.14}$$

is optimal for the problem ($\widetilde{P}^{\text{ret}}$). Moreover, the associated reward B_{μ^*} is

$$B_{\mu^*}(r) = \tau x^{\text{nom}} + \frac{c}{T}\left[(x^{\text{pi}})^2 - (m^*)^2\right] + q_{\mu^*}(r)\delta(m^*) \ , \tag{7.15}$$

and the associated retailer gain is

$$\pi^* = m^*\kappa'(m^*) - \kappa(m^*) + \left(\frac{m^* + x^{\text{pi}}}{2}\right)\delta(m^*) - \tau x^{\text{nom}} \ . \tag{7.16}$$

Corollary 7.2.2
Let Assumption 7.2.4 holds. Then, the fixed-point equation (7.14) admits a unique solution $m^* \in]0, x^{\text{pi}}]$. Moreover, the (unique) reward function is decreasing.

Proof. The increasingness of $\kappa'(\cdot)$ suffices to ensure that (7.14) admits a unique solution. Moreover, as $\delta(0) \geqslant 0 \geqslant \delta(x^{\mathrm{pi}})$, the root of the equation $\frac{T}{2c}\delta(m) - m + x^{\mathrm{pi}} = 0$ must belong to $]0, x^{\mathrm{pi}}]$. As a consequence, $\delta(m^*) \leqslant 0$ and the reward function B_{μ^*} is decreasing. □

The existence and uniqueness of a solution for (7.14) entails the existence and uniqueness of an optimal reward for $(\widetilde{P}^{\mathrm{ret}})$. The knowledge of the bounds for m^* along with the decreasingness of $\delta(\cdot)$ allows to use for instance a binary search algorithm to numerically find the optimal mean consumption in logarithmic time.

Remark 7.2.1

For quadratic function $s : m \mapsto \alpha_2 m^2 + \alpha_1 m + \alpha_0$, the fixed point of (7.14) is analytically known:

$$m^* = \left(1 + \frac{\alpha_2 T}{c}\right)^{-1} \left(x^{\mathrm{nom}} - \frac{(\alpha_1 + c_p)T}{2c}\right) .$$

The function $\delta(\cdot)$ is here interpreted as the *reduction desire* of the provider, as consumption reduction $x^{\mathrm{pi}} - m^*$ is proportional to $|\delta(m^*)|$, see (7.14). It expresses the marginal benefit coming from selling electricity (including the penalty function s provided by the regulator).

In the relaxed problem, we neglect that the reward is decreasing. However, this is directly ensured by Corollary 7.2.2: the reward provided in Theorem 7.2.4 is decreasing if and only if δ is negative at the optimum. Therefore, it is also optimal for the original retailer problem (P^{ret}).

The optimal reward obtained in Eq. (7.15) is defined through the quantile of μ^* and is therefore unbounded. From the application viewpoint (it is not realistic to give unbounded rewards to consumers) and for numerical issues, we now look at truncated reward. To this purpose, let us define for any $M > 0$ the truncated optimal equilibrium distribution μ_M through its p.d.f:

$$f_{\mu_M}(x) \propto h_M(x) := f^{\mathrm{nom}}(x) \exp\left(\frac{-y\kappa'(m_{\mu^*}) \wedge M \vee (-M)}{2c\sigma^2}\right) . \tag{7.17}$$

Theorem 7.2.5 (Bounded reward)

The total reward which leads to equilibrium μ_M and gives to agents the utility $V^{\mathrm{pi}} + \tau x^{\mathrm{nom}}$ is bounded for every consumption level and is defined as

$$\forall x \in \mathbb{R},\ R_{\mu_M}(x) = V^{\mathrm{pi}} + \tau x^{\mathrm{nom}} - 2c\sigma^2 \ln \int_{\mathbb{R}} h_M(y)dy + x\kappa'(m_{\mu^*}) \wedge M \vee (-M) . \tag{7.18}$$

Moreover, the mean consumption converges to the optimal one :

$$m_{\mu_M} = m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right) .$$

Proof. From Lemma A.2., the total reward associated to μ_M is

$$R_{\mu_M} = V^{\mathrm{pi}} + \tau x^{\mathrm{nom}} + 2c\sigma^2 \ln\left(f_{\mu_M}(y)/f^{\mathrm{nom}}(y)\right)$$

and satisfies the utility constraint by construction. The result is then obtained using the definition of f_{μ_M}. Besides, one can show (see [BZ21, Theorem 5.4]) that $\int_{\mathbb{R}} h_M dx = \alpha(\mu^*) + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$ and $\int_{\mathbb{R}} x h_M(x)dx = \alpha(\mu^*)m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$. As a consequence, $m_{\mu_M} = m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$. □

As the optimal (unbounded) total reward, its truncated analog obtained in (7.18) is linear in the terminal consumption (inside the bounds $[-M, M]$). This means that the consumers are rewarded proportionally to their consumption reduction. Moreover, for both the theoretical bonus (7.15) and the bounded one (7.18), τ only acts as a shift on the function in order to uplift or lower the bonus received by each agent. Consequently, it is possible to a posteriori choose τ in such a way that the bonus of a given ranking corresponds to a certain amount.

Heterogeneous population

We consider here the more general setting of a heterogeneous population, not studied yet in the ranking games literature, which consists in a finite number of clusters $K > 1$. The transformation which leads to (7.11) still applies, but the additional constraint Assumption 7.2.3 has to be imposed to ensure the unitary reward is identical for every sub-population[8].

As it will be seen below, we can recover explicitly solvable problems for a subclass of heterogeneous populations for which all agents of the overall population are similar up to a scaling factor.

Proposition 7.2.2 (Explicit characterization for a sub-class of heterogeneous population) Let suppose that the following statement holds:

$$\forall k \in [K], \quad \frac{x_k^{\mathrm{nom}}}{x_1^{\mathrm{nom}}} = \frac{\sigma_k}{\sigma_1} = \frac{c_1}{c_k} \quad (:= \theta_k) \ . \tag{7.19}$$

Then, any $\mu_1, \ldots, \mu_K$ equilibrium distributions associated to a common unitary reward β solution of (P^{ret}) satisfies $f_{\mu_k}(y) = \frac{1}{\theta_k} f_{\mu_1}\left(\frac{y}{\theta_k}\right)$ for all $k \in [K]$. Moreover, the retailer's profit problem simplifies to

$$\pi^* := \bar{\theta} \max_{\beta \in \mathcal{B}} \left\{ pm_{\mu_1} - \tilde{\kappa}(m_{\mu_1}) - x_1^{\mathrm{nom}} \int_0^1 \beta(r)dr \;\middle|\; \begin{array}{l} R_1(x,r) = x_1^{\mathrm{nom}}\beta(r) - px \\ \mu_1 = \epsilon_1(R_1) \\ V_1(R_1, \mu_1) \geqslant V_1^{\mathrm{pi}} + \tau x_1^{\mathrm{nom}} \end{array} \right\} \ , \tag{7.20}$$

with $\tilde{\kappa}(m) = \bar{\theta}^{-1}\kappa(\bar{\theta}m)$ and $\bar{\theta} = \sum_{k\in[K]} \rho_k\theta_k$.

Proof. Using the characterization of the equilibrium in (7.9), $q_{\mu_k}(r) = \theta_k q_{\mu_1}(r)$. Therefore, $F_{\mu_k}(y) = F_{\mu_1}\left(\frac{y}{\theta_k}\right)$ and $f_{\mu_k}(y) = \frac{1}{\theta_k} f_{\mu_1}\left(\frac{y}{\theta_k}\right)$. Moreover,

$$\begin{aligned} \gamma(\mu_k) &= \int_{\mathbb{R}} f_k^{\mathrm{nom}}(x) \exp\left(\frac{x_k^{\mathrm{nom}}\beta(F_{\mu_k}(x)) - px}{2c_k\sigma_k^2}\right) dx \\ &= \int_{\mathbb{R}} \frac{1}{\theta_k} f_1^{\mathrm{nom}}\left(\frac{x}{\theta_k}\right) \exp\left(\frac{x_1^{\mathrm{nom}}\beta\left(F_{\mu_1}\left(\frac{x}{\theta_k}\right)\right) - p\frac{x}{\theta_k}}{2c_1\sigma_1^2}\right) dx = \gamma(\mu_1) \ . \end{aligned}$$

Therefore, $V_k(R_k, \mu_k) = \theta_k V_1(R_1, \mu_1)$. As $V_k^{\mathrm{pi}} = \theta_k V_1^{\mathrm{pi}}$, the utility constraint is satisfied for every sub-population. □

Proposition 7.2.2 shows that in this specific case of heterogeneous population, the problem boils down to the homogeneous framework, up to a re-scaling of the cost function κ. Therefore, Theorems 7.2.4 and 7.2.5 and corollary 7.2.2 still apply, and in particular, the optimal distribution is

[8] Using Lemma 7.7.2, there exists a common unitary reward leading to equilibrium $\mu_1, \ldots, \mu_K$ if and only if there exists for all $k \in [K]$ a constant C_k such that $\frac{c_k\sigma_1^2}{x_k^{\mathrm{nom}}} \ln\left(\frac{f_{\mu_k}(x)}{f_k^{\mathrm{nom}}(x)}\right) = \frac{c_1\sigma_1^2}{x_1^{\mathrm{nom}}} \ln\left(\frac{f_{\mu_1}(x)}{f_1^{\mathrm{nom}}(x)}\right) + C_k$ for all $x \in \mathbb{R}$.

$\mu_1^* = \mathcal{N}(m_1^*, \sigma_1\sqrt{T})$ where μ_1^* is uniquely determined by the equation $m_1^* - x_1^{\text{pi}} = \frac{T}{2c_1}(p - \tilde{\kappa}'(m_1^*))$. The condition (7.19) corresponds to the case where (i) the volatility of the noise is proportional to the nominal consumption and where (ii) the price elasticity is identical for all sub-populations (see Section 7.4.1 and (7.25) for the link between the cost of effort c_k and the elasticity). The second statement (ii) may be more debatable, as the elasticity of a consumer intuitively depends on the equipment of the housing (for instance the type of heating).

7.3 Numerical resolution in the non-uniform heterogeneous case

To deal with the general case of a heterogeneous population, we develop a numerical algorithm to compute the optimal reward from the original problem P^{ret}. For a given $N \in \mathbb{N}$, we denote by Σ_N the uniform discretization of the interval $[0,1]$ by N points, such that $\Sigma_N := \{0 = \eta_1 < \eta_2 < \ldots < \eta_N = 1\}$. Let $M \in \mathbb{R}_+$, then we define the class of bounded piecewise linear rewards adapted to Σ_N as

$$\widehat{\mathcal{B}}_M^N := \left\{ r \in [0,1] \mapsto \sum_{i=1}^{N-1} \mathbb{1}_{r\in[\eta_i,\eta_{i+1}[} \left[b_i + \frac{b_{i+1}-b_i}{\eta_{i+1}-\eta_i}(r-\eta_i) \right] \;\middle|\; \begin{array}{l} b \in [-M,M]^N \\ b_1 \geqslant \ldots \geqslant b_N \end{array} \right\} .$$

The reward function obtained as a linear interpolation of a non-increasing vector b is denoted by $\hat{\beta}[b]$. For this special class of reward, the computation of some integrals can be simplified. The integral that appears in the equilibrium characterization (7.8) becomes

$$\begin{aligned} &\int_0^1 \exp\left(-\frac{x_k^{\text{nom}}\hat{\beta}[b](r)}{2c_k\sigma_k^2}\right) dr \\ &= 2c_k\sigma_k^2(x_k^{\text{nom}})^{-1} \sum_{i=1}^{N-1} \frac{\eta_{i+1}-\eta_i}{b_{i+1}-b_i} \left[\exp\left(-\frac{x_k^{\text{nom}} b_{i+1}}{2c_k\sigma_k^2}\right) - \exp\left(-\frac{x_k^{\text{nom}} b_i}{2c_k\sigma_k^2}\right) \right] \end{aligned}$$

and the integral of the bonus simplifies into

$$\int_0^1 \hat{\beta}[b](r)dr = \sum_{i=1}^{N-1} (\eta_{i+1}-\eta_i)\left(\frac{b_{i+1}+b_i}{2}\right) .$$

Box maximization. We define the following transformation:

$$\begin{array}{llll} \phi_M^N : & [\text{-}1,1]^N & \to [-M,M]^N \\ & z & \mapsto \quad b \end{array} \quad \text{where} \quad \begin{cases} b_1 = M z_1 \\ b_i = \frac{1}{2}(b_{i-1}-M) + \frac{1}{2}(b_{i-1}+M)z_i, \; i > 1 \end{cases} . \tag{7.21}$$

For any $M \in \mathbb{R}_+$ and $N \in \mathbb{N}$, the function ϕ_M^N is invertible and $\left(\phi_M^N\right)^{\text{-}1}$ is defined as:

$$\left(\phi_M^N\right)^{\text{-}1}(b) = \begin{cases} z_1 = \dfrac{1}{M} b_1 \\ z_i = \dfrac{2b_i - b_{i-1} + M}{b_{i-1}+M}, \; i > 1 \end{cases}$$

As an example, Figure 7.2 displays $(\eta_i, z_i)_{i\in[N]}$ and the corresponding bonus function $\hat{\beta}[\phi_M^N(z)]$.

We denote by $\pi_\lambda : \mathcal{B} \to \mathbb{R}$ the Lagrangian function of (P^{ret}), defined as

$$\pi_\lambda(\beta) := \left\{ \begin{array}{c|l} pm_\mu - \kappa(m_\mu) - \displaystyle\sum_{k\in[K]} \rho_k x_k^{\text{nom}} \int_0^1 \beta(r)dr & R_k(x,r) = x_k^{\text{nom}}\beta(r) - px \\ -\lambda \displaystyle\sum_{k\in[K]} \rho_k \left(V_k^{\text{pi}} + \tau x_k^{\text{nom}} - V_k(R_k,\mu_k)\right)^+ & \mu_k = \epsilon_k(R_k) \end{array} \right\} , \tag{7.22}$$

where $(\cdot)^+ := \max(0, \cdot)$. For fixed Lagrangian multiplier $\lambda > 0$, π_λ constitutes a relaxed version of the initial problem (P^{ret}), where violations of the utility condition are not fully forbidden but rather strongly penalized in the objective for large values of λ.

Proposition 7.3.1 (*Maximization with box constraints*)

$$\max_{z \in [-1,1]^N} \pi_\lambda(\widehat{\beta}[\phi_M^N(z)]) = \max_{\beta \in \mathcal{B}_M^N} \pi_\lambda(\beta) \ . \tag{7.23}$$

Proof. By definition of $\mathcal{B}_M^N$, $\max_{z \in [-1,1]^N} \pi_\lambda(\widehat{\beta}[\phi_M^N(z)]) \leqslant \max_{\beta \in \mathcal{B}_M^N} \pi_\lambda(\beta)$. As the map ϕ_M^N is invertible, for any reward $\beta \in \widehat{\mathcal{B}}_M^N$, there exists $z \in [-1,1]^N$ such that $\beta = \hat{\beta}[\phi_M^N(z)]$, hence the reverse inequality. Optimizing on $\widehat{\mathcal{B}}_M^N$ is then equivalent to optimize on $[-1,1]^N$ via the transformation ϕ_M^N. $\square$

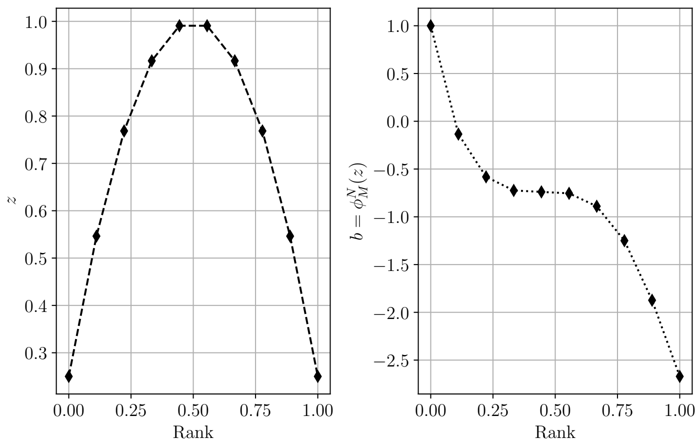


Figure 7.2: Example of transformation using function ϕ_M^N for $M = 4$ and $N = 10$

Algorithm 10 aims at maximizing the function π_λ. To this end, we do not directly search the optimal reward but, as described previously, we use the invertible map ϕ_N^M to search in the space $[-1,1]^N$, see Proposition 7.3.1. From a computational viewpoint, the search space is now independent of M, and the decreasingness of the bonus function is directly encoded in the transformation. Therefore, the only remaining constraints are the ones ensuring that the solution belongs to the unit box. The search is then achieved by black-box optimization, since the evaluation of π_λ can be explicitly done using (7.8)-(7.9)). In the numerical results, we use CMA-ES ([Han06]) as optimization solver through the C++ interface ([Fab13]). Convergence properties are analyzed in [HO97], and we display in Section 7.4 the numerical convergence of the objective along the iterations.

Remark 7.3.1 (i) The evaluation of π_λ linearly depends on the number of sub-populations (i.e., K) since, given a reward, the problem boils down to the computation of the equilibrium distributions for the K sub-populations.

(ii) The reward function found by Algorithm 10 is bounded and decreasing, but might violate the utility constraint "$V_k(R, \mu_k) \geqslant V_k^{\text{pi}} + \tau x_k^{\text{nom}}$" for small penalization values of λ. Note that if the optimizer for the discrete problem on a sufficiently precise grid is a global optimizer, then we get an ε-solution of the initial problem, see Theorem 7.2.5.

Algorithm 10 OPTIMIZATION OF THE REWARD

Require: M, N, λ, Σ_N, solver Π, initial point z^0,

Construct Θ as

$$\Theta : z \in]\text{-}1, 1]^N \mapsto \pi_\lambda \left(\hat{\beta}[\phi_M^N(z)]\right) \tag{7.24}$$

Apply Π to maximize Θ (starting from z^0) and get the final state z^Π.

return $\beta^\Pi = \hat{\beta}[\phi_M^N(z^\Pi)]$.

7.4 Application to Energy Savings

In this section, we develop a case study related to the French market of Energy Saving Certificates based on the use of realistic data. We compare the results with existing reward mechanisms, and analyze them in terms of consumption reduction (relatively to the target imposed by the European commission).

7.4.1 Instances

Consumers. We consider the case where the retailer aims at designing a reward for 4 types of consumers, listed in Table 7.1. Data on the average annual consumption correspond to the French case. The consumers are here distinguished according to the surface of the housing

	Distribution	Housing	Heating	Nb occupants	Consumption (mean/year)
Sub-pop. 1	26%	House 70 m^2	Electric	3	9.9 MWh
Sub-pop. 2	49%	House 70 m^2	Non-electric	3	1.5 MWh
Sub-pop. 3	9%	House 150 m^2	Electric	4	20 MWh
Sub-pop. 4	16%	House 150 m^2	Non-electric	4	2.2 MWh

Table 7.1: Annual electricity consumption by type of usage.
The consumption data are extracted from "Agence France Electricité"[9].

and the type of heating, which can represent up to 90% of the annual consumption. A more elaborated clustering might also take into account the location of the housing or the age of the occupants, but we focus here on the two main factors affecting the consumption. We suppose for simplicity that the overall population is composed of these four sub-populations, representing a total of 33 millions of households (current number of households in France). The distribution of the sub-populations is then computed by considering that there are thrice as many 70m^2-houses as 150m^2-houses (the mean surface in France[10] in around 90m^2) and that a 35%[11] of the French households is equipped with electric heating. This gives us a mean annual consumption of 5.46MWh, or a total annual consumption of 180TWh. In comparison, the French annual consumption for residential households is around 155TWh. This slight over-estimation is due to the fact that we only consider here houses with three or four occupants.

We suppose that the consumption levels displayed in Table 7.1 corresponds to customers having subscribed to a regulated offer, corresponding to a fixed price of electricity p. As showed

[10] https://www.lamaisonsaintgobain.fr/blog/insolites/metre-carre-et-confort-connaissez-vous-la-moyenne-francaise

[11] https://www.voltalis.com/comprendre-electricite/les-types-de-chauffage-preferes-des-foyers-francais-1772

[11] https://www.agence-france-electricite.fr/consommation-electrique/moyenne-par-jour/

in Corollary 7.2.1, nominal consumption (x^{nom}) and consumption under price p (x^{pi}) are linked by the relation $x^{\text{pi}} = x^{\text{nom}} - \frac{p}{2c}$ (we consider annual consumption in Table 7.1, i.e., $T = 1$).

In [NYK20], the authors used several utility concave utility function to model the price elasticity of the electricity demand. In particular, they studied a quadratic utility function similar to the cost of effort we consider: with $T = 1$ and constant effort, $V_k^{\text{pi}} = \max_{x\in\mathbb{R}}\{-px - c(x - x^{\text{nom}})^2\}$. This corresponds to the welfare maximization with quadratic utility, defined as $U(x, x^{\text{nom}}) = -c(x - x^{\text{nom}})^2$. For this type of utility function, the elasticity is defined as $\eta = 1 - \frac{x^{\text{nom}}}{x^{\text{pi}}}$, see e.g. [NYK20, Eq. 19]). As a consequence, using the relation between x^{pi} and x^{nom} and the definition of the elasticity, one can obtain the following relations:

$$c = \frac{-p}{2\eta x^{\text{pi}}} \ , \qquad x^{\text{nom}} = x^{\text{pi}}(1 - \eta) \ . \tag{7.25}$$

Several values of price elasticity are reported in [NYK20; Cse20], and we use here $\eta = -0.32$, which corresponds to the estimation of the long-run residential price elasticity made by [Bön+15] on the EPEX spot market between 2012 and 2014. Price elasticity is always studied at the scale of a country (or even broader), and therefore we take an estimate which is identical for all the agents (*uniform* elasticity). In the numerical results, we will analyze the influence of a non-uniform elasticity, see Section 7.4.

Regarding the volatility, in the Low Carbon London pricing study, Carmichael et al. [Car+14] reported a deviation of ±200 Watt for a demand of 1000 Watt. We take here a deviation $\sigma\sqrt{T}$ equals to 10% of the total consumption X_T under zero effort for each of the four sub-populations. Finally, we consider here for p the price of the regulated offer ("Tarif Bleu") in 2019, that is 145 €/MWh[12].

	c_k (€/MWh^2)	σ_k (MWh)
Sub-pop. 1	24	0.57
Sub-pop. 2	156	0.09
Sub-pop. 3	12	4.15
Sub-pop. 4	107	0.13

Table 7.2: Cost of effort and volatility parameters.

Retailer cost. We consider here the year 2019 (just before the energy crisis). We display in Table 7.3 the marginal cost and the annual production for each type of power plants.

Power plant	Marginal cost (€/MWh)	Production (TWh)
Hydro/Wind/Solar	0 to 15	115
Nuclear	30	380
Gas	70	30
Coal	86	7
Fuel	162	5

Table 7.3: Marginal price and annual production. Source: *RTE Bilan électrique 2019* and *Ademe*

[12] `https://prix-elec.com/tarifs/evolution/2019`

By aggregating the production capacities by increasing cost (as in merit order curves for day-ahead markets), we can obtain an estimate of the supply cost according the production, see Figure 7.3. The total cost is then obtained by dividing the supply cost by 0.35 as this approximately corresponds to the weight of supply in the total cost[13]. To fit with our situation where we only look at the residential part of the consumption, we shift the cost curve so that a residential consumption of 180TWh is “cleared” by a gas power plant (as it is often the case in the day-ahead market) and we regularize it to be differentiable.

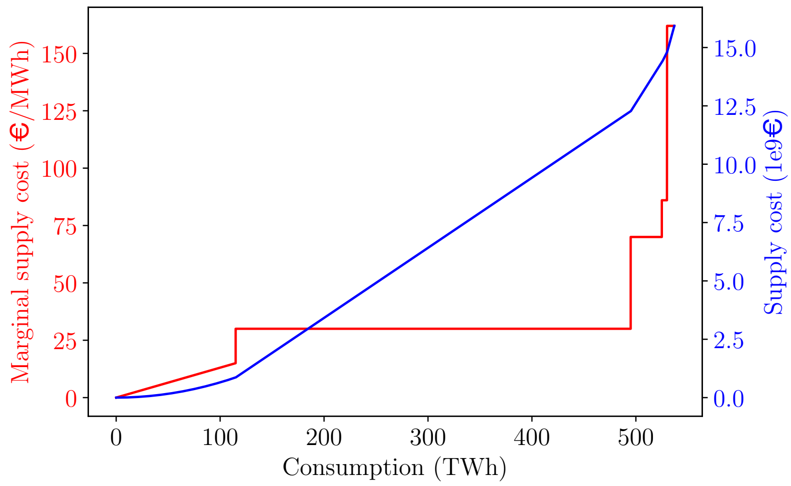


Figure 7.3: Estimation of supply cost through marginal costs

Valuation of energy savings. Electricity retailers are obliged by the French government[14] to reduce the global consumption of their customers, in the context of energy efficiency and sobriety. From 2024 to 2030, the European regulation will impose a reduction target of 1.49% of the annual consumption, and aspire to reach 1.9% by the end of 2030. If a retailer does not succeed in gathering a sufficient amount of Energy Saving Certificates, a penalty of 15€/MWh is applied (for “classic” certificates)[15]. In addition, each provider can buy (resp. sell) on a market a certain quantity of certificates if the quantity of energy consumption overshoots (resp. undershoots) the target. In 2023, the price of certificates is around 7.5€/MWh[16]. We consider here a target of 5% of consumption reduction over 3 years ($T = 3$), corresponding to a mean consumption of 15.6MWh for the three years. The valuation function is then defined as $s_\theta(m) = \mathrm{softplus}_\theta(15(m-15.6))$, where $\mathrm{softplus}_\theta = \theta^{-1}\log(1+\exp(\theta x))$. Figure 7.4 shows the two extreme cases : a purely liquid market ($\theta = 0$) and the absence of exchange ($\theta = \infty$). We choose here $\theta = 0.3$ to represent an intermediate case.

7.4.2 Numerical Results

We use $N = 20$ discretization points for the bonus description and $M = 0.1p$. This means that the maximal unitary bonus given to an agent cannot exceed 10% of the electricity price. We take $z^0 \equiv 1$ as initial guess. The main advantage of this initial guess is that it satisfies the

[13] https://www.ecologie.gouv.fr/commercialisation-lelectricite
[14] Loi POPE, 2005 : https://www.ecologie.gouv.fr/dispositif-des-certificats-deconomies-denergie
[15] https://www.calculcee.fr/les-primes-cee.php
[16] https://c2emarket.com/

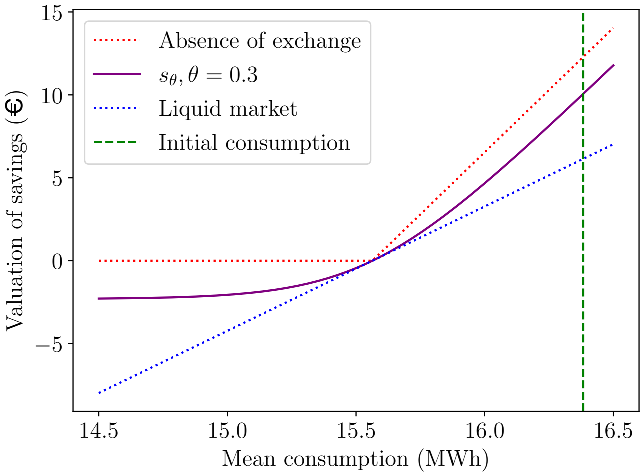


Figure 7.4: Penalty function $s(\cdot)$ given by the regulator.

utility constraint (if $\tau < M$). The initial standard deviation parameter of CMA-ES[17] was set to 5%. The numerical results[18] – parallelized on 10 threads – were obtained on a laptop i7-1065G7 CPU@1.30GHz.

Uniform elasticity. Figure 7.5 shows the results for the test case described in Section 7.4.1, where the price elasticity is identical for all the sub-populations. As a consequence, Proposition 7.2.2 applies and we can analyze in this setting the performance of the numerical solving procedure: in Figure 7.5a, the reward found by Algorithm 10 is very close to the (theoretical) optimal reward, showing that the solver successfully finds the global optimum. About the computational cost, the algorithm converged in approximately 3000 iterations (around 400 seconds), but succeeded in reducing the optimality gap to less than 0.5% in 100 iterations.

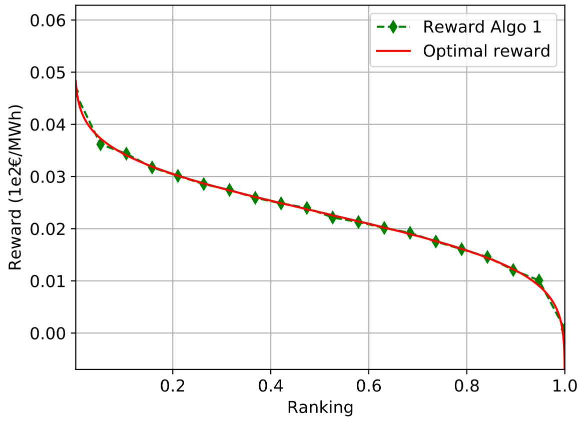
We depict in Figure 7.5 the distribution of the terminal consumption for the four sub-populations with and without the bonus. As shown in Corollary 7.2.1, the distribution without reward is a Gaussian process centered in x^{pi} (which corresponds to three times the annual consumption displayed in Table 7.1). The terminal distribution with the optimal reward is then a shift of this normal distribution – see Proposition 7.2.2. We observe that, as expected, the terminal distribution is also identical for the four sub-populations, up to a scaling ($f_{\mu_k^*}(x) = \theta_k^{-1} f_{\mu_1^*}(\theta_k^{-1} x)$). Here, the mean pluri-annual consumption on the whole population decreased from 16.38MWh to 15.7MWh, giving a saving ratio of 4.1%. This has to be compared with the initial objective of the regulator (a reduction of 5% of the pluri-annual consumption): the retailer found a compromise between the penalty imposed by the regulator, the cost to propose a reward mechanism and its natural willing to sell electricity.

The optimal bonus offered to customers takes negative values for the 1% consuming the most (we choose τ a posteriori in this sense) and goes up to more than 4€ per MWh, which corresponds to a bonus of 66€ in average over the three years. This should be compared for instance with the "Bonus Conso" proposed by TotalEnergies[19], where 30€ are proposed for a reduction of 5% over one year.

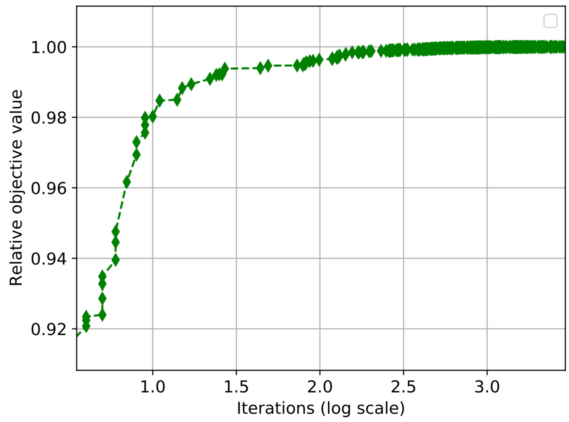
[17] We use the C++ implementation of CMA-ES, available at `https://github.com/CMA-ES/libcmaes`. Practical hints are provided for the choice of the parameters.

[18] The whole code is available on the GitHub repository: `https://github.com/jacquq/rk_games_electricity`.

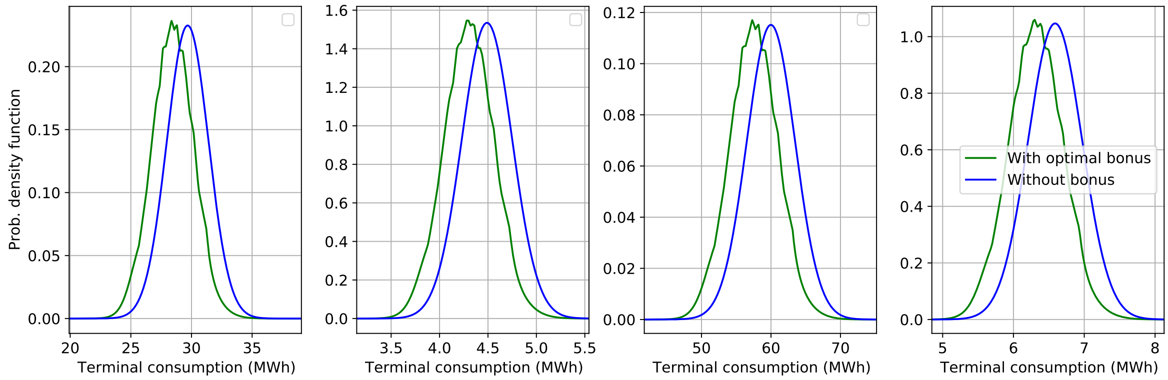
[19] `https://www.totalenergies.fr/bonus-conso`



(a) Analytic optimal reward in red, compared to the unitary bonus function found by Algorithm 10.



(b) Evolution of the relative objective value along the iterations.



(c) Terminal consumption distribution for the four sub-populations

Figure 7.5: Numerical results for the four populations described in Tables 7.1 and 7.2 (scalable case).

The N-players game. We now numerically illustrate the behavior of several individual consumers incentivized by the optimal bonus found in Figure 7.5a. The simulation of the trajectories is done using a Euler-Maruyama scheme, see e.g. [NT15] for details on the discretization, as for convergence rates.

Figure 7.6 displays the evolution of the forecasted consumption $X_1^{a_1^*}$, from which we subtracted the deviation coming from price in order to clearly distinguish the supplementary effort made through the influence of the bonus. This corresponds to the quantity

$$Y_1(t) = X_1^{a_1^*}(t) + \frac{p(t-T)}{2c_1} \quad ,$$

where a_1^* is the optimal effort in the presence of the bonus. We observe the same consumption decrease as in Figure 7.5c, and this reduction has a linear behavior. Indeed, we showed in (7.18) that the optimal total reward is linear in x, and for any reward $R_{k,\mu} = \alpha_0 - \alpha_1 x$, the corresponding effort is $a_k^*(t) = -\frac{\alpha_1}{2c_k}$ – see (7.6) – and the consumption reduction is then $\frac{\alpha_1}{2c_k}t$. This has a strong implication on the behavior of the model: the effort made at time $0 \leqslant t \leqslant T$ by a consumer is independent from his current situation, i.e., is not influenced by the hazard W_t. This means that a consumer will not stop/reduce his effort even if he is undergoing an adverse hazard.

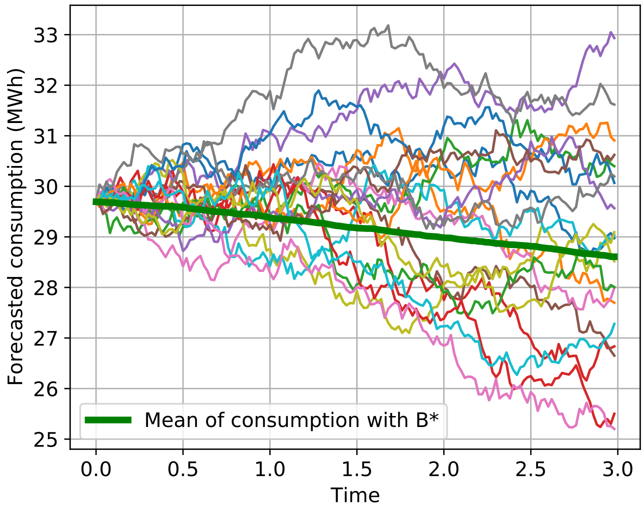


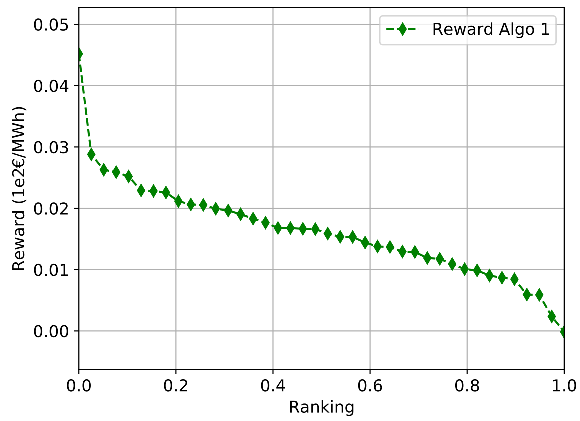
Figure 7.6: Deviation of the consumption from the no-bonus case
Trajectories for 20 consumers from sub-population 1.

Non-uniform price elasticity. We now slightly change the previous test case by considering that the price elasticity is not constant across the population, but rather depends on the characteristics of each agent. In particular, we consider here that the price elasticity of a consumer with electric heating is greater than someone with another heating technology. This greater specific adaptability is for instance exploited by some energy providers[20]. To see the influence of non uniform elasticity, we divide by two the elasticity of sub-populations 2 and 4 – as they do not have electric heating – and multiply by 1.5 the elasticity of sub-populations 1 and 3. In this setting, the scaling condition (7.19) is no longer satisfied, and so, contrary to the previous case, we are not able to find the theoretical optimal bonus function, but only able to perform a numerical optimization using Algorithm 10.

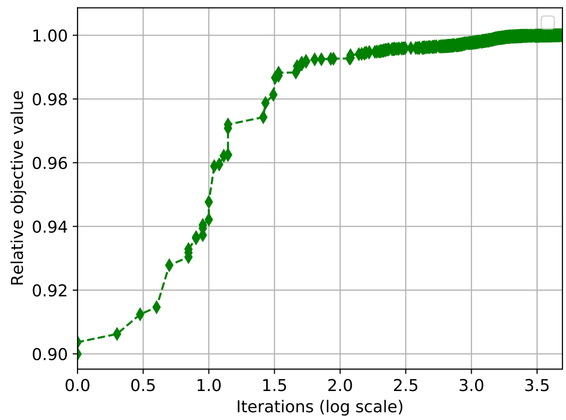
Figure 7.7 shows the results for the test case with modified elasticity parameters. We use here $N = 40$ discretization points and let the algorithm runs up to 5000 iterations. The convergence of Algorithm 10 is still fast since the gap between the solution at iteration 100 was already close to the final solution to less than 1%. About the terminal consumption distribution, we observe that the mean consumption for sub-populations 1 and 3 is reduced by 5.3% whereas the mean consumption for sub-populations 2 and 4 is reduced by 2.3%. Indeed, it reflects the increase (resp. decrease) of price-elasticity for 1 and 3 (resp. 2 and 4). This should be compared with the uniform consumption reduction of 4.1% in the previous setting.

The unitary bonus found by Algorithm 10 is lower than in Figure 7.5: for example, in the uniform-elasticity case, every agent with a ranking lower than 0.6 received a unitary bonus greater than 2€ per MWh, while in the non-uniform case, only consumers with ranking lower than 0.2 can claim this level of reward. This highlights the fact that the retailers does not need to propose a reward as huge as in the previous case since the reduction effort is mostly endorsed by users with electric heating, now more compliant to lower their consumption.

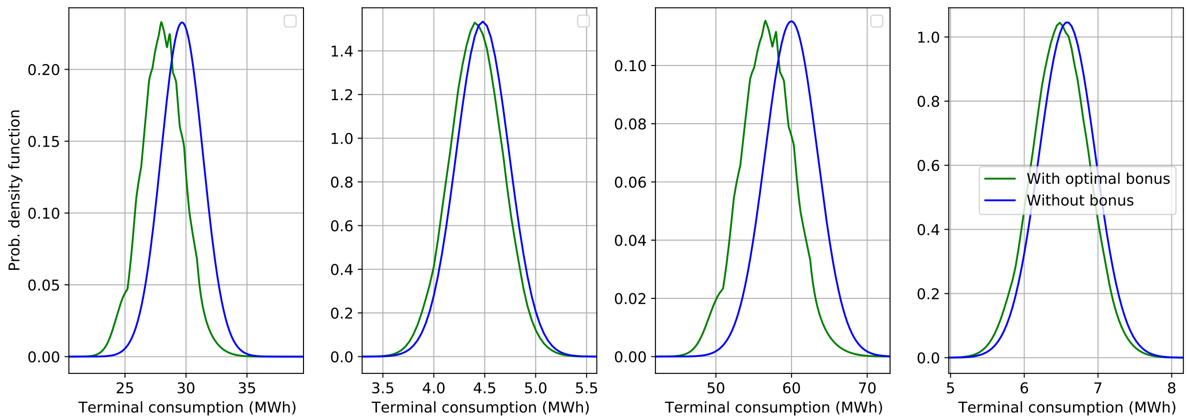
[20] `https://www.sowee.fr/`



(a) Unitary bonus function found by Algorithm 10.



(b) Evolution of the relative objective value along the iterations.



(c) Terminal consumption distribution for the four sub-populations

Figure 7.7: Numerical results for the four populations with different price elasticity.

7.5 Extensions

We propose in this section several extensions to fit with more general settings.

Energy consumption with common-noise. The add of common-noise is not rare in the modeling of electricity consumption. But in this present case, it does not impact the retailer problem. Intuitively, as the reward is determined by the ranking of the agents, an identical perturbation of the consumption will not modify the rankings, and so the effort made by the agents is independent of the common-noise.

Let us prove this intuitive behavior. To this purpose, we fix a sub-population $k \in [K]$, and suppose that the dynamics is now described as:

$$dX_k^a(t) = a_k(t)dt + \sigma_k dW_k(t) + \sigma^0 dW^0(t), \quad X_k(0) = x_k^{\text{nom}} \ . \tag{7.26}$$

Proposition 7.5.1 (Translation invariance of the effort)
Let R_k be the total reward for sub-population k (satisfying Assumption 7.2.2) and μ_k be the equilibrium distribution under R_k and without common-noise (given by (7.8)-(7.9)). Then

$$\mu_k^0 := x \mapsto \mu_k(x - \sigma^0 W^0(T)) \tag{7.27}$$

is a (random) equilibrium distribution under R_k and dynamics (7.26).

Proof. For all $x, q \in \mathbb{R}$ and $\mu \in \mathcal{P}(\mathbb{R})$, we have:

$$R_{k,\mu_k}(x+q) = B_k(F_{\mu_k}(x+q)) - p(x+q) = B_k(F_{\mu_k(\cdot+q)}(x)) - p(x+q) = R_{k,\mu_k(\cdot+q)}(x) - pq \ .$$

Therefore, according to the expression of the optimal effort in (7.6)

$$u_k(t, x+q, \mu) = q^{-1} u_k(t, x, \mu(\cdot + q))$$

and $a_k(t, x+q; \mu_k) = a_k(t, x; \mu(\cdot + q))$. Therefore, the drift is translation invariant, and the results of [LW15] apply: μ^0 defined in (7.27) is an equilibrium distribution for the dynamics with common-noise. □

In contrast with the purely rank-based case, total rewards satisfying Assumption 7.2.2 are not translation invariant. Nonetheless, the drift obtained through the optimal effort is translation invariant, enabling to use the results of [LW15]. For a common-noise W^0 such that $\mathbb{E}[W^0(\cdot)] = 0$, maximizing the (expected version of the) profit, defined in (P^{ret}), will boil down to the same problem, and so will lead to the same optimal unitary reward.

General reward $R(x, r)$. We consider here a more general form of reward, coupling the terminal consumption and the ranking. Therefore, Assumption 7.2.2 is no longer satisfied and the equilibrium cannot be explicitly computed with Theorem 7.2.4. Instead, one can used fixed-point resolution techniques to compute the equilibrium.To this purpose, let us denote by $W_1(f_1, f_2)$ the 1-Wasserstein metric for distribution $f_1, f_2 \in \mathcal{P}_1(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |x| d\mu(x) < \infty\}$. Algorithm 11 follows the standard way to numerically compute mean-field Nash equilibria – see [AL20] – by iteratively updating the distribution using the best response operator. Here, the operator is explicitly given by (7.5), which still applies for general forms of reward function, see [BZ21].

Instead of Picard iterates ($l_i = 1$), a decreasing damping $l_i = \left(\frac{1}{i+1}\right)^p$, $p \in \mathcal{N}$ can be used. The latter sequence of inertial parameters defines iterates of Krasnoselskii-Mann type, which has been proved to converge for pseudo-contractive map in Hilbert space, see [Raf07]. Such a damping has been used for example to solve Linear-Quadratic mean-field control problems in [Gra+16].

We then show that the uniqueness of the reward function is no longer true in the general setting, and there exists a family of equivalent reward function, going from purely rank-based rewards to purely consumption-based reward ones:

Proposition 7.5.2 (Invariance)
Let $R^*(x, r)$ be an optimal reward function for the following problem

$$\max_{R(x,r)} \left\{ -\kappa(m_\mu) - \int_{\mathbb{R}} R_\mu(x) f_\mu(x) dx \;\middle|\; \begin{array}{l} \mu = \epsilon(R) \\ V(R, \mu) \geqslant V^{\text{pi}} \end{array} \right\} \tag{7.28}$$

This equilibrium distribution obtained with R^* is denoted by μ^*. Then,

(i) the purely rank-based reward function $\hat{B} : r \mapsto R^*(q_{\mu^*}(r), r)$ is also an optimal reward,

(ii) the reward function $\hat{R} : x \mapsto R^*(x, F_{\mu^*}(x))$ is also an optimal reward.

Proof. By definition, the two reward functions $\hat{B}$ and $\hat{R}$ also satisfy the characterization of the equilibrium (7.7) with $\mu_k = \mu^*$. Therefore, under these rewards, agents reach the same equilibrium as with R, and their utility is identical. Morover, the objective in (7.28). $\square$

Algorithm 11 Fixed-point Resolution

Require:

- initial p.d.f. $f_{\mu_k^{(0)}}$ of cluster k,
- error tolerance ε,
- iteration maximum n_{max},
- sequence of damping coefficients $\{l_i\}_{i\in\mathbb{N}}$.

$d, i \leftarrow 2\varepsilon, 0$

while $d \geqslant \varepsilon$ or $n \leqslant n_{max}$ **do**

$f_{\mu_k^{(i+1/2)}} \leftarrow \Phi_k(f_{\mu_k^{(i)}})$ ▷ Best-response map defined in Definition 7.2.1

$f_{\mu_k^{(i+1)}} \leftarrow l_i f_{\mu_k^{(i+1/2)}} + (1-l_i) f_{\mu_k^{(i)}}$ ▷ damping l_i

$d \leftarrow W_1\left(f_{\mu_k^{(i)}}, f_{\mu_k^{(i+1)}}\right)$ ▷ distance between two iterates

$i \leftarrow i+1$

end while

In practice, Proposition 7.5.2 has very useful implications. It states that complicated reward policies simplify into simple rules. The first item shows that we can construct a purely *competitive* game in the sense that the consumers receive incentives only through their rank. The second item shows that we can construct a *decentralized* reward since the incentive of each customer only depends on their own consumption. Note that this notion of invariance applies at the equilibrium, and the equivalence of the reward is no longer true outside the equilibrium.

Time-dependent effort cost. In the context of the ecological transition, the consumers are more willing to contribute to the energy reduction, and therefore the effort cost c can be viewed as a time dependent parameter, modeling the change of customers' behavior.

In this case, with a cost profile $c_k(t)$, $t \in [0,T]$ for each cluster k, the consumer's problem becomes

$$V_k(R,\mu_k) := \sup_a \mathbb{E}\left[R_{\mu_k}(X_k^a(T)) - \int_0^T c_k(t) a_k^2(t) dt\right] . \tag{7.29}$$

As a direct extension of [BZ16], we have the following existence result:

Theorem 7.5.1

Assume that the cost profiles are bounded such that there exist $(\underline{c}_k, \overline{c}_k)$ verifying for all $t \leqslant T$

$$0 < \underline{c}_k \leqslant c_k(t) \leqslant \overline{c}_k \ .$$

Then, there exists at least one equilibrium.

Nonetheless, there is no more explicit formula (even for the best response of the agents) in presence of time-varying cost of effort, as the Schrödinger bridge method requires a quadratic cost of effort that is constant over time. To illustrate the behavior of the agents with a time-dependent cost of effort, we draw in Figure 7.8 the trajectories of the same 20 consumers as in Figure 7.6 obtained with the incentive depicted in Figure 7.5a and a cost of effort $c_k(t) = 24 - 1.5t$ €/MWh. As expected, the energy savings are greater than the previous case (the terminal consumption is now around 27.6MWh whereas it was around 28.5MWh with $c_k(t) = 24$€/MWh.

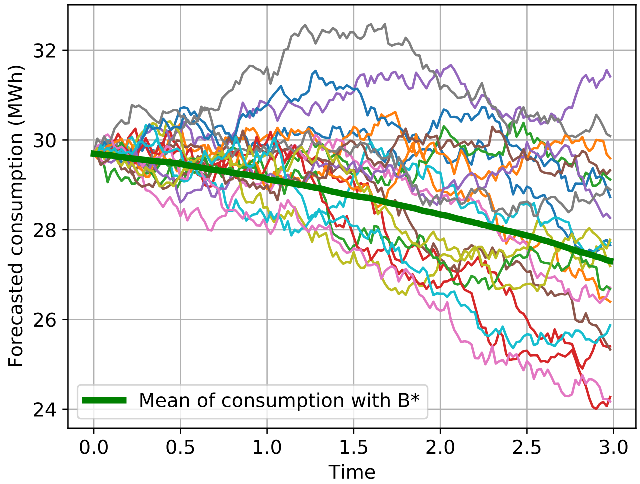


Figure 7.8: Deviation of the consumption from the no-bnous case.
Trajectories of 20 consumers from sub-population 1 obtained with the optimal control from the mean-field approximation and a time-dependent cost of effort.

7.6 Conclusion

In this work, we study a Principal-Agent mean-field game where the incentive designed by the principal is based on the ranking of each agent, initiating a competition between them. This specific framework allows us to derive explicit formula for the (unique) mean-field Nash equilibrium for the agents' problem. Incorporating this characterization in the principal profit maximization problem, we prove in the homogeneous setting that the optimal reward can be obtained by solving a convex reformulation of the problem in the distribution space. We exploit the optimality conditions of the latter to then get the optimal reward through a fixed-point equation. In the general case, we show that the problem can be recast as a finite-dimensional maximization over a box, which can be efficiently solved by numerical algorithms.

We apply the results to electricity markets where a provider aims at designing a reward for its consumers portfolio in order to incentivise them to energy sobriety. We construct realistic instances for the French market of Energy Saving Certificates, and numerically observe that the rank-based rewards can constitute efficient mechanisms to make substantial energy reduction, while staying sufficiently simple to be easily grasped by the consumers.

7.7 Appendix

In this section, we collect several results and proofs.

Lemma 7.7.1

$$f_k^{\mathrm{nom}}(x)\exp(\tau x) = \exp\left(\tau x_k^{\mathrm{nom}} + \frac{1}{2}\tau^2\sigma_k^2 T\right)\varphi\left(x; x_k^{\mathrm{nom}} + \tau\sigma_k^2 T, \sigma_k\sqrt{T}\right) \quad . \tag{7.30}$$

Proof.

$$\begin{aligned} f^{\mathrm{nom}}(x)\exp(\tau x) &= \frac{1}{\sigma\sqrt{T}\sqrt{2\pi}}\exp\left(-\frac{(x-x^{\mathrm{nom}})^2 - 2\tau\sigma^2 T x}{2\sigma^2 T}\right) \\ &= \frac{1}{\sigma\sqrt{T}\sqrt{2\pi}}\exp\left(-\frac{(x-[x^{\mathrm{nom}}+\tau\sigma^2 T])^2}{2\sigma^2 T} + \tau x^{\mathrm{nom}} + \frac{1}{2}\tau^2\sigma^2 T\right) \end{aligned}$$

□

Lemma 7.7.2 (Set of attainable equilibria) (i) For a given cluster k, the set of equilibria attainable by an additional reward function B is given by

$$\mathcal{E}_k = \{\mu \in \mathcal{P}^+(\mathbb{R}) : 2c_k\sigma_k^2 \ln\zeta_{k,\mu_k}(q_{\mu_k}(r)) + pq_{\mu_k}(r) \text{ is bounded and decreasing}\} \ ,$$

with $\zeta_{k,\mu} := f_\mu / f_k^{\mathrm{nom}}$.

(ii) If $\mu_k \in \mathcal{E}_k$, then

$$\epsilon_k^{-1}(\mu_k) = \left\{2c_k\sigma_k^2 \ln\zeta_{k,\mu_k}(q_{\mu_k}(r)) + pq_{\mu_k}(r) + C_k : C_k \in \mathbb{R}\right\}$$

(iii) Suppose that additional reservation "utility" constraint $V_k(R,\mu_k) \geqslant V_k^{\mathrm{pi}} + \tau x_k^{\mathrm{nom}}$ and budget constraint $\int_0^1 B(r)dr \leqslant K$, then the constant C_k in (ii) is restricted to

$$V_k^{\mathrm{pi}} + \tau x_k^{\mathrm{nom}} \leqslant C_k \leqslant K - 2c_k\sigma_k^2\int_0^1 \ln\zeta_{k,\mu_k}(q_{\mu_k}(r))dr - pm_{\mu_k} \quad .$$

In particular, such a C_k exists if and only if

$$2c_k\sigma_k^2\int_0^1 \ln\zeta_{k,\mu_k}(q_{\mu_k}(r))dr - pm_{\mu_k} \leqslant K - V_k^{\mathrm{pi}} - \tau x_k^{\mathrm{nom}} \quad .$$

Proof. Items (i) and (iii) directly comes from [BZ21]. For (ii), the condition of Theorem 7.2.2 is verified:

$$\int_0^r \exp\left(-\frac{R_\mu(q_\mu(z))}{2c\sigma^2}\right)dz = \int_0^r (\zeta_\mu(q_\mu(r)))^{-1}\,dz = \int_{-\infty}^{q_\mu(r)} f^{\mathrm{nom}}(z)dz \ .$$

As the uniqueness is concerned, suppose that B and B' lead to the same distribution μ with $p \neq 0$. Then, B and B' lead to the same distribution ν with $p = 0$, see Theorem 7.2.3. Therefore, as shown in [BZ21], B and B' are equal up to a constant. □

Proof of Theorem 7.2.3

We give here the proof for a given class and, for simplicity, we omit the dependence in k.

Characterization of an equilibrium. First, suppose that ν is an equilibrium distribution for the case $p = 0$. Let $\gamma \in \mathbb{R}$ whose value will be determined later. By definition of f_ν (see (7.5)), we get

$$\int_0^r \exp\left(-\frac{B(z) - p(q_\nu(z)+\gamma)}{2c\sigma^2}\right) dz = \int_{-\infty}^{q_\nu(r)} \exp\left(-\frac{B(F_\nu(x))}{2c\sigma^2} + \frac{p}{2c\sigma^2}(x+\gamma)\right) f_\nu(x)dx$$
$$= \frac{e^{\frac{p}{2c\sigma^2}\gamma}}{\gamma(\nu)} \int_{-\infty}^{q_\nu(r)} \exp\left(-\frac{B(F_\nu(x))}{2c\sigma^2} + \frac{p}{2c\sigma^2}x\right) f^{\text{nom}}(x) \exp\left(\frac{B(F_\nu(x))}{2c\sigma^2}\right) dx.$$

Using (7.30) with $\tau = \frac{p}{2c\sigma^2}$ and the change of variables $u = \frac{x-(x^{\text{nom}}+\frac{pT}{2c})}{\sigma\sqrt{T}}$, we deduce

$$\int_0^r \exp\left(-\frac{B(z) - p(q_\nu(z)+\gamma)}{2c\sigma^2}\right) dz = \frac{1}{\gamma(\nu)} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} \int_{-\infty}^{q_\nu(r)} \varphi\left(x; x^{\text{nom}} + \frac{pT}{2c}, \sigma\sqrt{T}\right) dx$$
$$= \frac{1}{\gamma(\nu)\sqrt{2\pi}} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} \int_{-\infty}^{\frac{q_\nu(r)-(x^{\text{nom}}+\frac{pT}{2c})}{\sigma\sqrt{T}}} \exp\left(-\frac{u^2}{2}\right) du$$
$$= \frac{1}{\gamma(\nu)} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} N\left(\frac{q_\nu(r) - (x^{\text{nom}} + \frac{pT}{2c})}{\sigma\sqrt{T}}\right) .$$

Therefore, taking $\gamma = -\frac{pT}{2c}$, we end up with

$$N\left(\frac{\left[q_\nu(r) - \frac{pT}{2c}\right] - x^{\text{nom}}}{\sigma\sqrt{T}}\right) = \frac{\int_0^r \exp\left(-\frac{B(z)-p\left[q_\nu(z)-\frac{pT}{2c}\right]}{2c\sigma^2}\right) dz}{\int_0^1 \exp\left(-\frac{B(z)-p\left[q_\nu(z)-\frac{pT}{2c}\right]}{2c\sigma^2}\right) dz} .$$

By setting $q_\mu(r) = q_\nu(r) - \frac{pT}{2c}$, we recover the characterization of an equilibrium (see Theorem 7.2.2).

Conversely, suppose now that μ is the equilibrium for $p \in \mathbb{R}$. Then, following the same steps,

$$N\left(\frac{\left[q_\mu(r) + \frac{pT}{2c}\right] - x^{\text{nom}}}{\sigma\sqrt{T}}\right) = \frac{\int_0^r \exp\left(-\frac{B(z)}{2c\sigma^2}\right) dz}{\int_0^1 \exp\left(-\frac{B(z)}{2c\sigma^2}\right) dz} .$$

The distribution ν defined as $q_\nu(r) = q_\mu(r) + \frac{pT}{2c}$ is a valid equilibrium.

Uniqueness of the equilibrium. Suppose that there exist two distinct equilibrium distributions μ and μ' such that $q_\mu \neq q_{\mu'}$. Then by the above proof, we derive the existence of two distinct equilibrium distributions ν and ν' for the case $p = 0$ satisfying $q_\nu \neq q_{\nu'}$. We get a contradiction by the uniqueness of the equilibrium for purely rank-based rewards.

Proof of Lemma 7.2.1

We apply the KKT conditions on $(\widetilde{P}^{\text{ret}})$ (relaxing the positivity assumption on f): for μ^*-almost every x in $\mathbb{R}$,

$$\begin{cases} 0 = x\kappa'(m_{\mu^*}) + 2c\sigma^2 \ln\left(\frac{f_{\mu^*}(x)}{f^{\text{nom}}(x)}\right) + \lambda, \\ \int_{-\infty}^{+\infty} f_{\mu^*}(y)dy = 1 \end{cases} , \lambda \in \mathbb{R}$$

From which we can deduce that $f_{\mu^*}(x) = f^{\text{nom}}(x)\exp\left(-\frac{x\kappa'(m_{\mu^*})+\lambda}{2c\sigma^2}\right)$, which is positive for all x. The Lagrange multiplier λ is then computed using the normalization condition on f_{μ^*}.

Proof of Theorem 7.2.4

Integrating (7.13) gives us

$$\begin{aligned} m_\mu = \int_{-\infty}^{+\infty} y f_\mu(y)dy &= \frac{1}{\alpha(\mu)}\int_{-\infty}^{+\infty} y f^{\text{nom}}(y)\exp\left(-y\frac{\kappa'(m_\mu)}{2c\sigma^2}\right)dy \\ &= \int_{-\infty}^{+\infty} y\phi\left(y; x^{\text{nom}} - \frac{T\kappa'(m_\mu)}{2c}, \sigma\sqrt{T}\right)dy \\ &= x^{\text{nom}} - \frac{T\kappa'(m_\mu)}{2c} = x^{\text{pi}} + \frac{T}{2c}\delta(m_\mu) \ , \end{aligned}$$

where we use Lemma 7.7.1 between the two first lines in order to recover a gaussian process.

We can now recover the reward:

$$\begin{aligned} B^*(r) &= V^{\text{pi}} + \tau x^{\text{nom}} + 2c\sigma^2\ln\left(\zeta_{\mu^*}(q_{\mu^*}(r))\right) + pq_{\mu^*}(r) \\ &= V^{\text{pi}} + \tau x^{\text{nom}} + q_{\mu^*}(r)\left[p - \kappa'(m_{\mu^*})\right] - 2c\sigma^2\ln\left(\int_{-\infty}^{+\infty} f^{\text{nom}}(y)\exp\left(-y\frac{\kappa'(m_{\mu^*})}{2c\sigma^2}\right)dy\right) \\ &= V^{\text{pi}} + \tau x^{\text{nom}} + \frac{c}{T}\left[(x^{\text{nom}})^2 - m^2\right] + q_{\mu^*}(r)\delta(m_{\mu^*}) \\ &= \tau x^{\text{nom}} + \frac{c}{T}\left[(x^{\text{pi}})^2 - m^2\right] + q_{\mu^*}(r)\delta(m_{\mu^*}) \ , \end{aligned}$$

where we use Lemma 7.7.1 to get the value of the integral. From the definition of the provider objective,

$$\begin{aligned} \pi &= pm - \kappa(m) - \int_0^1 B^*(r)dr \\ &= pm - \kappa(m) - \frac{c}{T}\left[(x^{\text{pi}})^2 - m^2\right] - m\left[p - \kappa'(m)\right] - \tau x^{\text{nom}} \\ &= m\kappa'(m) - \kappa(m) + \left(\frac{x^{\text{pi}} + m}{2}\right)\frac{2c}{T}\left(m - x^{\text{pi}}\right) - \tau x^{\text{nom}} \\ &= m\kappa'(m) - \kappa(m) + \left(\frac{x^{\text{pi}} + m}{2}\right)\delta(m) - \tau x^{\text{nom}} \ . \end{aligned}$$

Proof of Proposition 7.5.2

(i) By construction, the reward $\tilde{B}$ is also bounded and decreasing. Then, the cost induced by the additional reward is the same with R^* and $\hat{B}$:

$$\int_{-\infty}^{+\infty} R^*_{\mu^*}(x) f_{\mu^*}(x)dx = \int_0^1 \hat{B}(r)dr \ .$$

Finally, μ^* is also an equilibrium for the reward $\hat{B}$:

$$\frac{1}{\hat{\gamma}(\mu^*)} f^{\text{nom}}(x)\exp\left(\frac{\hat{B}(F_{\mu^*}(x))}{2c\sigma^2}\right) = \frac{1}{\gamma^*(\mu^*)} f^{\text{nom}}(x)\exp\left(\frac{R^*_{\mu^*}(x)}{2c\sigma^2}\right) = f_{\mu^*} \ ,$$

where $\hat{\gamma}$ and γ^* are computed respectively with $\hat{B}$ and R^*. The last equality comes from the characterization of an equilibrium. Therefore, the reward function $\hat{B}$ satisfies the constraints and produces the same objective value as R^*. It is also optimal.

(ii) The proof follows the same ideas as at the previous item.

Chapter 8

Tight Bound for Sum of Heterogeneous Random Variables: Application to Chance Constrained Programming

***

This chapter is based on the submitted chapter [JZ23].

Abstract. We study a tight Bennett-type concentration inequality for sums of heterogeneous and independent variables, defined as a one-dimensional minimization problem. We show that this refinement, which outperforms the standard known bounds, remains computationally tractable: we develop a polynomial-time algorithm to compute confidence bounds, proved to terminate with an ϵ-solution. From the proposed inequality, we deduce tight distributionally robust bounds to Chance-Constrained Programming problems. To illustrate the efficiency of our approach, we consider two use cases. First, we study the chance-constrained binary knapsack problem and highlight the efficiency of our cutting-plane approach by obtaining stronger solution than classical inequalities (such as Chebyshev-Cantelli or Hoeffding). Second, we deal with the Support Vector Machine problem, where the convex conservative approximation we obtain improves the robustness of the separating hyperplane, while staying computationally tractable.

8.1 Introduction

Concentration inequalities – such as Hoeffding [Hoe63], Bennett [Ben62] or McDiarmid [McD89] to cite a few – were originally introduced to quantify how a random variable deviates from its expectation. In this context, the probability to deviate is then estimated using information on the two first moments (mean and variance) or the length of the support of the distribution, depending on the concentration inequality that is considered. These inequalities have now a wide variety of applications, see e.g. [BLB04], including chance constrained programming or machine learning [NS07; PML22; WFP15; Kha+22].

Many refinements of Hoeffding and Bennett's inequalities have been proposed: all these works exploit Chernoff's inequality but differ in the estimation of the moment-generating function $t \mapsto \mathbb{E}[e^{t\xi}]$ associated to a random variable ξ. Figure 8.1 proposes a schematic classification of the literature. From and Swift [FS13] and Zheng [Zhe17] both use a linear approximation of $x \mapsto e^{tx}$, that is tighter than Hoeffding' bound [Hoe63] for variables in $[0, 1]$. They differ in the use of the arithmetic-geometric mean inequality. Jebara [Jeb18] exploits an inequality from [Ben62, (b)] to derive an analytic one-sided bound for sum of heterogeneous random variables. Finally, Cheng and Li [CL22] insert a multipoint approximation of $e^{t\xi}$ and compare their results with [Zhe17]. We emphazise that the classification we made – which is a contribution on its own – focuses on the crucial approximation done while tackling with Chernoff's inequalities, and does not directly compare the final bounds obtained in each work.

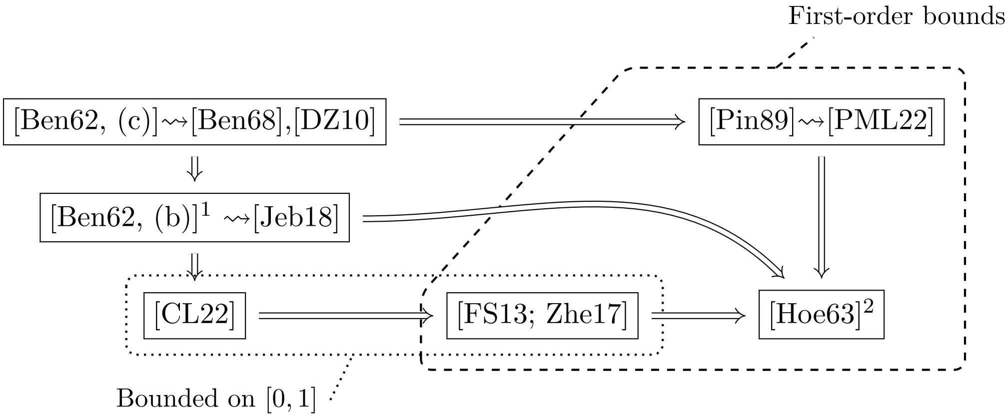


Figure 8.1: Classification of $t \mapsto \mathbb{E}[e^{t\xi}]$ estimations
[1] Bennett's inequality
[2] Hoeffding's inequality

If [a] $\Rightarrow$ [b], the upper estimator of the moment-generating function in [a] is tighter than in [b]. If [a] $\rightsquigarrow$ [b], [b] uses the same moment-generating estimator but improves / extends the results. Proofs of the different implications are provided in Section 8.6.1.

In this chapter, we focus on the moment-generating estimator introduced in [Ben62, (c)]: for a random variable ξ with mean μ, variance σ^2 and maximal positive deviation $b \in \mathbb{R}_{\geqslant 0}$ (i.e., $\xi - \mu \leqslant b$), we have the following inequality:

$$\forall t \geqslant 0,\ \mathbb{E}\left[e^{t\xi}\right] \leqslant e^{t\mu} \frac{\sigma^2 e^{tb} + b^2 e^{-\frac{t\sigma^2}{b}}}{b^2 + \sigma^2} \ . \tag{8.1}$$

This estimator is as tight as possible (knowing only μ, σ and b), since it has been proved to be exact for a particular Bernoulli distribution, see e.g. [Ben62]. Dembo and Zeitouni [DZ10] exploit this inequality to obtain a closed-form expression (involving a Kullback–Leibler divergence, see e.g. [BL15]) in the specific case of identically distributed variables. Bennett [Ben68] extends the results to non identically distributed variables, but, in order to obtain an explicit formula, further approximations have been made, leaving room for possible improvements. In contrast, we do not make additional approximations and directly construct the Chernoff bound using (8.1), see Theorem 8.2.1. Even if an analytic solution is not known in the heterogeneous setting, we prove that this bound can be used in many applications.

We first focus on the computation of confidence bound and introduce a double bisection algorithm (Algorithm 12). We prove that this algorithm computes a bound with arbitrary precision in polynomial time (Theorem 8.3.1). This algorithm belongs to the class of Probabilistic Bisection Algorithms (PBA), see e.g. [Hor63; WFH13], but instead of having a zero-mean noise, the error is bounded and controlled by a parameter.

We then apply this result on Chance-Constrained Programming (CCP) problems [CC59; MW65; Pré70; Hen04; Ack20], a very attractive tool for dealing with uncertainty in optimization problems in addition to stochastic [BL11; KW94; RS03] and robust [BEN09; BN00] optimization approaches. This approach relies upon the characterization of uncertainty by means of probabilistic information and tries to find a good solution in a probabilistic sense. A general CCP problem is expressed as:

$$\begin{array}{rl} \min_{x\in X} & \mathbb{E}[c(x,\xi)] \\ \text{s.t.} & \mathbb{P}[g_i(x,\xi) \geqslant 0, (i=1,...,m)] \geqslant p \end{array} \tag{8.2}$$

where $x \in X \subseteq \mathbb{R}^n$ denotes a decision vector and c is a cost function impacted by a random process (uncertainty) $\xi \in \mathbb{R}^m$. Also, $\mathbb{P}$ is the probability measure associated to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined ξ applied to a whole system of m stochastic inequalities $g_i(x,\xi) \geqslant 0, (i=1,...,m)$. The parameter $p \in]0,1[$ is then a given confidence level, typically close to 1.

Although CCP problems are in general very challenging, some specific cases lead to tractable algorithms. In particular, for individual chance-constrained optimization (i.e., $\mathbb{P}[g_i(x,\xi) \geqslant 0] \geqslant p_i,\ (i=1,...,m)$), and with Gaussian uncertainty (i.e., ξ follows a normal distribution), the CCP problem can be formulated as a (convex) Second-Order Cone Programming (SOCP) problem, see e.g. [Hen07]. More generally, for specific families of distributions, it is known that the set of a probabilistic constraint is convex, making possible the use of nonlinear methods, see e.g. [Pré95; LLS05; AM19]. However, the distributions are commonly unavailable in many applications, and even the evaluation of the constraint is not easy. A classical method is then to find a conservative approximation of the problem that is distributionally robust, see e.g. [Sha21]. In this case, the chance-constraint are satisfied for any distribution and an optimal solution of the approximate problem gives a feasible solution of the original chance-constrained problem. Concentration inequalities have already been used in this context: to the best of our knowledge, Pinter [Pin89] was the first to use concentration inequalities in optimization problem. Nemirovski and Shapiro [NS07] proved that the use of Chernoff's bounds provides tractable conservative approximation of chance-constrained problems. In particular, they detailed the convex approximation for several families of univariate distributions. Peng, Maggioni and Lisser [PML22] focus on SOCP conservative approximations of two types: distributionally robust formulations based on Hoeffding and Chebyshev's inequalities, and models which assume a normally distributed uncertainty. In particular, they deal with joint independent chance-constraints, which is known to be of a high complexity, see e.g. [NS07].

Here, we compare various formulations on knapsack problems [CG06; RP21] by specializing the study to second-order bounds (knowledge of means and variances). To this purpose, we introduce a new convex conservative approximation based on Bernstein's inequality (Proposition 8.4.1) and derive from the tight Bennett's inequality a strong approximation (Proposition 8.4.2). We show that the two formulations can be efficiently solved by a cutting-plane approach and lead to solution improvement on instances from the literature (Table 8.2). In particular, for a given budget, we improve the objective value compared to the SOCP formulations [PML22].

Finally, we focus on the Support Vector Machine (SVM) problem, see e.g. [CV95] in the deterministic setting. Here, under uncertainties, the main difficulty lies in the large number of probabilistic constraints. The distributionally robust version of the problem has been addressed in [Ben+10; WFP15; Kha+22; FMP22]. In particular, Ben-Tal et al. [Ben+10] first consider the same moment-generating estimator (8.1), but make additional approximations in order to obtain SOCP formulations. Using convex optimization tools, we numerically highlight that our approach increases further the quality of the separating hyperplane while staying tractable for instances of substantial size.

This chapter is organized as follows. In Section 8.2, we first derive properties of the proposed inequality, and numerically observe its asymptotic behavior. Then, we introduce in Section 8.3 an algorithm to compute confidence bounds. Finally, in Section 8.4, we apply the inequality to chance-constrained programming, focusing on the knapsack problem (Section 8.4.1) and on the Support Vector Machine problem (Section 8.4.2).

Notations. For two vectors a, b of $\mathbb{R}^N$, we denote by $\langle a, b\rangle$ the Euclidean scalar product. Moreover, $a \wedge b$ (resp. $a \vee b$) stands for the component-wise maximum (resp. the minimum) between a and b. Besides, for a discrete set X, $\mathrm{vex}(X)$ will read as the convex envelope of X.

8.2 On the tightest Cramér-Chernoff bound

We first recall Hoeffding's [Hoe63] and Bennett's [Ben62] inequalities:

Proposition 8.2.1 (Hoeffding)

Let $\xi_1, \dots, \xi_N$ be N independent random variables. If there exist $a, b \in \mathbb{R}^N$ such that $\mathbb{P}[a_k \leqslant \xi_k - \mathbb{E}[\xi_k] \leqslant b_k] = 1$ for all $k \in \{1, \dots, N\}$, then, for all $d \geqslant 0$,

$$\ln\ \mathbb{P}\left[\sum_{k=1}^{N} \xi_k - \mathbb{E}[\xi_k] \geqslant d\right] \leqslant -\frac{2d^2}{\sum_{k=1}^{N}(b_k - a_k)^2} \ . \tag{8.3}$$

Consequently, for all $\tau \in]0,1[$, $\mathbb{P}\left[\sum_{k=1}^{N} \xi_k - \mathbb{E}[\xi_k] \geqslant d_\tau\right] \leqslant \tau$ where $d_\tau = \|b-a\|_2\sqrt{-\ln\left(\sqrt{\tau}\right)}$.

Proposition 8.2.2 (Bennett)

Let $\xi_1, \dots, \xi_N$ be N independent random variables. If there exist $a, b, \sigma \in \mathbb{R}^N$ such that

(i) $\mathbb{P}[\xi_k - \mathbb{E}[\xi_k] \leqslant b] = 1$, $k \in \{1, \dots, N\}$,

(ii) $\sum_{k=1}^{N} \mathbb{E}[\xi_k^2] \leqslant \sigma^2$.

Then, with $g : u \mapsto (1+u)\ln(1+u) - u$, we get for all $d \geqslant 0$,

$$\ln\ \mathbb{P}\left[\sum_{k=1}^{N} \xi_k - \mathbb{E}[\xi_k] \geqslant d\right] \leqslant -\frac{\sigma^2}{b^2} g\left(\frac{bd}{\sigma^2}\right) \ . \tag{8.4}$$

Consequently, for all $\tau \in]0,1[$, $\mathbb{P}\left[\sum_{k=1}^{N} \xi_k - \mathbb{E}[\xi_k] \geqslant d_\tau\right] \leqslant \tau$ where $d_\tau = \frac{\sigma^2}{b} g^{-1}\left(\frac{b^2}{\sigma^2}\ln\left(\frac{1}{\tau}\right)\right)$.

Proposition 8.2.1 and Proposition 8.2.2 do not suppose the same *a priori* knowledge on the random variables: in the latter, information in second-moment is supposed whereas the former only needs knowledge on the mean of each random variable. We now focus on the tightest second-order Cramér-Chernoff bound, firstly introduced in [Ben62], and based on (8.1):

Theorem 8.2.1 (Refined Bennett's inequality [NS07],Table 2)
Let $\xi_1, \dots, \xi_N$ be N independent random variables. If there exist $b, \sigma \in \mathbb{R}^N$ such that such that

(i) $\mathbb{P}[\xi_k - \mathbb{E}[\xi_k] \leqslant b_k] = 1$, $k \in \{1, \dots, N\}$,

(ii) $\mathrm{Var}(\xi_k) \leqslant \sigma_k^2$, $k \in \{1, \dots, N\}$.

Then, introducing $\gamma_k := \frac{\sigma_k^2}{b_k^2}$, for all $d \geqslant 0$

$$\forall \lambda \in \mathbb{R}^N_{\geqslant 0}, \quad \ln\ \mathbb{P}\left[\langle \lambda, \xi - \mathbb{E}[\xi] \rangle \geqslant d\right] \leqslant \inf_{t \geqslant 0} \left\{ -td + \sum_{k=1}^{N} \ln \left(\frac{\gamma_k e^{t\lambda_k b_k} + e^{-t\lambda_k b_k \gamma_k}}{1+\gamma_k} \right) \right\} \ . \tag{8.5}$$

In addition, if $\mathbb{P}[\xi_k - \mathbb{E}[\xi_k] \geqslant -b_k] = 1$ for all $k \in \{1, \dots, N\}$,

$$\forall \lambda \in \mathbb{R}^N, \quad \ln\ \mathbb{P}\left[\langle \lambda, \xi - \mathbb{E}[\xi] \rangle \geqslant d\right] \leqslant \inf_{t \geqslant 0} \left\{ -td + \sum_{k=1}^{N} \ln \left(\frac{\gamma_k e^{t|\lambda_k| b_k} + e^{-t|\lambda_k| b_k \gamma_k}}{1+\gamma_k} \right) \right\} \ . \tag{8.6}$$

Proof. Using the Chernoff's inequality on the variable $\langle \lambda, \xi - \mathbb{E}[\xi] \rangle$, $\lambda \in \mathbb{R}^N$, we obtain

$$\mathbb{P}\left[\langle \lambda, \xi - \mathbb{E}[\xi] \rangle \geqslant d\right] \leqslant e^{-t(d + \langle \lambda, \mathbb{E}[\xi] \rangle)}\ \mathbb{E}\left[e^{t\langle \lambda, \xi \rangle}\right] \ .$$

By the independence of the variables ξ_k, we have $\mathbb{E}\left[e^{t\langle \lambda, \xi \rangle}\right] = \prod_{k=1}^N \mathbb{E}\left[e^{t\lambda_k \xi_k}\right]$. Finally, using (8.1), we obtain for all $t \geqslant 0$ and $\lambda \in \mathbb{R}^N_{\geqslant 0}$,

$$\mathbb{P}[\langle \lambda, \xi - \mathbb{E}[\xi] \rangle \geqslant d] \leqslant e^{-td} \prod_{k=1}^{N} \left(\frac{\gamma_k e^{t\lambda_k b_k} + e^{-t\lambda_k \gamma_k b_k}}{1+\gamma_k} \right) \ .$$

For $\lambda \in \mathbb{R}^N$ (possibly taking negative values), we boil down to the previous case by considering that $\langle \lambda, \xi - \mathbb{E}[\xi] \rangle = \langle |\lambda|, \chi - \mathbb{E}[\chi] \rangle$ where $\chi_k = \mathrm{sign}(\lambda_k)\xi_k$. By assumption, χ_k has also a maximal positive deviation less or equal than b. We conclude by applying the logarithm and by rearranging the terms.□ □

The right-hand sides of (8.5) and (8.6) correspond to the Cramér transform [DZ10, Section 2.2] of the Bernoulli distribution that achieves the equality in (8.1). The scope of Theorem 8.2.1 is slightly more general than Hoeffding and Bennett inequality since we allow to have sum of weighted heterogeneous random variables (positive or negative weights).

Under the assumptions of Theorem 8.2.1, and introducing $\tau^- := \prod_{k=1}^N \frac{\gamma_k}{1+\gamma_k}$, we get as an immediate corollary that for all $\alpha \geqslant 0$

$$\ln\ \mathbb{P}\left[\sum_{k=1}^{N} \xi_k - \mathbb{E}[\xi_k] \geqslant \alpha N\right] \leqslant \varphi_\alpha^* := \inf_{t \geqslant 0} \varphi_\alpha(t) \ , \tag{8.7}$$

where

$$\varphi_\alpha : t \geqslant 0 \mapsto \ln(\tau^-) + Nt\left(\bar{b} - \alpha\right) + \sum_{k=1}^{N} \ln\left(1 + \gamma_k^{-1} e^{-tb_k(1+\gamma_k)}\right) \tag{8.8}$$

and $\bar{b} = \frac{1}{N}\sum_{k=1}^{N} b_k$. The expression of φ_α is derived from (8.5) with $\lambda_k = 1$ for all $k \in \{1, \dots, N\}$. In the specific case where the coefficient $\mathbb{E}[\xi_k], \sigma_k$ and b_k are identical for all $k \in \{1 \dots N\}$ (*homogeneous* setting), the minimization in t that appears in (8.5) has an analytic solution (using Kullback-Leibler divergence), see e.g. [DZ10; RS13]. In the framework of this chapter, we allow for heterogeneous parameters, and therefore the minimum is no longer analytically known. Nonetheless, the following properties show that the one-dimensional minimization problem is well defined:

Proposition 8.2.3 (Study of φ_α)
Let $\alpha \geqslant 0$, then $\varphi_\alpha(0) = 0$. Moreover, the mapping φ_α is twice differentiable and its respective derivatives are

(i) $$\frac{d}{dt}\varphi_\alpha(t) = N\left(\bar{b} - \alpha\right) - \sum_{k=1}^{N} \frac{b_k(1+\gamma_k)}{1+\gamma_k e^{tb_k(1+\gamma_k)}} \quad \text{(in particular, } \tfrac{d}{dt}\varphi_\alpha(0) = -N\alpha\text{)},$$

(ii) $$\frac{d^2}{dt^2}\varphi_\alpha(t) = \sum_{k=1}^{N} b_k^2(1+\gamma_k)^2 \frac{\gamma_k e^{tb_k(1+\gamma_k)}}{\left(1+\gamma_k e^{tb_k(1+\gamma_k)}\right)^2} .$$
Moreover, $0 < \frac{d^2}{dt^2}\varphi_\alpha(t) \leqslant \Phi := \frac{1}{4}\sum_{k=1}^{N} b_k^2(1+\gamma_k)^2$.

Proof. Let us introduce for all $\gamma, d \in \mathbb{R}_{>0}$, $f(t) = \ln\left(1+\gamma^{-1}e^{-td}\right)$. It follows that $f'(t) = \frac{-d}{1+\gamma e^{td}}$ and $f''(t) = d^2\frac{\gamma e^{td}}{\left(1+\gamma e^{td}\right)^2} \in \left]0, \frac{d^2}{4}\right]$. To recover the results, note that $\varphi_\alpha(t)$ is composed of an affine part plus the sum over $k \in \{1, \dots, N\}$ of functions $f(\cdot)$ with $d = b_k(1+\gamma_k)$.□ □

We immediately deduce from Proposition 8.2.3 that the function $\varphi_\alpha(\cdot)$ is strictly convex, and thus the position of the minimum, denoted by $t_\alpha^* \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$, is unique. The following lemma lists useful properties of φ_α^* that will be used in the sequel.

Lemma 8.2.1 (Study of φ_α^*)

(i) For $0 < \alpha \leqslant \bar{b}$, the function $\alpha \mapsto \varphi_\alpha^*$ is decreasing and $t_\alpha^* > 0$.
Moreover, $\varphi_0^* = 0$, $\varphi_{\bar{b}}^* = \ln(\tau^-)$ and for $\alpha > \bar{b}$, $\varphi_\alpha^* = -\infty$.

(ii) For $\alpha < \min_k\{b_k\}$, $\min_k t_\alpha^{(k)} \leqslant t_\alpha^* \leqslant \max_k t_\alpha^{(k)}$, where $t_\alpha^{(k)} = \frac{1}{b_k(1+\gamma_k)} \ln\left(\frac{\alpha + b_k\gamma_k}{\gamma_k(b_k-\alpha)}\right)$.
Moreover, $t_\alpha^* \leqslant -\frac{1}{N(\bar{b}-\alpha)}\ln(\tau^-)$.

(iii) For $\alpha_1, \alpha_2 < \bar{b}$, $\left|\varphi_{\alpha_2}^* - \varphi_{\alpha_1}^*\right| \geqslant N\min\{t_{\alpha_1}^*, t_{\alpha_2}^*\}\,|\alpha_2 - \alpha_1|$.

Proof. (i) Let $0 < \alpha < \beta$. Then, there exists t_α^* such that $\varphi_\alpha^* = \varphi_\alpha(t_\alpha^*)$. By item (i) of Proposition 8.2.3, t_α^* must be positive as the derivative pf φ_α is negative in $t = 0$. Besides, $\varphi_\alpha(t_\alpha^*) = Nt_\alpha^*(\beta-\alpha) + \varphi_\beta(t_\alpha^*) > \varphi_\beta(t_\alpha^*)$. Here, we exploit the fact that for $\alpha > 0$, $t_\alpha^* > 0$. As $\varphi_\beta^* \leqslant \varphi_\beta(t_\alpha^*)$ by optimality, we easily conclude.
Then, if $\alpha = \bar{b}$, the infimum is reached for $t \to +\infty$ and is equal to $\ln(\tau^-)$. If now $\alpha = 0$, then the minimum is attained at $t = 0$ ($\frac{d\varphi_\alpha}{dt}(t) \geqslant 0$). Finally, if $\alpha > \bar{b}$, then $t \mapsto \varphi_\alpha(t)$ is decreasing, and diverges to $-\infty$ when $t \to \infty$.

(ii) $t_\alpha^{(k)}$ would be the minimum if there was only the k-th term in φ (the existence is guaranteed by the condition $\alpha \leqslant b_k$). We then get the property by using the fact that in one dimension,

the minimum of a sum of convex functions lies in between the minimum and maximum value of the minimizers of each function.
Besides, by (i), $\varphi_\alpha(t^*_\alpha) \leqslant \varphi_\alpha(0) = 0$. Therefore, $\ln(\tau^-) + Nt^*_\alpha(\bar{b} - \alpha) \leqslant 0$ (since the summation in (8.8) is positive), and so $Nt^*_\alpha \leqslant \frac{-1}{\bar{b}-\alpha} \ln(\tau^-)$.

(iii) Let $\alpha_1, \alpha_2 \leqslant \bar{b}$. Then, by definition,

$$\begin{aligned} \varphi^*_{\alpha_1} &= \varphi_{\alpha_1}(t^*_{\alpha_1}) = Nt^*_{\alpha_1}(\alpha_2 - \alpha_1) + \varphi_{\alpha_2}(t^*_{\alpha_1}) \\ \varphi^*_{\alpha_2} &= \varphi_{\alpha_2}(t^*_{\alpha_2}) = Nt^*_{\alpha_2}(\alpha_1 - \alpha_2) + \varphi_{\alpha_1}(t^*_{\alpha_2}) \end{aligned}$$

from which we deduce by optimality of $t^*_{\alpha_1}$ and $t^*_{\alpha_2}$:

$$\begin{aligned} \varphi^*_{\alpha_1} &\geqslant Nt^*_{\alpha_1}(\alpha_2 - \alpha_1) + \varphi^*_{\alpha_2} \\ \varphi^*_{\alpha_2} &\geqslant Nt^*_{\alpha_2}(\alpha_1 - \alpha_2) + \varphi^*_{\alpha_1} \end{aligned}$$

Consequently, $|\varphi^*_{\alpha_2} - \varphi^*_{\alpha_1}| \geqslant N \min\{t^*_{\alpha_1}, t^*_{\alpha_2}\} \, |\alpha_2 - \alpha_1|$.

□ □

The next theorem can be directly derived from Lemma 8.2.1 and provides an alternative confidence bound $\alpha_\tau N$ to the bound d_τ provided in Proposition 8.2.1 and Proposition 8.2.2.

Theorem 8.2.2

For all $\tau \in [\tau^-, 1]$, there exists a unique α_τ such that $\varphi^*_{\alpha_\tau} = \ln(\tau)$. Consequently, $\mathbb{P}\left[\sum_{k=1}^N \xi_k - \mathbb{E}[\xi_k] \geqslant \alpha_\tau N\right] \leqslant \tau$.

Numerical experiments. We aim to numerically compare the bounds developed in Section 8.2 with four inequalities: Hoeffding (8.3), Bennett (8.4), Cantelli (a one-sided improvement of Chebyshev's inequality, see e.g. [BLB04]) and the bound introduced by Jebara [Jeb18]. To this purpose, we follow the methodology of [Jeb18]: we search to bound $\ln\, \mathbb{P}\left[\sum_{k=1}^N \xi_k - \mathbb{E}[\xi_k] \geqslant \alpha N\right]$, where the parameters $\mathbb{E}[\xi_k], \sigma_k, a_k, b_k$ and α are randomly generated following the rules described in Table 8.1.

$\mathbb{E}[\xi_k]$	$\mathcal{U}(0,1)$
a_k	$\mathcal{U}(-1,0)$
b_k	$\mathcal{U}(0,1)$
σ_k	$\mathcal{U}\left(0, (b_k - a_k)/2\right)$
α	$\mathcal{U}\left(0, \bar{b}\right)$

Table 8.1: Definition of the random variables

In order to have a fast implementation of φ^*_α, we introduce a bisection algorithm, see Algorithm 12. Note that this bisection method is only valid because we have shown that φ_α is convex and t^*_α is bounded, see Lemma 8.2.1. The four other bounds are immediate to compute as they are analytically known.

The results are depicted in Figure 8.2 for 500 realizations of the uniformally distributed parameter α and are performed on a laptop `Intel Core i7 @2.20GHz` $\times$ 12. For each realization and for each of the four inequalities with which we are comparing ourselves, we report the corresponding value of φ^*_α (x-coordinate) and the log-probability obtained by the latter inequality (y-coordinate). The computation of the Bennett's and Hoeffding's bounds is almost immediate.

Algorithm 12 Bisection Search to compute φ^*_α

Require: N, α, b_k, σ_k, ϵ
  $t^-, t^+ \leftarrow 0, -\frac{1}{N(\bar{b}-\alpha)}\ln(\tau^-)$
  while $t^+ - t^- > \epsilon$ **do**
    $\hat{t} \leftarrow \frac{1}{2}(t^- + t^+)$
    $g \leftarrow \frac{d}{dt}\varphi_\alpha(\hat{t})$
    if $g \geqslant 0$ **then** $t^+ \leftarrow \hat{t}$ **else** $t^- \leftarrow \hat{t}$
  end while
  return $\hat{\varphi}$

The refined version of [Jeb18] takes around 1ms per instance for $N = 100$ (due to the computation of Lambert function), and φ^*_α takes around 5ms per instance for $N = 100$ for precision $\epsilon = $ 1e-6.

We recover the results proved in Section 8.6.1: φ^*_α always outperforms Bennett, Hoeffding and Jebara's inequalities. We observe that Cantelli's bound is better for large probability error (small α) – typically $\exp(\varphi^*_\alpha) \geqslant 20\%$ – but becomes rapidly dominated by the four other Chernoff's inequalities. In fact, Chebyshev's inequality has a quadratic decay in α when the Chernoff's bound has exponential behaviors . For $\varphi^*_\alpha \geqslant -5$, the bound from [Jeb18] may be less efficient. Possibly, this bound can exceed 1, because the minimizer that is used in the Chernoff's inequality has no guarantee to be optimal.

Remark 8.2.1

We did not display the results for Bernstein's bound, as it is known that this inequality is strictly looser than Bennett's inequality [Ben62], see e.g. [Jeb18] for a proof.

8.3 Computing confidence bounds

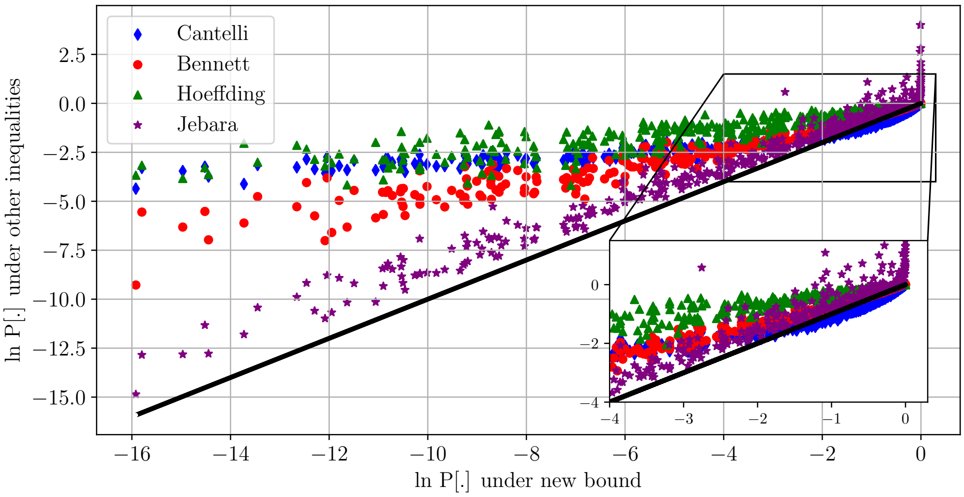
In this section, we aim to derive a confidence bound α_τ with a given maximum probability error τ such that $\varphi^*_{\alpha_\tau} = \ln(\tau)$. We first found an upper approximation of this confidence bound which is strictly lower than $\bar{b}$ for $\tau > \tau^-$ (note that for a probability error less or equal than τ^-, we can only certify a confidence bound of $\bar{b}$).

Lemma 8.3.1

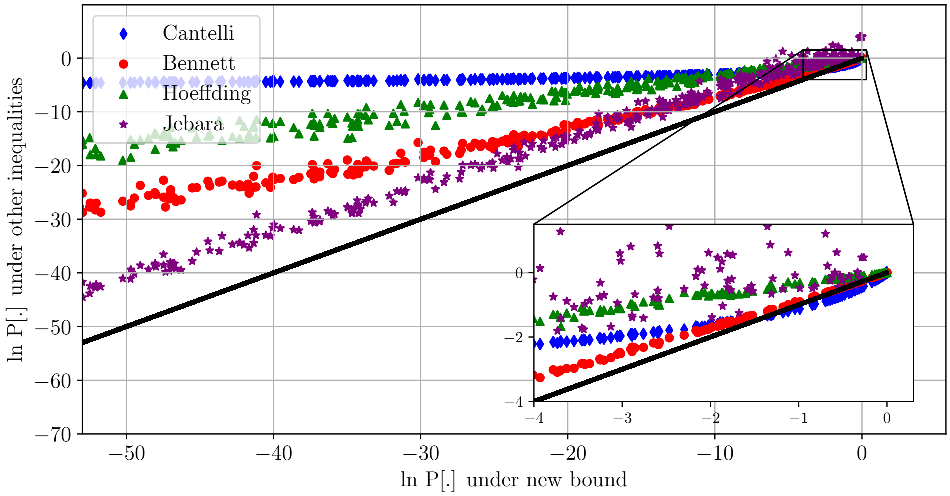
Suppose that $\tau \in [\tau^-, 1]$, then

$$\alpha_\tau \leqslant \bar{b} - \left(\frac{1}{N\Gamma}\ln(\tau/\tau^-)\right)^2 \quad ,$$

where $\Gamma = 1 + (\min_k \gamma_k \min_k (b_k(\gamma_k + 1)))^{-1}$.



(a) Random instances, made of $N = 10$ heterogeneous variables



(b) Random instances, made of $N = 100$ heterogeneous variables

Figure 8.2: Comparison of four bounds with φ_α^*.
If the marker is above the line, then the method gives looser bound compared to φ_α^*.

Proof. Let $\kappa \geqslant 0$ and $\alpha_\kappa = \overline{b} - \kappa$. Then,

$$\begin{aligned}
\varphi_{\alpha_\kappa}\left(\frac{1}{\sqrt{\kappa}}\right) &= \ln(\tau^-) + N\sqrt{\kappa} + \sum_{k=1}^{N} \ln\left(1 + \gamma_k^{-1} e^{-\frac{1}{\sqrt{\kappa}} b_k(\gamma_k+1)}\right) \\
&\leqslant \ln(\tau^-) + N\sqrt{\kappa} + \sum_{k=1}^{N} \gamma_k^{-1} e^{-\frac{1}{\sqrt{\kappa}} b_k(\gamma_k+1)} \\
&\leqslant \ln(\tau^-) + N\left[\sqrt{\kappa} + (\min_k \gamma_k)^{-1} e^{-\frac{1}{\sqrt{\kappa}} \min_k(b_k(\gamma_k+1))}\right] \\
&\leqslant \ln(\tau^-) + N\left[\sqrt{\kappa} + \frac{(\min_k \gamma_k)^{-1}\sqrt{\kappa}}{\sqrt{\kappa} + \min_k(b_k(\gamma_k+1))}\right] \\
&\leqslant \ln(\tau^-) + N\Gamma\sqrt{\kappa}\ .
\end{aligned}$$

The second last inequality is obtained using $e^{-x} \leqslant (1+x)^{-1}$ for $x > 0$. Therefore, for $\kappa^2 \leqslant$

$\frac{1}{N\Gamma}\ln(\tau/\tau^-)$, $\varphi_{\alpha_\kappa}\left(\frac{1}{\sqrt{\kappa}}\right) \leqslant \ln(\tau)$ and as a consequence $\varphi^*_{\alpha_\kappa} \leqslant \ln(\tau) = \varphi^*_{\alpha_\tau}$. As $\alpha \mapsto \varphi^*_\alpha$ is decreasing by Lemma 8.2.1, we obtain that $\alpha_\tau \leqslant \alpha_\kappa$. □ □

Note that, under its apparent simplicity, this property does not hold for other bounds such that Hoeffding or Bennett.

Double bisection search algorithm. We now present a fast algorithm to compute α_τ introduced in Theorem 8.2.2. Algorithm 13 consists of two nested bisection searches. The inner one is dedicated to find the minimum in t – see Algorithm 12 – to an arbitrary precision ϵ_1 and, as a consequence, to compute φ^* to a precision $\Phi\epsilon_1^2$, see Proposition 8.2.3. This estimation of φ^* constitutes the oracle for the outer bisection search. Therefore, the test is more elaborated as it checks whether the decision is sure or not: we only reduce the space by half when the oracle returns value far enough from the target $\ln(\tau)$, i.e., with a distance greater than $\Phi\epsilon_1^2$. If not, then, it means that we obtain at a certain iteration an estimation close enough to $\ln(\tau)$, so we stop at this point. This outer bisection search is a particular case of Probabilistic Bisection Algorithm (PBA) [Hor63; WFH13], where the error term is not necessarily of zero mean but takes values in a small bounded interval.

Algorithm 13 Double Bisection Search for confidence bound's computation

Require: τ, N, b_k, σ_k, ϵ_1, ϵ_2
  $\alpha^-,\ \alpha^+ \leftarrow 0,\ \bar{b} - (\ln(\tau/\tau^-)\,/\,(N\Gamma))^2$ ▷ Init α-bisection
  $tol \leftarrow$ `false`
  while $\alpha^+ - \alpha^- > \epsilon_2$ or $tol =$ `false` **do**
    $\hat{\alpha} \leftarrow \frac{1}{2}(\alpha^- + \alpha^+)$
    $t^-, t^+ \leftarrow 0, -\frac{1}{N(\bar{b}-\hat{\alpha})}\ln(\tau^-)$ ▷ Init t-bisection
    while $t^+ - t^- > \epsilon_1$ **do**
      $\hat{t} \leftarrow \frac{1}{2}(t^- + t^+)$
      $g \leftarrow \frac{d}{dt}\varphi_\alpha(\hat{t})$
      if $g \geqslant 0$ **then** $t^+ \leftarrow \hat{t}$ **else** $t^- \leftarrow \hat{t}$
    end while
    $\hat{\varphi} \leftarrow \varphi_{\hat{\alpha}}(\hat{t})$
    if $\hat{\varphi} > \ln(\tau) + \Phi\epsilon_1^2$ **then** $\alpha^- \leftarrow \hat{\alpha}$
    else if $\hat{\varphi} < \ln(\tau) - \Phi\epsilon_1^2$ **then** $\alpha^+ \leftarrow \hat{\alpha}$
    else $tol \leftarrow$ `true`
    end if
  end while
  return $\hat{\alpha}$

Termination guarantees. The following proposition proves that this algorithm is fast (log convergence) and provides a solution arbitrary close to the optimal solution.

Theorem 8.3.1

Let $\tau \in]\tau^-, 1]$. Algorithm 13 ends with a value $\hat{\alpha}$ such that

$$|\hat{\alpha} - \alpha_\tau| \leqslant \epsilon_2 \wedge \sqrt{\frac{2\Phi}{N\min_k m_k}}\epsilon_1 \wedge \frac{2\Phi}{N\min_k(b_k m_k)}\epsilon_1^2\ , \tag{8.9}$$

where $m_k := \frac{\ln(2+\gamma_k^{-1})}{b_k^2(1+\gamma_k)}$. Moreover, the total number of iterations I_τ is bounded:

$$I_\tau \leqslant \left\lceil \log_2\left(\frac{\bar{b}}{\epsilon_2}\right)\right\rceil \left\lceil \log_2\left(\frac{N\Gamma^2 \ln(1/\tau^-)}{(\ln(\tau/\tau^-))^2\, \epsilon_1}\right)\right\rceil \quad .$$

Proof. At the end of the algorithm, one obtain from the inner bisection that $t^- \leqslant t^*_{\hat{\alpha}}, \hat{t} \leqslant t^+$ and $|t^+ - t^-| \leqslant \epsilon_1$. Suppose that the algorithm ends with a value $\hat{\alpha}$ and $\hat{\varphi} = \varphi_{\hat{\alpha}}(\hat{t})$. Then, from the mean-value theorem, there exists $t \in [t^-, t^+]$ such that

$$\left|\frac{d}{dt}\varphi_\alpha(t^-) - \frac{d}{dt}\varphi_\alpha(t^+)\right| = \left|\frac{d^2}{dt^2}\varphi_\alpha(t)\right| (t^+ - t^-) \leqslant \Phi\epsilon_1 \quad .$$

As the derivative of φ_α is decreasing and positive (resp. negative) in t^- (resp. t^+), $|\frac{d}{dt}\varphi_\alpha(t)| \leqslant \Phi\epsilon_1$ for all $t \in [t^-, t^+]$, and using once again the mean-value theorem,

$$|\hat{\varphi} - \varphi^*_{\hat{\alpha}}| \leqslant \Phi\epsilon_1^2 \quad .$$

1st case: the algorithm ends with $tol = \texttt{true}$.
Therefore (criteria), $|\hat{\varphi} - \varphi^*_{\alpha_\tau}| \leqslant \Phi\epsilon_1^2$, and so $|\varphi^*_{\hat{\alpha}} - \varphi^*_{\alpha_\tau}| \leqslant 2\Phi\epsilon_1^2$. Using Lemma 8.2.1, item (iii), we obtain

$$|\hat{\alpha} - \alpha_\tau| \leqslant \frac{2\Phi\epsilon_1^2}{N \min\{t^*_{\alpha_\tau}, t^*_{\hat{\alpha}}\}} \quad .$$

Then, using Lemma 8.2.1, item (ii), for all α,

$$t^*_\alpha \geqslant \min_{k|b_k>\alpha} \left\{\frac{1}{b_k(1+\gamma_k)} \ln\left(\frac{\alpha + b_k\gamma_k}{\gamma_k(b_k - \alpha)}\right)\right\} \quad .$$

By concavity, $\ln(1+x) \geqslant \ln(1+z)\min\{x/z, 1\}$ for all $x, z \geqslant 0$. Therefore, for all k such that $b_k > \alpha$ (it exists otherwise $\alpha > \bar{b}$), we obtain:

$$\ln\left(\frac{\alpha + b_k\gamma_k}{\gamma_k(b_k-\alpha)}\right) = \ln\left(1 + \frac{\alpha(1+\gamma_k)}{\gamma_k(b_k-\alpha)}\right) \geqslant \ln\left(2 + \frac{1}{\gamma_k}\right)\min\{\frac{\alpha}{b_k - \alpha}, 1\} \geqslant \frac{1}{b_k}\ln\left(2+\frac{1}{\gamma_k}\right)\min\{\alpha, b_k\} \quad .$$

Then, $t^*_\alpha \geqslant \alpha \min_k\{m_k\} \vee \min_k\{b_k m_k\}$. Therefore, as $\hat{\alpha}$ and α_τ are positive quantities,

$$\begin{aligned} t^*_{\alpha_\tau} \vee t^*_{\hat{\alpha}} &\geqslant \alpha_\tau \min_k\{m_k\} \vee \hat{\alpha}\min_k\{m_k\} \vee \min_k\{b_k m_k\} \\ &\geqslant |\hat{\alpha} - \alpha_\tau| \min_k\{m_k\} \vee \min_k\{b_k m_k\} \quad . \end{aligned}$$

Finally,

$$|\hat{\alpha} - \alpha_\tau| \leqslant \max\left\{\sqrt{\frac{2\Phi}{N\min_k m_k}}\epsilon_1, \frac{2\Phi}{N\min_k(b_k m_k)}\epsilon_1^2\right\} \quad .$$

2nd case: the algorithm ends with $|\alpha^+ - \alpha^-| \leqslant \epsilon_2$.
Then, as $tol = \texttt{false}$, at each iteration, $|\hat{\varphi} - \ln(\tau)| \geqslant \Phi\epsilon_1^2$, and so α_τ lies in $[\alpha^-, \alpha^+]$. Therefore, $|\hat{\alpha} - \alpha_\tau| \leqslant |\alpha^+ - \alpha^-| \leqslant \epsilon_2$.

Besides, denoting by I_t the number of iterations for the inner bisection search, we have

$$I_t \leqslant \left\lceil \log_2\left(\frac{-\ln(\tau^-)}{N(\bar{b} - \alpha)\epsilon_1}\right)\right\rceil \quad .$$

Then, as $\bar{b} - \alpha \geqslant \left(\frac{1}{N\Gamma} \ln(\tau/\tau^-)\right)^2$ (see Lemma 8.3.1), $I_t \leqslant \left\lceil \log_2 \left(\frac{N\Gamma^2 \ln(1/\tau^-)}{(\ln(\tau/\tau^-))^2 \epsilon_1} \right) \right\rceil$. Furthermore, denoting by I_α the number of iterations for the outer bisection search, we have

$$I_\alpha \leqslant \left\lceil \log_2 \left(\frac{\bar{b}}{\epsilon_2} \right) \right\rceil \ .$$

□ □

Theorem 8.3.1 proves that Algorithm 13 is fast (log convergence), and provides a solution with an arbitrary precision. Note that the number of iterations is impacted by the distance of τ from the minimal value τ^-. In fact, very close to τ^-, the minimizer t^*_α tends to $+\infty$, and therefore the width of the bisection search space becomes large. Nonetheless, for reasonable error tolerance τ, the algorithm takes very few iterations. Meanwhile, the precision of $\hat{\alpha}$ does not depend on τ.

Numerical experiments. We use the instances developed in Table 8.1. The results are depicted in Figure 8.3 for 1000 realizations, and are fast to obtain (few seconds in total). Of course, the confidence bounds we obtain are larger than the value computed with normal distributions, and so all values are greater than 1. We recover the superiority of the studied bound compared to the standard Bennett's inequality. Besides, Chebyshev-Cantelli's bound is only valuable for low probability level, and becomes inefficient for probabilities close to 1.

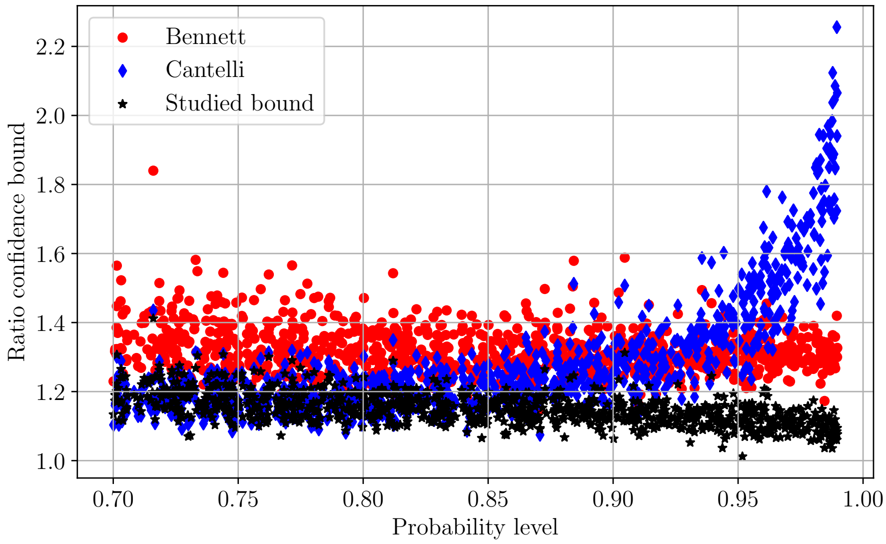


Figure 8.3: Random instances, made of $N = 10$ heterogeneous variables.
For different probability levels $(1 - \tau)$ and different inequalities, we display the value of the confidence bound normalized by the normal case, i.e., the (exact) value for normal distributions.

8.4 Application to Chance-Constrained Programming Problems

In the two next subsections, we will use the proposed concentration inequalities (8.5)-(8.6) into two CCP problems of the form (8.2) with individual bilinear constraints of the form $g(x, \xi) = \xi^T x - b$, with $b \in \mathbb{R}$. The first application (knapsack problems) is of a combinatorial nature and contains only one chance-constraint whereas the second one (Support Vector Machine problems) is a continuous problem but contains as many chance-constraints as training points.

8.4.1 Chance-constrained binary knapsack problem

Let us consider the knapsack problem with stochastic weights $\xi \in \mathbb{R}^N$, see e.g. [CG06; RP21]. To this purpose, we consider a measurable space of outcomes $(\Omega, \mathcal{F})$ and a probability measure F on this space. Then, the problem can be stated as the following chance-constrained problem:

$$\begin{aligned} \max_{y \in \{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \mathbb{P}_F \left[\xi^T y \geqslant c \right] \leqslant \tau \end{aligned} \tag{CKP}$$

where $\tau < 1$, $\pi \in \mathbb{R}^N$ denotes the utility of each item and $c \in \mathbb{R}$ is the maximum budget. The stochastic weights are here represented by a random vector $\xi : \Omega \to \Xi \subseteq \mathbb{R}^N$ and are supposed to be indepedent. If ξ follows a normal distribution, i.e., $F = \mathcal{N}(\mu, \sigma^2)$ with $\mu, \sigma \in \mathbb{R}^N$, problem (DKP) can be reformulated as a SOCP problem, see e.g. [Pré95, Theorem 10.4.1] :

$$\begin{aligned} \max_{y \in \{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \Phi^{-1}(1-\tau)\sqrt{y^T \Sigma y} + \mu^T y \leqslant c \end{aligned} \tag{CKP-N}$$

where Φ is the cumulative distribution for the standard normal distribution and $\Sigma = \operatorname{diag}(\sigma^2)$. In this specific setting, efficient algorithms have been developed in order to solve problem (CKP-N), see e.g. [Han+15].

Here, we rather consider that F is unknown but its first two moments (also denoted by $\mu \in \mathbb{R}^N$ and $\sigma^2 \in \mathbb{R}^N$) are available, as well as maximal deviation (denoted by $b \in \mathbb{R}^N_{\geqslant 0}$). Therefore, F is no longer a normal distribution but belongs to the distributional uncertainty set $\mathcal{D}(\mu, \sigma, b)$ defined as

$$\mathcal{D}(\mu, \sigma, b) = \left\{ F \;\middle|\; \begin{array}{l} \mathbb{P}_F \left[|\xi_i - \mu_i| \leqslant b_i \right] = 1, \} \\ \mathbb{E}_F[\xi_i] = \mu_i, \; i = \{1, \dots, N\} \\ \operatorname{Var}(\xi_i) \leqslant \sigma_i^2 \end{array} \right\} . \tag{8.10}$$

The distributionally robust version of (DKP) is then obtained by looking at the most constraining choice of $F \in \mathcal{D}(\mu, \sigma, b)$:

$$\begin{aligned} \max_{y \in \{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \sup_{F \in \mathcal{D}(\mu,\sigma,b)} \mathbb{P}_F \left[\xi^T y \geqslant c \right] \leqslant \tau \end{aligned} \tag{DKP}$$

This problem has been firstly studied by Calafiore and El Ghaoui [CG06], where they focused on Hoeffding-type *valid conservative approximations* (i.e., approximations whose feasible set is included in the one of (DKP)). Recently, Ryu and Park [RP21] proposed to repeatedly solve binary knapsack subproblems to deduce bounds on SOCP approximations which are classically obtained by considering Chebyshev-Cantelli's inequality. As we will compare the different approximations in the sequel, we first recall two classical results: let us define $B = \operatorname{diag}(b^2)$, then

(i) (Hoeffding) the problem (DKP-H) is a valid conservative approximation of (DKP)

$$\begin{aligned} \max_{y \in \{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \sqrt{2\ln(1/\tau)}\sqrt{y^T B y} + \mu^T y \leqslant c \end{aligned} \tag{DKP-H}$$

8. Bounds for CCP

(ii) (Chebyshev-Cantelli) the problem (DKP-C) is a valid conservative approximation of (DKP)

$$\begin{aligned} \max_{y\in\{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \sqrt{\tfrac{1}{\tau}-1}\sqrt{y^T\Sigma y}+\mu^T y \leqslant c \end{aligned} \tag{DKP-C}$$

This comparison is inspired by the work of Peng, Maggioni and Lisser [PML22] where first-order bounds (Hoeffding and an approximation of Bernstein bound) are used in the continuous knapsack problem, and compared to exact SOCP relaxation for normal variables.

Using Theorem 8.2.1 with $d := c - \mu^T y$, we obtain a tighter conservative approximation of (DKP) formulated as the following problem:

$$\begin{aligned} \max_{\substack{y\in\{0,1\}^N \\ t\geqslant 0}} \quad & \pi^T y \\ \text{s.t} \quad & t\left[\mu^T y - c\right] + \sum_{k=1}^N \ln\left(\frac{\gamma_k e^{t y_k b_k} + e^{-t y_k b_k \gamma_k}}{1+\gamma_k}\right) \leqslant \ln(\tau) \end{aligned} \tag{8.11}$$

The constraint contains bilinear terms ty_k. A naïve approach could be to consider a Fortet linearization of the bilinear terms, see e.g. [For60]. Here, we can reformulate the constraint by considering the change of variable $z := 1/t$ and by dividing the latter constraint by t:

$$(8.11) \iff \begin{cases} \max\limits_{\substack{y\in\{0,1\}^N \\ z\geqslant 0}} \quad \pi^T y \\ \text{s.t} \quad \mu^T y + \sum\limits_{k=1}^N z \ln\left(\dfrac{\gamma_k e^{\frac{y_k}{z} b_k} + e^{-\frac{y_k}{z} b_k \gamma_k}}{1+\gamma_k}\right) \leqslant c + z\ln(\tau) \end{cases} \tag{$\overline{\text{DKP}}$}$$

Note that this transformation is possible because $t > 0$ at the optimum as long as $\tau < 1$ (the left hand-side of the constraint in (8.11) is equal to zero when $t = 0$). Also, $z = 0$ never produces an optimal solution. Therefore, we can consider $z \geqslant 0$ in ($\overline{\text{DKP}}$).

Proposition 8.4.1

For every $\gamma, b \geqslant 0$, the function $\Psi^+_{\gamma,b} : \mathbb{R}\times\mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, defined as

$$\Psi^+_{\gamma,b}(y,z) := z\ln\left(\frac{\gamma e^{\frac{y}{z}b} + e^{-\frac{y}{z}b\gamma}}{1+\gamma}\right) \quad ,$$

is jointly convex. Therefore, the problem (8.11) – without the integrity condition – is convex.

Proof. From elementary calculation, the gradient and the Hessian of $\Psi^+_{\gamma,b}$ are respectively

$$\nabla\Psi^+_{\gamma,b}(y,z) = \left[b - \frac{b(1+\gamma)}{1+\gamma e^{\frac{y}{z}b(1+\gamma)}}, \quad \ln\left(\frac{\gamma + e^{-\frac{y}{z}b(1+\gamma)}}{1+\gamma}\right) + \frac{b(1+\gamma)y}{z\left(1+\gamma e^{\frac{y}{z}b(1+\gamma)}\right)}\right]^T$$

$$H_{\Psi^+_{\gamma,b}}(y,z) = \gamma e^{\frac{y}{z}b(1+\gamma)}\left(\frac{b(1+\gamma)}{1+\gamma e^{\frac{y}{z}b(1+\gamma)}}\right)^2 \begin{bmatrix} \frac{1}{z} & -\frac{y}{z^2} \\ -\frac{y}{z^2} & \frac{y^2}{z^3} \end{bmatrix}$$

As $\mathrm{Tr}\left(H_{\Psi^+_{\gamma,b}}(y,z)\right) \geqslant 0$ and $\det\left(H_{\Psi^+_{\gamma,b}}(y,z)\right) = 0$, the Hessian is always a positive semi-definite matrix, and the function is jointly convex. □ □

Remark 8.4.1

We can directly obtain the convexity of the function by noting that $\Psi^+_{\gamma,b}$ is the composition of the perspective function [Com17] with a particular log-sum-exp function. This proposition is generalized in [NS07] to a class of moment-generating function's estimators. Nonetheless, we make explicit the gradient and the Hessian of the function, since it will be necessary for numerical optimization.

By Proposition 8.4.1, Problem (DKP) reduces to a convex Mixed-Integer Non-Linear Programming (MINLP) problem, since there is a unique convex nonlinear constraint (the budget one). For such problems with a moderate degree of nonlinearity, cutting-plane methods [WP95] are known to be efficient. In this approach, the nonlinear term is approximated by a set of linear constraints (*cutting-plane*), incrementally built by adding at each iteration a new cutting-plane of the original constraint. The subproblem at iteration j is then expressed as:

$$\left(y^{(j+1)}, z^{(j+1)}\right) = \operatorname*{arg\,max}_{\substack{y\in\{0,1\}^N \\ z\geqslant 0}} \left\{ \pi^T y \;\middle|\; \mu^T y + \sum_{k=1}^{N} \left\langle s_k^{(i)}, \begin{pmatrix} y_k \\ z \end{pmatrix} \right\rangle \leqslant c + z\ln(\tau),\ 1 \leqslant i \leqslant j \right\} \ , \tag{8.12}$$

where $s_k^{(i)} := \nabla \Psi^+_{\gamma_k, b_k}\left(y_k^{(i)}, z^{(i)}\right)$. The convergence of this approach has been proved, see e.g. [Bon+06, Theorem 9.6], and is achieved in a finite number of steps for this particular problem (there is a finite number of knapsack-filling scenarios). Note that the cuts are dynamically added in the branch-and-bound at each node where an integer solution is found (*Lazy constraint*), so that the solver does not need to perform a complete Mixed-Integer Linear Programming (MILP) problem solving at each iteration.

As an alternative to the model (8.11), we also introduce a convex reformulation of the problem under Bernstein's inequality: a valid conservative estimation of (DKP) can be obtained by replacing the probabilistic constraint by Bernstein's inequality (see e.g. [BLB04] for more details on this inequality). We obtain the following formulation:

$$\begin{aligned} \max_{y\in\{0,1\}^N} \quad & \pi^T y \\ \text{s.t} \quad & \exp\left(- \frac{\frac{1}{2}(c-\mu^T y)^2}{\sum_{k=1}^N y_k^2 \sigma_k^2 + \frac{1}{3} z (c - \mu^T y)} \right) \leqslant \tau \\ & z \geqslant b_k y_k,\ 1 \leqslant k \leqslant N \end{aligned} \tag{DKP-B}$$

Proposition 8.4.2

Problem (DKP-B) is equivalent to the following problem:

$$\begin{aligned} \max_{\substack{y\in\{0,1\}^N \\ z\geqslant 0}} \quad & \pi^T y \\ \text{s.t} \quad & \tfrac{1}{3}\ln(1/\tau) z + \sqrt{(y^T\ z)\,\Lambda \begin{pmatrix} y \\ z \end{pmatrix}} + \mu^T y \leqslant c \\ & z \geqslant b_k y_k,\ 1 \leqslant k \leqslant N \end{aligned} \tag{8.13}$$

where $\Lambda = \operatorname{diag}\begin{bmatrix}(2\ln(1/\tau)\sigma_k^2)_{1\leqslant k\leqslant N} \\ \frac{1}{9}\ln^2(1/\tau)\end{bmatrix}$. Therefore, problem (DKP-B) – without the integrity condition – is convex.

Proof. We reformulate the constraint so that we end up with a convex reformulation:

$$\begin{aligned}
&\exp\left(-\frac{\frac{1}{2}t^2}{\sum_{i=1}^N y_k^2\sigma_k^2 + \frac{1}{3}zt}\right) \leqslant \tau, \quad t = c - \mu^T y,\ z = \max_k\{b_k y_k\} \\
\iff & \ln(1/\tau)\left[\sum_{i=1}^N y_k^2\sigma_k^2 + \tfrac{1}{3}zt\right] \leqslant \frac{1}{2}t^2 \\
\iff & \ln(1/\tau)\sum_{i=1}^N y_k^2\sigma_k^2 \leqslant \frac{1}{2}\left[t - \tfrac{1}{3}\ln(1/\tau)z\right]^2 - \tfrac{1}{18}\ln(1/\tau)^2 z^2 \\
\iff & \sqrt{2\ln(1/\tau)\sum_{i=1}^n y_k^2\sigma_k^2 + \tfrac{1}{9}\ln^2(1/\tau)z^2} \leqslant c - \mu^T y - \tfrac{1}{3}\ln(1/\tau)z \\
\iff & \sqrt{\begin{pmatrix} y^T & z\end{pmatrix}\Lambda\begin{pmatrix} y \\ z\end{pmatrix}} \leqslant c - \mu^T y - \tfrac{1}{3}\ln(1/\tau)z
\end{aligned}$$

Note that, in the optimization problem, it is sufficient to consider $z \geqslant \max_k\{b_k y_k\}$ as the optimization will search for the lowest value possible (z only appears on the constraint above), ans so the constraint will be naturally saturated. □ □

The SOCP formulation introduced in Proposition 8.4.2 provides an alternative conservative approximation, which can be directly compared to the classical Chebyshev approximation, as they both belong to the same class of problem. In contrast, the formulation (8.11) is not expressed as a cone programming, but we provide in Section 8.7 a reformulation where constraints are expressed via exponential cones.

Remark 8.4.2

We already know (proved theoretically above and highlighted by Figure 8.2) that (8.11) gives better solution than all other formulations (apart from exact one in the case of normally distributed weights), as the set of admissible solutions is larger. Note also that we did not provide optimization model for the bound developed in [Jeb18] and [Ben62], as a convex expression of the chance-constraint is all but immediate to obtain (if it exists).

Numerical results. In order to obtain chance-constrained instances, we adapted deterministic instances from the literature[1], see [Pis05], by adding a maximum standard deviation of 5% of the original weight (taken as mean value), and setting the maximum value to $b = 5\sigma$. Note that for normal distribution, the probability of exceeding $\mu + 3\sigma$ is 0.997. Finally, the maximum probability error τ is taken to 3%.

We use `Cplex v12.10` as a MILP solver and the tests are performed on a laptop `Intel Core i7 @2.20GHz` $\times$ `12`. The MIP gap tolerance is taken to 0.001% and the Integrity tolerance to 1e-8. The tests show that the cutting-plane method adds very few cuts. For instance, the solver added 190 cuts for the instance `1_10000`.

[1]The instances are extracted from the website `http://artemisa.unicauca.edu.co/~johnyortega/instances_01_KP/`. We use the set of instances `knapPI_{X}_1000_1` where `X` goes from `1_100` to `2_10000` (the second number stands for the number of items in the instances), see Table 8.2.

Instance	KP	(CKP-N)	(8.11)	Prob.	Time	(DKP-B)	(DKP-C)	(DKP-H)
1_100	9147	8842	8817	0.19	0.1	8719	8817	8150
1_200	11238	11227	10962	0.81	0.1	10682	10832	10353
1_500	28857	28606	28405	2.11	0.4	28152	28127	27924
1_1000	54503	54105	53836	1.58	0.65	53617	53267	52109
1_2000	110625	110130	109779	2.95	1.8	109621	109148	107228
1_5000	276457	275685	275220	2.99	33.4	275068	274151	*271160*
1_10000	563647	562560	561968	3.00	97.4	561809	560387	*556126*
2_100	1514	1513	1512	0.82	0.1	1456	1476	1395
2_200	1634	1619	1594	0.69	0.2	1558	1592	1508
2_500	4566	4537	4504	2.31	0.5	4472	4472	4348
2_1000	9052	9008	8970	2.87	1.52	8951	8927	8761
2_2000	18051	17991	17946	2.85	4.0	17925	17872	*17635*
2_5000	44356	44262	44201	2.86	32.7	44184	44073	*43696*
2_10000	90204	90071	89996	2.99	84.2	89975	89807	*89265*

Table 8.2: Results for knapsack instances. We compare the two new formulations (8.11) and (DKP-B) to the existing methods (DKP-H) and (DKP-C). The method KP corresponds to the deterministic case, and (CKP-N) corresponds to a normally-distributed uncertainty. For (8.11), we also provide the probability error and the computational time. When the objective is in italic, the solver does not succeed to prove the optimality in the given time.

The numerical tests show the efficiency of the proposed relaxation: the use of the second-order information leads to a substantial improvement of the optimal objective-function value, compared to the classical Hoeffding bound. Besides, this method appears to be easy tractable, as we were able to solve instances of 10000 items in less than two minutes. Note that the relaxation seems to be a bit more tractable than the Hoeffding bound, as the solver cannot prove the optimality of the solution with the desired precision in less than 10 minutes.

Remark 8.4.3

We present here the results for mixed-integer problems, but the results and the methodology does not exploit the integrity condition of the variables, and so the results and the methodology are still applicable on the (simpler) continuous problem. In particular, a cutting-plane approach still converges.

8.4.2 Distributionally Robust Support Vector Machine problem

Let us consider a dataset of M points (i.e., M pairs of features and labels) $\{\xi_i, l_i\}$ where each (stochastic) feature $\xi_i : \Omega \to \Xi_i \subseteq \mathbb{R}^N$ is distributed according to F_i. The features are categorized/classified into two classes, indexed by labels $l_i \in \{-1, +1\}$. The Support Vector Machine (SVM) problem consists in finding a good *separating* hyperplane for the dataset, i.e., in finding a pair $(w, w_0) \in \mathbb{R}^N \times \mathbb{R}$ such that the hyperplane $h : x \in \mathbb{R}^N \mapsto w^T x + w_0$ separates the "+1"-type features ($h(x_i) > 0$ for all i s.t. $l_i = +1$) from the "-1"-type features ($h(x_i) < 0$ for all i s.t. $l_i = -1$), see e.g. [SC08]. The chance-constrained formulation of the SVM problem (with soft margin), see e.g. [BM92; CV95], is then defined as follows:

$$\begin{aligned} \min_{w\in\mathbb{R}^N, w_0\in\mathbb{R}, \eta\in\mathbb{R}^M_{\geqslant 0}} \quad & \frac{1}{2}\|w\|_2^2 + r\sum_{i=1}^M \eta_i \\ \text{s.t.} \quad & \mathbb{P}_{F_i}\left[l_i(w^T\xi_i + w_0) \leqslant 1-\eta_i\right] \leqslant \tau_i,\ i=1,\dots,M \\ & \eta_i \geqslant 0,\ i=1,\dots,M \end{aligned} \tag{CSVM}$$

This chance-constrained version has been recently studied, see e.g. [WFP15; Kha+22]. Here, we focus on *independent* noises $\{\xi_k\}_{1\leqslant k\leqslant N}$ as in [Ben+10] and study the distributionally robust version of (CSVM) by supposing, once again, that for each F_i, $i\in\{1,\dots,N\}$, belongs to the distributional uncertainty set $\mathcal{D}(\mu_i,\sigma_i,b_i)$ with $\mu_i,\sigma_i,b_i\in\mathbb{R}^N$, see (8.10).

$$\begin{aligned} \min_{w\in\mathbb{R}^N, w_0\in\mathbb{R}, \eta\in\mathbb{R}^M_{\geqslant 0}} \quad & \frac{1}{2}\|w\|_2^2 + r\sum_{i=1}^M \eta_i \\ \text{s.t.} \quad & \inf_{F_i\in\mathcal{D}(\mu_i,\sigma_i,b_i)} \mathbb{P}_{F_i}\left[l_i(w^T\xi_i + w_0) \leqslant 1-\eta_i\right] \leqslant \tau_i,\ i=1,\dots,M \\ & \eta_i \geqslant 0,\ i=1,\dots,M \end{aligned} \tag{DSVM}$$

In contrast with the knapsack problem (DKP), the coordinates of the uncertain features are multiplied by the weight w, which can be either positive or negative. Therefore, we can no longer apply (8.5) and must use (8.6) which contains absolute values. Moreover, each training feature defines a (nonlinear) chance-constraint.

Proposition 8.4.3

Let $(\overline{\text{SVM}})$ be defined as

$$\begin{aligned} \min_{w\in\mathbb{R}^N, w_0\in\mathbb{R}, \eta\in\mathbb{R}^M_{\geqslant 0}} \quad & \frac{1}{2}\|w\|_2^2 + r\sum_{i=1}^M \eta_i \\ \text{s.t} \quad & -l_i(w_0 + w^T\mu_i) + \sum_{k=1}^N \Psi_{\gamma_{ik},b_{ik}}(w_k, z_i) \leqslant \eta_i - 1 + z_i\ln(\tau_i),\ 1\leqslant i\leqslant M \end{aligned} \tag{$\overline{\text{SVM}}$}$$

where $\Psi_{\gamma,b} : \mathbb{R}\times\mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is defined as $\Psi_{\gamma,b}(y,z) = \Psi^+_{\gamma,b}(|y|,z)$. This problem is a valid conservative approximation of (DSVM).

Proof. We follow the similar steps as for the knapsack case. In particular, we use Theorem 8.2.1 for each chance-constraint $i\in\{1,\dots,M\}$ with $\lambda = -l_i w$ and $d = \eta_i - 1 + l_i(w_0+\mu)$. Then, we apply the change of variable $z = 1/t$. □ □

The following proposition shows that the function Ψ keeps the same regularity as Ψ^+:

Proposition 8.4.4

The function $\Psi_{\gamma,b}$ is convex and twice continuously differentiable. Moreover,

$$\Psi_{\gamma,b}(y,z) = \begin{cases} \Psi^+_{\gamma,b}(y,z), & y\geqslant 0 \\ \Psi^+_{\gamma^{-1},b\gamma}(y,z), & y\leqslant 0 \end{cases} \tag{8.14}$$

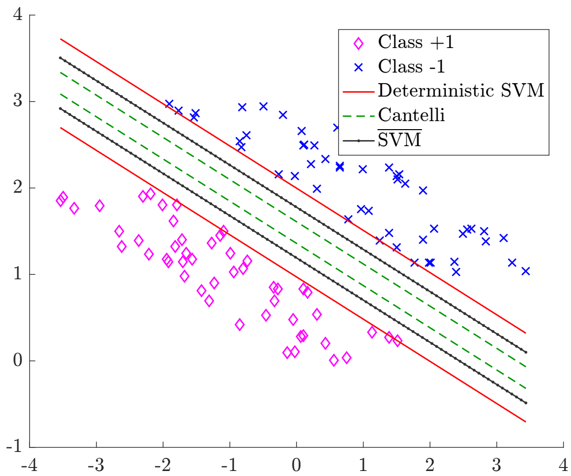
Consequently, problem ($\overline{\text{SVM}}$) is a convex conservative approximation of (DSVM).

Proof. For $y \geqslant 0$, we have the following direct equalities:

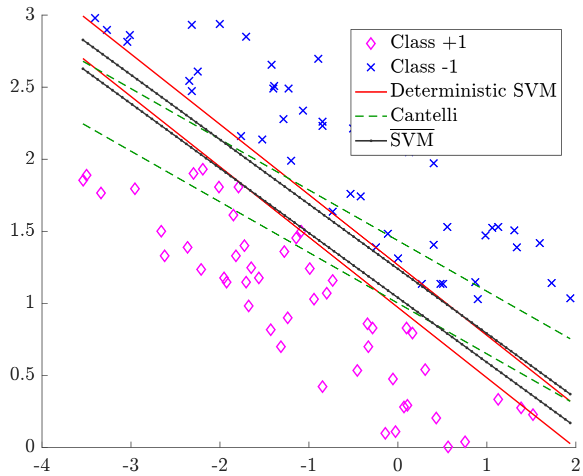
$$\Psi^+_{\gamma,b}(-y,z) = z\ln\left(\frac{e^{-\frac{y}{z}b} + \gamma^{-1}e^{\frac{y}{z}b\gamma}}{1+\gamma^{-1}}\right) = z\ln\left(\frac{e^{-\frac{y}{z}(b\gamma)\gamma^{-1}} + \gamma^{-1}e^{\frac{y}{z}(b\gamma)}}{1+\gamma^{-1}}\right) = \Psi^+_{\gamma^{-1},b\gamma}(y,z) \ .$$

Furthermore, to check the regularity property, it suffices to verify the condition in $y=0$. It holds that $\nabla\Psi^+_{\gamma,b}(0,z) = 0$ for all $\gamma \in \mathbb{R}$ and $b \in \mathbb{R}_{>0}$, so $\nabla\Psi_{\gamma,b}(0^-,z) = \nabla\Psi_{\gamma,b}(0^+,z) = 0$. Moreover, $H_{\Psi^+_{\gamma,b}}(0,z) = b^2\gamma\begin{bmatrix}\frac{1}{z} & 0\\ 0 & 0\end{bmatrix} = H_{\Psi^+_{\gamma^{-1},b\gamma}}(0,z)$. Therefore, $\Psi_{\gamma,b}$ is twice continuously differentiable in $y=0$. □ □

Numerical results. In the tests, ($\overline{\text{SVM}}$) is implemented using the interior-point nonlinear solver `IPOPT`[2]. The solver always returns the optimal solution as the problem has been proved to be convex, see Proposition 8.4.4. For comparison, we also implement the robust SVM approximation using Chebyshev-Cantelli inequality – see e.g. [WFP15] – which can be efficiently solved by any SOCP solver.



(a) Robust separation
(no slack variables used)

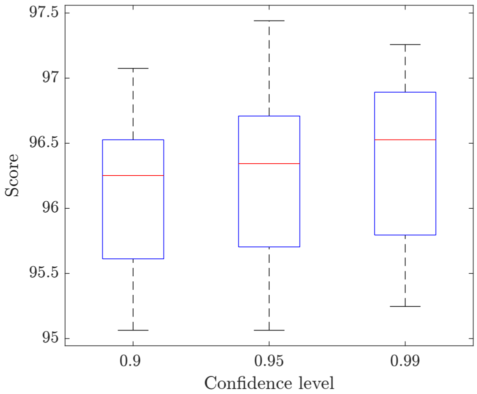


(b) Robust separation
(active slack variables for some points)

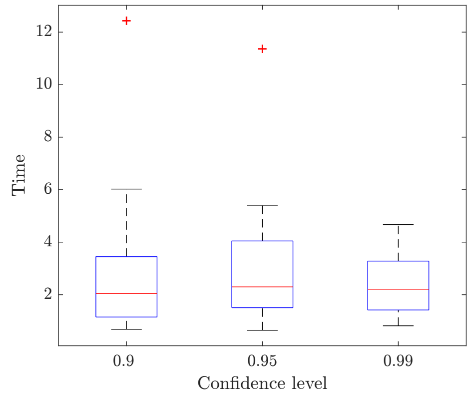
Figure 8.4: Two-dimensions SVM with $\tau = 0.02$, $M = 100$, $r = 100$.
We directly represent the margin around the hyperplane (centered between the two lines)

First, we construct 2D instances with linearly separable classes. To estimate the uncertainty of each feature, we follow the method of [WFP15] by calculating the standard deviation of the training features for each class and then divide by 10. This appears to be reasonable as an uncertainty set for each data point. The results are displayed in Figure 8.4. On the left (see Figure 8.4a), the classes are sufficiently distant so that it is not necessary to activate slack variables η_i. We observe that all the methods find the same hyperplane, but differ on the size of the margin width. As expected, the Chebyshev-Cantelli' inequality is more conservative on this example. On Figure 8.4b, we reduce the space between the two classes. The features are still linearly separable in the deterministic setting, but are not robustly separable both for Chebyshev and for the proposed method. Nonetheless, we numerically observe that Cantelli relaxation needs to activate more slack variables, and so the optimal value is greater than the one found by the proposed method.

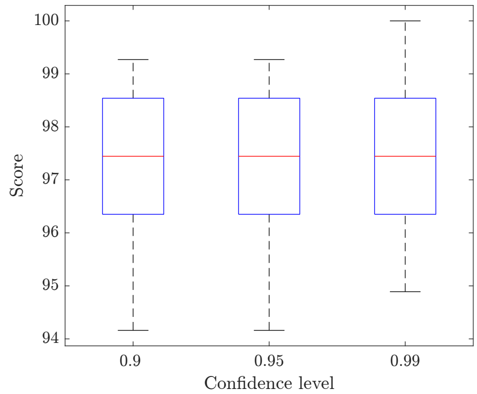
[2] `https://coin-or.github.io/Ipopt/`



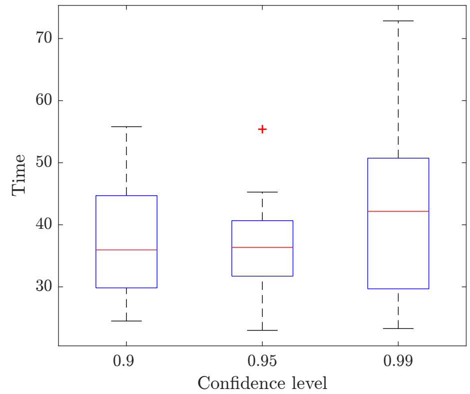
(a) Score with 20% of training points



(b) Time (s) with 20% of training points



(c) Score with 80% of training points



(d) Time (s) with 80% of training points

Figure 8.5: Two-dimensions SVM with $\tau = 0.02$, $M = 100$, $C = 100$

We then use the proposed method on instances from the literature. In particular, we use data on Breast Cancer in the Wisconsin[3]. This dataset contains 683 samples of dimension 10, see e.g. [WFP15; Kha+22] for more information on the dataset. Figure 8.5 displays the time and the score (the percentage of test data that satisfies the classification obtained with the training set) for two configurations. We observe that the mean score is always higher than 96% (same order as in [WFP15; Kha+22]), and the time stays reasonable even for a substantial number of features (more than 500 training points, see Figures 8.5b and 8.5d).

8.5 Conclusion and perspectives

In this chapter, we studied a refined Bennett-type inequality, originally developed in the homogeneous setting and extended here to the heterogeneous case. We have shown that this concentration inequality can be used in a wide range of applications. First, we introduce a double bisection search which computes (in logarithmic time) confidence bounds proved to be tighter than other classical approaches. In particular, it outperforms the standard Chebyshev's approach for high probability precision. Besides, we obtained tight distributionally robust bounds for individual CCP problems which can be formulated as convex problem. In particular, we highlighted that the inequality can be inserted into CCP binary knapsack problem while staying

[3]https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29

tractable (instances of 10 000 binary variables). Tests on SVM problems have also been performed, obtaining a better separability of the data on instances from the literature (containing up to 500 points).

Future works will be dedicated to the extension of the results to the independent joint probability constraint case. Moreover, we think that this inequality can be helpful in many concrete applications to estimate more precisely error bounds, especially we will focus on electricity bill estimates.

8.6 Proofs

8.6.1 Comparison of moment-generating function estimations from the literature

[Ben62, (c)]⇒[Ben62, (b)]. Suppose that $\xi - \mathbb{E}[\xi] \leqslant b$ and $\mathrm{Var}(\xi) \leqslant \sigma^2$. Then, in [Jeb18], the upper estimator of the moment-generating function is $J(t) = 1 + \gamma \left(e^{tb} - 1 - tb\right)$, where $\gamma = (\sigma/b)^2$. Besides, in [DZ10], the upper estimator is

$$D(t) := \frac{\gamma e^{tb} + e^{-t\gamma b}}{1+\gamma} \quad .$$

If now we consider D as a function of γ, i.e., $D(t,\gamma) = D(t)$, then, the second partial derivative w.r.t. γ is $\partial^2_\gamma D(t,\gamma) = \frac{2}{(1+\gamma)^3}\left[e^{\gamma t} - e^t\right] \leqslant 0$. Therefore, $D(t,\cdot)$ is concave for any fixed $t \geqslant 0$ and

$$D(t,\gamma) \leqslant D(t,0) + \gamma \partial_\gamma D(t,0) = J(t) \quad .$$

[Ben62, (c)]⇒[Pin89]⇒[Hoe63]. As $\gamma \mapsto D(t,\gamma)$ is increasing, then $D(t,\gamma) \geqslant D(t,1) = \cosh(tb)$, which is exactly the bound obtained by Pinter with $a = b$. As $\cosh(x) \leqslant \exp(x^2/2)$, we have $D(t,1) \leqslant e^{(tb)^2/2}$, which is exactly the Hoeffding's estimator.

[FS13; Zhe17]⇒[Hoe63]. Now, until the end of the proof, let us suppose that $\xi \in [0,1]$, i.e., $a = -\mathbb{E}[\xi]$ and $b = 1 - \mathbb{E}[\xi]$. We denote by $p = \mathbb{E}[\xi]$ the mean value and by σ^2 the variance. Then, in [Hoe63], the upper estimator of the moment-generating function is $H(t) = e^{tp+t^2/8}$. In [Zhe17] and [FS13], the upper estimator of the moment-generating function $\mathbb{E}[e^{t(\xi-p)}]$ is $Z(t) := 1 + p(e^t - 1)$. By basic algebra, $H'(t) - Z'(t) = \left(p + \frac{t}{4}\right) e^{tp+t^2/8} - pe^t$. Then

$$H'(t) - Z'(t) \geqslant 0 \iff \ln\left(1 + \frac{t}{4p}\right) + t(p-1) + t^2/8 \geqslant 0 \quad .$$

As $\ln(1+x) \geqslant \frac{x}{1+\frac{1}{2}x}$ for $x \geqslant 0$, $H'(t) - Z'(t) \geqslant 0$ if

$$\left[\frac{1}{4p} + p - 1\right] + t\left[\frac{1}{8} + \frac{p-1}{8p}\right] + t^2\left[\frac{1}{8^2 p}\right] \geqslant 0 \quad .$$

The above condition holds since the discriminant of this second-order equation is zero. Therefore, $H'(t) - Z'(t) \geqslant 0$, and since $H(0) = Z(0)$, we finally conclude that $H(t) \geqslant Z(t)$ for $t > 0$.

[CL22]⇒[FS13; Zhe17]. In [CL22], the upper estimator is a family of function C_k such that

$$C_k(t) := 1 + k\left(e^{t/k} - 1\right)(p - \sigma^2 - p^2) + (\sigma^2 + p^2)(e^t - 1) \quad ,$$

One can prove that $\{C_k(t)\}_k$ is decreasing $\forall t \in \mathbb{R}_{\geqslant 0}$, and

$$\lim_{k\to\infty} C_k(t) = C_\infty(t) := 1 + t(p-q) + q(e^t - 1) \quad ,$$

where $q := \sigma^2 + p^2 \leqslant p$. Hence, $C_\infty(t) - Z(t) = (p-q)\left(1 + t - e^t\right) \leqslant 0$.

[Ben62, (b)]⇒[CL22]. For $\xi \in [0,1]$, the upper estimator of [Jeb18] is $J(t) = 1+\gamma(e^{t(1-p)} - 1-t(1-p))$. Using the notation $\theta = (1-p)^2 \in [0,1]$, we express C_∞ and J in (θ,γ)-coordinates:

$$\begin{cases} C_\infty(t,\gamma,\theta) = 1+t(1\text{-}\sqrt{\theta}) + \left[\gamma\theta + (1\text{-}\sqrt{\theta})^2\right]\left[e^t - 1 - t\right] \\ J(t,\gamma,\theta) = 1+\gamma\left(e^{t\sqrt{\theta}} - 1 - t\sqrt{\theta}\right) \end{cases}$$

Now, the partial derivatives w.r.t γ are

$$\begin{cases} \partial_\gamma C_\infty(t,\gamma,\theta) = \theta\left[e^t - 1 - t\right] \\ \partial_\gamma J(t,\gamma,\theta) = e^{t\sqrt{\theta}} - 1 - t\sqrt{\theta} \end{cases}$$

The function $[0,1] \ni \theta \mapsto e^{t\sqrt{\theta}} - 1 - t\sqrt{\theta}$ is convex for any $t \geqslant 0$, and so $\partial_\gamma J \leqslant \partial_\gamma C_\infty$. As $J(t,0,\theta) = 1 \leqslant C_\infty(t,0,\theta)$, we conclude that $J(t) \leqslant C_\infty(t)$.

8.7 Conic reformulation

First, $\Psi_{\gamma,b}$ can also be written $\Psi_{\gamma,b}(y,z) = \max\left\{\Psi^+_{\gamma,b}(y,z), \Psi^+_{\gamma^{-1},b\gamma}(y,z)\right\}$. Therefore,

$$\sum_{k=1}^N \Psi_{\gamma_k,b_k}(y_k,z) \leqslant u \iff \begin{cases} \sum_{k=1}^N v_k \leqslant u \\ \Psi^+_{\gamma_k,b_k}(y_k,z) \leqslant v_k \\ \Psi^+_{\gamma_k^{-1},b_k\gamma_k}(y_k,z) \leqslant v_k \end{cases}$$

Now, denoting the exponential cone by $\mathcal{K}_{exp} = \{(x_1,x_2,x_3) : x_1 \geqslant x_2 e^{x_3/x_2}\}$,

$$\begin{aligned} \Psi^+_{\gamma_k,b_k}(y_k,z) \leqslant v_k &\iff \gamma_k e^{\frac{y_k}{z}b_k} + e^{-\frac{y_k}{z}b_k\gamma_k} \leqslant e^{\frac{v_k}{z}}(1+\gamma_k) \\ &\iff \gamma_k e^{\frac{y_kb_k - v_k}{z}b_k} + e^{\frac{-y_kb_k\gamma_k - v_k}{z}} \leqslant 1+\gamma_k \\ &\iff \begin{cases} \gamma_k\rho_k + \nu_k \leqslant (1+\gamma_k)z \\ (\rho_k, z, y_kb_k - v_k) \in \mathcal{K}_{exp} \\ (\nu_k, z, -y_kb_k\gamma_k - v_k) \in \mathcal{K}_{exp} \end{cases} \end{aligned}$$

This formulation has a number of variables and (conic) constraints of order $O(NM)$. Therefore, this conic reformulation is only valuable for small to medium instance sizes.

Chapter 9

Entropic Lower Bounds of ℓ_0-pseudo-norm for Sparse Optimization

***

Abstract. We introduce a family of cardinality's lower bounds, defined as ratios of norms. We prove that the tightest bound of the family is obtained as a limit case, and involves a Shannon entropy. We then use this entropic lower bound in sparse optimization problems to approximate cardinality requirements. This provides a nonlinear nonconvex relaxed problem, which can be efficiently solved by off-the-shelf nonlinear solvers. In the numerical study, we focus on the case where the optimization is performed on the simplex, and where the classical ℓ_1 penalization does not yield sparse solution. The Finance Index Tracking problem is taken as an example and illustrates the efficiency of the proposed approach.

9.1 Introduction

In numerous fields such as finance, energy or machine learning, decision makers aim to control the cardinality of the solution vector, i.e., the number of representative features (assets in portfolio optimization [BS07], shutdowns/start-ups in thermal power plants scheduling [BZE14a], Support Vectors in machine learning [Bi+03], . . .).

In optimization terminology, the cardinality of a solution encoded by a vector $x \in \mathbb{R}^n$ is the number of non-zero elements, i.e., $|\{i \in [n] : x_i \neq 0\}|$, and is often written $\mathrm{card}(x)$. Both correspond to the so-called *ℓ_0-pseudo-norm*, denoted $\|x\|_0$. This pseudo-norm is positively homogeneous of degree 0, meaning that for all $x \in \mathbb{R}^n$ and $\alpha \neq 0$, $\|\alpha x\|_0 = \|x\|_0$. Optimization under cardinality requirements is called *sparse optimization*.

Sparsity has mainly emerged from the signal processing and machine learning communities, under names such as compressed sensing [Don06] and sparse learning [Bi+03]. In machine learning, the Sparse Support Vector Machine aims at finding a minimal cardinality linear classifier which can separate two classes of labeled data. The sparsity of the solution helps for better interpretability of the solution which is crucial in automated analysis of large text corpora. One major case of using sparsity is the feature selection which refers to the necessity of selecting representative variables from datasets containing a large number of features, many of them being irrelevant or redundant. For example, in finance, feature selection is used to restrict asset allocation to a limited number of assets in the portfolio. Sparsity allows reducing a priori the dimension of a large-scale problem when performing a sparse regression that may be more efficient than the classical one, by selecting a small set of predictors in a least-squares sense.

Sparsity is also very useful in energy management where many problems involve cardinality constraints. Our original motivation is the intra-day problem, consisting in updating a day-ahead generation schedule by modifying a limited number of power units schedules ([BZE14a], [BZE14b]). Two other examples concern operation of power plants. During start-up, some components of thermal power plants go from 20°C to 1300 − 1900°C in a few seconds leading, over the long term, to damages, reducing their lifespan. Saving durability of these plants consists in limiting the number of shutdowns/start-ups. Finally, when operating nuclear power plants, it is necessary to limit the number of “deep” drops in power (because a nuclear reaction at low power for a long time generates unwanted isotopes that “poison the heart”) and also to limit the daily number of production variations (modulations) so as not to over-consume the boron (neutron absorber) because a reduction in the boron available in the core makes the plant more difficult to operate and leads to its premature shutdown for refueling.

Optimization problems involving the ℓ_0-norm of the decision vector belong to the class of sparse optimization problems and take one of the two following general form:

(i) $\|x\|_0$ in the objective function:

$$\min_{x\in X}\{f(x)+\lambda\|x\|_0 \;\mid\; g(x)\leqslant 0\} \tag{P_λ}$$

(ii) $\|x\|_0$ in constraints:

$$\min_{x\in X}\{f(x) \;\mid\; g(x)\leqslant 0,\ \|x\|_0\leqslant k\ (<n)\} \tag{P_k}$$

In both formulations, $X \subseteq \mathbb{R}^n$ is the set defining the constraints. The objective function f corresponds to a given criterion and is often considered as convex in machine learning applications (such as *least-squares problems* (LSQ)) while it may be nonconvex in energy applications. The parameter $\lambda \geqslant 0$ is viewed as a regularization parameter used to manage the trade-off between the criterion $f(x)$ and the sparsity of x. In selection problems, a stronger constraint on the decision vector arises: x must belongs to the probability simplex, i.e., $X \subseteq \Delta_n := \{x \in \mathbb{R}^n_+ \mid \sum_{i=1}^n x_i = 1\}$. In this specific case, the ℓ_1-norm is constrained to be one.

Since the $\|\cdot\|_0$ is lower semicontinuous on $\mathbb{R}^n$ and is discontinuous at any point belonging to an hyperplane $x_i = 0$, optimization problems involving the ℓ_0-norm are nonconvex and hence very challenging. They are inherently of combinatorial nature and hence, not solvable in polynomial-time in general [Bie96]. Huge research effort has been made in sparse optimization and several approaches have been proposed. Let us cite :

- **The convex approximation**. A typical example is the famous Least Absolute Shrinkage and Selection Operator (LASSO) penalty technique. It consists in replacing the nonconvex term $\|x\|_0$ by the convex approximation $\|x\|_1$. This approach has first been proposed for linear regression in [Tib96]. Since then, the ℓ_1-regularization technique has been extensively studied and improved ([GN03], [Zou06], [KF00], ...) This leads to very efficient and scalable algorithms in many cases. For example, the main approaches to sparse learning replace the (hard) cardinality requirements with some simpler (convex) functions such as the ℓ_1-norm, leading to tractable optimization problems. However, in several applications of great interest, in energy for instance, the solutions obtained in this way are generally far from the expected one. Moreover, replacing cardinality by the convex approximation based on ℓ_1-norm is pointless for optimization problems over the probability simplex (selection problems) i.e., when the variables are discrete probability distributions, since in this case the ℓ_1 norm is constant over the feasible set. Then, the now-standard approaches fail and some methods have been specifically dedicated to sparse optimization on simplex, finding alternative convex approximations, for e.g. based on the ℓ_∞-norm [PEC12].

- **The nonconvex approximation**. This approach consists in approximating $\|x\|_0$ by a continuous nonconvex function. Various functions have been proposed to approximate the ℓ_0 term ([BM98], [Fu98], [Wes+03]) and several types of algorithms have been designed to solve related optimization problems, including algorithms based on the Difference of Convex functions (DC) ([CXY10], [GRC09], [GG13], [OL13], [TPL08], [PL14]) or based on Successive or Local Linear Approximation ([BM98], [ZL08]). Nonconvex approximations can be better than convex relaxations by guarantying a higher sparsity level, but the related nonconvex optimization problems are more difficult to solve.

- **Heuristic approach**. In addition to the mathematical programming based approaches, heuristic methods have also been applied, especially greedy algorithms, designed to directly tackle cardinality minimization problem. Two noteworthy examples are the matching pursuit [MZ93] and the orthogonal matching pursuit [PRK93].

Table 9.1 gives some additional entries in the literature.

	Problem	Optimality	Resolution
[BS07]	Sparse LSQ Portfolio selection	Global	Branch & Bound
[Nad+20]	Sparse LSQ	Global	Branch & Bound
[BBN21]	Sparse LSQ	Global	Branch & Bound
[Tib96]	Sparse LSQ	Relaxation	Convex penalization (LASSO)
[SBA15]	Sparse LSQ	Relaxation	Continuous nonsmooth penalty
[HT19]	Sparsity	Relaxation	Nonconvex penalization
[AG19]	Sparse regression	Lower Bound	SDP (convex)
[CD19]	GSO	Lower Bound	Caprac conjugacy
[Sou+11]	Sparse LSQ	Heuristic	Penalization + Greedy

Table 9.1: Different approaches to sparse optimization.
LSQ: Least squares problem
GSO : General Sparse Optimization

In this context, we propose an approach based on constructing a set of lower bounds of ℓ_0-pseudo-norm expressed as ratios of norms (Theorem 9.2.1). In particular, we prove that the best lower bound we obtained is expressed as a function of Shannon entropy [Sha48] and ℓ_1-norm. In [SI16], the authors bring to light sharp extreme relations between Shannon entropy and ℓ_α-norm ($\alpha > 0$). Here, we obtain a relation for $\alpha = 0$. Then, we insert this new bound in sparse optimization problems, and show that the relaxed problem is a smooth nonlinear problem (yet non convex), see proposition 9.2.1. Then, a local solution can be obtained by using a nonlinear solvers like IPOPT [WB06]. Numerical experiments on the Finance Index Tracking problem illustrate the efficiency of the proposed approach (Section 9.4).

9.2 Entropic Lower Bound of $\|x\|_0$ and use in Sparse Optimization

9.2.1 Renyi's entropies

Recall that the Renyi's entropy [Rén+61] of order $\alpha \geqslant 0, \alpha \neq 1$, associated to a discrete distribution $p \in \mathbb{R}^n$, $p \geqslant 0$, $p_1 + \cdots + p_n = 1$, is the quantity:

$$H_\alpha(p) := \left(\frac{1}{1-\alpha}\right) \log \sum_{i=1}^{n} p_i^\alpha .$$

Depending on the value of parameter α, four important special cases of Renyi's entropies can be mentioned:

- Hartley's entropy [Har28] ($\alpha = 0$): $H_0(p) = \log \|x\|_0$.
- Shannon's entropy [Sha48] ($\alpha \to 1$): $H_1(p) = \lim\limits_{\alpha \to 1} H_\alpha(p) = -\sum_{i\in[n]} p_i \log p_i$.
- Collision entropy ($\alpha = 2$): $H_2(p) = -\log \sum_{i\in[n]} p_i^2 = -\log \|p\|_2^2$.
- Minimal entropy ($\alpha \to \infty$): $H_\infty(p) = \lim\limits_{\alpha\to\infty} H_\alpha(p) = -\log \|p\|_\infty$.

In the case of a uniform probability distribution, the Rényi entropies of all orders, the Hartley's entropy and the Shannon entropy coincide.

The natural logarithm of ℓ_0-*pseudo-norm* of a vector $x \in \mathbb{R}^n$ is the Hartley's entropy, a measure of uncertainty [Har28], corresponding to the information provided by selecting, randomly and uniformly, a sample from x.

9.2.2 A hierarchy of lower bounds

We define the ℓ_q-normof a vector $x \in \mathbb{R}^n$, $p \geqslant 1$, as:

$$\|x\|_q = \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}} .$$

We remind the known lower bounds of $\|x\|_0$ as ratios of norms ($\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$):

$$B_\infty(x) := \frac{\|x\|_1}{\|x\|_\infty} \leqslant \|x\|_0 \tag{9.1}$$

$$B_2(x) := \left(\frac{\|x\|_1}{\|x\|_2}\right)^2 \leqslant \|x\|_0 . \tag{9.2}$$

These lower bounds may be far from $\|x\|_0$ in practice.

We now introduce a family of bounds generalizing the two previous bounds: for $x \neq 0$, and $\alpha > 0$, define

$$B_\alpha(x) := \left(\frac{\|x\|_1}{\|x\|_\alpha}\right)^{\frac{\alpha}{\alpha-1}} = \exp H_\alpha(p(x)) = \left(\sum_{i\in[n]} p_i(x)^\alpha\right)^{\frac{1}{\alpha-1}} , \quad p(x) := |x|/\|x\|_1 .$$

In particular,

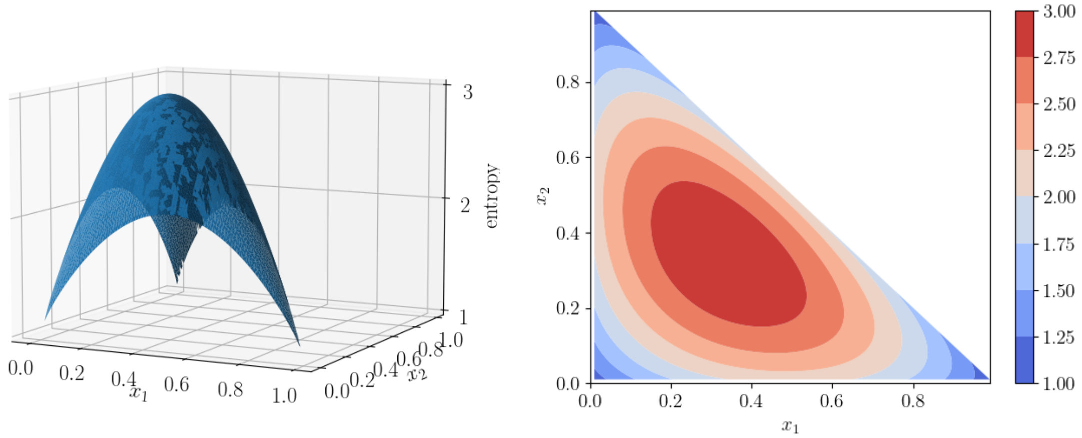
$$B_1(x) = \frac{\|x\|_1}{\prod_{i\in[n]} |x_i|^{|x_i|/\|x\|_1}} = \|x\|_1 \exp\left(-\frac{1}{\|x\|_1}\sum_{i\in[n]} |x|_i \log |x|_i\right) \ . \tag{9.3}$$

Theorem 9.2.1 recalls that the family $(B_\alpha)_{\alpha\in]0,+\infty[}$ is ordered in a decreasing fashion, so that the quality of the bound improves when α decreases.

Theorem 9.2.1 (Monotonicity according to order α, see e.g. [Cac97])

$$B_\infty(x) \leqslant \ldots \leqslant B_2 \leqslant \ldots \leqslant B_1 \leqslant \ldots \leqslant B_0 = \|x\|_0 \ . \tag{9.4}$$

In the case $\|x\|_1 = 1$, B_1 simplifies to the exponential of the Shannon entropy. We refer to Figure 9.1 for a numerical example of the bound B_1. This illustrates, in particular, the concavity of this nonlinear bound.



(a) Entropy in 3 dimensions (b) Entropy as a contour plot

Figure 9.1: Shannon entropy $H_1(x)$ for $x \in \Delta_3$. The two first dimensions x_1 and x_2 are displayed, and the third one is implicitly defined as $x_3 = 1 - x_1 - x_2$.

9.2.3 Sparse optimization and focus on Shannon entropy

We now focus on the integration of the previously defined entropic bound (9.3) in a sparse optimization problem: let us assume a generic problem of the form (P_k). The corresponding relaxation is then

$$\begin{aligned} \min_{x\in X} \quad & f(x) \\ \text{s.t.} \quad & g(x) \leqslant 0 \\ & B_1(x) \leqslant k \end{aligned} \tag{$\widetilde{P}_k$}$$

Proposition 9.2.1

The problem $(\widetilde{P}_k)$ can be equivalently reformulated as

$$\begin{aligned} \min_{x \in X} \quad & f(x) \\ \text{s.t.} \quad & g(x) \leqslant 0 \\ & \Gamma(x, \|x\|_1) \leqslant 0 \end{aligned} \tag{9.5}$$

where

(i) $\Gamma : (x_1, \dots, x_n, z) \in \mathbb{R}_+^{n+1} \mapsto z \log(z) - \sum_{i \in [n]} x_i \log(kx_i)$,

(ii) the Jacobian of Γ is defined as $\frac{\partial \Gamma}{\partial x_i}(x, z) = -1 - \log(kx_i)$, $\frac{\partial \Gamma}{\partial z}(x, z) = 1 + \log(z)$,

(iii) the Hessian of Γ is $H_\Gamma \coloneqq \operatorname{diag}(-1/x_1, \dots, -1/x_n, 1/z)$ for $(x, z) \in \mathbb{R}_{>0}^{n+1}$.

The proof is immediate.

The relaxation problem that we obtain is not convex (the function Γ is concave), and there is no guarantee in finding the global optimum of this relaxation. Nonetheless, this problem numerically leads to solutions which are both sparse and with satisfactory objective value, see Section 9.4.

9.3 Metric estimates between B_α and ϵ-cardinality

9.3.1 Majorization and Schur-convexity

Definition 9.3.1 (Majorization). For a vector $a \in \mathbb{R}_+^n$, we denote by $a^\downarrow \in \mathbb{R}_+^n$ the vector with the same components, but sorted in descending order. Given $a, b \in \mathbb{R}_+^n$, we say that a weakly majorizes (or dominates) b from below written $a \succ_w b$ iff

$$\sum_{i=1}^k a_i^\downarrow \geqslant \sum_{i=1}^k b_i^\downarrow \quad \text{for } k = 1, \dots, n \ .$$

If $a \succ_w b$ and in addition $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$, then we say that a majorizes b, written $a \succ b$.

Definition 9.3.2 (Schur-convexity/concavity). Let $\mathcal{A} \subset \mathbb{R}_+^n$. A real-valued function $\phi : \mathbb{R}_+^n \to \mathbb{R}$ is said to be *Schur-convex* (resp. *Schur-concave*) if $\phi(x) \leqslant \phi(y)$ (resp. $\phi(x) \geqslant \phi(y)$ for any $x, y \in \mathcal{A}$ satisfying $x \prec y$.

Proposition 9.3.1 ([MOA11], Appendix F.3.a (p.532))

The Rényi entropy of an arbitrary $\alpha > 0$ is Schur-concave; in particular, for $\alpha = 1$, the Shannon entropy is Schur-concave.

9.3.2 Extreme relation between ϵ-cardinality and entropy

We first show that no tight relation can be found between the cardinality and the bound B_α. To see this, let $k < n$, and $\epsilon < 1/n$. Define the probability distribution $v_n(k, \epsilon) \in \Delta_n$ as

$$[v_n(k, \epsilon)]_i = \begin{cases} 1 - (k-1)\epsilon, & i = 1 \\ \epsilon, & 2 \leqslant i \leqslant k \\ 0, & k+1 \leqslant i \leqslant n \end{cases} \tag{9.6}$$

and the associated entropy

$$\begin{aligned} H_\alpha^v(k,\epsilon) &:= H_\alpha(v_n(k,\epsilon)) = \frac{1}{1-\alpha}\ln\left((1-(k-1)\epsilon)^\alpha + (k-1)\epsilon^\alpha\right), && 0<\alpha<1 \\ H_\alpha^v(k,\epsilon) &:= H_\alpha\left(v_n(k,\epsilon)\right) = -\left(1-(k-1)\epsilon\right)\ln\left(1-(k-1)\epsilon\right) - (k-1)\epsilon\ln(\epsilon), && \alpha = 1 \end{aligned}$$

The following property is immediate.

Proposition 9.3.2 (Worst-case comparison between cardinality and B_α)
For any $\epsilon > 0$ and $0 < \alpha \leqslant 1$, $\mathrm{card}\left(v_n(n,\epsilon)\right) = n$ and $B_\alpha\left(v_n(n,\epsilon)\right) \xrightarrow[\epsilon\to 0]{} 1$. Therefore, the cardinality of a given probability distribution $p \in \Delta_n$ is not *controlled* by the estimation $B_\alpha(p)$. □

Nonetheless, we aim to find extreme relations between B_α and the ϵ-*cardinality*, defined as

$$\underset{\epsilon}{\mathrm{card}}(p) = \left|\{i\in[n] \mid p_i \geqslant \epsilon\}\right| \ . \tag{9.7}$$

The parameter ϵ is viewed as a filtering threshold.

Lemma 9.3.1
Viewing k as a real number, the function $k\in[1,n]\mapsto H_\alpha^v(k,\epsilon)$ is an increasing function for $\epsilon \leqslant \frac{1}{n}$ and $0<\alpha\leqslant 1$.

Proof. For $\alpha = 1$, $\frac{\partial H_\alpha^v}{\partial k}(k,\epsilon) = \epsilon\left[1+\ln\left(\frac{1}{\epsilon}-k+1\right)\right]$. As $\epsilon \leqslant \frac{1}{n}$ and $k \leqslant n$, we get that $k \mapsto H_v(k,\epsilon)$ is increasing. Now, for $0 \leqslant \alpha < 1$, $\frac{\partial}{\partial k}\exp\left((1-\alpha)H_\alpha^v(k,\epsilon)\right) = \epsilon^\alpha - \alpha\epsilon(1-(k-1)\epsilon)^\alpha \geqslant \epsilon^\alpha - \epsilon > 0$. □

Lemma 9.3.2
For any $\epsilon > 0$ and $0<\alpha\leqslant 1$, an optimal solution of the problem

$$\min_{p\in\Delta_n}\left\{H_\alpha(p) \mid \underset{\epsilon}{\mathrm{card}}(p) = k\right\} \tag{$P_{\alpha,\epsilon}^{k,n}$}$$

is $v_n(k,\epsilon)$, and corresponds to an objective value $H_\alpha^v(k,\epsilon)$.

Proof. Any ordered element $p\in\Delta_n$ satisfying $\mathrm{card}_\epsilon(p) = k$ can be represented as

$$p = \left(1 - \sum_{i=1}^{k-1}\alpha_i - \sum_{i=k}^{n}\beta_i, \dots, \alpha_{k-1}, \beta_k, \dots, \beta_n\right) \ ,$$

with $\alpha_1 \geqslant \dots \geqslant \alpha_{k-1} \geqslant \epsilon$ and $\epsilon > \beta_k \geqslant \dots \geqslant \beta_n \geqslant 0$. Then, for $1 \leqslant d \leqslant n$,

$$\sum_{i=1}^{d}[v_n(k,\epsilon)]_i - \sum_{i=1}^{d} p_i = \begin{cases} \sum_{i=d}^{k-1}\alpha_i - (k-d)\epsilon + \sum_{i=k}^{n}\beta_i, & d \leqslant k \\ \sum_{i=d}^{n}\beta_i, & d > k \end{cases} \ .$$

By using Proposition 9.3.1, we obtain that the minimum of the Rényi entropy is attained for $v_n(k,\epsilon)$. □

Finding the distribution giving the minimal Rényi entropy using majorization theory has been also performed in [Kog13] and [Sas18] for different set of constraints. Also, extreme relations between Rényi entropy and l_q-norm, $q > 0$, have been found in [SI16].

We introduce the invertible, increasing, function $\phi_{\alpha,\epsilon} : k \in [1, n] \mapsto \exp H^v_\alpha(k, \epsilon) \in [1, n]$.

Theorem 9.3.1 (ϵ-cardinality bounds)

Let $1 \leqslant b \leqslant n$, $\epsilon > 0$ and $0 \leqslant \alpha \leqslant 1$. For any vector $p \in \Delta_n$, if $B_\alpha(p) \leqslant b$, then $\mathrm{card}_\epsilon(p) \leqslant \lfloor \phi^{-1}_{\alpha,\epsilon}(b) \rfloor$.

Proof. By the resolution of $(P^{k,n}_{\alpha,\epsilon})$ (Lemma 9.3.2), we know that

$$\underset{\epsilon}{\mathrm{card}}(p) = k \Rightarrow B_\alpha(p) \geqslant \exp H^v_\alpha(k, \epsilon)$$

As $\phi_{\alpha,\epsilon}$ is increasing and invertible, we deduce that $\mathrm{card}_\epsilon(p) \geqslant k \Rightarrow B_\alpha(p) \geqslant \exp H^v_\alpha(k, \epsilon)$ and so

$$B_\infty(p) \leqslant b \Rightarrow \underset{\epsilon}{\mathrm{card}}(p) \leqslant \phi^{-1}_{\alpha,\epsilon}(b) \ .$$

□

Remark 9.3.1

The relation found in Theorem 9.3.1 is tight as it is attained for $p = v_n(\phi^{-1}_{\alpha,\epsilon}(b), \epsilon)$ if $\phi^{-1}_{\alpha,\epsilon}(b) \in \mathbb{N}$.

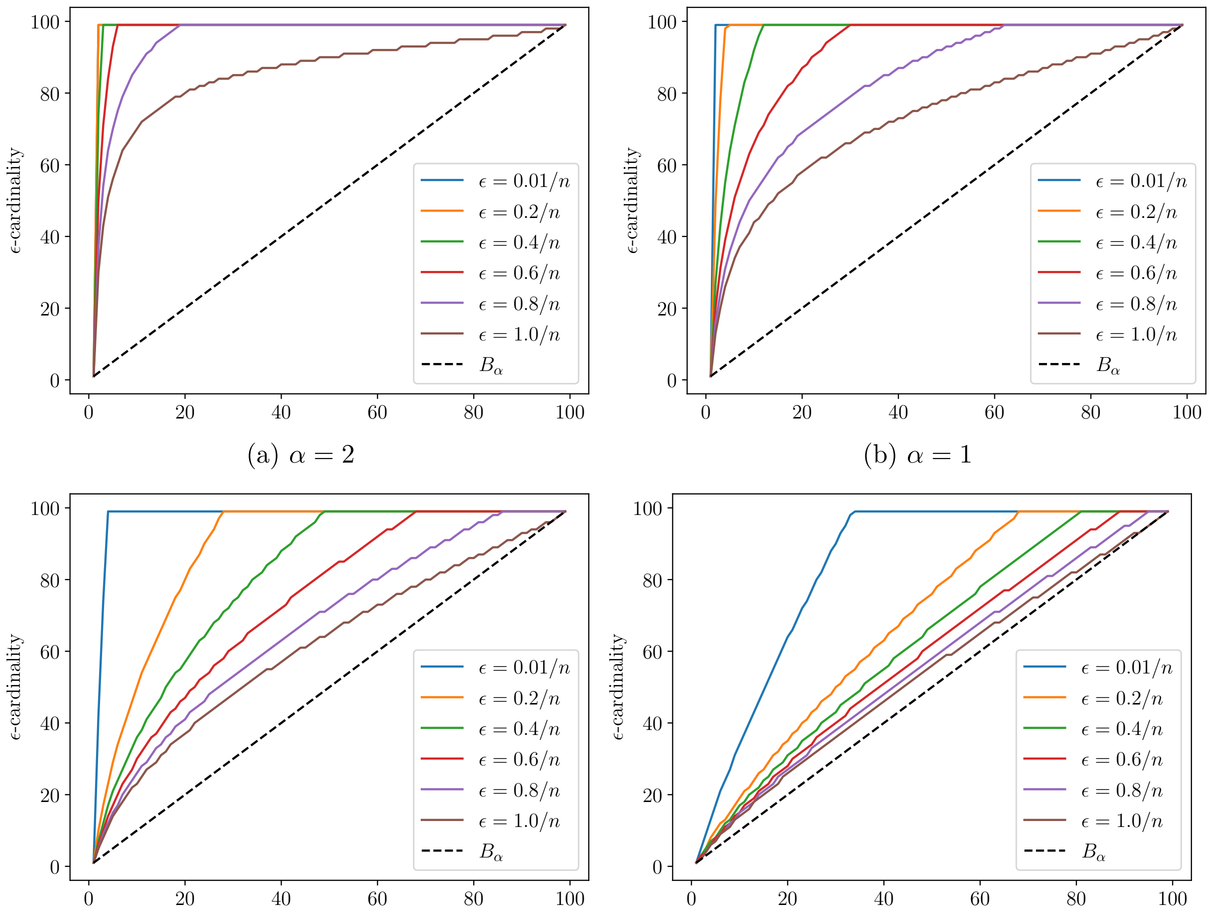
Theorem 9.3.1 provides sparsity guarantees for the solution. In fact, if one requires a maximum cardinality of b, the solution has an ϵ-cardinality of $\lfloor \phi^{-1}_{\alpha,\epsilon}(b) \rfloor$. Figure 9.2 shows that the tightness of the bound improves when ϵ grows and α decreases.

9.4 Numerical Experiments

As an illustration of our approach, we will consider a sparse regression problem on the simplex with a use case from finance (Index tracking).

A financial index is a number representing the value of the set of assets (stocks or bonds) which reflects the value of a specific market or a segment of it. Insofar as an index is not a financial instrument that we can directly trade, a stock or a bond market index is effectively equivalent to a hypothetical portfolio of assets. In order to gain access to an index, it is necessary to use financial instruments such as options, futures and exchange-traded funds, or to create a portfolio of assets that closely tracks a given index. For a given index, fund managers have the choice between two basic investment strategies. The active strategy assumes that the markets are not perfectly efficient so that fund managers, thanks to their know-how, makes specific investments and hope to add value by choosing high performing assets outperforming an investment benchmark index. On the contrary, the passive strategy assumes that the market cannot be beaten in the long run, so that fund managers expect a return that closely replicates the investment weighting and returns of a benchmark index.

Currently, passive strategies seem to attract more interest from investors. Index tracking, also known as index replication, is one of the most popular passive portfolio management strategy to use the market index to determine the portfolio weights by reproducing the performance of a market index, i.e., to match the performance of a theoretical portfolio as closely as possible. Index tracking allows to get the desired returns from the overall market growth with the lower



(a) $\alpha = 2$ (b) $\alpha = 1$

(c) $\alpha = 0.5$ (d) $\alpha = 0.2$

Figure 9.2: The ϵ-cardinality upper bound $b \mapsto \lfloor \phi_{\alpha,\epsilon}^{-1}(b) \rfloor$ for $1 \leqslant b \leqslant n = 100$. The approximation becomes tighter when ϵ increases and when α decreases.

variability and the lower expense ratio for the investment. The smaller the number of assets needed to mimic is, the smaller the incurred transaction costs will be. Nevertheless, the tracking error is likely to be higher when a small number of assets is used.

To create a tracking portfolio, the simplest technique, called full replication, is to buy appropriate amounts of all the assets that make up the index. Provided that the true index construction weights are available, it allows a perfect tracking. However, it has several disadvantages, one related to the fact that a portfolio can consist of thousands of stocks and the other to the fact that there can be many small or illiquid stocks. These last types of shares increase the risk associated with their sale, which is more difficult, and generate an arbitrage cost that is all the more significant as it is frequent. One of the ways to overcome these drawbacks is to construct a sparse index tracking portfolio ([BMC03], [JV02]) by limiting the number of assets to approximately replicate an index. It corresponds to tracking a signal using a sparse mixture of a given set of time series, see e.g. [BFP18]. A sparse portfolio simplifies the execution of the portfolio and tends to avoid illiquid stocks that usually correspond to the assets with small weights in an index, since in a sparse setting most of these assets are discarded. Furthermore, since only a small number of assets is used, the transaction costs are reduced significantly due to the reduction of the fixed (minimum) costs in the commission fees. For more details, see ([BFP18], [CE14]).

Formulation as a Sparse Regression Problem. Following [CE14], we give the main steps leading to formulate the Index Tracking problem as Sparse Regression Problem. Let a single financial asset j on which we invest a sum S_j at the beginning of a period. If the rate of return (or return) of this single asset is denoted r_j, we will earn $S_{j,end} = (1 + r_j)S_j$ at the end of the period with $r_j = \dfrac{S_{j,end} - S_j}{S_j}$. For n assets, we define a vector $r \in \mathbb{R}^n$ where the j-th component is the rate of return of the j-th asset. $r(k) \in \mathbb{R}^n$ represents the vector of simple returns of the components assets during the k-th period of time $[(k-1)\Delta, k\Delta]$, where Δ is a fixed duration.

Let the entries of $x \in \mathbb{R}^n$ are the fractions of an investor's total wealth invested in each of n different assets. Investing at the beginning of the period a total sum S over all assets is made by allocating a fraction $x_j,\ j = 1, \dots, n$ of S in the j-th asset. The non-negative vector $x \in \mathbb{R}^n_+$ represents the portfolio "mix", and its components sum to one. At the end of the period, the total value of the portfolio is $S_{end} = \sum_{j=1}^{n}(1 + r_j)x_j S$. The rate of return of the portfolio is the relative increase in wealth $\dfrac{S_{end} - S}{S} = \sum_{j=1}^{n}(1 + r_j)x_j - 1 = \sum_{j=1}^{n} x_j - 1 + \sum_{j=1}^{n} r_j x_j = r^T x$; i.e., the standard inner product between the vector r of individual returns $r_j,\ j = 1, \dots, n$ and the vector of the portfolio allocation weights x. The $m \times n$ matrix R gives the (close price) data of the component assets. The component y_k of the vector $y \in \mathbb{R}^T$ represents the return of some target financial index over the j-th period, for $j = 1, \dots, n$. Vector y is the close price of the target index. Then , the so-called *index tracking* problem is to construct a portfolio x so as to track as close as possible the "benchmark" index returns y. Since the vector of portfolio returns over the considered time horizon is :

$$z = Rx, \quad R \in \mathbb{R}^{m \times n} \quad .$$

We may seek for the portfolio x with minimum Least Squares tracking error, by minimizing $\|Rx - y\|_2^2$. However, we need to take into account the fact that the elements of x represent relative *weights*, that is they are non-negative and they sum up to one. In addition, a cardinality constraint is added for constructing a sparse index tracking portfolio. For given $R \in \mathbb{R}^{m \times n}$ and $y \in \mathbb{R}^m$, this leads to the following sparse regression problem :

$$(P_k) \quad \min_{x \in \mathbb{R}^n_+} \left\{ \|y - Rx\|^2 \;\middle|\; \sum_{i=1}^{n} x_i = 1,\ \mathrm{card}(x) \leqslant k \right\} \quad .$$

Problem (P_n) is then the problem *without sparsity requirement.* The constraint $x \geqslant 0, 1^T x = 1$ makes the use of LASSO penalty (constant over the feasible set) irrelevant.

Numerical results. We conducted two experiments with data from [Cal21]. The results have been obtained on a laptop i7-1065G7 CPU@1.30GHz.

In the first experiment, we consider the following sparse techniques with a limited index tracking data set index with $n = 50$ assets over a period of $m = 229$ time steps (the limited number of assets being the limiting dimension that the SDP method can accept) :

(i) *Greedy heuristic*: solve (P_n), take the k greatest value of x and renormalize

(ii) *Reversed greedy heuristic*:

(iii) *SDP approach*: computation of method `sdp2` of [AG19]

(iv) *Mixed-integer programming*: exact solving using `CPLEX`

Algorithm 14 Reversed greedy heuristic

$x \leftarrow$ Solution of (P_n)
while card$(x) > k$ **do**
  $i \leftarrow \arg\min_{1 \leqslant j \leqslant n} x_j$
  Add the constraint $x_i = 0$ to (P_n)
  $x \leftarrow$ Solution of (P_n)
end whilereturn x

(v) *Entropy lower bound*: solve the problem $\min_{x \in \mathbb{R}^n_+} \left\{ \|y - Rx\|^2 \;\middle|\; \sum_{i=1}^n x_i = 1,\ B_1(x) \leqslant k \right\}$

Remark 9.4.1

The plain method based on $\|x\|_\infty \geqslant 1/k$ [PEC12] has also been tested, but it does not produce solutions with significant sparsity for this specific problem.

We aim at finding a vector $x \in \mathbb{R}^n$ with sparsity $k = 10$. Figures 9.3 to 9.5 illustrate the obtained results. For the different methods tested, we carried out various simulations by varying the desired cardinality along the x-axis in an interval ranging from 5 (a high degree of sparsity is required with only 5 non-zero values out of the 50) to 45 (the desired vector is practically dense). The quality of the solution to the problem can be assessed according to two criteria: the value of the objective function at the optimum and the respect of the cardinality constraint.
The main comments that can be made from these results are listed below:

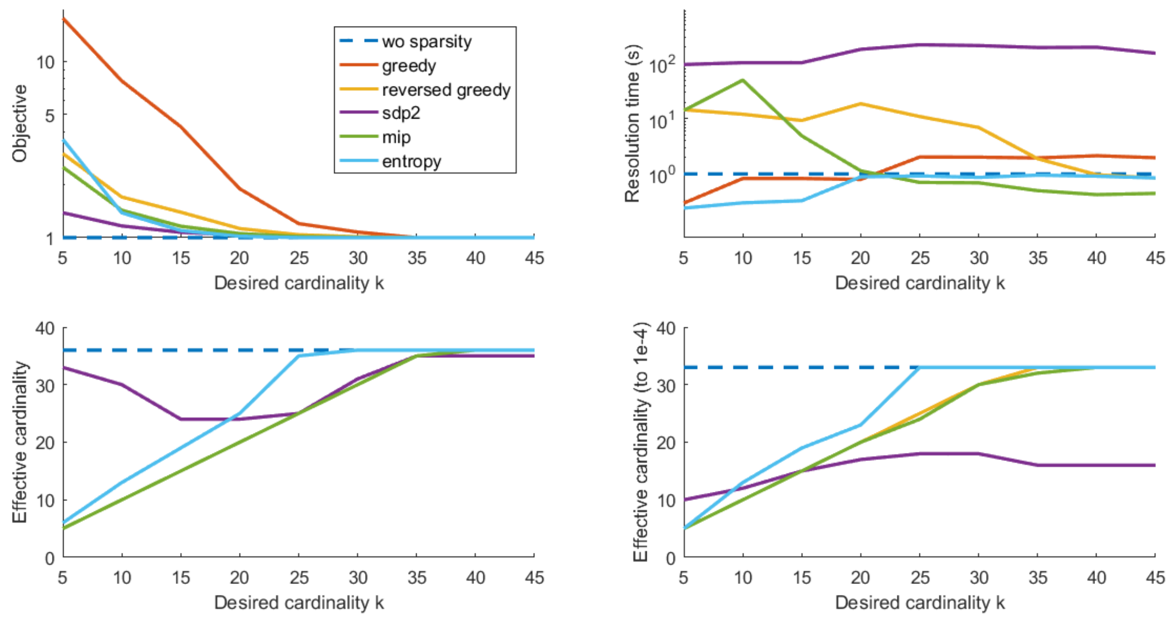


Figure 9.3: Index tracking for $n = 50$.
At the bottom left, the cardinality constraint is satisfied below the diagonal $y = x$.
At the bottom right, we display the cardinality for the solution filtered with a threshold of 10^{-4}.

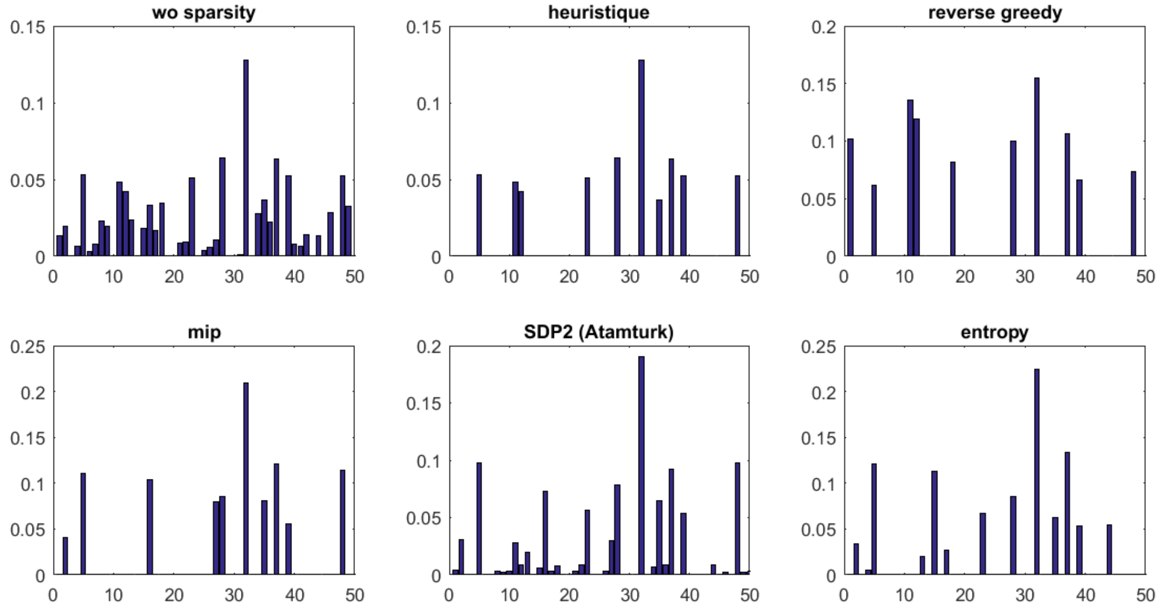


Figure 9.4: Index tracking for $n = 50$ and $k = 10$

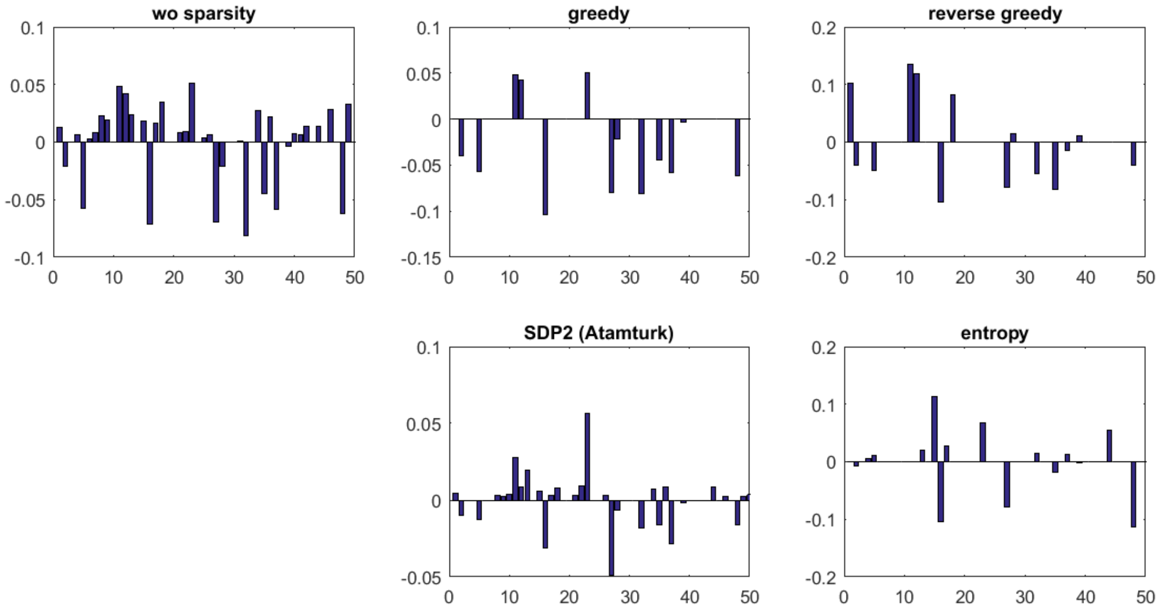


Figure 9.5: Index tracking for $n = 50$ and $k = 10$.
The bars represent the difference with exact solution.

(i) Compliance with the cardinality constraint is all the more difficult to satisfy when the desired cardinality is low (see e.g. the time of the exact solver). Beyond a certain degree of sparsity (here, about 30), the problem becomes easy to solve for all the methods tested.

(ii) Concerning the value of the objective function, we note that the "greedy" method is the least efficient of all, while the "reversed greedy" method is competitive. From a desired cardinality of 10, the results of the entropic method are very close to the MIP method (exact resolution of the problem).

(iii) Regarding the respect of the cardinality constraint, we observe that, for strong sparsity requirements, the SDP2 technique absolutely does not respect the desired cardinality unlike the entropic method. The fact of filtering at 10^{-4} the values of cardinalities obtained does not change the fact that the SDP2 method cannot calculate a solution which respects the desired cardinality. The reversed greedy method provides solutions that respect the cardinality constraint.

(iv) Concerning computation times, unsurprisingly the SDP2 method is the most expensive by far, even for very low sparsity requirements. The exact MIP method is also expensive but the computation time becomes logically lower as the sparsity requirement weakens. The "reversed greedy" method requires a computation time that remains quite high, regardless of the level of cardinality requirement. The entropic method, on the other hand, makes it possible to calculate solutions in short times, even when the cardinality requirement is very strong. This is an important point in practice, especially for large problem instances, in which the entropic approach is more adapted than the SDP approach or the exact approach.

In a second experiment, we illustrate the possibility to compute a sparse solution via entropic bound even in high dimension ($n = 430$ assets) and hard cardinality requirement ($k = 6$). Figure 9.6 compares the solution obtained without cardinality constraint (left subfigure) with the relaxed problem (right subfigure). Our technique is highly scalable since its computational time is low (around 1 second for our technique against around 3 seconds for the problem without sparsity requirement). Moreover, the sparsity requirement is almost fully satisfied, as the effective cardinality of the solution is 7 (the target was $k = 6$).

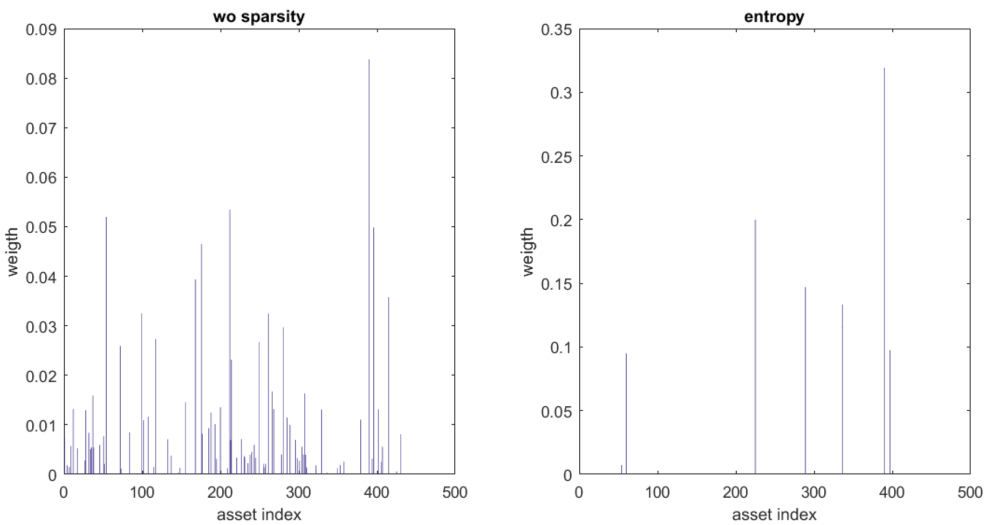


Figure 9.6: Index Tracking for $n = 430$ and $k = 6$.

9.5 Conclusions and Perspectives

By using ratios of norms, we proposed a new lower bound of cardinality, based on Shannon entropy. Despite its non-convexity, the use of this entropic bound in a sparse optimization problem is easy, and a local solution can be found very rapidly by using nonlinear solvers. Early results obtained on Index Tracking Finance problem are good regarding other approaches (heuristics, SDP,...) and the proposed approach seems promising.

Among the various perspectives opened to future investigation, we can mention the search for efficient bounds and estimates of cardinality (results on estimates can be found in [BZE15]). Extensive simulations on various applications, including Machine Learning, in order to evaluate the efficiency of our approach would be worth considering. Finally, a close look on the relations between Shannon entropy and ℓ_0-pseudonorm should also be done to possibly get approximation guarantees in the sparse optimization problem.

Perspectives

***

We tackle here some perspectives and possible extensions of the work presented in this PhD dissertation.

Extensions to prosumers and multi-leader games

In the first part of this thesis, and especially in Chapter 4 we consider that the competitors are static, i.e., the prices of their offers are supposed to be known in advance. This is justified when it comes to reacting to a market situation, but not if one aims at finding the equilibrium in the whole retail market. To this purpose, a natural extension would be to consider a multi-leader-common-follower game, as in [LM10] and [Aus+20]. In this setting, given the leaders strategies, the analyze of the customers reaction that we conducted still applies. Nonetheless, the computation of all the strategies of the leaders requires the solving of an Equilibrium Problem with Equilibrium Constraints (EPEC) of high complexity, and where only "stationary points" (which constitute a superset of Nash equilibria) can be found in general.

We have decided in this thesis to focus on inertia and elasticity respectively in Chapter 5 and Chapter 6, but the follower problem can be complexified in other directions. Among them, followers can be in reality prosumers, i.e., consumers who can both consume and produce electricity, and even store electricity. Therefore, in this setting, the decision of the customer does not only relate to the choice of a contract, but also involves the choice of installing or not capacities (e.g. solar panels). The description by tropical geometry and cell arrangements does not straightforwardly extend to this more complex case.

Links with Weak-KAM theory for non-controllable system

We provide at the end of Chapter 5 links between weak-KAM theory and turnpike property. In particular, we restrict to completely controllable systems to prove that under strict-dissipativity condition, the process converges to an Aubry set, which can be either restricted to a single state (standard turnpike property) or a set of states constituting a periodic stationary strategy. However, this study was conducted by supposing that it is always possible to attain a state from another one (controllability). This assumption does not hold in the pricing model we solved in Chapter 5. Therefore, a natural extension could be the development of similar convergence results with weaker condition, in the spirit of [AL10].

Quantization of monopolist problem for partial participation.

In Chapter 6, we present a pruning procedure to obtain a quantized solution (a menu of prescribed number of offers) from the solution of the monopolist problem (nonlinear pricing). To do so, we require that the solution satisfies the full-participation condition, i.e., that the whole population contracts with the retailer (monopoly situation). This situation easily extends to the case where the targeted agents in the population are known. However, the retailer may aim at optimizing in the same time the tariff menu and the set of customers with whom she will contract. This partial-participation case was first studied by Jullien [Jul00], and Carlier and Zhang [CZ20] demonstrated the existence of a solution for the continuum contract. Nonetheless, no numerical method has been developed for this specific setting. The reverse greedy algorithm introduced in this work must be adapted so that the procedure ensures at each iteration that each targeted agent is effectively chosen and that the others are not.

Chance-constrained programing and sparse optimization for electricity markets

Chance-constrained programming problems naturally appear in a various range of applications, especially in energy management, see e.g. [Ack+11]. In a pricing context, a major interest is the estimate of electricity invoice. Indeed, the consumption is of stochastic nature and depends on multiple factors such as the weather. In the model developed in Chapter 4, we consider quadratic regularization or logit distribution as probability measure of the quantities $y_{sw} = \mathbb{P}[U_{sw} \leqslant U_{sw'}]$. Another possibility could be to design distributionally robust approximation by considering Bennett-type bounds, knowing estimations of mean and variances of energy consumption. In Chapter 6, the question of the optimal number of contracts naturally belongs to sparse optimization problem, in which we require the cardinality of the menu to be bounded. The use of entropic bounds (such as Shannon entropy) could lead to sparse menu of offers with sufficient revenue for the retailer.

Bibliography

***

[ARR21] A. Abate, R. Riccardi, and C. Ruiz. "Retailer-consumers model in electricity market under demand response". May 2021. DOI: 10.13140/RG.2.2.36840.80647.

[Abd17] O. A. Abdelrahman. "The Relationship between Pricing and Consumers Switching Costs: Comparisons between the Myopic and Perfect-foresight Equilibria". In: *International Business Research* 10.3 (Feb. 2017), p. 221. DOI: 10.5539/ibr.v10n3p221.

[AL20] Y. Achdou and M. Laurière. "Mean Field Games and Applications: Numerical Aspects". In: *Lecture Notes in Mathematics*. Springer International Publishing, 2020, pp. 249–307. DOI: 10.1007/978-3-030-59837-2_4.

[Ack+18] W. van Ackooij, I. Danti Lopez, A. Frangioni, F. Lacalandra, and M. Tahanan. "Large-scale unit commitment under uncertainty: an updated literature survey". In: *Annals of Operations Research* 271.1 (Sept. 2018), pp. 11–85. DOI: 10.1007/s10479-018-3003-z.

[AM19] W. van Ackooij and J. Malick. "Eventual convexity of probability constraints with elliptical distributions". In: *Mathematical Programming* 175.1 (2019), pp. 1–27. DOI: 10.1007/s10107-018-1230-3.

[AO22] W. van Ackooij and W. de Oliveira. "Addendum to the paper 'Nonsmooth DC-constrained optimization: constraint qualification and minimizing methodologies'". In: *Optimization Methods and Software* 37.6 (May 2022), pp. 2241–2250. DOI: 10.1080/10556788.2022.2063861.

[Ack20] W. van Ackooij. "A discussion of probability functions and constraints from a variational perspective". In: *Set-Valued and Variational Analysis* 28.4 (2020), pp. 585–609. DOI: 10.1007/s11228-020-00552-2.

[Ack+11] W. van Ackooij, R. Zorgati, R. Henrion, and A. Mller. "Chance Constrained Programming and Its Applications to Energy Management". In: *Stochastic Optimization - Seeing the Optimal for the Uncertain*. InTech, Feb. 2011. DOI: 10.5772/15438.

[Afş+16] S. Afşar, L. Brotcorne, P. Marcotte, and G. Savard. "Achieving an optimal trade-off between revenue and energy peak within a smart grid environment". In: *Renewable Energy* 91 (2016), pp. 293–301. DOI: 10.1016/j.renene.2016.01.055.

[AL10] A. Agrachev and P. W. Y. Lee. "Continuity of optimal control costs and its application to weak KAM theory". In: 39.1-2 (Feb. 2010), pp. 213–232. DOI: 10.1007/s00526-010-0308-4.

[AKT23] R. Aïd, A. Kemper, and N. Touzi. *A Principal-Agent Framework for Optimal Incentives in Renewable Investments*. 2023. arXiv: 2302.12167 [econ.GN].

[APT22] R. Aïd, D. Possamaï, and N. Touzi. "Optimal Electricity Demand Response Contracting with Responsiveness Incentives". In: *Mathematics of Operations Research* 47.3 (Aug. 2022), pp. 2112–2137. DOI: 10.1287/moor.2021.1201.

[AGL08] M. Akian, S. Gaubert, and A. Lakhoua. "The Max-Plus Finite Element Method for Solving Deterministic Optimal Control Problems: Basic Properties and Convergence Analysis". In: *SIAM Journal on Control and Optimization* 47.2 (Jan. 2008), pp. 817–848. DOI: 10.1137/060655286.

[AGN11] M. Akian, S. Gaubert, and R. Nussbaum. "A Collatz-Wielandt characterization of the spectral radius of order-preserving homogeneous maps on cones". 2011. eprint: 1112.5968.

[AGN15] M. Akian, S. Gaubert, and R. Nussbaum. "Uniqueness of the fixed point of nonexpansive semidifferentiable maps". In: *Transactions of the American Mathematical Society* 368.2 (Feb. 2015), pp. 1271–1320. DOI: 10.1090/s0002-9947-2015-06413-7.

[AGW09] M. Akian, S. Gaubert, and C. Walsh. "The max-plus Martin boundary". In: *Doc. Math* 14 (2009), pp. 195–240.

[ABM20] C. Alasseur, I. Ben Tahar, and A. Matoussi. "An extended mean field game for storage in smart grids". In: *Journal of Optimization Theory and Applications* 184.2 (2020), pp. 644–670.

[Ala+20] C. Alasseur, I. Ekeland, R. Élie, N. H. Santibáñez, and D. Possamaï. "An Adverse Selection Approach to Power Pricing". In: *SIAM Journal on Control and Optimization* 58.2 (Jan. 2020), pp. 686–713. DOI: 10.1137/19m1260578.

[AFS20] C. Alasseur, H. Farhat, and M. Saguan. "A Principal-agent approach to capacity remuneration mechanisms". In: *International Journal of Theoretical and Applied Finance* 23.08 (Dec. 2020), p. 2050052. DOI: 10.1142/s0219024920500521.

[AE08] M. Albadi and E. El-Saadany. "A summary of demand response in electricity markets". In: *Electric Power Systems Research* 78.11 (Nov. 2008), pp. 1989–1996. DOI: 10.1016/j.epsr.2008.04.002.

[Ale+19] E. Alekseeva, L. Brotcorne, S. Lepaul, and A. Montmeat. "A bilevel approach to optimize electricity prices". In: *Yugoslav Journal of Operations Research* 29.1 (2019), pp. 9–30. DOI: 10.2298/yjor171115002a.

[AT14] H. Allcott and R. Todd. "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation". In: *American Economic Review* 104.10 (2014), pp. 3003–37.

[AR12] W. J. Allender and T. J. Richards. "Brand Loyalty and Price Promotion Strategies: An Empirical Analysis". In: *Journal of Retailing* 88.3 (Sept. 2012), pp. 323–342. DOI: 10.1016/j.jretai.2012.01.001.

[And+97] P. L. Anderson, R. D. McLellan, J. P. Overton, and G. L. Wolfram. "Price elasticity of demand". In: *McKinac Center for Public Policy. Accessed October* 13.2 (1997).

[AF19] S. Andrieu and A. Ferrari. "Green electricity offers: how to guarantee their origin?" In: (2019).

[ABG11] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic Control of Diffusion Processes*. Cambridge University Press, Nov. 2011. DOI: 10.1017/cbo9781139003605.

[Arm96] M. Armstrong. "Multiproduct Nonlinear Pricing". In: *Econometrica* 64.1 (Jan. 1996), p. 51. DOI: 10.2307/2171924.

[AG19] A. Atamturk and A. Gomez. *Rank-one Convexification for Sparse Regression*. 2019. DOI: 10.48550/ARXIV.1901.10334.

[ACR20] S. Auray, V. Caponi, and B. Ravel. "Price Elasticity of Electricity Demand in France". In: *Economics and Statistics* 513 (Apr. 2020), pp. 91–103. DOI: 10.24187/ecostat.2019.513.2002.

[Aus+20] D. Aussel, L. Brotcorne, S. Lepaul, and L. von Niederhäusern. "A trilevel model for best response in energy demand-side management". In: *European Journal of Operational Research* 281.2 (2020), pp. 299–315. DOI: https://doi.org/10.1016/j.ejor.2019.03.005.

[Avi00] D. Avis. "A Revised Implementation of the Reverse Search Vertex Enumeration Algorithm". In: *Polytopes — Combinatorics and Computation*. Ed. by G. Kalai and G. M. Ziegler. Basel: Birkhäuser Basel, 2000, pp. 177–198. ISBN: 978-3-0348-8438-9. DOI: 10.1007/978-3-0348-8438-9_9.

[BMP13] L. Bai, J. Mitchell, and J.-S. Pang. "On convex quadratic programs with linear complementarity constraints". In: *Computational Optimization and Applications* 54 (Apr. 2013). DOI: 10.1007/s10589-012-9497-4.

[BK19] E. Baldwin and P. Klemperer. "Understanding preferences:"demand types", and the existence of equilibrium with indivisibilities". In: *Econometrica* 87.3 (2019), pp. 867–932.

[Ban+05] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh. "Clustering with Bregman Divergences". In: *Journal of Machine Learning Research* 6.58 (2005), pp. 1705–1749.

[Bar13] J. F. Bard. *Practical bilevel optimization: algorithms and applications*. Vol. 30. New York: Springer Science & Business Media, 2013.

[Bar38] C. Barnard. "The Functions of the Executive". In: *Harvard University Press* (1938).

[Bäu23] N. Bäuerle. "Mean Field Markov Decision Processes". In: *Applied Mathematics and Optimization* 88.1 (Apr. 2023). DOI: 10.1007/s00245-023-09985-1.

[BL15] G. Bayraksan and D. K. Love. "Data-Driven Stochastic Programming Using Phi-Divergences". In: *The Operations Research Revolution*. INFORMS, Sept. 2015, pp. 1–19. DOI: 10.1287/educ.2015.0134.

[BCZ19] E. Bayraktar, J. Cvitanić, and Y. Zhang. "Large tournament games". In: *The Annals of Applied Probability* 29.6 (Dec. 2019). DOI: 10.1214/19-aap1490.

[BZ16] E. Bayraktar and Y. Zhang. "A rank-based mean field game in the strong formulation". In: *Electronic Communications in Probability* 21 (2016), pp. 1–12.

[BZ21] E. Bayraktar and Y. Zhang. "Terminal Ranking Games". In: *Mathematics of Operations Research* 46.4 (Nov. 2021), pp. 1349–1365. DOI: 10.1287/moor.2020.1107.

[BMC03] J. Beasley, N. Meade, and T. Chang. "An evolutionary heuristic for the index tracking problem". In: *European Journal of Operational Research* 148.3 (2003), pp. 621–643.

[BLS23] Y. Beck, I. Ljubić, and M. Schmidt. "A survey on bilevel optimization under uncertainty". In: *European Journal of Operational Research* (Feb. 2023). DOI: 10.1016/j.ejor.2023.01.008.

[Bel54] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Dec. 1954. DOI: 10.1515/9781400835386.

[BBN21] R. Ben Mhenni, S. Bourguignon, and J. Ninin. "Global optimization for sparse solution of least squares problems". In: *Optimization Methods and Software* (2021), pp. 1–30. DOI: 10.1080/10556788.2021.1977809.

[Ben+10] A. Ben-Tal, S. Bhadra, C. Bhattacharyya, and J. S. Nath. "Chance constrained uncertain classification via robust optimization". In: *Mathematical Programming* 127.1 (Oct. 2010), pp. 145–173. DOI: 10.1007/s10107-010-0415-1.

[BEN09] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Vol. 28. Princeton university press, 2009.

[BN00] A. Ben-Tal and A. Nemirovski. "Robust solutions of linear programming problems contaminated with uncertain data". In: *Mathematical programming* 88.3 (2000), pp. 411–424.

[BCS17] J.-D. Benamou, G. Carlier, and F. Santambrogio. "Variational mean field games". In: *Active Particles, Volume 1: Advances in Theory, Models, and Applications* (2017), pp. 141–171.

[BFP18] K. Benidis, Y. Feng, and D. P. Palomar. "Sparse Portfolios for High-Dimensional Financial Index Tracking". In: *IEEE Transactions on Signal Processing* 66.1 (2018), pp. 155–170. DOI: 10.1109/tsp.2017.2762286.

[Ben62] G. Bennett. "Probability Inequalities for the Sum of Independent Random Variables". In: *Journal of the American Statistical Association* 57.297 (Mar. 1962), pp. 33–45. DOI: 10.1080/01621459.1962.10482149.

[Ben68] G. Bennett. "A one-sided probability inequality for the sum of independent, bounded random variables". In: *Biometrika* 55.3 (1968), pp. 565–569. DOI: 10.1093/biomet/55.3.565.

[BM92] K. P. Bennett and O. L. Mangasarian. "Robust linear programming discrimination of two linearly inseparable sets". In: *Optimization Methods and Software* 1.1 (Jan. 1992), pp. 23–34. DOI: 10.1080/10556789208805504.

[BYZ21] D. Bergemann, E. Yeh, and J. Zhang. "Nonlinear pricing with finite information". In: *Games and Economic Behavior* 130 (Nov. 2021), pp. 62–84. DOI: 10.1016/j.geb.2021.08.004.

[Ber04] A. Berizzi. "The Italian 2003 blackout". In: *IEEE Power Engineering Society General Meeting, 2004*. IEEE, 2004. DOI: 10.1109/pes.2004.1373159.

[BP94] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial and Applied Mathematics, Jan. 1994. DOI: 10.1137/1.9781611971262.

[Ber12] D. Bertsekas. *Dynamic programming and optimal control: Volume I*. Vol. 1. Athena scientific, 2012.

[Ber98] D. P. Bertsekas. "A New Value Iteration method for the Average Cost Dynamic Programming Problem". In: *SIAM Journal on Control and Optimization* 36.2 (Mar. 1998), pp. 742–759. DOI: 10.1137/s0363012995291609.

[BS07] D. Bertsimas and R. Shioda. "Algorithm for cardinality-constrained quadratic optimization". In: *Computational Optimization and Applications* 43.1 (Nov. 2007), pp. 1–22. DOI: 10.1007/s10589-007-9126-9.

[Bi+03] J. Bi, K. Bennett, M. Embrechts, C. Breneman, and M. Song. "Dimensionality Reduction via Sparse Support Vector Machines." In: *Journal of Machine Learning Research* 3 (Mar. 2003), pp. 1229–1243. DOI: 10.1162/153244303322753643.

[BZE14a] A. Bialecki, R. Zorgati, and L. El Ghaoui. "Intra-Day Unit-Commitment : A Group Sparsity Approach". In: *Euro Mini Conference on Stochastic Programming and Energy Applications, Institut Henri Poincaré (IHP), Paris, September 24-26* (2014).

[BZE14b] A. Bialecki, R. Zorgati, and L. El Ghaoui. "Intra-Day Unit-Commitment : A Group Sparsity Approach". In: *COPI'14 : Conference on Optimization and Practices in Industry, Palaiseau, October 28-31* (2014).

[BZE15] A. Bialecki, R. Zorgati, and L. El Ghaoui. "Estimating the cardinality of a vector for optimization problems". In: *SIAM Conference on Applied Linear Algebra LA'15, Atlanta, October 26-30* (2015).

[Bie96] D. Bienstock. "Computational study of a family of mixed-integer quadratic programming problems". In: *Mathematical Programming* 74.2 (Aug. 1996), pp. 121–140. DOI: 10.1007/bf02592208.

[BL11] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer New York, 2011. DOI: 10.1007/978-1-4614-0237-4.

[Bis15] A. Biswas. *Mean Field Games with Ergodic cost for Discrete Time Markov Processes.* 2015. DOI: 10.48550/ARXIV.1510.08968. arXiv: 1510.08968 [math.OC].

[BNN10] J.-D. Boissonnat, F. Nielsen, and R. Nock. "Bregman Voronoi Diagrams". In: *Discrete and Computational Geometry* 44.2 (Apr. 2010), pp. 281–307. DOI: 10.1007/s00454-010-9256-1.

[Bon+06] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical optimization: theoretical and practical aspects.* Springer Science and Business Media, 2006.

[Bön+15] W. Bönte, S. Nielen, N. Valitov, and T. Engelmeyer. "Price elasticity of demand in the EPEX spot market for electricity—New empirical evidence". In: *Economics Letters* 135 (Oct. 2015), pp. 5–8. DOI: 10.1016/j.econlet.2015.07.007.

[Bor62] K. Borch. "Equilibrium in a Reinsurance Market". In: *Econometrica* 30.3 (July 1962), p. 424. DOI: 10.2307/1909887.

[BKS15] T. Börgers, D. Krähmer, and R. Strausz. *An Introduction to the Theory of Mechanism Design.* Oxford University PressNew York, May 2015. DOI: 10.1093/acprof:oso/9780199734023.001.0001.

[BLB04] S. Boucheron, G. Lugosi, and O. Bousquet. "Concentration Inequalities". In: *Advanced Lectures on Machine Learning.* Springer Berlin Heidelberg, 2004, pp. 208–240. DOI: 10.1007/978-3-540-28650-9_9.

[BM98] P. Bradley and O. Mangasarian. "Feature Selection via Concave Minimization and Support Vector Machines". In: *Machine Learning Proceedings of the Fifteenth International Conference(ICML98).* Morgan Kaufmann, 1998, pp. 82–90.

[Bré+19] M. Brégère, P. Gaillard, Y. Goude, and G. Stoltz. "Target tracking for contextual bandits: Application to demand side management". In: *International Conference on Machine Learning.* PMLR. 2019, pp. 754–763.

[Bro+06] L. Brotcorne, J.-P. Côté, P. Marcotte, and G. Savard. "A BILEVEL APPROACH TO OPTIMAL PRICING". In: *IFAC Proceedings Volumes* 39.3 (2006), pp. 443–448. DOI: 10.3182/20060517-3-fr-2903.00234.

[Bro+00] L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. "A Bilevel Model and Solution Algorithm for a Freight Tariff-Setting Problem". In: *Transportation Science* 34.3 (Aug. 2000), pp. 289–302. DOI: 10.1287/trsc.34.3.289.12299.

[BLS22] L. Brotcorne, H. Le Cadre, and J. Sepúlveda. "A Reverse Stackelberg Game Model for Grid Usage Pricing". In: *INFORMS 2022 - INFORMS Annual Meeting*. Indianapolis, United States, Oct. 2022.

[BNW06] R. H. Byrd, J. Nocedal, and R. A. Waltz. "Knitro: An Integrated Package for Nonlinear Optimization". In: *Large-Scale Nonlinear Optimization*. Ed. by G. Di Pillo and M. Roma. Boston, MA: Springer, 2006, pp. 35–59. DOI: 10.1007/0-387-30065-1_4.

[Cab09] L. Cabral. "Small switching costs lead to lower prices". In: *Journal of Marketing Research* 46.4 (2009), pp. 449–451.

[Cac97] C. Cachin. "Entropy measures and unconditional security in cryptography". PhD thesis. ETH Zurich, 1997.

[CG06] G. C. Calafiore and L. E. Ghaoui. "On Distributionally Robust Chance-Constrained Linear Programs". In: *Journal of Optimization Theory and Applications* 130.1 (Dec. 2006), pp. 1–22. DOI: 10.1007/s10957-006-9084-x.

[Cal21] G. Calafiore. *Index Tracking Data*. Tech. rep. Personnal Communication: Politchnico Torino, 2021.

[CE14] G. Calafiore and L. El Ghaoui. *Optimization Models*. Control systems and optimization series. Cambridge University Press, 2014.

[CGG14] V. Calvez, P. Gabriel, and S. Gaubert. "Non-linear eigenvalue problems arising from growth maximization of positive linear dynamical systems". In: *Proceedings of the 53rd IEEE Annual Conference on Decision and Control (CDC), Los Angeles*. 2014, pp. 1600–1607.

[Cam+19] N. Cammardella, A. Busic, Y. Ji, and S. Meyn. "Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, Dec. 2019. DOI: 10.1109/cdc40024.2019.9029512.

[Cam+21] S. Campbell, Y. Chen, A. Shrivats, and S. Jaimungal. *Deep Learning for Principal-Agent Mean Field Games*. 2021. arXiv: 2110.01127 [cs.LG].

[CCY12] A. Capponi, J. Cvitanić, and T. Yolcu. "Optimal contracting with effort and misvaluation". In: *Mathematics and Financial Economics* 7.1 (Sept. 2012), pp. 93–128. DOI: 10.1007/s11579-012-0088-z.

[Car+15] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon. "Second order mean field games with degenerate diffusion and local coupling". In: *Nonlinear Differential Equations and Applications NoDEA* 22 (2015), pp. 1287–1317.

[Car01] G. Carlier. "A general existence result for the principal-agent problem with adverse selection". In: *Journal of Mathematical Economics* 35.1 (Feb. 2001), pp. 129–150. DOI: 10.1016/s0304-4068(00)00057-4.

[CD17] G. Carlier and X. Dupuis. "An iterated projection approach to variational problems under generalized convexity constraints". In: *Applied Mathematics and Optimization* 76.3 (2017), pp. 565–592.

[CZ20] G. Carlier and K. S. Zhang. "Existence of solutions to principal–agent problems with adverse selection under minimal assumptions". In: *Journal of Mathematical Economics* 88 (May 2020), pp. 64–71. DOI: 10.1016/j.jmateco.2020.03.002.

[Car+14] R. Carmichael, J. Schofield, M. Woolf, M. Bilton, R. Ozaki, and G. Strbac. "Residential consumer attitudes to time varying pricing". In: (Sept. 2014). DOI: 10.13140/RG.2.1.4138.1280.

[CFH09] R. Carmona, M. Fehr, and J. Hinz. "Optimal Stochastic Control and Carbon Price Formation". In: *SIAM Journal on Control and Optimization* 48.4 (Jan. 2009), pp. 2168–2190. DOI: 10.1137/080736910.

[CW21] R. Carmona and P. Wang. "Finite-State Contract Theory with a Principal and a Field of Agents". In: *Management Science* 67.8 (Aug. 2021), pp. 4725–4741. DOI: 10.1287/mnsc.2020.3760.

[CD19] J.-P. Chancelier and M. De Lara. *Lower Bound Convex Programs for Exact Sparse Optimization.* 2019. DOI: 10.48550/ARXIV.1902.04813.

[CC59] A. Charnes and W. Cooper. "Chance-Constrained Programming". In: *Management Science* 6 (Oct. 1959), pp. 73–79.

[CXY10] X. Chen, F. Xu, and Y. Ye. "Lower Bound Theory of Nonzero Entries in Solutions of L2-Lp Minimization". In: *SIAM J. Sci. Comp.* 32.5 (2010), pp. 2832–2852.

[CGP15] Y. Chen, T. T. Georgiou, and M. Pavon. "On the Relation Between Optimal Transport and Schrödinger Bridges: A Stochastic Control Viewpoint". In: *Journal of Optimization Theory and Applications* 169.2 (Sept. 2015), pp. 671–691. DOI: 10.1007/s10957-015-0803-z.

[CL22] X. Cheng and Y. Li. "An improved Hoeffding's inequality for sum of independent random variables". In: *Statistics and Probability Letters* 183 (Apr. 2022), p. 109349.

[Chu60] K. L. Chung. *Markov Chains with Stationary Transition Probabilities.* Springer Berlin Heidelberg, 1960. DOI: 10.1007/978-3-642-49686-8.

[Coc+98] J. Cochet-Terrasson, G. Cohen, S. Gaubert, M. McGettrick, and J.-P. Quadrat. "Numerical Computation of Spectral Elements in Max-Plus Algebra ☆". In: *IFAC Proceedings Volumes* 31.18 (July 1998), pp. 667–674.

[Com17] P. L. Combettes. "Perspective Functions: Properties, Constructions, and Examples". In: *Set-Valued and Variational Analysis* 26.2 (Apr. 2017), pp. 247–264. DOI: 10.1007/s11228-017-0407-x.

[Con16] L. Condat. "Fast Projection onto the Simplex and the ℓ_1 Ball". In: *Mathematical Programming, Series A* 158.1 (July 2016), pp. 575–585. DOI: 10.1007/s10107-015-0946-6.

[Cor+96] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth. "On the Lambert W Function". In: *Advances in Computational Mathematics* 5 (Jan. 1996), pp. 329–359. DOI: 10.1007/BF02124750.

[CV95] C. Cortes and V. Vapnik. "Support-vector networks". In: *Machine Learning* 20.3 (Sept. 1995), pp. 273–297. DOI: 10.1007/bf00994018.

[Cou38] A.-A. Cournot. *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot.* L. Hachette, 1838.

[CRE04] CRE. "Electricity Tariff Structure Review: International Comparisons. An Information Paper". 2004.

[CRE21] CRE. *Rapport d'évalulation des tarifs réglementés de vente d'électricité.* 2021.

[Cse20] Z. Csereklyei. "Price and income elasticities of residential and industrial electricity demand in the European Union". In: *Energy Policy* 137 (Feb. 2020), p. 111079. DOI: 10.1016/j.enpol.2019.111079.

[CPT17] J. Cvitanić, D. Possamaï, and N. Touzi. "Dynamic programming approach to principal–agent problems". In: *Finance and Stochastics* 22.1 (Oct. 2017), pp. 1–37. DOI: 10.1007/s00780-017-0344-4.

[CMM98] J. Czyzyk, M. P. Mesnier, and J. J. Moré. “The NEOS server”. In: *IEEE Computational Science and Engineering* 5.3 (1998), pp. 68–75.

[Dam+14] T. Damm, L. Grüne, M. Stieler, and K. Worthmann. “An Exponential Turnpike Theorem for Dissipative Discrete Time Optimal Control Problems”. In: *SIAM Journal on Control and Optimization* 52.3 (Jan. 2014), pp. 1935–1957. DOI: 10.1137/120888934.

[Dav+16] A. Davini, A. Fathi, R. Iturriaga, and M. Zavidovique. “Convergence of the solutions of the discounted Hamilton–Jacobi equation”. In: *Inventiones mathematicae* 206.1 (Jan. 2016), pp. 29–55. DOI: 10.1007/s00222-016-0648-6.

[DZ10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer Berlin Heidelberg, 2010. DOI: 10.1007/978-3-642-03311-7.

[DB01] S. Dempe and J. F. Bard. “Bundle Trust-Region Algorithm for Bilinear Bilevel Programming”. In: *Journal of Optimization Theory and Applications* 110.2 (Aug. 2001), pp. 265–288. DOI: 10.1023/a:1017571111854.

[DD12] S. Dempe and J. Dutta. “Bilevel programming a special Case of a Mathematical Program with Complementarity Constraints?” In: *Math. Program.* 131 (Feb. 2012), pp. 37–48. DOI: 10.1007/s10107-010-0342-1.

[DK16] S. Dempe and F. M. Kue. “Solving discrete linear bilevel optimization problems using the optimal value reformulation”. In: *Journal of Global Optimization* 68.2 (Oct. 2016), pp. 255–277. DOI: 10.1007/s10898-016-0478-5.

[Dem02] S. Dempe. *Foundations of Bi-Level Programming*. Jan. 2002.

[Dem20] S. Dempe. “Bilevel Optimization: Theory, Algorithms, Applications and a Bibliography”. In: *Bilevel Optimization*. Springer International Publishing, 2020, pp. 581–672. DOI: 10.1007/978-3-030-52119-6_20.

[Dem+15] S. Dempe, D. V. Kalashnikov, G. Perez-Valdes, and N. Kalashnykova. *Bilevel Programming Problems*. Jan. 2015. ISBN: 978-3-662-45826-6. DOI: 10.1007/978-3-662-45827-3.

[DF68] E. V. Denardo and B. L. Fox. “Multichain Markov Renewal Programs”. In: *SIAM Journal on Applied Mathematics* 16.3 (May 1968), pp. 468–487. DOI: 10.1137/0116038.

[Den+17] Z. Deng, Y. Tian, C. Lu, and W. Xing. “Globally solving Quadratic programs with convex objective and complementarity constraints via completely positive programming”. In: *Journal of Industrial and Management Optimization* 13 (June 2017), pp. 64–64. DOI: 10.3934/jimo.2017064.

[Dew+08] S. Dewez, M. Labbé, P. Marcotte, and G. Savard. “New formulations and valid inequalities for a bilevel pricing problem”. In: *Operations Research Letters* 36.2 (Mar. 2008), pp. 141–149. DOI: 10.1016/j.orl.2007.03.005.

[DM15] P. Dolan and R. Metcalfe. “Neighbors, knowledge, and nuggets: two natural field experiments on the role of incentives on energy conservation”. In: *Becker Friedman Institute for Research in Economics Working Paper* (2015).

[Don06] D. L. Donoho. “Compressed sensing”. In: *IEEE Transactions on information theory* 52.4 (2006), pp. 1289–1306.

[DW19] L. Dressler and S. Weiergraber. “Alert the inert! switching costs and limited awareness in retail electricity markets”. 2019.

[DW23] L. Dressler and S. Weiergraeber. “Alert the Inert? Switching Costs and Limited Awareness in Retail Electricity Markets”. In: *American Economic Journal: Microeconomics* 15.1 (Feb. 2023), pp. 74–116. DOI: 10.1257/mic.20190163.

[DHR09] J.-P. Dubé, G. J. Hitsch, and P. E. Rossi. “Do Switching Costs Make Markets Less Competitive?” In: *Journal of Marketing Research* 46.4 (Aug. 2009), pp. 435–445. DOI: 10.1509/jmkr.46.4.435.

[DHR10] J.-P. Dubé, G. J. Hitsch, and P. E. Rossi. “State dependence and alternative explanations for consumer inertia”. In: *The RAND Journal of Economics* 41.3 (Aug. 2010), pp. 417–445. DOI: 10.1111/j.1756-2171.2010.00106.x.

[Dub+08] J.-P. Dubé, G. J. Hitsch, P. E. Rossi, and M. A. Vitorino. “Category Pricing with State-Dependent Utility”. In: *Marketing Science* 27.3 (May 2008), pp. 417–429. DOI: 10.1287/mksc.1070.0305.

[Dus+17] J.-P. Dussault, M. Haddou, A. Kadrani, and T. Migot. “How to Compute a M-stationary point of the MPCC”. working paper or preprint. Feb. 2017.

[Dus+20] J.-P. Dussault, M. Haddou, A. Kadrani, and T. Migot. “On approximate stationary points of the regularized mathematical program with complementarity constraints”. In: *Journal of Optimization Theory and Applications* 186.2 (2020), pp. 504–522.

[EM09] I. Ekeland and S. Moreno-Bromberg. “An algorithm for computing solutions of variational problems with global convexity constraints”. In: *Numerische Mathematik* 115.1 (Oct. 2009), pp. 45–69. DOI: 10.1007/s00211-009-0270-2.

[ET99] I. Ekeland and R. Temam. *Convex analysis and variational problems*. SIAM, 1999.

[EMP19] R. Elie, T. Mastrolia, and D. Possamaï. “A Tale of a Principal and Many, Many Agents”. In: *Mathematics of Operations Research* 44.2 (May 2019), pp. 440–467. DOI: 10.1287/moor.2018.0931.

[Éli+20] R. Élie, E. Hubert, T. Mastrolia, and D. Possamaï. “Mean–field moral hazard for optimal energy demand response management”. In: *Mathematical Finance* 31.1 (Oct. 2020), pp. 399–473. DOI: 10.1111/mafi.12291.

[ES17] M. Erlei and H. Schenk-Mathes. “Bounded Rationality in Principal-Agent Relationships”. In: *German Economic Review* 18.4 (Dec. 2017), pp. 411–443. DOI: 10.1111/geer.12111.

[Eyt18] J.-B. Eytard. “A tropical geometry and discrete convexity approach to bilevel programming: application to smart data pricing in mobile telecommunication networks”. PhD thesis. Université Paris-Saclay, 2018.

[Fab13] A. Fabisch. *CMA-ESpp*. https://github.com/AlexanderFabisch/CMA-ESpp. 2013.

[FMP22] D. Faccini, F. Maggioni, and F. A. Potra. “Robust and Distributionally Robust Optimization Models for Linear Support Vector Machine”. In: *Computers and Operations Research* 147 (Nov. 2022), p. 105930. DOI: 10.1016/j.cor.2022.105930.

[Fat97] A. Fathi. “Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens”. In: *Comptes Rendus De L Academie Des Sciences Serie I-mathematique* 324 (1997), pp. 1043–1046.

[Fat08] A. Fathi. “The weak-KAM theorem in Lagrangian dynamics, Preliminary Version Number 10”. https://www.math.u-bordeaux.fr/~pthieull/Recherche/KamFaible/Publications/Fathi2008_01.pdf. 2008.

[FS04] A. Fathi and A. Siconolfi. “Existence of C 1 critical subsolutions of the Hamilton-Jacobi equation”. In: *Inventiones Mathematicae* 155.2 (Feb. 2004), pp. 363–388. DOI: 10.1007/s00222-003-0323-6.

[FS05] A. Fathi and A. Siconolfi. “PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians”. In: *Calculus of Variations* 22.2 (Feb. 2005), pp. 185–228. DOI: 10.1007/s00526-004-0271-z.

[Fer+13] C. G. Fernandes, C. E. Ferreira, Á. J. P. Franco, and R. C. S. Schouery. *The Unit-Demand Envy-Free Pricing Problem.* 2013. arXiv: 1310.0038 [cs.GT].

[Fer+16] C. G. Fernandes, C. E. Ferreira, A. J. Franco, and R. C. Schouery. “The envy-free pricing problem, unit-demand markets and connections with the network pricing problem”. In: *Discrete Optimization* 22 (2016), pp. 141–161. DOI: 10.1016/j.disopt.2015.09.003.

[FTT20] O. Féron, P. Tankov, and L. Tinsi. “Price Formation and Optimal Trading in Intraday Electricity Markets with a Major Player”. In: *Risks* 8.4 (Dec. 2020), p. 133. DOI: 10.3390/risks8040133.

[Fes18] A. Festa. “Domain decomposition based parallel Howard’s algorithm”. In: *Mathematics and Computers in Simulation* 147 (May 2018), pp. 121–139. DOI: 10.1016/j.matcom.2017.04.008.

[Fis+17] M. Fischetti, I. Ljubić, M. Monaci, and M. Sinnl. “On the use of intersection cuts for bilevel optimization”. In: *Mathematical Programming* 172.1-2 (Sept. 2017), pp. 77–103. DOI: 10.1007/s10107-017-1189-5.

[Fis80] C. Fisk. “Some developments in equilibrium traffic assignment”. In: *Transportation Research Part B: Methodological* 14.3 (1980), pp. 243–255. DOI: https://doi.org/10.1016/0191-2615(80)90004-1.

[FL04a] R. Fletcher and S. Leyffer. “Solving mathematical programs with complementarity constraints as nonlinear programs”. In: *Optimization Methods and Software* 19.1 (2004), pp. 15–40. DOI: 10.1080/10556780410001654241.

[FKM18] C. L. Floch, E. C. Kara, and S. Moura. “PDE Modeling and Control of Electric Vehicle Fleets for Ancillary Services: A Discrete Charging Case”. In: *IEEE Transactions on Smart Grid* 9.2 (Mar. 2018), pp. 573–581. DOI: 10.1109/tsg.2016.2556643.

[Fly79] J. Flynn. “Steady State Policies for Deterministic Dynamic Programs”. In: *SIAM Journal on Applied Mathematics* 37.1 (Aug. 1979), pp. 128–147. DOI: 10.1137/0137009.

[For60] R. Fortet. “Applications de l’algèbre de boole en recherche opérationelle”. In: *Revue Français de Recherche Opérationelle* (1960).

[For95] S. Fortune. “Voronoi Diagrams and Delaunay Triangulations”. In: *Lecture Notes Series on Computing.* World Scientific, Jan. 1995, pp. 225–265. DOI: 10.1142/9789812831699_0007.

[FM81] J. Fortuny-Amat and B. McCarl. “A Representation and Economic Interpretation of a Two-Level Programming Problem”. In: *Journal of the Operational Research Society* 32.9 (Sept. 1981), pp. 783–792. DOI: 10.1057/jors.1981.156.

[FH08] M. Freimer and D. Horsky. “Try It, You Will Like It—Does Consumer Learning Lead to Competitive Price Promotions?” In: *Marketing Science* 27.5 (Sept. 2008), pp. 796–810. DOI: 10.1287/mksc.1080.0368.

[FS13] S. G. From and A. W. Swift. “A refinement of Hoeffding’s inequality”. In: *Journal of Statistical Computation and Simulation* 83.5 (May 2013), pp. 977–983. DOI: 10.1080/00949655.2011.644290.

[Fru19] R. Fruit. “Exploration-exploitation dilemma in Reinforcement Learning under various form of prior knowledge”. PhD thesis. Université de Lille 1, Sciences et Technologies; CRIStAL UMR 9189, 2019.

[Fu98] W. Fu. “Penalized regression: the bridge versus the lasso”. In: *J. Comp. Graph. Stat.* 7 (1998), pp. 397–416.

[FL04b] M. Fukushima and G.-H. Lin. “Smoothing methods for mathematical programs with equilibrium constraints”. In: *International Conference on Informatics Research for Development of Knowledge Society Infrastructure*. Apr. 2004, pp. 206–213. ISBN: 0-7695-2150-9. DOI: 10.1109/ICKS.2004.1313426.

[GT11] E. Garibaldi and P. Thieullen. “Minimizing orbits in the discrete Aubry–Mather model”. In: *Nonlinearity* 24 (2011), pp. 563–611.

[GPY13] J. Garnier, G. Papanicolaou, and T.-W. Yang. “Large Deviations for a Mean Field Model of Systemic Risk”. In: *SIAM Journal on Financial Mathematics* 4.1 (Jan. 2013), pp. 151–184. DOI: 10.1137/12087387x.

[GRC09] G. Gasso, A. Rakotomamonjy, and S. Canu. “Recovering sparse signals with a certain family of nonconvex penalties and DC programming”. In: *IEEE Trans. Sign. Proc.* 57.12 (2009), pp. 4686–4698.

[GG10] N. Gast and B. Gaujal. “A mean field approach for optimization in discrete time”. In: *Discrete Event Dynamic Systems* 21.1 (Oct. 2010), pp. 63–101. DOI: 10.1007/s10626-010-0094-3.

[Gau96] S. Gaubert. “On the burnside problem for semigroups of matrices in the (max, +) algebra”. In: *Semigroup Forum* 52.1 (Dec. 1996), pp. 271–292. DOI: 10.1007/bf02574104.

[GMQ11] S. Gaubert, W. McEneaney, and Z. Qu. “Curse of dimensionality reduction in max-plus based approximation methods: Theoretical estimates and improved pruning algorithms”. In: *IEEE Conference on Decision and Control and European Control Conference*. IEEE, Dec. 2011. DOI: 10.1109/cdc.2011.6161386.

[GQS14] S. Gaubert, Z. Qu, and S. Sridharan. “Bundle-based pruning in the max-plus curse of dimensionality free method”. In: *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems July 7-11, 2014. Groningen, The Netherland*. 2014, pp. 166–172.

[GS20] S. Gaubert and N. Stott. “A convergent hierarchy of non-linear eigenproblems to compute the joint spectral radius of nonnegative matrices”. In: *Mathematical Control & Related Fields* 10.3 (2020), pp. 573–590. DOI: 10.3934/mcrf.2020011.

[GMS15] F. Gilbert, P. Marcotte, and G. Savard. “A Numerical Study of the Logit Network Pricing Problem”. In: *Transportation Science* 49 (Jan. 2015), p. 150105061815001. DOI: 10.1287/trsc.2014.0560.

[Gra+16] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros. “Decentralized Convergence to Nash Equilibria in Constrained Deterministic Mean Field Control”. In: *IEEE Transactions on Automatic Control* 61.11 (Nov. 2016), pp. 3315–3329. DOI: 10.1109/tac.2015.2513368.

[GN03] R. Gribonval and M. Nielsen. "Sparse representation in union of bases." In: *IEEE Trans. on Information Theory* 49 (2003), pp. 3320–3325.

[Gri+21] V. Grimm, G. Orlinskaya, L. Schewe, M. Schmidt, and G. Zöttl. "Optimal design of retailer-prosumer electricity tariffs using bilevel optimization". In: *Omega* 102 (July 2021), p. 102327. DOI: `10.1016/j.omega.2020.102327`.

[GSH12] N. Groot, B. D. Schutter, and H. Hellendoorn. "Reverse Stackelberg games, Part I: Basic framework". In: *Proceedings of the 2012 IEEE International Conference on Control Applications*. Dubrovnik, Croatia, Oct. 2012, pp. 421–426.

[Grü+21] L. Grüne, M. A. Müller, C. M. Kellett, and S. R. Weller. "Strict dissipativity for discrete time discounted optimal control problems". In: *Mathematical Control and Related Fields* 11.4 (2021), p. 771. DOI: `10.3934/mcrf.2020046`.

[GG13] W. Guan and A. Gray. "Sparse high-dimensional fractional-norm support vector machine via DC programming". In: *Computational Statistics and Data Analysis* 67 (2013), pp. 136–148.

[GL84] R. Guesnerie and J.-J. Laffont. "A Complete Solution to a Class of Principal-Agent Probems with an Application to the Control of a Self-Managed Firm". In: *Journal of public Economics* 25.3 (1984), pp. 329–369.

[Gur+05] V. Guruswami, J. D. Hartline, A. R. Karlin, D. Kempe, C. Kenyon, and F. McSherry. "On profit-maximizing envy-free pricing." In: *SODA*. Vol. 5. 2005, pp. 1164–1173.

[HT19] M. Haddou and M. T. "A smoothing method for sparse optimization over convex sets". In: *Optimization Letters* 14.5 (2019), pp. 1053–1069. DOI: `10.1007/s11590-019-01408-x`.

[Han+15] J. Han, K. Lee, C. Lee, K.-S. Choi, and S. Park. "Robust optimization approach for a chance-constrained binary knapsack problem". In: *Mathematical Programming* 157.1 (July 2015), pp. 277–296. DOI: `10.1007/s10107-015-0931-0`.

[Han06] N. Hansen. "The CMA evolution strategy: a comparing review". In: *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*. New York: Springer, 2006, pp. 75–102.

[Han+10] N. Hansen, A. Auger, R. Ros, S. Finck, and P. Posik. "Comparing results of 31 algorithms from the black-box optimization benchmarking BBOB-2009". In: *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation*. 2010, pp. 1689–1696.

[HO97] N. Hansen and A. Ostermeier. "Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The (/I,)-ES". In: *Eufit* 97 (1997), pp. 650–654.

[Har28] R. Hartley. "Transmission of Information". In: *Bell System Technical Journal* 7.3 (1928), pp. 535–563.

[Hau00] M. Hauskrecht. "Value-Function Approximations for Partially Observable Markov Decision Processes". In: *Journal of Artificial Intelligence Research* 13 (Aug. 2000), pp. 33–94. DOI: `10.1613/jair.678`.

[Hei+10] G. Heilporn, M. Labbé, P. Marcotte, and G. Savard. "A parallel between two classes of pricing problems in transportation and marketing". In: *Journal of Revenue and Pricing Management* 9 (Jan. 2010), pp. 110–125. DOI: `10.1057/rpm.2009.39`.

[Hen07] R. Henrion. “Structural properties of linear probabilistic constraints”. In: *Optimization* 56.4 (Aug. 2007), pp. 425–440. DOI: 10.1080/02331930701421046.

[Hen04] R. Henrion. “Perturbation Analysis of Chance-constrained Programs under Variation of all Constraint Data”. In: *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, 2004, pp. 257–274. DOI: 10.1007/978-3-642-55884-9_13.

[HR14] C. Herbes and I. Ramme. “Online marketing of green electricity in Germany—A content analysis of providers’ websites”. In: *Energy Policy* 66 (2014), pp. 257–266.

[HL96] O. Hernández-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes*. Springer New York, 1996. DOI: 10.1007/978-1-4612-0729-0.

[Hoe63] W. Hoeffding. “Probability Inequalities for Sums of Bounded Random Variables”. In: *Journal of the American Statistical Association* 58.301 (Mar. 1963), pp. 13–30. DOI: 10.1080/01621459.1963.10500830.

[Hoh20] S. Hohberger. “Dynamic pricing under customer choice behavior for revenue management in passenger railway networks”. PhD thesis. Universität Mannheim, 2020.

[HM91] B. Holmstrom and P. Milgrom. “Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design”. In: *The Journal of Law, Economics, and Organization* 7.special_issue (Jan. 1991), pp. 24–52. DOI: 10.1093/jleo/7.special_issue.24.

[HP10] D. Horsky and P. Pavlidis. “Brand Loyalty Induced Price Promotions: An Empirical Investigation”. In: *SSRN Electronic Journal* (2010). DOI: 10.2139/ssrn.1674765.

[Hor63] M. Horstein. “Sequential transmission using noiseless feedback”. In: *IEEE Transactions on Information Theory* 9.3 (1963), pp. 136–143. DOI: 10.1109/TIT.1963.1057832.

[HMP17] A. Hortaçsu, S. A. Madanizadeh, and S. L. Puller. “Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market”. In: *American Economic Journal: Economic Policy* 9.4 (Nov. 2017), pp. 192–226. DOI: 10.1257/pol.20150235.

[HCM07] M. Huang, P. E. Caines, and R. P. Malhame. “Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized Nash Equilibria”. In: *IEEE Transactions on Automatic Control* 52.9 (2007), pp. 1560–1571.

[HMC06] M. Huang, R. P. Malhame, and P. E. Caines. “Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Commun. Inf. Syst.* 6.3 (2006), pp. 221–252.

[Hua22] Y. Huang. “Tied Goods and Consumer Switching Costs”. In: *Marketing Science* 41.1 (Jan. 2022), pp. 93–114. DOI: 10.1287/mksc.2021.1312.

[Hur+15] T. Huraux, N. Sabouret, Y. Haradji, and F. Sempé. “Simulations multi-agents de l’activité humaine: application dans le contexte énergétique résidentiel français”. In: *APIA* (2015).

[IBM09] IBM ILOG. “Cplex V12. 1: User’s Manual for CPLEX”. In: *International Business Machines Corporation* 46.53 (2009), p. 157.

[JV02] R. Jansen and R. Van Dijk. “Optimal benchmark tracking with small portfolios”. In: *The Journal of Portfolio Management* 28.2 (2002), pp. 33–39.

[Jar+20] F. Jara-Moroni, J. Mitchell, J.-S. Pang, and A. Wächter. "An enhanced logical benders approach for linear programs with complementarity constraints". In: *Journal of Global Optimization* 77 (May 2020). DOI: 10.1007/s10898-020-00905-z.

[Jeb18] T. Jebara. *A refinement of Bennett's inequality with applications to portfolio optimization*. 2018. DOI: 10.48550/ARXIV.1804.05454.

[Jer85] R. G. Jeroslow. "The polynomial hierarchy and a simple model for competitive analysis". In: *Mathematical Programming* 32.2 (June 1985), pp. 146–164. DOI: 10.1007/bf01586088.

[Jul00] B. Jullien. "Participation Constraints in Adverse Selection Models". In: *Journal of Economic Theory* 93.1 (July 2000), pp. 1–47. DOI: 10.1006/jeth.1999.2641.

[KW94] P. Kall and S. Wallace. *Stochastic Programming*. Vol. 46. Jan. 1994. DOI: 10.2307/2584504.

[Kha+22] R. Khanjani-Shiraz, A. Babapour-Azar, Z. Hosseini-Nodeh, and P. M. Pardalos. "Distributionally robust joint chance-constrained support vector machines". In: *Optimization Letters* (Mar. 2022). DOI: 10.1007/s11590-022-01873-x.

[KLM20] Y. Kim, S. Leyffer, and T. Munson. "Mpec methods for bilevel optimization problems". In: *Bilevel Optimization*. New York: Springer, 2020, pp. 335–360. DOI: 10.1007/978-3-030-52119-6_12.

[KM13] A. C. Kizilkale and R. P. Malhame. "Mean field based control of power system dispersed energy storage devices for peak load relief". In: *52nd IEEE Conference on Decision and Control*. IEEE, Dec. 2013. DOI: 10.1109/cdc.2013.6760669.

[Kle+21] T. Kleinert, M. Labbé, I. Ljubić, and M. Schmidt. "A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization". In: *EURO Journal on Computational Optimization* 9 (2021), p. 100007. DOI: https://doi.org/10.1016/j.ejco.2021.100007.

[Kle+20] T. Kleinert, M. Labbé, F. Plein, and M. Schmidt. "Technical Note—There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization". In: *Operations Research* 68.6 (Nov. 2020), pp. 1716–1721. DOI: 10.1287/opre.2019.1944.

[KS23] T. Kleinert and M. Schmidt. "Why there is no need to use a big-M in linear bilevel optimization: a computational study of two ready-to-use approaches". In: *Computational Management Science* 20.1 (Feb. 2023). DOI: 10.1007/s10287-023-00435-5.

[KF00] K. Knight and W. Fu. "Asymptotics for lasso-type estimators". In: *Ann. Stat.* 28 (2000), pp. 1356–1378.

[Kog13] H. Koga. "Characterization of the smooth Rényi Entropy Using Majorization". In: *2013 IEEE Information Theory Workshop (ITW)*. 2013, pp. 1–5. DOI: 10.1109/ITW.2013.6691332.

[KM97] V. N. Kolokoltsov and V. P. Maslov. *Idempotent analysis and its applications*. Vol. 401. Mathematics and its Applications. Dordrecht: Kluwer Academic Publishers Group, 1997, pp. xii+305. ISBN: 0-7923-4509-6.

[Kov16] A. Kovacs. "Bilevel programming approach to optimizing a time-variant electricity tariff for demand response". In: *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, Nov. 2016. DOI: 10.1109/smartgridcomm.2016.7778839.

[KK23] G. Kozanidis and E. Kostarelou. "An Exact Solution Algorithm for Integer Bilevel Programming with Application in Energy Market Optimization". In: *Journal of Optimization Theory and Applications* (Mar. 2023). DOI: 10.1007/s10957-023-02166-8.

[Kur89] M. Kurano. "The Existence of a Minimum Pair of State and Policy for Markov Decision Processes under the Hypothesis of Doeblin". In: *SIAM Journal on Control and Optimization* 27.2 (Mar. 1989), pp. 296–307. DOI: 10.1137/0327016.

[LMS98] M. Labbé, P. Marcotte, and G. Savard. "A bilevel model of taxation and its application to optimal highway pricing". In: *Management science* 44 (1998), pp. 1608–1622.

[LV15] M. Labbé and A. Violin. "Bilevel programming and price setting problems". In: *Annals of Operations Research* 240.1 (Sept. 2015), pp. 141–169. DOI: 10.1007/s10479-015-2016-0.

[LW15] D. Lacker and K. Webster. "Translation invariant mean field games with common noise". In: *Electronic Communications in Probability* 20 (Jan. 2015). DOI: 10.1214/ecp.v20-3822.

[LM09] J.-J. Laffont and D. Martimort. *The Theory of Incentives*. Princeton University Press, Dec. 2009. DOI: 10.2307/j.ctv7h0rwr.

[LLS05] C. M. Lagoa, X. Li, and M. Sznaier. "Probabilistically Constrained Linear Programs and Risk-Adjusted Controller Design". In: *SIAM Journal on Optimization* 15.3 (Jan. 2005), pp. 938–951. DOI: 10.1137/s1052623403430099.

[LBD22] N. Lanzetti, S. Bolognani, and F. Dörfler. "First-order conditions for optimization in the Wasserstein space". In: *arXiv preprint arXiv:2209.12197* (2022).

[LP98] T. Larsson and M. Patriksson. "Side Constrained Traffic Equilibrium Models—Traffic Management Through Link Tolls". In: *Equilibrium and Advanced Transportation Modelling*. Springer US, 1998, pp. 125–151. DOI: 10.1007/978-1-4615-5757-9_7.

[LL06a] J.-M. Lasry and P.-L. Lions. "Jeux à champ moyen. I-Le cas stationnaire". In: *Comptes Rendus Mathématique* 343.9 (2006), pp. 619–625.

[LL06b] J.-M. Lasry and P.-L. Lions. "Jeux à champ moyen. II–Horizon fini et contrôle optimal". In: *Comptes Rendus Mathématique* 343.10 (2006), pp. 679–684.

[LL07] J.-M. Lasry and P.-L. Lions. "Mean field games". In: *Japanese Journal of Mathematics* 2.1 (2007), pp. 229–260.

[LN09] B. Lemmens and R. Nussbaum. *Nonlinear Perron–Frobenius Theory*. Cambridge University Press, 2009. DOI: 10.1017/cbo9781139026079.

[Ley06] S. Leyffer. "Complementarity constraints as nonlinear equations: Theory and numerical experience". In: 2 (Jan. 2006), pp. 169–208. DOI: 10.1007/0-387-34221-4_9.

[LM10] S. Leyffer and T. Munson. "Solving multi-leader–common-follower games". In: *Optimization Methods and Software* 25.4 (2010), pp. 601–623. DOI: 10.1080/10556780903448052.

[LH11] H. Li and W. Huh. "Pricing Multiple Products with the Multinomial Logit and Nested Logit Models: Concavity and Implications". In: *Manufacturing and Service Operations Management* 13 (Oct. 2011), pp. 549–563. DOI: 10.1287/msom.1110.0344.

[Li+19] H. Li, S. Webster, N. Mason, and K. Kempf. "Product-Line Pricing Under Discrete Mixed Multinomial Logit Demand". In: *Manufacturing and Service Operations Management* 21 (2019), pp. 14–28. DOI: 10.1287/msom.2017.0675.

[Lij07] M. G. Lijesen. "The real-time price elasticity of electricity". In: *Energy Economics* 29.2 (Mar. 2007), pp. 249–258. DOI: 10.1016/j.eneco.2006.08.008.

[LB16] C. Liu and M. Belkin. "Clustering with Bregman Divergences: an Asymptotic Analysis". In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett. Vol. 29. Curran Associates, Inc., 2016.

[Liv72] A. N. Livšic. "COHOMOLOGY OF DYNAMICAL SYSTEMS". In: *Mathematics of the USSR-Izvestiya* 6.6 (Dec. 1972), pp. 1278–1301. DOI: 10.1070/im1972v006n06abeh001919.

[Llo82] S. Lloyd. "Least squares quantization in PCM". In: *IEEE Transactions on Information Theory* 28.2 (Mar. 1982), pp. 129–137. DOI: 10.1109/tit.1982.1056489.

[Lov91] W. S. Lovejoy. "Computationally Feasible Bounds for Partially Observed Markov Decision Processes". In: *Operations Research* 39.1 (Feb. 1991), pp. 162–175. DOI: 10.1287/opre.39.1.162.

[Lun+20] J. P. Luna, C. Sagastizábal, J. Filiberti, S. A. Gabriel, and M. Solodov. *Regularized equilibrium problems with equilibrium constraints with application to energy markets*. 2020.

[MS21] D. Maclagan and B. Sturmfels. *Introduction to tropical geometry*. Vol. 161. American Mathematical Society, 2021.

[MMV23] M. Magnani, F. M. Manenti, and P. Valbonesi. "Measuring Switching Costs in the Italian Residential Electricity Market". In: *SSRN Electronic Journal* (2023). DOI: 10.2139/ssrn.4350592.

[MC85] R. Malhame and C.-Y. Chong. "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system". In: *IEEE Transactions on Automatic Control* 30.9 (Sept. 1985), pp. 854–860. DOI: 10.1109/tac.1985.1104071.

[MZ93] S. Mallat and Z. Zhang. "Matching Pursuit in a Time-Frequency Dictionary". In: *IEEE Trans. Signal Processing* 41.12 (1993), pp. 3397–3415.

[MN02] J. Mallet-Paret and R. Nussbaum. "Eigenvalues for a Class of Homogeneous Cone Maps Arising from Max-Plus Operators". In: *Discrete and Continuous Dynamical Systems* 8.3 (2002), pp. 519–562.

[MOA11] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer New York, 2011. DOI: 10.1007/978-0-387-68276-1.

[MR18] J. Martin and A. Réveillac. "A revisit of the Borch rule for the Principal-Agent Risk-Sharing problem". working paper or preprint. Sept. 2018.

[MF94] J. N. Mather and G. Forni. "Action minimizing orbits in Hamiltonian systems". In: *Transition to Chaos in Classical and Quantum Mechanics*. Springer Berlin Heidelberg, 1994, pp. 92–186. DOI: 10.1007/bfb0074076.

[McD89] C. McDiarmid. "On the method of bounded differences". In: *Surveys in Combinatorics, 1989*. Cambridge University Press, Aug. 1989, pp. 148–188. DOI: 10.1017/cbo9781107359949.008.

[McE07] W. M. McEneaney. "A Curse-of-Dimensionality-Free Numerical Method for Solution of Certain HJB PDEs". In: *SIAM Journal on Control and Optimization* 46.4 (Jan. 2007), pp. 1239–1276. DOI: 10.1137/040610830.

[MDG08] W. M. McEneaney, A. Deshpande, and S. Gaubert. "Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs". In: *2008 American Control Conference*. IEEE, June 2008. DOI: 10.1109/acc.2008.4587234.

[MD15] W. M. McEneaney and P. M. Dower. "The Principle of Least Action and Fundamental Solutions of Mass-Spring and N-Body Two-Point Boundary Value Problems". In: *SIAM Journal on Control and Optimization* 53.5 (Jan. 2015), pp. 2898–2933. DOI: 10.1137/130921908.

[McF74] D. McFadden. *Conditional logit analysis of qualitative choice behavior*. 1974.

[MN10] N. Merentes and K. Nikodem. "Remarks on strongly convex functions". In: *Aequationes mathematicae* 80.1-2 (Sept. 2010), pp. 193–199. DOI: 10.1007/s00010-010-0043-0.

[MW65] B. L. Miller and H. M. Wagner. "Chance Constrained Programming with Joint Constraints". In: *Operations Research* 13.6 (Dec. 1965), pp. 930–945. DOI: 10.1287/opre.13.6.930.

[Mir16] J.-M. Mirebeau. "Adaptive, anisotropic and hierarchical cones of discrete convex functions". In: *Numerische Mathematik* 132.4 (2016), pp. 807–853.

[Mir76] J. Mirrlees. "Optimal tax theory". In: *Journal of Public Economics* 6.4 (Nov. 1976), pp. 327–358. DOI: 10.1016/0047-2727(76)90047-5.

[Mir99] J. A. Mirrlees. "The theory of moral hazard and unobservable behaviour: Part I". In: *The Review of Economic Studies* 66.1 (1999), pp. 3–21.

[MT12] K. Mizobuchi and K. Takeuchi. "Using Economic Incentives to Reduce Electricity Consumption: A field Experiment in Matsuyama, Japan". In: *International Journal of Energy Economics and Policy* 2.4 (Aug. 2012), pp. 318–332.

[MB90] J. T. Moore and J. F. Bard. "The Mixed Integer Linear Bilevel Programming Problem". In: *Operations Research* 38.5 (Oct. 1990), pp. 911–921. DOI: 10.1287/opre.38.5.911.

[Mor+23] B. M. Moreno, M. Brégère, P. Gaillard, and N. Oudjane. "A mirror descent approach for Mean Field Control applied to Demande-Side management". In: *arXiv preprint arXiv:2302.08190* (2023).

[MP22] M. Motte and H. Pham. "Mean-field Markov decision processes with common noise and open-loop controls". In: *The Annals of Applied Probability* 32.2 (Apr. 2022). DOI: 10.1214/21-aap1713.

[MZ16] M. Mulder and S. P. Zomer. "Contribution of green labels in electricity retail markets to fostering renewable energy". In: *Energy Policy* 99 (Dec. 2016), pp. 100–109. DOI: 10.1016/j.enpol.2016.09.040.

[MR78] M. Mussa and S. Rosen. "Monopoly and product quality". In: *Journal of Economic theory* 18.2 (1978), pp. 301–317.

[Mys+18] M. Mysliwski, F. Sanches, D. S. Junior, and S. Srisuma. "Implications of Consumer Loyalty for Price Dynamics when Price Adjustment is Costly". In: (2018).

[Nad+20] N. Nadisic, A. Vandaele, N. Gillis, and J. Cohen. "Exact Sparse Nonnegative Least Squares". In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May 2020. DOI: 10.1109/icassp40776.2020.9053295.

[Nas53] J. Nash. "Two-Person Cooperative Games". In: *Econometrica* 21.1 (Jan. 1953), p. 128. DOI: 10.2307/1906951.

[NMS19] T. Ndebele, D. Marsh, and R. Scarpa. "Consumer switching in retail electricity markets: Is price all that matters?" In: *Energy Economics* 83 (Sept. 2019), pp. 88–103. DOI: 10.1016/j.eneco.2019.06.012.

[NS07] A. Nemirovski and A. Shapiro. "Convex Approximations of Chance Constrained Programs". In: *SIAM Journal on Optimization* 17.4 (Jan. 2007), pp. 969–996. DOI: 10.1137/050622328.

[NT15] H.-L. Ngo and D. Taguchi. "Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients". In: *Mathematics of Computation* 85.300 (Oct. 2015), pp. 1793–1819. DOI: 10.1090/mcom3042.

[NC20] A. Niromandfam and S. P. Choboghloo. *Modeling electricity demand, welfare function and elasticity of electricity demand based on the customers risk aversion behavior*. 2020. arXiv: 2010.07600 [eess.SY].

[NYK20] A. Niromandfam, A. S. Yazdankhah, and R. Kazemzadeh. "Modeling demand response based on utility function considering wind profit maximization in the day-ahead market". In: *Journal of Cleaner Production* 251 (Apr. 2020), p. 119317. DOI: 10.1016/j.jclepro.2019.119317.

[Ols09] G. J. Olsder. "Phenomena in Inverse Stackelberg Games, Part 1: Static Problems". In: *Journal of Optimization Theory and Applications* 143.3 (May 2009), pp. 589–600. DOI: 10.1007/s10957-009-9573-9.

[OL13] C. Ong and H. Le Thi. "Learning sparse classifiers with difference of convex functions algorithms". In: *Optimization Methods and Software* 28.4 (Aug. 2013), pp. 830–854. DOI: 10.1080/10556788.2011.652630.

[Out90] J. V. Outrata. "On the numerical solution of a class of Stackelberg problems". In: *ZOR Zeitschrift Operations Research Methods and Models of Operations Research* 34.4 (July 1990), pp. 255–277. DOI: 10.1007/bf01416737.

[Pag87] F. H. Page. "The existence of optimal contracts in the principal-agent model". In: *Journal of Mathematical Economics* 16.2 (Jan. 1987), pp. 157–167. DOI: 10.1016/0304-4068(87)90005-x.

[Pag15] G. Pagès. "Introduction to vector quantization and its applications for numerics". In: *ESAIM: Proceedings and Surveys* 48 (Jan. 2015). Ed. by N. Champagnat, T. Lelièvre, and A. Nouy, pp. 29–79. DOI: 10.1051/proc/201448002.

[Pas09] V. T. Paschos. "An overview on polynomial approximation of NP-hard problems". In: *Yugoslav Journal of Operations Research* 19.1 (2009), pp. 3–40. DOI: 10.2298/YJOR0901003P.

[PRK93] Y. Pati, R. Rezaifar, and P. Krishnaprasa. "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition". In: *27th Asilomar Conf. on Signals Systems and Computers* (1993), pp. 40–44.

[PE17] P. Pavlidis and P. B. Ellickson. "Implications of parent brand inertia for multiproduct pricing". In: *Quantitative Marketing and Economics* 15.4 (July 2017), pp. 369–407. DOI: 10.1007/s11129-017-9187-8.

[PML22] S. Peng, F. Maggioni, and A. Lisser. "Bounds for probabilistic programming with application to a blend planning problem". In: *European Journal of Operational Research* 297.3 (Mar. 2022), pp. 964–976. DOI: 10.1016/j.ejor.2021.09.023.

[PL14] T. Pham Dinh and H. Le Thi. "Recent Advances in DC Programming and DCA". In: *Transactions on Computational Collective Intelligence* 8342 (2014), pp. 1–37.

[Pie06] L. Pierre-Louis. "Théorie des jeux de champ moyen et applications (mean field games)". In: vol. 343. 9. 2006, ttp://www. college-de–france. fr/default/EN/all/equ der/audio video. jsp, 2009 (2007).

[PEC12] M. Pilanci, L. El Ghaoui, and V. Chandrasekaran. "Recovery of Sparse Probability Measures via Convex Programming". In: *Proceedings Conference on Neural Information Processing Systems*. 2012.

[Pin12] R. S. Pindyck. "Uncertain outcomes and climate change policy". In: *Journal of Environmental Economics and Management* 63.3 (May 2012), pp. 289–303. DOI: 10.1016/j.jeem.2011.12.001.

[PM19] S. Pineda and J. M. Morales. "Solving Linear Bilevel Problems Using Big-Ms: Not All That Glitters Is Gold". In: *IEEE Transactions on Power Systems* 34.3 (May 2019), pp. 2469–2471. DOI: 10.1109/tpwrs.2019.2892607.

[Pin89] J. Pinter. "Deterministic approximations of probability inequalities". In: *Zeitschrift für Operations-Research* 33.4 (1989), pp. 219–239.

[Pis05] D. Pisinger. "Where are the hard knapsack problems?" In: *Computers and Operations Research* 32.9 (Sept. 2005), pp. 2271–2284. DOI: 10.1016/j.cor.2004.03.002.

[Pra+18] A. Prasanna, J. Mahmoodi, T. Brosch, and M. K. Patel. "Recent experiences with tariffs for saving electricity in households". In: *Energy Policy* 115 (Apr. 2018), pp. 514–522. DOI: 10.1016/j.enpol.2018.01.044.

[Pré70] A. Prékopa. "On probabilistic constrained programming". In: *Proceedings of the Princeton Symposium on Mathematical Programming*. Princeton University Press, 1970, pp. 113–138. DOI: 10.1515/9781400869930-009.

[Pré95] A. Prékopa. *Stochastic Programming*. Springer Netherlands, 1995. DOI: 10.1007/978-94-017-3087-7.

[Put94] M. L. Puterman. *Markov Decision Processes*. Wiley, Apr. 1994. DOI: 10.1002/9780470316887.

[RDP19] A. M. Radoszynski, V. Dvorkin, and P. Pinson. "Accommodating Bounded Rationality in Pricing Demand Response". In: *2019 IEEE Milan PowerTech*. IEEE, June 2019. DOI: 10.1109/ptc.2019.8810419.

[Raf07] A. Rafiq. "On Mann iteration in Hilbert spaces". In: *Nonlinear Analysis: Theory, Methods and Applications* 66.10 (May 2007), pp. 2230–2236. DOI: 10.1016/j.na.2006.03.012.

[RS13] M. Raginsky and I. Sason. "Concentration of Measure Inequalities in Information Theory, Communications, and Coding". In: *Foundations and Trends in Communications and Information Theory* 10.1-2 (2013), pp. 1–246. ISSN: 1567-2190. DOI: 10.1561/0100000064.

[RH18] H. Ren and T. Huang. "Modeling customer bounded rationality in operations management: A review and research opportunities". In: *Computers and Operations Research* 91 (Mar. 2018), pp. 48–58. DOI: 10.1016/j.cor.2017.11.002.

[Rén+61] A. Rényi et al. "On measures of entropy and information". In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability.* Vol. 1. 547-561. Berkeley, California, USA. 1961.

[RSM16] C. Roberts, E. M. Stewart, and F. Milano. "Validation of the Ornstein-Uhlenbeck process for load modeling based on μPMU measurements". In: *2016 Power Systems Computation Conference (PSCC).* IEEE. 2016, pp. 1–7.

[Roc87] J.-C. Rochet. "A necessary and sufficient condition for rationalizability in a quasilinear context". In: *Journal of Mathematical Economics* 16.2 (Jan. 1987), pp. 191–200. DOI: 10.1016/0304-4068(87)90007-3.

[RC98] J.-C. Rochet and P. Choné. "Ironing, sweeping, and multidimensional screening". In: *Econometrica* (1998), pp. 783–826.

[RW09] R. Rockafellar and R. Wets. *Variational Analysis.* 2009. DOI: 10.1007/978-3-642-02431-3.

[Rog85] W. P. Rogerson. "The First-Order Approach to Principal-Agent Problems". In: *Econometrica* 53.6 (Nov. 1985), p. 1357. DOI: 10.2307/1913212.

[Rus+23] G. D. Russo, A. Lytle, S. Hoffenson, L. Wu, and C. Mahoney. "An experimental study of consumer attitudes and intentions in electricity markets". In: *Cleaner and Responsible Consumption* 9 (June 2023), p. 100116. DOI: 10.1016/j.clrc.2023.100116.

[RS03] A. Ruszczyński and A. Shapiro. "Stochastic Programming Models". In: *Handbooks in Operations Research and Management Science.* Elsevier, 2003, pp. 1–64. DOI: 10.1016/s0927-0507(03)10001-1.

[RP21] J. Ryu and S. Park. *Robust solutions for stochastic and distributionally robust chance-constrained binary knapsack problems.* 2021. DOI: 10.48550/ARXIV.2105.11875.

[SI16] Y. Sakai and K. Iwata. "Extremal relations between Shannon entropy and ℓ_α-norm". In: *2016 International Symposium on Information Theory and Its Applications (ISITA).* 2016, pp. 428–432.

[Sal05] B. Salanié. *The economics of contracts: a primer.* MIT press, 2005.

[Sal22] E. Salies. *A measure of switching Costs in the GB electricity retail market.* Working Papers hal-03592459. HAL, Mar. 2022.

[Sam+12] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. 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. DOI: 10.1109/TSG.2012.2203341.

[San08] Y. Sannikov. "A Continuous-Time Version of the Principal–Agent Problem". In: *Review of Economic Studies* 75.3 (July 2008), pp. 957–984. DOI: 10.1111/j.1467-937x.2008.00486.x.

[Sas18] I. Sason. "Tight Bounds on the Rényi Entropy via Majorization with Applications to Guessing and Compression". In: *Entropy* 20.12 (Nov. 2018), p. 896. DOI: 10.3390/e20120896.

[Sch+16] D. A. Schiro, T. Zheng, F. Zhao, and E. Litvinov. "Convex Hull Pricing in Electricity Markets: Formulation, Analysis, and Implementation Challenges". In: *IEEE Transactions on Power Systems* 31.5 (Sept. 2016), pp. 4068–4075. DOI: 10.1109/tpwrs.2015.2486380.

[Sch85] P. J. Schweitzer. "On undiscounted Markovian decision processes with compact action spaces". In: *RAIRO-Operations Research* 19.1 (1985), pp. 71–86.

[Sco00] J. Scott. "Rational choice theory". In: *Understanding contemporary society: Theories of the present* 129 (2000), pp. 126–138.

[Sha48] C. E. Shannon. "A mathematical theory of communication". In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.

[SK20] H. Shao and A. J. Kleywegt. *Tractable Constrained Optimization over Multiple Product Attributes under Discrete Choice Models*. 2020. arXiv: 2007.09193 [math.OC].

[Sha21] A. Shapiro. "Tutorial on risk neutral, distributionally robust and risk averse multistage stochastic programming". In: *European Journal of Operational Research* 288.1 (Jan. 2021), pp. 1–13. DOI: 10.1016/j.ejor.2020.03.065.

[STH07] R. Shioda, L. Tunçel, and B. Hui. "Applications of deterministic optimization techniques to some probabilistic choice models for product pricing using reservation prices". In: *Pacific Journal of Optimization* 10 (Mar. 2007).

[STM11] R. Shioda, L. Tunçel, and T. Myklebust. "Maximum utility product pricing models and algorithms based on reservation price". In: *Computational Optimization and Applications* 48 (Mar. 2011), pp. 157–198. DOI: 10.1007/s10589-009-9254-5.

[SFJ21] A. Shrivats, D. Firoozi, and S. Jaimungal. *Principal agent mean field games in REC markets*. 2021. DOI: 10.48550/ARXIV.2112.11963.

[SFJ22] A. V. Shrivats, D. Firoozi, and S. Jaimungal. "A mean-field game approach to equilibrium pricing in solar renewable energy certificate markets". In: *Mathematical Finance* 32.3 (Feb. 2022), pp. 779–824. DOI: 10.1111/mafi.12345.

[SC73] M. Simaan and J. B. Cruz. "On the Stackelberg strategy in nonzero-sum games". In: *Journal of Optimization Theory and Applications* 11.5 (1973), pp. 533–555. DOI: 10.1007/bf00935665.

[SBA15] E. Soubies, L. Blanc-Féraud, and G. Aubert. "A Continuous Exact l0 Penalty (CEL0) for Least Squares Regularized Problem". In: *SIAM Journal on Imaging Sciences* 8.3 (2015), pp. 1607–1639. DOI: 10.1137/151003714.

[Sou+11] C. Soussen, J. Idier, D. Brie, and J. Duan. "From Bernoulli–Gaussian Deconvolution to Sparse Signal Restoration". In: *IEEE Transactions on Signal Processing* 59.10 (2011), pp. 4572–4584. DOI: 10.1109/tsp.2011.2160633.

[Sta52] H. von Stackelberg. "Theory of the Market Economy". In: (1952).

[Sta86] R. P. Stanley. "Two poset polytopes". In: *Discrete & Computational Geometry* 1.1 (Mar. 1986), pp. 9–23. DOI: 10.1007/bf02187680.

[SC08] I. Steinwart and A. Christmann. *Support vector machines*. Springer Science & Business Media, 2008.

[Stu02] J. T. Sturluson. "Consumer search and switching costs in electricity retailing". In: *Topics in the Industrial Organization of Electricity Markets* 3 (2002).

[SSC17] H. Sun, C.-L. Su, and X. Chen. "SAA-regularized methods for multiproduct price optimization under the pure characteristics demand model". In: *Mathematical Programming* 165.1 (2017), pp. 361–389.

[TPL08] M. Thiao, T. Pham Dinh, and H. Le Thi. "DC Programming Approach for a Class of Nonconvex Programs Involving L0 Norm". In: *Modelling, Computation and Optimization in Information Systems and Management Sciences: Second International Conference MCO 2008, Metz, France - Luxembourg, September 8-10, 2008. Proceedings*. Springer Berlin Heidelberg, 2008, pp. 348–357.

[Tib96] R. Tibshirani. "Regression Shrinkage and Selection Via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (1996), pp. 267–288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

[Tra09] K. Train. *Discrete Choice Methods With Simulation*. Vol. 2009. Jan. 2009. ISBN: 9780521766555. DOI: 10.1017/CBO9780511805271.

[TY19] N. M. Tran and J. Yu. "Product-Mix Auctions and Tropical Geometry". In: *Mathematics of Operations Research* 44.4 (Nov. 2019), pp. 1396–1411. DOI: 10.1287/moor.2018.0975.

[Tré22] E. Trélat. "Linear turnpike theorem". working paper or preprint. June 2022.

[Tun08] L. Tunçel. "Optimization based approaches to product pricing". In: *Selected Proceedings of ICBME* 2 (2008), pp. 93–102.

[VM47] J. Von Neumann and O. Morgenstern. "Theory of games and economic behavior, 2nd rev". In: (1947).

[WB06] A. Wächter and L. Biegler. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming". In: *Mathematical programming* 106.1 (2006), pp. 25–57.

[WFH13] R. Waeber, P. I. Frazier, and S. G. Henderson. "Bisection search with noisy responses". In: *SIAM Journal on Control and Optimization* 51.3 (Jan. 2013), pp. 2261–2279. DOI: 10.1137/120861898.

[WFP15] X. Wang, N. Fan, and P. M. Pardalos. "Robust chance-constrained support vector machines with second-order moment information". In: *Annals of Operations Research* 263.1-2 (Oct. 2015), pp. 45–68. DOI: 10.1007/s10479-015-2039-6.

[Was14] A. Washburn. *Two-Person Zero-Sum Games*. Springer US, 2014. DOI: 10.1007/978-1-4614-9050-0.

[WP95] T. Westerlund and F. Pettersson. "An extended cutting plane method for solving convex MINLP problems". In: *Computers and Chemical Engineering* 19 (June 1995), pp. 131–136. DOI: 10.1016/0098-1354(95)87027-x.

[Wes+03] J. Weston, A. Elisseeff, B. Scholkopf, and M. Tipping. "Use of the Zero-Norm with Linear Models and Kernel Methods". In: *Journal of Machine Learning Research* 3 (2003), pp. 1439–1461.

[Wie+13] W. Wiesemann, A. Tsoukalas, P.-M. Kleniati, and B. Rustem. "Pessimistic Bilevel Optimization". In: *SIAM Journal on Optimization* 23.1 (Jan. 2013), pp. 353–380. DOI: 10.1137/120864015.

[Zav12] M. Zavidovique. "Strict sub-solutions and Mañé potential in discrete weak KAM theory". In: *Commentarii Mathematici Helvetici* (2012), pp. 1–39. DOI: 10.4171/cmh/247.

[Zav23] M. Zavidovique. *Discrete and Continuous Weak KAM Theory: an introduction through examples and its applications to twist maps*. 2023. arXiv: 2308.06356 [math.DS].

[Zhe17] S. Zheng. “A refined Hoeffding’s upper tail probability bound for sum of independent random variables”. In: *Statistics and Probability Letters* 131 (Dec. 2017), pp. 87–92. DOI: 10.1016/j.spl.2017.08.012.

[ZX19] J. Zhou and Z. Xu. “A simultaneous diagonalization based SOCP relaxation for convex quadratic programs with linear complementarity constraints”. In: *Optimization Letters* 13 (Oct. 2019). DOI: 10.1007/s11590-018-1337-8.

[Zou06] H. Zou. “The adaptive lasso and its oracle properties”. In: *J. Amer. Stat. Ass.* 101 (2006), pp. 1418–1429.

[ZL08] H. Zou and R. Li. “One-Step Sparse Estimates in Nonconcave Penalized Likelihood Models”. In: *Ann. Statist.* 36.4 (2008), pp. 1509–1533.

[Zug+13] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen. “A bilevel model for electricity retailers’ participation in a demand response market environment”. In: *Energy Economics* 36 (Mar. 2013), pp. 182–197. DOI: 10.1016/j.eneco.2012.12.010.

List of Figures

***

List of Tables

***

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Jeux de Stackelberg, tarification optimale et application aux marchés de l'électricité

Mots clés : Tarification, Optimisation bi-niveau, Géométrie tropicale, Jeux à champ moyen, Marchés de l'électricité

Résumé : Dans cette thèse, nous combinons des outils d'optimisation bi-niveaux, de jeux à champ moyen et de contrôle ergodique pour résoudre des problèmes complexes de gestion des prix, en particulier dans la conception optimale de contrats pour les marchés de détail de l'électricité. Tout d'abord, nous formulons le problème comme une interaction meneur-suiveur (jeu de Stackelberg), dans laquelle la décision du client est de nature probabiliste (rationalité limitée). Du point de vue de la géométrie tropicale, nous analysons le choix du client en interprétant ce dernier comme un complexe polyédral (arrangement cellulaire). Nous développons un nouvel algorithme qui exploite cette géométrie sous-jacente et fournissons des résultats sur des instances réalistes du problème de tarification rencontré sur les marchés de détail de l'électricité. Nous étendons ensuite ce modèle de deux manières : premièrement, nous intégrons l'inertie dans la décision des clients, modélisée comme un processus décisionnel de Markov dans lequel les transitions correspondent à des problèmes bi-niveaux. Nous prouvons que le problème de contrôle ergodique associé admet une solution, qui peut être obtenue par la résolution d'un problème aux valeurs propres. Dans un second temps, nous étendons le modèle en optimisant non seulement les coefficients de prix mais aussi la structure du menu tarifaire : la question du nombre optimal de contrats est vue comme la quantification optimale d'un menu d'offres de taille infinie, ce dernier étant décrit comme un programme convexe. Nous développons à cet effet de nouvelles procédures d'élagage, héritées des méthodes max-plus utilisées en contrôle optimal. Par ailleurs, nous étudions les mécanismes incitatifs par le biais de récompenses monétaires en définissant une interaction Principal-Agent entre un détaillant et un continuum d'agents, où chacun agent rivalise pour être le plus économe en énergie (jeux de classement). Nous explicitons à la fois l'équilibre de Nash atteint par les agents et la fonction de récompense optimale à offrir en présence d'une élasticité-prix uniforme au sein d'une population hétérogène. Nous présentons des résultats numériques sur le cas général, et montrons le potentiel de ce jeu de classement comme levier de sobriété. Enfin, nous nous intéressons à deux théories qui apparaissent dans la gestion des prix, à savoir l'optimisation sous contrainte en probabilité et l'optimisation parcimonieuse. Pour le premier, nous analysons la pertinence et le passage à l'échelle d'approximations convexes basées sur les inégalités de concentration. Pour le second, nous introduisons une famille de bornes entropiques pour laquelle nous prouvons la capacité à contrôler la cardinalité de la solution, et que nous intégrons dans des problèmes d'optimisation parcimonieux pour obtenir des approximations non linéaires de ces derniers.

Title : Stackelberg games, optimal pricing and application to electricity markets

Keywords : Pricing, Bilevel optimization, Tropical geometry, Mean-field games, Electricity markets

Abstract : In this PhD dissertation, we combine tools from bilevel optimization, mean-field games and ergodic control to tackle challenging issues in pricing management, especially in the optimal design of contracts for retail electricity markets. First of all, we formulate the problem as a leader-follower interaction (Stackelberg game), in which the customers decision is of probabilistic nature (bounded rationality). Through the tropical geometry viewpoint, we analyze the customers choice by interpreting the latter as a polyhedral complex (cell arrangement). We develop a new algorithm that exploits this underlying geometry, and provide results on realistic instances from the pricing problem faced in retail electricity markets. We then extend this model in two ways: firstly, we incorporate inertia in the customers decision, modeled as a Markov Decision Process in which transitions correspond to bilevel problems. We prove that the associated ergodic control problem admits a solution, that can be obtained through the solving of an eigenvalue problem. Secondly, we extend the model by optimizing not only the price coefficients but also the structure of the tariff menu : the question of the optimal number of contracts is viewed as the optimal quantization of an infinite-size menu of offers, the latter being described as a convex program. We develop to this purpose new pruning procedures, inherited from max-plus based methods used in optimal control. Besides, we study incentive mechanisms through monetary rewards by defining a Principal-Agent interaction between a retailer and a field of agents, where each agent competes with similar ones to be the most energy-compliant customer (ranking games). We make explicit both the Nash equilibrium achieved by the agents and the optimal reward function to offer to a heterogeneous population with uniform price elasticity. We present numerical results on the general case, and show the potential of this ranking game as a sobriety lever. Finally, we study two frameworks that appear in pricing management, that is chance-constrained programming and sparse optimization. For the former, we analyze the tractability of convex conservative approximations based on concentration inequalities. For sparsity concerns, we introduce a family of entropic bounds – proved to control the cardinality requirement – that we embed into sparse optimization problems to derive nonlinear approximations to the latter.

Institut Polytechnique de Paris
91120 Palaiseau, France