ETH zürich Research Collection

Doctoral Thesis

Distributed Control and Game Design: From Strategic Agents to Programmable Machines

Author(s):
Paccagnan, Dario

Publication Date:
2018-12

Permanent Link:

https://doi.org/10.3929/ethz-b-000314981 →

Rights / License:

In Copyright - Non-Commercial Use Permitted →

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.

ETH Library

Dissertation ETH Zurich No. 25597

Distributed control and game design:

From strategic agents to programmable machines

A dissertation submitted to attain the degree of

Doctor of Sciences of ETH Zurich

(Dr. sc. ETH Zurich)

presented by

Dario Paccagnan

Dott. Magistrale, Università degli studi di Padova, Italy
M.Sc. Technical University of Denmark, Denmark

born 19.01.1990 in Treviso
citizen of Italy

accepted on the recommendation of
Prof. Dr. John Lygeros, examiner
Prof. Dr. Andreas Krause, co-examiner
Prof. Dr. Jason R. Marden, co-examiner

2018

ETH Zurich
IfA - Automatic Control Laboratory
ETL, Physikstrasse 3
8092 Zurich, Switzerland

© Dario Paccagnan, 2018
All Rights Reserved

ISBN 978-3-906916-47-7
DOI 10.3929/ethz-b-000314981

Alla mia famiglia

Fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza.

– Dante Alighieri

Acknowledgements

First and foremost, I would like to express my gratitude to my PhD Advisor, Prof. John Lygeros. While this thesis would not be in its current shape without his guidance and support, his contribution is far broader than that. Most of all, I would like to thank him for his unconditional trust, freedom and balance. On one hand, he allowed me to pursue any research direction I found exciting and relevant, on the other he has always been present to counsel me on a range of different topics - technical and non. A special thanks goes for always believing in me and for helping me to set and pursue my goals. Looking back at when I moved my first steps at our institute, I can safely say that I have grown not only on the technical side, but all around as a researcher. Let me keep it simple: thank you for everything!

It was an incredible luck to work with Prof. Jason Marden who agreed to host me in Santa Barbara, without ever meeting with me before. I still remember when I walked in your office for the first time and I told you "I have ten ideas to discuss with you". You listened to me extremely carefully for one hour. As a result, few days later we begun working on a completely different topic. I am not sure whether you simply did not like any of the ideas or maybe thought I should have listened to yours first. Nevertheless, from that point onwards, it has been an amazing journey and our collaboration has been one of the most exciting I have ever had. I would like to thank you for all the time you dedicated me, for you sharp-thinking, for your patience, and, amongst everything, for the enthusiasm you instilled in me: I hope this is just the beginning.

Thanks to Prof. Andreas Krause for agreeing to serve on my PhD committee and for meeting with me to discuss the content of this thesis, regardless of the tight schedule.

I feel truly blessed and honored to have met and interacted with Prof. Maria Elena Valcher. Not only has she given me continuous support during the years, but she has shown me what it means to be an unparalleled advisor. Energy, passion, crystal-clear reasoning, unrestricted support to her students are only a few of her everyday's ingredients. Most importantly, I would like to thank her for being the first person that believed in me, for introducing me to control theory and for allowing me to take off at DTU. Thank you for being a lighthouse.

I am deeply indebted with the new Professors of our institute Prof. Florian Dörfler and Prof. Maryam Kamgarpour. Thank you Florian for "taking me climbing", for all the adventures in the mountains we had together, for the advices you gave me when I was only a newbie at IfA, and for putting me in contact with Prof. Marden. Maryam, a

sincere thanks goes to you for your support and collaboration during the early stages of my PhD. I remember vividly how much help and enthusiasm you put in teaching me how to write my "first article". If all of this is possible, it is in good part also thanks to you.

The atmosphere we breath at IfA is absolutely incredible: a great mix of hard-working, positive energy, and friendship. Sometimes, bent over our papers, we seem to forget how lucky we are. Fortunately (or not), as soon as we step out of the ETH domain, we get reminded of what a privilege this is. I would like to express my greatest gratitude to all those people that made this place what it is. In this regard, there are many colleagues I should thank, but I limit myself to the essential. First of all, I would like to thank Basilio and Francesca for sharing with me a good half of their PhDs, and countless hours on the board conjecturing results that would be proved wrong only a few hours later. I wish both of you a brilliant career as you truly deserve. A very special thanks is due to Nicolò for sharing innumerable moments inside and outside of the office, for the many jokes we have been laughing of, and for the unmissable extra slice of cake (Lucia's). A big thanks goes to the Italian Gang - Basilio, Francesca, Nicolò, Giampaolo, Marcello, Saverio - for the countless adventures, and for making me feel as IfA was the extension of my time in high school. Thank you Tony for being such a positive, cheerful and selfless person. Thank you Peyman for the numerous research discussions we had: I feel there is a lot to learn form you. Thanks to the dream team - Juan, Marcello, Paul - for all the mountains adventures we shared: I will hardly forget when we set foot on the Dammazwillinge for the first time. I am confident you will not forget either. A special thanks goes to Paul for being a terrific office mate, and for sharing the first years of your PhD with me, often out in the Alps. At this point, I feel I owe a big thanks to our secretaries, Sabrina and Tanja, for making everything run incredibly smooth and for being extremely helpful with really anything. "If there is a solution, we will find it" Sabrina told me once, and this pretty much says it all.

At a personal level, I would like to thank all my friends in Zurich and elsewhere, in particular Chiara for the gazillions hours we spent climbing together, Federica for being a terrific housemate, Martina and Alessandro for hosting me after my return from the US. A line of his own is needed (and probably not enough) to thank Alessio: a very special thought goes to you for all the things we went through together, way too many to recall. I will never forget the 20-hours day on the Allighesi ferrata, our first climbing route in the dolomites, or simply the usual giro di Ca' Tron. Amongst everything, I would like to thank you for teaching me to be ambitious, to fight, to never give up and to endure pain. Last, I want to thank you for always finding time to meet with me (the usual run) when I come back from Zurich, regardless of our busy lives.

At this stage, it feels quite difficult to find the right words to thank those who contributed the most to all of this. Thanks to the bottom of my heart to my family and in particular to my parents and brother. Mum and Dad, you have been a guiding force throughout all these years. I would like to thank you for the curiosity you instilled in

me, for giving me a great education and making me understand its importance. I am incredibly grateful for the unconditional support I received, for the freedom you gave me in pursuing my own path, and for teaching me that great things happen only if you dare trying. You have set the bar quite high, and if I'll ever be able to be half as good as you have been, I'll consider myself satisfied. A warm thanks goes to my brother for showing me the way, for all the opportunities you gave me, and for being an example of what one can achieve with a "little" amount of constant effort. Perhaps I should thank you for those initial english lesson you gave me when I was only a newbie, but I am confident the idea of showing me your beloved "green english textbook" already made up for that.

Finally, I owe a huge, sincere and heartfelt thanks to Erisa. First and foremost for teaching me (slowly but surely) the meaning of *we*, for your love and patience, and for bearing with me "doing science", probably the hardest amongst all. On the latter topic, I am indebted with you for reading through my papers many times, for teaching me that "things need to get done" and, more in general, for being a great counsellor. Quite a few things have changed since the first time you "forced" me to buy the Zurich-Copenhagen ticket, but I am very excited about all that come next, *together*.

Dario Paccagnan
Zurich, November 2018

Ringraziamenti

Innanzitutto, vorrei esprimere la mia gratitudine al mio PhD Advisor, il Prof. John Lygeros. Anche se questa tesi non sarebbe nella sua forma attuale senza la sua guida, il suo contributo non si limita certamente a questo. Vorrei ringraziarlo, in particolare, per la fiducia datami, la libertà e l'equilibrio. Da un lato, mi ha permesso di intraprendere qualsiasi direzione di ricerca trovassi interessante, dall'altro è sempre stato presente per consigliarmi su una serie di argomenti diversi - tecnici e non. Devo un ringraziamento speciale al Prof. Lygeros per aver sempre creduto in me e per avermi aiutato a fissare e raggiungere i miei obiettivi. Rivolgendo lo sguardo a quando ho mosso i primi passi nel nostro istituto, posso dire di essere cresciuto non solo dal punto di vista tecnico, ma anche come ricercatore: grazie di tutto!

E' stata una fortuna incredibile lavorare con il Prof. Jason Marden, il quale ha accettato di ospitarmi a Santa Barbara senza avermi incontrato prima. Ricordo bene quando sono entrato nel tuo ufficio per la prima volta dicendo "Ho dieci idee che vorrei discutere con te". Mi hai ascoltato con estrema attenzione per più di un'ora. Pochi giorni dopo, stavamo lavorando su un argomento completamente diverso. Forse non ti sono piaciute le mie idee, o probabilmente avrei dovuto ascoltare prima le tue. Tuttavia, da quel momento in poi, è cominciato un bellissimo viaggio e la nostra collaborazione è stata una delle più entusiasmanti che abbia mai avuto. Vorrei ringraziarti per tutto il tempo che mi hai dedicato, per la tua pazienza e, sopratutto, per l'entusiasmo che mi hai infuso: spero questo sia solo l'inizio.

Vorrei ringraziare il Prof. Andreas Krause per aver accettato il ruolo di esaminatore esterno e per avermi incontrato per discutere il contenuto di questa tesi.

Sono davvero onorato di aver incontrato e interagito con la Prof. Maria Elena Valcher. Non solo mi ha dato un sostegno continuo nel corso degli anni, ma mi ha anche mostrato cosa significa essere un Advisor senza pari. Energia, passione, ragionamento cristallino, sostegno ai suoi studenti, sono solo alcuni degli ingredienti della sua vita quotidiana. Vorrei ringraziarla, in particolare, per essere la prima persona che ha creduto in me, per avermi introdotto alla teoria del controllo e per avermi spinto ad andare al DTU.

Sono profondamente in debito con i nuovi professori del nostro istituto, il Prof. Florian Dörfler e la Prof. Maryam Kamgarpour. Grazie Florian per "portarmi a scalare", per tutte le avventure in montagna che abbiamo avuto insieme, per i consigli che mi hai dato quando ero solo un novizio all'IfA, e per avermi messo in contatto con il Prof. Marden. Maryam, un sincero ringraziamento va a te per il tuo sostegno e la tua collabo-

razione durante le prime fasi del mio dottorato. Ricordo molto bene quanto entusiasmo ed energia hai messo nell'insegnarmi a scrivere il mio "primo articolo". Se tutto questo è possibile, è in buona parte anche grazie a te.

L'atmosfera che respiriamo all'IfA è assolutamente incredibile: un perfetto mix di lavoro, energia positiva e amicizia. A volte, piegati sui nostri articoli, finiamo per dimenticarci di quanto siamo coccolati. Fortunatamente (o forse no), non appena usciamo dal nostro ufficio, ci accorgiamo immediatamente di che privilegio sia questo. Vorrei esprimere la mia più grande gratitudine a tutte le persone che hanno reso questo luogo quello che è. A questo proposito, mentre dovrei ringraziare molti colleghi, mi limiterò solamente all'essenziale. Prima di tutto, vorrei ringraziare Basilio e Francesca per aver condiviso con me una buona metà dei loro dottorati, e innumerevoli ore alla lavagna a congetturare risultati che si sarebbero rivelati errati solo poche ore dopo. Auguro ad entrambi una brillante carriera. Un ringraziamento speciale va a Nicolò per aver condiviso moltissimi momenti dentro e fuori dall'ufficio, per le tante battute di cui abbiamo riso e per l'immancabile fetta di torta in più (confezionata da Lucia). Un sentito ringraziamento va all'Italian Gang - Basilio, Francesca, Nicolò, Giampaolo, Marcello, Saverio - per le innumerevoli avventure, e per avermi fatto sentire come se IfA fosse la continuazione del liceo. Grazie Tony per essere una persona così positiva, allegra ed altruista. Grazie Peyman per le numerose discussioni che abbiamo avuto: ho l'impressione che ci sia molto da imparare da te. Grazie al Dream Team - Juan, Marcello, Paul - per tutte le avventure in montagna che abbiamo vissuto insieme: non dimenticherò mai quando mettemmo piede sul Dammazwillinge per la prima volta. Sono convinto non lo dimenticherete neanche voi. Un ringraziamento speciale va a Paul per essere stato un fantastico compagno di ufficio e per aver condiviso con me, spesso sulle Alpi, i primi anni in IfA. A questo punto, è d'obbligo un ringraziamento alle nostre segretarie, Sabrina e Tanja, per aver fatto funzionare tutto in maniera impeccabile: "se c'è una soluzione, la troveremo" ha detto Sabrina più di una volta.

A livello personale, vorrei ringraziare tutti i miei amici di Zurigo e non solo, in particolare Chiara per le innumerevoli ore passate assieme arrampicando, Federica per essere stata una fantastica compagna di casa, Martina e Alessandro per avermi ospitato dopo il ritorno dagli Stati Uniti. Devo un pensiero molto speciale (e comunque non sufficiente) ad Alessio per tutte le cose che abbiamo vissuto insieme, troppe per ricordarle tutte. Non dimenticherò mai la giornata di 20 ore sulla ferrata Allighesi, la nostra prima via d'arrampicata nelle Dolomiti, o semplicemente il solito giro di Ca' Tron. Vorrei ringraziarti sopratutto per avermi insegnato ad essere ambizioso, a combattere, a non arrendermi mai e a sopportare la fatica. Infine, ti devo ringraziare per aver sempre trovato il tempo di incontrarmi al mio ritorno da Zurigo (la solita corsa), indipendentemente dagli innumerevoli impegni.

Giunto a questo punto, risulta difficile trovare le parole giuste per ringraziare coloro che hanno contribuito maggiormente a tutto questo. Grazie di cuore alla mia famiglia,

in particolare ai miei genitori ed a mio fratello. Mamma e papà, siete stati una forza trainante in tutti questi anni. Vorrei ringraziarvi per la curiosità che mi avete instillato, per avermi dato un'educazione invidiabile e per avermene fatto capire l'importanza. Sono incredibilmente grato per il sostegno incondizionato che mi avete dato in tutti questi anni, per la libertà che ho ricevuto nel seguire la mia strada, e per avermi insegnato che le cose più belle accadono solo a chi ha il coraggio di provare. Avete posto l'asticella molto in alto, e se mai riuscirò a raggiungerne la metà, mi considererò soddisfatto. Un caloroso ringraziamento va a Diego per avermi indicato la strada, per tutte le opportunità che mi hai offerto, e per essere un esempio di ciò che si può ottenere con un "piccolo" sforzo quotidiano. Forse dovrei ringraziarti per quelle prime lezioni di inglese che mi hai dato quando ero solo un principiante, ma sono sicuro che il piacere di mostrarmi il tuo amato "libro verde" ti ha già ripagato.

Infine, devo un enorme, sincero e sentito grazie ad Erisa. Innanzitutto per avermi insegnato il significato della parola *noi*, per il tuo amore e pazienza, e per sopportarmi quando sono "impegnato con la scienza", probabilmente la più difficile tra tutte le prove. Ti devo un ringraziamento speciale per aver riletto più e più volte i miei articoli, per avermi insegnato che "è inutile perdere tempo a cincischiare, bisogna finire il lavoro" e, più in generale, per essere un grande consigliere. Molte cose sono cambiate dalla prima volta che mi hai "costretto" a comprare il biglietto Zurigo-Copenhagen, e sono elettrizzato per tutto ciò che ci aspetta, *insieme*.

Dario Paccagnan
Zurigo, Novembre 2018

Abstract

Large scale systems are forecasted to greatly impact our future lives thanks to their wide ranging applications including cooperative robotics, mobility on demand, resource and task allocation, supply chain management, and many more. While technological developments have paved the way for the realization of such futuristic systems, we have a limited grasp on how to coordinate the behavior of their individual components to achieve the desired global objective.

With the objective of advancing our understanding, this thesis focus on the analysis and coordination of large scale systems without the need of a centralized authority. At a high level, we distinguish these systems depending on wether they are composed of *cooperative* or *non-cooperative* subsystems. In regard to the first class, a key challenge is the design of local decision rules for the individual components to guarantee that the collective behavior is desirable with respect to a global objective. Non-cooperative systems, on the other hand, require a more careful thinking in that the designer needs to take into account the self-interested nature of the agents. In both cases, the need for distributed protocols stems from the observation that centralized decision making is prohibited due to the scale and privacy requirement associated with typical systems.

In the first part of this thesis, we focus on the coordination of a large number of *non-cooperative* agents. More specifically, we consider strategic decision making problems where each agent's objective is a function of the aggregate behavior of the population. Examples are ubiquitous and include social and traffic networks, demand-response markets, vaccination campaigns, to name just a few. We present two cohesive contributions. First, we compare the performance of an equilibrium allocation with that of an optimal allocation, that is an allocation where a common welfare function is maximized. We propose conditions under which all Nash equilibrium allocations are efficient, i.e., are desirable from a macroscopic standpoint. In the journey towards this goal, we prove a novel result bounding the distance between the strategies at a Nash and at a Wardrop equilibrium that might be of independent interest. Second, we show how to derive scalable algorithms that guide agents towards an equilibrium allocation, i.e., a stable configuration where no agent has any incentive to deviate. When the corresponding equilibria are efficient, these algorithms attain the global objective and respect the agents' selfish nature.

In the second part of this thesis, we focus on the coordination of *cooperative* agents. We consider large-scale resource allocation problems, where a number of agents need to be

allocated to a set of resources, with the goal of jointly maximizing a given submodular or supermodular set function. Applications include sensor allocation problems, distributed caching, data summarization, and many more. Since this class of problems is computationally intractable, we aim at deriving tractable algorithms for attaining approximate solutions, ideally with the best possible approximation ratio. We approach the problem from a game-theoretic perspective and ask the following question: how should we design agents' utilities so that any equilibrium configuration recovers a large fraction of the optimum welfare? In order to answer this question, we introduce a novel framework providing a tight expression for the worst-case performance (price of anarchy) as a function of the chosen utilities. Leveraging this result, we show how to design utility functions so as to optimize the price of anarchy by means of a tractable linear program. The upshot of our contribution is the design of algorithms that are distributed, efficient, and whose performance is certified to be on par or better than that of existing (and centralized) schemes.

Sommario

I sistemi tecnologici su larga scala promettono di migliorare sensibilmente la qualità della nostra vita futura grazie alle loro numerose applicazioni, tra cui la robotica cooperativa, la mobilità su richiesta, l'allocazione di risorse, la gestione della supply chain. Nonostante gli sviluppi tecnologici abbiano aperto la strada alla realizzazione di questi sistemi futuristici, abbiamo una conoscenza limitata su come coordinare i singoli componenti per ottenere l'obiettivo macroscopico desiderato.

Questa tesi si concentra sull'analisi e il coordinamento di sistemi su larga scala privi di un'autorità centralizzata, con l'obiettivo di migliorarne la comprensione ed il funzionamento. Ad alto livello, distinguiamo questi sistemi a seconda che essi siano cooperativi o meno. Una sfida chiave in relazione ai sistemi cooperativi è la progettazione di algoritmi di controllo per le singole componenti che garantiscano il raggiungimento di un predeterminato obiettivo globale. I sistemi non cooperativi, d'altra parte, richiedono una maggiore attenzione in quanto è necessario tenere in considerazione la natura egoistica degli agenti. In entrambi i casi, l'utilizzo di protocolli distribuiti è reso necessario dalle dimensioni di tali sistemi e dai requisiti di privacy che vi sono associati.

Nella prima parte di questa tesi, ci concentriamo sul coordinamento di sistemi non cooperativi. Più specificamente, consideriamo problemi strategici in cui l'obiettivo di ciascun agente è influenzato del comportamento aggregato della popolazione. Esempi di tali sistemi comprendono i social networks, le reti stradali, i mercati azionari. Nel seguito presentiamo due risultati coesivi. In primo luogo, confrontiamo la performance di un'allocazione di equilibrio con la performance di un'allocazione ottimale, cioè di un'allocazione in cui viene massimizzata una funzione obiettivo comune. Proponiamo poi condizioni che garantiscono l'efficienza di tutte le allocazioni di equilibrio. Nel percorso verso questo obiettivo, otteniamo un risultato che delimita la distanza tra gli equilibri di Nash e Wardrop e che potrebbe essere di interesse indipendente. In secondo luogo, progettiamo algoritmi scalabili che guidano gli agenti verso un'allocazione di equilibrio, cioè una configurazione stabile in cui nessun agente ha alcun incentivo a deviare. Quando tali equilibri sono efficienti, questi algoritmi raggiungono l'obiettivo globale e rispettano la natura individualistica degli agenti.

Nella seconda parte di questa tesi, ci concentriamo sul controllo di sistemi cooperativi. In particolare, consideriamo problemi di allocazione delle risorse su larga scala, dove un insieme di risorse deve essere assegnato ad un fissato numero di agenti, con l'obiettivo di massimizzare una funzione obiettivo globale, submodulare o supermodu-

lare. Le applicazioni includono problemi di allocazione dei sensori, caching distribuito, data summarization e molto altro ancora. Poiché questa classe di problemi è intrattabile dal punto di vista computazionale, ci prefiggiamo di ricavare soluzioni approssimate con algoritmi efficienti, idealmente con il miglior rapporto di approssimazione possibile. Formuliamo questo problema con il linguaggio della teoria dei giochi e ci poniamo la seguente domanda: come progettare le funzioni obiettivo da assegnare agli agenti in modo che ogni configurazione di equilibrio produca la massima frazione del valore ottimo? Per rispondere a questa domanda, introduciamo un nuovo metodo per calcolare in maniera esatta la qualità di un equilibrio in relazione alle funzioni obiettivo scelte (price of anarchy). Sfruttando questo risultato, mostriamo come costruire tali funzioni obiettivo in modo da massimizzare la performance dei corrispondenti equilibri grazie ad un programma lineare ausiliario. Il risultato finale è la progettazione di algoritmi distribuiti ed efficienti, il cui rapporto di approssimazione è alla pari o superiore a quello di molti schemi (centralizzati) comunemente usati.

Contents

Acknowledgements i

Ringraziamenti v

Abstract ix

Sommario xi

Notation xvii

1 Overview 1

1.1 Outline and contributions . . . 2

1.1.1 Part I: strategic agents . . . 2

1.1.2 Part II: programmable machines . . . 3

1.2 Publications . . . 5

1.2.1 Part I: strategic agents . . . 5

1.2.2 Part II: programmable machines . . . 6

1.2.3 Other publications . . . 6

I Strategic agents: aggregative games 9

2 Introduction 11

2.1 Equilibrium efficiency and algorithms . . . 13

2.2 Related works . . . 14

3 Mathematical preliminaries 17

3.1 Operator properties . . . 17

3.2 Variational inequalities . . . 18

4 Nash and Wardrop equilibria in aggregative games 23

4.1 Equilibria with coupling constraints . . . 23
4.2 Variational reformulations . . . 25
4.3 Sufficient conditions for monotonicity . . . 28
4.3.1 Linear price function . . . 29
4.3.2 Diagonal price function . . . 30
4.4 Appendix . . . 32
4.4.1 Proofs of the results presented in Sections 4.2 and 4.3 . . . 32

5 Equilibria and efficiency in large populations 37
5.1 Distance between Nash and Wardrop equilibria . . . 37
5.2 Equilibrium efficiency: the price of anarchy . . . 40
5.2.1 Linear price function . . . 41
5.2.2 Diagonal price function . . . 42
5.3 Appendix . . . 45
5.3.1 Proofs of the results presented in Section 5.1 . . . 45
5.3.2 Proofs of the results presented in Section 5.2 . . . 47

6 Decentralized algorithms 53
6.1 Best-response algorithm for Wardrop equilibrium . . . 55
6.2 Gradient-based algorithm for Nash and Wardrop equilibria . . . 57
6.3 Appendix . . . 59
6.3.1 Proofs of the results presented in Sections 6.1 and 6.2 . . . 59

7 Applications 63
7.1 Charging of electric vehicles . . . 63
7.1.1 Theoretical guarantees . . . 64
7.1.2 Numerical analysis . . . 66
7.2 Route choice in a road network . . . 70
7.2.1 Theoretical guarantees . . . 73
7.2.2 Numerical analysis . . . 74
7.3 Appendix . . . 76
7.3.1 Proofs of the results presented in Sections 7.1 and 7.2 . . . 76

II Programmable machines: game design **79**

8 Introduction **81**
8.1 The game-design framework . . . 81
8.2 The general multiagent maximum coverage . . . 83
8.2.1 Applications . . . 84
8.3 Related work . . . 85

9 Mathematical preliminaries **89**
9.1 Strategic-form games and equilibrium concepts . . . 89
9.2 Potential games and congestion games . . . 91
9.3 Price of anarchy and smoothness . . . 94
9.3.1 The question of tightness . . . 96
9.4 Complexity of computing equilibria . . . 97

10 Tight price of anarchy and utility design: a linear program approach **101**
10.1 The price of anarchy as performance metric . . . 101
10.2 The limitations of the smoothness framework . . . 105
10.3 A tight price of anarchy . . . 107
10.3.1 Primal formulation . . . 109
10.3.2 Dual formulation . . . 111
10.3.3 Related works . . . 113
10.4 Optimal utility design via linear programming . . . 114
10.5 Appendix . . . 115
10.5.1 Proofs of the results presented in Section 10.2 . . . 115
10.5.2 Proofs of the results presented in Section 10.3 . . . 116
10.5.3 Proofs of the results presented in Section 10.4 . . . 127

11 Submodular, supermodular, covering problems **131**
11.1 The case of submodular welfare function . . . 132
11.1.1 Improved approximation and comparison with existing result . . . 136
11.2 Covering problems . . . 137
11.3 The case of supermodular welfare function . . . 139
11.4 Appendix . . . 141
11.4.1 Proofs of the results presented in Section 11.1 . . . 141

11.4.2 Proofs of the results presented in Section 11.2 145
11.4.3 Proofs of the results presented in Section 11.3 147

12 Applications **149**
12.1 The vehicle target allocation problem 149
12.2 Distributed caching 152

III Conclusion **157**

13 Conclusions and outlook **159**
13.1 Part I: strategic agents 159
13.1.1 Further research directions 159
13.2 Part II: programmable machines 161
13.2.1 Further research directions 162

Bibliography **165**

Notation

Acronyms

ACCE$^{\text{opt}}$	average coarse correlated equilibrium
BR	best-response dynamics
CCE	coarse correlated equilibrium
GMMC	general multiagent weighted maximum coverage
KKT	Karush-Kuhn-Tucker
LP	linear program
MMC	multiagent weighted maximum coverage
MNE	mixed Nash equilibrium
NE	Nash equilibrium
PoA	price of anarchy

Symbols

$\coloneqq$	equal by definition
$\mathbb{N}$, $\mathbb{N}_0$	set of natural numbers, set of natural numbers including zero
$\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$	set of real, positive real, non negative real numbers
$[p]$, $[p]_0$	set of integers $\{1, \dots, p\}$, set of integers v$\{0, 1, \dots, p\}$
$[a, b]$	interval of real numbers $x \in \mathbb{R}$ with $a \leq x \leq b$
$\mathbb{1}_n$, $\mathbb{0}_n$, $e_i \in \mathbb{R}^n$	vector of unit entries, vector of zero entries, i^{th} canonical vector
I_n	identity matrix $I_n \in \mathbb{R}^{n \times n}$
$A \succ 0$ ($\succeq 0$)	positive definite (semi-) $A \in \mathbb{R}^{n \times n}$, i.e., $x^\top A x > 0$ (≥ 0), $\forall x \neq 0$
$\|x\|$	2-norm of $x \in \mathbb{R}^n$
$\|A\|$	induced 2-norm of $A \in \mathbb{R}^{n \times n}$, $\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
$\lambda_{\min}(A)$, $\lambda_{\max}(A)$	minimum, maximum eigenvalue of the symmetric matrix $A \in \mathbb{R}^{n \times n}$
$[A]_{ij} = A_{ij}$,	element in position (i, j) of the matrix A
$A \otimes B$	Kronecker product of the matrices A, B
$[x^i]_{i=1}^m$	stacked vector $[x^i]_{i=1}^m \coloneqq [(x^1)^\top, \dots, (x^m)^\top]^\top = [x^1; \dots; x^m]$, $x^i \in \mathbb{R}^{n \times 1}$

$\Pi_{\mathcal{X}}(y)$	metric projection of $y \in \mathbb{R}^n$ onto $\mathcal{X} \subseteq \mathbb{R}^n$, see Definition 5
$f(x) = \mathcal{O}(g(x))$	big O notation: $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$
$f'(x)$	derivative of differentiable $f : \mathbb{R} \to \mathbb{R}$
$\nabla_x f(x) \in \mathbb{R}^{n\times m}$	Jacobian of differentiable $f : \mathbb{R}^n \to \mathbb{R}^m$, i.e., $[\nabla_x f(x)]_{ij} \coloneqq \frac{\partial g_j(x)}{\partial x^i}$
$\sum_{i=1}^n \mathcal{X}^i$	Minkowski sum of the sets $\{\mathcal{X}_i\}_{i=1}^n$
$\mathrm{VI}(\mathcal{X}, F)$	variational inequality with set $\mathcal{X}$ and operator F, see Definition 4
$\mathcal{U}[a, b]$	uniform distribution on the real interval $[a, b]$
$\mathbb{1}_{\{f(x)\geq 0\}}$	indicator function of the set $\{x \in \mathbb{R}^n \text{ s.t. } f(x) \geq 0\}$, $f : \mathbb{R}^n \to \mathbb{R}$
$\lvert S\rvert$	cardinality of the (finite) set S

Reserved symbols

Part I

M	number of players
n	dimension of players' strategy vectors
$x^i \in \mathbb{R}^n$	strategy vector of player $i \in [M]$
$x^{-i} \in \mathbb{R}^{n(M-1)}$	strategy vector of all players but $i \in [M]$
$\mathcal{X}^i \subseteq \mathbb{R}^n$	local constraint set of player $i \in [M]$
$\mathcal{X}$	product of local constraint sets $\mathcal{X} \coloneqq \mathcal{X}^1 \times \cdots \times \mathcal{X}^M$
$\mathcal{C} \subseteq \mathbb{R}^{Mn}$	coupling constraint set
$\sigma(x) \in \mathbb{R}^n$	average of strategies $\sigma(x) \coloneqq \frac{1}{M}\sum_{i=1}^M x^i$
$J^i(x^i, \sigma(x))$	cost function of player $i \in [M]$, $J^i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
$\mathcal{G}$	aggregative game defined in (4.3)
x_{N}, x_{W}	Nash, Wardrop equilibrium according to Definitions 6 and 7
$J_S(\sigma(x))$	social cost function, $J_S : \mathbb{R}^n \to \mathbb{R}$, see Definition 8

Part II

e	Euler's number
$\mathcal{R}$	set of resources $\mathcal{R} = \{r_1, \ldots, r_m\}$
v_r	value of resource $r \in \mathcal{R}$, $v_r \in \mathbb{R}_{\geq 0}$
n	number of agents
a_i	strategy of player $i \in [n]$
$\mathcal{A}_i$	strategy set of player $i \in [n]$
$\mathcal{A}$	product of agents' strategy sets $\mathcal{A} \coloneqq \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$
$W(a)$	welfare function $W : 2^{\mathcal{R}} \times \cdots \times 2^{\mathcal{R}} \to \mathbb{R}_{\geq 0}$
$\lvert a\rvert_r$	number of agents selecting resource r in allocation a
$w(j)$	welfare basis function $w : [n] \to \mathbb{R}_{\geq 0}$
$u_i(a)$	utility function of agent $i \in [n]$, $u_i : \mathcal{A} \to \mathbb{R}$
$f(j)$	distribution rule $f : [n] \to \mathbb{R}_{\geq 0}$
a^{ne}	pure Nash equilibrium strategy according to Definition 10
f_{G}, f_{SV}, f_{MC}	distribution rules introduced in (10.5) and Definition 24

Complexity classes

$\mathcal{P}$	deterministic polynomial time class
$\mathcal{NP}$	nondeterministic polynomial time class
$\mathcal{PLS}$	polynomial local search class
$\mathcal{PPAD}$	polynomial parity arguments on directed graphs class

CHAPTER 1

Overview

Large scale systems have enormous potential for solving many of the current societal challenges. Robotic networks can operate in post-disaster environments and reduce the impact of nuclear, industrial or natural calamities [Kun+12; Kit+99]. Fleets of autonomous cars are forecasted to revolutionize the future mobility and to reduce traffic congestion as well as pollutant emissions [Spi+14]. Demand-response schemes have the potential to allow for the integration of a large share of renewable resources [Mot+16]. On a smaller scale, swarms of "microbots" promise groundbreaking results in medicine by means of local drug delivery [Ser+15] or microsurgery [ISA02].

While all the above-mentioned systems (and many more) can be thought of as a collection of multiple *subsystems* or *agents*, we distinguish them in two categories depending on wether the corresponding subsystems *are* or *are not* cooperative. An example of cooperative system is that of a drones swarm performing a rescue mission. On the other hand, privately-owned self driving cars are non-cooperative, since each car's objective is that of reaching its destination as swiftly as possible, while respecting the traffic rules. Another example of non-cooperative large scale system is the electricity reserve market, where generators sell their ability to increase or decrease their electricity production to the system operator, whose ultimate objective is that of balancing production and consumption.

One of the *main challenges* in the operation of both types of systems is the design of *local decision rules* for the individual subsystems to guarantee that the collective behavior is desirable with respect to a *global objective* [LM13]. With this respect, non-cooperative systems pose an additional layer of difficulty, in that preexisting local objectives might not be aligned with the system-level goal. As a concrete example consider on-demand ridesharing platforms such as Uber, Lyft or Didi, where agents are represented by human drivers hovering around different city neighbors (Figure 1.1). While drivers' might position themselves in neighbors

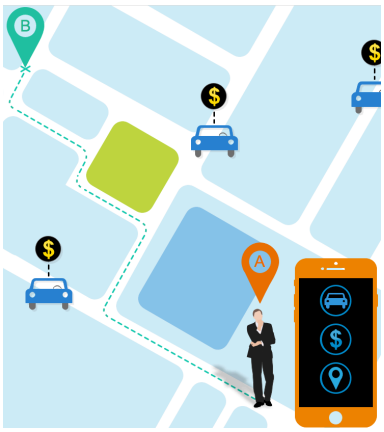

Figure 1.1

that maximize their own profit, the system-operator might have a different goal, e.g., to guarantee a minimum coverage of the city. For this class of systems, it is not sufficient to design local decision rules to be executed from each subsystem, but it is equally important to *incentivize* their adoption.

The spatial distribution, privacy requirements, scale, and quantity of information associated with typical systems do not allow for centralized communication and decision making, but require instead the use of *distributed protocols*. In addition to the above requirements, designing such protocols is non-trivial due to the presence of heterogenous decision makers and informational constraints. To complete the overview, we note that the quality of a control architecture is usually gauged by several metrics including the satisfaction of the global objective, the robustness to external disturbances, as well as the amount of information propagated through the communication network.

The **goal** of this thesis is to address the challenges previously discussed, with particular attention to the design of local decision rules in relation to their corresponding system-level performance. While in the first part of the thesis we focus on large scale *non-cooperative* system, we exploit some of the insight obtained therein to address, in the second part of the thesis, large scale *cooperative* system and propose novel efficient distributed algorithms.

1.1 Outline and contributions

1.1.1 Part I: strategic agents

In the first part of the thesis we focus on large scale systems composed of *strategic* agents. Specifically, we consider the framework of *average aggregative games*, where each agent aims at minimizing a cost function that depends both on his decision and on the average population strategy. Our objective is twofold. First, we wish to understand to what extent selfish decision making reduces the system performance. We do this using the notion of price of anarchy. Second, we aim at the design of scalable and decentralized algorithms that provably converge to a Nash or a Wardrop equilibrium. We achieve this leveraging the theory of variational inequalities.

Outline. Chapter 2 provides an informal introduction to the framework of aggregative games and describes how such games can be used to model applications pertaining to various fields. At the end of the chapter we review the existing literature and connect our work with it. In Chapter 3 we review the mathematical tools needed throughout Part I of this thesis. In Chapter 4 we formalize the notions of Nash and Wardrop equilibria in the presence of coupling constraints, and use the language of variational inequalities to reformulate these problems. We conclude the chapter studying the monotonicity

properties of the variational inequality operator associated with the corresponding equilibrium problems. In the first part of Chapter 5 we study the relation between Nash and Wardrop equilibria with particular attention to the distance between the corresponding strategies. In the second part of this chapter, we leverage these results to bound the performance degradation incurred when moving from a centralized solution to strategic decision making. In Chapter 6 we present a best-response algorithm and a gradient-based algorithm that provably converge to a Nash or Wardrop equilibrium. We conclude Part I with Chapter 7, where we demonstrate the results previously obtained to two large scale applications.

Contributions. The main contributions of Part I of this thesis are contained in Chapters 4, 5, 6 and 7, and are detailed in the following.

(a) In Chapter 4 we introduce the notion of Wardrop equilibrium as a condition on the agents' strategies, rather than a condition on the aggregate behaviour. This allows to address a larger class of equilibrium problems, compared to the existing literature. We then study the relation between Nash and Wardrop equilibrium strategies and show that, in a game with M players, their euclidean distance is upper bounded by $\mathcal{O}(1/\sqrt{M})$ when one of the corresponding variational inequality is strongly monotone (Theorem 1). This allows us to provide guarantees on the efficiency of Nash equilibria by studying the efficiency of the corresponding Wardrop equilibria (Theorems 2, 3 and 4).

(b) In Chapter 6 we present a best-response and a gradient-based algorithm that allow to compute a Nash or a Wardrop equilibrium in the presence of coupling constraints (Theorems 5 and 6).

(c) In Chapter 7 we apply the theoretical results previously derived to i) coordinate the charging profile of a population of electric vehicles, and ii) to predict the travel time distribution for a road traffic network. The results we obtain both in terms of equilibrium efficiency and algorithmic convergence are novel.

1.1.2 Part II: programmable machines

In the second part of the thesis we focus on the control of large scale systems composed of multiple *cooperative* subsystems. We assume that each subsystem (agent) is endowed with computation and communication capabilities, and we aim at achieving a global objective through local coordination of the agents. More specifically, we consider a class of combinatorial allocation problems, where each agent selects a subset of resources with the goal of jointly maximizing a given welfare function, additive over the resources. Since this class of problems is intractable, we seek distributed algorithms that run in polynomial time and give provable approximation guarantees. Rather than directly

specifying a decision making process, we adopt the *game design* approach, and assign to each agent a local utility function. The fundamental question we seek to answer in this part of the thesis is how to design local utility functions so that their selfish maximization recovers a large fraction of the optimal welfare.

Outline. In Chapter 8 we introduce the problem considered, discuss potential applications as well as related works. In Chapter 9 we review the mathematical tools needed for the development of our work. In Chapter 10 we formulate the utility design question and tackle it in two steps. First, we provide performance certificates for a given set of utility functions; second, we show how to design utilities that maximize the corresponding worst-case performance. In Chapter 11 we specialize the results to a class of submodular, supermodular and maximum coverage problems. Finally, in Chapter 12 we present two applications: the vehicle-target assignment problem and a coverage problem arising in distributed caching for mobile networks.

Contributions. The main contributions of Part II of this thesis are contained in Chapters 10, 11 and 12 and are detailed in the following.

(a) In Chapter 10 we pose the utility design problem and adopt the notion of price of anarchy as the worst-case performance metric. We show that traditional approaches used to quantify such performance metric are rather conservative and are not suited for the design problem considered here (Theorem 7). Motivated by this shortcoming, we propose a novel framework to *compute* (Theorems 8 and 9) and *optimize* (Theorem 10) the price of anarchy as a function of the given utilities. In particular, we show that the utility design problem can be reformulated as a tractable linear program. The upshot of this contributions is the possibility to apply the game design procedure to a broad class of problems. To the best of our knowledge, this is the first approach that allows to systematically compute and optimize the price of anarchy.

(b) In Chapter 11 we specialize the previous results to the case of submodular, maximum coverage, and supermodular problems. Relative to the submodular case, we obtain a novel and fully explicit expression for the price of anarchy (Theorem 11). We further apply this result to determine the exact price of anarchy for the Shapley value and marginal contribution design methodologies (Corollary 4). These results are compared with previous (non tight and fragmented) results from the literature, and are placed in the larger context of submodular maximization subject to matroid constraints. Relative to the class of problems considered, we show how optimally-designed utilities provide an approximation ratio superior to the best known ratio $1 - c/e$ of [SVW17].[1] Relative to the case of multiagent maximum

[1] The result of [SVW17] improves on the $(1 - e^{-c})/c$ of [CC84], where c is the (total) curvature of the welfare function [CC84] and e the Euler's number.

coverage problems, we obtain a novel analytical expression for the price of anarchy (Theorem 12), and subsume previous results in [Gai09; RPM17]. Optimally designed utilities achieve a $1 - 1/e$ approximation, the best possible [Fei98]. We conclude the chapter providing a tight expression for the price of anarchy in the case of supermodular welfare function (Theorem 13), and show that our result complements [JM18; PM17b]. Limitedly to this case, we observe that optimally-designed utility functions provide a rather poor approximation ratio.

(c) In Chapter 12 we test the performance of the proposed algorithms on a task-allocation problem, and on a coverage problem arising in distributed mobile networks. We provide thorough simulation results and show the theoretical and numerical advantages of our approach.

1.2 Publications

This thesis contains a subset of the results derived during the author's studies as PhD student at ETH Zurich, all of which have already been published or submitted for publication. The corresponding articles on which this thesis is based are listed below.

1.2.1 Part I: strategic agents

The relations between Nash and Wardrop equilibria presented in Chapter 5, the algorithms developed in Chapter 6 as well as the numerical simulations included in Chapter 7 were developed in collaboration with B. Gentile, F. Parise, M. Kamgarpour and J. Lygeros. The results on the equilibrium efficiency featured in Chapter 5 were derived with the help of F. Parise and J. Lygeros.

[Pac+18] **D. Paccagnan**, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros. "Nash and Wardrop Equilibria in Aggregative Games with Coupling Constraints". In: *IEEE Transactions on Automatic Control* (2018). *Early access*.

[PPL18] **D. Paccagnan**, F. Parise, and J. Lygeros. "On the Efficiency of Nash Equilibria in Aggregative Charging Games". In: *IEEE Control Systems Letters* 2.4 (Oct. 2018), pp. 629–634.

[Pac+16] **D. Paccagnan**, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros. "Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints". In: *2016 IEEE 55th Conference on Decision and Control*. Dec. 2016, pp. 6123–6128.

[PKL16] **D. Paccagnan**, M. Kamgarpour, and J. Lygeros. "On aggregative and mean field games with applications to electricity markets". In: *2016 European Control Conference (ECC)*. June 2016, pp. 196–201.

1.2.2 Part II: programmable machines

The utility design approach presented in Chapter 8, the characterization and optimization of the price of anarchy presented in Chapters 10 and 11 were developed in collaboration with J.R. Marden, with the additional help of R. Chandan limitedly to Theorem 7. The approach, the theoretical findings as well as the numerical studies (Chapter 12) are published in the following papers.

[PCM18] **D. Paccagnan**, R. Chandan, and J. R. Marden. "Distributed resource allocation through utility design – Part I: optimizing the performance certificates via the price of anarchy". In: *arXiv preprint arXiv: 1807.01333* (2018). *Submitted for journal publication.*

[PM18b] **D. Paccagnan** and J. R. Marden. "Distributed resource allocation through utility design – Part II: applications to submodular, supermodular and set covering problems". In: *arXiv preprint arXiv: 1807.01343* (2018). *Submitted for journal publication.*

1.2.3 Other publications

The following papers were published by the author during his doctoral studies, but are not included in this dissertation:

Aggregative games and applications

[Gen+18] B. Gentile, F. Parise, **D. Paccagnan**, M. Kamgarpour, and J. Lygeros. "A game theoretic approach to decentralized charging of plug-in electric vehicles". In: *Challenges in Engineering and Management of Cyber-Physical Systems of Systems*. River Publishers, 2018.

[Bur+17] G. Burger, **D. Paccagnan**, B. Gentile, and J. Lygeros. "Guarantees of convergence to a dynamic user equilibrium for a single arc network". In: Elsevier, July 2017, pp. 9674–9679.

[Gen+17] B. Gentile, **D. Paccagnan**, B. Ogunsola, and J. Lygeros. "A novel concept of equilibrium over a network". In: *2017 IEEE 56th Annual Conference on Decision and Control*. Dec. 2017, pp. 3829–3834.

Utility Design

[PM18a] **D. Paccagnan** and J. Marden. "The Importance of System-Level Information in Multiagent Systems Design: Cardinality and Covering Problems". In: *IEEE Transactions on Automatic Control* (2018). *Early access.*

[PM17a] **D. Paccagnan** and J. R. Marden. "The risks and rewards of conditioning noncooperative designs to additional information". In: *2017 55th Annual Allerton Conference on Communication, Control, and Computing*. Oct. 2017, pp. 958–965.

[RPM17] V. Ramaswamy, **D. Paccagnan**, and J. R. Marden. "The Impact of Local Information on the Performance of Multiagent Systems". In: *arXiv preprint arXiv:1710.01409* (2017). *Submitted for journal publication*.

Others

[PKL15] **D. Paccagnan**, M. Kamgarpour, and J. Lygeros. "On the range of feasible power trajectories for a population of thermostatically controlled loads". In: *2015 54th IEEE Conference on Decision and Control*. Dec. 2015, pp. 5883–5888.

[Jör+14] M. J. Jörgensen, **D. Paccagnan**, N. K. Poulsen, and M. B. Larsen. "IMU calibration and validation in a factory, remote on land and at sea". In: *2014 IEEE/ION Position, Location and Navigation Symposium*. 2014, pp. 1384–1391.

Part I

Strategic agents: aggregative games

CHAPTER 2

Introduction

In the first part of the thesis we consider large scale systems composed of *mutual influencing* and *strategic* agents. We use the term "mutual influencing" to describe the fact that agents' actions have influence on one another, while the term "strategic" captures the self-interested nature of the agents. As an example, consider that of traders in the stock exchange market. In a simplistic setup, each trader's goal is to maximize his profit by carefully buying and selling various financial products. At the same time, the value of one such product depends on what action the other traders take, making the final outcome difficult to predict. While this is only one example, similar scenarios arise in a number of real-life applications ranging from road traffic network to opinion dynamics and even missile defense or racing cars. A setup in which multiple agents behave strategically and influence each others' objectives is typically referred to as a game, and the corresponding field of study termed *game theory*.[1]

Game theory originated as a set of tools to model the interaction of selfish decision makers and has been given formal recognition as an independent research area thanks to the pioneering work of Von Neumann [Neu28] and to the celebrated existence result of Nash [Nas50]. With the modern terminology of game theory, a game is fully specified by four elements:

- **players or agents:** these are the decision makers, e.g., the traders in the stock exchange market. In the following we identify each agent with an index $i \in \{1, \dots, M\}$.

- **strategy or action sets**: these are the actions available to each agent, e.g., which financial products a trader can buy/sell, in what amount, and when. In the following we denote with $\mathcal{X}^i$ the set containing all the possible actions available to agent $i \in \{1, \dots, M\}$.

- **utilities or cost functions**: a measure that quantifies whether the goal of each agent has been satisfied and to what extent. This is typically captured through

[1] The terminology derives from the fact that chess, poker, go, and many other board games are prototypical examples of strategic decision making.

a function mapping an element of the joint action space $\mathcal{X} := \mathcal{X}^1 \times \cdots \times \mathcal{X}^M$ to a real number. In the following we concentrate on cost minimization games and thus introduce the function $J^i : \mathcal{X} \to \mathbb{R}$ representing the cost incurred by player $i \in \{1, \ldots, M\}$, e.g., the negative profit incurred by each trader in the stock market.[2]

- **equilibrium concept**: while player i aims at minimizing his cost function J^i, this function depends on both $x^i \in \mathcal{X}^i$ and the choices of all the other agents, typically referred to as x^{-i}. Thus, we need to define what is a descriptive outcome of the game. This concept is captured by the notion of equilibrium, the most celebrated of which is known as Nash equilibrium. Informally, a joint strategy $x_{\mathrm{N}} \in \mathcal{X}$ is a (pure) Nash equilibrium, if no agent can lower his cost by unilaterally changing his action.

Building on these foundation, we wish to introduce an additional dimension to the problem and to address games with a large number of players. This is motivated by the observation that a relevant number of applications are indeed large scale, e.g., stock exchange markets, road traffic networks, online advertising, and many more. The first difficulty that one is faced with, when thinking about these *large system*, is that of complexity or more formally *tractability*. In order to alleviate this issue, in the remainder of Part I of this thesis, we will consider *aggregative games*. Aggregative games are games where the cost function of each agent does not directly depend on the choice of all the other players, but instead is a function of the *aggregate* players' behaviour. As a purely conceptual example consider the following.

Example 1 (Guess 2/3 of the average)**.** *During the first lecture of the course in algorithmic game theory each student is asked to pick an integer number in the interval* $\{0, \ldots, 100\}$*. The student(s) that selects the number closest to* $2/3$ *of the average wins. What number should you pick?*

This puzzle can be modeled as a game where the players are the students, and the action set of each player is $\{0, \ldots, 100\}$. Further, each player's cost function is captured by the distance from his selection to the $2/3$ of the average. According to the previous definition, this game is aggregative in that the cost function of every player *does not* depend on which number each player selected, but only on an aggregate measure, i.e., the average in this case.[3]

[2]Observe that any cost minimization game can be transformed in a utility maximization game upon reversing the sign of the cost functions.

[3]The answer to this puzzle is more subtle than what it might appear at first, and is more of an exercise in behavioral psychology than a question related to game theory. Indeed, it immediate to observe that the only pure Nash equilibrium of the game consists in all players selecting the number 0. Nevertheless, the fundamental question we need to answer is different: is the notion of Nash equilibrium an appropriate equilibrium concept for the given setup? Real world experiments show that this is not the case, as the average of the players' actions is usually much higher than 0.

Besides Example 1, the aggregative structure arises in various real world applications: in a stock exchange market, the price of a product depends on the total demand and supply, but not on the specific choice of each trader. Similarly, in a road traffic network, the travel time on each link depends (ideally) only on the total number of vehicles on that link.

2.1 Equilibrium efficiency and algorithms

The notion of Nash equilibrium describes a strong stability condition, requiring no agent to be capable of improving by means of unilateral deviations.[4] On the other hand, the quality of an allocation is often measured at the system level with a single scalar cost function $J_{\mathrm{S}} : \mathcal{X} \to \mathbb{R}$.[5] As an example, consider that of a road traffic network, where agents move from origin to destination with the goal of minimizing their own travel time. In this scenario, each agent's cost function captures the time spent on the the road. Nevertheless, a system-level measure describing how well the infrastructure is used is the sum of all users' travel time. Thus, of great interest from a system's perspective is to further understand to what extent equilibrium strategies are efficient. Formally, given a game and a social cost function J_{S}, the efficiency of a Nash equilibrium x_{N} is measured by the ratio between $J_{\mathrm{S}}(x_{\mathrm{N}})$ and the minimum possible social cost, i.e., $\min_{x \in \mathcal{X}} J_{\mathrm{S}}(s)$. The worst-case (best-case) efficiency over all possible equilibria is known as price of anarchy (price of stability). While non uniqueness of the equilibrium set means that these quantities can be quite different, the notion of price of anarchy has received greater attention. Indeed, knowledge of the price of anarchy can be used to bound the efficiency of any possible Nash equilibrium. Additionally, the system regulator can exploit knowledge of the price of anarchy to influence or design better-performing systems. For example, in relation to the road traffic network mentioned previously, the system operator could impose tolls on specific streets or dynamically modify the speed limit so as to improve the efficiency of the overall system. Following this research direction, the first objective of Part I of this thesis is to study the price of anarchy relative to a class of aggregative games.

Once the equilibrium efficiency problem has been settled (and measures have been taken in case of non-satisfactory performance), of fundamental importance is the problem of coordinating the agents towards an equilibrium of the underlying game. With this respect, we are particularly interested in the use of decentralized algorithms. The advantage in using this class of algorithms includes privacy-preserving features and computational tractability. In this spirit, the second objective we pursue in Part I of this

[4] In order to make sense of the following discussion, we assume existence of a Nash equilibrium. We will formally tackle the existence question in Chapter 4.

[5] In some cases the function J_{S} is simply the sum of each agents' cost function. This need not be the general case.

thesis is the development of decentralized algorithms for a class of aggregative games.

To summarize, the **goal** of Part I of this thesis is twofold.

- First, we wish to provide guarantees on the efficiency of Nash and Wardrop equilibria as formally defined in Chapter 4.
- Second, we want to devise decentralized algorithms to coordinate the agents toward one such equilibrium.

2.2 Related works

In this section, we limit ourselves to connect our work and the aggregative game framework with other research threads and models. In particular, we do *not* provide a comparison between the contributions presented in Part I of this thesis and the existing literature. On the contrary, we postpone this task after the presentation of the results in each of the Chapters 5, 6 and 7. This allows us to provide a sharper literature comparison.

Aggregative games

While well-known models studied in game theory belong to the class of *aggregative games* (e.g., the classical Cournot model of competition [Cou38], or the traffic user equilibrium of Wardrop [War52]), a systematic study of this class was initiated only at the turn of the last century, with significant effort coming from the economic literature [Cor94; DHZ06]. Early studies have been devoted to proving existence of the equilibria, and to the analysis of parametric equilibrium problems. A particular class of which is that of comparative statics, where the goal is to predict how the modification of a parameter in the game would alter the set of equilibria [AJ13]. Additional results include convergence analysis for best-response like algorithms, but their scope is generally limited to scalar valued aggregate functions [Jen10; CH12]. Within the engineering and control community there has been a recent surge of interest in the class of aggregative games, in particular because of their potential applications to road traffic dispatch, wireless network routing, and demand-response schemes [Pac+18; Scu+12; MCH13]. Under technical assumptions, gradient-based algorithms have been proposed to coordinate the agents towards a Nash equilibrium, for example in [KNS16; Che+14].

Mean field games

Mean field games are a class of continuous-time dynamic games, where the evolution of each agent's trajectory is governed by a stochastic controlled differential equation. In the

simplest setup, agents are coupled purely through the cost function, which is assumed to depend *only* on the average state of the agents [HCM07]. The analysis is carried out in the limiting regime of large populations, since the problem "simplifies" to a system comprising a Hamilton-Jacobi equation (backward in time, capturing the optimality condition) and a Fokker-Planck equation (forward in time, capturing the distribution of the agents in the state space) [LL07]. While there are some elements of contact between mean field and aggregative games (e.g., the dependence of each agent's cost on the average), some fundamental differences prevent from deeming one class of problems a subset of the other. In particular, the presence of input constraints in aggregative games does not allow for a reformulation in terms of mean field games. The converse is also true, for example due to the fundamental role played by stochasticity in the realm of mean field games.

Population games

A population game consists of a game played by a splittable unit mass of players. To facilitate the comprehension, one can think of this model as a game with infinitely many identical agents. By choosing an action from a finite and common set, each agent receives a payoff that depends on the chosen action and on the total mass of agents selecting the same strategy [SA10]. We note that this class of games differs from that of aggregative games for at least two reasons. First, aggregative games are a modeling language capable of describing games with any number of agents, in contrast to population games. Additionally, the result available for aggregative games are not confined to the limiting case of infinite number of players, but the analysis is possible without restoration to the limit. Second, in aggregative games the strategy sets are typically thought of as continuous sets, while this is not the case for population games. Classical results in population games include, amongst others, convergence analysis of evolutionary dynamics including the replicator dynamics and extension thereof [Bom83; CT14].

CHAPTER 3

Mathematical preliminaries

In this chapter we introduce the mathematical tools required for the development of the first part of this thesis. We begin discussing and connecting useful properties of finite dimensional operators. We then turn our attention to variational inequalities, discuss existence and uniqueness of the solution and present two classical algorithms. While all the material is already available in the literature, we redirect the reader to [FP07] for a comprehensive treatment.

3.1 Operator properties

In this section we introduce some useful properties of finite dimensional operators. Our interest stems from the key role they play in the study of variational inequalities.

Definition 1 (Lipschitz, nonexpansive, contractive)**.** *The operator $F : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz with Lipschitz constant $L > 0$ if*

$$||F(x) - F(y)|| \leq L||x - y|| \qquad \forall x, y \in \mathcal{X}. \tag{3.1}$$

The operator F is non-expansive if (3.1) holds with $L = 1$. The operator F is contractive if (3.1) holds with $L < 1$.

Definition 2 (Monotone and strongly monotone [FP07])**.** *The operator $F : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is strongly monotone with monotonicity constant $\alpha > 0$ if*

$$(F(x) - F(y))^\top (x - y) \geq \alpha \|x - y\|^2 \qquad \forall x, y \in \mathcal{X}. \tag{3.2}$$

The operator F is monotone if (3.2) holds for $\alpha = 0$. □

An example of monotone operator is that of the gradient of a convex function, as detailed in the next proposition.

Proposition 1 (Convex functions have monotone gradients [BC+11, Prop. 17.10])**.** *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be convex, and consider $f : \mathcal{X} \to \mathbb{R}$ a continuously differentiable and (strongly) convex function. The operator $F : \mathcal{X} \to \mathbb{R}^n$ defined by $F(x) = \nabla_x f(x)$ is (strongly) monotone.*

Definition 3 (Co-coercive)**.** *The operator $F : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is co-coercive with constant $\eta > 0$ if*

$$(F(x) - F(y))^\top (x - y) \geq \eta ||F(x) - F(y)||^2 \qquad \forall x, y \in \mathcal{X}.$$

The notion of co-coercivity sits in between that of strong monotonicity and monotonicity. In particular, for a given Lipschitz continuous operator it is possible to show that

$$\text{strong monotonicity} \implies \text{co-coercivity} \implies \text{monotonicity}. \tag{3.3}$$

These results follow directly from the corresponding definitions and can be found in [FP07, p. 164]. The following figure is typically employed to give a visual interpretation of the properties just defined.

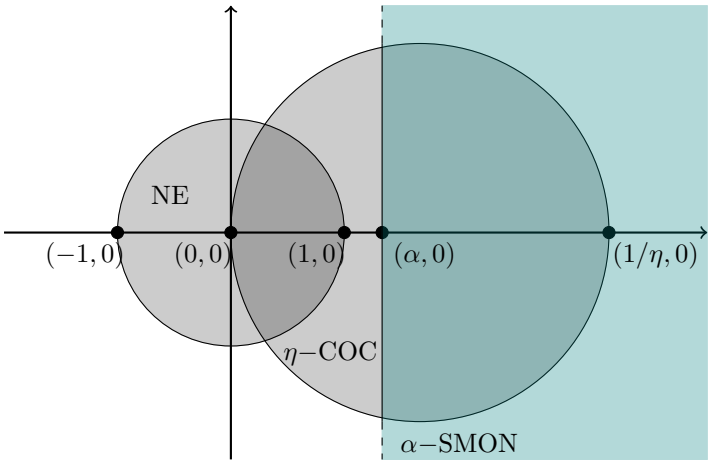

Figure 3.1: Two dimensional representation of nonexpansive operator (NE), co-coercive operator with constant η ($\eta-$COC), and strongly monotone operator with constant α ($\alpha-$SMON). For each of these properties, the corresponding colored region represents the locus of points where $F(1,0)$ must lie, under the assumption that $\mathbb{0}_2$ is a fixed point of F, i.e., that $F(\mathbb{0}_2) = \mathbb{0}_2$. The regions can be easily derived from the corresponding definitions.

3.2 Variational inequalities

Variational inequalities are fundamental mathematical tools that lend their power from their chameleonic nature. Indeed, surprisingly different problems can be formulated and studied using the language of variational inequality. Examples include systems of equations, optimization problems, Nash equilibrium problems, contact problem in mechanics, options pricing. While the term "variational inequality" was coined by Stampacchia in relation to partial differential equation [HS66], in the following we focus on the treatment of finite dimensional variational inequalities, as defined next.

Definition 4 (Variational inequality)**.** *Consider a set $\mathcal{X} \subseteq \mathbb{R}^n$ and an operator $F : \mathcal{X} \to \mathbb{R}^n$. A point $\bar{x} \in \mathcal{X}$ is a solution of the variational inequality VI$(\mathcal{X}, F)$ if*

$$F(\bar{x})^\top (x - \bar{x}) \geq 0 \quad \forall x \in \mathcal{X}. \tag{3.4}$$

Proposition 2 (Existence and uniqueness [FP07, Cor. 2.2.5, Thm. 2.3.3])**.** *Consider the variational inequality VI$(\mathcal{X}, F)$, where $\mathcal{X}$ is compact convex and F continuous.*

(a) The solution set of VI$(\mathcal{X}, F)$ is nonempty and compact.

(b) If the operator F is strongly monotone, the solution of VI$(\mathcal{X}, F)$ is unique.

Connection to convex optimization

The variational inequality problem is tightly connected with that of mathematical programming. In a mathematical program we are given a set $\mathcal{X} \subseteq \mathbb{R}^n$ and a real valued function $f : \mathbb{R}^n \to \mathbb{R}$. Our goal is to select an element of $\mathcal{X}$ that minimizes f over such set. The next proposition makes this connection clear.

Proposition 3 (Minimum principle [BT89, Prop 3.1])**.** *Given $\mathcal{X} \subseteq \mathbb{R}^n$ closed convex and $f : \mathcal{X} \to \mathbb{R}$ continuously differentiable, consider the problem of minimizing f over $\mathcal{X}$.*[1]

(a) If $\bar{x} \in \mathcal{X}$ is a local minimizer of f, then $\bar{x}$ solves VI$(\mathcal{X}, \nabla_x f)$

(b) If f is convex on $\mathcal{X}$, then any solution to VI$(\mathcal{X}, \nabla_x f)$ is a global minimizer of f.

In a nutshell, a convex optimization problem is equivalent to a variational inequality where the operator F represents the gradient of the original function and the set captures the constraint set $\mathcal{X}$. It is important to observe that the converse does *not* hold. Indeed, there are variational inequalities that do not represent the first order condition for any optimization problem. To convince ourselves of this, it suffices to observe that not all operators $F : \mathbb{R}^n \to \mathbb{R}^n$ can be written as the gradient of some underlying function. We also note that the gradient of a strongly convex function is strongly monotone (see Proposition 1), so that existence and uniqueness of the solution is already guaranteed by the corresponding result on variational inequalities presented in Proposition 2. A geometric interpretation of condition (3.4) and the corresponding interpretation in terms of mathematical program is illustrated in the following figure.

[1] In the following, we say that f is continuously differentiable in a *closed* set $\mathcal{X}$ if there exists an *open* set $\mathcal{Y} \supset \mathcal{X}$ where f is continuously differentiable.

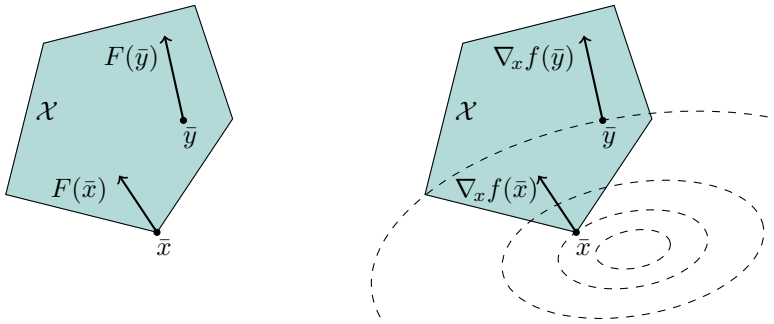

Figure 3.2: On the left: illustration of the condition (3.4) for a general variational inequality. The point $\bar{x}$ is a solution of $\mathrm{VI}(\mathcal{X}, F)$ since the scalar product of $F(\bar{x})$ with any other vector attached to $\bar{x}$ and pointing inside the set $\mathcal{X}$ is non-negative. With a similar reasoning, it is immediate to note that the point $\bar{y}$ is *not* a solution of $\mathrm{VI}(\mathcal{X}, F)$. On the right: the special case of variational inequality $\mathrm{VI}(\mathcal{X}, \nabla_x f)$ corresponding to the convex optimization program $\min_{x\in\mathcal{X}} f(x)$. Similarly to the case on the left, $\bar{x}$ is a solution of $\mathrm{VI}(\mathcal{X}, \nabla_x f)$ and thus a global minimizer of f (see Proposition 3), while $\bar{y}$ is not.

Projection based algorithms

In the following we introduce two classical algorithms for the solution of variational inequalities with a strongly monotone (Algorithm 1) or monotone (Algorithm 2) operator. Before doing so, we recall the definition of metric projection of a point onto a convex set.

Definition 5 (Metric projection)**.** *Given $\mathcal{X} \subseteq \mathbb{R}^n$, we define the metric projection of x onto $\mathcal{X}$ as the map $\Pi_{\mathcal{X}} : \mathbb{R}^n \to \mathbb{R}^n$ with*

$$\Pi_{\mathcal{X}}(x) = \arg\min_{y\in\mathcal{X}} ||y - x||. \tag{3.5}$$

Informally, the projection of x onto the convex set $\mathcal{X}$ is the closest point in $\mathcal{X}$ to x. From the computational point of view, computing the projection of a point onto a convex set amounts to solving the program in (3.5). We observe that the program (3.5) reduces to a quadratic program if $\mathcal{X}$ is a polytope. Since quadratic programs can be solved efficiently, (3.5) can be used as subroutine in the following algorithms.

Algorithm 1 (Projection algorithm)

1: **Initialise** $k = 0$, $\tau > 0$, $x_{(0)} \in \mathbb{R}^n$
2: **while** not converged **do**
3: $\quad x_{(k+1)} = \Pi_{\mathcal{X}}(x_{(k)} - \tau F(x_{(k)}))$
4: $\quad k \leftarrow k + 1$
5: **end while**

Proposition 4 ([FP07, Thm. 12.1.8]). *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be compact convex and $F : \mathbb{R}^n \to \mathbb{R}^n$ be co-coercive with constant η. Then Algorithm 1 converges to a solution of VI$(\mathcal{X}, F)$ for any choice of $\tau < 2\eta$ and $x_{(0)}$.*

Since any strongly monotone and Lipschitz operator is also co-coercive as seen in (3.3), the previous proposition applies in particular to the special case of strongly monotone operators. Observe that strongly convex optimization problems are equivalent to strongly monotone variational inequalities with the corresponding gradient as operator as discussed in Proposition 3. Thus, the previous proposition gives an alternative proof for the convergence of the well-known gradient projection algorithm for strongly convex programs.

If the operator F is not strongly monotone, Algorithm 1 might not converge in general (it does if we restrict ourselves to variational inequalities representing convex optimization problems). A counterexample is provided in [FP07, Ex. 12.1.3]. It is possible to recover convergence of the algorithm at the price of one extra projection per each iteration, as detailed next.

Proposition 5 ([FP07, Thm. 12.1.11]). *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be compact convex and $F : \mathbb{R}^n \to \mathbb{R}^n$ be monotone and Lipschitz with constant L. Then Algorithm 2 converges to a solution of VI$(\mathcal{X}, F)$ for any choice of $\tau < 1/L$ and $x_{(0)}$.*

Algorithm 2 (Extragradient algorithm)

1: **Initialize** $k = 0$, $\tau > 0$, $x_{(0)} \in \mathbb{R}^n$
2: **while** not converged **do**
3: $\quad y_{(k+1)} = \Pi_{\mathcal{X}}(x_{(k)} - \tau F(x_{(k)}))$
4: $\quad x_{(k+1)} = \Pi_{\mathcal{X}}(x_{(k)} - \tau F(y_{(k+1)}))$
5: $\quad k \leftarrow k + 1$
6: **end while**

CHAPTER 4

Nash and Wardrop equilibria in aggregative games

In the first section of this chapter we introduce the class of average aggregative games as well as the notions of Nash and Wardrop equilibrium. In Section 4.2 we show how these can be reformulated as variational inequalities. We conclude the chapter discussing the monotonicity properties of the operators associated to the Nash and Wardrop problems in Section 4.3. All the proofs are reported in the Appendix (Section 4.4). The formulation presented in this chapter has been published in [Pac+18].

4.1 Equilibria with coupling constraints

We consider a population of M agents, where each agent $i \in \{1, \dots, M\}$ can choose a strategy x^i in his individual constraint set $\mathcal{X}^i \subset \mathbb{R}^n$. In addition to the constraint $x^i \in \mathcal{X}^i$, each agent's strategy has to satisfy a coupling constraint, which involves the decision variables of other agents. Upon stacking together the strategies of all players as in $x \coloneqq [x^1; \dots; x^M] \in \mathbb{R}^{Mn}$, the coupling constraint takes the form

$$x \in \mathcal{C} \coloneqq \{x \in \mathbb{R}^{Mn} \,|\, g(x) \leq \mathbb{0}_m\} \subset \mathbb{R}^{Mn}, \qquad g : \mathbb{R}^{Mn} \to \mathbb{R}^m. \tag{4.1}$$

We assume that the cost function of agent i depends on his own strategy x^i and on the strategies of the other agents via the average population strategy $\sigma(x) \coloneqq \frac{1}{M}\sum_{j=1}^{M} x^j$, as typical of aggregative games [Jen10]. The cost function of agent i is identified with $J^i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and takes the form

$$J^i(x^i, \sigma(x)). \tag{4.2}$$

The cost and constraints introduced above give rise to the game $\mathcal{G}$ identified with

$$\mathcal{G} \coloneqq \begin{cases} \text{agents} & \{1, \dots, M\} \\ \text{cost of agent } i & J^i(x^i, \sigma(x)) \\ \text{individual constraint} & \mathcal{X}^i \\ \text{coupling constraint} & \mathcal{C} \end{cases}, \tag{4.3}$$

which is the focus of Part I of this thesis. We denote for convenience $\mathcal{X} \coloneqq \mathcal{X}^1 \times \ldots \times \mathcal{X}^M$ and define

$$\mathcal{Q}^i(x^{-i}) \coloneqq \{x^i \in \mathcal{X}^i \mid g(x) \leq \mathbb{0}_m\}, \qquad \mathcal{Q} \coloneqq \mathcal{X} \cap \mathcal{C}. \tag{4.4}$$

Note that $\mathcal{Q}^i(x^{-i})$ represents the feasible set of player i, given that the other players have selected the strategy x^{-i}, while $\mathcal{Q}$ represents the feasible set for the stacked strategy profile x. We consider two notions of equilibrium for the game $\mathcal{G}$ in (4.3). The first is a generalization of the celebrated Nash equilibrium concept [Nas50] to games with coupling constraints [AD54; Ros65].

Definition 6 (Nash Equilibrium)**.** *A set of strategies $x_{\mathrm{N}} = [x_{\mathrm{N}}^1; \ldots; x_{\mathrm{N}}^M] \in \mathbb{R}^{Mn}$ is an ε-Nash equilibrium of the game $\mathcal{G}$, if $x_{\mathrm{N}} \in \mathcal{Q}$ and for all $i \in \{1, \ldots, M\}$ and all $x^i \in \mathcal{Q}^i(x_{\mathrm{N}}^{-i})$*

$$J^i(x_{\mathrm{N}}^i, \sigma(x_{\mathrm{N}})) \leq J^i\left(x^i, \tfrac{1}{M}x^i + \tfrac{1}{M}\textstyle\sum_{j \neq i} x_{\mathrm{N}}^j\right) + \varepsilon\,. \tag{4.5}$$

If (4.5) holds with $\varepsilon = 0$ then x_{N} is a Nash equilibrium. □

Intuitively, a feasible set of strategies $\{x_{\mathrm{N}}^i\}_{i=1}^M$ is a Nash equilibrium if no agent can lower his cost by unilaterally deviating from his strategy, assuming that the strategies of the other agents are fixed. If no coupling constraint is present, i.e., if $\mathcal{C} = \mathbb{R}^{Mn}$, the previous definition reduces to the well known notion of Nash equilibrium introduced in [Nas50]. In order to differentiate the two definitions, a Nash equilibrium for a game with coupling constraints is usually referred to in the literature as *generalized Nash equilibrium* [FK07]. Nevertheless, in Part I of this thesis we refer to one such equilibrium simply as a Nash equilibrium.

Note that on the right-hand side of (4.5) the decision variable x^i appears in both arguments of $J^i(\cdot, \cdot)$. However, as the number of agents grows the contribution of agent i to $\sigma(x)$ decreases. This motivates the definition of Wardrop equilibrium.

Definition 7 (Wardrop Equilibrium)**.** *A set of strategies $x_{\mathrm{W}} = [x_{\mathrm{W}}^1; \ldots; x_{\mathrm{W}}^M] \in \mathbb{R}^{Mn}$ is a Wardrop equilibrium of the game $\mathcal{G}$ if $x_{\mathrm{W}} \in \mathcal{Q}$ and for all $i \in \{1, \ldots, M\}$ and all $x^i \in \mathcal{Q}^i(x_{\mathrm{W}}^{-i})$*

$$J^i(x_{\mathrm{W}}^i, \sigma(x_{\mathrm{W}})) \leq J^i(x^i, \sigma(x_{\mathrm{W}})).$$ □

Intuitively, a feasible set of strategies $\{x_{\mathrm{W}}^i\}_{i=1}^M$ is a Wardrop equilibrium if no agent can lower his cost by unilaterally deviating from his strategy, assuming that the average strategy is fixed (i.e., he does not influence the average $\sigma(x_{\mathrm{W}})$). Similarly to the terminology introduced to indicate a Nash equilibrium, in the following we will refer to a *generalized Wardrop Equilibrium* simply as a Wardrop equilibrium. The term "Wardrop equilibrium" originates from the fact that Definition 7 can be used to model, amongst others, the equilibrium concept introduced in [War52] in relation to the study of road traffic networks and often referred to as *traffic user equilibrium* or *Wardrop equilibrium*.

Remark 1 (On the definition of Wardrop equilibrium)**.** *Even though the notion of Wardrop equilibrium is thought of as a classical concept, the existing literature defines a Wardrop equilibrium only in terms of the aggregate behaviour* $\sigma(x)$ *[War52; ABS02; AW04; MW95; DN87], while Definition 7 is presented in terms of the agents' strategies* $\{x^i\}_{i=1}^M$*. It is important to observe that Definition 7 can be reformulated as a condition on the aggregate* $\sigma(x)$ only *in specific cases (e.g., for applications in transportation networks [War52] or competitive markets [DN87]), while there are games for which one such aggregate reformulation is just not possible. Thus, Definition 7 is not a mere revisitation of the classical notion of Wardrop equilibrium, but instead can be used to address a larger class of equilibrium problems. Additionally, all the aforementioned works define a Wardrop equilibrium in relation to a specific application, and thus restrict themselves to specific constraint sets or cost functions. On the other hand Definition 7 does not pose any such limitation.*

To the best of our knowledge, the first formulation of a Wardrop equilibrium in terms of agents' strategies appears in [MCH13; Gra+16], where however it is not recognized as an equilibrium concept on its own, but rather characterized as an ε*-Nash equilibrium for an appropriate value of* ε*.*

4.2 Variational reformulations

In this section we show how Nash and Wardrop equilibria introduced in Definitions 6 and 7 can be obtained by solving a corresponding variational inequality. The connection we will draw between these equilibrium notions and the theory of variational inequalities is *fundamental* for the development of Part I of this thesis. As a matter of fact, most of the results we will derive in relation to the concepts of Nash and Wardrop equilibria are based on the analysis of their corresponding variational inequalities.

Recall from Definition 4 that a variational inequality is fully specified by its constraint set $\mathcal{X}$ and operator F (see Chapter 3 for a brief introduction to the theory of variational inequalities). Towards this goal, we introduce the operators $F_{\mathrm{N}},\ F_{\mathrm{W}} : \mathcal{X} \to \mathbb{R}^{Mn}$, where

$$F_{\mathrm{N}}(x) \coloneqq [\nabla_{x^i} J^i(x^i, \sigma(x))]_{i=1}^M \,, \tag{4.6a}$$

$$F_{\mathrm{W}}(x) \coloneqq [\nabla_{x^i} J^i(x^i, z)_{|z=\sigma(x)}]_{i=1}^M \,. \tag{4.6b}$$

The operator F_{N} is obtained by stacking together the gradients of each agent's cost with respect to his decision variable. F_{W} is obtained similarly, but considering $\sigma(x)$ as fixed when differentiating. The following proposition provides a sufficient characterization of the Nash and Wardrop equilibria introduced in Definitions 6 and 7 as solutions of two variational inequalities. Both variational inequalities feature the same constraint set $\mathcal{Q}$, defined in (4.4), but different operators F_{N} and F_{W}, defined in (4.6a) and (4.6b).

Assumption 1. *For all $i \in \{1, \dots, M\}$, the constraint set $\mathcal{X}^i$ is closed and convex. The set $\mathcal{Q}$ in (4.4) is non-empty. The cost functions $J^i(x^i, \sigma(x))$ are convex in x^i for any fixed $x^j \in \mathcal{X}^j$, $j \neq i$. The cost functions $J^i(x^i, z)$ are convex in x^i for any $z \in \frac{1}{M}\sum_{j=1}^{M} \mathcal{X}^j$. The cost functions $J^i(z_1, z_2)$ are continuously differentiable in $[z_1; z_2]$ for any $z_1 \in \mathcal{X}^i$ and $z_2 \in \frac{1}{M}\sum_{j=1}^{M} \mathcal{X}^j$. The function g in (4.1) is convex.*

Proposition 6. *Under Assumption 1, the following hold.*

(a) Any solution $\bar{x}_{\mathrm{N}}$ of VI$(\mathcal{Q}, F_{\mathrm{N}})$ is a Nash equilibrium of the game $\mathcal{G}$ in (4.3).

(b) Any solution $\bar{x}_{\mathrm{W}}$ of VI$(\mathcal{Q}, F_{\mathrm{W}})$ is a Wardrop equilibrium of the game $\mathcal{G}$ in (4.3).

Proposition 6 states that any solution of the variational inequality VI$(\mathcal{Q}, F_{\mathrm{N}})$ is a Nash equilibrium and, similarly, any solution of VI$(\mathcal{Q}, F_{\mathrm{W}})$ is a Wardrop equilibrium. The converse *does not hold* in general, in that there might be strategy profiles that are Nash equilibria but do not satisfy the corresponding variational inequality. This is due to the presence of the coupling constraint $\mathcal{C}$. Indeed, if $\mathcal{C} = \mathbb{R}^{Mn}$, then $\mathcal{Q} = \mathcal{X}$ and one can show that x_{N} solves the VI$(\mathcal{X}, F_{\mathrm{N}})$ *if and only if* it is a Nash equilibrium of $\mathcal{G}$ [FK07, Cor. 1]. A similar result holds in the case of Wardrop equilibrium. The equilibria that can be obtained as solution of the corresponding variational inequality are called *variational equilibria* [FK07, Def. 3] and are here denoted with $\bar{x}_{\mathrm{N}}, \bar{x}_{\mathrm{W}}$ instead of $x_{\mathrm{N}}, x_{\mathrm{W}}$ (indicating any equilibrium satisfying Definition 6 or Definition 7). We next provide sufficient conditions for the existence and uniqueness of variational equilibria by exploiting two well-known results in the theory of variational inequalities.

Lemma 1. [FP07, Cor. 2.2.5, Thm. 2.3.3] *Let Assumption 1 hold.*

(a) If $\mathcal{Q}$ is bounded, then there exist a variational Nash equilibrium and a variational Wardrop equilibrium.[1]

(b) If F_{N} is strongly monotone on $\mathcal{Q}$, then the variational Nash equilibrium is unique. If F_{W} is strongly monotone on $\mathcal{Q}$ then the variational Wardrop equilibrium is unique.

In light of Proposition 6, the proof of the first statement in Lemma 1 amounts to showing that Assumption 1 ensures the existence of a solution to VI$(\mathcal{Q}, F_{\mathrm{N}})$ and VI$(\mathcal{Q}, F_{\mathrm{W}})$. This is guaranteed if the constraint set $\mathcal{Q}$ is compact and convex, and the operator is continuous [FP07, Cor. 2.2.5]. Such conditions follow immediately form Assumption 1. Similarly, the proof of the second statement relies on the fact that the solution of a variational inequality is unique if the constraint set $\mathcal{Q}$ is compact and convex, and the operator is continuous and strongly monotone [FP07, Thm. 2.3.3]. The proofs are not reported here, but can be found in the above-mentioned references.

[1] The convexity of the cost functions required by Assumption 1 is not needed for the first statement of Lemma 1, continuity is enough.

Since any variational Nash equilibrium is a Nash equilibrium, the first claim in Lemma 1 guarantees the existence of a Nash equilibrium. A similar conclusion hold for the existence of a Wardrop equilibrium.

A hierarchy of equilibria: variational and normalized equilibria

The notion of games with coupling constraints has been introduced in the seminal works [AD54; Ros65]. In [Ros65] the author defines the concept of *normalized equilibria* to describe the fact that one should expect a manifold of equilibria when the agents are subject to a coupling constraint, even under strong monotonicity conditions. Formally, the strategy profile x_{N} is a normalized Nash equilibrium if there exists a vector of weights $r \in \mathbb{R}^M_{>0}$, such that x_{N} solves the $\mathrm{VI}(\mathcal{Q}, F^r_{\mathrm{N}})$ where $F^r_{\mathrm{N}}(x) \coloneqq [r_i \nabla_{x^i} J^i(x^i, \sigma(x))]_{i=1}^M$. It is proven that any normalized Nash equilibrium is a Nash equilibrium in the sense of Definition 6. Additionally, [Ros65] shows that different choices of r correspond to a different division of the burden of satisfying the coupling constraints $\mathcal{C}$ among the agents. In the context of aggregative games, however, each agent contributes equally to the average. Therefore it is typically assumed that the burden of satisfying the coupling constraint should also be split equally among the agents by selecting $r = \mathbb{1}_M$, see [FK07; PP09; FFP07]. It is immediate to see that the subclass of normalized equilibria for which this property holds is the class of *variational equilibria* introduced in the previous section. Nonetheless we note that our results could be easily extended to normalized equilibria by using the operator F^r_{N} instead of F_{N}. We conclude observing that the set of Nash equilibria, normalized Nash equilibria and variational Nash equilibria are all nested as in Figure 4.1. A similar result holds for Wardrop equilibria.

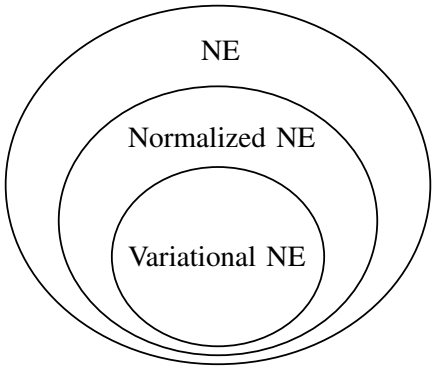

Figure 4.1: The set of Nash equilibria (NE), normalized NE and variational NE are all nested.

In the following we exemplify how the presence of the coupling constraint $\mathcal{C}$ is typically associated with a manifold of equilibria, regardless of the monotonicity properties of the operators F_{N} or F_{W}.

Example 2 (Coupling constraints and manifold of equilibria)**.** Consider the aggregative game $\mathcal{G}$ defined as in (4.3) where there are only two players, and

$$\begin{aligned} \mathcal{X}^1 &= \{x^1 \in \mathbb{R} \,|\, 0 \leq x^1 \leq 1\}, \\ \mathcal{X}^2 &= \{x^2 \in \mathbb{R} \,|\, 0 \leq x^2 \leq 1\}, \\ J^1(x^1, \sigma(x)) &= \frac{3}{2}(x^1)^2 - 2\sigma(x)x^1, \\ J^2(x^2, \sigma(x)) &= 2\sigma(x)x^2. \end{aligned} \tag{4.7}$$

We first study the case where there is no coupling constraint, i.e., $\mathcal{C} = \mathbb{R}^2$, and observe that for such game Assumption 1 is satisfied. Thus, any Nash equilibrium is a solution of $\mathrm{VI}(\mathcal{X}, F_{\mathrm{N}})$ and *vice versa* as discussed immediately after Proposition 6. The operator F_{N} and the corresponding $\nabla_x F_{\mathrm{N}}(x)$ are obtained from (4.6a) as

$$F_{\mathrm{N}}(x^1, x^2) = \begin{bmatrix} x^1 - x^2 \\ x^1 + 2x^2 \end{bmatrix} \qquad \nabla_x F_{\mathrm{N}}(x^1, x^2) = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}.$$

Lemma 2 in Section 4.3 ensures that F_{N} is strongly monotone since $\nabla_x F_{\mathrm{N}}(x^1, x^2) + \nabla_x F_{\mathrm{N}}(x^1, x^2)^\top \succ 0$. Thus, the solution of the variational inequality $\mathrm{VI}(\mathcal{X}, F_{\mathrm{N}})$ is unique (thanks to Lemma 1), and so is the Nash equilibrium. It is immediate to verify that the unique Nash equilibrium is given by $(x^1, x^2) = (0, 0)$.

Let us now consider the same game defined in (4.7) and introduce the additional coupling constraint

$$\mathcal{C} = \{(x^1, x^2) \in \mathbb{R}^2 \,|\, x^1 + x^2 \geq 1\}.$$

Assumption 1 is still satisfied so that any solution of the variational inequality $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{N}})$ is a Nash equilibrium, but the reverse *does not hold* in this case, due to the presence of $\mathcal{C}$. As a matter of fact, the solution of $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{N}})$ (i.e., the variational equilibrium) is unique thanks to the strong monotonicity of F_{N}. On the contrary, it can be verified that any point in the set $\{x \in \mathbb{R}^2 \,|\, x^1 + x^2 = 1,\ x^1 \geq 1/2\}$ is a Nash equilibrium as no player can improve by means of unilateral deviations.

4.3 Sufficient conditions for monotonicity

In this section we derive sufficient conditions that guarantee the monotonicity or strong monotonicity of the operators F_{N}, F_{W} associated with the Nash and Wardrop equilibrium problems. The importance in assessing whether these operators posses any monotonicity property stems from the following three observations.

i) Uniqueness of the variational equilibrium is guaranteed by the strong monotonicity of the corresponding operator, as already discussed in Lemma 1.

ii) Strong monotonicity is crucial to control the behaviour of the variational equilibria and their corresponding efficiency in large populations regimes (Chapter 5).

iii) Monotonicity of F_{N}, F_{W} allows to compute the corresponding equilibria using tractable algorithms and to bound their distance (Chapter 6).

To verify whether F_{N}, F_{W} are monotone or strongly monotone one can exploit the following equivalent characterizations.

Lemma 2. [FP07, Prop. 2.3.2] *A continuously differentiable operator $F : \mathcal{K} \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is strongly monotone with monotonicity constant α (resp. monotone) if and only if $\nabla_x F(x) \succeq \alpha I$ (resp. $\nabla_x F(x) \succeq 0$) for all $x \in \mathcal{K}$. Moreover, if $\mathcal{K}$ is compact, there exists $\alpha > 0$ such that $\nabla_x F(x) \succeq \alpha I$ for all $x \in \mathcal{K}$ if and only if $\nabla_x F(x) \succ 0$ for all $x \in \mathcal{K}$.*

In the following we specialize this result to the case when the cost functions (4.2) reduce to

$$J^i(x^i, \sigma(x)) := v^i(x^i) + p(\sigma(x))^\top x^i. \tag{4.8}$$

The cost functions in (4.8) can describe, for example, applications where x^i denotes the usage level of a certain commodity, whose negative utility is modeled by $v^i : \mathcal{X}^i \to \mathbb{R}$ and whose per-unit cost $p : \frac{1}{M}\sum_{i=1}^M \mathcal{X}^i \to \mathbb{R}^n$ depends on the average usage level of the entire population. Cost functions of the form (4.8) are widely used in the applications, see [Che+14; MCH13]. We refer to p in the following as to the *price function*. The operators in (4.6) become

$$F_{\mathrm{W}}(x) = [\nabla_{x^i} v^i(x^i)]_{i=1}^M + [p(\sigma(x))]_{i=1}^M, \tag{4.9a}$$
$$F_{\mathrm{N}}(x) = F_{\mathrm{W}}(x) + \tfrac{1}{M}[\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M. \tag{4.9b}$$

Lemma 3 (Sufficient conditions for strong monotonicity of (4.9))**.**

(a) Suppose that for each agent $i \in \{1, \dots, M\}$ the function v^i in (4.8) is convex and that p is monotone; then F_{W} is monotone. Under the further assumption that p is affine and strongly monotone, F_{N} is strongly monotone.

(b) Suppose that for each agent $i \in \{1, \dots, M\}$ the function v^i in (4.8) is strongly convex and that p is monotone. Then F_{W} is strongly monotone. □

4.3.1 Linear price function

In the following we refine the sufficient conditions of Lemma 3 to the important class of aggregative games with cost functions of the following form

$$J^i(x^i, \sigma(x)) := \frac{1}{2}(x^i)^\top Q x^i + (C\sigma(x) + c^i)^\top x^i\,, \tag{4.10}$$

where $Q \in \mathbb{R}^{n\times n}$ is symmetric, $C \in \mathbb{R}^{n\times n}$ (not necessarily symmetric), $c^i \in \mathbb{R}^n$. We observe that (4.10) is a special case of (4.8), obtained setting $v^i(x^i) = (x^i)^\top Q x^i + (c^i)^\top x^i$ and $p(\sigma(x)) = C\sigma(x)$. We refer to this case as to the case of *linear price function*. The cost functions in (4.10) have been used for example in [HCM07; Gra+16; BP13]. Since the operators $F_{\mathrm{N}}, F_{\mathrm{W}}$ defined in (4.6) are obtained by differentiating quadratic functions, their expression is affine, and given by

$$F_{\mathrm{W}}(x) = \left(I_M \otimes Q + \tfrac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right) x + c, \tag{4.11a}$$

$$F_{\mathrm{N}}(x) = F_{\mathrm{W}}(x) + \tfrac{1}{M}(I_M \otimes C^\top)x, \tag{4.11b}$$

where $c = [c^1; \dots; c^M]$. The following lemma exploits the structure in (4.11) to derive sufficient conditions for strong monotonicity of F_{W}, F_{N}.

Lemma 4 (Sufficient conditions for strong monotonicity of (4.11))**.**

(a) If $Q \succ 0$, $C \succeq 0$ then F_{W} in (4.11a) is strongly monotone.

(b) If $Q \succ 0$, $Q - C^\top Q^{-1} C \succ 0$ then F_{W} in (4.11a) is strongly monotone.

(c) If $Q \succ 0$, $C \succeq 0$ or if $Q \succeq 0$, $C \succ 0$ then F_{N} in (4.11b) is strongly monotone.

4.3.2 Diagonal price function

In the following we consider the case when the price function $p(\sigma)$ has diagonal structure, i.e., the t-th component of p depends only on the corresponding component of the average. Formally, we assume that $p(\sigma(x))$ can be decomposed as $p(\sigma(x)) = [p_t(\sigma_t(x_t))]_{t=1}^n$, with $p_t : \mathbb{R} \to \mathbb{R}$ for all t, $\sigma_t(x_t) = \frac{1}{N}\sum_{i=1}^M x_t^i$ and $x_t := [x_t^1, \dots, x_t^n]$. This corresponds to cost functions of the following form

$$J^i(x^i, \sigma(x)) = v^i(x^i) + \sum_{t=1}^n p_t(\sigma_t(x_t)) x_t^i \,. \tag{4.12}$$

Cost functions as in (4.12) are typically used in the literature to describe congestion costs of road traffic networks ([War52; CSS04] and Section 7.2) or the charging of electric vehicles ([MCH13; Gra+16] and Section 7.1). We refer to this case as to the case of *diagonal price function*. A sufficient condition ensuring the monotonicity or strong monotonicity of F_{W} can be obtained directly exploiting the structure of (4.12) and the result in Lemma 2. The situation is more complicated when we turn our attention to F_{N} due to the presence of the additional term $[\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M$ in (4.9). The following lemma provides a sufficient condition.

Lemma 5. *Let $\mathcal{X}$ be closed and convex. Assume that $v^i(x^i)$ in (4.12) is convex for each agent $i \in [M]$ and that p_t is continuously differentiable and strictly increasing for all*

$t \in [n]$. *Further, suppose that* $\mathcal{X}^i \subseteq [0, x^0]^n$ *for each* $i \in [M]$. *If*

$$\min_{\substack{t\in\{1,\dots,n\}\\ z\in[0,x^0]}} \left(p_t'(z) - \frac{\tilde{x}^0 p_t''(z)}{8} \right) > 0, \tag{4.13}$$

then the operator F_{N} *is strongly monotone.*

We note that the positivity requirement on the agent strategies is satisfied in many applications such as those studied in Chapter 7. Nevertheless, the previous lemma can be extended adjusting the condition (4.13) to the case where $\mathcal{X}^i \subseteq [-x^0, x^0]^n$, see [Gen18].

An immediate consequence of the previous lemma is that, when p_t is continuously differentiable, strictly increasing and *concave* for all t, the operator F_{N} is strongly monotone. It is worth noting that [YSM11] considers a similar setup to what studied in this section. In [YSM11, Lem. 3] the authors exploit the structure in (4.12) and give conditions for $\nabla_x F_{\mathrm{N}}(x)$ to be a P-matrix, which in turn guarantees uniqueness of the Nash equilibrium in the absence of coupling constraints. This is, to the best of our knowledge, the only work providing sufficient conditions for equilibrium uniqueness and convergence of the algorithms. It is interesting to note that uniqueness in [YSM11] holds assuming $p_t' > 0, p_t'' > 0$, whereas our result holds if the opposite condition is satisfied, namely if $p_t' > 0, p_t'' < 0$.

4.4 Appendix

4.4.1 Proofs of the results presented in Sections 4.2 and 4.3

Proof of Proposition 6

Proof.

(a) The proof of the first statement can be also found in [FFP07, Thm. 2.1].
By definition $\bar{x}_{\mathrm{N}}$ is a solution of $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{N}})$, that is

$$F_{\mathrm{N}}(\bar{x}_{\mathrm{N}})^{\top}(x - \bar{x}_{\mathrm{N}}) \geq 0, \qquad \forall x \in \mathcal{Q}\,. \tag{4.14}$$

In the following we fix the strategies of all the players but i to $x^{-i} = \bar{x}_{\mathrm{N}}^{-i}$, so that all the summands in (4.14) vanish, except for the i-th term

$$\nabla_{x^i} J^i(\bar{x}_{\mathrm{N}}^i, \sigma(\bar{x}_{\mathrm{N}}))^{\top}(x^i - \bar{x}_{\mathrm{N}}^i) \geq 0, \qquad \forall x^i \in \mathcal{Q}(\bar{x}_{\mathrm{N}}^{-i})\,.$$

Consider the function $x^i \mapsto J^i(x^i, \frac{1}{M}x^i + \frac{1}{M}\sum_{j\neq i}\bar{x}_{\mathrm{N}}^j)$ and observe that $J^i : \mathcal{Q}(\bar{x}_{\mathrm{N}}^{-i}) \to \mathbb{R}$ is convex by assumption. Since $\mathcal{Q}(\bar{x}_{\mathrm{N}}^{-i})$ is also convex by assumption, it follows from [BT89, Prop. 3.1] that $\bar{x}_{\mathrm{N}}^i$ must be a minimizer of $J^i : \mathcal{Q}(\bar{x}_{\mathrm{N}}^{-i}) \to \mathbb{R}$, i.e., that

$$J^i(\bar{x}_{\mathrm{N}}^i, \bar{x}_{\mathrm{N}}^{-i}) \leq J^i\left(x^i, \frac{1}{M}x^i + \frac{1}{M}\sum_{j\neq i}\bar{x}_{\mathrm{N}}^j\right), \qquad \forall x^i \in \mathcal{Q}(\bar{x}_{\mathrm{N}}^{-i})\,.$$

Since this holds for all $i \in \{1, \ldots, M\}$ and since $\bar{x}_{\mathrm{N}} \in \mathcal{Q}$ by definition of variational inequality, it follows that $\bar{x}_{\mathrm{N}}$ is a Nash equilibrium of $\mathcal{G}$.

(b) We rewrite the operator $F_{\mathrm{W}}(x)$ as $\tilde{F}_{\mathrm{W}}(x, \sigma(x))$, where $\tilde{F}_{\mathrm{W}}(x, z) \coloneqq [\nabla_{x^i} J^i(x^i, z)]_{i=1}^M$. Fix $\bar{z} = \sigma(\bar{x}_{\mathrm{W}})$. By definition, if $\bar{x}_{\mathrm{W}}$ solves $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{W}})$ then $F_{\mathrm{W}}(\bar{x}_{\mathrm{W}})^{\top}(x - \bar{x}_{\mathrm{W}}) \geq 0$ for all $x \in \mathcal{Q}$, i.e.,

$$\tilde{F}_{\mathrm{W}}(\bar{x}_{\mathrm{W}}, \bar{z})^{\top}(x - \bar{x}_{\mathrm{W}}) \geq 0, \qquad \forall x \in \mathcal{Q}. \tag{4.15}$$

Consider $i \in \{1, \ldots, M\}$, set $x^{-i} = \bar{x}_{\mathrm{W}}^{-i}$ in (4.15) and consider an arbitrary $x^i \in \mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i})$; then all the summands in (4.15) vanish except the i^{th} one and (4.15) reads

$$\nabla_{x^i} J^i(\bar{x}_{\mathrm{W}}^i, \bar{z})^{\top}(x^i - \bar{x}_{\mathrm{W}}^i) \geq 0, \qquad \forall x^i \in \mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i}). \tag{4.16}$$

Consider the convex function $J^i(\cdot, \bar{z}) : \mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i}) \to \mathbb{R}$. Since $\mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i})$ is a convex set, by (4.16) and [BT89, Prop. 3.1] we have that $\bar{x}_{\mathrm{W}}^i \in \arg\min_{x^i \in \mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i})} J^i(x^i, \bar{z})$. Substituting $\bar{z} = \sigma(\bar{x}_{\mathrm{W}})$, one has $J^i(\bar{x}_{\mathrm{W}}^i, \sigma(\bar{x}_{\mathrm{W}})) \leq J^i(x^i, \sigma(\bar{x}_{\mathrm{W}}))$ for all $x^i \in \mathcal{Q}^i(\bar{x}_{\mathrm{W}}^{-i})$. Since this holds for all $i \in \{1, \ldots, M\}$ and since $\bar{x}_{\mathrm{W}} \in \mathcal{Q}$, it follows that $\bar{x}_{\mathrm{W}}$ is a Wardrop equilibrium of $\mathcal{G}$.

□

Proof of Lemma 3

Proof.

(a) Let us first show that F_{W} is monotone. Since v^i is convex, then $\nabla_{x^i} v^i(x^i)$ is monotone in x^i by [Scu+12, Sec. 4.2.2]. Hence $[\nabla_{x^i} v^i(x^i)]_{i=1}^M$ is monotone. Moreover, for any x_1, x_2

$$\begin{aligned} &([p(\sigma(x_1))]_{i=1}^M - [p(\sigma(x_2))]_{i=1}^M)^\top (x_1 - x_2) \\ &= M(p(\sigma(x_1)) - p(\sigma(x_2)))^\top (\sigma(x_1) - \sigma(x_2)) \geq 0, \end{aligned} \tag{4.17}$$

where the last inequality follows from the fact that p is monotone. By (4.9a) and the fact that the sum of two monotone operators is monotone, one can conclude that F_{W} is monotone.

To show that F_{N} is strongly monotone, we write the affine expression of p as $p(x) = Cx + c$, where there exists $\alpha > 0$ such that $C \succ \alpha I_n$ by Lemma 2. Then the term $\frac{1}{M}[\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M$ in (4.9b) equals $\frac{1}{M}(I_M \otimes C^\top)x$. Since $\nabla_x(\frac{1}{M}(I_M \otimes C^\top)x) \succ \frac{\alpha}{M} I_{Mn}$, then $\frac{1}{M}[\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M$ is strongly monotone by Lemma 2. Having already shown that F_{W} is monotone, the proof is concluded upon noting that the sum of a monotone operator and a strongly monotone operator is strongly monotone.

(b) Strong convexity of v^i is equivalent to strong monotonicity of $\nabla_{x^i} v^i(x^i)$ in x^i [Scu+12, Sec. 4.2.2]. Then $[\nabla_{x^i} v^i(x^i)]_{i=1}^M$ is strongly monotone. Monotonicity of $[p(\sigma(x))]_{i=1}^M$ in (4.9a) can be shown as in (4.17).

□

Proof of Lemma 4

Proof.

(a) By Lemma 2, strong monotonicity of F_{W} in (4.11a) is equivalent to $\nabla_x F_{\mathrm{W}}(x) = \left(I_M \otimes Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right)^\top \succ 0$, which is independent from x. If $Q \succ 0$ and $C \succeq 0$, it holds $\left(I_M \otimes Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right)^\top \succ 0$, proving the statement.

(b) Since Q is symmetric, $Q \succ 0$, and $Q - C^\top Q^{-1} C \succ 0$, by Schur's Complement we have

$$\begin{bmatrix} Q & C^\top \\ C & Q \end{bmatrix} \succ 0\,.$$

It follows that

$$\begin{bmatrix} x \\ x \end{bmatrix}^\top \begin{bmatrix} Q & C^\top \\ C & Q \end{bmatrix} \begin{bmatrix} x \\ x \end{bmatrix} = x^\top \left(2Q + C + C^\top\right) x \geq 0\,,$$

from which it must be $Q + C \succ 0$. We conclude the proof by showing that $Q \succ 0$ symmetric and $Q + C \succ 0$ imply $\nabla_x F_{\mathrm{W}}(x) = \left(I_M \otimes Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right)^\top \succ 0$. Recall that $\nabla_x F_{\mathrm{W}}(x) \succ 0$ is equivalent to showing positive definiteness of

$$I_M \otimes 2Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes (C + C^\top). \tag{4.18}$$

To prove the latter inequality, let us consider λ_j an eigenvalue of the matrix appearing in (4.18) with corresponding eigenvector $v_j \neq \mathbb{0}_{Mn}$. It must be

$$\left(I_M \otimes 2Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes (C + C^\top)\right) v_j = \lambda_j v_j \iff 2Qv_j^i + \frac{C + C^\top}{M}\sum_{i=1}^M v_j^i = \lambda_j v_j^i, \tag{4.19}$$

for all $i \in [M]$. Summing the previous expressions over i gives

$$\left(2Q + C + C^\top\right)\sum_{i=1}^M v_j^i = \lambda_j \sum_{i=1}^M v_j^i\,.$$

Thus, if $\sum_{i=1}^M v_j^i \neq \mathbb{0}_n$, λ_j is also an eigenvalue of $2Q + C + C^\top$ and it must be $\lambda_j > 0$ since $Q + C \succ 0$. If, on the contrary, $\sum_{i=1}^M v_j^i = \mathbb{0}_n$, it follows from (4.19) that $2Qv_j = \lambda_j v_j$, i.e., λ_j is also an eigenvalue of $2Q$ and it must be $\lambda_j > 0$ since $Q \succ 0$ and symmetric. We conclude, as required, that the matrix appearing in (4.18) is positive definite, since all its eigenvalues are strictly positive.

(c) Similarly to the first point, strong monotonicity of F_{N} in (4.11b) is equivalent by Lemma 2 to $\left(I_M \otimes Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right)^\top + \frac{1}{M}(I_M \otimes C^\top)^\top \succ 0$. If $Q \succ 0$ and $C \succeq 0$ or if $Q \succeq 0$ and $C \succ 0$, it follows that $\left(I_M \otimes Q + \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^\top \otimes C\right)^\top + \frac{1}{M}(I_M \otimes C^\top)^\top \succ 0$, completing the proof.

□

Proof of Lemma 5

Proof. First, observe that the operator $p : \mathbb{R}^n \to \mathbb{R}^n$ is monotone. Indeed, since p_t is strictly increasing it holds for all y, z that

$$(p(y) - p(z))^\top (y - z) = \sum_{t=1}^n (p_t(y_t) - p_t(z_t))(y_t - z_t) > 0\,.$$

Thanks to Lemma 3, we conclude that F_{W} is also monotone. According to (4.9b), to show strong monotonicity of F_{N} it is sufficient to show that the term $[\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M$ is strongly monotone for all $x \in \mathcal{X}$. The latter is equivalent to proving $\nabla_x [\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M \succ 0$ for all $x \in \mathcal{X}$ by Lemma 2. We have

$$\nabla_x [\nabla_z p(z)_{|z=\sigma(x)} x^i]_{i=1}^M = I_M \otimes \nabla_z p(z)_{|z=\sigma(x)} + \frac{1}{M}\mathbb{1}_M \otimes \left([\mathrm{diag}\{p_t''(\sigma_t)x_t^i\}_{t=1}^n]_{i=1}^M\right)^\top, \tag{4.20}$$

where $\mathrm{diag}\{p_t''(\sigma_t)x_t^i\}_{t=1}^n$ is the diagonal matrix whose entry in position (t,t) is $p_t''(\sigma_t)x_t^i$. The permutation matrix $P = [[\mathbf{e}_{t+(i-1)n}^\top]_{i=1}^M]_{t=1}^n$ ($\mathbf{e}_i$ denotes the i-th vector of the canonical basis) permutes (4.20) into block-diagonal form

$$P\nabla_x[\nabla_z p(z)_{|z=\sigma(x)}x^i]_{i=1}^M P^\top = \tag{4.21}$$

$$\begin{bmatrix} p_1'(\sigma_1)I_M & & \\ & \ddots & \\ & & p_n'(\sigma_n)I_M \end{bmatrix} + \frac{1}{M}\begin{bmatrix} p_1''(\sigma_1)x_1\mathbb{1}_M^\top & & \\ & \ddots & \\ & & p_n''(\sigma_n)x_n\mathbb{1}_M^\top \end{bmatrix}$$

where $x_t = [x_t^i]_{i=1}^M$. To conclude, it suffices to show that $p_t'(\sigma_t)I_M + \frac{1}{M}p_t''(\sigma_t)x_t\mathbb{1}_M^\top \succ 0$ for all t. Lemma 6 (reported at the end of this proof) guarantees that $\lambda_{\min}\left(x_t\mathbb{1}_M^\top + \mathbb{1}_M x_t^\top\right)/2 \geq -\frac{\tilde{x}^0 M}{8}$, which terminates the proof. □

Lemma 6. *For all $M \in \mathbb{N}$, it holds*

$$\min_{y\in[0,1]^M} \lambda_{\min}\left(y\mathbb{1}_M^\top + \mathbb{1}_M y^\top\right) \geq -\frac{M}{4}. \tag{4.22}$$

Proof. The statement is trivially true for $M = 1$. For $M > 1$, the left hand side of (4.22) is equivalent to

$$\min_{\substack{y\in[0,1]^M \\ \|v\|=1}} v^\top\left(y\mathbb{1}_M^\top + \mathbb{1}_M y^\top\right)v = \min_{\substack{y\in[0,1]^M \\ \|v\|=1}} 2\left(v^\top y\right)\left(\mathbb{1}_M^\top v\right). \tag{4.23}$$

Let us consider a pair $y^\star, v^\star$ minimizing (4.23). If $\mathbb{1}_M^\top v^\star = 0$, the bound (4.23) is trivially satisfied. We are left with two cases, $\mathbb{1}_M^\top v^\star > 0$ and $\mathbb{1}_M^\top v^\star < 0$. Let us start from the case of $\mathbb{1}_M^\top v^\star > 0$. To minimize $2\left(v^\top y\right)\left(\mathbb{1}_M^\top v\right)$, it must be

$$y_i^\star = \begin{cases} 0 & \text{if } v_i^\star > 0 \\ 1 & \text{if } v_i^\star < 0, \end{cases} \quad \text{for all } i \in \{1,\dots,M\}. \tag{4.24}$$

Without loss of generality, we can assume $y_i^\star \in \{0,1\}$ if $v_i^\star = 0$. Hence we conclude that $y^\star \in \{0,1\}^M$ and (4.22) reduces to

$$\min_{p\in\{0,\dots,M\}} \lambda_{\min}\left[\begin{array}{c|c} 2(\mathbb{1}_p\mathbb{1}_p^\top) & \mathbb{1}_p\mathbb{1}_{(M-p)}^\top \\ \hline \mathbb{1}_{(M-p)}\mathbb{1}_p^\top & \mathbb{0}_{(M-p)}\mathbb{0}_{(M-p)}^\top \end{array}\right], \tag{4.25}$$

where without loss of generality we assumed the first p components of $y^\star$ to be 1 and the remaining to be 0. Note that the matrix in (4.25) features p identical rows followed by $M-p$ other identical rows. Hence any of its eigenvectors must have p identical components followed by $M-p$ other identical components. With this observation and the definition of eigenvalue, algebraic calculations show that the matrix in (4.25) has only

two distinct eigenvalues, the minimum of the two being $p-\sqrt{Mp}$. The function $p-\sqrt{Mp}$ is minimized over the reals for $p = M/4$ with corresponding minimum $\lambda_{\min} = -M/4$, as it can be seen by using the change of variables $p = q^2$ and minimizing the quadratic function $q^2 - \sqrt{M}q$. Since $p \in \{0, \dots, M\}$ in (4.25), the value $-M/4$ is a lower bound for the minimum eigenvalue, and it is attained only if M is a multiple of 4. We conclude by noting that the derivation for the case $\mathbb{1}_M^\top v^\star < 0$ is identical to the derivation for the case $\mathbb{1}_M^\top v^\star > 0$ just shown, upon switching 0 and 1 in (4.24). □

CHAPTER 5

Equilibria and efficiency in large populations

Many real world applications where agents behave strategically feature the interaction of a large population of individuals. As an example, consider that of drivers moving on the road network of a city, with the objective of reaching their destination as swiftly as possible. As an alternative example consider that of traders in a stock market. Motivated by this observation, the current chapter is dedicated to the study of aggregative games with a large number of players. The chapter is divided in two parts. In Section 5.1 we provide bounds on the distance between Wardrop and Nash equilibria, while in Section 5.2 we study the efficiency of these equilibria, i.e., we study how much selfish behaviour degrades the performance of a centrally controlled system. All the proofs are reported in the Appendix (Section 5.3). The results presented in this chapter have been published in [Pac+18; PPL18].

Specifically, we consider a sequence of games $(\mathcal{G}_M)_{M=1}^{\infty}$. For fixed M, the game $\mathcal{G}_M$ is played among M agents and is defined as in (4.3) with an arbitrary coupling constraint $\mathcal{C}$, arbitrary costs $\{J^i(x^i, \sigma(x))\}_{i=1}^M$ and arbitrary local constraints $\{\mathcal{X}^i\}_{i=1}^M$. For the sake of readability, we avoid the explicit dependence on M in denoting these quantities and in denoting x_{N}, x_{W}, F_{N}, F_{W}.

5.1 Distance between Nash and Wardrop equilibria

As we have learnt from the previous chapter, the monotonicity properties of the operators F_{N} and F_{W} might not coincide. For example, F_{N} might be strongly monotone for a given game, while for the same game F_{W} might not be. Unfortunately, if the operator associated to a variational inequality is not monotone, determining the corresponding solution is in general an intractable problem. Motivated by this shortcoming, in this section we provide bounds on the distance between $\bar{x}_{\mathrm{N}}$ and $\bar{x}_{\mathrm{W}}$, so that, should one of the two equilibria be difficult to compute (e.g., due to the lack of monotonicity), we might be able to compute the other and still be able to learn something about the former.

Assumption 2. *There exists a convex, compact set $\mathcal{X}^0 \subset \mathbb{R}^n$ such that $\cup_{i=1}^M \mathcal{X}^i \subseteq \mathcal{X}^0$ for each $\mathcal{G}_M$ in the sequence $(\mathcal{G}_M)_{M=1}^{\infty}$. Let $R := \max_{y \in \mathcal{X}^0}\{\|y\|\}$. For each M and*

$i \in \{1, \dots, M\}$, the function $J^i(z_1, z_2)$ is Lipschitz with respect to z_2 in $\mathcal{X}^0$ with Lipschitz constant L_2 independent from M, i and $z_1 \in \mathcal{X}^i$.

We note that Assumption 2 implies that $\sigma(x) \in \mathcal{X}^0$ for any M and any $x \in \mathcal{X}^1 \times \cdots \times \mathcal{X}^M$. Furthermore, if the cost function (4.2) takes the specific form (4.8), then p being Lipschitz in $\mathcal{X}^0$ with constant L_p implies $J^i(z_1, z_2)$ being Lipschitz with respect to z_2 in $\mathcal{X}^0$ with constant $L_2 = RL_p$, as by the Cauchy-Schwartz inequality

$$\begin{aligned} &\|J^i(z_1, z_2) - J^i(z_1, z_2')\| = \|(p(z_2) - p(z_2'))^\top z_1\| \\ &\leq \|p(z_2) - p(z_2')\|\|z_1\| \leq RL_p\|z_2 - z_2'\|. \end{aligned} \tag{5.1}$$

The next proposition shows that every Wardrop equilibrium is an ε-Nash equilibrium, with ε vanishing as M grows.

Proposition 7. *Let the sequence of games $(\mathcal{G}_M)_{M=1}^\infty$ satisfy Assumption 2. For each $\mathcal{G}_M$, every Wardrop equilibrium is an ε-Nash equilibrium, with $\varepsilon = \frac{2RL_2}{M}$.*

Proposition 7 is a strong result as it guarantees that for relatively large M a Wardrop equilibrium is *almost* stable in the sense of the Nash equilibrium definition. In particular, at any given Wardrop equilibrium no player can improve upon its cost by more than an additive factor ε, considering the strategies of the others fixed. Unfortunately, Proposition 7 provides no information on the distance between the set of strategies constituting a Nash and a Wardrop equilibrium. This question is addressed in the following theorem.

Theorem 1. *Let the sequence of games $(\mathcal{G}_M)_{M=1}^\infty$ satisfy Assumption 2, and each $\mathcal{G}_M$ satisfy Assumption 1. Then:*

(a) If the operator F_{N} relative to $\mathcal{G}_M$ is strongly monotone on $\mathcal{Q}$ with monotonicity constant $\alpha_{_M} > 0$, then there exists a unique variational Nash equilibrium $\bar{x}_{\mathrm{N}}$ of $\mathcal{G}_M$. Moreover, for any variational Wardrop equilibrium $\bar{x}_{\mathrm{W}}$

$$\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \leq \frac{L_2}{\alpha_{_M}\sqrt{M}}. \tag{5.2}$$

As a consequence, if $\alpha_{_M}\sqrt{M} \to \infty$ as $M \to \infty$, then $\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \to 0$ as $M \to \infty$.

(b) If the operator F_{W} relative to $\mathcal{G}_M$ is strongly monotone on $\mathcal{Q}$ with monotonicity constant $\alpha_{_M} > 0$, then there exists a unique variational Wardrop equilibrium $\bar{x}_{\mathrm{W}}$ of $\mathcal{G}_M$. Moreover, for any variational Nash equilibrium $\bar{x}_{\mathrm{N}}$

$$\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \leq \frac{L_2}{\alpha_{_M}\sqrt{M}}. \tag{5.3}$$

As a consequence, if $\alpha_{_M}\sqrt{M} \to \infty$ as $M \to \infty$, then $\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \to 0$ as $M \to \infty$.

(c) *If in each game $\mathcal{G}_M$ the cost function $J^i(x^i, \sigma(x))$ takes the form (4.8), with $v^i = 0$ and p being strongly monotone on $\mathcal{X}^0$ with monotonicity constant α, then there exists a unique $\bar{\sigma}$ such that $\sigma(\bar{x}_{\mathrm{W}}) = \bar{\sigma}$ for any variational Wardrop equilibrium $\bar{x}_{\mathrm{W}}$ of $\mathcal{G}_M$. Moreover, for any variational Nash equilibrium $\bar{x}_{\mathrm{N}}$ of $\mathcal{G}_M$ and for any variational Wardrop equilibrium $\bar{x}_{\mathrm{W}}$ of $\mathcal{G}_M$*

$$\|\sigma(\bar{x}_{\mathrm{N}}) - \sigma(\bar{x}_{\mathrm{W}})\| \leq \sqrt{\frac{2RL_2}{\alpha M}}. \tag{5.4}$$

Hence, $\|\sigma(\bar{x}_{\mathrm{N}}) - \sigma(\bar{x}_{\mathrm{W}})\| \to 0$ as $M \to \infty$.[1]

We point out that (5.2) and (5.3) can be used to derive a bound on the average strategies similar to (5.4).

Related Works

Proposition 7 ensures that, under minimal assumptions, any Wardrop equilibrium is an ε-Nash equilibrium. Such result follows directly from the aggregative structure of the game, and from the Lipschitz continuity of the cost functions. A similar idea is used to prove analogous results in various previous contributions. For example, the case of potential games is investigated in [AW04; Alt+06], routing games are considered in [Alt+11], flow control and routing in communication networks are discussed in [ABS02], while a similar argument is used in [Gra+16] for the case of average aggregative games with no coupling constraints. Proposition 7 is a direct extension of these works to generic aggregative games with coupling constraints.

Theorem 1 show that it is possible to derive bounds on the Euclidean distance between Nash and Wardrop equilibria at the price of introducing further assumptions. More precisely, strong monotonicity of either the Nash or Wardrop operator ensures that the actual strategies $\bar{x}_{\mathrm{N}}$ and $\bar{x}_{\mathrm{W}}$ converge to each other as M grows large. A weaker requirement, i.e., the strong monotonicity of p ensures instead convergence in the aggregate. To the best of our knowledge, the only result bounding the Euclidean distance between the two equilibria is obtained in [HM85]. Therein a similar bound to (5.4) is derived limitedly to routing/congestion games. However, [HM85] requires the population to increase by means of identical replicas of the agents. We here prove that a similar argument can be used to address the case of generic new agents. In addition, the first two results of Theorem 1 address a more general class of aggregative games (i.e., not necessarily congestion games) by employing a new type of argument, based on a sensitivity analysis result for variational inequalities with perturbed strongly monotone operators [Nag13, Thm. 1.14]. We note that the works [DN87; AW04; Alt+06] guarantee convergence of

[1] If p is Lipschitz with constant L_p, then in (5.4) L_2 can be replaced by RL_p, as by (5.1). This is used in the application in Sections 7.1 and 7.2.

Nash to Wardrop in terms of Euclidean distance, but do not provide a bound on the convergence rate.

Finally, we observe that our results are derived in relation to variational equilibria. Nevertheless, if there is no coupling constraint as in all the above-mentioned works, then any equilibrium is a variational equilibrium. Hence our results subsume the previous.

5.2 Equilibrium efficiency: the price of anarchy

In this section we study the efficiency of Nash and Wardrop equilibria by means of the concept of price of anarchy. The notion of equilibrium efficiency was first formalized in [KP99] and is used to describe the performance degradation incurred when moving from a centralized solution to distributed and strategic decision making. The motivations that lead us to the study of the price of anarchy are essentially two. The first is analytical: given an optimization problem and the corresponding competitive counterpart, we wish to know how inefficient an equilibrium might be. The second stems from the possibility to engineer the behaviour of a large population of strategic thinkers. For example, in the application considered in Section 7.1, the system operator has the freedom to select the price function p. In these cases we wish to understand how to modify the game so as to make it as efficient as possible.

Similarly to previous section, we consider a sequence of games $(\mathcal{G}_M)_{M=1}^{\infty}$, where each game $\mathcal{G}_M$ is defined as in (4.3) with arbitrary constraint sets $\{\mathcal{X}^i\}_{i=1}^M$, and cost functions of the following form

$$J^i(x^i, \sigma(x)) \coloneqq p(\sigma(x) + d)^\top x^i, \qquad d \in \mathbb{R}^n. \tag{5.5}$$

In order to simplify the exposition, throughout this section we consider the case when no coupling constraint is present, i.e., $\mathcal{C} = \mathbb{R}^{Mn}$.[2] We observe that the cost functions in (5.5) have a similar structure to those in (4.8). More precisely, it is possible to reduce (5.5) to (4.8), upon setting $v^i(x^i) = 0$ in the latter equation and introducing an additional player whose constraint set is given by $\{x \in \mathbb{R}^n \mid x = d \cdot M\}$. Since we are interested in the case of large population, we do not purse this approach because the unboundedness of this set (as $M \to \infty$) will complicate the analysis. The costs in (5.5) can be used to describe applications where x^i denotes the usage level of a certain commodity, whose per-unit cost p depends on the average usage level plus some inflexible normalized usage level d [MCH13; Che+14]. As the notion of equilibrium efficiency relates the behaviour of an equilibrium allocation with that of a socially optimal one, we begin with the following definition.

Definition 8 (Social optimizer)**.** *A set of actions $x_{\mathrm{S}} = [x_{\mathrm{S}}^1; \dots; x_{\mathrm{S}}^M] \in \mathbb{R}^{Mn}$ is a social optimizer of $\mathcal{G}_M$ if $x_{\mathrm{S}} \in \mathcal{X}$ and it minimizes the cost $J_{\mathrm{S}}(\sigma(x)) \coloneqq p(\sigma(x)+d)^\top(\sigma(x)+d)$.*

[2] Most of the results hold with minor adaptations in the presence of coupling constraints too.

Note that the cost J_S is the sum of all the players costs, divided by M, and the additional term $p(\sigma(x) + d)^\top d$. The reason why the latter term is included is that we want to compute the total cost of buying the commodity for both the flexible ($\sigma(x)$) and inflexible (d) users. This cost was first introduced in [MCH13] and successively used in [GGL15; DMP17; DAS17]. For a given a game $\mathcal{G}_M$, we quantify the efficiency of equilibrium allocations using the notion of price of anarchy [KP99]

$$\mathrm{PoA}_M := \frac{\max_{x_N \in \mathrm{NE}_M} J_S(\sigma(x_N))}{J_S(\sigma(x_S))},$$

where $\mathrm{NE}_M \subseteq \mathcal{X}$ is the set of Nash equilibria of $\mathcal{G}_M$ and x_S is a social optimizer of $\mathcal{G}_M$. The price of anarchy captures the ratio between the cost at the worst Nash equilibrium and the optimal cost; by definition $\mathrm{PoA}_M \geq 1$. In the following we study the behavior of PoA_M, for three different classes of admissible price functions p.

5.2.1 Linear price function

Throughout this subsection we consider cost functions of the form (5.5), where the price functions p is linear as detailed in Assumption 8. Linear price functions have been used in [GGL15; DMP17] to model, e.g., the competitive charging of electric vehicles.

Assumption 3. *The cost functions are as in (5.5), where the price function p takes the form $p(z + d) = C(z + d)$, with $C = C^\top \in \mathbb{R}^{n \times n}$, $C \succ 0$.*

Under Assumption 8, let L_s, L_p be the Lipschitz constants of J_S, p, and α the monotonicity constant of p.[3] The following theorem shows that, under minimal assumptions, any Wardrop equilibrium is also socially optimum irrespective of the population size M. This is no longer the case for Nash equilibria, which nevertheless recover this property when the population size grows.

Theorem 2 (PoA_M bound and convergence to 1)**.** *Let Assumption 8 hold.*

(a) Let each of the constraint set $\{\mathcal{X}^i\}_{i=1}^M$ be closed, convex, non empty. Then, for any fixed game $\mathcal{G}_M$ in the sequence $(\mathcal{G}_M)_{M=1}^\infty$, every Wardrop equilibrium x_W is a social optimizer, i.e., $J_S(\sigma(x_W)) \leq J_S(\sigma(x))$, $\forall x \in \mathcal{X}$.

(b) Assume, in addition, that there exists a convex, compact set $\mathcal{X}^0 \subset \mathbb{R}^n$ such that $\cup_{i=1}^M \mathcal{X}^i \subseteq \mathcal{X}^0$ for each $\mathcal{G}_M$ in $(\mathcal{G}_M)_{M=1}^\infty$. Define the constant $c = RL_S\sqrt{2L_p\alpha^{-1}}$, where $R = \max_{y \in \mathcal{X}^0}\{\|y\|\}$. Then,

$$J_S(\sigma(x_S)) \leq J_S(\sigma(x_N)) \leq J_S(\sigma(x_S)) + c/\sqrt{M}, \tag{5.6}$$

[3] The function $p(z + d) = C(z + d)$ is strongly monotone since $C \succ 0$ with monotonicity constant given by the smallest eigenvalue of C.

for any fixed game $\mathcal{G}_M$ in the sequence. Thus, if there exists $\hat{J} \geq 0$ s.t. $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) > \hat{J}$ for every game in the sequence $(\mathcal{G}_M)_{M=1}^{\infty}$, one has

$$1 \leq \mathrm{PoA}_M \leq 1 + c/\big(\hat{J}\sqrt{M}\big) \quad \textit{and} \quad \lim_{M\to\infty} \mathrm{PoA}_M = 1\,.$$

Remark 2. *The previous theorem extends the results of [MCH13; GGL15; DMP17; DAS17] simultaneously allowing for arbitrary convex constraints, finite populations, and non diagonal price function. Note that the condition $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) > \hat{J} \geq 0$ is merely technical and required to properly define PoA_M. This condition is trivially satisfied in the most of the applications considered, see, e.g., Section 7.1. Even if the latter condition does not hold, the cost at any Nash equilibrium converges to the minimum cost as $M \to \infty$, see (5.6).*

5.2.2 Diagonal price function

In the following we study the efficiency of Nash and Wardrop equilibria when the cost functions take the form (5.5) and the price function $p(z+d)$ has diagonal structure, i.e., the t-th component of p depends only on the corresponding component of the average. We distinguish two cases depending on wether p_t has the same structure for different values of t, or not. Towards this goal, we first introduce two useful assumptions.

Assumption 4. *For $i \in \{1,\dots,M\}$, the constraint set $\mathcal{X}^i$ is closed, convex, non empty. For $z \in \frac{1}{M}\sum_{i=1}^{M}\mathcal{X}^i$, the function $z \mapsto p(z+d)$ is continuously differentiable and strongly monotone while $z \mapsto p(z+d)^\top(z+d)$ is strongly convex. Let L_{S}, L_p be the Lipschitz constant of J_{S}, p, and α be the monotonicity constant of p.*

Assumption 5. *There exists a convex, compact set $\mathcal{X}_0 \subset \mathbb{R}^n$ s.t. $\cup_{i=1}^{M}\mathcal{X}^i \subseteq \mathcal{X}_0$ for each game $\mathcal{G}_M$ in $(\mathcal{G}_M)_{M=1}^{\infty}$. Moreover, $J^i(x^i,\sigma(x))$ is convex in $x^i \in \mathcal{X}^i$ for all fixed $x^{-i} \in \mathcal{X}^{-i}$, for all $i \in \{1,\dots,M\}$. We let $R = \max_{y\in\mathcal{X}_0}||y||$.*

Homogeneous price function

In this section we consider $p(z+d)$ to be a nonlinear function, and assume its t-th component to depend only on the t-th component $z_t + d_t$, for all $t \in \{1,\dots,n\}$. Additionally, we assume that the functions p_t have the same structure for all the values of t. This describes, for example, electricity markets where the unit cost of electricity at every instant of time is captured by a time invariant function depending on the total consumption at that same instant.

Assumption 6. *The price function p takes the form*

$$p(z+d) = \big[f(z_1+d_1),\dots,f(z_n+d_n)\big]^\top,$$

with $f(y) : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$. Further $\mathcal{X}^i \subseteq \mathbb{R}^n_{\geq 0}$ and $d \in \mathbb{R}^n_{>0}$.

If $f(y)$ is not linear, a simple check shows that, in general, $\nabla_{x^j}(\nabla_{x^i} J^i(x^i, \sigma(x))) \neq \nabla_{x^i}(\nabla_{x^j} J^j(x^j, \sigma(x)))$ when $i \neq j$. Consequently, the game is not potential, [FP07, Theorem 1.3.1]. Hence methods to bound the PoA based on the existence of an underlying potential function [GGL15; DMP17], can not be used here. The following theorem provides a necessary and sufficient condition on the structure of f that ensures the efficiency of the resulting equilibria.

Theorem 3 (PoA_M convergence and counterexample)**.** *Suppose that Assumptions 4, 5 and 6 hold. Further assume that $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) > \hat{J}$ for some $\hat{J} \geq 0$, for every game in $(\mathcal{G}_M)_{M=1}^{\infty}$.*

(a) If $f(y) = \alpha y^k$ with $\alpha > 0$ and $k > 0$, it holds

$$1 \leq \mathrm{PoA}_M \leq 1 + c/\left(\hat{J}\sqrt{M}\right) \quad \text{and} \quad \lim_{M\to\infty} \mathrm{PoA}_M = 1\,,$$

with $c = RL_{\mathrm{S}}\sqrt{2L_p\alpha^{-1}}$ constant.

(b) For $n \geq 2$, if $f(y)$ satisfies the assumptions, but does not take the form αy^k for some $\alpha > 0$ and $k > 0$, it is possible to construct a sequence of games $(\mathcal{G}_M)_{M=1}^{\infty}$ for which $\lim_{M\to\infty} \mathrm{PoA}_M > 1$.

The counterexample relative to the second claim is constructed using $\mathcal{X}^i = \bar{\mathcal{X}}$. In other words our impossibility result holds also for the case of homogeneous populations. This is not in contrast with the result in [MCH13] or [DAS17], because therein the sets $\bar{\mathcal{X}}$ were assumed to be simplexes with upper bounds constraints. Here we claim that there exists a convex set $\bar{\mathcal{X}}$ (not a simplex with upper bounds) such that PoA_M does not converge to 1.

Remark 3. *The previous theorem is of fundamental importance in applications where the system operator has the possibility to freely set the price function. In these cases, Theorem 3 suggests the use of monomial price functions to guarantee the highest achievable efficiency (all Nash equilibria become social optimizers for large M). If different price functions are chosen, it is always possible to construct a problem instance such that the worst Nash equilibrium is* not *a social optimizer.*

Heterogeneous price function

In the previous subsection we showed that if the price function is not a monomial, then PoA_M may not converge to one. In this section we derive upper bounds for PoA_M when the price function belongs to a general class of functions, as formalized next.

Assumption 7. *The price function p takes the form*

$$p(z + d) = \left[l_1(z_1 + d_1), \ldots, l_n(z_n + d_n)\right]^{\top},$$

where $l_t(y) : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $l_t \in \mathcal{L}$ *for all* t *and* $\mathcal{L}$ *is a given set of continuous and nondecreasing price functions. Further let* $\mathcal{X}^i \subseteq \mathbb{R}^n_{\geq 0}$ *be non empty, closed and convex.*

Note that Assumption 7 is *less restrictive* than Assumption 6 as we let the price l_t depend on the time instant t. The key idea in this case is to show that standard results derived for Wardrop equilibria in routing games [Rou03], [CSS04] can be applied to the setup studied here. The resulting bounds on PoA_M can then be derived using the convergence result in Theorem 1. Formally, given a game $\mathcal{G}_M$ with cost functions as in (5.5), we consider an equivalent nonatomic routing game over a parallel network with a number of links equal to n, the dimension of decision variables x^i. To present our next result, we first introduce the quantity

$$\beta(\mathcal{L}) := \sup_{l \in \mathcal{L}} \sup_{v \geq 0} \left(\frac{1}{vl(v)} \max_{w \geq 0} [(l(v) - l(w))w] \right).$$

defined in [CSS04, Eq 3.8]. Therein, the authors show that $\beta(\mathcal{L}) \leq 1$ and $[1 - \beta(\mathcal{L})]^{-1} = \alpha(\mathcal{L})$. The quantity $\alpha(\mathcal{L})$ describes, essentially, the worst-case price of anarchy over all possible cost functions in the set $\mathcal{L}$, as detailed in the following theorem. The key is to show that the games considered here are $(1, \beta(\mathcal{L}))$-smooth, as defined in [Rou09, Def. 1.1].

Theorem 4 (PoA_M for heterogeneous price function)**.**

(a) Suppose that Assumption 7 holds. Then for any fixed game $\mathcal{G}_M$ *and any Wardrop equilibrium* x_W *it holds*

$$J_S(\sigma(x_W)) \leq J_S(\sigma(x_S))\alpha(\mathcal{L}) \tag{5.7}$$

(b) Further suppose Assumptions 4 and 5 hold, and there exists $\hat{J} \geq 0$ *s.t.* $J_S(\sigma(x_S)) > \hat{J}$ *for every game in* $(\mathcal{G}_M)_{M=1}^{\infty}$*. Then, for any game* $\mathcal{G}_M$ *in the sequence*

$$J_S(\sigma(x_S)) \leq J_S(\sigma(x_N)) \leq J_S(\sigma(x_S))\alpha(\mathcal{L}) + c/\sqrt{M},$$

and $1 \leq \mathrm{PoA}_M \leq \alpha(\mathcal{L}) + c/\left(\hat{J}\sqrt{M}\right)$*, thus implying* $\lim_{M\to\infty} \mathrm{PoA}_M \leq \alpha(\mathcal{L})$*, with* $c = RL_s\sqrt{2L_p\alpha^{-1}}$*.*

Remark 4. *In [Rou03, Table 1],* $\alpha(\mathcal{L})$ *is computed for classes of functions such as affine, quadratic, polynomials. If* $\mathcal{L}$ *contains constant functions, then* (5.7) *is tight (see [Rou03] and the application discussed in Section 7.1). This is not a contradiction of Theorems 2 and 3 because therein either constant functions are not allowed or the price function* p_t *is assumed to be independent of* t*. Theorems 2 and 3 can be seen as refinements of Theorem 4 and guarantee that* $\lim_{M\to\infty} \mathrm{PoA}_M = 1$ *by restricting the admissible class of price functions.*

5.3 Appendix

5.3.1 Proofs of the results presented in Section 5.1

Proof of Proposition 7

Proof. Consider any Wardrop equilibrium x_{W} of $\mathcal{G}_M$ (not necessarily a variational one). By Definition 7, $x_{\mathrm{W}} \in \mathcal{Q}$ and for each agent i

$$J^i(x^i_{\mathrm{W}}, \sigma(x_{\mathrm{W}})) \leq J^i(x^i, \sigma(x_{\mathrm{W}})), \qquad \forall x^i \in \mathcal{Q}^i(x^{-i}_{\mathrm{W}}).$$

It follows that for each agent i and for all $x^i \in \mathcal{Q}^i(x^{-i}_{\mathrm{W}})$

$$\begin{aligned}
&J^i(x^i_{\mathrm{W}}, \sigma(x_{\mathrm{W}})) - J^i\left(x^i, \frac{1}{M}\left(x^i + \sum_{j\neq i} x^j_{\mathrm{W}}\right)\right) \\
&= \underbrace{J^i(x^i_{\mathrm{W}}, \sigma(x_{\mathrm{W}})) - J^i(x^i, \sigma(x_{\mathrm{W}}))}_{\leq 0} + J^i(x^i, \sigma(x_{\mathrm{W}})) - J^i\left(x^i, \frac{1}{M}\left(x^i + \sum_{j\neq i} x^j_{\mathrm{W}}\right)\right) \\
&\leq L_2 \left\| \sigma(x_{\mathrm{W}}) - \left(\frac{1}{M}\left(x^i + \sum_{j\neq i} x^j_{\mathrm{W}}\right)\right)\right\| = \frac{L_2}{M}\left\| \left(x^i_{\mathrm{W}} + \sum_{j\neq i} x^j_{\mathrm{W}}\right) - \left(x^i + \sum_{j\neq i} x^j_{\mathrm{W}}\right)\right\| \\
&= \frac{L_2}{M}\|x^i_{\mathrm{W}} - x^i\| \leq \frac{2RL_2}{M}.
\end{aligned}$$

Hence x_{W} is an ε-Nash equilibrium of $\mathcal{G}_M$ with $\varepsilon = \frac{2RL_2}{M}$. □

Proof of Theorem 1

Proof.

(a) We first bound the distance between the operators F_{N} and F_{W} in terms of M. By (4.6) it holds

$$\begin{aligned}
\|F_{\mathrm{N}}(x) - F_{\mathrm{W}}(x)\|^2 &= \left\|[\nabla_{x^i} J^i(x^i, \sigma(x))]_{i=1}^M - [\nabla_{x^i} J^i(x^i, z)_{|z=\sigma(x)}]_{i=1}^M\right\|^2 \\
&= \sum_{i=1}^M \left\| \frac{1}{M}\nabla_z J^i(x^i, z)_{|z=\sigma(x)}\right\|^2 \leq \frac{1}{M^2}\sum_{i=1}^M L_2^2 = \frac{L_2^2}{M},
\end{aligned}$$

where the inequality follows from the fact that $J^i(z_1, z_2)$ is Lipschitz in z_2 on $\mathcal{X}^0$ with constant L_2 by Assumption 2 and hence the term $\|\nabla_z J^i(x^i, z)_{|z=\sigma(x)}\|$ is bounded by L_2 by definition of derivative. Taking the square root, it follows that

$$\|F_{\mathrm{N}}(x) - F_{\mathrm{W}}(x)\| \leq \frac{L_2}{\sqrt{M}}. \tag{5.8}$$

for all $x \in \mathcal{X}^0$. We exploit (5.8) to bound the distance between Nash and Wardrop strategies. Since F_{N} is strongly monotone on $\mathcal{Q}$ by assumption, $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{N}})$ has a unique solution $\bar{x}_{\mathrm{N}}$ by Lemma 1. Moreover, the distance between the solutions of two variational inequalities differing in the operator used can be bounded using [Nag13]. Formally, for all solutions $\bar{x}_{\mathrm{W}}$ of $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{W}})$ [Nag13, Thm. 1.14] shows that

$$\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \leq \frac{1}{\alpha_{_M}} \|F_{\mathrm{N}}(\bar{x}_{\mathrm{W}}) - F_{\mathrm{W}}(\bar{x}_{\mathrm{W}})\|.$$

Combining this with equation (5.8) yields the result.

(b) As in the above, with Nash in place of Wardrop and vice versa.

(c) Any solution $\bar{x}_{\mathrm{W}}$ to the $\mathrm{VI}(Q, F_{\mathrm{W}})$ satisfies

$$\begin{aligned} &F_{\mathrm{W}}(\bar{x}_{\mathrm{W}})^\top (x - \bar{x}_{\mathrm{W}}) \geq 0, \ \forall x \in Q \Leftrightarrow \\ &\sum_{i=1}^{M} p(\sigma(\bar{x}_{\mathrm{W}}))^\top (x^i - \bar{x}^i_{\mathrm{W}}) \geq 0, \ \forall x \in Q \Leftrightarrow \\ &p(\sigma(\bar{x}_{\mathrm{W}}))^\top (\sigma(x) - \sigma(\bar{x}_{\mathrm{W}})) \geq 0, \ \forall x \in Q. \end{aligned} \tag{5.9}$$

Any solution $\bar{x}_{\mathrm{N}}$ to the $\mathrm{VI}(Q, F_{\mathrm{N}})$ satisfies

$$\begin{aligned} &F_{\mathrm{N}}(\bar{x}_{\mathrm{N}})^\top (x - \bar{x}_{\mathrm{N}}) \geq 0, \ \forall x \in Q \Leftrightarrow \\ &p(\sigma(\bar{x}_{\mathrm{N}}))^\top (\sigma(x) - \sigma(\bar{x}_{\mathrm{N}})) + \frac{1}{M^2} \sum_{i=1}^{M} (\nabla_z p(z)_{|z=\sigma(\bar{x}_{\mathrm{N}})} \bar{x}^i_{\mathrm{N}})^\top (x^i - \bar{x}^i_{\mathrm{N}}) \geq 0, \ \forall x \in Q. \end{aligned} \tag{5.10}$$

Exploiting the strong monotonicity of p on $\mathcal{X}^0$, one has

$$\begin{aligned} \alpha \|\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})\|^2 &\leq (p(\sigma(\bar{x}_{\mathrm{W}})) - p(\sigma(\bar{x}_{\mathrm{N}})))^\top (\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})) \\ &= p(\sigma(\bar{x}_{\mathrm{W}}))^\top (\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})) - p(\sigma(\bar{x}_{\mathrm{N}}))^\top (\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})) \\ &\underset{\text{by (5.9)}}{\leq} -p(\sigma(\bar{x}_{\mathrm{N}}))^\top (\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})) \\ &\underset{\text{by (5.10)}}{\leq} \frac{1}{M^2} \sum_{i=1}^{M} (\bar{x}^i_{\mathrm{N}})^\top (\nabla_z p(z)_{|z=\sigma(\bar{x}_{\mathrm{N}})})^\top (\bar{x}^i_{\mathrm{W}} - \bar{x}^i_{\mathrm{N}}) \\ &= \frac{1}{M^2} \sum_{i=1}^{M} (\bar{x}^i_{\mathrm{N}})^\top (\nabla_z J^i(\bar{x}^i_{\mathrm{W}}, z)_{|z=\sigma(\bar{x}_{\mathrm{N}})} - \nabla_z J^i(\bar{x}^i_{\mathrm{N}}, z)_{|z=\sigma(\bar{x}_{\mathrm{N}})}) \\ &\leq \frac{1}{M^2} \sum_{i=1}^{M} \|\bar{x}^i_{\mathrm{N}}\| (\|\nabla_z J^i(\bar{x}^i_{\mathrm{W}}, z)_{|z=\sigma(\bar{x}_{\mathrm{N}})}\| + \|\nabla_z J^i(\bar{x}^i_{\mathrm{N}}, z)_{|z=\sigma(\bar{x}_{\mathrm{N}})}\|) \\ &\leq \frac{2L_2}{M^2} \sum_{i=1}^{M} \|\bar{x}^i_{\mathrm{N}}\| \leq \frac{2L_2}{M^2} \sum_{i=1}^{M} R \leq \frac{1}{M} 2RL_2, \end{aligned}$$

where we have used the Chauchy-Schwartz inequality, the triangular inequality, and the Lipschitzianity of $J^i(z_1, z_2)$ in addition to (5.9) and (5.10). We conclude that $\|\sigma(\bar{x}_{\mathrm{W}}) - \sigma(\bar{x}_{\mathrm{N}})\| \leq \sqrt{\frac{2RL_2}{\alpha M}}$.

□

5.3.2 Proofs of the results presented in Section 5.2

Before proving any of the claims in Section 5.2, we provide a lemma that will be useful in the forthcoming analysis. Throughout the following proofs, we denote with $\Sigma := \frac{1}{M}\sum_{i=1}^{M} \mathcal{X}^i$.

Lemma 7 (Equivalent characterizations of x_{W}, x_{S})**.** *Let the cost functions be given as in (5.5), and each of the constraint set $\{\mathcal{X}^i\}_{i=1}^M$ be closed, convex, non empty. Additionally, assume that the function $z \mapsto p(z+d)$ is continuously differentiable and strongly monotone while $z \mapsto p(z+d)^\top (z+d)$ is strongly convex, for all $z \in \Sigma$. The following holds.*

(a) Given x_{W} a Wardrop equilibrium, its average $\sigma(x_{\mathrm{W}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$, with $F_{\mathrm{W}} : \mathbb{R}^n \to \mathbb{R}^n$, $F_{\mathrm{W}}(z) := p(z+d)$. The $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$ admits a unique solution σ_{W}. Let us define $\mathcal{X}_{\mathrm{W}} := \{x \in \mathcal{X} \text{ s.t. } \frac{1}{M}\sum_{j=1}^{M} x^j = \sigma_{\mathrm{W}}\}$. Then any vector of strategies $x_{\mathrm{W}} \in \mathcal{X}_{\mathrm{W}}$ is a Wardrop equilibrium.

(b) Given x_{S} a social optimizer, its average $\sigma(x_{\mathrm{S}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$, with $F_{\mathrm{S}} : \mathbb{R}^n \to \mathbb{R}^n$, $F_{\mathrm{S}}(z) := p(z+d) + [\nabla_z p(z+d)](z+d)$. The $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$ admits a unique solution σ_{S}. Define $\mathcal{X}_{\mathrm{S}} := \{x \in \mathcal{X} \text{ s.t. } \frac{1}{M}\sum_{j=1}^{M} x^j = \sigma_{\mathrm{S}}\}$. Then any vector of strategies $x_{\mathrm{S}} \in \mathcal{X}_{\mathrm{S}}$ is a social optimizer.

Proof.

(a) The sets $\mathcal{X}^i$ are convex and closed by assumption; further, for fixed $z \in \Sigma$, the functions $J^i(x^i, z)$ are linear and thus convex in $x^i \in \mathcal{X}^i$ for all $i \in \{1, \dots, M\}$. It follows by Proposition 6 that a Wardrop equilibrium x_{W} satisfies[4]

$$[\mathbb{1}_M \otimes p(\sigma(x_{\mathrm{W}}) + d)]^\top (x - x_{\mathrm{W}}) \geq 0, \quad \forall x \in \mathcal{X}. \tag{5.11}$$

Rearranging and dividing by M we get $p(\sigma(x_{\mathrm{W}})+d)^\top (\frac{1}{M}\sum_{j=1}^{M} x^i - \frac{1}{M}\sum_{j=1}^{M} x^i_{\mathrm{W}}) \geq 0$, for all $x \in \mathcal{X}$, or equivalently $p(\sigma(x_{\mathrm{W}}) + d)^\top (z - \sigma(x_{\mathrm{W}})) \geq 0, \ \forall z \in \Sigma$, that is, $\sigma(x_{\mathrm{W}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$.

By assumption $F_{\mathrm{W}}(z) = p(z+d)$ is strongly monotone and Σ is closed, convex (since the sets $\mathcal{X}^i$ are closed, convex), hence by [FP07, Thm. 2.3.3] $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$ has

[4] Proposition 6 goes in both directions here as there is no coupling constraint, see the discussion in Section 4.2.

a unique solution σ_{W}. By definition of variational inequality, for any $z \in \Sigma$ it holds $p(\sigma_{\mathrm{W}} + d)^\top (z - \sigma_{\mathrm{W}}) \geq 0$. By definition of $x_{\mathrm{W}} \in \mathcal{X}_{\mathrm{W}}$, we have $\sigma(x_{\mathrm{W}}) = \sigma_{\mathrm{W}}$. It follows that $p(\sigma(x_{\mathrm{W}}) + d)^\top (z - \sigma(x_{\mathrm{W}})) \geq 0$ for any $z \in \Sigma$. By definition of Σ, we conclude that (5.11) holds for all $x \in \mathcal{X}$. By Proposition 6, we conclude that x_{W} is a Wardrop equilibrium.

(b) By assumption the set $\mathcal{X}$ is convex and closed and $J_{\mathrm{S}}(\sigma(x))$ is convex. Hence, any social optimizer x_{S} satisfies the first order condition in Proposition 3

$$\nabla_x [p(\sigma(x) + d)(\sigma(x) + d)]^\top_{|x=x_{\mathrm{S}}} (x - x_{\mathrm{S}}) \geq 0 \quad \forall x \in \mathcal{X}\,. \tag{5.12}$$

Note that $M \nabla_{x^i} (p(\sigma(x) + d)^\top (\sigma(x) + d)) = p(\sigma(x_{\mathrm{S}}) + d) + [\nabla_z p(\sigma(x_{\mathrm{S}}) + d)](\sigma(x_{\mathrm{S}}) + d)$ for all $i \in \{1, \ldots, M\}$. Consequently, (5.12) is equivalent to $[p(\sigma(x_{\mathrm{S}}) + d) + \nabla_z p(\sigma(x_{\mathrm{S}}) + d)(\sigma(x_{\mathrm{S}}) + d)]^\top (\sigma(x) - \sigma(x_{\mathrm{S}})) \geq 0$. Thus $\sigma(x_{\mathrm{S}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$. The remaining claims are shown similarly to those for x_{W}.

□

Proof of Theorem 2

Proof.

(a) Note that Assumption 8 implies strong monotonicity of $z \mapsto p(z + d)$, and strong convexity of $z \mapsto p(z + d)^\top (z + d)$. Thus the assumptions of Lemma 7 are satisfied. Let x_{W} be a Wardrop equilibrium. By Lemma 7 part 1, $\sigma(x_{\mathrm{W}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$. Thanks to Assumption 8, $F_{\mathrm{S}}(z) = C(z + d) + C^\top (z + d) = 2C(z + d) = 2F_{\mathrm{W}}(z)$. Since the two operators $F_{\mathrm{W}}(z)$ and $F_{\mathrm{S}}(z)$ are parallel for each $z \in \Sigma$, it follows from the definition of variational inequality that $\sigma(x_{\mathrm{W}})$ must solve $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$ too. Using Lemma 7 part 2 we conclude that x_{W} must be a social optimizer.

(b) By definition $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) \leq J_{\mathrm{S}}(\sigma(x_{\mathrm{N}}))$ and so $1 \leq \mathrm{PoA}_M$. Observe that the assumptions on the sets $\{\mathcal{X}^i\}_{i=1}^M$ together with Assumption 8 imply Assumption 1 and ensures that $J^i(z_1, z_2)$ is Lipschitz with respect to z_2 in $\mathcal{X}^0$. Thus, the assumptions of Theorem 1 part 3 are satisfied. It follows that for any Nash equilibrium x_{N} and Wardrop equilibrium x_{W} of the game $\mathcal{G}_M$, it holds $||\sigma(x_{\mathrm{W}}) - \sigma(x_{\mathrm{N}})|| \leq \sqrt{2R^2 L_p \alpha^{-1} M^{-1}}$. Thus, using the Lipschitz property of J_{S} one has that $|J_{\mathrm{S}}(\sigma(x_{\mathrm{N}})) - J_{\mathrm{S}}(\sigma(x_{\mathrm{W}}))| \leq L_{\mathrm{S}} R \sqrt{2L_p \alpha^{-1} M^{-1}} = c\sqrt{M^{-1}}$. Since every Wardrop equilibrium is socially optimum (previous point of this proof), one has $|J_{\mathrm{S}}(\sigma(x_{\mathrm{N}})) - J_{\mathrm{S}}(\sigma(x_{\mathrm{S}}))| \leq c\sqrt{M^{-1}}$ and thus $J_{\mathrm{S}}(\sigma(x_{\mathrm{N}})) \leq J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) + c\sqrt{M^{-1}}$. The final result regarding the price of anarchy follows from the latter inequality upon dividing both sides by $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) > \hat{J} \geq 0$.

□

Proof of Theorem 3

Proof.

(a) We first show that any Wardrop equilibrium is a social optimizer. To do so, observe that the function $f(y) = \alpha y^k$ satisfies all the assumptions required by Lemma 7 (see Lemma 9 in the Appendix). Let x_{W} be a Wardrop equilibrium of $\mathcal{G}_M$. By Lemma 7, $\sigma(x_{\mathrm{W}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$. Thanks to Assumption 6 and the choice of $f(y)$,

$$F_{\mathrm{S}}(z) = (k+1)[\alpha(z_1 + d_1)^k, \ldots, \alpha(z_n + d_n)^k]^\top = (k+1)F_{\mathrm{W}}(z)\,.$$

Hence $\sigma(x_{\mathrm{W}})$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$ too. Using Lemma 7 we conclude that x_{W} must be a social optimizer. The proof is now identical to the proof of the second part of Theorem 2.

(b) If $f(y)$ does not take the form αy^k for some $\alpha > 0$ and $k > 0$, by Lemma 8 there exists a point $\bar{z} \in \mathbb{R}^n_{>0}$ for which $F_{\mathrm{W}}(\bar{z})$ and $F_{\mathrm{S}}(\bar{z})$ are not aligned, i.e., for which $F_{\mathrm{S}}(\bar{z}) \neq hF_{\mathrm{W}}(\bar{z})$ for all $h \in \mathbb{R}$. We intend to construct a sequence of games $\mathcal{G}_M$ so that for every $\mathcal{G}_M$ in the sequence the unique average at the Wardrop equilibrium is exactly $\bar{z}$, that is $\bar{z}$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$, but $\bar{z}$ does not solve $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$. This fact indeed proves, by Lemma 7, that for any game $\mathcal{G}_M$ the Wardrop equilibria of $\mathcal{G}_M$ are not social minimizers. By Theorem 1, $\sigma(x_{\mathrm{N}}) \to \sigma(x_{\mathrm{W}})$ as $M \to \infty$. Thus, PoA cannot converge to 1.

In the following we construct a sequence of games with the above mentioned properties. To this end let us define $\mathcal{X}^i := \bar{\mathcal{X}} \subseteq \mathbb{R}^n$, so that $\Sigma = \bar{\mathcal{X}}$ with $\bar{\mathcal{X}} := \{\bar{z} + \alpha v_1 + \beta v_2 \;\; \alpha, \beta \in [0\ 1]\} \cap \mathbb{R}^n_{\geq 0}$, where $v_1 := \bar{F}_{\mathrm{W}}$, $v_2 := (\bar{F}_{\mathrm{W}}^\top \bar{F}_{\mathrm{S}})\bar{F}_{\mathrm{W}} - (\bar{F}_{\mathrm{W}}^\top \bar{F}_{\mathrm{W}})\bar{F}_{\mathrm{S}}$ and $\bar{F}_{\mathrm{W}} := F_{\mathrm{W}}(\bar{z})$, $\bar{F}_{\mathrm{S}} := F_{\mathrm{S}}(\bar{z})$; see Figure 5.1. The intuition is that $-v_2$ is the component of $\bar{F}_S$ that lives in the same plane as $\bar{F}_S$ and $\bar{F}_W$ and is orthogonal to $\bar{F}_W$, so that $\bar{F}_W^\top v_2 = 0$. Observe that $\Sigma = \bar{\mathcal{X}}$ is the intersection of a bounded and convex set with the positive orthant and thus satisfies Assumptions 1, 5 and 6. It is easy to verify that $\bar{z} \in \bar{\mathcal{X}}$ and that $F_{\mathrm{W}}(\bar{z})^\top(z - \bar{z}) = \alpha ||F_{\mathrm{W}}(\bar{z})||^2 \geq 0$ for all $z \in \Sigma = \bar{\mathcal{X}}$, so that $\bar{z}$ solves $\mathrm{VI}(\Sigma, F_{\mathrm{W}})$. Let us pick $\hat{z} = \bar{z} + \beta v_2$. Note that since $\bar{z} > 0$, for β small enough $\hat{z}$ belongs to $\mathbb{R}^n_{>0}$ as well and thus to $\bar{\mathcal{X}}$. Then $F_{\mathrm{S}}(\bar{z})^\top(\hat{z} - \bar{z}) = \beta(\bar{F}_{\mathrm{S}}^\top \bar{F}_{\mathrm{W}})^2 - \beta ||\bar{F}_{\mathrm{S}}||^2 ||\bar{F}_{\mathrm{W}}||^2 < 0$. The inequality is strict because $\bar{F}_{\mathrm{W}}$, $\bar{F}_{\mathrm{S}}$ are neither parallel nor zero (Lemma 8). Thus, $\bar{z}$ does not solve $\mathrm{VI}(\Sigma, F_{\mathrm{S}})$.

□

Lemma 8. *For $n \geq 2$, if $f(y)$ satisfies Assumptions 1, 5 and 6, but does not take the form αy^k for some $\alpha > 0$ and $k > 0$, then there exists $\bar{z} \in \mathbb{R}^n_{>0}$ such that $F_{\mathrm{S}}(\bar{z}) \neq hF_{\mathrm{W}}(\bar{z})$, $\forall h \in \mathbb{R}$. Moreover, $F_{\mathrm{S}}(\bar{z}) \neq 0$, $F_{\mathrm{W}}(\bar{z}) \neq 0$.*

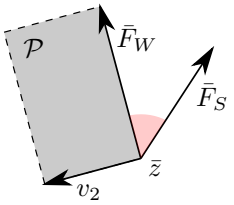

Figure 5.1: Construction of the set $\bar{\mathcal{X}}$.

Proof. Let us consider the first statement. By contradiction, assume there exists $\beta(z) : \mathbb{R}^n_{>0} \to \mathbb{R}$ such that $F_{\mathrm{S}}(z) = \beta(z) F_{\mathrm{W}}(z)$ for all $z \in \mathbb{R}^n_{>0}$. This implies

$$f'(z_t + d_t)(z_t + d_t) = (\beta(z_1, \dots, z_n) - 1) f(z_t + d_t)\,, \tag{5.13}$$

for all $t \in \{1, \dots, n\}$ and for all $z \in \mathbb{R}^n_{>0}$, $d \in \mathbb{R}^n_{>0}$. By Assumption 6, $f(z_t + d_t) > 0$. Hence one can divide (5.13) for $f(z_t + d_t)$ without loss of generality, and conclude that $\beta(z_1, \dots, z_n) = \beta_1(z_1) = \cdots = \beta_n(z_n)$ with $\beta_i : \mathbb{R} \to \mathbb{R}$ for all $z \in \mathbb{R}^n_{>0}$. For $n \geq 2$ the last condition implies $\beta(z_1, \dots, z_n) = b$ constant. Equation (5.13) reads as $f'(y)y = (b-1)f(y) \quad \forall y > 0$, whose continuously differentiable solutions are all and only $f(y) = ay^{b-1}$. Note that if $a \leq 0$ or $b \leq 1$, Assumption 1 is not satisfied, while if $a > 0$ and $b > 1$ we contradicted the assumption that $f(y)$ did not take the form αy^k for some $\alpha > 0$ and $k > 0$. Setting $h = 0$ in the previous claim gives $F_{\mathrm{S}}(\bar{z}) \neq 0$. Since $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, one has $F_{\mathrm{W}}(\bar{z}) := [f(\bar{z}_t + d_t)]_{t=1}^n \neq 0$. □

Lemma 9. *Suppose that the price function p is as in Assumption 6 with $f(y) = \alpha y^k$, $\alpha > 0, k > 0$. Then p satisfies Assumptions 4 and 5.*

Proof. Note that $\nabla_z p(z+d)$ is a diagonal matrix with entry $f'(z_t + d_t)$ in position (t,t). Since $f'(y) = \alpha k y^{k-1} > 0$ for all $y > 0$ and since $z_t + d_t$ is positive by assumption for all t, we get that $p(z+d)$ is continuously differentiable and that $\nabla_z p(z+d) \succ 0$, i.e., that $z \mapsto p(z+d)$ is strongly monotone. Similarly, one can show that the Hessian of $p(z+d)^\top (z+d)$ and the Hessian of $J^i(x^i, \sigma(x))$ with respect to x^i are positive definite. Thus, $z \mapsto p(z+d)^\top (z+d)$ and $x^i \mapsto J^i(x^i, \sigma(x))$ are strongly convex. □

Proof of Theorem 4

Proof. We prove only the first claim as the second can be shown as in Theorem 2. To do so, we define $C^{\sigma_1}(\sigma_2) := p(\sigma_1 + d)^\top (\sigma_2 + d)$ so that $J_S(\sigma) = C^\sigma(\sigma)$. Let x_W be any Wardrop equilibrium. Then, the average $\bar{\sigma} := \sigma_W$ solves $\mathrm{VI}(\Sigma, F_W)$, i.e., $F_W(\bar{\sigma})^\top (\sigma - \bar{\sigma}) \geq 0, \ \forall \sigma \in \Sigma$. This can be seen following the proof of Lemma 7, and observing that only convexity and closedness of $\mathcal{X}^i$ are required. Equivalently, $J_S(\bar{\sigma}) \leq C^{\bar{\sigma}}(\sigma), \ \forall \sigma \in \Sigma$.

However,

$$\begin{aligned} C^{\bar{\sigma}}(\sigma) &= \sum_t l_t(\bar{\sigma}_t + d_t)(\sigma_t + d_t) \\ &= J_S(\sigma) + \sum_t [l_t(\bar{\sigma}_t + d_t) - l_t(\sigma_t + d_t)](\sigma_t + d_t) \\ &= J_S(\sigma) + \sum_t \frac{[l_t(v_t) - l_t(w_t)]w_t}{l_t(v_t)v_t} l_t(v_t)v_t \\ &\leq J_S(\sigma) + \sum_t \beta(\mathcal{L}) l_t(v_t)v_t \\ &= J_S(\sigma) + \beta(\mathcal{L}) J_S(\bar{\sigma}) \end{aligned}$$

where we used $v_t := \bar{\sigma}_t + d_t \geq d_t$, $w_t := \sigma_t + d_t \geq d_t$ and $d_t \geq 0$. The previous relation holds for all $\sigma \in \Sigma$. Selecting $\sigma = \sigma_S$ (the optimum average), we get $J_S(\bar{\sigma}) \leq J_S(\sigma_S) + \beta(\mathcal{L}) J_S(\bar{\sigma})$. Rearranging we obtain (5.7). □

CHAPTER 6

Decentralized algorithms

In this chapter we are interested in the design of algorithms that converge to a Nash or a Wardrop equilibrium of a given game $\mathcal{G}_M$, formally defined in (4.3). All the proofs are reported in the Appendix (Section 6.3). The results presented in this chapter have been published in [Pac+16; Pac+18]. Throughout the following sections we assume that no agent i wishes to disclose information about his cost function J^i or individual constraint set $\mathcal{X}^i$, to other agents, or to a central operator. Thus, we turn our attention to *decentralized algorithms*. The advantage in using such algorithms is not limited to privacy-preserving issues, but decentralized algorithms are generally preferred when dealing with large scale systems for various reasons, including that of computational tractability. For a comprehensive list of advantages and shortcomings in the use of distributed computing, we redirect the reader to the monograph [BT89]. In the following we assume the presence of a central operator able to measure only aggregate quantities, such as the population average $\sigma(x)$, and to broadcast aggregate signals to the agents. Figure 6.1 describes the setup more clearly, in relation to Algorithm 4. Based on this information structure, we focus on the design of decentralized algorithms to obtain a solution of either $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{N}})$ or $\mathrm{VI}(\mathcal{Q}, F_{\mathrm{W}})$. As the techniques are the same for Nash and Wardrop equilibrium, we consider the general problem $\mathrm{VI}(\mathcal{Q}, F)$, where F can be replaced with F_{N} or F_{W}.

Throughout this chapter we assume linearity of the coupling constraints as by Assumption 8, and observe that this property arises in a range of applications, as detailed, e.g., in [FK07, p. 188] and [YP17].

Assumption 8. *The coupling constraint in (4.1) is of the form*

$$x \in \mathcal{C} := \{x \in \mathbb{R}^{Mn} \mid Ax \leq b\} \subset \mathbb{R}^{Mn}, \tag{6.1}$$

with $A := [A_{(:,1)}, \ldots, A_{(:,M)}] \in \mathbb{R}^{m \times Mn}$, $A_{(:,i)} \in \mathbb{R}^{m \times n}$ *for all* $i \in \{1, \ldots, M\}$, $b \in \mathbb{R}^m$. *Moreover, for all* $i \in \{1, \ldots, M\}$, *the set* $\mathcal{X}^i$ *can be expressed as* $\mathcal{X}^i = \{x^i \in \mathbb{R}^n | g^i(x^i) \leq 0\}$, *where* $g^i : \mathbb{R}^n \to \mathbb{R}^{p_i}$ *is continuously differentiable. The set* Q, *which can thus be expressed as* $\mathcal{Q} = \{x \in \mathbb{R}^{Mn} | g^i(x^i) \leq 0, \forall i, Ax \leq b\}$, *satisfies Slater's constraint qualification [BV04, Eq. (5.27)]. Each agent* i *has information on the sub-matrix* $A_{(:,i)}$ *in (6.1), i.e., he is aware of his influence on the coupling constraint.*

If the operator F associated with the variational inequality $\mathrm{VI}(\mathcal{Q}, F)$ is integrable[1] and monotone on $\mathcal{Q}$, that is, if there exists a convex function $E(x) : \mathbb{R}^{Mn} \to \mathbb{R}$ such that $F(x) = \nabla_x E(x)$ for all $x \in \mathcal{Q}$, then $\mathrm{VI}(\mathcal{Q}, F)$ is equivalent to the convex optimization problem [FP07, Sec. 1.3.1]

$$\arg\min_{x \in \mathcal{Q}} E(x).$$

Therefore a solution of $\mathrm{VI}(\mathcal{Q}, F)$ and thus a variational equilibrium can be found by applying any of the decentralized optimization algorithms available in the literature of convex optimization [BT89]; the decentralized structure arises because each agent can evaluate $\nabla_{x^i} E(x)$ by knowing only his strategy x^i and $\sigma(x)$. Since the integrability assumption guarantees that $\mathcal{G}$ is a *potential game* with potential function $E(x)$ [MS96], decentralized convergence tools for potential games such as [DHZ06; MAS09] can also be employed.

In light of this observation, our objective is to determine a solution of $\mathrm{VI}(\mathcal{Q}, F)$ when F is not necessarily integrable, so that the previous methods do not apply. In oder to construct a decentralized scheme, we begin by reformulating $\mathrm{VI}(\mathcal{Q}, F)$ in an extended space $[x; \lambda]$ following the spirit of primal-dual methods used in optimization. The variable λ represents the Lagrange multipliers associated to the coupling constraint $\mathcal{C}$. The following two reformulations will be used to propose two corresponding decentralized algorithms. Formally, for any given $\lambda \in \mathbb{R}^m_{\geq 0}$, we define the λ-dependent game as

$$\mathcal{G}(\lambda) := \begin{cases} \text{agents} & \{1, \dots, M\} \\ \text{cost of agent } i & J^i(x^i, \sigma(x)) + \lambda^\top A(:, i) x^i \\ \text{individual constraint} & \mathcal{X}^i \\ \text{coupling constraint} & \mathbb{R}^{Mn} \end{cases}, \tag{6.2}$$

and introduce the extended $\mathrm{VI}(\mathcal{Y}, T)$, where

$$\mathcal{Y} := \mathcal{X} \times \mathbb{R}^m_{\geq 0}, \quad T(x, \lambda) := \begin{bmatrix} F(x) + A^\top \lambda \\ -(Ax - b) \end{bmatrix}.$$

The following proposition draws a connection between $\mathrm{VI}(\mathcal{Q}, F)$, the game $\mathcal{G}(\lambda)$ and $\mathrm{VI}(\mathcal{Y}, T)$.

Proposition 8. [Scu+12, Sec. 4.3.2] *Let Assumptions 1 and 8 hold. The following statements are equivalent.*

(a) The vector $\bar{x}$ is a solution of VI$(\mathcal{Q}, F)$.

(b) There exists $\bar{\lambda} \in \mathbb{R}^m_{\geq 0}$ such that $\bar{x}$ is a variational equilibrium of $\mathcal{G}(\bar{\lambda})$ and $0 \leq \bar{\lambda} \perp b - A\bar{x} \geq 0$.

[1] A necessary and sufficient condition for the integrability of the operator F is that $\nabla_x F(x) = \nabla_x F(x)^\top$ for all $x \in \mathcal{Q}$ [FP07, Thm. 1.3.1].

(c) There exists $\bar{\lambda} \in \mathbb{R}^m_{\geq 0}$ such that the vector $[\bar{x}; \bar{\lambda}]$ is a solution of VI$(\mathcal{Y}, T)$. □

While the proof is an adaptation of [Scu+12, Sec. 4.3.2], we provide a sketch of it for completeness at the end of this chapter. In the following Sections 6.1 and 6.2 we exploit the equivalence between the statements in Proposition 8 to propose two algorithms that converge to a Wardrop or Nash equilibrium. A numerical comparison of their performance can be found in Chapter 12. We summarize in Table 6.1 the main conditions that guarantee their convergence.

	Nash	Wardrop
Best-response (Algorithm 3)	-	F_{W} strongly monotone and Assumption 9
Gradient-based (Algorithm 4)	F_{N} strongly monotone	F_{W} strongly monotone

Table 6.1: Range of applicability of the presented algorithms, under Assumption 1 and Assumption 2.

6.1 Best-response algorithm for Wardrop equilibrium

Based on the equivalence between the first two statements of Proposition 8, we introduce Algorithm 3. The algorithm features i) an outer loop, in which the central operator updates and broadcasts to the agents the dual variables $\lambda_{(k)}$ based on the current constraint violation, and ii) an inner loop, in which the agents update their strategies to reach a Wardrop equilibrium of the game $\mathcal{G}(\lambda_{(k)})$. Since $\mathcal{G}(\lambda_{(k)})$ is a game *without* coupling constraints, the Wardrop equilibrium can be found, e.g., via the iterative algorithm proposed in [Gra+16, Alg. 1]. In order to ease the forthcoming notation, we define for each agent $i \in \{1, \dots, M\}$ the best-response map to $z \in \frac{1}{M}\sum_{i=1}^{M} \mathcal{X}^i$ and dual variables $\lambda \in \mathbb{R}^m_{\geq 0}$ as

$$x^i_{\mathrm{br}}(z, \lambda) \coloneqq \operatorname*{arg\,min}_{x^i \in \mathcal{X}^i} \; J^i(x^i, z) + \lambda^\top A(:, i) x^i. \tag{6.3}$$

Assumption 9. *For all $i \in \{1, \dots, M\}$ and $\lambda \in \mathbb{R}^m_{\geq 0}$, the mapping $z \mapsto x^i_{\mathrm{br}}(z, \lambda)$ is single valued and Lipschitz with constant L. Moreover, one of the following holds.*

(a) For each $i \in \{1, \dots, M\}$ and $\lambda \in \mathbb{R}^m_{\geq 0}$, the mapping $z \mapsto x^i_{\mathrm{br}}(z, \lambda)$ is non-expansive (see Chapter 3).

(b) For each $i \in \{1, \dots, M\}$ and $\lambda \in \mathbb{R}^m_{\geq 0}$, the mapping $z \mapsto z - x^i_{\mathrm{br}}(z, \lambda)$ is strongly monotone. □

Algorithm 3 (Best-response algorithm for Wardrop equilibrium)

1: **Initialise** $k = 0$, $\tau > 0$, $x_{(0)} \in \mathbb{R}^{nM}$, $\lambda_{(0)} \in \mathbb{R}^m_{\geq 0}$
2: **while** not converged **do**
3: $\quad h = 0$, $\tilde{x}^i_{(0)} = x^i_{(k)}$, $z_{(0)} \in \mathbb{R}^n$.
4: $\quad$ **while** not converged **do**
5: $\quad\quad \tilde{x}^i_{(h+1)} = x^i_{\mathrm{br}}(z_{(h)}, \lambda_{(k)}) \quad \forall i \in \{1, \dots, M\}$
6: $\quad\quad \tilde{\sigma}_{(h+1)} = \frac{1}{M} \sum_{j=1}^{M} \tilde{x}^j_{(h+1)}$
7: $\quad\quad z_{(h+1)} = (1 - \frac{1}{h}) z_{(h)} + \frac{1}{h} \tilde{\sigma}_{(h+1)}$
8: $\quad\quad h \leftarrow h + 1$
9: $\quad$ **end while**
10: $\quad x_{(k+1)} = \tilde{x}_{(h)}$
11: $\quad \lambda_{(k+1)} = \Pi_{\mathbb{R}^m_{\geq 0}} \left(\lambda_{(k)} - \tau(b - A x_{(k+1)}) \right)$
12: $\quad k \leftarrow k + 1$
13: **end while**

Convergence of the inner loop to a Wardrop equilibrium of the game $\mathcal{G}(\lambda_{(k)})$ is guaranteed by Assumption 9 in [Gra+16, Thm. 3 and Cor. 1]. Additionally, [Gra+16] provides sufficient conditions for Assumption 9 to hold, relative to cost functions of the form (4.10). More precisely, it is shown that $Q \succ 0$ and $C = C^\top \succ 0$ or $Q \succ 0$ and $Q - C^\top Q^{-1} C \succ 0$ imply Assumption 9 [Gra+16, Thm.2].

Theorem 5 (Convergence of Algorithm 3)**.** *Suppose that the operator F_{W} in (4.6b) is strongly monotone on $\mathcal{X}$ with constant α, that Assumptions 1, 8 and 9 hold, and that $\mathcal{X}^i$ is bounded for all $i \in \{1, \dots, M\}$. If $\tau < \frac{2\alpha}{\|A\|^2}$, then $x_{(k)}$ in Algorithm 3 converges to a variational Wardrop equilibrium of $\mathcal{G}$.*

Two observations on Theorem 5 follow. First, we note that the convergence result of Theorem 5 holds in the ideal case when, for every fixed $\lambda_{(k)}$, the inner loop converges to the *exact* Wardrop equilibrium. Since this assumption is hardly satisfied due to the finite precision offered by traditional computers, one would like to obtain a guarantee on the convergence of the overall algorithm even if the internal loop provides only an approximate solution. We do not further pursue this direction and instead leave this as a future work. Second, we observe that the convergence speed of Algorithm 3 is, to the best of our knowledge, an open question. Nevertheless, it is possible to characterize the convergence rate in each of the two levels separately. In particular, if in Assumption 9 it holds that $z \mapsto z - x^i_{\mathrm{br}}(z, \lambda)$ is strongly monotone, then it is possible to modify line 7 with $z_{(h+1)} \leftarrow (1 - \frac{1}{\mu}) z_{(h)} + \frac{1}{\mu} \tilde{\sigma}_{(h+1)}$ and guarantee geometric convergence for $\mu \in [0, 1]$ small enough, see [Ber07, Thm. 3.6 (iii)]. The outer loop on the other hand has geometric convergence under the additional assumption that the mapping Φ as defined in the proof of Theorem 5 is not only co-coercive but also strongly monotone.

To the best of our knowledge Algorithm 3 is the first algorithm that guarantees convergence to a Wardrop equilibrium in games with coupling constraints using a best-response algorithm. We note that, for the case of specific costs (4.10), [Gra17] proposes a best-response algorithm that converges to a pair $(\bar{x}, \bar{\lambda})$ such that $\bar{x}$ is a Wardrop equilibrium of the game $\mathcal{G}(\bar{\lambda})$ satisfying the coupling constraint $\mathcal{C}$. However such point is not a Wardrop equilibrium because the complementarity condition $0 \leq \bar{\lambda} \perp b - A\bar{x} \geq 0$ is not guaranteed. A *gradient-step* algorithm based on two nested loops for Nash equilibrium with coupling constraints has been proposed in [Pan+10, Alg. 2] and in [Pav07, Sec. 4].

6.2 Gradient-based algorithm for Nash and Wardrop equilibria

In this section we devise a decentralized algorithm to achieve a Nash or a Wardrop equilibrium using the reformulation of $\mathrm{VI}(\mathcal{Q}, F)$ as a variational inequality in the the extended space $\mathcal{Y}$, see Proposition 8.

Algorithm 4 proceeds as in the following. After an initialization phase, the agents communicate their current decision variables to the central operator, which in turn broadcasts the initial average and dual variable $\sigma_{(0)}$, $\lambda_{(0)}$ to all agents. At every subsequent iteration the agents update their decision variable and communicate their updated strategy to the central operator, which in turn updates the dual variable to $\lambda_{(k+1)}$ and broadcasts $\sigma_{(k+1)}$, $\lambda_{(k+1)}$ to the agents. Figure 6.1 describes the flow of information for Algorithm 4.

Algorithm 4 (Gradient-based algorithm for Nash equilibrium)

1: **Initialise** $k = 0$, $\tau > 0$, $x_{(0)} \in \mathbb{R}^{nM}$, $\lambda_{(0)} \in \mathbb{R}^m_{\geq 0}$
2: **while** not converged **do**
3: $\quad \sigma_{(k)} = \frac{1}{M} \sum_{i=1}^{M} x^i_{(k)}$
4: $\quad x^i_{(k+1)} = \Pi_{\mathcal{X}^i} \left(x^i_{(k)} - \tau \left(\nabla_{x^i} J^i(x^i_{(k)}, \sigma(x_{(k)})) + A^\top_{(:,i)} \lambda_{(k)} \right) \right) \quad \forall i \in \{1, \dots, M\}$
5: $\quad \lambda_{(k+1)} = \Pi_{\mathbb{R}^m_{\geq 0}} \left(\lambda_{(k)} - \tau(b - 2Ax_{(k+1)} + Ax_{(k)}) \right)$alg:eqouter
6: $\quad k \leftarrow k + 1$
7: **end while**

Remark 5. *While Algorithm 4 is presented here for the computation of a Nash equilibrium, the same algorithm can be used to compute a Wardrop equilibrium upon replacing* $\nabla_{x^i} J^i(x^i_{(k)}, \sigma(x_{(k)}))$ *with* $\nabla_{x^i} J^i(x^i_{(k)}, z)_{|z=\sigma(x)}$ *in line 4.*

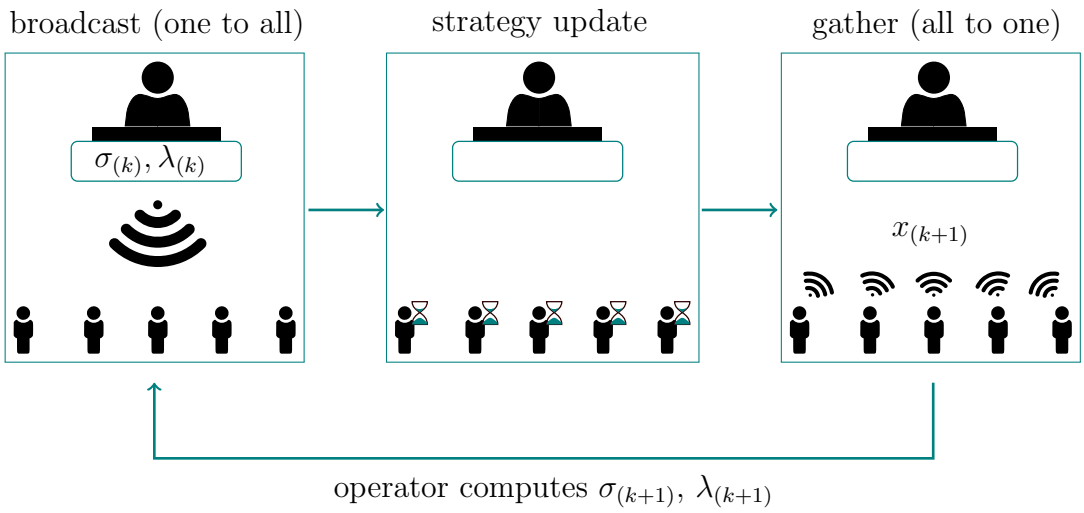

Figure 6.1: Information flow for Algorithm 4

The fundamental ingredient that guarantees the convergence of Algorithm 3 is the strong monotonicity of the operator associated to the corresponding variational inequality, as formalized next.

Theorem 6. *Let Assumption 1 and Assumption 8 hold. Then*

(a) Let F_{N} in (4.6a) be strongly monotone on $\mathcal{X}$ with constant α and Lipschitz on $\mathcal{X}$ with constant L_F. Let $\tau > 0$ s.t.

$$\tau < \frac{-L_F^2+\sqrt{L_F^4+4\alpha^2\|A\|^2}}{2\alpha\|A\|^2} \ . \tag{6.4}$$

Then $x_{(k)}$ in Algorithm 4 converges to a variational Nash equilibrium of $\mathcal{G}$ in (4.3).

(b) Let F_{W} in (4.6b) be strongly monotone and Lipschitz on $\mathcal{X}$ with constants α, L_F, respectively. Let τ satisfies (6.4). Then Algorithm 4 with $\nabla_{x^i} J^i(x^i_{(k)}, z)_{|z=\sigma(x)}$ in place of $\nabla_{x^i} J^i(x^i_{(k)}, \sigma(x_{(k)}))$ in line 4 converges to a variational Wardrop equilibrium.

Remark 6 (Convergence rate)**.** *If the operator F is not only monotone but also affine, and the set $\mathcal{X}$ is a polyhedron, then Algorithm 4 converges R-linearly for τ sufficiently small, i.e., $\limsup_{k\to\infty}(\|y_{(k)} - \bar{y}\|)^{\frac{1}{k}} < 1$, [Pac+16, Prop. 1].*

We conclude this section observing that, while there are other gradient-based algorithms that allow to solve $\mathrm{VI}(\mathcal{Y}, T)$ in a decentralized fashion, they typically require a higher number of gradient steps in each iteration. For example, the extragradient algorithm [FP07, Alg. 12.1.9] requires two updates for both x and λ at each iteration.

6.3 Appendix

6.3.1 Proofs of the results presented in Sections 6.1 and 6.2

Proof of Proposition 8

Proof. Under Assumptions 1 and 8 the set $\mathcal{Q}$, and consequently the sets $\{\mathcal{X}^i\}_{i=1}^M$, $\mathcal{X}$ and $\mathcal{Y}$, are convex and satisfy Slater's constraint qualification. The VI$(\mathcal{Q}, F)$ is therefore equivalent to its KKT system [FP07, Prop. 1.3.4]. Moreover, since $\mathcal{X}^i$ satisfies Slater's constraint qualification, the optimization problem of agent i in the game (6.2) is equivalent to its KKT system, for each i. Finally, by [FP07, Prop. 1.3.4], the VI$(\mathcal{Y}, T)$ is equivalent to its KKT system. We do not report the three KKT systems here, but it can be seen by direct inspection that they are identical [Scu+12, Section 4.3.2]. □

Proof of Theorem 5

Proof. We split the proof of the theorem into two parts. First we show convergence of the inner loop and then of the outer loop.

Inner loop. Using the same approach of [Gra+16, Thm. 3 and Cor. 1], it is possible to show that under Assumption 9 for any $\lambda_{(k)} \in \mathbb{R}^m_{\geq 0}$ the sequences of $z_{(h)}$ and of $\tilde{x}(h)$ converge respectively to $\bar{z}$ and to $\bar{x}$ such that $\bar{z} = \frac{1}{M}\sum_{i=1}^M x^i_{\text{or}}(\bar{z}, \lambda_{(k)}) =: \frac{1}{M}\sum_{i=1}^M \bar{x}^i = \sigma(\bar{x})$. In [Gra+16, Thm. 1] it is shown that the set $\{\bar{x}^i\}_{i=1}^M$ is an ε-Nash equilibrium for the game $\mathcal{G}(\lambda_{(k)})$, with $\varepsilon = \mathcal{O}(\frac{1}{M})$. In the following, we show that $\{\bar{x}^i\}_{i=1}^M$ is actually a Wardrop equilibrium of $\mathcal{G}(\lambda_{(k)})$Indeed, for each agent i, by the definition of optimal response in (6.3), one has

$$J^i(\bar{x}^i, \bar{z}) + \lambda_{(k)}^\top A_{(:,i)}\bar{x}^i \leq J^i(x^i, \bar{z}) + \lambda_{(k)}^\top A_{(:,i)}x^i, \forall x^i \in \mathcal{X}^i\,.$$

Using the fact that $\bar{z} = \sigma(\bar{x})$, we get

$$J^i(\bar{x}^i, \sigma(\bar{x})) + \lambda_{(k)}^\top A_{(:,i)}\bar{x}^i \leq J^i(x^i, \sigma(\bar{x})) + \lambda_{(k)}^\top A_{(:,i)}x^i,$$

for all $x^i \in \mathcal{X}^i$ and for all $i \in \{1, \ldots, M\}$. Thus $\{\bar{x}^i\}_{i=1}^M$ is a Wardrop equilibrium of $\mathcal{G}(\lambda_{(k)})$ by Definition 7.

Outer loop. We follow the steps of the proof of [Pan+10, Proposition 8]. For each $\lambda \in \mathbb{R}^m_{\geq 0}$ define $F_{\text{W}}(x; \lambda) := F_{\text{W}}(x) + A^\top\lambda$. Such operator is strongly monotone in x on $\mathcal{Q}$ with the same constant α as $F_{\text{W}}(x)$. It follows by Lemma 1, that $\mathcal{G}(\lambda)$ has a unique variational Wardrop equilibrium which we denote by $\bar{x}_{\text{W}}(\lambda)$. Note that the outer loop update can be written as

$$\lambda_{(k+1)} = \Pi_{\mathbb{R}^m_{\geq 0}}[\lambda_{(k)} - \tau(b - A\bar{x}_{\text{W}}(\lambda_{(k)}))],$$

which is a step of the projection algorithm [FP07, Alg. 12.1.4] applied to $\mathrm{VI}(\mathbb{R}^m_{\geq 0}, \Phi)$, with $\Phi(\lambda) \coloneqq b - A\bar{x}_{\mathrm{W}}(\lambda)$. To conclude, it suffices to show that $\lambda_{(k)}$ converges to a solution $\bar{\lambda}$ of such VI, because by [FP07, Prop. 1.1.3], $\bar{\lambda}$ solves $\mathrm{VI}(\mathbb{R}^m_{\geq 0}, \Phi)$ if and only if $0 \leq \bar{\lambda} \perp (b - A\bar{x}_{\mathrm{W}}(\bar{\lambda})) \geq 0$. Having already proved convergence of the inner loop, the conclusion then follows from the second statement of Proposition 8.

To show that the sequence $\lambda_{(k)}$ converges to a solution of the $\mathrm{VI}(\mathbb{R}^m_{\geq 0}, \Phi)$, we prove that the mapping Φ is co-coercive (see Chapter 3) with co-coercitivity constant $c_\Phi = \alpha/\|A\|^2$ and apply [FP07, Thm. 12.1.8] to conclude the proof. Note that [FP07, Thm. 12.1.8] requires $\mathrm{VI}(\mathbb{R}^m_{\geq 0}, \Phi)$ to have at least a solution; this is guaranteed by the equivalence between the first two statements in Proposition 8 upon noting that a solution of $\mathrm{VI}(Q, F)$ exists by Lemma 1.

To show co-coercitivity of Φ, consider $\lambda_1, \lambda_2 \in \mathbb{R}^m_{\geq 0}$ and the corresponding unique solutions $x_1 \coloneqq \bar{x}_{\mathrm{W}}(\lambda_1)$ of $\mathrm{VI}(\mathcal{X}, F_{\mathrm{W}} + A^\top \lambda_1)$ and $x_2 \coloneqq \bar{x}_{\mathrm{W}}(\lambda_2)$ of $\mathrm{VI}(\mathcal{X}, F_{\mathrm{W}} + A^\top \lambda_2)$. By definition

$$(x_2 - x_1)^\top (F_{\mathrm{W}}(x_1) + A^\top \lambda_1) \geq 0\,, \tag{6.5a}$$
$$(x_1 - x_2)^\top (F_{\mathrm{W}}(x_2) + A^\top \lambda_2) \geq 0\,. \tag{6.5b}$$

Adding (6.5a) and (6.5b) we obtain $(x_2 - x_1)^\top (F_{\mathrm{W}}(x_1) - F_{\mathrm{W}}(x_2) + A^\top(\lambda_1 - \lambda_2)) \geq 0$, i.e., $(x_2 - x_1)^\top A^\top (\lambda_1 - \lambda_2) \geq (x_2 - x_1)^\top (F_{\mathrm{W}}(x_2) - F_{\mathrm{W}}(x_1))$. Since F_{W} is strongly monotone, it follows from the last inequality that

$$(Ax_2 - Ax_1)^\top (\lambda_1 - \lambda_2) \geq \alpha \|x_2 - x_1\|^2\,. \tag{6.6}$$

Since by definition $\|A(x_2 - x_1)\| \leq \|A\| \|x_2 - x_1\|$, then

$$\|x_2 - x_1\|^2 \geq \frac{\|A(x_2 - x_1)\|^2}{\|A\|^2}\,. \tag{6.7}$$

Combining (6.6), (6.7), and adding and subtracting b, we obtain

$$(b - Ax_2 - (b - Ax_1))^\top (\lambda_2 - \lambda_1) \geq \frac{\alpha}{\|A\|^2} \|b - Ax_2 - (b - Ax_1)\|^2,$$

hence Φ is co-coercive in λ with constant $c_\Phi = \alpha/\|A\|^2$. □

Proof of Theorem 6

Proof. We give the proof for a strongly monotone operator F, which is to be interpreted as F_{N} in the first statement and F_{W} in the second statement. We divide the proof into two parts: i) we prove that Algorithm 4 is a particular case of a class of algorithms known as asymmetric projection algorithms (APA) [FP07, Alg. 12.5.1] applied to $\mathrm{VI}(\mathcal{Y}, T)$; ii) we prove that our algorithm satisfies a convergence condition for APA. It can be shown

that if τ satisfies (6.4) then also $\tau < 1/\|A\|$ holds.
i) The APA are parametrized by the choice of a matrix $D \succ 0$. For a fixed D a step of the APA for $\mathrm{VI}(\mathcal{Y}, T)$ is

$$y_{(k+1)} = \text{solution of } \mathrm{VI}(\mathcal{Y}, T_D^k), \tag{6.8}$$

where $y_{(k)}$ is the state at iteration k and $T_D^k(y) \coloneqq T(y_{(k)}) + D(y - y_{(k)})$. Every step of the APA requires the solution of a different variational inequality that depends on the operator T, on a fixed matrix D and on the previous strategies' vector $y_{(k)}$. We choose

$$D \coloneqq \begin{bmatrix} \frac{1}{\tau} I_{Mn} & 0 \\ -2A & \frac{1}{\tau} I_m \end{bmatrix}, \tag{6.9}$$

which by using the Schur complement condition can be shown to positive definite because $\tau < 1/\|A\|$. It is shown in [FP07, Sec. 12.5.1] that with the choice (6.9) the update (6.8) coincides with the steps of Algorithm 4.
ii) As illustrated in the previous point, Algorithm 4 is the specific APA associated with the choice of D given in (6.9). According to [FP07, Prop. 12.5.2], this algorithm converges if the mapping $G(y) = D_s^{-1/2} T(D_s^{-1/2} y) - D_s^{-1/2}(D - D_s) D_s^{-1/2} y$ is co-coercive with constant 1, where $D_s = (D + D^\top)/2$ and $D_s^{-1/2}$ denotes the principal square root of the symmetric positive definite matrix D_s^{-1} and is therefore symmetric positive definite. Let us rename $L \coloneqq D_s^{-1/2}$ and $Ly = \begin{bmatrix} v \\ w \end{bmatrix}$ and simplify the expression of $G(y)$

$$\begin{aligned} G(y) &= LT(Ly) - L(D - D_s)Ly \\ &= L\left(\begin{bmatrix} F(v) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & A^\top \\ -A & 0 \end{bmatrix} Ly + \begin{bmatrix} 0 \\ b \end{bmatrix} \right) - L \begin{bmatrix} 0 & A^\top \\ -A & 0 \end{bmatrix} Ly \\ &= L\left(\begin{bmatrix} F(v) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \right). \end{aligned} \tag{6.10}$$

We now prove that $G(y)$ is co-coercive with constant 1, i.e., that

$$(y_1 - y_2)^\top (G(y_1) - G(y_2)) - \|G(y_1) - G(y_2)\|^2 \geq 0. \tag{6.11}$$

Let us substitute (6.10) in the left-hand side of (6.11)

$$
\begin{aligned}
&(y_1 - y_2)^\top (G(y_1) - G(y_2)) - \|G(y_1) - G(y_2)\|^2 \\
&= (y_1 - y_2)^\top (L \begin{bmatrix} F(v_1) \\ 0 \end{bmatrix} - L \begin{bmatrix} F(v_2) \\ 0 \end{bmatrix}) - \|L \begin{bmatrix} F(v_1) \\ 0 \end{bmatrix} - L \begin{bmatrix} F(v_2) \\ 0 \end{bmatrix} \|^2 \\
&= (Ly_1 - Ly_2)^\top (\begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}) - \|L \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix} \|^2 \\
&= (\begin{bmatrix} v_1 - v_2 \\ w_1 - w_2 \end{bmatrix})^\top (\begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}) - \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}^\top L^2 \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix} \\
&= (F(v_1) - F(v_2))^\top [(v_1 - v_2) - [L^2]_{11}(F(v_1) - F(v_2))] \\
&\geq \alpha \|v_1 - v_2\|^2 - \|[L^2]_{11}\| \|F(v_1) - F(v_2))\|^2 \\
&\geq \left(\alpha - \|[L^2]_{11}\| L_F^2\right) \|v_1 - v_2\|^2 =: K \|v_1 - v_2\|^2,
\end{aligned}
$$

The proof is concluded if $K \geq 0$. Let us compute $[L^2]_{11} = [D_s^{-1}]_{11}$. By inverting the block matrix D_s we get

$$[L^2]_{11} = \tau(I - \tau^2 A^\top A)^{-1} \succ 0. \tag{6.12}$$

Since $\tau^2 A^\top A$ is symmetric positive semidefinite, $\lambda_{\max}(\tau^2 A^\top A) = \tau^2 \|A\|^2 < 1$ because $\tau < 1/\|A\|$ and $\rho(\tau^2 A^\top A) < 1$, i.e., the matrix is convergent. Hence, the Neumann series $\sum_{k=0}^{\infty} (\tau^2 A^\top A)^k$ converges to $(I - \tau^2 A^\top A)^{-1}$. Substituting in (6.12) yields

$$[L^2]_{11} = \tau \sum_{k=0}^{\infty} (\tau^2 A^\top A)^k \succeq 0 \quad \text{and} \quad \|[L^2]_{11}\| \leq \tau \sum_{k=0}^{\infty} (\tau^2 \|A\|^2)^k = \frac{\tau}{1 - \tau^2 \|A\|^2},$$

where we used the fact that the geometric series converges since $\tau^2 \|A\|^2 < 1$. Therefore $K \geq \alpha - \frac{\tau}{1-\tau^2\|A\|^2} L_F^2$. By condition (6.4) we get $\alpha\tau^2\|A\|^2 + \tau L_F^2 < \alpha$ and thus

$$K \geq \frac{\alpha - \alpha\tau^2\|A\|^2 - \tau L_F^2}{1 - \tau^2 \|A\|^2} > 0.$$

□

CHAPTER 7

Applications

In this chapter we verify the theoretical results derived in the previous two chapters. In particular, we consider a coordination problem arising in the charging of electric vehicles, and a selfish routing model used in road traffic network. All the proofs are reported in the Appendix (Section 7.3). The results presented in this chapter have been published in [PKL16; Pac+16; Pac+18].

7.1 Charging of electric vehicles

Electric-vehicles (EV) are foreseen to significantly penetrate the market in the coming years [NB+10], therefore coordinating their charging schedules can provide useful services for the operation of the grid, e.g., peak shaving, ancillary services [GTL13]. In the following we model this problem as a game, where vehicles owners wish to minimize their total electricity bill, while requiring a sufficient final state of charge. By assuming that the electricity price depends on the aggregate consumption, [MCH13; Gra+16] formulate the EV charging problem as an aggregative game and propose decentralized schemes, in the absence of coupling constraints. In this section, we show how the results derived in the previous chapters can be used to study this problem. In particular, our formulation extends the existing literature by introducing coupling constraints and by relaxing the assumptions required for the convergence of the corresponding algorithms.[1] In addition, we study the performance degradation of an equilibrium configuration, when compared to the centralized optimal solution. Finally, we establish uniqueness of the dual variables associated to the violation of the coupling constraints.

In the remainder of this section, we consider a population of M electric vehicles and identify with agent i the corresponding vehicle $i \in \{1, \dots, M\}$. Additionally, we identify with s_t^i the state of charge of vehicle i at time t. The time evolution of s_t^i is specified by the discrete-time system $s_{t+1}^i = s_t^i + b^i x_t^i \,, t = 1, \dots, n$, where x_t^i is the charging input and the parameter $b^i > 0$ captures the charging efficiency.

[1] Coupling constraints model limits on the aggregate peak consumption or on the local consumption of EVs connected to the same transformer.

Constraints

We assume that the charging input cannot take negative values and that at time t it cannot exceed $\tilde{x}_t^i \geq 0$. The final state of charge is constrained to $s_{n+1}^i \geq \eta^i$, where $\eta^i \geq 0$ is the desired state of charge of agent i. Denoting with $x^i = [x_1^i, \dots, x_n^i]^\top \in \mathbb{R}^n$, the individual constraint of agent i can be expressed as

$$x^i \in \mathcal{X}^i := \left\{ x^i \in \mathbb{R}^n \;\middle|\; \begin{array}{l} 0 \leq x_t^i \leq \tilde{x}_t^i, \quad \forall\, t = 1, \dots, n \\ \sum_{t=1}^n x_t^i \geq \theta^i \end{array} \right\}, \tag{7.1}$$

where $\theta^i := (b^i)^{-1}(\eta^i - s_1^i)$, with $s_1^i \geq 0$ the state of charge at the beginning of the time horizon. Besides the individual constraints $x^i \in \mathcal{X}^i$, we introduce the coupling constraint

$$x \in \mathcal{C} := \left\{ x \in \mathbb{R}^{Mn} \;\middle|\; \frac{1}{M} \sum_{i=1}^M x_t^i \leq K_t,\ \forall\, t = 1, \dots, n \right\}, \tag{7.2}$$

indicating that at time t the grid cannot deliver more than MK_t units of power to the vehicles. In compact form (7.2) reads as $(\mathbb{1}_M^\top \otimes I_n)x \leq MK$, where $K := [K_1, \dots, K_n]^\top$.

Cost function

The cost function of each vehicle represents its electricity bill, which we model as

$$J^i(x^i, \sigma(x)) = \sum_{t=1}^n p_t \left(\frac{d_t + \sigma_t(x)}{\kappa_t} \right) x_t^i =: p(\sigma(x))^\top x^i, \tag{7.3}$$

where we have assumed that the energy price for each time interval $p_t : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ depends on the ratio between total consumption and total capacity $(d_t + \sigma_t(x))/\kappa_t$, where d_t and $\sigma_t(x) := \frac{1}{M} \sum_{i=1}^M x_t^i$ are the non-EV and EV demand at time t divided by M and κ_t is the total production capacity divided by M as in [MCH13, Eq. (6)]. The quantity κ_t is in general not related to K_t.

7.1.1 Theoretical guarantees

We define the game $\mathcal{G}_M^{\text{EV}}$ as in (4.3), with $\mathcal{X}^i$, $\mathcal{C}$ and $J^i(x^i, \sigma(x))$ as in (7.1), (7.2) and (7.3) respectively. In the following corollary we refine the main results of Chapter 4, 5 and Chapter 6 for the EV application.

Corollary 1. *Consider a sequence of games $(\mathcal{G}_M^{\text{EV}})_{M=1}^\infty$. Assume that there exists $\tilde{x}^0$ such that $\tilde{x}_t^i \leq \tilde{x}^0$ for all $t \in \{1, \dots, n\}, i \in \{1, \dots, M\}$ and for each game $\mathcal{G}_M^{\text{EV}}$. Moreover, assume that for each game $\mathcal{G}_M^{\text{EV}}$ the set $\mathcal{Q} = \mathcal{C} \cap \mathcal{X}$ is non-empty and that for each t the price function p_t in (7.3) is twice continuously differentiable, strictly increasing and Lipschitz in $[0, \tilde{x}^0]$ with constant L_p. Then:*

(a) *A Wardrop and a Nash equilibrium exist for each game* $\mathcal{G}_M^{\mathrm{EV}}$ *of the sequence. Furthermore, every Wardrop equilibrium is an* ε*-Nash equilibrium with* $\varepsilon = \frac{2n(\tilde{x}^0)^2 L_p}{M}$.

(b) *The function* p *is strongly monotone, hence for each game* $\mathcal{G}_M^{\mathrm{EV}}$ *there exists a unique* $\bar{\sigma}$ *such that* $\sigma(\bar{x}_{\mathrm{W}}) = \bar{\sigma}$ *for any variational Wardrop equilibrium* $\bar{x}_{\mathrm{W}}$ *of* $\mathcal{G}_M^{\mathrm{EV}}$. *Moreover for any variational Nash equilibrium* $\bar{x}_{\mathrm{N}}$ *of* $\mathcal{G}_M^{\mathrm{EV}}$, $\|\sigma(\bar{x}_{\mathrm{N}}) - \sigma(\bar{x}_{\mathrm{W}})\| \leq \tilde{x}^0 \sqrt{\frac{2nL_p}{\alpha M}}$, *where* α *is the monotonicity constant of* p.

(c) *Assume that there is no coupling constraint, i.e.,* $\mathcal{C} = \mathbb{R}^{Mn}$, *that* $d_t > 0$ *for all* t, *and that* $\sum_{i=1}^{M} \theta^i > 0$. *If* $p_t\left(\frac{d_t + \sigma_t(x)}{\kappa_t}\right) = \alpha \left(\frac{d_t + \sigma_t(x)}{\kappa_t}\right)^k$ *with* $\alpha > 0$, $k > 0$, *then*

$$1 \leq \mathrm{PoA}_M \leq 1 + \mathcal{O}\left(1/\sqrt{M}\right) \quad \text{and} \quad \lim_{M \to \infty} \mathrm{PoA}_M = 1\,.$$

(d) *Assume that*

$$\min_{\substack{t \in \{1,\dots,n\} \\ z \in [0, \tilde{x}^0]}} \left(p_t'(z) - \frac{\tilde{x}^0 p_t''(z)}{8} \right) > 0. \tag{7.4}$$

For each game $\mathcal{G}_M^{\mathrm{EV}}$ *the operator* F_{N} *is strongly monotone. Hence, if Assumption 8 holds, Algorithm 4 converges to a variational Nash equilibrium of* $\mathcal{G}_M^{\mathrm{EV}}$.

We note that the previous corollary provides guarantees on the equilibrium efficiency for the case of polynomial price functions. Nevertheless, different results can be obtained in the case of affine or diagonal price function by applying the bounds derived in Theorems 2 and 4. In this respect, the third statement of Corollary 1 is purely exemplificative.

Uniqueness of dual variables

Corollary 1 shows that under condition (7.4) the operator F_{N} of $\mathcal{G}_M^{\mathrm{EV}}$ is strongly monotone, hence the game $\mathcal{G}_M^{\mathrm{EV}}$ admits a unique variational Nash equilibrium (Lemma 1). We study here the uniqueness of the associated dual variables $\bar{\lambda}_{\mathrm{N}}$ introduced in Proposition 8. Guaranteeing unique dual variables is important to convince the vehicles owners to participate in the proposed scheme, as $\bar{\lambda}_{\mathrm{N}}$ represent the penalty price associated to the coupling constraint. Define $R^{\mathrm{tight}} \subseteq \{1, \dots, n\}$ as the set of instants in which $\mathcal{C}$ is active. We provide a sufficient condition for uniqueness of the dual variables which relies on a modification of the linear-independence constraint qualification [Wac13].

Lemma 10. *Assume that condition (7.4) holds and consider the unique variational Nash equilibrium* $\bar{x}_{\mathrm{N}}$ *of* $\mathcal{G}_M^{\mathrm{EV}}$. *If there exists a vehicle* i *such that* $\bar{x}_{\mathrm{N},t}^i \notin \{0, \tilde{x}_t^i\}$ *for all* $t \in R^{\mathrm{tight}}$ *and* $\bar{x}_{\mathrm{N},t'}^i \notin \{0, \tilde{x}_{t'}^i\}$ *for some* $t' \notin R^{\mathrm{tight}}$, *then the dual variables* $\bar{\lambda}_{\mathrm{N}}$ *associated to the coupling constraint (7.2) are unique.*

We note that the sufficient condition of Lemma 10 has to be verified a-posteriori as it depends on the primal solution $\bar{x}_{\mathrm{N}}$. In the numerical analysis presented in the following such sufficient condition always holds. Uniqueness of the dual variables associated to the coupling constraint of an aggregative game has been studied also in [YSM11, Thm. 4], where the conditions in the bullets of Lemma 10 are not required, but p is restricted to be affine.

7.1.2 Numerical analysis

The numerical study is conducted on a heterogeneous population of agents. We set the price function to $p_t(z_t) = 0.15\sqrt{z_t}$ and $n = 24$. The agents differ in θ^i, randomly chosen according to $\mathcal{U}[0.5, 1.5]$; they also differ in $\tilde{x}_t^i$, which is chosen such that the charge is allowed in a connected interval, with left and right endpoints uniformly randomly chosen. Within this interval, $\tilde{x}_t^i$ is constant and randomly chosen for each agent according to $\mathcal{U}[1, 5]$, while outside this interval $\tilde{x}_t^i = 0$. The demand d_t is taken as the typical (non-EV) base demand over a summer day in the United States [MCH13, Fig. 1]; $\kappa_t = 12$ kW for all t, and the upper bound $K_t = 0.55$ kW is chosen such that the coupling constraint (7.2) is active in the middle of the night. Note that with these choices all the assumptions of Corollary 1 are met. In particular, for the given choice of p condition (7.4) holds because $p_t''(z) < 0$ for all z and all t. Figure 7.1 presents the aggregate consumption at the Nash equilibrium found by Algorithm 4, with stopping criterion $\|(x_{(k+1)}, \lambda_{(k+1)}) - (x_{(k)}, \lambda_{(k)})\|_\infty \leq 10^{-4}$. Note that without the coupling constraint the quantity $\bar{\sigma}+d$ would be constant overnight, as shown in [MCH13].

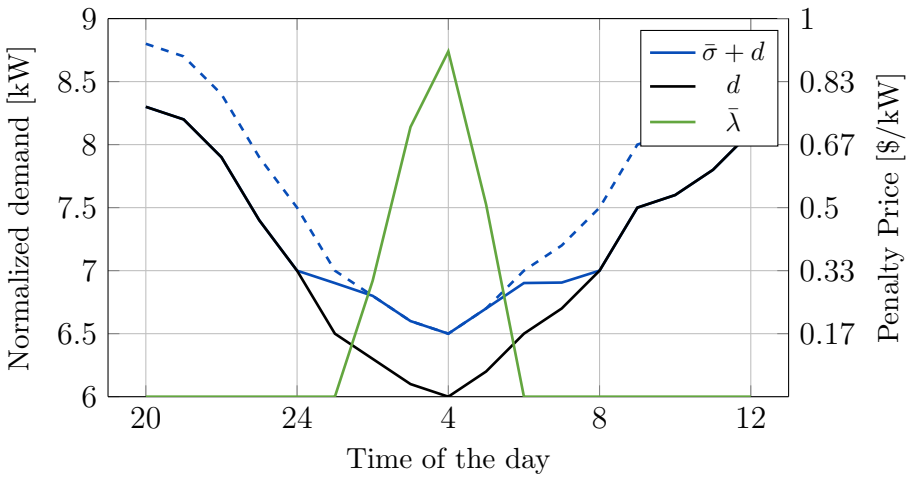

Figure 7.1: Aggregate EV demand $\sigma(\bar{x}_{\mathrm{N}})$ and dual variables $\bar{\lambda}_{\mathrm{N}}$ for $M = 100$, subject to $\sigma(x) \leq 0.55$ kW. The region below the dashed line satisfies $\sigma(x) + d \leq 0.55$ kW$+d$.

Figure 7.2 illustrates the bound $\|\sigma(\bar{x}_{\mathrm{N}}) - \sigma(\bar{x}_{\mathrm{W}})\| \leq \tilde{x}^0 \sqrt{\frac{2nL_p}{\alpha M}}$ of the second statement of Corollary 1. The Wardrop equilibrium is computed with the extragradient algorithm with stopping criterion $\|(x_{(k+1)}, \lambda_{(k+1)}) - (x_{(k)}, \lambda_{(k)})\|_\infty \leq 10^{-4}$. The framework introduced above can also be used to enforce local coupling constraints, i.e., constraints on a subset of all the vehicles. These can for instance be used to model capacity limits for local substations as we discuss in [Pac+16, Fig. 4].

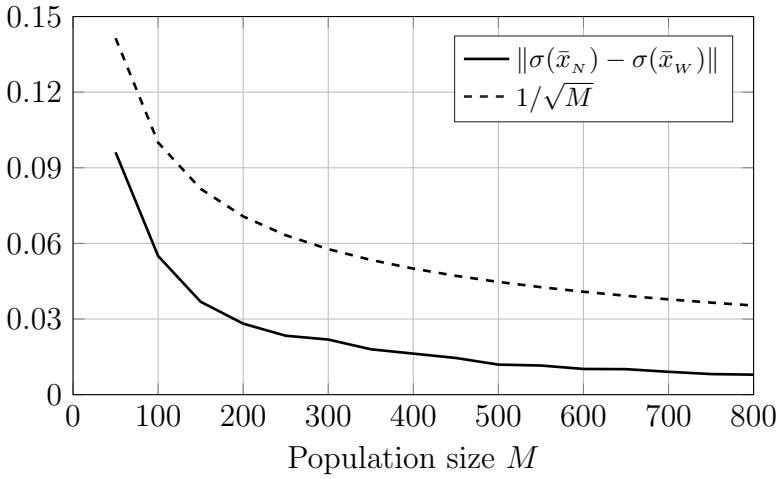

Figure 7.2: Distance between the aggregates at the Nash and Wardrop equilibrium (solid line). Corollary 1 ensures that such distance is upper bounded by $\tilde{x}^0 \sqrt{2nL_p \alpha^{-1}/M}$. The dotted line shows $1/\sqrt{M}$ proving that our bound captures the correct trend.

The case of linear price function

Different works in the EV literature [Gra+16; KCM11] use the cost (4.10), with $Q \succ 0$ and $C \succ 0$, diagonal. Existence of a Nash and of a Wardrop equilibrium is guaranteed by Lemma 1, while Proposition 7 gives the ε-Nash property. Further, Lemma 4 shows that the resulting operators F_{N} and F_{W} are strongly monotone with monotonicity constant independent from M. Theorem 1 ensures then that $\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \leq L_2/(\alpha\sqrt{M})$, with $L_2 = R \cdot \lambda_{\mathrm{M}}$, where λ_{M} represents the largest eigenvalue of C. A Nash equilibrium can be found using Algorithm 4, while a Wardrop equilibrium can be achieved using both Algorithms 3 and 4. Figure 7.3 presents a comparison between the two algorithms in terms of iteration count, where $Q = 0.1I_n$, $C = I_n$, $c^i = d$ for all i. Figure 7.3 (top) represents the number of strategy updates required to converge, i.e., the number of times line 5 in Algorithm 3 or line 4 in Algorithm 4 is used. Figure 7.3 (bottom) depicts the number of dual variables updates, i.e., the number of times line 11 in Algorithm 3 or line 5 in Algorithm 4 is used. For both algorithms the number of iterations does not seem to increase with the population size. Algorithm 4 requires fewer primal iterations, while Algorithm 3 needs much fewer dual iterations.

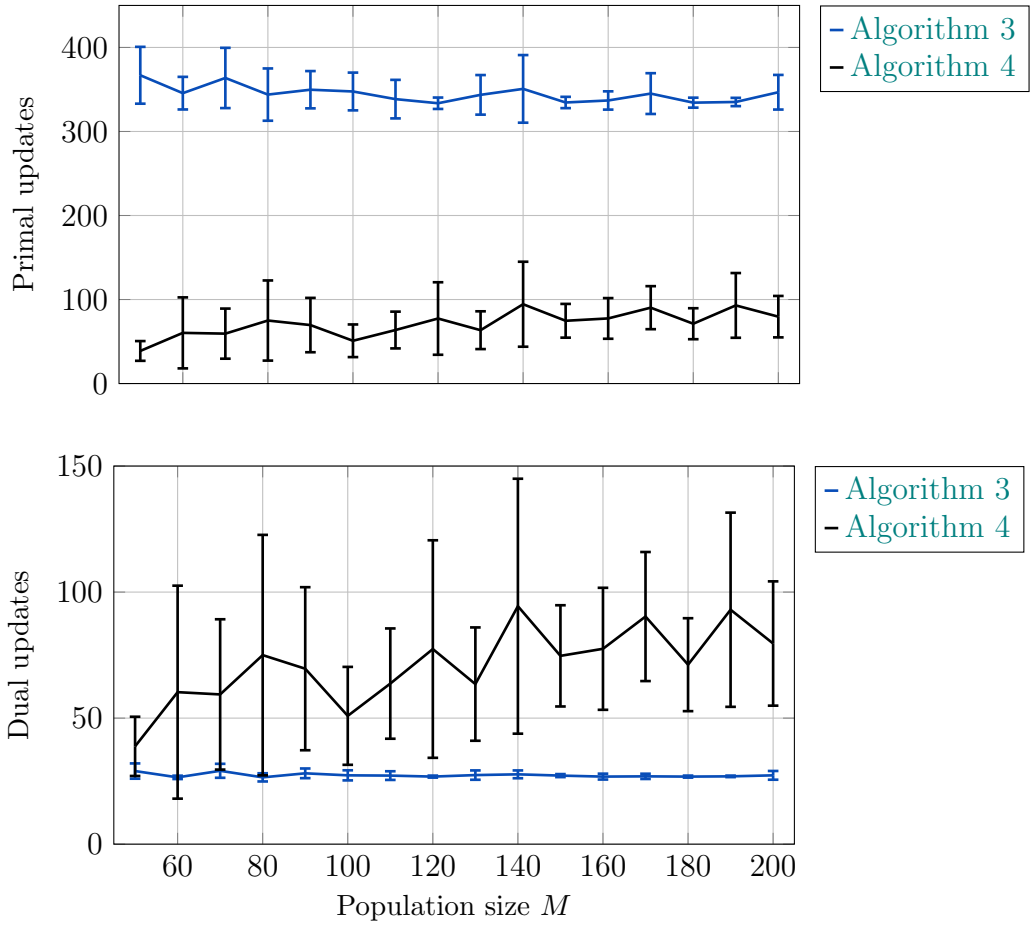

Figure 7.3: Primal (top) and dual (bottom) updates required to converge; mean and standard deviation for 10 repetitions. As Algorithm 4 performs one primal and one dual update in each iteration, the black lines appearing in the two figures coincide.

Equilibrium efficiency

In this section we verify the theoretical results on the efficiency of equilibria obtained in Corollary 1, by means of numerical simulations. We consider four cases as follows.

Case 1. We set $p_t(y) = 0.15y^3$ and choose $\tilde{x}_t^i$ to allow charging in $[t^i_{\min}, t^i_{\max}]$, with $t^i_{\min}, t^i_{\max}$ uniformly randomly distributed between 5pm and 10am; $\theta^i \sim \mathcal{U}[5, 15]$ and d_t as in [MCH13, Fig. 1].

Cases 2-4. We set $p_t(y) = 0.15$ from 5pm to 1am and $p_t(y) = 0.15y$ from 2am to 10am. For all vehicles, we choose $\tilde{x}_t^i$ to allow charging from 5pm to 10am. Cases 2-4 differ in θ^i, d_t as in the following table.

Case	θ^i	d_t
2	9	$\mathbb{0}_n$
3	9	as in [MCH13, Fig. 1]
4	$\mathcal{U}[5, 13]$	$\mathbb{0}_n$

For each case, we report the (numerical) price of anarchy as a function of M in Figure 7.4 (top). Observe that case 1 and 4 feature heterogenous charging needs. For these cases, we have randomly extracted 100 games $\mathcal{G}_M^{\text{EV}}$ (for any fixed M) and report the worst PoA amongst the 100 realization. In order to plot the price of anarchy, we computed the ratio between *one* (instead of the *worst*) Nash equilibrium of $\mathcal{G}_M^{\text{EV}}$ and the social optimum. This choice is imposed by the fact that computing all Nash equilibria of $\mathcal{G}_M^{\text{EV}}$ is in general a hard problem.[2] In Figure 7.4 (bottom) we plot the difference between the cost at the Nash and at the social optimizer, relative to case 1.

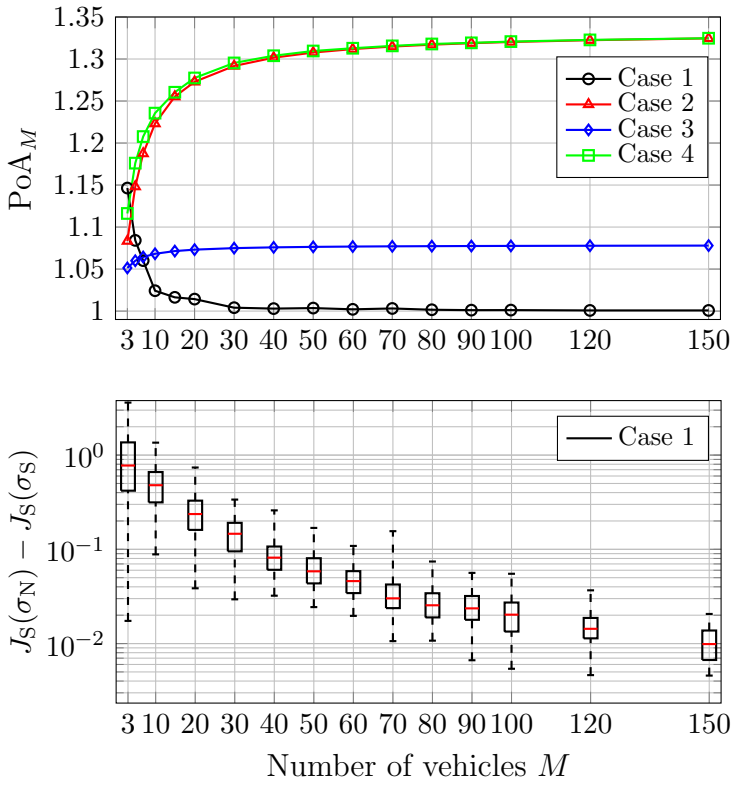

Figure 7.4: Price of anarchy (top), and cost difference between Nash and social optimum (bottom) as a function of M.

[2]This is due to the fact that the operator associated with the variational inequality of the Nash problem is not guaranteed to be strongly monotone since condition (7.4) does not hold due to the choices of $p(t)$ in Cases 1-4. To compute a Nash equilibrium we applied the extragradient algorithm [FP07], which is though not guaranteed to converge. We thus verified a posteriori that the point where the algorithm stopped was a Nash equilibrium.

Thanks to the choice of parameters and price function, the third statement in Corollary 1 guarantees that $\lim_{M\to\infty} \mathrm{PoA}_M = 1$. The numerical results reported in Figure 7.4 (top, black line) are consistent with it: the ratio between the cost at the Nash and the cost at the social optimum converges to one. In addition to this, Figure 7.4 (bottom) shows that also the difference between these costs converges to zero, as guaranteed in the proof of Theorem 3 by the boundedness of $\mathcal{X}_0$. Case 2 has been constructed so that the corresponding Wardrop equilibrium features the worst possible asymptotic price of anarchy within the class of affine cost functions (for which $\alpha(\mathcal{L}) = 4/3$, see [Rou03]). The numerics of Figure 7.4 (top, red line) show that PoA_M (i.e., the efficiency of Nash equilibria) converges to $1.33 \approx 4/3 = \alpha(\mathcal{L})$. Cases 3 and 4 are a modification of case 2. While the presence of base demand (case 3) helps in lowering the price of anarchy, the impact of heterogeneity (case 4) on the asymptotic price of anarchy is minor (blue and green plots in Figure 7.4).

7.2 Route choice in a road network

As second application we consider that of traffic routing in a road network. Traffic congestion is a well-recognized issue in densely populated cities, and the corresponding economic costs are significant [AS94]. Since every driver seeks his own interest (e.g., minimizing the travel time) and is affected by the others' choices via congestion, a classic approach is to model the traffic problem as a game [Daf80]. In the following we focus on a stationary model that aims at capturing the basic interactions among the vehicles flow during rush hours. Building upon our theoretical findings, we derive results specific for the route choice game. Moreover, we perform a realistic numerical analysis based on the data set of the city of Oldenburg in Germany [Bri02]. Specifically, we investigate via simulation the effect of road access limitations, expressed as coupling constraints [San75].

We consider a strongly-connected directed graph $(\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V} = \{1, \dots, V\}$, representing geographical locations, and directed edge set $\mathcal{E} = \{1, \dots, E\} \subseteq \mathcal{V} \times \mathcal{V}$, representing roads connecting the locations. Each agent $i \in \{1, \dots, M\}$ represents a driver who wants to drive from his origin $o^i \in \mathcal{V}$ to his destination $d^i \in \mathcal{V}$.

Constraints

Let us introduce the vector $x^i \in [0,1]^E$ to describe the strategy (route choice) of agent i, with $[x^i]_e$ representing the probability that agent i transits on edge e [DP05]. To guarantee that agent i leaves his origin and reaches his destination with probability 1, the strategy x^i has to satisfy

$$\sum_{e\in\mathrm{in}(v)} [x^i]_e - \sum_{e\in\mathrm{out}(v)} [x^i]_e = \begin{cases} -1 & \text{if} \quad v = o^i \\ 1 & \text{if} \quad v = d^i \\ 0 & \text{otherwise,} \end{cases} \qquad \forall\, v \in \mathcal{V},$$

where $\text{in}(v)$ and $\text{out}(v)$ represent the set of in-edges and the set of out-edges of node v. We denote the graph incidence matrix by $B \in \mathbb{R}^{V \times E}$, so that $[B]_{ve} = 1$ if edge e points to vertex v, $[B]_{ve} = -1$ if edge e exits vertex v and $[B]_{ve} = 0$ otherwise. The individual constraint set of agent i is then

$$\mathcal{X}^i \coloneqq \{x \in [0,1]^E \mid Bx = b^i\}, \tag{7.5}$$

where $b^i \in \mathbb{R}^V$ is such that $[b^i]_v = -1$ if $v = o^i$, $[b^i]_v = 1$ if $v = d^i$ and $[b^i]_v = 0$ otherwise. We introduce the constraint

$$x \in \mathcal{C} \coloneqq \{x \in \mathbb{R}^{ME} \mid \tfrac{1}{M} \textstyle\sum_{i=1}^{M} x_e^i \leq K_e, \, \forall \, e = 1, \dots, E\}, \tag{7.6}$$

expressing the fact that the number of vehicles on edge e cannot exceed MK_e. Such coupling constraint can be imposed by authorities to decrease the congestion in a specific road or neighborhood, with the goal of reducing noise or pollution.

Cost function

We assume that each driver $i \in \{1, \dots, M\}$ wants to minimize his travel time and, at the same time, does not want to deviate too much from a preferred route $\tilde{x}^i \in \mathcal{X}^i$. We model this objective with the following cost function

$$J^i(x^i, \sigma(x)) = \frac{\gamma^i}{2} \|x^i - \tilde{x}^i\|^2 + \sum_{e=1}^{E} t_e(\sigma_e(x_e)) x_e^i, \tag{7.7}$$

with $\gamma^i \geq 0$ a weighting factor, $x_e \coloneqq [x_e^1, \dots, x_e^M]^\top$, $\sigma_e(x_e) = \frac{1}{M} \sum_{i=1}^{M} x_e^i$ and $t_e(\sigma_e(x_e))$ the travel time on edge e.

Travel time

This subsection is devoted to the derivation of the analytical expression of the travel time $t_e(\sigma_e(x_e))$. The reader not interested in the technical details of the derivation can proceed to the expression of $t_e(\sigma_e(x_e))$ in (7.10), which is illustrated in Figure 7.5.

In the following, we introduce the quantity $D_e(x_e) = \sum_{i=1}^{M} x_e^i$ to describe the total demand on edge e and consider a rush-hour interval $[0, h]$. We assume that the instantaneous demand equals $D_e(x_e)/h$ at any time $t \in [0, h]$ and zero for $t > h$. Additionally, we assume that edge e can support a maximum flow F_e (vehicles per unit of time) and features a free-flow travel time $t_{e,\text{free}}$. As we are interested in comparing populations of different sizes, we further assume that the peak hour duration h is independent from the population size M and that the road maximum capacity flow F_e scales linearly with the population size, i.e., $F_e(M) = f_e \cdot M$, with f_e constant in M. The consideration underpinning this last assumption is that the road infrastructure scales with the number of

vehicles to accommodate the increasing demand, similarly as what assumed in [MCH13] for the energy infrastructure.

If $D_e(x_e)/h \leq F_e$ then every car has instantaneous access to edge e and no queue accumulates, hence the travel time equals $t_{e,\text{free}}$. We focus in the rest of this paragraph on the case $D_e(x_e)/h > F_e$. An increasing queue forms in the interval $[0, h]$ and decreases at rate F_e for $t > h$. The number of vehicles $q_e(t)$ queuing on edge e at time t obeys then the dynamics

$$\dot{q}_e(t) = \begin{cases} \frac{D_e(x_e)}{h} \cdot \mathbf{1}_{[0,h]}(t) - F_e & \text{if } q_e(t) \geq 0 \\ 0 & \text{otherwise,} \end{cases} \qquad q_e(0) = 0, \tag{7.8}$$

where $\mathbf{1}_{[0,h]}$ is the indicator function of $[0, h]$. The solution $q_e(t)$ to (7.8) is hence

$$q_e(t) = \begin{cases} \left(\frac{D_e(x_e) - F_e h}{h}\right) t & \text{if } 0 \leq t \leq h \\ D_e(x_e) - F_e\, t & \text{if } h \leq t \leq D_e(x_e)/F_e \\ 0 & \text{if } t \geq D_e(x_e)/F_e. \end{cases} \tag{7.9}$$

As a consequence, the total queuing time at edge e (i.e, the queuing times summed over all vehicles) is the integral of $q_e(t)$, which equals $D_e(x_e)(D_e(x_e) - F_e h)/(2F_e)$; the queuing time is then $(D_e(x_e) - F_e h)/(2F_e)$.

Since $\sigma_e(x_e) = \frac{1}{M}\sum_{i=1}^{M} x_e^i = \frac{1}{M} D_e(x_e)$, the travel time is

$$t_e^{\text{PWA}}(\sigma_e(x_e)) = \begin{cases} t_{e,\text{free}} & \text{if } \sigma_e(x_e) \leq f_e h \\ t_{e,\text{free}} + \frac{\sigma_e(x_e) - f_e h}{2 f_e} & \text{otherwise,} \end{cases}$$

and is reported in Figure 7.5. Note that t_e^{PWA} is a continuous and piece-wise affine function of $\sigma_e(x_e)$, but it is not continuously differentiable, hence Assumption 1 would not hold. Therefore, we define t_e appearing in (7.7) as the smoothed version of t_e^{PWA}

$$t_e(\sigma_e(x_e)) = \begin{cases} t_{e,\text{free}} & \text{if } \sigma_e(x_e) \leq f_e h - \Delta_e \\ t_{e,\text{free}} + \frac{\sigma_e(x_e) - f_e h}{2 f_e} & \text{if } \sigma_e(x_e) \geq f_e h + \Delta_e \\ a\sigma_e(x_e)^2 + b\sigma_e(x_e) + c & \text{otherwise,} \end{cases} \tag{7.10}$$

where the values of Δ_e, a, b, c are such that t_e is continuously differentiable[3], as illustrated in Figure 7.5. We note that the function $t_e(\sigma_e(x_e))$ is used within a stationary traffic model but includes the average queuing time which is based on the dynamic function (7.9). A thorough analysis of a dynamic traffic model is subject of future work.

Finally, we remark that a travel time with similar monotonicity properties can be derived from the piecewise affine fundamental diagram of traffic [LZ11, Fig. 7], but $t_e(\sigma_e(x_e))$ would present a vertical asymptote which is absent here.

[3]The values are $\Delta_e = 0.5(\sqrt{(f_e h)^2 + 4 f_e h} - f_e h)$, $a = 1/(8 f_e \Delta_e)$, $b = 1/(4 f_e) - h/(4\Delta_e)$, $c = t_{e,\text{free}} + (f_e h)^2/(8 f_e \Delta_e) - h/4 - (\Delta_e)/(8 f_e)$.

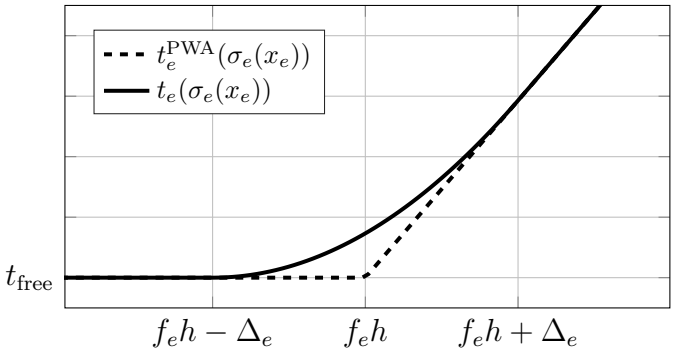

Figure 7.5: Piece-wise affine travel time $t_e^{\mathrm{PWA}}(\sigma_e(x_e))$ and its smooth approximation $t_e(\sigma_e(x_e))$ as functions of $\sigma_e(x_e)$.

7.2.1 Theoretical guarantees

We define the route-choice game $\mathcal{G}_M^{\mathrm{RC}}$ as in (4.3), with $\mathcal{X}^i$ as in (7.5), $\mathcal{C}$ as in (7.6) and $J^i(x^i, \sigma(x))$ as in (7.7), (7.10). In the following we summarize the main results from the previous chapters.

Corollary 2. *Consider the sequence of games $(\mathcal{G}_M^{\mathrm{RC}})_{M=1}^{\infty}$. Assume that for each game $\mathcal{G}_M^{\mathrm{RC}}$ the set $\mathcal{Q} = \mathcal{C} \cap \mathcal{X}$ is non-empty, that $h > 0$ and $t_{\mathrm{e,free}}, f_e > 0$ for each $e \in \mathcal{E}$. Moreover, assume that there exists $\hat{\gamma} > 0$ such that $\gamma^i \geq \hat{\gamma}$ for all $i \in \{1, \dots, M\}$, for all M. Then:*

(a) The operator F_{W} is strongly monotone, hence each game $\mathcal{G}_M^{\mathrm{RC}}$ admits a unique variational Wardrop equilibrium. For every M satisfying

$$M > \max_{e \in \mathcal{E}} \frac{1}{32 f_e \Delta_e \hat{\gamma}} \tag{7.11}$$

the operator F_{N} is strongly monotone, hence each game $\mathcal{G}_M^{\mathrm{RC}}$ admits a unique variational Nash equilibrium. Every Wardrop equilibrium is an ε-Nash equilibrium with $\varepsilon = \frac{E}{M f_{min}}$, where $f_{\min} = \min_{e \in \mathcal{E}} f_e$.

(b) For any variational Nash equilibrium $\bar{x}_{\mathrm{N}}$ of $\mathcal{G}_M^{\mathrm{RC}}$, the unique variational Wardrop equilibrium $\bar{x}_{\mathrm{W}}$ of $\mathcal{G}_M^{\mathrm{RC}}$ satisfies

$$\|\bar{x}_{\mathrm{N}} - \bar{x}_{\mathrm{W}}\| \leq \frac{\sqrt{E}}{2 f_{\min} \hat{\gamma} \sqrt{M}}.$$

(c) For any M, Algorithm 4 with operator F_{W} converges to a variational Wardrop equilibrium of $\mathcal{G}_M^{\mathrm{RC}}$. For M satisfying (7.11), Algorithm 4 with operator F_{N} converges to a variational Nash equilibrium of $\mathcal{G}_M^{\mathrm{RC}}$. □

7.2.2 Numerical analysis

For the numerical analysis we use the data set of the city of Oldenburg [Bri02], whose graph features 175 nodes, 213 undirected edges, and is reported in Figure 7.6.[4] For each agent i the origin o^i and the destination d^i are chosen uniformly at random. Regarding the cost (7.7), $t_{e,\text{free}}$ is computed as the ratio between the road length, which is provided in the data set, and the free-flow speed. Based on the road topology, we divide the roads into main roads, where the free-flow speed is 50 km/h, and secondary roads, where the free-flow speed is 30 km/h. Moreover, we assume a peak hour duration h of 2 hours, and for all $e \in \mathcal{E}$, we set $f_e = 4 \cdot 10^{-3}$ vehicles per second, which corresponds to 1 vehicle every 4 seconds for a population of $M = 60$ vehicles. Finally, the parameter γ^i is picked uniformly at random in $[0.5, 3.5]$ and $\tilde{x}^i$ is such that $\tilde{x}^i_e = 1$ if e belongs to the shortest path from o^i to d^i, while $\tilde{x}^i_e = 0$ otherwise. The shortest path is computed based on $\{t_{e,\text{free}}\}_{e=1}^E$. Note that with the above values the bound (7.11) becomes $M > 16.14$, which is satisfied for relatively small-size populations.

We compute the Wardrop equilibrium with Algorithm 4 relatively to a population of $M = 60$ drivers without coupling constraint, i.e., with $K_e = 1$ for all $e \in \mathcal{E}$. We report in Figure 7.6 the corresponding queuing time $t_e(\sigma_e(x_e)) - t_{e,\text{free}}$ as by (7.10).

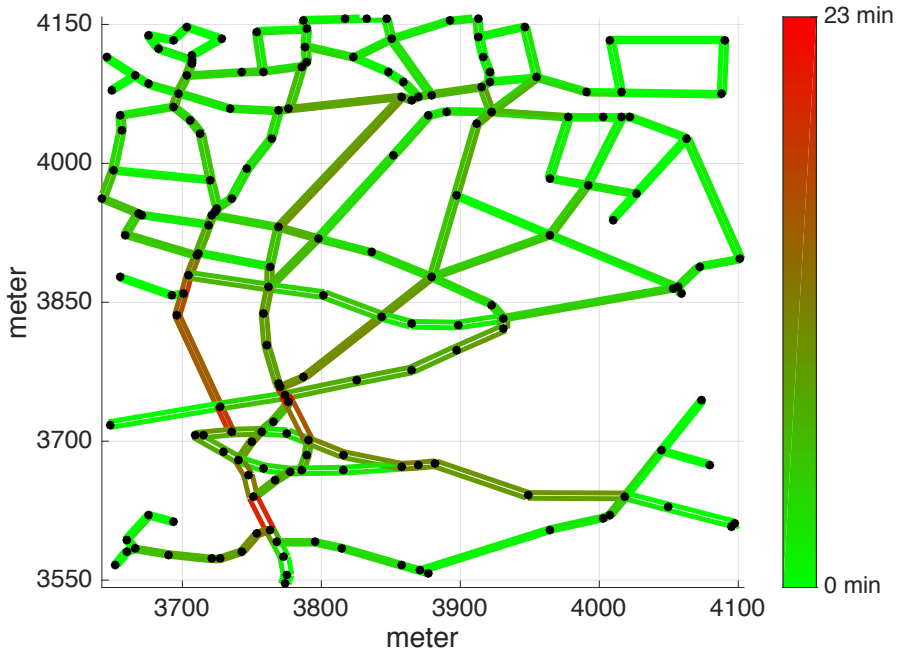

Figure 7.6: The queuing time reported in green-red color scale. Note that this pattern changes if one modifies the pairs origin-destination.

[4]The graph in the original data set features 6105 vertexes and 7035 undirected edges. We reduce it by excluding all the nodes that are outside the rectangle $[3619, 4081] \times [3542, 4158]$ and all the edges that do not connect two nodes in the rectangle. The resulting graph is strongly connected.

We illustrate in Figure 7.7 the change in the queuing time of an entire neighborhood when introducing a coupling constraint that upper bounds the total number of cars on a single edge, relatively to a Wardrop equilibrium with $M = 60$. Finally, we illustrate the second statement of Corollary 2 by reporting in Figure 7.8 the distance between the unique variational Wardrop equilibrium and the variational Nash equilibrium found by Algorithm 4.

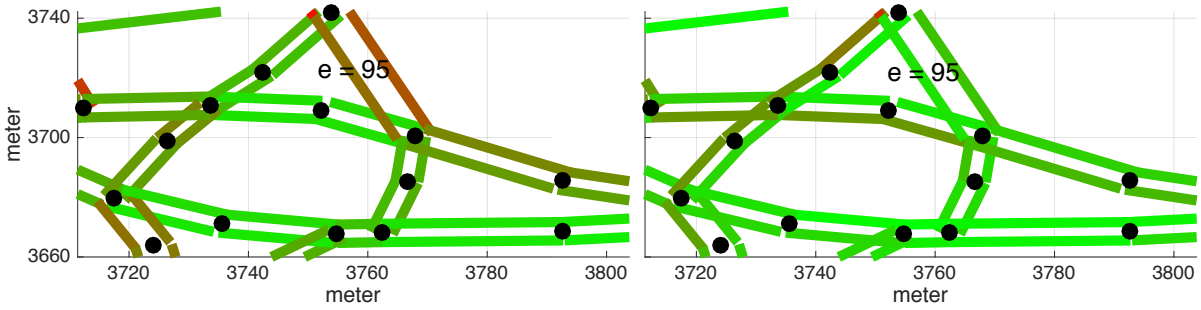

Figure 7.7: On the left, the queuing time in a neighborhood without any coupling constraints; 10% of the population transits on edge 95, and the queuing time is 7.28 minutes. On the right, the queuing time in presence of a coupling constraint allowing at most 3% of the entire population on edge 95; the queuing time is reduced to 1.42 minutes, but it visibly increases on the edges of the alternative route.

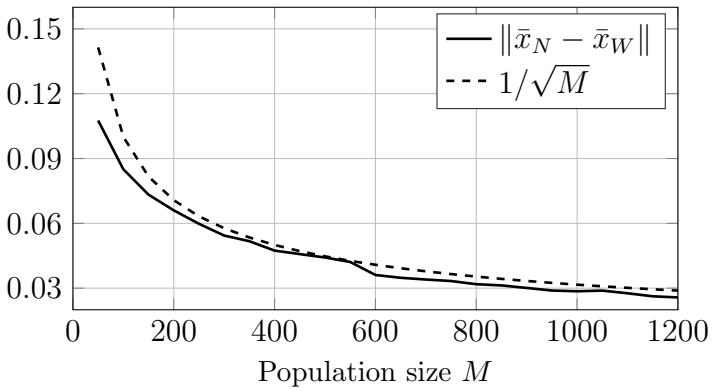

Figure 7.8: Distance between Nash and Wardrop variational equilibria. The quantity $1/\sqrt{M}$ illustrates the trend of the bound in Corollary 2 and not the specific constant.

7.3 Appendix

7.3.1 Proofs of the results presented in Sections 7.1 and 7.2

Proof of Corollary 1

Proof.

(a) First, we show that Assumption 1 holds. Indeed the sets $\mathcal{X}^i$ in (7.1) are convex and compact, the function g in (4.1) is affine and hence convex, and $\mathcal{Q}$ is non-empty by assumption. For each z fixed, the function $J^i(x^i, z)$ is linear hence convex in x^i. We prove in the last statement that F_{N} is strongly monotone. This is equivalent to $\nabla_x F_{\mathrm{N}}(x) \succ 0$ by Lemma 2, which by definition of $F_{\mathrm{N}}(x)$ implies $\nabla_{x^i}(\nabla_{x^i} J^i(x^i, \sigma(x))) \succ 0$, which implies convexity of $J^i(x^i, \sigma(x))$. Finally, $J^i(z_1, z_2)$ is continuously differentiable in $[z_1; z_2]$ because p_t is twice continuously differentiable. Having verified Assumption 1, Lemma 1 guarantees the existence of a Nash and of a Wardrop equilibrium. The ε-Nash property is guaranteed by Proposition 7 upon verifying Assumption 2. This holds because: i) $\cup_{i=1}^M \mathcal{X}^i \subseteq [0, \tilde{x}^0]^n$, ii) $J^i(z_1, z_2)$ is Lipschitz in z_2 on $[0, \tilde{x}^0]^n$ with Lipschitz constant $L_2 = RL_p$, iii) (5.1) holds and iv) p_t is assumed Lipschitz in $[0, \tilde{x}^0]$ with Lipschitz constant L_p for all t. We conclude by noting that $R = \tilde{x}^0\sqrt{n}$.

(b) The fact that each p_t is strictly increasing in $[0, \tilde{x}^0]$ implies that $\nabla_z p(z) \succ 0$ in $[0, \tilde{x}^0]^n$, where $p(z) \coloneqq \left[p_1(\frac{d_1+z_1}{\kappa}), \ldots, p_n(\frac{d_n+z_n}{\kappa})\right]^\top$. In turn $\nabla_z p(z) \succ 0$ guarantees strong monotonicity of p in $[0, \tilde{x}^0]^n$ by Lemma 2. This, together with Assumption 1 and Assumption 2 verified above, allows us to use the third result in Theorem 1.

(c) Given the special form of the sets $\{\mathcal{X}^i\}_{i=1}^M$ and the price $p_t\left(\frac{d_t+\sigma_t(x)}{\kappa_t}\right)$, Assumptions 4, 5 and 6 are satisfied. In addition since $\sum_{i=1}^M \theta^i > 0$, it must be that $J_{\mathrm{S}}(\sigma(x_{\mathrm{S}})) > \hat{J}$ for some $\hat{J} \geq 0$. Thus, the assumptions of Theorem 3.

(d) The strong monotonicity of F_{N} follows immediately thanks to Lemma 5. Additionally, Assumption 1 holds as previously shown. Since Assumption 8 holds, we can directly employ Theorem 6 and conclude the proof.

□

Proof of Lemma 10

Proof. The constraints in (7.1), (7.2) can be expressed as $\Gamma x \leq \gamma$ with

$$\Gamma = \begin{bmatrix} I_{Mn} \\ -I_{Mn} \\ -I_M \otimes \mathbb{1}_n^\top \\ \mathbb{1}_M^\top \otimes I_n \end{bmatrix}, \quad \gamma = \begin{bmatrix} \tilde{x} \\ 0 \\ -\theta \\ MK \end{bmatrix},$$

where $\theta = [\theta^1, \dots, \theta^M]^\top$, and $\tilde{x} = [[\tilde{x}_t^i]_{t=1}^n]_{i=1}^M$. Let us partition the constraint matrix Γ into its individual part Γ_1 and coupling part Γ_2

$$\Gamma = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}, \ \Gamma_1 = \begin{bmatrix} I_{Mn} \\ -I_{Mn} \\ -I_M \otimes \mathbb{1}_n^\top \end{bmatrix}, \ \Gamma_2 = \begin{bmatrix} \mathbb{1}_M^\top \otimes I_n \end{bmatrix} \tag{7.12}$$

and $\gamma = [\gamma_1^\top, \gamma_2^\top]^\top$ accordingly. The KKT conditions for $\mathrm{VI}(\mathcal{Q}, F_\mathrm{N})$ at the primal solution $\bar{x}_\mathrm{N}$ are [FP07, Prop. 1.3.4]

$$\begin{aligned} &F_\mathrm{N}(\bar{x}_\mathrm{N}) + \Gamma_1^\top \mu + \Gamma_2^\top \lambda = 0, \\ &0 \leq \mu \perp \gamma_1 - \Gamma_1 \bar{x}_\mathrm{N} \geq 0, && \text{(7.13a)} \\ &0 \leq \lambda \perp \gamma_2 - \Gamma_2 \bar{x}_\mathrm{N} \geq 0. && \text{(7.13b)} \end{aligned}$$

Define $\tilde{\mu}$ and $\tilde{\lambda}$ as the dual variables corresponding to the active constraints (the other dual variables must be zero due to (7.13a) and (7.13b)). The KKT system (7.13) in $\tilde{\mu}, \tilde{\lambda}$ only reads

$$\begin{aligned} &\tilde{\Gamma}_1^\top \tilde{\mu} + \tilde{\Gamma}_2^\top \tilde{\lambda} = -F_\mathrm{N}(\bar{x}_\mathrm{N}), \\ &\tilde{\mu}, \tilde{\lambda} \geq 0\,, \end{aligned} \tag{7.14}$$

where $\tilde{\Gamma}_1, \tilde{\Gamma}_2$ contain the subset of rows of Γ_1, Γ_2 corresponding to active constraints. To conclude the proof we need to show that (7.14) has a unique solution $\tilde{\lambda}$. To this end we apply the subsequent Lemma 11. To verify its assumption, we note that its negation is equivalent, given the expressions of $\tilde{\Gamma}_1, \tilde{\Gamma}_2$ in (7.12), to the existence of $R' \subseteq R^\mathrm{tight}$ such that for each vehicle i it holds $\bar{x}_{\mathrm{N},t}^i \in \{0, \tilde{x}_r^i\}$ for all $t \in R'$ or $\bar{x}_{\mathrm{N},t}^i \in \{0, \tilde{x}_t^i\}$ for $t \in \{1, \dots, n\} \setminus R'$ and such R' cannot exist by assumption. □

Lemma 11. *Consider $A_1 \in \mathbb{R}^{m \times n_1}$, $A_2 \in \mathbb{R}^{m \times n_2}$, $b \in \mathbb{R}^m$. If the implication $A_1 x_1 + A_2 x_2 = 0 \ \Rightarrow \ x_1 = 0$ holds, then the linear system of equations $A_1 x_1 + A_2 x_2 = b$ has at most one solution in x_1.*

Proof. Assume $A\tilde{x} = b$ and $A\hat{x} = b$, then $A_1\tilde{x}_1 + A_2\tilde{x}_2 = b$ and $A_1\hat{x}_1 + A_2\hat{x}_2 = b$ imply $A_1(\hat{x}_1 - \tilde{x}_1) + A_2(\hat{x}_2 - \tilde{x}_2) = 0$, which by assumption implies $\hat{x}_1 = \tilde{x}_1$. □

Proof of Corollary 2

Proof.

(a) Satisfaction of Assumption 1 and the consequent existence of a variational Nash and of a variational Wardrop equilibrium for any M can be shown as in Corollary 1. The operator F_{W} for the cost (7.7) reads

$$F_{\mathrm{W}}(x) = [\gamma^i(x^i - \hat{x}^i) + t(\sigma(x))]_{i=1}^{M}.$$

where $t(\sigma(x)) \coloneqq [t_e(\sigma_e(x_e))]_{e=1}^{E}$. Since $t_e(\sigma_e(x_e))$ in (7.10) is a monotone function of $\sigma_e(x_e)$, the operator $t(\sigma(x))$ is monotone. Then F_{W} is strongly monotone with constant $\hat{\gamma}$ because it is the sum of a monotone and a strongly monotone operator with constant $\hat{\gamma}$. As a consequence, each $\mathcal{G}_M^{\mathrm{RC}}$ admits a unique variational Wardrop equilibrium.

To prove strong monotonicity of F_{N} we use the result of Lemma 2.[5] We first note that each t_e only depends on the corresponding σ_e, hence $\nabla_x F_{\mathrm{N}}(x)$ can be permuted into diagonal form similarly to what done in (4.21). It then suffices to show $\hat{\gamma} I_M + \frac{1}{M} t_e'(\sigma_e) I_M + \frac{1}{M^2} t_e''(\sigma_e) x_e \mathbb{1}_M^\top \succ 0$ for all σ_e and for all e. This matrix is indeed positive definite if $\sigma_e(x_e) \notin [f_e h - \Delta_e, f_e h + \Delta_e]$, because then $t_e'(\sigma_e) \geq 0$ and $t_e''(\sigma_e) = 0$ by (7.10). For $\sigma_e(x_e) \in [f_e h - \Delta_e, f_e h + \Delta_e]$ it suffices to show $\hat{\gamma} I_M + \frac{1}{M^2 4 f_e \Delta_e} x_e \mathbb{1}_M^\top \succ 0$, because $t_e'(\sigma_e) \geq 0$ and $t_e''(\sigma_e) = \frac{1}{4 f_e \Delta_e}$. By Lemma 6, $\lambda_{\min}\left(x_e \mathbb{1}_M^\top + \mathbb{1}_M x_e^\top\right)/2 \geq -\frac{M}{8}$, which proves strong monotonicity of F_{N} under (7.11). Consequently, if M satisfies (7.11) then $\mathcal{G}_M^{\mathrm{RC}}$ admits a unique variational Nash equilibrium. Finally, we verify Assumption 2 in order to use Proposition 7. We have $\mathcal{X}^0 = [0,1]^E$ and t is continuously differentiable and hence Lipschitz in $\mathcal{X}^0$, with constant $L_p = 1/(2 f_{\min})$. Moreover, $R \coloneqq \max_{y \in \mathcal{X}^0}\{\|y\|\} = \sqrt{E}$. Using (5.1) concludes the proof.

(b) Since all the assumptions of Theorem 1 have just been verified, it is a direct consequence of its second statement.

(c) As Assumption 8 holds trivially (the others have already been verified), we apply Theorem 6 and conclude the proof.

□

[5] Lemma 2 requires F_{N} to be continuously differentiable, which is not the case here. The more general result [Sch+96, Prop. 2.1] extends the statement of Lemma 2 to operators which are not continuously differentiable. It then suffices to show $\nabla_x F_{\mathrm{N}}(x) \succ 0$ for $\sigma(x)$ in each of the three intervals defined by (7.10), because in each of them F_{N} is continuously differentiable.

Part II

Programmable machines: game design

CHAPTER 8

Introduction

In this part of the thesis we focus on large scale cooperative systems composed of programmable machines, which we refer as multiagent systems. As discussed in the overview of Chapter 1, one of the main challenges in the control of large scale cooperative systems rests in the design of control algorithms that achieve a given global objective by relying solely on local information. The problem of designing local control algorithms is typically posed as an optimization problem (finite or infinite dimensional), where the system-level objective is captured by an *objective function* (functional), while physical laws and informational availability are incorporated as *constraints* on the decision variables. The design is complete once a distributed decision making algorithm has been found, satisfying the constraints and maximizing the objective function [DD03; Cor+02]. A well-established approach to tackle this problem consists in the design of a centralized maximization algorithm, that is later distributed by leveraging the structure of the problem considered. Examples in continuous optimization include algorithms such as distributed gradient ascent, primal-dual and Newton's method [BT89; NO09; WOJ13]. While the existing approaches, including the above-mentioned one, have produced a variety of algorithms for the control of distributed systems, the design question has not been entirely solved. A perspective article recently appearing in Science Robotics summarizes the difficulties: "*There are currently no systematic approaches for designing such multidimensional feedback loops*" [Yan+18, p. 6].

8.1 The game-design framework

A promising approach, termed *game design*, has recently emerged as a tool to complement the partial understanding offered by more traditional techniques [Sha07]. The game design approach is tightly connected with the notion of equilibrium in game theory, and its origin stems from a novel engineering perspective on this field. While game theory has originated as a set of tools to model the interaction of multiple decision makers (players) [VM07], its relevance to distributed control stems from the observation that players of a game are required to take local decisions based on partial information of the entire

system. Motivated by this consideration, the seminal works of [MS07; AMS07] proposed the use of game theoretic tools to tackle distributed optimization problems arising in the area of multi-agent systems. Rather than using game theory to describe existing interactions, [AMS07] suggested a *paradigm shift* and proposed the use of game theory to *design* control architectures with the aim of meeting a given system level objective.

In lieu of directly specifying a decision-making process, the game design approach consists in assigning local utility functions to the agents, so that their selfish maximization translates in the achievement of the system level objective. The potential of this technique stems from the possibility to inherit a pool of algorithms from the literature of learning in games [Blu93; FL98; MS12; YP17] that are *distributed* by nature, *asynchronous*, and *resilient* to external disturbance [AMS07].

Given an optimization problem we wish to solve distributedly, the *game design* procedure proposed in [Sha07; MW13] is summarized in Figure 8.1 and consists in the following steps:[1]

1) **Utility design**: assign utility functions to each agent and an equilibrium concept for the corresponding game.

2) **Algorithm design**: devise a distributed algorithm to guide agents to the chosen equilibrium concept.

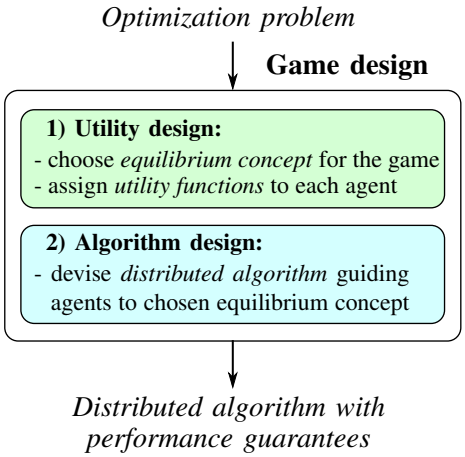

Figure 8.1: Game theoretic approach for the design of distributed control systems.

The objective of the game design procedure is to obtain an efficient and distributed algorithm for the solution of the original optimization problem. While the introduction of an auxiliary equilibrium problem might seem artificial at first, this approach has recently produced a host of new results [Gai09; MAS09; RPM17; Geb+16]. Observe that, in order

[1]In the following, we identify the agents of the optimization problem and their local constraint sets with the players of the game and their action sets.

for the game design procedure to be relevant to the original optimization problem, the utility functions need to be carefully designed so that the chosen equilibrium (equilibria) coincide with the global optimizer(s) of the original problem, or is provably close to.

Within the boundaries of the game design procedure discussed above, it is important to highlight that agents are *not* modeled as competing units, but the system operator is rather designing their utilities to distribute the global objective. For this purpose, agents are considered as purely programmable machines endowed with computational and communication capabilities. Game theory represents, in this context, a mere set of tools that can be exploited to derive distributed algorithms with provable performance certificates, and *not* a modeling language describing the behaviour of egoistic agents.

While the field of learning in games offers *readily available algorithms* to coordinate agents towards an equilibrium in a distributed fashion (i.e., it addresses the second step of Figure 8.1), the *utility design* problem is much less tracked.

The **goal** of Part II of this thesis is to provide a framework to compute the equilibrium efficiency as a function of the given utility functions, and to optimally select utilities so as to maximize such efficiency.

8.2 The general multiagent maximum coverage

In this section we introduce the problem considered in Part II of this thesis.

Consider $\mathcal{R} = \{r_1, \dots, r_m\}$ a finite set of resources, where each resource $r \in \mathcal{R}$ is associated with a value $v_r \geq 0$ describing its importance. Further let $N = \{1, \dots, n\}$ be a finite set of agents. Every agent $i \in N$ selects a_i, a subset of the resources, from the given collection $\mathcal{A}_i \subseteq 2^{\mathcal{R}}$, i.e., $a_i \in \mathcal{A}_i$. The welfare of an allocation $a = (a_1, \dots, a_n) \in \mathcal{A} \coloneqq \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ is given by

$$W(a) \coloneqq \sum_{r \in \cup_{i \in N} a_i} v_r w(|a|_r), \tag{8.1}$$

where $W : 2^{\mathcal{R}} \times \cdots \times 2^{\mathcal{R}} \to \mathbb{R}$, $|a|_r = |\{i \in N \,\text{s.t.}\, r \in a_i\}|$ captures the number of agents selecting resource r in allocation a, and $w : [n] \to \mathbb{R}_{\geq 0}$ is called the welfare basis function. Informally, w scales the value of each resource depending on how many agents selected it. The goal is to find a feasible allocation maximizing the welfare, i.e.,

$$a^{\text{opt}} \in \arg\max_{a \in \mathcal{A}} W(a)\,. \tag{8.2}$$

We refer to the above problem as to the *general multiagent weighted maximum coverage* (GMMC) problem, due to its connections with coverage problems as discussed in the forthcoming Section 8.3.

Observe that we have not posed any constraint on the structure of the sets $\{\mathcal{A}_i\}_{i\in N}$ in the sense that they are not required to represent matroid constraints, knapsack constraints, etc. At this stage, they are a mere collection of subsets of $\mathcal{R}$.

Since the problem in (8.2) is $\mathcal{NP}$-hard (see the discussion in Section 8.3), we seek an *efficient* algorithm (or a class of algorithms) to determine an approximate solution, ideally with the best possible approximation ratio. Additionally, we request the algorithm to be distributed as detailed in Section 10.1. We pursue this goal by means of the game theoretic approach previously introduced.

8.2.1 Applications

In the following we discuss two classes of problems that can be solved by using the techniques discussed in the second part of this thesis.

Multiagent task assignment problems

In a multiagent task assignment problem we are given a list of tasks to be performed, as well as a list of agents. The goal is to match agents and tasks so as to maximize a given welfare function representing the quality of the matching. Such function is typically additive over the tasks and some of the tasks may require a minimum number of agents to be completed. It is typically assumed that the more agents participate in the execution of a task, the higher the welfare generated from that task, and that the problem exhibits diminishing returns. Practical examples of problems belonging to this class include vehicle-target assignment [Mur00; AMS07], sensor deployment [CL05; MAS09], satellite assignment [QBL15] problems.

Distributed maximum coverage

In a distributed maximum coverage problem we are given a list of resources with their respective value and a list of agents. Each agent has access to a collection of subsets of the resources, while different agents typically have access to different collections (due to, e.g., geographical or other limitations). The goal is to allocate the agents so as to maximize the total value of covered resources. A large number of problems can be cast into this framework. Examples include staff scheduling [Ern+04], facility location [Far+12] and wireless scheduling [CK08] (see [Hoc97] for an overview of the applications). More recent applications include, among others, distributed caching in wireless networks [Goe+06; De +17], multi-topic searches [SG09], influence maximization [Kar+17], vehicle scheduling in mobility-on-demand platforms [SC17; Aga+12].

8.3 Related work

The work presented in Part II of this thesis is multidisciplinary in that it sits at the interface between approximation theory, distributed optimization and game theory. In the followings we present the most relevant connections to each of these areas and corresponding related works.

Maximum coverage and approximation guarantees

The general multiagent maximum coverage (GMMC) problem defined in Section 8.2 and studied in Part II of this thesis is tightly connected with the maximum coverage problem defined in [Fei98].

In a maximum coverage problem we are given a ground set of elements, and a collection of subsets of the ground set. The objective is to select n subsets from the collection, so as to maximize the total number of covered elements. The greedy algorithm achieves a $1-1/e$ approximation in polynomial time, and no polynomial algorithm can approximate the solution within any ratio better than $1-1/e+\epsilon$ (for all $\epsilon > 0$) unless $\mathcal{P} = \mathcal{NP}$, [Fei98]. This inapproximability result applies to all extensions discussed next (including the GMMC problem), since they hold the maximum coverage problem as a special case.

A generalization of the maximum coverage problem is the weighted maximum coverage problem. In a weighted maximum coverage we are given a ground set of elements, and a collection of subsets of the ground set. Every element in the ground set is given a weight. The goal is to select n subsets from the collection in order to maximize the total weight of covered elements. The greedy algorithm gives the best possible polynomial approximation ratio of $1-1/e$. The proof is no longer based on the result of [Fei98], but on the more general result in submodular maximization subject to cardinality constraints [NWF78].

Algorithms based on a continuous relaxation of the previous problems are also available. In particular the result in [Cal+11] applies to the problem of monotone submodular maximization subject to matroid constraints, and thus provides a solution for the weighted maximum coverage. The algorithm of [Cal+11] computes a non integer solution, which is then rounded using the pipage algorithm producing a $1-1/e$ approximation. Relative to the problem of monotone submodular maximization over a matroid constraint, a more refined result is available when the objective function has known (total) curvature c. The notion of curvature has been introduced in [CC84] and describes how far a given function is from being modular. In this case, [SVW17] has recently provided a $1-c/e$ approximation and has showed that no polynomial time algorithm can give a better approximation. The latter work improves upon the $(1-e^{-c})/c$ of [CC84]. Observe that the maximum coverage problem has $c = 1$, so that [SVW17]

matches [Fei98].

Multiagent versions of the weighted maximum covering problem have been introduced independently in [CK04] with the name of *maximum coverage problem with group budget constraints* and in [Gai09] with the name of *general covering problems*. This class of problems subsumes the previous ones; we refer to it as to the class of *multiagent weighted maximum coverage* (MMC) problems. In a MMC problem we are given not one, but n collections of subsets. The objective is to select one set from each collection so as to maximize the total weight of covered elements. Relative to MMC problems, the greedy algorithm provides a $1/2$ approximation [CK04], and the local search algorithm proposed in [Gai09] achieves the optimal $1 - 1/e$, under technical assumptions.

The GMMC problem studied in Part II of this thesis is a generalization of the MMC problem in that we allow for a function w to rescale the weight of each element depending on how many agents cover such element. Any MMC problem can be recovered by the corresponding GMMC problem by setting $w(j) = 1$ in (8.1). Any weighted maximum coverage problem can be recovered from a GMMC problem, upon setting $w(j) = 1$ in (8.1) and $\mathcal{A}_i = \mathcal{A}_j$ for all i, j. Further classes of problems such as the multiple-choice knapsack problem or the standard knapsack problem [Pis95] can be obtained from the MMC problem (and thus from the GMMC problem). The former problem can be recovered assuming $\mathcal{A}_i$ to represent knapsack constraints. The latter problem is obtained by additionally imposing $\mathcal{A}_i = \mathcal{A}_j$ for all i, j. Observe that when the welfare basis w is increasing and concave (in the discrete sense), the welfare function W defined in (8.1) is monotone submodular. Submodular functions are subject of intense study due to their ability to model engineering problems that feature diminishing returns. Similarly, if w is increasing and convex, W is monotone supermodular. Figure 8.2 summarizes the main classes of problems discussed.

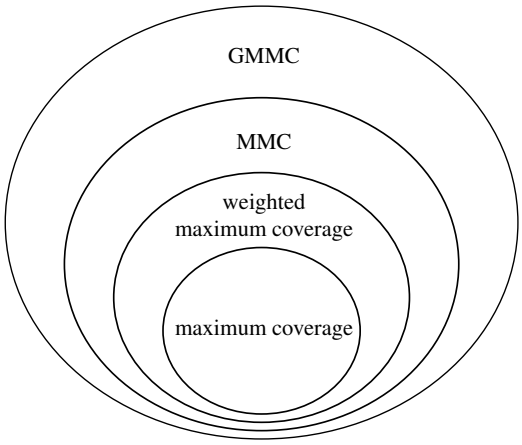

Figure 8.2: Classes of problems discussed in Section 8.3.

Distributed combinatorial optimization

While distributed algorithms have been studied since the early nineties in the context of continuous (and convex) optimization [BT89], the interest in their combinatorial counterpart is more recent.

Particular attention has been devoted to the problem of maximizing a submodular function subject to various form of constraints such as cardinality, matroid or knapsack constraints. This is due to the potential applications of submodular maximization in different fields featuring "large-scale" systems. A non-exhaustive list includes sensor allocation [SCL16; KSG08], data summarization [Mir+16], task-assignment problems [QBL15]. While centralized algorithms are available to produce good approximations (e.g., the greedy algorithm and its variations [Fei98]), their sequential implementation makes them unsuited for parallel and distributed execution. In this respect, there has been recent effort in distributing such algorithms using the so called *MapReduce* programming approach [DG08]. In [Mir+16; Bar+15] and references therein, the authors propose to divide the original optimization problem into smaller parts and to solve each of them on a different machine. The solution is determined by patching together the partial results and is certified to achieve a competitive approximation ratio. Nevertheless, the approach still requires a central coordinator.

Other classes of combinatorial problems for which distributed algorithms have been recently proposed include graph coloring, maximum coverage, and multiple-choice knapsack [BE13; Gai09; MYR17]. Finally, [MHK18] provides distributed algorithms for submodular maximization problems, but admissible objective functions are required to be the sum of agents' individual contributions (unlike here).

Game design and utility design approach

The problem of designing local utility functions so as to maximize the efficiency of the emerging equilibria find its roots in the economic literature relative to the design of optimal taxations [Ram27]. The approach has been applied to the design of engineering systems only recently. More in details, the use of game theoretic learning algorithms for the distributed solution of optimization problems has been proposed in [AMS07], and since then a number of works have followed a similar approach [SSR09; Gai09; Cha+11; SWL11]. We redirect the reader to [MS18] for a general overview on equilibrium learning algorithms in distributed control. What has been less understood so far, is how to provide performance certificates for a given set of utility functions, and more fundamental how to select utility functions so as to maximize such performance certificates.

The performance degradation of an equilibrium allocation compared to the optimal solution has been subject of intense research in the field of algorithmic game theory (through the notions of price of anarchy, price of stability [KP99; SM03]). Nevertheless,

the results available therein are not helpful for the design problem studied here. The widely used smoothness framework proposed in [Rou09] has brought a number of different results under a common language and has produced tightness guarantees for different problems [Rou09; RST17]. Unfortunately the latter framework requires the sum of the utility functions to be equal (or less equal) to the welfare function (budget-balance condition). While this assumption is well justified for a number of problems modeled through game theory (e.g., cost sharing games [MS01]), it has little bearing on the design of local utility functions studied here.

The utility design problem considered here has been addressed limitedly to specific applications, e.g., concave cost sharing, reverse carpooling problems [MP17; ME12] or confined to particular design methodologies such as the Shapley value or marginal contribution [MR14; PSM16].

CHAPTER 9

Mathematical preliminaries

In this chapter we introduce the mathematical tools required to move forward and present the results of Chapter 10 and Chapter 11.

9.1 Strategic-form games and equilibrium concepts

Definition 9 (Strategic-form game)**.** *A strategic form game* $G = (N, \{\mathcal{A}_i\}_{i=1}^N, \{u_i\}_{i=1}^N)$ *is a tuple where* $N = \{1, \dots, n\}$ *is a finite set of players,* $\mathcal{A}_i$ *is the action set of player* $i \in N$*, and* $u_i : \mathcal{A} \to \mathbb{R}$ *is the utility function of player* $i \in N$*, where* $\mathcal{A} \coloneqq \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$*. A strategic-form game is called finite if the set* $\mathcal{A}$ *is finite.*

Informally, a game is fully specified in its strategic form if every player is given an action set and a utility function depending on the choice of all the players. We refer to $a \coloneqq (a_1, \dots, a_n) \in \mathcal{A}$ as to an allocation. We will often represent an allocation as $a = (a_i, a_{-i})$, where $a_{-i} \coloneqq (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$ denotes the allocations of all players but $i \in N$.

In the following we consider strategic-form games only. We do not repeat this in the forthcoming statements for ease of presentation.

Definition 10 (*Nash equilibrium* (NE), [Nas50])**.** *A feasible allocation* $a^{\text{ne}} \in \mathcal{A}$ *is a pure Nash equilibrium for the game* G*, if no player can increase his utility function by unilaterally deviating from his equilibrium allocation, i.e., if*

$$u_i(a^{\text{ne}}) \geq u_i(a_i, a_{-i}^{\text{ne}}) \qquad \forall a_i \in \mathcal{A}_i, \quad \forall i \in N.$$

We denote with $\text{ne}(G)$ *the set of pure Nash equilibria of* G*.*

In the remaining of this thesis we will refer to a pure Nash equilibrium as just a Nash equilibrium, if no confusion arises. It is not difficult to show that Nash equilibria may not exist. This and many other reasons motivate the definition of mixed Nash equilibria. Towards this goal, we first introduce the concept of *mixed strategy*. A mixed strategy σ_i

is a probability distribution over the action space of player i, i.e., $\sigma_i \in \Delta(\mathcal{A}_i)$. A mixed strategy profile $\sigma := (\sigma_1, \dots, \sigma_n)$ is a distribution $\sigma \in \Sigma := \times_{i \in N} \Delta(\mathcal{A}_i)$.

Definition 11 (*Mixed Nash equilibrium* (MNE), [Nas50])**.** *A mixed strategy profile $\sigma^{\text{mne}} \in \Sigma$ is a mixed Nash equilibrium for the game G if no player can increase his expected utility by deviating to a pure strategy, i.e., if*

$$\mathbb{E}_{a \sim \sigma^{\text{mne}}} [u_i(a)] \geq \mathbb{E}_{a \sim \sigma^{\text{mne}}} [u_i(a'_i, a_{-i})] \qquad \forall a'_i \in \mathcal{A}_i, \quad \forall i \in N.$$

We denote with $\text{mne}(G)$ *the set of mixed Nash equilibria of G.*

In the previous definition player i compares $\mathbb{E}_{a \sim \sigma^{\text{mne}}} [u_i(a)]$ with the expected value of his utility when he deviates and selects the *pure* strategy a'_i. It is possible to show that this is equivalent to requiring player i not to improve even if selecting a *mixed strategy* $\sigma_i \in \Delta(\mathcal{A}_i)$. Thus, an equivalent definition could be given with respect to deviations in mixed strategies. Additionally, observe that the set of mixed Nash equilibria contains the set of pure Nash equilibria.

Mixed Nash equilibria are guaranteed to exist in any game where the actions sets are finite, as shown in the celebrated paper by John Nash, [Nas50].

Proposition 9 (Existence of MNE, [Nas50])**.** *Any finite game admits a MNE.*

Despite the fact that existence of mixed Nash equilibria is guaranteed, the problem of computing a MNE is, in general, intractable [DGP06]. For this reason, we consider the following enlarged class of equilibria.

Definition 12 (*Coarse correlated equilibrium* (CCE), [MV78])**.** *A probability distribution $\sigma^{\text{cce}} \in \Delta(\mathcal{A})$ is a coarse correlated equilibrium for the game G if no player can increase his expected utility by deviating to a pure strategy, i.e., if*

$$\mathbb{E}_{a \sim \sigma^{\text{cce}}} [u_i(a)] \geq \mathbb{E}_{a \sim \sigma^{\text{cce}}} [u_i(a'_i, a_{-i})] \qquad \forall a'_i \in \mathcal{A}_i, \quad \forall i \in N. \tag{9.1}$$

We denote with $\text{cce}(G)$ *the set of CCE of G.*

The only difference in the definitions of mixed Nash equilibrium and corse correlated equilibrium is in that σ^{mne} is required to be a product distribution $\sigma^{\text{mne}} \in \times_{i \in N} \Delta(\mathcal{A}_i)$, while $\sigma^{\text{cce}} \in \Delta(\mathcal{A})$ is not. It follows that the set of coarse correlated equilibria is a superset of the set of mixed Nash equilibria. The interest in CCE stems from the fact that, unlike MNE and NE, they are computationally tractable [LW94; Nis+07]. We will return to this in Section 9.4.

We conclude introducing the last equilibrium concept. To do so, we first consider a welfare function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$ and define the allocation $a^{\text{opt}} \in \mathcal{A}$ as an allocation such that $W(a^{\text{opt}}) \geq W(a)$ for all $a \in \mathcal{A}$.

Definition 13 (*Average coarse correlated equilibrium* (ACCE$^{\text{opt}}$),[NR10])**.** *Given a game G and a function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$, a probability distribution $\sigma^{\text{acce}} \in \Delta(\mathcal{A})$ is an average coarse correlated equilibrium with respect to the allocation $a^{\text{opt}} \in \mathcal{A}$ if*

$$\mathbb{E}_{a\sim\sigma^{\text{acce}}}\left[\sum_i u_i(a)\right] \geq \mathbb{E}_{a\sim\sigma^{\text{acce}}}\left[\sum_i u_i(a_i^{\text{opt}}, a_{-i})\right].$$

We denote with $\text{acce}^{\text{opt}}(G)$ *the set of ACCE$^{\text{opt}}$ of G.*

Average coarse correlated equilibria are a superset of coarse correlated equilibria. This is because the previous condition can be obtained from Definition 12 by summing the condition (9.1) over all players, and selecting $a' = a^{\text{opt}}$.

It follows that the equilibrium sets previously defined are all nested

$$\text{ne}(G) \subseteq \text{mne}(G) \subseteq \text{cce}(G) \subseteq \text{acce}^{\text{opt}}(G).$$

While NE, MNE, CCE are well studied and regularly used equilibrium concepts, the notion of ACCE$^{\text{opt}}$ is rather novel. The latter equilibrium concept will be used here as a purely conceptual tool in connection with the study of equilibrium efficiency, see Section 9.3.1.

9.2 Potential games and congestion games

In the previous section we have introduced three fundamental equilibrium concepts. Additionally, we have commented on their existence and on their tractability (or lack thereof). In this section we refine the analysis to potential games and congestion games.

Potential games

Definition 14 (Potential game, [MS96])**.** *A strategic-form game is a potential game if there exists a function $\varphi : \mathcal{A} \to \mathbb{R}$ such that*

$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = \varphi(a_i, a_{-i}) - \varphi(a_i', a_{-i}) \qquad \forall a \in \mathcal{A}, \quad \forall a_i' \in \mathcal{A}_i, \quad \forall i \in N.$$

The function φ is called potential function.

Informally, a game is potential if the variation in each player's utility experienced when deviating from a_i to a_i' can be captured by the function φ, and such function is the *same* for all the players $i \in N$. The condition is reminiscent of the notion of conservative force and corresponding potential field taken form Physics. Indeed, for games with continuous action space, the two notions coincide.

An immediate consequence of the previous definition is the existence of a pure NE.

Proposition 10 (Existence of pure NE in potential games, [MS96])**.** *Any finite potential game admits a pure Nash equilibrium.*

Proof. Consider $a^\star \in \mathcal{A}$ a global maximizer of the potential. Since the actions sets are finite, $a^\star$ is guaranteed to exist. By definition of maximizer and of potential game, it is

$$u_i(a_i^\star, a_{-i}^\star) - u_i(a_i', a_{-i}^\star) = \varphi(a_i^\star, a_{-i}^\star) - \varphi(a_i', a_{-i}^\star) \geq 0 \qquad \forall a_i' \in \mathcal{A}_i, \quad \forall i \in N\,.$$

Thus, $a^\star$ is a pure Nash equilibrium. □

It can be similarly shown that any local maximizer of the potential function φ is a pure Nash equilibrium.

Definition 15. *(Local maximizer) Given a function $\varphi : \mathcal{A} \to \mathbb{R}$ with $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$, an allocation $a^\star \in \mathcal{A}$ is a local maximizer of φ if $\varphi(a^\star) \geq \varphi(a_i', a_{-i}^\star)$ for all $a_i' \in \mathcal{A}_i$ and for all $i \in N$.*

The previous observation builds a *fundamental bridge* between optimization problems and equilibrium problems as it suggests a seemingly simple technique to compute a NE for the class of potential games: determine a local maximizer of the potential function. Additionally, it suggests a natural dynamics to compute one such equilibrium.

Definition 16 (*Best-response dynamics* (BR))**.** *Let $t \in \mathbb{N}_0$ indicate the time step of the algorithm and $a^t \in \mathcal{A}$ the corresponding allocation. The best-response dynamics is presented in Algorithm 5. Ties are broken according to a pre-specified rule (any rule).*

In the best-response dynamics, players take ordered turns and update their choice by selecting their best action, given the current actions of the others. While the BR dynamics is not guaranteed to converge for a general game, this is the case if we restrict to the class of potential games.

Algorithm 5 Best-response dynamics (round-robin)

Initialise $a^0 \in \mathcal{A}$; $t \leftarrow 0$
while not converged **do**
▷ Best response
  $i \leftarrow (t \mod n) + 1$
  $a_i^{t+1} \leftarrow \arg\max_{a_i \in \mathcal{A}_i} u_i(a_i, a_{-i}^t)$
  $a^{t+1} \leftarrow (a_i^{t+1}, a_{-i}^t)$
  $t \leftarrow t + 1$
end while

Proposition 11 (Convergence of the BR dynamics, [MS96])**.** *The best-response dynamics converges, for any initial condition $a^0 \in \mathcal{A}$, to a NE in a finite number of steps, for any potential and finite game.*

Proof. The proof is based on the use of the potential function as a Lyapunov function. After every round of the BR dynamics either no player improved his utility, in which case we are at a Nash equilibrium, or at least the utility of one player has increased. In the latter case, φ has increased too. Since φ is upper bounded by its maximum value, the BR dynamics must converge. Additionally, since φ strictly increases in every round, the best-response dynamics can not return to an allocation visited in the past. Thus, convergence in a finite number of steps follows by the finiteness of $\mathcal{A}$. □

Three important comments follow. First, we considered here a round-robin best response algorithm, i.e., an algorithm where the players revise their decision in a given order. Similar statements to those in Proposition 11 can be made *almost surely* if the players updating their decision are uniformly randomly selected. This will produce a totally asynchronous algorithm. Second, note that the claim in Proposition 11 holds even if the players were to update their actions using a better-response dynamics, instead of a best-response dynamics. In the better-response dynamics, players update their previous choice by selecting an action that improves their utility, but need not be the best. Third, observe that the best-response dynamics (better-response dynamics) might be slow to converge, in that it could visit *all* the allocations in $\mathcal{A}$ before settling to a NE. Additionally, the task of finding a best-response (line 4 in Algorithm 5) might also be intractable. We return to this in Section 9.4.

Congestion games

Congestion games are defined as follows.

Definition 17 (Congestion game, [Ros73])**.** *Consider $\mathcal{R}$ a finite set of resources and for every resource $r \in \mathcal{R}$ a function $w_r : \mathbb{N} \to \mathbb{R}_{\geq 0}$. A congestion game is a normal-form finite game where $N = \{1, \dots, n\}$ is the set of players, $\mathcal{A}_i \subseteq 2^{\mathcal{R}}$ and $u_i(a) = \sum_{r \in a_i} w_r(|a|_r)$ are the action set and utility function of player i, respectively. The quantity $|a|_r$ represents the number of players selecting resource r in allocation a,*

$$|a|_r := \{i \in N \; s.t. \; r \in a_i\}$$

The next proposition shows that congestion games are a subclass of potential games. Thus, existence of a pure Nash equilibrium is guaranteed as well as convergence of the best-response dynamics, see Proposition 10 and Proposition 11.

Proposition 12 (Congestion games are potential, [Ros73])**.** *Congestion games are potential games with potential function φ given by*

$$\varphi(a) = \sum_{\substack{r \in \mathcal{R} \\ |a|_r \geq 1}} \sum_{j=1}^{|a|_r} w_r(|a|_r). \tag{9.2}$$

Thus, a pure Nash equilibrium is guaranteed to exist.

The potential function in (9.2) is often referred to as *Rosenthal's potential*.

9.3 Price of anarchy and smoothness

The notions of price of anarchy and price of stability have been introduced to quantify the efficiency of the equilibrium allocations with respect to centralized optimal allocations [KP99; SM03]. Let us consider a strategic-form game $G = (N, \{\mathcal{A}_i\}_{i=1}^N, \{u_i\}_{i=1}^N)$ and a corresponding welfare function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$. The function W measures the quality of a given allocation, and can be used to model the achievement of a global objective. The price of anarchy represents the ratio between the welfare at the worst performing equilibrium and the optimal welfare. Consequently, it provides a bound on the efficiency for *all* the equilibria. In the following we assume that $W(a^{\text{opt}}) > 0$ so that the notion of price of anarchy is well posed.

Definition 18 (*Price of anarchy* (PoA), [KP99])**.** *Consider the strategic-form game G and the welfare function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$.*

(a) The price of anarchy for the class of NE is defined as

$$\text{PoA}^{\text{ne}} := \min_{a \in \text{ne}(G)} \frac{W(a)}{W(a^{\text{opt}})}.$$

(b) The price of anarchy for the class of MNE is defined as

$$\text{PoA}^{\text{mne}} := \min_{\sigma \in \text{mne}(G)} \frac{\mathbb{E}_{a \sim \sigma}[W(a)]}{W(a^{\text{opt}})}.$$

Replacing the set $\text{mne}(G)$ *with* $\text{cce}(G)$ *or* $\text{acce}^{\text{opt}}(G)$*, one obtains the corresponding definitions for* PoA^{cce} *and* PoA^{acce}*.*

Observe that the expression $W(a^{\text{opt}})$ also depends on the game instance G considered but we do not indicate it explicitly, for ease of presentation. By definition, the price of anarchy is bounded between zero and one. The higher the price of anarchy, the

more efficient the worst performing equilibrium. Since $\text{ne}(G) \subseteq \text{mne}(G) \subseteq \text{cce}(G) \subseteq \text{acce}^{\text{opt}}(G)$ as seen in Section 9.1, it follows that

$$\text{PoA}^{\text{ne}} \geq \text{PoA}^{\text{mne}} \geq \text{PoA}^{\text{cce}} \geq \text{PoA}^{\text{acce}},$$

i.e., the efficiency degrades as we move to a richer class of equilibria.

As a prototypical example to clarify the concept of price of anarchy, consider that of a road traffic network where a large number of drivers traveling from a certain origin to their corresponding destination. If each driver was to minimize his own travel time, this will result in an equilibrium configuration such as the NE. Instead, if the system operator was to instruct the drivers on which route to take, he will try to minimize the total travel time, i.e., the sum over all the drivers' individual travel time. The price of anarchy precisely capture the ratio between these two quantities.

While we present results relative to welfare maximization problems, analogous definitions and claims are available in case of cost minimization.

Definition 19 (Smooth game, [Rou09])**.** *Consider the strategic-form game G and the welfare function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$. The pair (G, W) is (λ, μ)-smooth if for some $\lambda, \mu \geq 0$ it holds*

$$\sum_{i \in N} u_i(a'_i, a_{-i}) \geq \lambda W(a') - \mu W(a), \quad \forall a, a' \in \mathcal{A}. \tag{9.3}$$

The following proposition provides a lower bound on the ratio between the expected welfare at any CCE and the optimum, i.e., it gives a bound on the price of anarchy relative to the specific instance G considered.

Proposition 13 (PoA bound,[Rou09])**.** *Consider a (λ, μ)-smooth game with $\sum_{i \in N} u_i(a) \leq W(a)$ for all $a \in \mathcal{A}$. Then, for any coarse correlated equilibrium σ^{cce} of G it holds*

$$\frac{\mathbb{E}_{a \sim \sigma^{\text{cce}}}[W(a)]}{W(a^{\text{opt}})} \geq \frac{\lambda}{1+\mu}.$$

Proof. Consider σ^{cce} any CCE of G. Setting $a'_i = a_i^{\text{opt}}$ in Definition 19 it is

$$0 \leq \mathbb{E}_{a \sim \sigma^{\text{cce}}}[u_i(a)] - \mathbb{E}_{a \sim \sigma^{\text{cce}}}\left[u_i(a^{\text{opt}}, a_{-i})\right] \qquad \forall i \in N.$$

Summing over the agents one obtains

$$\begin{aligned}
0 &\leq \mathbb{E}_{a \sim \sigma^{\text{cce}}}\left[\sum_i u_i(a)\right] - \mathbb{E}_{a \sim \sigma^{\text{cce}}}\left[\sum_i u_i(a_i^{\text{opt}}, a_{-i})\right] \\
&\leq \mathbb{E}_{a \sim \sigma^{\text{cce}}}\left[\sum_i u_i(a)\right] - \lambda W(a^{\text{opt}}) + \mu \mathbb{E}_{a \sim \sigma^{\text{cce}}}[W(a)] \\
&\leq -\lambda W(a^{\text{opt}}) + (1+\mu)\mathbb{E}_{a \sim \sigma^{\text{cce}}}[W(a)],
\end{aligned}$$

where we used the linearity of the expectation, the definition of (λ, μ)-smooth game, and the assumption for which $\sum_{i \in N} u_i(a) \leq W(a)$. The claim follows from $W(a) \geq 0$

$$\frac{\mathbb{E}_{a \sim \sigma^{\text{cce}}}[W(a)]}{W(a^{\text{opt}})} \geq \frac{\lambda}{1+\mu}.$$

□

The smoothness framework has proved useful in bringing a number of different results under a common language and has produced tight bounds on the price of anarchy for different problems [Rou09; RST17]. Its strength amongst others, lies in the recipe it provides to obtain performance bounds for a large class of equilibria. Indeed, as seen in the previous section, pure NE (and MNE) are a subclass of CCE. Thus, for any pure Nash equilibrium a^{ne} of G it also holds

$$\frac{W(a^{\text{ne}})}{W(a^{\text{opt}})} \geq \frac{\lambda}{1+\mu}.$$

The proof presented in Proposition 13 shows that once a game has been shown to be (λ, μ)-smooth, the corresponding bound on the price of anarchy follows easily (only linearity of the expectation is additionally used). Thus, the main difficulty in proving bounds on the price of anarchy using a smoothness argument resides in proving the smoothness property itself, i.e., in selecting λ and μ so that (9.3) holds for all $a, a^\star \in \mathcal{A}$. These parameters have been determined for certain classes of games. A non-comprehensive list include scheduling games [CDT12], location and valid utility games [Vet02], affine congestion games [Rou09], first-price auctions [KZ12; ST13], second-price auctions [CKS08].

9.3.1 The question of tightness

An important question we discuss in this section is the capability of the smoothness framework to give good (ideally tight) bounds on the price of anarchy. Given a game G, we define the best bound on the price of anarchy attainable via a smoothness argument as

$$\begin{aligned} \text{SPoA} \coloneqq \sup_{\lambda, \mu \geq 0} \ & \frac{\lambda}{1+\mu} \\ & \text{s.t. } (\lambda, \mu) \text{ satisfy (9.3)}. \end{aligned}$$

Observe that SPoA $\leq$ PoA as Proposition 13 provides only a bound on the equilibrium efficiency. The next proposition shows that SPoA is tight for the class of ACCE$^{\text{opt}}$.

Proposition 14 (Smoothness is tight for ACCE$^{\text{opt}}$, [NR10])**.** *For any given game G it is*

$$\text{SPoA} = \text{PoA}^{\text{acce}} = \min_{\sigma \in \text{acce}^{\text{opt}}(G)} \frac{\mathbb{E}_{a \sim \sigma}[W(a)]}{W(a^{\text{opt}})}.$$

The previous proposition provides a positive result, in that it shows that SPoA matches the "true" price of anarchy for the class of ACCE$^{\text{opt}}$. Nevertheless, this result is rather weak. Indeed, it has been shown by means of counterexamples that the best bound on the price of anarchy achievable using a smoothness argument is *not tight* in the class of CCE, [NR10; PCM18]. That is, there are instances G where SPoA $<$ PoA$^{\text{cce}}$ and SPoA provides a rather weak bound on PoA$^{\text{cce}}$, [PCM18]. It follows that the best smoothness bound can not be tight for any of the subclasses of CCE including MNE and NE.

9.4 Complexity of computing equilibria

The goal of this section is to present an overview on the complexity issues related to the equilibrium computation problem. We do not delve in the details of different complexity classes, but simply try to highlight which equilibrium concepts are "hard" to compute and which are "easy".

We divide the presentation in four parts. First, we present an intractability result for pure and mixed Nash equilibria. Second, we restrict our attention to congestion games and show that the best-response dynamics converges in polynomial time under structural assumptions on the actions sets $\{\mathcal{A}_i\}_{i\in N}$. Third, we show that coarse correlated equilibria are tractable in general. We conclude discussing the tradeoff between equilibrium efficiency and computational tractability.

Pure and mixed Nash equilibria are intractable

We begin with a negative result showing that the problem of computing a pure NE is intractable, even if we restrict to the class of congestion games. In the following $\mathcal{PLS}$ represents the complexity class known as *polynomial local search*. Loosely speaking the $\mathcal{PLS}$ class models the difficulty of finding a local optimum solution in the sense of Definition 15. The $\mathcal{PLS}$ class lives in between the classes $\mathcal{P}$ and $\mathcal{NP}$ and there is strong evidence suggesting that $\mathcal{PLS} \not\subseteq \mathcal{P}$, where $\mathcal{P}$ is the class of problems that can be solved polynomially. As a matter of fact, many concrete problems including the local Max-Cut problem are in the $\mathcal{PLS}$ class and no efficient algorithm is available. We redirect the reader to [Rou16] for an introduction to the polynomial local search class.

Proposition 15 (Computing a pure NE is $\mathcal{PLS}$-complete, [FPT04])**.** *The problem of computing a pure NE in a congestion game is $\mathcal{PLS}$-complete.*

It follows that for a general strategic form and finite game (not necessarily a congestion game), computing a pure Nash equilibrium is as hard as the hardest problem in the $\mathcal{PLS}$ class. Any modification of the original problem (e.g., determining if a game G has a pure NE, determining the NE that maximizes a given welfare function) makes it $\mathcal{NP}$-complete [GGS05; CS02].

Computing a mixed Nash equilibrium is also an intractable problem. Its complexity has been settled in [DGP06; CD06] with the introduction of the $\mathcal{PPAD}$ complexity class. Similarly to the $\mathcal{PLS}$ class, the $\mathcal{PPAD}$ class lives in between the classes $\mathcal{P}$ and $\mathcal{NP}$ and despite the great interest in the topic, there are currently no known efficient algorithms to tackle these problems [DGP06].

Proposition 16 (Computing a MNE is $\mathcal{PPAD}$-complete, [DGP06; CD06])**.** *The problem of computing a MNE in a strategic-form finite game is $\mathcal{PPAD}$-complete.*

Nash equilibria are tractable in matroid congestion games

While computing a (pure) Nash equilibrium is intractable even if restricting to the class of congestion games, it is possible to obtain a more positive result by imposing structural constraints on the actions sets.

Definition 20 (Matroid, [Wel10])**.** *A tuple $\mathcal{M} = (\mathcal{R}, \mathcal{I})$ is a matroid if $\mathcal{R}$ is a finite set, $\mathcal{I} \subseteq 2^{\mathcal{R}}$ is a collection of subsets of $\mathcal{R}$, and the following two properties hold:*

- *If $B \in \mathcal{I}$ and $A \subseteq B$, then $A \in \mathcal{I}$;*
- *If $A \in \mathcal{I}$, $B \in \mathcal{I}$ and $|B| > |A|$, then there exists an element $r \in B \setminus A$ s.t. $A \cup \{r\} \in \mathcal{I}$.*

Definition 21 (Basis of a matroid, [Wel10])**.** *A set $S \in \mathcal{I}$ such that for all $r \in \mathcal{R} \setminus S$, $(S \cup r) \notin \mathcal{I}$ is called a basis of the matroid.*

It can be shown that all basis have the same number of elements, which is known as the rank of the matroid and indicated with $\text{rank}(\mathcal{M})$, [Wel10]. An example of matroid is that of uniform matroid defined as follows.

Definition 22 (Uniform matroid, [Wel10])**.** *Given a finite set $\mathcal{R}$ with $|\mathcal{R}| = m$, let $\mathcal{I} \subseteq 2^{\mathcal{R}}$ be the collection of all subsets with a number of elements $k \leq m$. $\mathcal{M} = (\mathcal{R}, \mathcal{I})$ is a matroid, $\text{rank}(\mathcal{M}) = k$ and $\mathcal{M}$ is called the uniform matroid of rank k.*

The following proposition provides sufficient conditions under which the best-response dynamics of Algorithm 5 has polynomial running time for the class of congestion games. The main assumption amounts to requiring each of the player's allocation set to coincide with the set of bases of some matroid.

Proposition 17. *[ARV08, Thm. 2.5] Consider a congestion game G and assume the action sets $\mathcal{A}_i$ are the set of bases for a matroid $\mathcal{M}_i = (\mathcal{R}, \mathcal{I}_i)$ over the set $\mathcal{R}$. Then, players reach a (pure) Nash equilibrium after at most $n^2 m \max_{i \in N} \text{rank}(\mathcal{M}_i)$ best responses.*

Example 3. The case when $\mathcal{A}_i$ contains only sets with a single element (singletons) does satisfy the assumptions of the previous proposition, even if a player does not have access to all the possible resources. One such example is the following: $\mathcal{R} = \{r_1, \ldots, r_m\}$, $m > 2$, $\mathcal{A}_i = \{\{r_1\}, \{r_2\}\}$. Define $\mathcal{I}_i = \{\emptyset, \{r_1\}, \{r_2\}\}$. We have that $\mathcal{M}_i \coloneqq (\mathcal{R}, \mathcal{I}_i)$ is a matroid of rank 1 and that $\mathcal{A}_i$ is a set of bases for $\mathcal{M}_i$.

On the negative side, a few examples that *do not* satisfy the requirements are presented next. Consider $\mathcal{R} = \{r_1, \ldots, r_m\}$, $m \geq 3$ and $\mathcal{A}_i = \{\{r_1\}, \{r_2, r_3\}\}$. The set $\mathcal{A}_i$ can not form the set of bases for any matroid $\mathcal{M}_i$, as all bases must have the same number of elements while $\{r_1\}$ and $\{r_2, r_3\}$ do not have this property. A more involved example that does not satisfy the requirements is the following: $\mathcal{R} = \{r_1, \ldots, r_m\}$, $m \geq 4$ and $\mathcal{A}_i = \{\{r_1, r_2\}, \{r_3, r_4\}\}$. For the given $\mathcal{A}_i$ to be the set of bases of a matroid $\mathcal{M}_i = (\mathcal{R}, \mathcal{I}_i)$, it must be that $\{r_1, r_2\} \in \mathcal{I}_i$ and $\{r_3, r_4\} \in \mathcal{I}_i$. But due to definition of matroid, it must also be $\{r_1\} \in \mathcal{I}_i$ (Definition Definition 20, first property), so that also $\{r_1, r_3, r_4\} \in \mathcal{I}_i$ (Definition 20, second property). Thus $\mathcal{A}_i$ can not be the set of bases for a matroid $\mathcal{M}_i = (\mathcal{R}, \mathcal{I}_i)$, as any possible choice of $\mathcal{I}_i$ will contain at least one set with more elements than $\{r_1, r_2\} \in \mathcal{A}_i$.

Remark 7. *The previous theorem gives conditions under which the maximum number of best responses required to converge to a Nash equilibrium is polynomially bounded in the number of players and resources. If it is possible to compute a single best response polynomially in the number of resources, then it is possible to compute a NE in polynomial time using the best-response algorithm.*

Coarse correlated equilibria

Contrary to NE and MNE, (approximate) coarse correlated equilibria can be computed in polynomial time. We limit ourself to report this result in the following proposition.

Formally, an ε-CCE is defined as a distribution $\sigma \in \Delta(\mathcal{A})$ such that the equilibrium condition in Definition 12 holds up to an additive $\varepsilon \geq 0$ term, i.e.,

$$\mathbb{E}_{a\sim\sigma}\left[u_i(a)\right] + \varepsilon \geq \mathbb{E}_{a\sim\sigma}\left[u_i(a_i', a_{-i})\right] \qquad \forall a_i' \in \mathcal{A}_i, \quad \forall i \in N.$$

Proposition 18 (ε-CCE can be computed efficiently, [LW94; Rou16])**.** *For every $\varepsilon > 0$, an ε-CCE can be computed polynomially using the multiplicative-weight algorithm.*

The tradeoff between tractability and efficiency

This section connects the efficiency result presented in Section 9.3 with the complexity results presented in Section 9.4. In the former section we have seen that $\text{PoA}^{\text{ne}} \geq \text{PoA}^{\text{mne}} \geq \text{PoA}^{\text{cce}}$ and thus the equilibrium efficiency (PoA) degrades by moving from NE to MNE and from MNE to CCE. In the latter section we have seen that NE and

MNE are tractable in limited cases, while CCE are tractable in general. This shows a fundamental tradeoff between equilibrium efficiency and corresponding tractability: the larger the class of equilibria we consider, the easier to compute one, but the lower the corresponding efficiency. This is depicted in Figure 9.1.

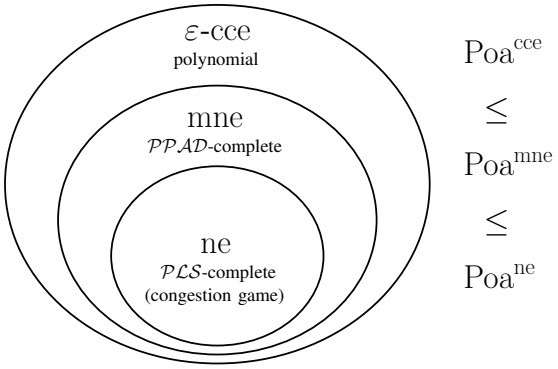

Figure 9.1: Hierarchy of equilibria, corresponding complexity and efficiency.

CHAPTER 10

Tight price of anarchy and utility design: a linear program approach

We seek approximation algorithms for the solution of GMMC problems defined in Section 8.2. Towards this goal, we adopt the game design approach discussed in Chapter 8 and consisting of two separate steps: utility design and algorithm design. In this chapter we formulate and solve the utility design problem. More precisely, in Section 10.1 we pose the utility design problem and introduce the game-theoretic notion of *price of anarchy*. We observe that any algorithm capable of computing an equilibrium, will inherit an approximation ratio matching the price of anarchy. Thus, in a quest to construct good approximating algorithms, we turn our attention to *quantifying* the price of anarchy. In Section 9.3 we show that standard approaches used to characterize the price of anarchy are rather conservative and not suited for the design problems we are interested in (Theorem 7). Motivated by this observation, in Section 10.3 we provide a novel technique based on a linear programming reformulation to characterize the price of anarchy (Theorems 8 and 9) as a function of the utility functions assigned to the agents. This result is provably tight. We conclude the chapter by addressing the utility design question in Section 10.4. In particular, we show how the problem of designing utility functions so as to *optimize* the price of anarchy can be posed as a tractable linear program in $n+1$ variables (Theorem 10).

All the proofs are reported in the Appendix (Section 10.5). The results presented in this chapter have been published in [PCM18; PM18b].

10.1 The price of anarchy as performance metric

Within the combinatorial framework considered, finding a solution to the GMMC problem, i.e., determining a feasible allocation that maximizes the welfare function

$$W(a) = \sum_{r \in \cup_{i \in N} a_i} v_r w(|a|_r),$$

defined in (8.1) is an $\mathcal{NP}$-hard problem. Based on such observation, we focus on deriving *efficient* and *distributed algorithms* for attaining approximate solutions to the maximization of W, ideally with the best possible ratio. In the following, each agent is assumed to have information only regarding the resources that he can select, i.e., regarding the resources $r \in \mathcal{A}_i \subseteq \mathcal{R}$. Agents are requested to make independent choices in response to this local piece of information. Rather than directly specifying a decision-making process, we adopt the *game design* approach discussed in Chapter 8 and depicted in Figure 8.1. The idea is to carefully define an auxiliary problem, namely an equilibrium problem, which will guide the search and serve as a proxy for the original maximization of W. The motivations and advantages of this approach have been discussed in Chapter 8.

In the following we focus on the first component of the game design approach: the *utility design* problem.

The **utility design problem** amounts to the choice of local utility functions that adhere to the above mentioned informational constraints, and whose corresponding equilibria offer the highest achievable performance.

We naturally identify the agents of the original optimization problem and their local constraint sets $\{\mathcal{A}_i\}_{i\in N}$ with the players of the game and their action sets. In the following we will use the terms agents and players interchangeably.

In order to tackle the utility design problem, each agent is assigned a local utility function $u_i : \mathcal{A} \to \mathbb{R}_{\geq 0}$ of the form

$$u_i(a) := \sum_{r\in a_i} v_r w(|a|_r) f(|a|_r)\,, \tag{10.1}$$

where $f : [n] \to \mathbb{R}_{\geq 0}$ describes the fractional benefit that each agent receives by selecting resource r in allocation a. The function f constitutes our *design choice*; we refer to it as to the *distribution rule* or simply the *distribution*. Observe that each utility function in (10.1) satisfies the required informational constraints in that it only depends on the value of the resources that the agent selected, the distribution rule f and the number of agents that selected the very same resource.

Remark 8 (On the choice of the utility functions)**.** *In principle one needs not to restrict himself to utility functions of the form* (10.1)*. The reasons for choosing utilities as in* (10.1) *are as follows. First, the utility functions* (10.1) *satisfy the required informational constraints, as just discussed. Second, restricting ourselves to the above mentioned utilities reduces the design problem to a hopefully tractable problem. Indeed, the utilities* (10.1) *are fully determined if the distribution rule f is so. While designing the distribution rule f amounts to choosing n real numbers, the problem in its full generality consists in choosing the value of $u_i(a)$ for all $a \in \mathcal{A}$ and for all $i \in N$, clearly a large number of decision variables (exponential in the worst case in both the number of agents and in*

the number of resources since $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ and $\mathcal{A}_i \subseteq 2^{\mathcal{R}}$). Third, utilities of the form (10.1) will ensure equilibrium existence and convergence of the best-response dynamics, as explained after this remark. Fourth, even when restricting to this special class of utilities, we will obtain performance certificates that are competitive with the state of the art approximation algorithms. We will return to this in the next chapter in Section 11.1.1 and Remark 15. Finally, we observe that a different and apparently less restrictive choice of utilities might entail assigning different distribution rules f_i to different players $i \in N$. However, it is possible to show that working in this larger set of admissible utility functions will not improve the best achievable performance.[1] *For all these reasons, in the followings we focus on utility functions of the form (10.1).*

The game introduced above and identified with the agents set N, the actions sets $\{\mathcal{A}_i\}_{i\in N}$ and the utilities $\{u_i\}_{i\in N}$ in (10.1) is a normal-form finite game, according to Definition 9. Additionally, such game belongs to the class of congestion games due to the special structure of the actions sets and utilities, see Definition 17. Thus, a pure Nash equilibrium is *guaranteed to exist* for any choice of f thanks to Proposition 12.

In the forthcoming analysis we focus on the solution concept of pure Nash equilibrium, which we will refer to as just an equilibrium. Recall that an allocation $a^{\text{ne}} \in \mathcal{A}$ is a pure Nash equilibrium if $u_i(a^{\text{ne}}) \geq u_i(a_i, a^{\text{ne}}_{-i})$ for all alternative allocations $a_i \in \mathcal{A}_i$ and for all agents $i \in N$ (see Definition 10). We identify one instance of the game introduced above with the tuple

$$G = (\mathcal{R}, \{v_r\}_{r\in\mathcal{R}}, N, \{\mathcal{A}_i\}_{i\in N}, f)\,, \tag{10.2}$$

and for ease of notation remove the subscripts of the above sets, e.g., use $\{\mathcal{A}_i\}$ instead of $\{\mathcal{A}_i\}_{i\in N}$.

In the following we require a system operator to *robustly* design a distribution rule, that is to design f *without* any prior information regarding the resource set $\mathcal{R}$, the value of the resources $\{v_r\}$ or the action sets of the agents $\{\mathcal{A}_i\}$. The *only* datum available to the system designer is an upper bound on the number of players in the game, i.e., $|N| \leq n$. This request stems from the observation that the previous pieces of information may be unreliable, or unavailable to the system designer due to, e.g., communication restrictions or privacy concerns. Formally, given a distribution rule f, we introduce the following family of games

$$\mathcal{G}_f := \{(\mathcal{R}, \{v_r\}, N, \{\mathcal{A}_i\}, f) \text{ s.t. } |N| \leq n\}\,,$$

containing all possible games G where the number of agents is bounded by n. In the forthcoming analysis, we restrict our attention to the class of games where the number of players is *exactly* n. This is without loss of generality. Indeed the latter class of games and the class of games where the number of players is upper bounded by n have the same price of anarchy. To see this, note that the price of anarchy of any game with l players

[1] While we do not provide a proof of this statement, a similar conclusions was found in [Gai09].

$1 < l < n$ can be obtained as the price of anarchy of a game with n players where we add a resource valued $v_0 = 0$ and set $\mathcal{A}_i = \{v_0\}$ for the additional $n - l$ players.

We measure the performance of a distribution rule f adapting the concept of *price of anarchy* introduced in [KP99] and reported in Definition 18 as

$$\mathrm{PoA}(f) \coloneqq \inf_{G \in \mathcal{G}_f} \left(\frac{\min_{a \in \mathrm{ne}(G)} W(a)}{\max_{a \in \mathcal{A}} W(a)} \right), \tag{10.3}$$

where $\mathrm{ne}(G)$ denotes the set of Nash equilibria of G. While the optimal value at the denominator of (10.3) also depends on the instance G considered, we do not indicate it explicitly, for ease of presentation. The quantity $\mathrm{PoA}(f)$ characterizes the efficiency of the worst-performing Nash equilibrium relative to the corresponding optimal allocation over all instances in the class $\mathcal{G}_f$. According to the previous definition, $0 \leq \mathrm{PoA}(f) \leq 1$ and the higher the price of anarchy, the better performance certificates we can offer.

It is important to highlight that whenever an algorithm is available to compute one such equilibrium, the price of anarchy also represents the *approximation ratio* of the corresponding algorithm over all instances $G \in \mathcal{G}_f$. For this reason, the price of anarchy defined in (10.3) will serve as the *performance metric* in all the forthcoming analysis.

Remark 9 (On the choice of pure NE as equilibrium concept)**.** *The choice of pure NE as equilibrium concept has the benefit of providing us with potentially better performance guarantees compared to that offered by, e.g., mixed Nash equilibria or coarse correlated equilibria, as* $\mathrm{PoA}^{\mathrm{ne}} \geq \mathrm{PoA}^{\mathrm{mne}} \geq \mathrm{PoA}^{\mathrm{cce}}$*, see Section 9.3. The drawback of this choice is the general intractability of pure Nash equilibria. Indeed, computing a pure Nash equilibrium is hard (*$\mathcal{PLS}$*-complete, as discussed in Proposition 15) even when limited to the class of games considered here (congestion games). Nevertheless we have seen that under structural assumptions on the sets* $\{\mathcal{A}_i\}$ *similar to those used in combinatorial optimization, computing a pure NE is a polynomial task (Proposition 17). Finally, the approximation guarantees offered by* $\mathrm{PoA}^{\mathrm{ne}}$ *are deterministic, while the bounds provided by* $\mathrm{PoA}^{\mathrm{mne}}$ *and* $\mathrm{PoA}^{\mathrm{cce}}$ *are in expected value. An antipodal choice might entail using the notion of CCE instead of NE as computing one such equilibrium is known to be tractable in general. The price to pay for this is a potentially worsened performance certificate since* $\mathrm{PoA}^{\mathrm{cce}} \leq \mathrm{PoA}^{\mathrm{ne}}$*.*

The utility design problem can be decomposed in two tasks:

i) providing a bound (or ideally an exact characterization) of the price of anarchy as a function of f;

ii) optimizing this expression over the admissible distribution rules.

In Section 10.3 we address i), while in Section 10.4 we tackle ii).

10.2 The limitations of the smoothness framework

In this section we recall the definition of smooth games introduced in Section 9.3, and show that the corresponding best achievable bounds on the price of anarchy are not tight, but rather conservative when applied to utility design problems.

Before delving in the details of the smoothness framework, we introduce the notion of budget-balanced and sub budget-balanced utility functions.

Definition 23 (Budget-balanced utility functions)**.** *Consider a strategic-form game with actions sets $\{\mathcal{A}_i\}$, utilities $\{u_i\}$, and a welfare function $W : \mathcal{A} \to \mathbb{R}_{\geq 0}$. The utility functions are budget-balanced if for all $a \in \mathcal{A}$*

$$\sum_{i \in N} u_i(a) = W(a).$$

The utility functions are sub budget-balanced if $\sum_{i \in N} u_i(a) \leq W(a)$ for all $a \in \mathcal{A}$.

The notion of smooth game has been introduced in [Rou09] and has been successively employed to obtain tight bounds on the price of anarchy for different classes of games. Recall from Definition 19 that the game (10.2) together with the welfare function (8.1) are (λ, μ)-smooth if for some $\lambda, \mu \geq 0$ it holds

$$\sum_{i \in N} u_i(a'_i, a_{-i}) \geq \lambda W(a') - \mu W(a), \quad \forall a', a \in \mathcal{A}. \tag{10.4}$$

Proposition 13 showed that the price of anarchy of a (λ, μ)-smooth game G is bounded. More precisely, given G a (λ, μ)-smooth game with $\sum_{i \in N} u_i(a) \leq W(a)\ \forall a \in \mathcal{A}$, the ratio between the total welfare at any coarse correlated equilibrium and the optimum is lower bounded by

$$\frac{\mathbb{E}_{a \sim \sigma^{\text{cce}}}[W(a)]}{W(a^{\text{opt}})} \geq \frac{\lambda}{1+\mu}, \qquad \forall \sigma^{\text{cce}} \in \text{cce}(G).$$

Since $\text{ne}(G) \subseteq \text{cce}(G)$, it follows immediately

$$\frac{W(a^{\text{ne}})}{W(a^{\text{opt}})} \geq \frac{\lambda}{1+\mu}, \qquad \forall a^{\text{ne}} \in \text{ne}(G).$$

Note that the smoothness framework forces us to restrict the attention to utilities satisfying $\sum_{i \in N} u_i(a) \leq W(a)$, else no guarantee is provided by Proposition 13. This corresponds to requesting $f(j) \leq 1/j$. Thus, in the remaining of this section *only*, we consider utilities satisfying this constraint.

The next lemma shows that when we are allowed to freely choose the players' utilities (i.e., if we are interested in design problems), the best achievable smoothness guarantee is obtained when the assigned utilities are budget-balanced.

Lemma 12. *Suppose* $\sum_{i\in N} u_i(a) = W(a)$. *Consider a different set of utilities* $\tilde{u}_i(a)$ *such that* $\sum_{i\in N} \tilde{u}_i(a) \leq \sum_{i\in N} u_i(a)$ *for all* $a \in \mathcal{A}$. *If the game with utilities* $\tilde{u}_i(a)$ *is* (λ, μ)*-smooth, then the game with utilities* $u_i(a)$ *is also* (λ, μ)*-smooth.*

Proof. By assumption the game with utilities $\tilde{u}_i(a)$ is (λ, μ)-smooth and $\sum_i u_i(a) \geq \sum_i \tilde{u}_i(a)$, so that for all $a, a' \in \mathcal{A}$

$$\sum_{i\in N} u_i(a'_i, a_{-i}) \geq \sum_{i\in N} \tilde{u}_i(a'_i, a_{-i}) \geq \lambda W(a') - \mu W(a).$$

Thus, the game with utilities $u_i(a)$ is (λ, μ)-smooth too. □

Observe that the statement of Lemma 12 holds true in general and does not depend on the specific form of the utility functions or of the welfare considered here.

Lemma 12 suggests to design utilities that are budget-balanced, as sub budget-balanced utilities can never be advantageous with regards to the performance guarantees associated with smoothness. This observation turns out to be misleading, in that there are utility functions that are sub budget-balanced, but give a better performance certificate compared to what the smoothness argument can offer, as shown next.

Consider f a distribution rule satisfying $f(j) \leq 1/j$ for all $j \in [n]$, the best bound on the price of anarchy (10.3) that can be obtained via smoothness, is given by the solution to the following program

$$\begin{aligned} \mathrm{SPoA}(f) \coloneqq \sup_{\lambda,\mu\geq 0} \ & \frac{\lambda}{1+\mu} \\ & \text{s.t. } (\lambda, \mu) \text{ satisfy (10.4) for all } G \in \mathcal{G}_f. \end{aligned}$$

Observe that $\mathrm{SPoA}(f) \leq \mathrm{PoA}(f)$ as Proposition 13 provides only a bound on the equilibrium efficiency. In the following we show that the best smoothness bound captured by $\mathrm{SPoA}(f)$ is not representative of the "true" price of anarchy $\mathrm{PoA}(f)$ defined in (10.3). To do so, we illustrate the *gap* between these two quantities in the special case of multi-agent weighted maximum coverage (MMC) problems (see Section 8.2). MMC problems are a special class of the resource allocation problems considered here. They are obtained setting $w(j) = 1$ for all $j \in [n]$. Before stating the result, we introduce the distribution rule

$$f_{\mathrm{G}}(j) = (j-1)! \frac{\frac{1}{(n-1)(n-1)!} + \sum_{i=j}^{n-1} \frac{1}{i!}}{\frac{1}{(n-1)(n-1)!} + \sum_{i=1}^{n-1} \frac{1}{i!}}, \quad j \in [n]\,. \tag{10.5}$$

as originally defined in [Gai09, Eq. (5)].

Theorem 7 (Limitations of the smoothness framework). *Consider the class of MMC problems, i.e., fix* $w(j) = 1$ *for all* $j \in [n]$.

(a) *For any choice of f, the best bound on the price of anarchy that can be achieved using a smoothness argument is*

$$\mathrm{SPoA}(f) \leq \frac{1}{2-1/n} \coloneqq b(n) \quad \xrightarrow{n\to\infty} \quad \frac{1}{2}\,.$$

(b) *The distribution (10.5) satisfies $f_{\mathrm{G}}(j) \leq 1/j$ and achieves*

$$\mathrm{PoA}(f_{\mathrm{G}}) = 1 - \frac{1}{\frac{1}{(n-1)(n-1)!} + \sum_{i=0}^{n} \frac{1}{i!}} \quad \xrightarrow{n\to\infty} \quad 1 - \frac{1}{e}\,, \tag{10.6}$$

where e is Euler's number.

(c) *For all $n > 2$, $\mathrm{SPoA}(f_{\mathrm{G}}) < \mathrm{PoA}(f_{\mathrm{G}})$.*

Remark 10 (The limitations of smoothness are structural)**.** *While the previous theorem compares the performance guarantees offered by $\mathrm{SPoA}(f)$ and $\mathrm{PoA}(f)$ we recall that $\mathrm{SPoA}(f)$ bounds the equilibrium efficiency for any coarse correlated equilibrium, while $\mathrm{PoA}(f)$ provides a certificate limitedly to pure NE. Thus, one might think that the result of the previous theorem is simply an artifact due to this observation and to the fact that $\mathrm{ne}(G) \subseteq \mathrm{cce}(G)$. This is not the case and the limitations of the smoothness framework are structural. Indeed, it can be shown that f_{G} has the same price of anarchy of (10.6) even in the larger set of CCE.*[2]

The quantity $b(n)$ bounding the best possible performance certificate offered by the smoothness framework, and the guarantee offered by the "true" price of anarchy for f_{G} are presented in Figure 10.1 (left). Additionally, the distribution rules $f_{\mathrm{G}}(j)$ and $1/j$ are depicted in Figure 10.1 (right). The gap between $b(n)$ and $\mathrm{PoA}(f_{\mathrm{G}})$ is significant: for a system with, e.g., $n = 20$ agents, $\mathrm{PoA}(f_{\mathrm{G}})$ produces a performance certificate that is at least 25% higher than what $\mathrm{SPoA}(f_{\mathrm{G}})$ can offer. Thus, the smoothness framework is not the right tool to study the utility design problems considered here. First, it restricts the set of admissible distribution rule to $f(j) \leq 1/j$. Second, even for distribution rules satisfying this assumption, it provides performance certificates that are too conservative. Finally, we observe that the notion of local smoothness (a refinement of the original notion introduced in [RS15]) will not be useful here in improving SPoA.

10.3 A tight price of anarchy

In the previous section we have highlighted the limitations of the smoothness framework when applied to utility design problems. In this section we propose a novel approach for

[2] While [Gai09, Thm. 3] provides a proof limitedly to mixed Nash equilibria, it is not difficult to extend such proof to CCE.

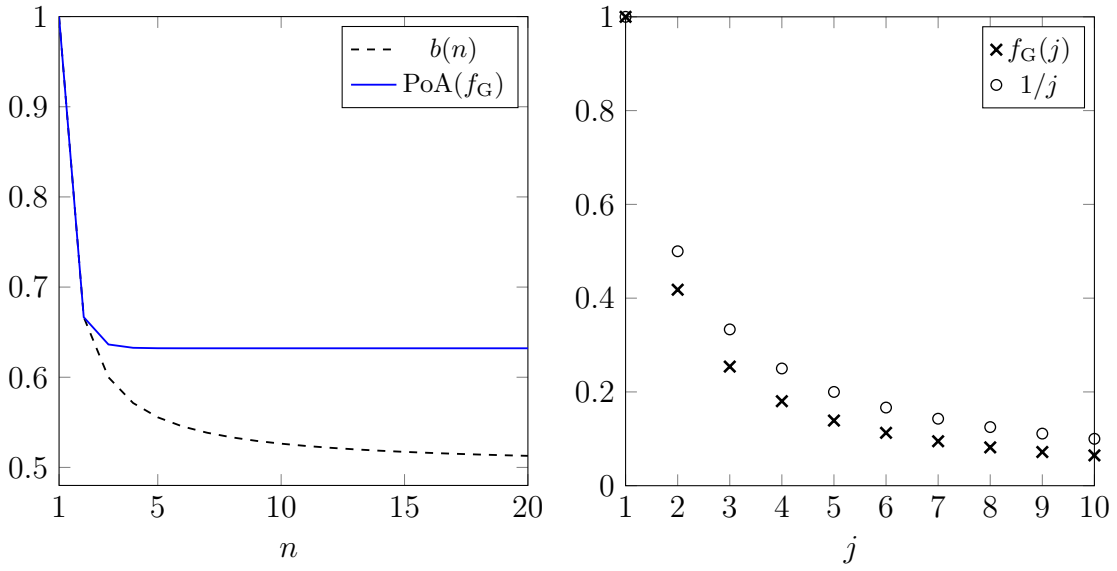

Figure 10.1: Left: best achievable bound $b(n)$ on the price of anarchy using a smoothness argument, and actual price of anarchy $\mathrm{PoA}(f_{\mathrm{G}})$ for the distribution f_{G} in (10.5). Right: distribution rule f_{G} and $1/j$ for $n = 10$.

the exact characterization of $\mathrm{PoA}(f)$ as defined in (10.3). More precisely, we reformulate the problem of *computing* the price of anarchy as a *tractable linear program* (LP) involving the components of w and of f (Theorems 8 and 9). This section is dedicated to the problem of characterizing the price of anarchy in its full generality, while in Chapter 11 we specialize the results to a class of submodular and supermodular problems.

In all the forthcoming analysis we make the following regularity assumptions on admissible welfare basis functions and distribution rules.

Standing Assumptions. *The sets $\mathcal{A}_i \subseteq 2^{\mathcal{R}}$ are nonempty and $\mathcal{A}_i \setminus \emptyset \neq \emptyset$ for all $i \in N$. Further, $\exists r \in \mathcal{R}$ s.t. $v_r > 0$ and $r \in a_i \in \mathcal{A}_i$ for some $i \in N$. The welfare basis function $w : [n] \to \mathbb{R}_{>0}$ satisfies $w(j) > 0$ for all $j \in [n]$. A distribution rule $f : [n] \to \mathbb{R}_{\geq 0}$ satisfies $f(1) \geq 1$, $f(j) \geq 0$ for all $j \in [n]$. The latter is equivalent to $f \in \mathcal{F}$, with*

$$\mathcal{F} \coloneqq \{f : [n] \to \mathbb{R}_{\geq 0} \text{ s.t. } f(1) \geq 1,\ f(j) \geq 0\ \forall j \in [n]\}\,.$$

The non-emptiness of $\mathcal{A}_i$ ensures feasibility of the welfare maximization introduced in Section 8.2. The assumptions $\mathcal{A}_i \setminus \emptyset \neq \emptyset$ for all $i \in N$ and $\exists r \in \mathcal{R}$ s.t. $v_r > 0$ ensure that the problem is non degenerate, in that every agent has the possibility to select at least one resource, and not all the resources have a value of zero. Finally, observe that the assumption $f(1) \geq 1$ is without loss of generality for all distributions with $f(1) > 0$. Indeed, If $f(1) \neq 1$, but $f(1) > 0$, it is possible to scale the value of the resources and reduce to the case $f(1) = 1$.

10.3.1 Primal formulation

An informal introduction

While (10.3) corresponds to the *definition* of price of anarchy, it also describes a (seemingly difficult) *optimization problem*. The aim of this section, is to transform the definition of price of anarchy into a finite dimensional LP that can be efficiently solved. Towards this goal, we provide here an informal introduction (based on four steps), as we believe the reader will benefit from it. The formal derivation and justification of each of these steps is *postponed to Theorem 8* and its proof.

Step 1: we observe that the price of anarchy computed over the family of games $G \in \mathcal{G}_f$ is the same of the price of anarchy over the reduced family of games $\hat{\mathcal{G}}_f$, where the feasible set of every player only contains two allocations: (worst) equilibrium and optimal allocation, that is $\hat{\mathcal{A}}_i = \{a_i^{\text{ne}}, a_i^{\text{opt}}\}$, $\hat{\mathcal{G}}_f \coloneqq \{(\mathcal{R}, \{v_r\}, N, \{\hat{\mathcal{A}}_i\}, f)\}$. Thus, definition (10.3) reduces to

$$\begin{aligned} \text{PoA}(f) = \inf_{G \in \hat{\mathcal{G}}_f} & \left(\frac{W(a^{\text{ne}})}{W(a^{\text{opt}})} \right), \\ & \text{s.t.} \quad u_i(a^{\text{ne}}) \geq u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \quad \forall i \in N, \end{aligned}$$

where we have constrained a^{ne} to be an equilibrium. We do not include the additional constraints requiring a^{ne} to be the *worst* equilibrium and a^{opt} to provide the *highest* welfare. This is because the infimum over $\hat{\mathcal{G}}_f$ and the parametrization we will introduce to describe an instance G (in step 4) will implicitly ensure this.

Step 2: we assume without loss of generality that $W(a^{\text{ne}}) = 1$ and get

$$\begin{aligned} \text{PoA}(f) = \inf_{G \in \hat{\mathcal{G}}_f} & \frac{1}{W(a^{\text{opt}})}, \\ & \text{s.t.} \quad u_i(a^{\text{ne}}) \geq u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \quad \forall i \in N, \\ & \qquad W(a^{\text{ne}}) = 1. \end{aligned} \tag{10.7}$$

Step 3: we relax the previous program as in the following

$$\begin{aligned} \text{PoA}(f) = \inf_{G \in \hat{\mathcal{G}}_f} & \frac{1}{W(a^{\text{opt}})}, \\ & \text{s.t.} \quad \sum_{i \in N} u_i(a^{\text{ne}}) - u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \geq 0, \\ & \qquad W(a^{\text{ne}}) = 1, \end{aligned} \tag{10.8}$$

where the n equilibrium constraints (one per each player) have been substituted by their sum. We show that the relaxation gives the same price of anarchy of (10.7).

Step 4: for a given instance in the reduced family $\hat{\mathcal{G}}_f$, computing the efficiency amounts

to identifying an optimal allocation and the corresponding worst Nash equilibrium. The additional difficulty appearing in (10.8) is in how to describe a generic instance $G \in \hat{\mathcal{G}}_f$ and on how to compute the infimum over all such (infinite) instances. To do so, we introduce an efficient parametrization that fully describes the objective function and the decision variables of the previous problem. This allows to reduce (10.8) and obtain the result in the following Theorem 8.

The linear program

The following theorem makes the reasoning presented in Section 10.3.1 formal and constitutes the second result of this manuscript.

In order to capture all the instances in $\hat{\mathcal{G}}_f$, we use a parametrization inspired by [War12] and introduce the variables $\theta(a, x, b) \in \mathbb{R}$ defined for any tuple of integers $(a, x, b) \in \mathcal{I}$, where

$$\mathcal{I} \coloneqq \{(a, x, b) \in \mathbb{N}^3_{\geq 0} \text{ s.t. } 1 \leq a + x + b \leq n\},$$
$$\mathcal{I}_R \coloneqq \{(a, x, b) \in \mathcal{I} \text{ s.t. } a \cdot x \cdot b = 0 \text{ or } a + x + b = n\}.$$

Note that $\mathcal{I}_R$ contains all the integer points on the planes $a = 0$, $b = 0$, $x = 0$, $a+x+b = n$ bounding $\mathcal{I}$. The set $\mathcal{I}$ is depicted in Figure 10.2 for the case of $n = 3$.

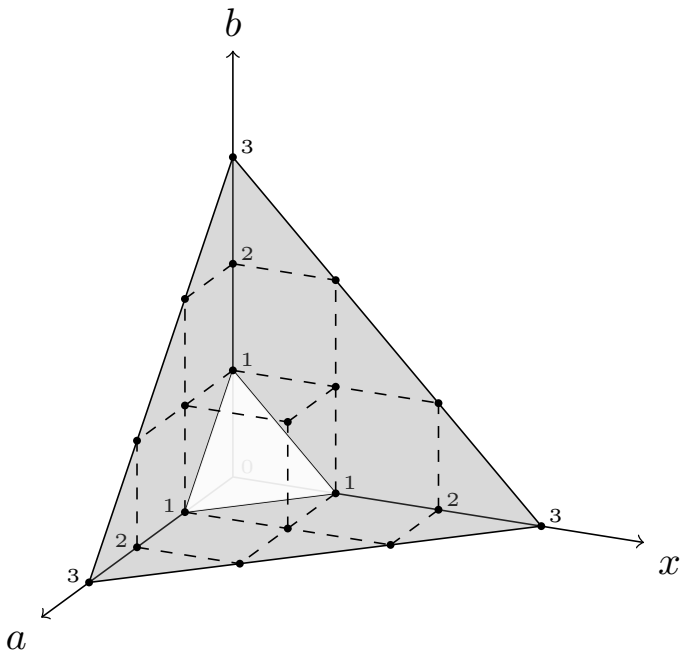

Figure 10.2: The black circles represent all the points belonging to $\mathcal{I}$, $n = 3$.

In the remainder we write $\sum_{a,x,b}$ instead of $\sum_{(a,x,b)\in\mathcal{I}}$, for readability. Additionally, given a distribution rule $f : [n] \to \mathbb{R}_{\geq 0}$, and a welfare basis function $w : [n] \to \mathbb{R}_{>0}$, we extend their definition, with slight abuse of notation, to $f : [n+1]_0 \to \mathbb{R}_{\geq 0}$ and

$w : [n+1]_0 \to \mathbb{R}_{\geq 0}$, where we set the first and last components to be identically zero, i.e., $f(0) = w(0) = 0$, $f(n+1) = w(n+1) = 0$.[3]

Theorem 8 (PoA as a linear program)**.** *Given $f \in \mathcal{F}$, the price of anarchy (10.3) is*

$$\mathrm{PoA}(f) = \frac{1}{W^\star},$$

where $W^\star$ is the value of the following (primal) linear program in the unknowns $\theta(a,x,b) \in \mathbb{R}_{\geq 0}$, $(a,x,b) \in \mathcal{I}$

$$\begin{aligned} W^\star = \max_{\theta(a,x,b)} & \sum_{a,x,b} \mathbb{1}_{\{b+x\geq 1\}} w(b+x)\theta(a,x,b) \\ s.t. & \sum_{a,x,b} [af(a+x)w(a+x) - bf(a+x+1)w(a+x+1)]\theta(a,x,b) \geq 0 \\ & \sum_{a,x,b} \mathbb{1}_{\{a+x\geq 1\}} w(a+x)\theta(a,x,b) = 1 \\ & \theta(a,x,b) \geq 0 \quad \forall (a,x,b) \in \mathcal{I}\,. \end{aligned} \tag{10.9}$$

The proof is based on the four steps previously discussed.

Given a distribution rule f, the solution to the previous program returns both the price of anarchy, and the corresponding worst case instance (encoded in $\theta(a,x,b)$, see the proof of the Step 4 in Section 10.5). Observe that the number of decision variables in (10.9) is $|\mathcal{I}| = \frac{1}{2}\sum_{j=0}^{n}(j+2)(j+1) - 1 = n(n+1)(n+2)/6 - 1 \sim \mathcal{O}(n^3)$, while only two scalar constraints are present (neglecting the positivity constraint). The previous program can thus already be solved efficiently. Nevertheless, we are only interested in the expression of $\mathrm{PoA}(f)$ (i.e., ultimately in the *value* of the program), and therefore consider the dual counterpart of (10.9) in the following.

10.3.2 Dual formulation

Thanks to strong duality, it suffices to solve the dual program of (10.9) to compute the price of anarchy (10.3). While the dual program should feature two scalar decision variables and $\mathcal{O}(n^3)$ constraints, the following theorem shows how to reduce the number of constraints to only $|\mathcal{I}_R| = 2(n^2+1) - 1 \sim \mathcal{O}(n^2)$. The overarching goal is to progress towards an *explicit expression* for $\mathrm{PoA}(f)$.

[3] This adjustment does not play any role, but is required to avoid the use of cumbersome notation in the forthcoming expressions. Else, e.g., $f(a+x+1)$ and $w(a+x+1)$ in (10.9) will not be defined for $a+x=n$.

Theorem 9 (Dual reformulation of PoA)**.** *Given $f \in \mathcal{F}$, the price of anarchy (10.3) is* $\mathrm{PoA}(f) = 1/W^\star$*, where $W^\star$ is the value of the following (dual) program*

$$\begin{aligned} W^\star = \min_{\lambda \in \mathbb{R}_{\geq 0}, \mu \in \mathbb{R}} \ & \mu \\ s.t. \ & \mathbb{1}_{\{b+x \geq 1\}} w(b+x) - \mu \mathbb{1}_{\{a+x \geq 1\}} w(a+x) + \\ & + \lambda [a f(a+x) w(a+x) - b f(a+x+1) w(a+x+1)] \leq 0 \\ & \qquad \forall (a, x, b) \in \mathcal{I}_R \end{aligned} \tag{10.10}$$

The proof of the previous theorem (see Section 10.5) suggests that a further simplification can be made when $f(j)w(j)$ is non-increasing for all j. In this case the number of constraints reduces to exactly n^2, as detailed in the following corollary.

Corollary 3. *Consider a given $f \in \mathcal{F}$.*

(a) Assume $f(j)w(j)$ non increasing for $j \in [n]$. Then $\mathrm{PoA} = 1/W^\star$*, where*

$$\begin{aligned} W^\star = \min_{\lambda \in \mathbb{R}_{\geq 0}, \mu \in \mathbb{R}} \ & \mu \\ s.t. \ & \mu w(j) \geq w(l) + \lambda [j f(j) w(j) - l f(j+1) w(j+1)] \\ & \qquad \forall j, l \in [0, n], \quad 1 \leq j + l \leq n, \\ & \mu w(j) \geq w(l) + \lambda [(n-l) f(j) w(j) - (n-j) f(j+1) w(j+1)] \\ & \qquad \forall j, l \in [0, n], \qquad j + l > n. \end{aligned} \tag{10.11}$$

(b) If additionally $f(j) \geq \frac{1}{j} f(1) w(1) \min_{l \in [n]} \frac{l}{w(l)}$, then

$$\lambda^\star = \max_{l \in [n]} \frac{w(l)}{l} \frac{1}{f(1) w(1)} .$$

Mimicking the proof of the previous corollary, it is possible to obtain a similar result when $f(j)w(j)$ is instead non-decreasing. While the requirements on $f(j)w(j)$ being non increasing might seem restrictive at first, similar assumptions were made relative to a simpler class of problems in [MR14; Gai09]. We remark that this requirement is added to obtain an explicit expression for the price of anarchy. If this is not the goal, one can compute $\mathrm{PoA}(f)$ using Theorem 9 without imposing any additional assumption.

Remark 11 (Explicit expression of $\mathrm{PoA}(f)$)**.** *Observe that, if the optimal value $\lambda^\star$ is known a priori, as in the second statement from the previous corollary, the quantity $W^\star$ (and consequently the price of anarchy) can be computed* explicitly *from (10.11) as the maximum between n^2 real numbers depending on all the entries of f and w. To see this, divide both sides of the constraints in (10.11) by $w(j)$ for $1 \leq j \leq n$, and observe*

that the solution $\mu^\star$ is then found as the maximum of the resulting right hand side. The corresponding value of $W^\star$ is given by the following expression.

$$W^\star = \max \begin{cases} \max\limits_{\substack{j\neq 0 \\ 1\leq j+l\leq n \\ j,l\in[0,n]}} \frac{w(l)}{w(j)} + \lambda^\star[jf(j) - lf(j+1)\frac{w(j+1)}{w(j)}] \\ \\ \max\limits_{\substack{j\neq 0 \\ j+l>n \\ j,l\in[0,n]}} \frac{w(l)}{w(j)} + \lambda^\star[(n-l)f(j) - (n-j)f(j+1)\frac{w(j+1)}{w(j)}] \end{cases} \tag{10.12}$$

Equation (10.12) is reminiscent of the result obtained using a very different approach in [MR14, Thm. 6] (limited to Shapley value) and [Gai09, Thm. 3] (limited to set covering problems and sub budget-balanced utilities).

Finally, observe that for the case of MMC problems discussed in Section 10.2 (it is $w(j) = 1$ for all $j \in [n]$) the assumption required in the first statement of the previous corollary reduces to $f(j)$ non increasing. That is, the previous corollary gives us an expression for the PoA(f) *also for utilities that do not satisfy $\sum_{i\in N} u_i(a) \leq W(a)$, as instead required in [Gai09]. We discuss further connections with these works and others in Chapter 11.*

10.3.3 Related works

The idea of using an auxiliary linear program to study the equilibrium efficiency has appeared in few works in the literature [NR10; Bil12; KM14; Tha17]. Note that, *all* the aforementioned works assume the budget-balance condition to hold true. In [NR10], the authors pose the problem in an abstract form and the corresponding linear program is used as a conceptual tool, rather than as a machinery to explicitly compute the price of anarchy. While [Bil12] provides result for polynomial latency functions in weighted congestion games, the techniques proposed in [Bil12; KM14; Tha17] require an ad-hoc bound on the dual objective to obtain a bound on the price of anarchy. This is not the case with our approach. Additionally, we note that the linear programming reformulations of [NR10] capture the price of anarchy for a *given* problem instance, while in this work we consider the *worst case* instance over an admissible class of problems. This additional requirement complicates the analysis, but will produce algorithms that are provably robust to the presence of uncertainty, and are thus better suited for engineering implementation. Finally, we observe that a direct transposition of the approach in, e.g., [NR10] to our setting would produce a linear program whose size grows exponentially in the number of resources, making it impossible to solve for real world applications.

10.4 Optimal utility design via linear programming

Given w and a distribution rule f, Theorem 9 and Corollary 3 have reduced the computation of the price of anarchy to the solution of a tractable linear program. Nevertheless, determining the distribution rule maximizing $\mathrm{PoA}(f)$, i.e., giving the best performance guarantees, is *also a tractable linear program*. The following theorem makes this clear.

Theorem 10 (Optimizing $\mathrm{PoA}(f)$ is a linear program)**.** *For a given welfare basis w, the design problem*

$$\arg\max_{f \in \mathcal{F}} \mathrm{PoA}(f)$$

is equivalent to the following LP in $n+1$ scalar unknowns

$$\begin{aligned} f^\star \in \arg\min_{f \in \mathcal{F},\, \mu \in \mathbb{R}} \ & \mu \\ s.t. \ \ & \mathbb{1}_{\{b+x \geq 1\}} w(b+x) - \mu \mathbb{1}_{\{a+x \geq 1\}} w(a+x) + \\ & + a f(a+x) w(a+x) - b f(a+x+1) w(a+x+1) \leq 0 \\ & \qquad \forall (a,x,b) \in \mathcal{I}_R \end{aligned} \tag{10.13}$$

The corresponding optimal price of anarchy is

$$\mathrm{PoA}(f^\star) = \frac{1}{\mu^\star},$$

where $\mu^\star$ is the value of the program (10.13).

Remark 12. *The importance of this results stems from its applicability for the* game design procedure *outlined in Chapter 8. More precisely, the previous theorem provides a solution to the utility design problem introduced in Section 10.1. As a matter of fact, Theorem 10 allows to compute the optimal distribution rule, for any given welfare basis function (satisfying the Standing Assumptions), and thus to solve the utility design problem. Applications of these results are presented in Chapters 11 and 12.*

10.5 Appendix

10.5.1 Proofs of the results presented in Section 10.2

Proof of Theorem 7

Proof. We prove the first and third claims only, as the second statement is shown in [Gai09, Thm. 3].

(a) The claim in Proposition 13 requires $f(j) \leq 1/j$, so that we need to restrict to this class of admissible utility functions to apply any smoothness argument. We proceed dividing the proof in two parts. First, we consider the valid distribution rule f_{SV} defined for all $j \in [n]$ as $f_{\mathrm{SV}}(j) \coloneqq 1/j$, and show that the best smoothness parameters are $(1, 1-1/n)$ so that

$$\mathrm{SPoA}(f_{\mathrm{SV}}) = \frac{1}{2-1/n} = b(n).$$

Second, we show that for any distribution with $f(j) \leq f_{\mathrm{SV}}(j)$ for all $j \in [n]$ it holds $\mathrm{SPoA}(f) \leq \mathrm{SPoA}(f_{\mathrm{SV}})$. From this, we conclude $\mathrm{SPoA}(f) \leq b(n) = \frac{1}{2-1/n}$ for all admissible distribution rules.

Part 1: with the special choice of f_{SV}, the proof of [Gai09, Thm. 2] shows that for any pair of feasible a, a' and any $G \in \mathcal{G}_f$, it holds

$$\sum_{i \in N} u_i(a'_i, a_{-i}) \geq W(a') - \chi_{\mathrm{SV}} W(a)\,,$$

$$\text{where} \quad \chi_{\mathrm{SV}} = \max_{j \in [n-1]} \{j f_{\mathrm{SV}}(j) - f_{\mathrm{SV}}(j+1), (n-1) f_{\mathrm{SV}}(n)\}\,,$$

from which $\chi_{\mathrm{SV}} = 1 - 1/n$. Thus the game is $(1, 1-1/n)$-smooth and it follows that $\mathrm{SPoA}(f_{\mathrm{SV}}) \geq \frac{1}{2-1/n}$. To show that there is no better pair (λ, μ) we show that the price of anarchy is exactly $\frac{1}{2-1/n}$. To do so, we consider the instance G proposed in [RPM17, Fig. S2] and observe that $W(a^{\mathrm{opt}}) = 2 - 1/n$ while $W(a^{\mathrm{ne}}) = 1$. Thus, $\mathrm{SPoA}(f_{\mathrm{SV}}) \leq \mathrm{PoA}(f_{\mathrm{SV}}) \leq \frac{1}{2-1/n}$. Since the lower and upper bounds obtained for $\mathrm{SPoA}(f_{\mathrm{SV}})$ match, we conclude that $\mathrm{SPoA}(f_{\mathrm{SV}}) = \frac{1}{2-1/n}$.

Part 2: Consider any distribution rule such that $f(j) \leq f_{\mathrm{SV}}(j)$ for all $j \in [n]$. Let us define the set

$$A(f) \coloneqq \Bigg\{ (\lambda, \mu) \text{ s.t. for all } a, a' \in \mathcal{A}, \text{ for all } G \in \mathcal{G}_f \\ \sum_{\substack{i \in N \\ r \in a_i}} v_r f(|(a'_i, a_{-i})|_r) w(|(a'_i, a_{-i})|_r) \geq \lambda W(a') - \mu W(a) \Bigg\},$$

and analogously for $A(f_{\mathrm{SV}})$. With this notation, the claim we intend to prove reduces to

$$\sup_{(\lambda,\mu)\in A(f)} \frac{\lambda}{1+\mu} \leq \sup_{(\lambda,\mu)\in A(f_{\mathrm{SV}})} \frac{\lambda}{1+\mu}. \tag{10.14}$$

To show the latter, we prove that $A(f) \subseteq A(f_{\mathrm{SV}})$. Consider a feasible tuple $(\lambda,\mu) \in A(f)$; by definition of $A(f)$ it is

$$\sum_{\substack{i\in N\\ r\in a_i}} v_r f(|(a'_i, a_{-i})|_r) w(|(a'_i, a_{-i})|_r) \geq \lambda W(a') - \mu W(a),$$

$\forall a, a' \in \mathcal{A}$, $G \in \mathcal{G}_f$. Since $f_{\mathrm{SV}}(j) \geq f(j)$, it follows that

$$\sum_{\substack{i\in N\\ r\in a_i}} v_r f_{\mathrm{SV}}(|(a'_i, a_{-i})|_r) w(|(a'_i, a_{-i})|_r) \geq \lambda W(a') - \mu W(a),$$

$\forall a, a' \in \mathcal{A}$, $G \in \mathcal{G}_f$. Thus $(\lambda,\mu) \in A(f_{\mathrm{SV}})$ too, from which we conclude that $A(f) \subseteq A(f_{\mathrm{SV}})$ and (10.14) must hold.

(c) Follows from the previous claims upon noticing that $b(n) < \mathrm{PoA}(f_{\mathrm{G}})$ for $n > 2$ (while $b(n) = \mathrm{PoA}(f_{\mathrm{G}})$ for $n = 2$).

□

10.5.2 Proofs of the results presented in Section 10.3

Proof of Theorem 8

Proof. The proof formalizes the steps introduced in Section 10.3.1.

Step 1: We intend to show that the price of anarchy computed over $G \in \mathcal{G}_f$ is the same of the price of anarchy computed over a reduced set of games. Consider a game $G \in \mathcal{G}_f$ and denote with a^{ne} the corresponding worst equilibrium (as measured by W) and with a^{opt} an optimal allocation of G. For every such game G, we construct a new game $\hat{G}$, where $\hat{G} \coloneqq (\mathcal{R}, \{v_r\}, N, \{\hat{\mathcal{A}}_i\}, f)$ and $\hat{\mathcal{A}}_i = \{a_i^{\mathrm{ne}}, a_i^{\mathrm{opt}}\}$ for all $i \in N$. That is, the feasible set of every player in $\hat{G}$ contains only two allocations: an optimal allocation, and the (worst) equilibrium of G. With slight abuse of notation we write $\hat{G}(G)$ to describe the game $\hat{G}$ constructed from G as just discussed. Observe that G and $\hat{G}$ have the same price of anarchy, i.e.,

$$\frac{\min_{a\in \mathrm{ne}(G)} W(a)}{\max_{a\in\mathcal{A}} W(a)} = \frac{\min_{a\in \mathrm{ne}(\hat{G})} W(a)}{\max_{a\in\hat{\mathcal{A}}} W(a)}.$$

Denote with $\hat{\mathcal{G}}_f$ the class of games $\hat{\mathcal{G}}_f \coloneqq \{\hat{G}(G)\ \forall G \in \mathcal{G}_f\}$. Observe that $\hat{\mathcal{G}}_f \subseteq \mathcal{G}_f$ (by definition) and since for every game $G \in \mathcal{G}_f$, it is possible to construct a game $\hat{G} \in \hat{\mathcal{G}}_f$ with the same price of anarchy, it follows that (10.3) can be computed as

$$\mathrm{PoA}(f) = \inf_{\hat{G}\in\hat{\mathcal{G}}_f} \left(\frac{\min_{a\in \mathrm{ne}(\hat{G})} W(a)}{\max_{a\in\mathcal{A}} W(a)} \right).$$

Step 2: Lemma 13 ensures for any game G, every equilibrium configuration has strictly positive welfare. Thus, we assume without loss of generality that $W(a^{\mathrm{ne}}) = 1$, where a^{ne} represents the worst equilibrium of G.[4] The price of anarchy reduces to

$$\begin{aligned} \mathrm{PoA}(f) = \inf_{\hat{G}\in\hat{\mathcal{G}}_f} \frac{1}{W(a^{\mathrm{opt}})}\,, \\ \text{s.t.} \quad u_i(a^{\mathrm{ne}}) \geq u_i(a_i^{\mathrm{opt}}, a_{-i}^{\mathrm{ne}}) \quad \forall i \in N\,, \\ W(a^{\mathrm{ne}}) = 1\,. \end{aligned}$$

Steps 3, 4: While in Section 10.3.1 these steps have been introduced separately for ease of exposition, their proof is presented jointly here. First observe, from the last equation, that $\mathrm{PoA}(f) = 1/W^\star$, where

$$\begin{aligned} W^\star \coloneqq \sup_{\hat{G}\in\hat{\mathcal{G}}_f} W(a^{\mathrm{opt}})\,, \\ \text{s.t.} \quad u_i(a^{\mathrm{ne}}) \geq u_i(a_i^{\mathrm{opt}}, a_{-i}^{\mathrm{ne}}) \quad \forall i \in N\,, \\ W(a^{\mathrm{ne}}) = 1\,. \end{aligned} \tag{10.15}$$

We relax the previous program as in the following

$$\begin{aligned} V^\star \coloneqq \sup_{\hat{G}\in\hat{\mathcal{G}}_f} W(a^{\mathrm{opt}})\,, \\ \text{s.t.} \quad \sum_{i\in N} u_i(a^{\mathrm{ne}}) - u_i(a_i^{\mathrm{opt}}, a_{-i}^{\mathrm{ne}}) \geq 0\,, \\ W(a^{\mathrm{ne}}) = 1\,, \end{aligned} \tag{10.16}$$

where the n equilibrium constraints (one per each player) have been substituted by their sum. Thus, $V^\star \geq W^\star$, but it also holds $V^\star \leq W^\star$ as Lemma 14 proves, so that $V^\star = W^\star$.

In the following we show how to transform (10.16) in (10.9) by introducing the variables $\theta(a, x, b)$, $(a, x, b) \in \mathcal{I}$. This parametrization has

[4] If, for a given game G, this is not the case, it is possible to construct a new game (by simply rescaling the value of the resources) such that $W(a^{\mathrm{ne}}) = 1$. Note that the new game has the same game price of anarchy of G.

been introduced to study covering problems in [War12], and will be used here to efficiently represent the quantities appearing in (10.16). To begin with, recall that each feasible set is composed of only two allocations, that is $\hat{\mathcal{A}}_i = \{a_i^{\text{ne}}, a_i^{\text{opt}}\}$. For any given triplet (a, x, b) in $\mathcal{I}$, we thus define $\theta(a, x, b) \in \mathbb{R}_{\geq 0}$ as the total value of resources that belong to precisely $a + x$ of the sets a_i^{ne}, $b + x$ of the sets a_j^{opt}, for which exactly x sets have the same index (i.e., $i = j$). These $\mathcal{O}(n^3)$ variables suffice to fully describe the terms appearing in (10.16). Indeed, extending the formulation of [War12] to the welfare defined in (8.1) and the utilities defined in (10.1), we can write

$$W(a^{\text{opt}}) = \sum_{(a,x,b)\in\mathcal{I}} \mathbb{1}_{\{b+x\geq 1\}} w(b+x)\theta(a,x,b)\,,$$
$$W(a^{\text{ne}}) = \sum_{(a,x,b)\in\mathcal{I}} \mathbb{1}_{\{a+x\geq 1\}} w(a+x)\theta(a,x,b)\,.$$

The relaxed equilibrium constraint

$$\sum_{i\in N} u_i(a^{\text{ne}}) - u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \geq 0$$

reduces to

$$\begin{aligned}
&\sum_{i\in N} u_i(a^{\text{ne}}) - u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \\
&= \sum_{(a,x,b)\in\mathcal{I}} [(a+x)f(a+x)w(a+x) - bf(a+x+1)w(a+x+1) \\
&\qquad\qquad - xf(a+x)w(a+x)]\theta(a,x,b) \\
&= \sum_{(a,x,b)\in\mathcal{I}} [af(a+x)w(a+x) - bf(a+x+1)w(a+x+1)]\theta(a,x,b) \geq 0\,.
\end{aligned}$$

Substituting the latter expressions in (10.16), one gets

$$\begin{aligned}
W^\star = &\sup_{\theta(a,x,b)} \sum_{a,x,b} \mathbb{1}_{\{b+x\geq 1\}} w(b+x)\theta(a,x,b) \\
\text{s.t.} &\sum_{a,x,b} [af(a+x)w(a+x) - bf(a+x+1)w(a+x+1)]\theta(a,x,b) \geq 0 \\
&\sum_{a,x,b} \mathbb{1}_{\{a+x\geq 1\}} w(a+x)\theta(a,x,b) = 1 \\
&\theta(a,x,b) \geq 0 \quad \forall (a,x,b) \in \mathcal{I}\,.
\end{aligned}$$

To transform the latter expression in (10.9) (i.e., the desired result) it suffices to show that the supremum is attained. To see this observe that the

decision variables $\theta(a, x, b)$ live in a compact space. Indeed $\theta(a, x, b)$ are constrained to the positive orthant for all $(a, x, b) \in \mathcal{I}$. Additionally, the decision variables with $a + x \neq 0$ must be bounded due to the constraint $W(a^{\text{ne}}) = 1$

$$\sum_{\substack{(a,x,b)\in\mathcal{I}\\ a+x\geq 1}} w(a+x)\theta(a,x,b) = 1\,,$$

where $w(j) \neq 0$ by assumption. Finally, the decision variables left, i.e., those of the form $\theta(0, 0, b)$, $b \in [n]$ are bounded due to the equilibrium constraint, which can be rewritten as

$$\sum_{b\in[n]} bf(1)w(1)\theta(0,0,b) \leq$$
$$\sum_{\substack{(a,x,b)\in\mathcal{I}\\ a+x\geq 1}} [af(a+x)w(a+x) - bf(a+x+1)w(a+x+1)]\theta(a,x,b),$$

where $f(1)w(1) \neq 0$ by assumption.

□

Lemma 13. *For any game $G \in \mathcal{G}_f$, it holds*

$$W(a^{\text{ne}}) > 0 \quad \textit{for all} \quad a^{\text{ne}} \in \text{ne}(G)\,.$$

Proof. Let us consider a fixed game $G \in \mathcal{G}_f$. By contradiction, let us assume that $W(a^{\text{ne}}) = 0$ for some $a^{\text{ne}} \in \text{ne}(G)$. It follows that all the players must have distributed themselves on resources that are either valued zero, or have selected the empty set allocation (since $w(j) > 0$). Thus, their utility function must also evaluate to zero. However, by Standing Assumptions, there exists a player p and a resource $r \in a_p \in \mathcal{A}_p$ with $v_r > 0$. Observe that no other player is currently selecting this resource, else $W(a^{\text{ne}}) > 0$. If player p was to deviate and selected instead a_p, his utility would be strictly positive (since $f(1) > 0$). Thus a^{ne} is not an equilibrium: a contradiction. Repeating the same reasoning for all games $G \in \mathcal{G}_f$ yields the claim. □

Lemma 14. *Consider $W^\star$ and $V^\star$ defined respectively in (10.15) and (10.16). It holds that $V^\star \leq W^\star$.*

Proof. Since (10.16) is equivalent to (10.9) as shown in the proof of Theorem 8, we will work with (10.9) to prove $V^\star \leq W^\star$. To do so, for any $\theta(a, x, b)$, $(a, x, b) \in \mathcal{I}$ feasible solution of (10.9) with value v, we will construct an instance of game $\hat{G}$ satisfying the constraints of the original problem (10.15) too. This allows to conclude that $V^\star \leq W^\star$. To ease the notation we will use $\sum_{a,x,b}$ in place of $\sum_{(a,x,b)\in\mathcal{I}}$.

Consider $\theta(a,x,b)$, $(a,x,b) \in \mathcal{I}$ a feasible point for (10.9) with value v. For every $(a,x,b) \in \mathcal{I}$ and for each $i \in N$ we create a resource $r(a,x,b,i)$ and assign to it the value of $\theta(a,x,b)/n$, i.e., $v_{r(a,x,b,i)} = \theta(a,x,b)/n \ \forall i \in N$. We then construct the game $\hat{G}$ by defining $\forall i \in N$, $\hat{\mathcal{A}}_i = \{a_i^{\text{ne}}, a_i^{\text{opt}}\}$ and assigning the resources as follows

$$\begin{aligned} a_i^{\text{ne}} &= \cup_{j=1}^{n} \left\{ r(a,x,b,j) \text{ s.t. } a + x \geq 1 + g(i,j) \right\}, \\ a_i^{\text{opt}} &= \cup_{j=1}^{n} \left\{ r(a,x,b,j) \text{ s.t. } b + x \geq 1 + h(i,j) \right\}, \end{aligned}$$

where

$$\begin{aligned} g(i,j) &\coloneqq (j - 1 + (n-1)(i-1)) \mod n, \\ &= (j - i) \mod n \\ h(i,j) &\coloneqq (j + (n-1)(i-1)) \mod n \\ &= (j - i + 1) \mod n. \end{aligned}$$

We begin by showing $W(a^{\text{ne}}) = 1$ and $W(a^{\text{opt}}) = v$. Aside from the cumbersome definition of g and h, it is not difficult to verify that for any fixed resource (i.e., for every fixed tuple (a,x,b,j)), there are exactly $a + x$ (resp. $b + x$) players selecting it while at the equilibrium (resp. optimum) allocation. It follows that

$$\begin{aligned} W(a^{\text{ne}}) &= \sum_{j \in [n]} \sum_{a+x>0} v_{r(a,x,b,j)} w(a+x) \\ &= \sum_{j \in [n]} \sum_{a+x>0} \frac{\theta(a,x,b)}{n} w(a+x) \\ &= \sum_{a,x,b} \mathbb{1}_{\{a+x \geq 1\}} w(a+x) \theta(a,x,b) = 1, \end{aligned}$$

With an identical reasoning, one shows that

$$\begin{aligned} W(a^{\text{opt}}) &= \sum_{j \in [n]} \sum_{b+x>0} v_{r(a,x,b,j)} w(b+x) \\ &= \sum_{a,x,b} \mathbb{1}_{\{b+x \geq 1\}} w(b+x) \theta(a,x,b) = v. \end{aligned}$$

Finally, we prove that a^{ne} is indeed an equilibrium, i.e., it satisfies $u_i(a^{\text{ne}}) - u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \geq 0$ for all $i \in N$. Towards this goal, we recall that the game under consideration is a congestion game with potential $\varphi : \mathcal{A} \to \mathbb{R}_{\geq 0}$

$$\varphi(a) = \sum_{r \in \mathcal{R}} \sum_{j=1}^{|a|_r} v_r w(j) f(j)$$

It follows that $u_i(a^{\text{ne}}) - u_i(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) = \varphi(a^{\text{ne}}) - \varphi(a_i^{\text{opt}}, a_{-i}^{\text{ne}})$ and so we equivalently prove that

$$\varphi(a^{\text{ne}}) - \varphi(a_i^{\text{opt}}, a_{-i}^{\text{ne}}) \geq 0 \quad \forall i \in N.$$

Thanks to the previous observation, according to which every resource (a, x, b, j) is covered by exactly $a + x$ players at the equilibrium, we have

$$\begin{aligned}\varphi(a^{\mathrm{ne}}) &= \sum_{j\in[n]} \sum_{a,x,b} \frac{\theta(a,x,b)}{n} \sum_{j=1}^{a+x} w(j) f(j) \\ &= \frac{1}{n} \sum_{a,x,b} n\,\theta(a,x,b) \sum_{j=1}^{a+x} w(j) f(j)\,.\end{aligned}$$

Additionally, observe that there are b resources selected by one extra agent and a resources selected by one less agent when moving from a^{ne} to $(a_i^{\mathrm{opt}}, a_{-i}^{\mathrm{ne}})$. The remaining resources are chosen by the same number of agents. It follows that

$$\begin{aligned}\varphi(a^{\mathrm{ne}}) - \varphi(a_i^{\mathrm{opt}}, a_{-i}^{\mathrm{ne}}) &= \frac{1}{n} \sum_{a,x,b} n\,\theta(a,x,b) \sum_{j=1}^{a+x} w(j) f(j) \\ &\quad - \frac{1}{n} \sum_{a,x,b} \theta(a,x,b) \left(b \sum_{j=1}^{a+x+1} w(j) f(j) + a \sum_{j=1}^{a+x-1} w(j) f(j) + (n-a-b) \sum_{j=1}^{a+x} w(j) f(j) \right) \\ &= \frac{1}{n} \sum_{a,x,b} \theta(a,x,b)\,(a\,w(a+x) f(a+x) - b\,w(a+x+1) f(a+x+1)) \geq 0\,,\end{aligned}$$

where the inequality holds because $\theta(a, x, b)$ is assumed feasible for (10.9). This concludes the proof. □

Proof of Theorem 9

Proof. We divide the proof in two steps. In the first step we write the dual of the original program in (10.9). With the second step we show that only the constraints obtained for $(a, x, b) \in \mathcal{I}_R$ are binding.

Step 1. Upon stacking the decision variables $\theta(a, x, b)$ in the vector $y \in \mathbb{R}^\ell$, $\ell = |\mathcal{I}|$, and after properly defining the coefficients c, d, $e \in \mathbb{R}^\ell$, the program (10.9) can be compactly written as

$$\begin{aligned} W^\star = \max_y\ & c^\top y \\ \text{s.t.}\quad & -e^\top y \leq 0\,, \quad (\lambda) \\ & d^\top y - 1 = 0\,, \quad (\mu) \\ & -y \leq 0\,. \quad (\nu)\end{aligned}$$

The Lagrangian function is defined for $\lambda \geq 0$, $\nu \geq 0$ as $\mathcal{L}(y, \lambda, \mu, \nu) = c^\top y - \lambda(-e^\top y) - \mu(d^\top y - 1) - \nu^\top(-y) = (c^\top + \lambda e^\top + \nu - \mu d^\top) y + \mu$, while the dual function reads as

$$g(\lambda, \mu, \nu) = \mu \quad \text{if} \quad c^\top + \lambda e^\top + \nu^\top - \mu d^\top = 0\,,$$

and it is unbounded elsewhere. Hence the dual program takes the form

$$\begin{aligned} \min_{\lambda\in\mathbb{R}_{\geq 0},\,\mu\in\mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & c+\lambda e-\mu d\leq 0\,, \end{aligned}$$

which corresponds, in the original variables, to

$$\begin{aligned} \min_{\lambda\in\mathbb{R}_{\geq 0},\,\mu\in\mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \mathbb{1}_{\{b+x\geq 1\}}w(b+x)-\mu\mathbb{1}_{\{a+x\geq 1\}}w(a+x)+ \\ & +\lambda[af(a+x)w(a+x)-bf(a+x+1)w(a+x+1)]\leq 0 \qquad \forall(a,x,b)\in\mathcal{I}\,. \end{aligned} \tag{10.17}$$

By strong duality[5], the value of (10.9) matches (10.17).

Step 2. In this step we show that only the constraints with $(a,x,b)\in\mathcal{I}_R$ are necessary in (10.17), thus obtaining (10.10).

Observe that when $(a,x,b)\in\mathcal{I}$ and $a+x=0$, b can take any value $1\leq b\leq n$, and these indices are already included in $\mathcal{I}_R$. Similarly for the indices $(a,x,b)\in\mathcal{I}$ with $b+x=0$. Thus, we focus on the remaining constraints, i.e., those with $a+x\neq 0$ and $b+x\neq 0$. We change the coordinates from the original indices (a,x,b) to (j,x,l), $j\coloneqq a+x$, $l\coloneqq b+x$. The constraints in (10.17) now read as

$$\begin{aligned} \mu w(j) &\geq w(l)+\lambda[(j-x)f(j)w(j)-(l-x)f(j+1)w(j+1)]\,, \\ &= w(l)+\lambda[jf(j)w(j)-lf(j+1)w(j+1)+x(f(j+1)w(j+1)-f(j)w(j))] \end{aligned} \tag{10.18}$$

where $(j,x,l)\in\hat{\mathcal{I}}$ and $\hat{\mathcal{I}}=\{(j,x,l)\in\mathbb{N}^3_{\geq 0}\text{ s.t. }1\leq j-x+l\leq n,\ j\geq x,\ l\geq x,\ j,l\neq 0\}$. In the remaining of this proof we consider j fixed, while l, x are free to move within $\hat{\mathcal{I}}$. This corresponds to moving the indices in the rectangular region defined by the blue and green patches in Figures 10.3 and 10.4. Observe that for $j=n$ it must be $l=x$ (since $-x+l\leq 0$ and $l-x\geq 0$), i.e., in the original coordinates $b=0$, which represents the segment on the plane $b=0$ with $a+x=n$. These indices already belong to $\mathcal{I}_R$. Thus, we consider the case $j\neq n$ and divide the reasoning in two parts.

a) Case of $f(j+1)w(j+1)\leq f(j)w(j)$.

The term $f(j+1)w(j+1)-f(j)w(j)$ is non-positive and so the most binding constraint in (10.18) is obtained picking x as small as possible. In the following we fix l as well (recall that we have previously fixed j). This corresponds to considering points on a black line on the plane $j=$const in Figure 10.3). Since it must be $x\geq 0$ and $x\geq j+l-n$, for fixed j and l we set $x=\max\{0,j+l-n\}$. In the following we show that these constraints are already included in (10.10).

[5]The primal LP (10.9) is always feasible, since $\theta(0,1,0)=1/w(1)$, $\theta(a,x,b)=0\,\forall\,(a,x,b)\in\mathcal{I}\backslash(0,1,0)$ satisfies all the constraints in (10.9).

- If $j + l \leq n$, i.e., if $a + b + 2x \leq n$, we set $x = 0$. These indices correspond to points on the plane $x = 0$, $(1 \leq a + b \leq n)$ bounding the pyramid and so they are already included in $\mathcal{I}_R$.

- If $j + l > n$, i.e., if $a + b + 2x > n$, we set $x = j + l - n$, i.e., $a + b + x = n$. These indices correspond to points on the plane $a + b + x = n$, and so they are included in $\mathcal{I}_R$ too.

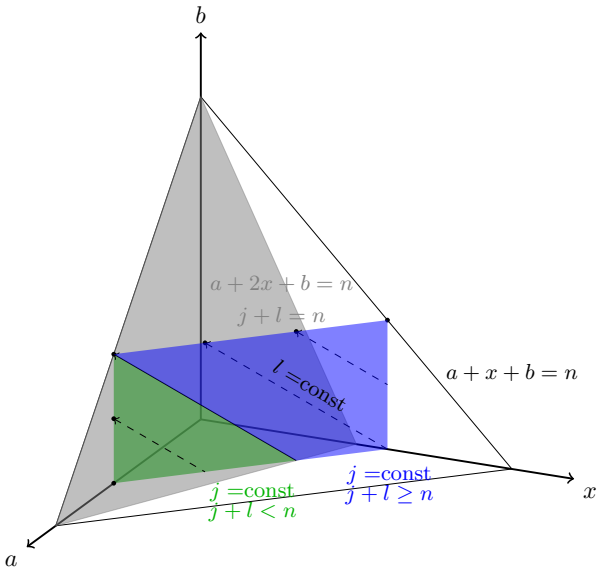

Figure 10.3: Indices representation for case a).

b) Case of $f(j+1)w(j+1) > f(j)w(j)$.
The term $f(j+1)w(j+1) - f(j)w(j)$ is positive and so the most binding constraint in (10.18) is obtained picking x as large as possible. In the following (after having fixed j) we fix l as well (this means we are moving on a black line on the plane j =const in Figure 10.4). Since it must be $x \leq l$, $x \leq j$ and $x \leq j + l - 1$, we set $x = \min\{j, l\}$. In the following we show that these constraints are already included in (10.10).

- If $j \leq l$, i.e., if $a \leq b$, we set $x = j$, i.e., $a = 0$. These indices correspond to points on the plane $a = 0$, $(1 \leq x + b \leq k)$ and so they are included in $\mathcal{I}_R$.

- If $j > l$, i.e., if $a > b$, then we set $x = l$, i.e., $b = 0$. These indices correspond to points on the plane $b = 0$, $(1 \leq a + b \leq k)$ and so they are included in $\mathcal{I}_R$.

□

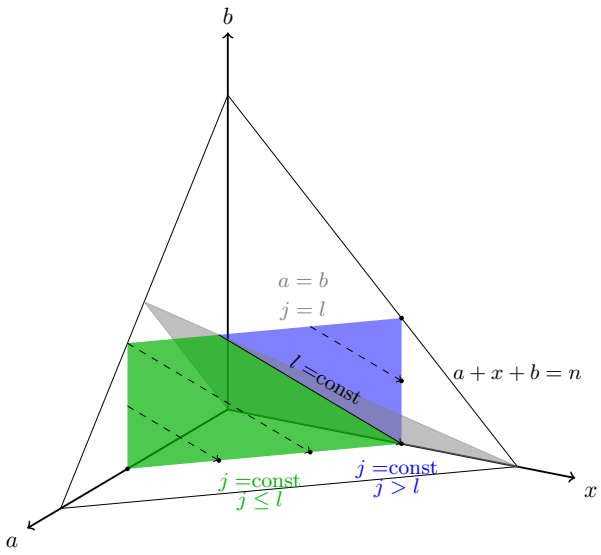

Figure 10.4: Indices representation for case b).

Proof of Corollary 3

Proof.

(a) Following the proof of Theorem 9 (second step, case a)), we note that if $f(j)w(j)$ is non increasing for $j \in N$, the only binding indices are those lying on the the two surfaces $x = 0$, $1 \leq a + b \leq n$ and $a + x + b \leq n$. The surface $x = 0$, $1 \leq a + b \leq n$ gives

$$\mu w(j) \geq w(l) + \lambda[jf(j)w(j) - lf(j+1)w(j+1)] \tag{10.19}$$

for $1 \leq j + l \leq n$ and $j, l \in [n]_0$, where we used j, l instead of a, b. The surface $a + x + b = n$ gives

$$\mu w(n-b) = \lambda[af(n-b)w(n-b) - bf(n-b+1)w(n-b+1)] + w(n-a)$$

which can be written as

$$\mu w(j) \geq w(l) + \lambda[(n-l)f(j)w(j) - (n-j)f(j+1)w(j+1)] \tag{10.20}$$

for $j + l > n$ and $j, l \in [n]_0$, where we have used the same change of coordinates of the proof of Theorem 9, i.e., $j = a + x = n - b$, $l = b + x = n - a$. Thus, we conclude that (10.19) and (10.20) are sufficient to describe the constraints in (10.10), and the result follows.

(b) First, observe that for $j = 0$, it must be $l \in [n]$. Additionally, note that the second set of constraints (those with $j + l > n$) is empty. The first set of constraints yields $\lambda \geq \frac{w(l)}{l} \frac{1}{f(1)w(1)}$ for $l \in [n]$. Define

$$\lambda^\star = \max_{l \in [n]} \frac{w(l)}{l} \frac{1}{f(1)w(1)},$$

and observe that any feasible λ must satisfy $\lambda \geq \lambda^\star$. Second, observe that for $l = 0$, it must be $j \in [n]$. Additionally, the second set of constraints (those with $j + l > n$) is empty. The first set of constraints yields $\mu \geq \lambda j f(j)$ for $j \in [n]$.

In the following we show that the most binding constraint amongst all those in (10.11) is of the form $\mu \geq \alpha\lambda + \beta$, with $\alpha \geq 0$ (i.e., the most binding constraint is a straight line in the (λ, μ) plane pointing north-east). Consequently, the best choice of λ so as to satisfy the constraints and minimize μ is to select λ as small as possible, i.e., $\lambda = \lambda^\star$. See Figure 10.5 for an illustrative plot.

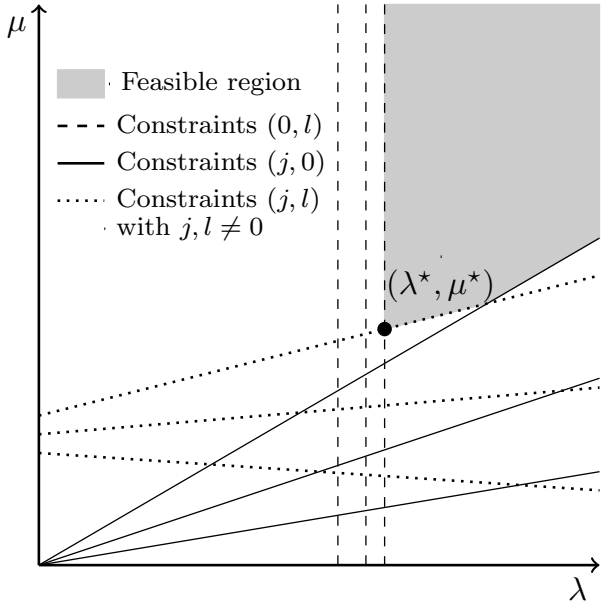

Figure 10.5: Illustration of the three classes of constraints used in the proof of Corollary 3.

As shown previously, the constraints with $j = 0$ are straight lines parallel to the μ axis, while the constraints with $l = 0$ are straight line of the form $\mu \geq \lambda j f(j)$ (and thus point north-east in the (λ, μ) plane). We are thus left to check the constraints with $j \neq 0$ and $l \neq 0$.

To do so, we prove that if one such constraint (identified by the indices (j, l)) has negative slope, the constraint identified with $(j, 0)$ is more binding. Since the constraint $(j, 0)$ is of the form $\mu \geq j\lambda j f(j)$ (and thus has non-negative slope), this will conclude the proof. We split the reasoning depending on wether $1 \leq j + l \leq n$ or $j + l > n$ as the constraints in (10.11) have a different expression.

- Case of $1 \leq j + l \leq n$: to complete the reasoning, in the following we assume that $j f(j) - l f(j+1)\frac{w(j+1)}{w(j)} < 0$, and show that the constraint $(j, 0)$ is more binding, i.e., that

$$\lambda j f(j) \geq \frac{w(l)}{w(j)} + \lambda j f(j) - \lambda l f(j+1)\frac{w(j+1)}{w(j)},$$

which is equivalent to showing

$$\frac{w(l)}{w(j)} - \lambda l f(j+1)\frac{w(j+1)}{w(j)} \leq 0\,. \tag{10.21}$$

Since $jf(j) - lf(j+1)\frac{w(j+1)}{w(j)} < 0$ it must be

$$l > j\frac{f(j)w(j)}{f(j+1)w(j+1)} \geq j\,,$$

by non-increasingness of $f(j)w(j)$. Thus it must be $l \geq j+1$. Consequently, by non-increasingness of $f(j)w(j)$ it is $w(l) \leq f(j+1)w(j+1)/f(l)$ and we can bound the left hand side of (10.21) as

$$\begin{aligned}
&\frac{w(l)}{w(j)} - \lambda l f(j+1)\frac{w(j+1)}{w(j)} \\
&\leq \frac{f(j+1)w(j+1)}{w(j)f(l)} - \lambda l\frac{f(j+1)w(j+1)}{w(j)} \\
&= \frac{f(j+1)w(j+1)}{f(j)w(j)}\left(\frac{1}{f(l)} - \lambda l\right) f(j)\,.
\end{aligned}$$

The claim (10.21) is shown upon noticing that $f(l) \geq \frac{1}{l}\min_{l\in[n]}\frac{l}{w(l)}f(1)w(1) = \frac{1}{l\lambda^\star}$ (by assumption), and thus

$$\frac{f(j+1)w(j+1)}{f(j)w(j)}\left(\frac{1}{f(l)} - \lambda l\right) f(j) \leq \frac{f(j+1)w(j+1)}{f(j)w(j)}\left(\lambda^\star - \lambda\right) l f(j) \leq 0\,,$$

since we have already shown that $\lambda \geq \lambda^\star$ for every feasible λ.

- Case of $j + l > n$: to complete the proof we proceed in a similar fashion to what seen in the previous case. In particular, we assume that $(n-l)f(j) - (n-j)f(j+1)\frac{w(j+1)}{w(j)} < 0$, and show that the constraints $(j, 0)$ is more binding, i.e., that

$$\frac{w(l)}{w(j)} + \lambda(n-l-j)f(j) - \lambda(n-j)f(j+1)\frac{w(j+1)}{w(j)}) \leq 0\,. \tag{10.22}$$

Since $(n-l)f(j) - (n-j)f(j+1)\frac{w(j+1)}{w(j)} < 0$, it must be

$$n - j > (n-l)\frac{f(j)w(j)}{f(j+1)w(j+1)} \geq n - l$$

by non-increasingness of $f(j)w(j)$. Thus it must be $l \geq j+1$. Consequently, by non-increasingness of $f(j)w(j)$ we can bound the left hand side of (10.22)

as

$$\begin{aligned}
&\frac{w(l)}{w(j)} + \lambda(n-l-j)f(j) - \lambda(n-j)f(j+1)\frac{w(j+1)}{w(j)} \\
&\leq \frac{f(j+1)w(j+1)}{w(j)f(l)} + \lambda(n-l-j)f(j) - \lambda(n-j)f(j+1)\frac{w(j+1)}{w(j)} \\
&= \frac{f(j+1)w(j+1)}{w(j)f(j)}\left(\frac{1}{f(l)} - \lambda(n-j)\right)f(j) + \lambda(n-l-j)f(j) \\
&\leq \frac{f(j+1)w(j+1)}{w(j)f(j)}\left(\frac{1}{f(l)} - \lambda(n-j) + \lambda(n-l-j)\right)f(j) \\
&\leq \frac{f(j+1)w(j+1)}{w(j)f(j)}\left(\frac{1}{f(l)} - \lambda l\right)f(j) \leq 0\,,
\end{aligned}$$

where the chain of inequality is proven similarly to what done in the case of $1 \leq j + l \leq n$, using the non-decreasingness of $f(j)w(j)$ and the fact that $f(l) \geq \frac{1}{l\lambda^\star}$ by assumption.

□

10.5.3 Proofs of the results presented in Section 10.4

Proof of Theorem 10

Proof. We first observe that the problem $\arg\max_{f\in\mathcal{F}} \mathrm{PoA}(f)$ is well posed, in the sense that the supremum $\sup_{f\in\mathcal{F}} \mathrm{PoA}(f)$ is attained for some $f \in \mathcal{F}$. A proof of this is reported in the following Lemma 15.

The (well posed) problem $\arg\max_{f\in\mathcal{F}} \mathrm{PoA}(f)$ is equivalent to finding the distribution rule that minimizes $W^\star$ given in Theorem 9, i.e.,

$$\begin{aligned}
f^\star \in \arg\min_{f\in\mathcal{F}} \min_{\lambda\in\mathbb{R}_{\geq 0},\mu\in\mathbb{R}} \quad & \mu \\
\text{s.t. } & \mathbb{1}_{\{b+x\geq 1\}}w(b+x) - \mu\mathbb{1}_{\{a+x\geq 1\}}w(a+x) + \\
& + \lambda[af(a+x)w(a+x) - bf(a+x+1)w(a+x+1)] \leq 0 \qquad \forall (a,x,b) \in \mathcal{I}_R\,.
\end{aligned}$$

The previous program is non linear, but the decision variables λ and f always appear multiplied together. Thus, we define $\tilde{f}(j) \coloneqq \lambda f(j)$ for all $j \in [n+1]_0$ and observe that the constraint obtained in (10.10) for $(a,x,b) = (0,0,1)$ gives $\tilde{f}(1) = \lambda f(1) \geq 1$, which also implies $\lambda \geq 1/f(1) > 0$ since $f(1) > 0$ (by assumption of $f \in \mathcal{F}$). Folding the min operators gives

$$\begin{aligned}
\tilde{f}^\star \in \underset{\substack{\tilde{f}\in\mathbb{R}^n_{\geq 0},\mu\in\mathbb{R} \\ \tilde{f}(1)\geq 1}}{\arg\min} \quad & \mu \\
\text{s.t. } & \mathbb{1}_{\{b+x\geq 1\}}w(b+x) - \mu\mathbb{1}_{\{a+x\geq 1\}}w(a+x) + \\
& + a\tilde{f}(a+x)w(a+x) - b\tilde{f}(a+x+1)w(a+x+1) \leq 0 \qquad \forall (a,x,b) \in \mathcal{I}_R\,.
\end{aligned} \tag{10.23}$$

Finally, observe that $\tilde{f}^\star$ is feasible for the original program, since $\tilde{f}^\star \in \mathcal{F}$. Additionally, we note that $\tilde{f}^\star$ and $f^\star$ give the same price of anarchy (since $\tilde{f}(j) = \lambda f(j)$, $\lambda > 0$ and the equilibrium conditions are invariant to rescaling). Thus $\tilde{f}^\star$ solving (10.23) must be optimal. The optimal price of anarchy value follows. □

Lemma 15. *The supremum* $\sup_{f\in\mathcal{F}} \mathrm{PoA}(f)$ *is attained in* $\mathcal{F}$.

Proof. Recall that $\mathcal{F}$ is defined as follows

$$\mathcal{F} \coloneqq \{f : [n] \to \mathbb{R}_{\geq 0} \text{ s.t. } f(1) \geq 1,\ f(j) \geq 0\ \forall j \in [n]\}.$$

To conclude, we show that any distribution $f^\star$ achieving a performance equal to

$$\sup_{f\in\mathcal{F}} \mathrm{PoA}(f)$$

is bounded (i.e., all the components are bounded), so that it must be $f^\star \in \mathcal{F}$. To do so, consider a fixed distribution $f \in \mathcal{F}$, and construct from it f_M. The distribution f_M is defined as follows: $f_M(j) = M$, with $M \in \mathbb{R}_{\geq 0}$ for some fixed $j \in [n]$, while it exactly matches f for the remaining components. In the following we show that there exists $\hat{M} \in \mathbb{R}_{\geq 0}$ such that $\mathrm{PoA}(f_M) < \mathrm{PoA}(f)$ for all $M \geq \hat{M}$. Thus f_M can not attain $\sup_{f\in\mathcal{F}} \mathrm{PoA}(f)$ for $M \geq \hat{M}$ as the corresponding f would give a better price of anarchy. Repeating this reasoning for any $f \in \mathcal{F}$, one concludes that the distribution rule achieving $\sup_{f\in\mathcal{F}} \mathrm{PoA}(f)$ must be bounded along the j-th component. Repeating the reasoning for all possible $j \in [n]$, one obtains the claim.

To conclude we are left to show that $\exists \hat{M} \in \mathbb{R}_{\geq 0}$ such that $\mathrm{PoA}(f_M) < \mathrm{PoA}(f)$ for all $M \geq \hat{M}$. To do so, observe that the price of anarchy of $f \in \mathcal{F}$ can be computed as $\mathrm{PoA}(f) = 1/W^\star$, where $W^\star$ is the solution to the primal problem in (10.9). Since the decision variables of (10.9) live in a compact space (and the primal is feasible, see the footnote in the proof of Theorem 9), we have $W^\star < +\infty$ and so $\mathrm{PoA}(f) > 0$, i.e., $\mathrm{PoA}(f)$ is bounded away from zero. On the other hand, thanks to Theorem 9, the price of anarchy of f_M can be computed for any M as $\mathrm{PoA}(f_M) = 1/W^\star_M$, where

$$\begin{aligned} W^\star_M = \min_{\lambda\in\mathbb{R}_{\geq 0},\, \mu\in\mathbb{R}}\ & \mu \\ \text{s.t. } & \mathbb{1}_{\{b+x\geq 1\}} w(b+x) - \mu \mathbb{1}_{\{a+x\geq 1\}} w(a+x) + \\ & + \lambda[a f_M(a+x) w(a+x) - b f_M(a+x+1) w(a+x+1)] \leq 0 \qquad \forall (a,x,b) \in \mathcal{I}_R \end{aligned}$$

First, observe that for any feasible λ, it must be $\lambda \geq \frac{1}{f_M(1)}$, else the constraints obtained form the previous linear program with $a = x = 0$, $b = 1$ would be infeasible. Further, consider the constraints with $b = 0$, $x = 0$, $a = j \geq 1$. They amount to

$$\mu \geq \lambda j f_M(j) \geq \frac{j}{f_M(1)} f_M(j) = \frac{jM}{f_M(1)},$$

where $f_M(1) > 0$ by Standing Assumptions and $f_M \in \mathcal{F}$. It follows that

$$\mathrm{PoA}(f_M) = \frac{1}{W_M^\star} \leq \frac{f_M(1)}{jM}\,.$$

Thus, it is possible to make $\mathrm{PoA}(f_M)$ arbitrarily close to zero, by selecting M sufficiently large. It follows that $\exists \hat{M} \in \mathbb{R}_{\geq 0}$ such that $\mathrm{PoA}(f_M) < \mathrm{PoA}(f)$ for all $M \geq \hat{M}$, since $\mathrm{PoA}(f)$ is bounded away from zero, as previously argued. This concludes the proof. □

CHAPTER 11

Submodular, supermodular, covering problems

In the previous chapter we have addressed the problem of characterizing and optimizing the price of anarchy as a function of the chosen utilities. In this chapter we specialize the general result of Theorems 9 and 10 to the case when W is monotone submodular, supermodular, or a coverage function. We show how previously fragmented results from other authors can now be obtained as special case of the more general Theorem 9.

Relative to the submodular case, in Section 11.1 we give an explicit expression for the price of anarchy (Theorem 11), and apply the result to obtain the efficiency of the Shapley value and marginal contribution distribution rule (Corollary 4). This is, to the best of our knowledge, the first exact characterization of the price of anarchy in the submodular settings, and the first exact characterization of the performance associated to the Shapley value and marginal contribution distribution rule. Additionally, we show how the distribution rule designed maximizing the price of anarchy outperforms the very recent $1-c/e$ approximation of [SVW17], relative to submodular maximization problems.

In Section 11.2 we consider the special case of MMC problems (see Section 8.2 for their definition) and obtain a tight expression (Theorem 12) for the price of anarchy relying solely on the Standing Assumptions. We further show how the expression subsumes previous results obtained under the additional assumptions therein required (Corollary 5). The distribution rule designed to maximize the price of anarchy achieves a $1 - 1/e$ approximation.

In Section 11.3 we consider the case when W is supermodular and obtain an explicit expression for the price of anarchy (Theorem 13), extending previous results. Finally, we show that the Shapley value distribution rule is optimal, but observe that the utility design approach provides very poor approximation guarantees limitedly to this case.

Throughout this chapter we assume that the Standing Assumptions introduced in Chapter 10 continue to hold. All the proofs are reported in the Appendix (Section 11.4). The results presented in this chapter have been published in [PCM18; PM18b].

11.1 The case of submodular welfare function

In this section we focus on the case when the welfare basis function w is non-decreasing and concave (in the discrete sense). This results in the welfare function in (8.1) being monotone submodular. Submodular functions model problems with diminishing returns and are used to describe a wide range of engineering applications such as satellite assignment problems [QBL15], Adwords for e-commerce [DJ12], and combinatorial auctions [LLN06], among others. For the considered class of problems, we show (Theorem 11) that characterizing the price of anarchy reduces to computing the maximum between $n(n+1)/2 \sim \mathcal{O}(n^2)$ numbers. Using this result, we give an explicit expression of the price of anarchy for the well known Shapley value and marginal contribution distribution rule (Corollary 4). We then show how to design f so as to maximize the performance measured by $\mathrm{PoA}(f)$. Finally, we compare our performance certificates with existing approximation results.

We begin by formally introducing two distribution rules that have attracted the researchers' attention due to their simple interpretation and to their special properties: the Shapley value distribution rule and the marginal contribution distribution rule [FH13].

Definition 24. *The Shapley value and marginal contribution distribution rules are identified with f_{SV} an f_{MC}, respectively. For $j \in [n]$, they are given by*

$$f_{\mathrm{SV}}(j) = \frac{1}{j}\,,$$
$$f_{\mathrm{MC}}(j) = 1 - \frac{w(j-1)}{w(j)}\,.$$

Observe that the Shapley value distribution rule is the only distribution rule for which the sum of all the players utility exactly matches the total welfare. The marginal contribution distribution rule takes its name from the observation that (10.1) reduces to

$$\begin{aligned} u_i(a) &= \sum_{r \in a_i} v_r w(|a|_r) f_{\mathrm{MC}}(|a|_r) \\ &= \sum_{r \in a_i} v_r (w(|a|_r) - w(|a|_r - 1)) = W(a) - W(\emptyset, a_{-i})\,, \end{aligned}$$

i.e., player's i utility function represent its marginal contribution to the total welfare, that is the difference between $W(a)$ and the welfare generated when player i is removed from the game.

Assumption 10. *Throughout this section we assume that the function w is non-decreasing and concave, in the following sense*

$$w(j+1) \geq w(j)\,,$$
$$w(j+1) - w(j) \leq w(j) - w(j-1) \quad \forall j \in [n-1]\,.$$

Further we assume that $w(1) = 1$.

The requirement $w(1) = 1$ is without loss of generality. Indeed, If $w(1) \neq 1$, it is possible to normalize its value and reduce to the case $w(1) = 1$ (since $w(1) > 0$ by Standing Assumptions). As a consequence of Assumption 10, the function $W(a)$ is monotone and submodular, i.e., it satisfies the following:

Monotonicity:

$$\forall\, a, b \in \mathcal{A} \text{ s.t. } a_i \subseteq b_i \;\forall i \in N \implies W(a) \leq W(b)\,.$$

Submodularity:

$$\begin{aligned}&\forall\, a, b \in \mathcal{A} \text{ s.t. } a_i \subseteq b_i \;\forall i \in N,\\ &\forall\, c \in 2^{\mathcal{R}n} \text{ s.t. } a'_i := a_i \cup c_i \in \mathcal{A}_i,\; b'_i := b_i \cup c_i \in \mathcal{A}_i \;\forall i \in N,\\ &\qquad\implies W(a') - W(a) \geq W(b') - W(b)\,.\end{aligned}$$

While Theorem 9 gives a general answer on how to determine the price of anarchy, it is possible to exploit the additional properties given by Assumption 10 to obtain an explicit expression of $\mathrm{PoA}(f)$.

Theorem 11 (PoA for submodular welfare)**.** *Consider f a distribution rule such that $f(j)w(j)$ is non increasing and $f(j) \geq f_{\mathrm{MC}}(j)$ for all $j \in [n]$. Then,* $\mathrm{PoA}(f) = 1/W^\star$*,*

$$W^\star = \max_{l \leq j \in [n]} \left\{ \frac{w(l)}{w(j)} + \min(j, n-l) f(j) - \min(l, n-j) f(j+1) \frac{w(j+1)}{w(j)} \right\}, \tag{11.1}$$

or equivalently

$$\begin{aligned} W^\star = &\min_{\mu \in \mathbb{R}} \mu \\ &s.t.\;\; \mu w(j) \geq w(l) + j f(j) w(j) - l f(j+1) w(j+1) \\ &\qquad\qquad \forall j, l \in [n]_0 \;\; s.t. \;\; j \geq l \;\; \textit{and} \;\; 1 \leq j + l \leq n, \\ &\qquad \mu w(j) \geq w(l) + (n-l) f(j) w(j) - (n-j) f(j+1) w(j+1) \\ &\qquad\qquad \forall j, l \in [n]_0 \;\; s.t. \;\; j \geq l \;\; \textit{and} \;\; \quad j + l \geq n\,. \end{aligned} \tag{11.2}$$

The proof amounts to showing that λ appearing in Corollary 3 can be computed a priori, and takes the value $\lambda^\star = 1$. The requirements on $f(j)w(j)$ being non increasing and $f(j) \geq f_{\mathrm{MC}}(j)$ might seem restrictive at first. Nevertheless, similar assumptions where made in [MR14; Gai09] relative to a simpler class of problems. Additionally, the Shapley value and marginal contribution distribution rules (and many others) satisfy these assumptions. Thus, a direct application of Theorem 11 returns the exact price of anarchy of f_{SV} and f_{MC}, as detailed next.

Corollary 4 (Tight PoA for f_{SV} and f_{MC})**.**

(a) The PoA *for the Shapley value distribution rule is* $\mathrm{PoA}(f_{\mathrm{SV}}) = 1/W^\star_{\mathrm{SV}}$*, where*

$$W^\star_{\mathrm{SV}} = \max_{l \leq j \in [n]} \left\{ \frac{w(l)}{w(j)} + \min(j, n-l) \frac{1}{j} - \min(l, n-j) \frac{w(j+1)}{(j+1)w(j)} \right\}. \tag{11.3}$$

(b) The PoA *for the marginal contribution distribution rule is* $\mathrm{PoA}(f_{\mathrm{MC}}) = 1/W^{\star}_{\mathrm{MC}}$, *where*

$$W^{\star}_{\mathrm{MC}} = 1 + \max_{j\in[n]}\left\{\frac{1}{w(j)}\min(j, n-j)[2w(j) - w(j-1) - w(j+1)]\right\} \tag{11.4}$$

The previous Corollary shows that the price of anarchy of the Shapley value and marginal contribution distribution rule can be computed as the maximum of $n(n+1)/2$ and n numbers, respectively.

Remark 13 (Connection with [MR14])**.** *The quantity* (11.3) *can be equivalently written as*

$$W^{\star}_{\mathrm{SV}} = 1 + \max_{l\le j\in[n]}\left\{\frac{w(l)}{w(j)} - \frac{1}{j}[\max\{j+l-n, 0\} + \min\{l, n-j\}\beta(j)]\right\}, \tag{11.5}$$

$$\text{where}\quad \beta(j) \coloneqq \frac{j}{j+1}\frac{w(j+1)}{w(j)}\,.$$

The previous expression partially matches the result in [MR14, Thm. 6], where the authors used a different approach to obtain a bound on the price of anarchy for the larger class of coarse correlated equilibria, but limitedly to f_{SV} *and singleton problems. More precisely, [MR14, Thm. 6] provides a bound of the price of anarchy relative to* f_{SV}*, as the minimum between two expression. While their first expression exactly matches* (11.5)*, the second one is not present here. Nevertheless, it is possible to show that such additional expression is redundant, as the first one is always the most constraining.*[1] *This allows us to conclude that the bound obtained in [MR14, Thm. 6] precisely matches the one in* (11.5)*. Additionally, since our result is provably tight for the class of Nash equilibria, and the result in [MR14] provides a lower bound for CCE, such bound is tight as well (in the set of CCE) and the worst performing coarse correlated equilibrium is, simply, a pure Nash equilibrium.*

For the submodular welfare case considered here, it is still possible to determine the distribution rule $f^{\star}$ that maximizes $\mathrm{PoA}(f)$ as the solution of a tractable linear program either directly employing the more general result in Theorem 10 or using the following linear program derived from (11.2), which additionally constrains the admissible distri-

[1] This statement is not formally shown here, in the interest of space. Its proof amounts to showing that the second expression appearing in [MR14, Thm. 6] is always upper bounded by (11.5), thanks to the concavity of w.

butions f to satisfy $f(j) \geq f_{\text{MC}}(j)$ and $f(j)w(j)$ to be non increasing,

$$
\begin{aligned}
f^\star \in & \underset{f \in \mathcal{F}_s,\, \mu \in \mathbb{R}}{\arg\min}\ \mu \\
\text{s.t. } & \mu w(j) \geq w(l) + j f(j) w(j) - l f(j+1) w(j+1) \\
& \qquad \forall j, l \in [n]_0 \ s.t.\ j \geq l \ \text{ and } \ 1 \leq j + l \leq n, \\
& \mu w(j) \geq w(l) + (n-l) f(j) w(j) - (n-j) f(j+1) w(j+1) \\
& \qquad \forall j, l \in [n]_0 \ s.t.\ j \geq l \ \text{ and } \quad j + l \geq n\,,
\end{aligned}
\tag{11.6}
$$

where $\mathcal{F}_s = \{f \in \mathcal{F} \,|\, f(j) \geq f_{\text{MC}}(j),\ f(j+1)w(j+1) \leq f(j)w(j)\ \forall j \in [n]\}$. Extensive numerical simulations have shown that both these approaches return the same optimal value, so that the additional constraints $f \in \mathcal{F}_s$ required in (11.6) do not rule out the optimal distribution derived solving the linear program in Theorem 10. This statement can be formally proved, for example, by showing that any distribution rule satisfying the *Karush-Kuhn-Tucker* (KKT) system of the LP in (11.6) is also a solution to the KKT system of the LP in Theorem 10. We do not further purse this direction here.

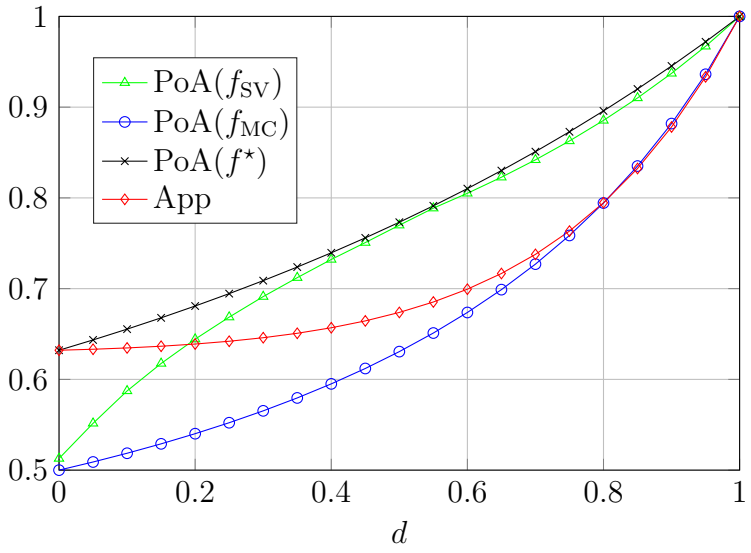

Figure 11.1: Comparison between the approximation ratio (11.11) and the price of anarchy of the optimal distribution rule $f^\star$ (determined as the solution of the LP in Theorem 10), Shapley value f_{SV} and marginal contribution f_{MC} distribution rules. The problems considered features $|N| \leq n = 20$ agents and a welfare basis of the form $w(j) = j^d$ with $d \in [0, 1]$ represented over the x-axis.

Figure 11.1 compares the price of anarchy (and thus the approximation ratio of any algorithm capable of computing a Nash equilibrium) of the Shapley value, marginal contribution and optimal distribution rule $f^\star$, in the case when $w(j) = j^d$ with $d \in [0, 1]$, $|N| \leq 20$. They have been computed using respectively (11.3), (11.4), where $f^\star$ has been

determined as the solution to the LP appearing in Theorem 10. For values of $d \in [0.5, 1]$ the Shapley value distribution rule performs close to the optimal, but its performance degrades for $d \in [0, 0.5]$ and for $d = 0$ it reaches the lower bound of $1/2$, as predicted for the class of valid utility games defined in [Vet02, Thm. 5]. The marginal contribution rule instead, performs the worst amongst the considered distribution rules. While $f^\star$ will always perform better or equal than any other distribution, it is unclear if, and to what extent, f_{SV} outperforms f_{MC} in the general settings. The expressions in (11.3) and (11.4) can nevertheless be used to provide an answer to this question.

11.1.1 Improved approximation and comparison with existing result

In this section we compare the approximation guarantees offered by the utility design approach with other recent result in the maximization of submodular functions.

A monotone submodular maximization problem is defined as follows. We are given a set X, and a collection of subsets $S \subseteq 2^X$. Given a monotone and submodular set function $g : 2^X \to \mathbb{R}_{\geq 0}$, the objective is to find a set $s \in S$ maximizing g. If the collection of subsets S is a matroid, we term the problem a monotone submodular maximization problem subject to matroid constraints. For the latter class of problems, the best approximation ratio achievable in polynomial time has been very recently shown to be [SVW17]

$$1 - \frac{c}{e}, \tag{11.7}$$

where c represents the curvature of the welfare function and e the Euler's number. The curvature is formally defined as [CC84]

$$c \coloneqq 1 - \min_{\substack{e \in X \\ g(\{e\}) - g(\emptyset) \neq 0}} \frac{g(X) - g(X \setminus \{e\})}{g(\{e\}) - g(\emptyset)} \tag{11.8}$$

For this class of problems, no polynomial time algorithm can do better than (11.7) on all instances, even if the matroid is the uniform matroid, i.e., in the case of cardinality constraints [SVW17]. The GMMC problems studied here differs from the problem of maximizing a submodular function subject to matroid constraints, in that we are given not one, but n collections of sets. Thus, to compare the approximation results, in the following we restrict to GMMC problems where $\mathcal{A}_i = \mathcal{A}_j = \bar{\mathcal{A}} \subseteq 2^{\mathcal{R}}$. We allow for some set to appear multiple times in $\bar{\mathcal{A}}$ so as to cover the case when different agents select the same set. The objective is to select n subsets from $\bar{\mathcal{A}}$ so as to maximize W as defined in (8.1). The problem can be transformed in a monotone submodular maximization problem subject to cardinality constraints. To do so, let us enumerate all the subsets as in $\bar{\mathcal{A}} = \{A_1, \ldots, A_k\}$. We set $X \coloneqq [k]$, $S = 2^{[k]}$ and identify with $s = (s_1, \ldots, s_l) \in [k]^l$

an element of S (note that $l \leq k$). We define $g : 2^X \to \mathbb{R}_{\geq 0}$ for any $s \in S$ as

$$g(s) \coloneqq \sum_{r \in (\cup_j A_{s_j})} v_r w(|s|_r)\,, \tag{11.9}$$

where $|s|_r = |\{i \text{ s.t. } r \in A_{s_i}\}|$. Selecting n subsets ($n \leq k$) from $\bar{\mathcal{A}}$ to maximize $W(a)$ is then equivalent to solving

$$\max_{s \in S,\, |s| \leq n} g(s)\,. \tag{11.10}$$

The problem in (11.10) belongs to the class of monotone submodular maximization subject to cardinality constraints. Indeed, $g(s)$ is monotone and submodular due to Assumption 11. Additionally the number of elements in s is constrained to be less or equal to n. Thus, the approximation ratio (11.7) holds for (11.10). The curvature can be determined using (11.8), and amounts to $c = 1 + w(n-1) - w(n)$.[2] In Figure 11.1 we plot the approximation ratio (11.7) for the class of problems considered here, with the choice of $w(j) = j^d$, i.e., we plot (red curve) the quantity

$$\text{App} = 1 - \frac{1 + w(n-1) - w(n)}{e}\,, \tag{11.11}$$

for $d \in [0, 1]$. We observe that the optimal distribution rule $f^\star$ outperforms (11.11) for different values of d, so that, when there exists an algorithm capable of computing a Nash equilibrium in polynomial time (see Proposition 17), the approach presented here gives improved guarantees compared to (11.7).

Remark 14. *It is important to note that this is not in contradiction with the inapproximability result presented in [SVW17], as we are not solving a general submodular maximization problem, but the welfare function in* (11.9) *has a* special form.

11.2 Covering problems

In this section we specialize the previous results to the case of multiagent weighted maximum coverage (MMC) problems introduced in Section 8.3, a generalization of weighted maximum coverage problems. In a MMC problem we are given a ground set of elements $\mathcal{R}$ and n collections of subsets of the ground sets: $\mathcal{A}_i$ for $i \in N$. The goal is to select n subsets, one from each collection, so as to maximize the *total value* of covered elements. The corresponding welfare is

$$W(a) = \sum_{r \in \cup_{i \in N} a_i} v_r\,,$$

[2] This value of the curvature constitutes the worst case value amongst all possible GMMC problems introduced in Section 8.2. The reason to consider the worst case curvature over all problem instances (i.e., over all possible choices of $\bar{\mathcal{A}}$ and $\{v_r\}_{r \in \mathcal{R}}$) is that we wish to compare the approximation ratio (11.7) with the game theoretic approximation, and the latter gives a certificate over *all possible instances*.

which is obtained with the choice of $w(j) = 1$ for all j in (8.1) and (10.1). MMC problems are a subclass of GMMC submodular problems (they satisfy Assumption 10), and are used to model engineering problems such as vehicle-target assignment [AMS07] and sensor allocation problems [MW08]. Due to their importance in the applications, we treat their study separately.

Relative to MMC problems, we provide a general expression for the price of anarchy as a function of f (Theorem 12) and show how this reduces to the results obtained in [Gai09; RPM17], under the additional assumptions therein required.

Theorem 12 (PoA for multiagent maximum coverage)**.** *Consider MMC problems, i.e., fix $w(j) = 1\ \forall j \in [n]$. The price of anarchy is* $\mathrm{PoA}(f) = 1/W^\star$ *where*

$$W^\star = 1 + \max_{j\in[n-1]} \{(j+1)f(j+1) - 1, jf(j) - f(j+1), jf(j+1)\}, \tag{11.12}$$

or equivalently

$$\begin{aligned} W^\star = &\min_{\mu\in\mathbb{R}} \mu \\ &s.t.\ \mu \geq (j+1)f(j+1) \\ &\qquad \mu \geq 1 + jf(j) - f(j+1) \\ &\qquad \mu \geq 1 + jf(j+1) \qquad\qquad \forall j \in [n-1]. \end{aligned} \tag{11.13}$$

The previous theorem gives a simple and explicit way to compute the price of anarchy (10.3) as the maximum between $3(n-1)$ numbers. Observe that no assumptions are required other than the Standing Assumptions. Theorem 12 thus extends the previous bounds derived in [Gai09; RPM17]. In the latter works, the authors required the distribution rules to be non increasing and sub budget-balanced, i.e., $jf(j) \leq 1$ for all $j \in [n]$.

The next corollary shows how the result in the previous theorem matches the results in [Gai09; RPM17], simply requiring f to be non increasing (this is a less restrictive assumption than what asked for in [Gai09; RPM17]).

Corollary 5. *Consider f a non increasing distribution rule. The value of (11.12) is given by*

$$W^\star = 1 + \max_{j\in[n-1]} \{jf(j) - f(j+1), (n-1)f(n)\}\,. \tag{11.14}$$

In [Gai09, Thm. 2] the author provides a bound matching the expression in (11.14). Tightness of the previous bound is shown in [RPM17, Thm. 1]. Additionally, [Gai09, Eq. 5] also determines the distribution rule maximizing the price of anarchy (11.14). The optimal distribution, denoted with f_{G}, has already been introduced in (10.5) and is reported in the following for completeness

$$f_{\mathrm{G}}(j) = (j-1)!\frac{\frac{1}{(n-1)(n-1)!} + \sum_{i=j}^{n-1} \frac{1}{i!}}{\frac{1}{(n-1)(n-1)!} + \sum_{i=1}^{n-1} \frac{1}{i!}}, \quad j \in [n]\,.$$

In all the above mentioned results the feasible set of distribution rules is limited to $jf(j) \leq 1$ and f non increasing. Using the result provided here in Theorem 12 it is possible to determine the optimal distribution without imposing these additional constraints on f by solving the following LP derived from (11.13)

$$\begin{aligned} &\arg\min_{f\in\mathcal{F},\,\mu\in\mathbb{R}} \mu \\ &\text{s.t. } \mu \geq (j+1)f(j+1) \\ &\qquad \mu \geq 1 + jf(j) - f(j+1) \\ &\qquad \mu \geq 1 + jf(j+1) \qquad\qquad \forall j \in [n-1]. \end{aligned} \tag{11.15}$$

Numerical simulations have shown that the optimal distribution rule obtained optimizing (11.12) precisely matches the one derived in [Gai09], so that removing the additional assumption required therein does not improve the best achievable price of anarchy.[3]

Remark 15 (Matching the $1 - 1/e$ of [NWF78])**.** *Relative to MMC problems, [Gai09] explicitly determines the value of the price of anarchy for the optimal distribution f_{G}. It's value amounts to (see Theorem 7)*

$$\mathrm{PoA}(f_{\mathrm{G}}) = 1 - \frac{1}{\frac{1}{(n-1)(n-1)!} + \sum_{i=0}^{n} \frac{1}{i!}} \xrightarrow{n\to\infty} 1 - \frac{1}{e}\,.$$

This shows that for MMC problems (a generalization of weighted maximum coverage problems) one can obtain the same approximation guarantee achievable for weighted maximum coverage problems and first shown in [NWF78].

11.3 The case of supermodular welfare function

In this section we consider welfare basis functions that are non-decreasing and convex, resulting in a monotone and supermodular total welfare $W(a)$. Applications featuring this property include clustering and image segmentation [SK10], power allocation in multiuser networks [Yas+17]. In the following we explicitly characterize the price of anarchy for the class of supermodular resource allocation problems as a function of f (Theorem 13), extending [JM18; PM17b]. Additionally, we show that the Shapley value distribution rule maximizes this measure of efficiency (recovering the result in [JM18; PM17b]), but *is not the only one.*

Assumption 11. *Throughout this section we assume that $f(1) = w(1) = 1$ and that w is a non-decreasing and convex function, i.e.,*

$$\begin{aligned} w(j+1) &\geq w(j)\,, \\ w(j+1) - w(j) &\geq w(j) - w(j-1) \quad \forall j \in [n-1]\,. \end{aligned}$$

[3] This statement con be proved, by showing that the distribution f_{G} solves the KKT system of (11.15).

Theorem 13 (PoA for supermodular welfare)**.** *Consider a distribution rule f such that $f(j)w(j) \geq 1 \ \forall j \in [n]$. It holds*

$$\mathrm{PoA}(f) = \frac{n}{w(n)} \frac{1}{\max_{j \in [n]} jf(j)} .$$

Additionally, f_{SV} is optimal amongst $f \in \mathcal{F}$ and achieves

$$\mathrm{PoA}(f_{\mathrm{SV}}) = \frac{n}{w(n)} .$$

Observe that the Shapley value and all the distribution rules for which $jf(j) \geq 1$ satisfy the conditions of Theorem 13. Indeed $f(j)w(j) \geq jf(j) \geq 1$ by convexity and Standing Assumptions. Further note that the Shapley value distribution rule is *not* the unique maximizer of $\mathrm{PoA}(f)$. Indeed, all the distribution rules with $1/w(j) \leq f(j) \leq 1/j$ are optimal, as the previous theorem applies and they achieve a price of anarchy of $n/w(n)$ since it is $\max_{j \in [n]} jf(j) = 1$ (due to $f(1) = 1$ and $f(j) \leq 1/j$). Figure 11.2 compares the price of anarchy of the Shapley value, marginal contribution and optimal distribution rule, in the case when $w(j) = j^d$ with $d \in [1, 2]$, $|N| \leq 20$. First, we observe that any optimal distribution rule and f_{SV} give the same performance, as predicted from the previous theorem. Additionally, we observe that the quality of the approximation quickly degrades as the welfare basis w gets steeper (d gets larger). This is due to the fact that if $w(n)$ grows much faster than n, the quantity $n/w(n)$ quickly decreases.

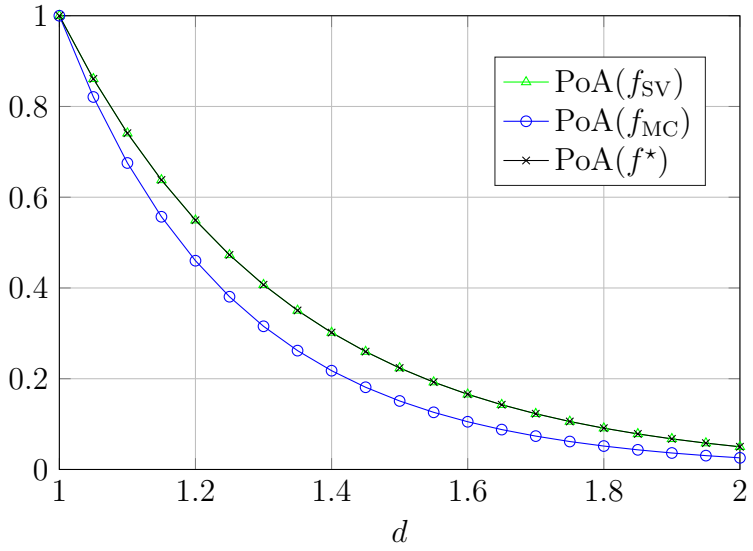

Figure 11.2: Price of anarchy comparison between the optimal distribution rule $f^\star$ determined as the solution of the LP in Theorem 10, Shapley value f_{SV} and marginal contribution f_{MC} distribution rules. The problems considered features $|N| \leq 20$ agents and a welfare basis of the form $w(j) = j^d$ with $d \in [1, 2]$ represented over the x-axis.

11.4 Appendix

11.4.1 Proofs of the results presented in Section 11.1

Proof of Theorem 11

Proof. Observe that the value of $W^\star$ in (11.1) can be equivalently reformulated as in the following program, upon observing that for $j+l \leq n$ it holds $\min(j, n-l) = j$ and $\min(l, n-j) = l$, while for $j+l > n$ it holds $\min(j, n-l) = n-l$ and $\min(l, n-j) = n-j$,

$$
\begin{aligned}
W^\star = \min_{\mu \in \mathbb{R}} \ & \mu \\
\text{s.t. } & \mu w(j) \geq w(l) + j f(j) w(j) - l f(j+1) w(j+1) \\
& \qquad \forall j, l \in [n]_0 \ s.t. \ j \geq l \text{ and } 1 \leq j + l \leq n, \\
& \mu w(j) \geq w(l) + (n-l) f(j) w(j) - (n-j) f(j+1) w(j+1) \\
& \qquad \forall j, l \in [n]_0 \ s.t. \ j \geq l \text{ and } \quad j + l > n \, .
\end{aligned}
$$

In the following we prove that the latter program follows from Corollary 3 by showing that only the constraints with $l \leq j$ are required, and that the decision variable λ in (10.11) takes the value $\lambda^\star = 1$. First, notice that $f(j)w(j)$ is assumed to be non increasing, and so $W^\star$ can be correctly computed using Corollary 3. For $j = 0$, the constraints in (10.11) read as

$$
\lambda \geq \frac{w(l)}{l} \quad \forall \, l \in [n] \, ,
$$

and the most binding amounts to $\lambda \geq 1$, due the to concavity of w. For $j \neq 0$, we intend to show that the constraints with $l > j$ appearing in (10.11) are not required since those with $j = l$ are more binding. The following figure explains this more clearly.

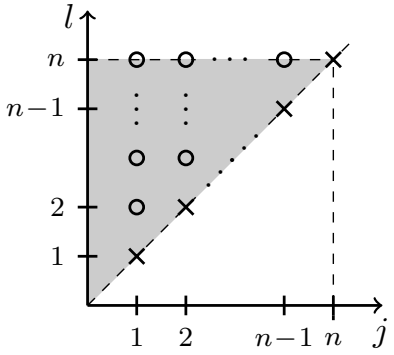

Figure 11.3: The proof amounts to showing that for any constraint identified with the indices (j, l) and $l > j$ (circles), the constraint (j, j) is more binding (crosses).

To do so, we divide the discussion in two cases: $l + j \leq n$ and $l + j > n$.
Case 1. When $1 \leq j + l \leq n$ we want to show that for any $l > j$ and $\lambda \geq 1$

$$1 + \lambda \frac{j}{w(j)}[f(j)w(j) - f(j+1)w(j+1)] \geq$$
$$\frac{w(l)}{w(j)} + \lambda \left[\frac{j}{w(j)} f(j)w(j) - \frac{l}{w(j)} f(j+1)w(j+1) \right],$$

where the left hand side is obtained setting $l = j$. This is equivalent to showing

$$w(l) - w(j) + \lambda(j - l)f(j+1)w(j+1) \leq 0\,. \tag{11.16}$$

By concavity of w and $l > j$, one observes that

$$w(l) \leq w(j+1) + (w(j+1) - w(j))(l - j - 1) = w(j) + (w(j+1) - w(j))(l - j)$$

and since $l - j > 0$, $w(j+1) - w(j) \geq 0$, $\lambda \geq 1$, it holds

$$w(l) \leq w(j) + \lambda(w(j+1) - w(j))(l - j). \tag{11.17}$$

Using inequality (11.17), one can show that (11.16) has to hold

$$\begin{aligned} &w(l) - w(j) + \lambda(j - l)f(j+1)w(j+1) \\ &\leq w(j) + \lambda(w(j+1) - w(j))(l - j) - w(j) + \lambda(j - l)f(j+1)w(j+1) \\ &= \lambda(l - j)(w(j+1) - w(j) - f(j+1)w(j+1)) \leq 0\,, \end{aligned}$$

where the last inequality holds because $f(j+1)w(j+1) \geq w(j+1) - w(j)$ (by assumption) and $l > j$. Observe that the previous inequality is never evaluated for $j = n$, as there is no $l \in [n]$ with $l > j = n$.
Case 2. We now consider the case $j + l > n$. Here we intend to prove that for any $l > j$ and $\lambda \geq 1$

$$1 + \lambda \frac{n - j}{w(j)}[f(j)w(j) - f(j+1)w(j+1)] \geq$$
$$\frac{w(l)}{w(j)} + \lambda \left[\frac{n - l}{w(j)} f(j)w(j) - \frac{n - j}{w(j)} f(j+1)w(j+1) \right],$$

where the left hand side is obtained setting $l = j$. The latter is equivalent to

$$w(l) - w(j) + \lambda(j - l)f(j)w(j) \leq 0.$$

Similarly to (11.17), one can show that

$$w(l) \leq w(j) + \lambda(w(j) - w(j-1))(l - j),$$

and get the desired result as follows

$$\begin{aligned}
&w(l) - w(j) + \lambda(j-l)f(j)w(j) \\
&\leq w(j) + \lambda(w(j) - w(j-1))(l-j) - w(j) + \lambda(j-l)f(j)w(j) \\
&= \lambda(l-j)(w(j) - w(j-1) - f(j)w(j)) \leq 0\,,
\end{aligned}$$

where the last inequality holds because $f(j)w(j) \geq w(j) - w(j-1)$ (by assumption) and $l > j$.

The two cases just discussed showed that $W^\star$ in (10.11) can be computed as

$$\begin{aligned}
W^\star = \min_{\lambda \in \mathbb{R}_{\geq 0}, \mu \in \mathbb{R}} \ & \mu \\
\text{s.t. } & \mu w(j) \geq w(l) + \lambda[jf(j)w(j) - lf(j+1)w(j+1)] \\
& \qquad \forall j, l \in [n]_0 \text{ s.t. } j \geq l \text{ and } 1 \leq j + l \leq n, \\
& \mu w(j) \geq w(l) + \lambda[(n-l)f(j)w(j) - (n-j)f(j+1)w(j+1)] \\
& \qquad \forall j, l \in [n]_0 \text{ s.t. } j \geq l \text{ and } \quad j + l \geq n,
\end{aligned}$$

Every constraint appearing in the previous program is indexed by (j, l) and can be compactly written as $\mu w(j) \geq b_{jl} + a_{jl}\lambda$, upon defining $b_{jl} \coloneqq w(l)$ and consequently

$$a_{jl} \coloneqq \begin{cases} jf(j)w(j) - lf(j+1)w(j+1) & 1 \leq j + l \leq n, \\ (n-l)f(j)w(j) - (n-j)f(j+1)w(j+1) & j + l \geq n. \end{cases}$$

Consequently $W^\star$ can be computed as

$$\begin{aligned}
W^\star = \min_{\lambda \in \mathbb{R}_{\geq 0}, \mu \in \mathbb{R}} \ & \mu \\
\text{s.t. } & \mu w(j) \geq b_{jl} + a_{jl}\lambda \qquad \forall j, l \in [n]_0, \text{ s.t. } j \geq l,\ j + l \geq 1\,.
\end{aligned}$$

As previously seen, for $j = 0$ the most binding constraint is $\lambda \geq 1$. Observe that, when $j \geq 1$ and $j \geq l$, it holds $a_{jl} \geq 0$. Indeed, since $f(j)w(j)$ is non increasing, for $1 \leq j + l \leq n$ one has $a_{jl} = jf(j)w(j) - lf(j+1)w(j+1) \geq (j-l)f(j)w(j) \geq 0$. Similarly for $j + l \geq n$. Thus, the optimal choice is to pick λ as small as possible, i.e., $\lambda^\star = 1$. □

Proof of Corollary 4

Proof. The proof is an application of Theorem 11.

(a) Observe that f_{SV} satisfies the assumptions of Theorem 11 in that $f(j)w(j) = w(j)/j$ is non increasing (due to concavity of w) and $f_{\text{SV}}(j) = 1/j \geq 1 - w(j)/w(j-1)$ $\iff$ $w(j-1) + j(w(j) - w(j-1)) \geq 0$ (due to positivity and non-decreasingness of w). Hence the result of Theorem 11 applies and substituting $f(j) = 1/j$ gives $W^\star_{\text{SV}}$ as in the claim.

(b) Observe that f_{MC} satisfies the assumption of Theorem 11 in that $f(j)w(j) = w(j) - w(j-1)$ is non increasing (due to concavity of w) and $f(j) = 1 - \frac{w(j-1)}{w(j)}$ so the second condition is satisfied too.

We conclude by proving that the constraints indexed with $l < j \in [n]$ are not needed and it is enough to consider $j = l \in [n]$, so that $W^{\star}_{\mathrm{MC}}$ is as given (11.4). To do so, we show that for any constraint with $l < j$ the constraint with $l = j$ is more binding.

For $l < j$ and $j + l \leq n$ we want to prove that

$$1 + \lambda \frac{j}{w(j)}[f(j)w(j) - f(j+1)w(j+1)] \geq$$
$$\frac{w(l)}{w(j)} + \lambda \left[\frac{j}{w(j)} f(j)w(j) - \frac{l}{w(j)} f(j+1)w(j+1) \right],$$

where the left hand side is obtained setting $l = j$. The previous is equivalent to

$$w(l) - w(j) + \lambda(j-l)f(j+1)w(j+1) \leq 0\,,$$

and since $f(j+1)w(j+1) = w(j+1) - w(j)$, it reduces to

$$w(l) - w(j) + (j-l)(w(j+1) - w(j)) \leq 0\,. \tag{11.18}$$

By concavity of w and $l < j$, it holds that $w(j) \geq w(l) + (j-l)(w(j+1) - w(j))$ and thus (11.18) follows.

In the case of $l < j$ and $j + l > n$ we intend to show

$$1 + \lambda \frac{n-j}{w(j)}[f(j)w(j) - f(j+1)w(j+1)] \geq$$
$$\frac{w(l)}{w(j)} + \lambda \left[\frac{n-l}{w(j)} f(j)w(j) - \frac{n-j}{w(j)} f(j+1)w(j+1) \right],$$

which reduces to

$$w(l) - w(j) + (j-l)(w(j) - w(j-1))\,.$$

The latter follows by concavity of w. Hence, the price of anarchy of f_{MC} is governed by $W^{\star}$ as in Theorem 11, where we set $f = f_{\mathrm{MC}}$ and fix $j = l$. This gives the following expression

$$W^{\star}_{\mathrm{MC}} = 1 + \max_{j \in [n]} \left\{ \min(j, n-j) \left[f_{\mathrm{MC}}(j) - f_{\mathrm{MC}}(j+1) \frac{w(j+1)}{w(j)} \right] \right\},$$

which reduced to the expression for $W^{\star}_{\mathrm{MC}}$ in the claim, upon substituting f_{MC} with its definition.

□

11.4.2 Proofs of the results presented in Section 11.2

Proof of Theorem 12

Proof. The proof is a specialization of the general result obtained in Theorem 9 to the case of set covering problems. We divide the study in three distinct cases, as in the following

$$C_1 : \begin{cases} a + x = 0 \\ b + x \neq 0 \end{cases} \qquad C_2 : \begin{cases} a + x \neq 0 \\ b + x = 0 \end{cases} \qquad C_3 : \begin{cases} a + x \neq 0 \\ b + x \neq 0 \end{cases}$$

In case C_1 it must be $a = x = 0$, $b \neq 0$ and the constraints read as

$$\lambda \geq \frac{1}{b}\,.$$

The most binding one is obtained for $b = 1$, i.e., it suffices to have $\lambda \geq 1$ in order to guarantee $\lambda \geq 1/b$. In case C_2 it must be $b = x = 0$, $a \neq 0$. The constraints read as

$$\mu \geq \lambda a f(a) \quad \forall\ a \in [n].$$

In case C_3, since $a + x \neq 0$ and $b + x \neq 0$, the constraints become

$$\mu \geq 1 + \lambda[a f(a + x) - b f(a + x + 1)]\,.$$

If $x = 0$, then a, $b > 0$ and the previous inequality reads

$$\mu \geq 1 + \lambda[a f(a) - b f(a + 1)] \quad a + b \in [n],\,.$$

The most constraining inequality is obtained for b taking the smallest possible value, that is $b = 1$. Thus $0 < a \leq n - 1$. Consequently when $x = 0$, it suffices to have

$$\mu \geq 1 + \lambda[a f(a) - f(a + 1)] \quad \forall a \in [n - 1]\,.$$

If $x \neq 0$, the most binding constraint is obtained for $b = 0$. In such case, $0 < a + x \leq n$ and the constraints read as

$$\mu \geq 1 + \lambda a f(a + x) \quad \forall a \in [n]\,.$$

For ease of readability, we introduce the variable $j \coloneqq a + x$ and use j and x instead of a and x. With this new system of indices the feasible region becomes $0 < j \leq n$ and $j - x \geq 0$, $x > 0$. The latter set of constraints read as

$$\mu \geq 1 + \lambda(j - x) f(j)$$

and the most binding is trivially obtained for $x = 1$, reducing the previous to

$$\mu \geq 1 + \lambda(j - 1) f(j) \quad \forall\ j \in [n]\,.$$

This guarantees that the program in (10.10) is equivalent to

$$\begin{aligned} W^\star = \min_{\lambda\in\mathbb{R}_{\geq 0},\, \mu\in\mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \lambda \geq 1 \\ & \mu \geq \lambda j f(j) \quad j \in [n] \\ & \mu \geq 1 + \lambda(j f(j) - f(j+1)) \quad j \in [n-1] \\ & \mu \geq 1 + \lambda(j-1) f(j) \quad j \in [n]\,. \end{aligned}$$

Amongst the last three sets of constraints, the tightest constraint always features a positive coefficient multiplying λ. Indeed the only term multiplying λ that could take negative values is $jf(j) - f(j+1)$, but every time this is negative, the constraints $\mu \geq 1 + \lambda(j-1)f(j)$ are tighter. It follows that the solution consists in picking λ as small as possible, that is in choosing $\lambda^\star = 1$. The program becomes

$$\begin{aligned} W^\star = \min_{\mu\in\mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \mu \geq j f(j) \quad j \in [n] \\ & \mu \geq 1 + j f(j) - f(j+1) \quad j \in [n-1] \\ & \mu \geq 1 + (j-1) f(j) \quad j \in [n]\,. \end{aligned}$$

We conclude with a little of cosmetics: the first and third set of inequalities run over $j \in [n]$, while the second one has $j \in [n-1]$. Observe that the first and the third condition evaluated at $j = 1$ read both as $\mu \geq 1$. This condition is implied by the last set of condition with $j = 2$, indeed it reads as $\mu \geq 1 + f(2) \geq 1$ since we assumed f non negative. Thus the first and third conditions can be reduced to $j \in \{2, \dots, n\}$. Shifting the indices down by one, we get

$$\begin{aligned} W^\star = \min_{\mu\in\mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \mu \geq (j+1) f(j+1) \quad j \in [n-1] \\ & \mu \geq 1 + j f(j) - f(j+1) \quad j \in [n-1] \\ & \mu \geq 1 + j f(j+1) \quad j \in [n-1]\,, \end{aligned}$$

from which we get the analytic expression in (11.12), i.e.,

$$W^\star = 1 + \max_{j\in[n-1]} \{(j+1)f(j+1) - 1, j f(j) - f(j+1), j f(j+1)\}\,.$$

□

Proof of Corollary 5

Proof. Thanks to Theorem 12, the value $W^\star$ and consequently the price of anarchy can be computed as

$$W^\star = \max_{j\in[n-1]} \{(j+1)f(j+1),\, 1 + j f(j) - f(j+1),\, 1 + j f(j+1)\}\,.$$

We will show that when f is non-increasing, fewer constraints are required, producing exactly (11.14).

First observe that f being non-increasing implies $(j+1)f(j+1) = f(j+1)+jf(j+1) \leq f(1) + jf(j+1) = 1 + jf(j+1)$, so that the first set of conditions is implied by the third. Hence

$$W^\star = 1 + \max_{j\in[n-1]} \{jf(j) - f(j+1),\, jf(j+1)\}\,.$$

We now verify that the first set of remaining conditions implies all the conditions in the second set, but not the last one:

$$\mu \geq 1 + jf(j) - f(j+1) \geq 1 + jf(j) - f(j) = 1 + (j-1)f(j)\,,$$

$\forall j \in [n-1]$ that is, all conditions $\mu \geq jf(j+1)$ are satisfied for $j \in [n-2]$. Thus, it suffices to require $\mu - 1 \geq jf(j) - f(j+1)$ and $\mu - 1 \geq (n-1)f(n)$ for all $j \in [n]$ and the result in (11.14) follows. □

11.4.3 Proofs of the results presented in Section 11.3

Proof of Theorem 13

Proof. The proof is a specialization of the general result obtain in Theorem 9. We divide the study in the same three cases used for the proof of Theorem 12.

In case C_1, the constraints read as

$$w(b) - \lambda b \leq 0 \iff \lambda \geq \frac{w(b)}{b},$$

the most constraining of which is given for $b = n$ as $w(b)$ is convex. Thus it must be

$$\lambda \geq \frac{w(n)}{n}\,.$$

In case C_2, the constraints read as

$$\lambda a f(a) w(a) \leq \mu w(a) \iff \mu \geq \lambda a f(a).$$

In case C_3, the constraints read as

$$\mu \geq \frac{w(b+x)}{w(a+x)} + \lambda\left[af(a+x) - bf(a+x+1)\frac{w(a+x+1)}{w(a+x)}\right].$$

In order to conclude, we will show that the constraints obtained from C_1 and C_2 imply

all the conditions stemming from C_3. To do so observe that

$$\begin{aligned}
&\frac{w(b+x)}{w(a+x)} + \lambda\left[af(a+x) - bf(a+x+1)\frac{w(a+x+1)}{w(a+x)}\right] \\
=&\frac{1}{w(a+x)}\left[w(b+x) - \lambda b w(a+x+1)f(a+x+1)\right] + \lambda a f(a+x) \\
\leq&\frac{1}{w(a+x)}\left[\lambda(b+x) - \lambda b \cdot f(1)w(1)\right] + \lambda a f(a+x) = \frac{1}{w(a+x)}x\lambda + \lambda a f(a+x) \\
\leq&\lambda x f(a+x) + \lambda a f(a+x) = \lambda(a+x)f(a+x)
\end{aligned}$$

From first to second line is rearrangement. From second to third is due to $f(a+x+1)w(a+x+1) \geq w(1)f(1) = 1$ and to $w(b+x) \leq \frac{w(n)}{n}(b+x) \leq \lambda(b+x)$ where the first inequality holds because of convexity of w and the second inequality follows from C_1, i.e., from $\lambda \geq \frac{w(n)}{n}$. From third to fourth is rearrangement. From fourth to fifth is due to $w(a+x)f(a+x) \geq f(1)w(1) = 1 \implies f(a+x) \geq \frac{f(1)w(1)}{w(a+x)}$.

The previous series of inequalities have demonstrated that if $\mu \geq \lambda a f(a)$ as required by condition C_2, and if $\lambda \geq \frac{w(n)}{n}$ as required by condition C_1, then $\mu \geq \lambda(a+x)f(a+x) \geq \frac{w(b+x)}{w(a+x)} + \lambda\left[af(a+x) - bf(a+x+1)\frac{w(a+x+1)}{w(a+x)}\right]$, i.e., conditions C_3 are all satisfied.

It follows that $W^\star$ and consequently the price of anarchy is easily obtained as

$$\begin{aligned}
W^\star = &\min_{\lambda\in\mathbb{R}_{\geq 0},\,\mu\in\mathbb{R}} \ \mu \\
\text{s.t. } &\mu \geq \lambda j f(j) \quad \forall j \in [n] \\
&\lambda \geq \frac{w(n)}{n}\,.
\end{aligned}$$

The solution is given by $\lambda^\star = \frac{w(n)}{n}$, $\mu^\star = \lambda^\star \max_{j\in[n]} jf(j)$, which gives a price of anarchy of

$$\mathrm{PoA}(f) = \frac{n}{w(n)}\frac{1}{\max_{j\in[n]} j \cdot f(j)}\,.$$

Amongst all the distribution rules satisfying $f(j)w(j) \geq 1$, the distribution f_{SV} is optimal. This follows from the fact that $\max_{j\in[n]} j \cdot f_{\mathrm{SV}}(j) = 1$ is the smallest achievable value since $f(1) = 1$. To conclude that f_{SV} is optimal not only over all distributions with $f(j)w(j) \geq 1$ but also over all distributions $f \in \mathcal{F}$ it suffices to observe that [JM18, Lem. 7.2] constructs an instance showing that $n/w(n)$ is the best attainable price of anarchy independently of what f is used. □

CHAPTER 12

Applications

As set forward in the introduction of Chapter 8, our objective was to obtain efficient and distributed algorithms for the solution of GMMC problems. We decided to follow a game theoretic approach and studied the utility design problem in Chapter 10 and Chapter 11. More precisely we have developed a general theory to compute and optimize the price of anarchy as a function of the chosen utility functions. In the following we *do not* tackle the algorithm design component (the second component of the game design approach of Figure 8.1), as there are readily available algorithms capable of determining a Nash equilibrium in a distributed fashion (see Section 9.2 for the best-response algorithm, and its complexity). In this chapter we demonstrate the applicability of our results to the vehicle target allocation problem (Section 12.1), and to the problem of distributed caching in mobile networks (Section 12.2). We provide thorough simulation results and show the theoretical and numerical advantages of our approach. The results presented in this chapter have been published in [PM18b].

12.1 The vehicle target allocation problem

In this section we consider the vehicle target assignment problem introduced in [Mur00] and studied, e.g., in [AMS07; MR14]. We are given a finite set of targets $\mathcal{R}$, and for each target $r \in \mathcal{R}$ its relative importance $v_r \geq 0$. Additionally, we are given a finite set of vehicles $N = \{1, \dots, n\}$, and for each vehicle a set of feasible target assignments $\mathcal{A}_i \in 2^{\mathcal{R}}$. The goal is to distributedly compute a feasible allocation $a \in \mathcal{A}$ so as to maximize the joint probability of successfully destroying the selected targets, expressed as

$$W(a) = \sum_{r \in \cup_{i \in N} a_i} v_r (1 - (1-p)^{|a|_r}),$$

where $(1-(1-p)^{|a|_r})$ is the probability that $|a|_r$ vehicles eliminate the target v_r and the scalar quantity $0 < p \leq 1$ is a parameter representing the probability that a vehicle will successfully destroy a target. In the forthcoming presentation, it is assumed that the success probability p is the same for all vehicles, else one would have to define a different

p_i for every vehicle $i \in N$. Observe that the welfare considered here has the form (8.1) with welfare basis $(1-(1-p)^{|a|_r})$. We normalize this quantity (without affecting the problem's solution) so that $w(1) = 1$ and thus define

$$w(j) = \frac{1-(1-p)^j}{1-(1-p)}. \tag{12.1}$$

Observe that (12.1) satisfies the Standing Assumptions, and Assumption 11 in that $w(j) > 0$ and $w(j)$ is increasing and concave. Thus, it is possible to compute the performance of any set of utility functions of the form (10.1) using Theorem 9, and to further determine the optimal distribution rule $f^\star \in \mathcal{F}$ by solving a corresponding linear program.

Figure 12.1 shows the achievable approximation ratios for the Shapley value, marginal contribution, optimal distribution, as well as the approximation bound in (11.11). We observe that the optimal distribution rule significantly outperforms all the others as well as the bound (11.11) for non trivial values of p. For the extreme case of $p = 1$, $f^\star$ matches (11.11), while for small p all the design methodologies offer a similarly high performance guarantee. Figure 12.2 shows the distribution rules f_{SV}, f_{MC} and $f^\star$ for the choice of $p = 0.5$.

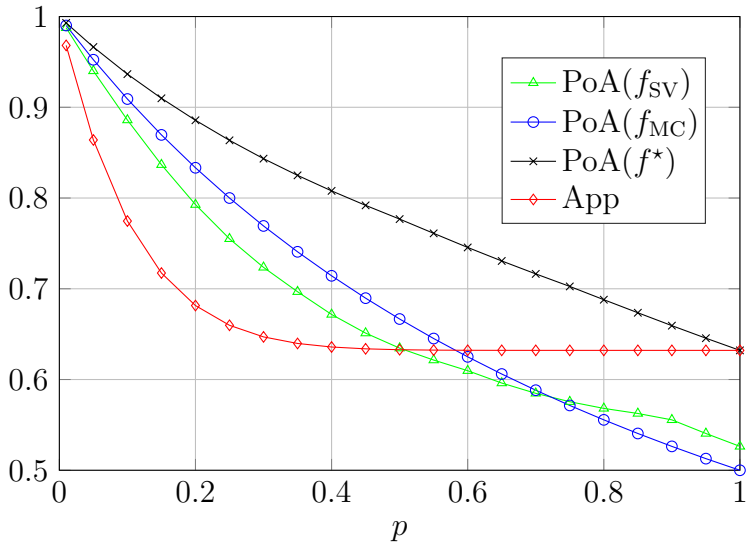

Figure 12.1: Price of anarchy and approximation ratio comparison between the optimal distribution rule $f^\star$, the Shapley value distribution rule f_{SV}, the marginal contribution distribution rule f_{MC}, and (11.11). The problems considered feature $|N| \leq n = 10$ vehicles and $w(j) = \frac{1-(1-p)^j}{1-(1-p)}$ with $0 < p \leq 1$ represented over the x-axis.

In both Figures 12.1 and 12.2 we have set the number of agents to be relatively small[1], i.e., $|N| \leq n = 10$. This choice was purely made so as to perform an exhaustive

[1]Similar trends and conclusions can be obtained with larger values of n.

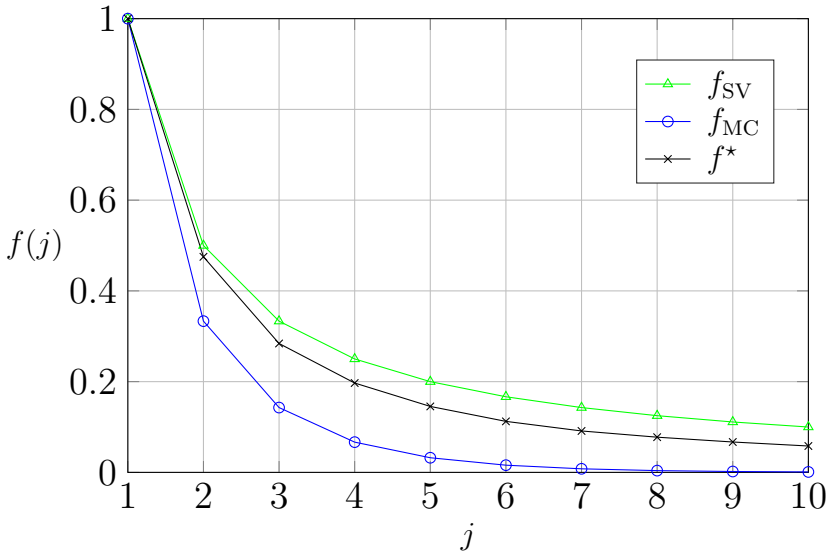

Figure 12.2: Distributions f_{SV}, f_{MC} and optimal distribution $f^\star$ obtained solving the LP in Theorem 10 for the specific choice of $w(j)$ in (12.1) with $|N| \leq n = 10$ and $p = 0.5$.

search simulation in order to test the provided bounds displayed in Figure 12.1. More specifically, we considered 10^5 random instances of the vehicle target assignment problem. Each instance features $n = 10$ agents, $n + 1$ resources and fixed $p = 0.8$. Each agent is equipped with an action set with only two allocations, whose elements are singletons, i.e., $|a_i| = 1$. We believe this is not restrictive in assessing the performance, as the structure of some worst case instances is of this form [RPM17].

Observe that any constraint set $\mathcal{A}_i$ where feasible allocations are singletons is the bases of a uniform matroid of rank one, see Example 3. Further note that computing a single best response is a polynomial operation in the number of resources. Thus, the best response algorithm will converge polynomially to a Nash equilibrium (see Proposition 17) and so the performance guarantees offered by PoA are *easy* to achieve.

The structure of the constraints sets $\mathcal{A}_i$ and the values of the resources are randomly generated, the latter with uniform distribution in the interval $[0, 1]$. For this class of problems considered, the theoretical worst case performance is $\mathrm{PoA}(f_{\mathrm{SV}}) \approx 0.568$, $\mathrm{PoA}(f_{\mathrm{MC}}) \approx 0.556$, $\mathrm{PoA}(f^\star) \approx 0.688$ (see Figure 12.1 with $p = 0.8$). For each instance G generated, we performed an exhaustive search so as to compute the welfare at the worst equilibrium $\min_{a \in \mathrm{ne}(G)} W(a)$ and the value $W(a^{\mathrm{opt}})$. The ratio between these quantities (their empirical cumulative distribution) is plotted across the 10^5 samples in Figure 12.3, for f_{SV}, f_{MC}, $f^\star$. In the same figure the vertical dashed lines represent the theoretical bound on the price of anarchy, while the markers represent the worst case performance occurred during the simulations.

First, we observe that no instance has performed worse than the corresponding price

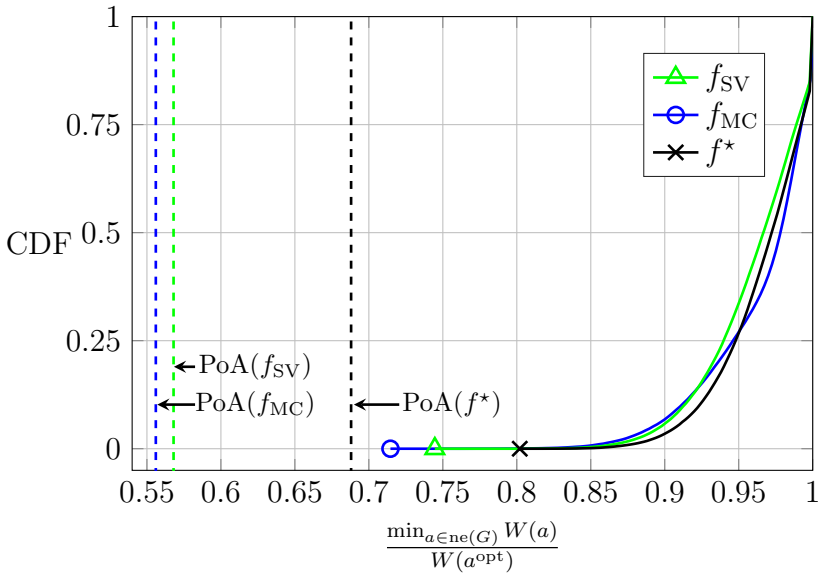

Figure 12.3: Cumulative distribution of the ratio $\min_{a\in \mathrm{ne}(G)} W(a)/W(a^{\mathrm{opt}})$ for f_{SV}, f_{MC}, $f^\star$ across 10^5 samples. The dashed lines represent the theoretical value of PoA(f_{SV}), PoA(f_{MC}), PoA($f^\star$) while the corresponding markers identify the worst case performance encountered during the simulations.

of anarchy, as predicted by Theorem 9. Second, we note that the worst case performance encountered in the simulation is circa 15% better than the true worst case instance.[2] Further, the optimal distribution $f^\star$ has outperformed the others also in the simulations. Its worst case performance is indeed superior to the others (markers in Figure 12.3). Additionally, the cumulative distribution of $f^\star$ lies below the cumulative distributions of f_{SV} and f_{MC} (for abscissas smaller than 0.95). This means that, for any given approximation ratio $r \in [0, 0.95]$, there is a smaller fraction of problems on which $f^\star$ performs worse or equal to r, compared to f_{SV} and f_{MC}. Observe that this is not obvious a priori, as $f^\star$ is designed to maximize the worst case performance and not the average performance.

12.2 Distributed caching

In this section we consider the problem of distributed data caching introduced in [Goe+06] as a technique to reduce peak traffic in mobile data networks. In order to alleviate the growing radio congestion caused by the recent surge of mobile data traffic, the latter work suggested to store popular and spectrum intensive items (such as movies or songs) in geographically distributed stations. The approach has the advantage of bringing the

[2]Recall that our results in Theorems 8 and 9 are provably tight: there exists at least one instance achieving exactly an efficiency equal to the price of anarchy.

content closer to the customer, and to avoid recurring transmission of large quantities of data. Similar offloading techniques, aiming at minimizing the peak traffic demand by storing popular items at local cells, have been recently proposed and studied in the context of modern 5G mobile networks [And13; De +17]. The fundamental question we seek to answer in this section is how to geographically distribute the popular items across the nodes of a network so as to maximize the total number of queries fulfilled. In the following we borrow the model introduced in [Goe+06] and show how the utility design approach presented here yields improved theoretical and practical performances.

We consider a rectangular grid with $n_x \times n_y$ bins and a finite set $\mathcal{R}$ of data items. For each item $r \in \mathcal{R}$, we are given its query rate $q_r \geq 0$ as well as its position in the grid O_r and a radius ρ_r. A circle of radius ρ_r centered in O_r represents the region where the item r is requested. Additionally we consider a set of geographically distributed nodes N (the local cells), where each node $i \in N$ is assigned to a position in the grid P_i. A node is assigned a set of feasible allocations $\mathcal{A}_i$ according to the following rules:

i) $\mathcal{A}_i \subseteq 2^{\mathcal{R}_i}$, where $\mathcal{R}_i \coloneqq \{r \in \mathcal{R} \text{ s.t. } ||O_r - P_i||_2 \leq \rho_r\}$. That is, $r \in \mathcal{R}_i$ if the (euclidean) distance between the position of node i and item r is smaller equal to ρ_r.

ii) $|\mathcal{A}_i| \leq k_i$, for some integer $k_i \geq 1$.

In other words, node i can include the resource r in his allocation a_i only if he is in the region where the item r is requested (first rule), while we limit the number of stored items to k_i for reasons of physical storage (second rule).[3] The situation is exemplified in Figure 12.4.

The objective is to select a feasible allocation for every node so as to jointly maximize the total amount of queries fulfilled

$$\max_{a \in \mathcal{A}} \sum_{r \in \cup_{i \in N} a_i} q_r \,.$$

In order to obtain a distributed algorithm, [Goe+06] proposes a game theoretic approach where each agent is given a Shapley value utility function, i.e., they assign to agents utilities of the form (10.1), where $f(j) = f_{\text{SV}}(j) = 1/j$.

In the following we compare the results of numerical simulations obtained using f_{SV} or the optimal distribution $f^\star = f_{\text{G}}$ defined in (10.5). The following parameters are employed. We choose $n_x = n_y = 800$, $|N| = 100$, $|\mathcal{R}| = 1000$. The nodes and the data items are uniformly randomly placed in the grid. The query rate of data items is chosen according to a power law (Zipf distribution) $q_r = 1/r^\alpha$ for $r \in [1000]$.[4] The

[3] Similarly to what discussed for the application in Section 12.1, it is possible to reduce the problem to the case where $\mathcal{A}_i$ are the bases of a matroid $\mathcal{M}_i$, so that Proposition 17 applies here too. Once more computing the best response is a polynomial task (it amounts to sorting $q_r w(|a|_r) f(|a|_r)$ and picking the k_i first items). Thus the best-response dynamics introduced in Algorithm 5 converges in polynomial time.

[4] Typical query rate curves has been shown to follow this distribution, with $\alpha \in [0.6, 0.9]$, see [Bre+99].

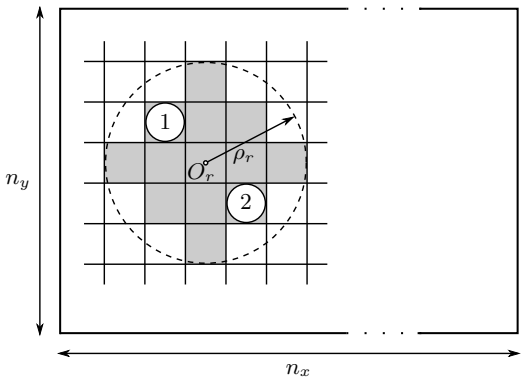

Figure 12.4: The nodes 1 and 2 can include the item r in any allocation, i.e., $r \in \mathcal{R}_1$ and $r \in \mathcal{R}_2$ since the distance from nodes 1 and 2 to O_r is less than ρ_r.

radii of interests are set to be identical for all items $\rho_r = \rho = 200$. We let α vary in $[0.7, 0.9]$. We consider 10^5 instances of such problem, and for every instance compute a Nash equilibrium by means of the best response algorithm. Given the size of the problem, it is not possible to compute the optimal allocation and thus the price of anarchy. As a surrogate for the latter we use the ratio $W(a^{\text{ne}})/W_{\text{tot}}$, where a^{ne} is the Nash equilibrium determined by the algorithm and

$$W_{\text{tot}} \coloneqq \sum_{r \in \mathcal{R}} q_r$$

is the sum of all the query rates and thus is an upper bound for $W(a^{\text{opt}})$. Observe that W_{tot} is a constant for all the simulations with fixed α, indeed $W_{\text{tot}} = \sum_{r \leq 1000} \frac{1}{r^\alpha}$ and thus serves as a mere scaling factor. The theoretical price of anarchy is $\text{PoA}(f_{\text{SV}}) = 0.5$ (tight also when the query rates are Zipf distributed [Goe+06]) and $\text{PoA}(f^\star) = 1 - 1/e \approx 0.632$, see Theorem 7. Figure 12.5 compares the quantity $W(a^{\text{ne}})/W_{\text{tot}}$ for the choice of f_{SV} and $f^\star$, across different values of α. First we observe that the worst cases encountered in the simulations are at least 10% better than the theoretical counterparts. Further, for each fixed value of α, there is a good separation between the performance of f_{SV} and $f^\star$, in favor of the latter. This holds true, not only in the worst case sense (markers in Figure 12.5), but also on average. As α increases from 0.6 to 0.9, the worst case performance seems to degrade for both f_{SV} and $f^\star$. Nevertheless, since we are using $W(a^{\text{ne}})/W_{\text{tot}}$ as a surrogate for the true price of anarchy, it is unclear if the previous conclusion also holds for $W(a^{\text{ne}})/W(a^{\text{opt}})$. Figure 12.6 presents a more detailed comparison between f_{SV} and $f^\star$ for a fixed value of $\alpha = 0.7$ over all the 10^5 instances. Relative to this case, Figure 12.7 describes the (distribution of) number of best response rounds required for the algorithm to converge. Quick convergence is achieved, with a number of best response rounds equal to 11 in the worst case. Observe that in every best response round all players have a chance to update their decision variable, so that a total number of n_{BR} rounds amounts to $n \cdot n_{\text{BR}}$ individual best responses.

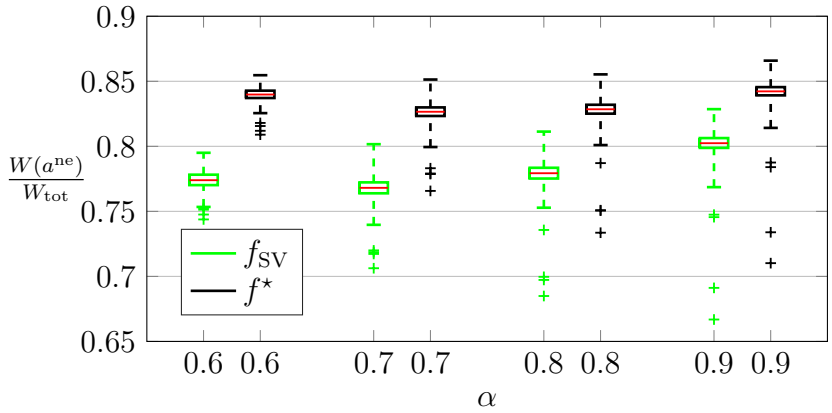

Figure 12.5: Box plot comparing the performance of the best response algorithm on 10^5 instances for the choice of distributions f_{SV} and $f^\star$, across different values of α. On each plot, the median is represented with a red line, and the corresponding box contains the 25th and 75th percentiles. The (four) worst cases are represented with crosses.

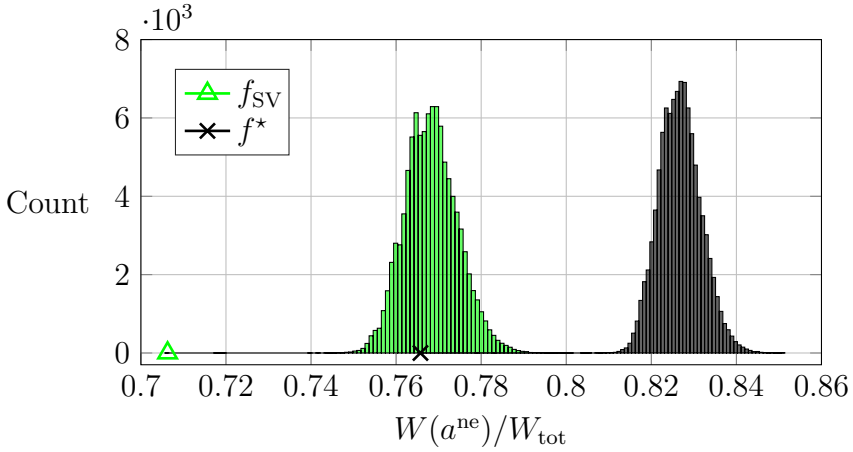

Figure 12.6: Distribution of $W(a^{\text{ne}})/W_{\text{tot}}$ on 10^5 instances for fixed $\alpha = 0.7$.

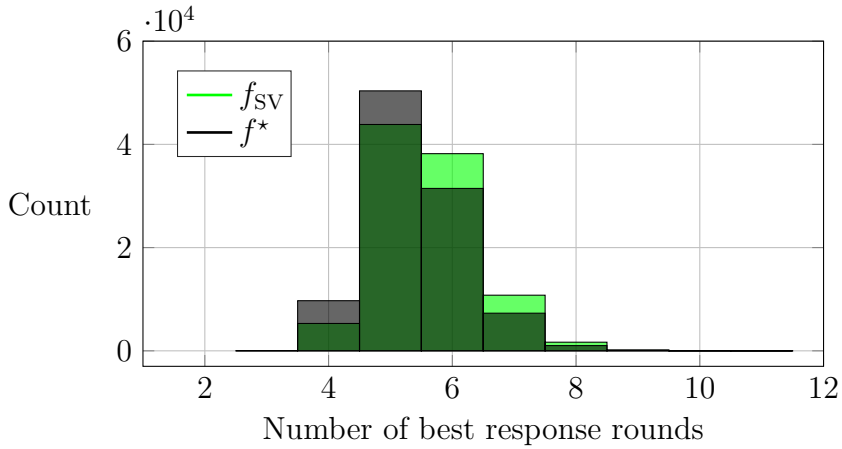

Figure 12.7: Distribution of the number of best response rounds required for convergence on 10^5 instances, $\alpha = 0.7$.

Part III

Conclusion

CHAPTER 13

Conclusions and outlook

13.1 Part I: strategic agents

In the first part of the thesis we considered large scale systems composed of self-interested agents and modeled their strategic interaction using the language of game theory. Motivated by the special structure arising in different real-world applications, we focused on average aggregative games, i.e., games where the cost function of each agent depends solely on his decision and on the average population strategy. The setup considered allows for multidimensional decision variables, heterogenous private constraints, and global constraints coupling the decision variables of the entire population.

Our research agenda was aimed at i) understanding the performance degradation due to selfish decision making, and ii) designing scalable algorithms to guide agents towards an equilibrium configuration. Towards these goals, we first exploited the theory of variational inequalities to reduce both the Nash equilibrium problem and the Wardrop equilibrium to common ground. This allowed to study the efficiency of a Nash equilibrium allocation through the analysis of the corresponding Wardrop equilibrium counterpart. In this respect, we provided conditions on the agents' cost functions that either guarantee the efficiency of the equilibria, or provide meaningful bounds on the efficiency loss. We concluded Part I proposing two decentralized schemes to coordinate the agents towards a Nash or Wardrop equilibrium and discussed under which conditions their convergence is guaranteed. Our findings have been tested on a coordination problem arising in the charging of electric vehicles and on a selfish routing model used in road traffic network.

13.1.1 Further research directions

Non average aggregative games

In Part I of this thesis we focused on *average* aggregative games. While this class of games has recently attracted the attention of the researchers, we believe that many problems within the general framework of (non average) aggregative games are still open. As an

example, the problem of designing distributed algorithms for network aggregative games has been considered only very recently.[1] More broadly, it is unclear to what extent the aggregative structure helps in providing results such as existence and uniqueness of the equilibria under *weaker assumptions* than what usually imposed on non aggregative games. There are few works addressing this question and their results are limited in their scope, for example to scalar valued aggregator functions [Jen10].

Uncertain games and receding-horizon implementations

Within the framework studied in this thesis, we focused on the case of deterministic games. Nevertheless, there has been recent interest both in the areas of optimization and equilibrium theory to incorporate the effect of uncertainty. This desire stems from the observation that a large portion of nowadays decision making happens in face of uncertainty. As a concrete example, consider that of a car driver on a road network. While his goal might entail reaching the desired destination as swiftly as possible, his decisions are based on uncertain knowledge of the congestion he will encounter further ahead on the network. In this respect one can envision at least two future research directions.

First, one could consider stochastic aggregative games where the aggregate function is subject to common uncertainty. The fundamental question one needs to ask is what it means to be an equilibrium configuration. In the simplest scenario, one can think of an equilibrium as a stable configuration of the game constructed with the *expected* costs. Most of the results presented in connection with the variational reformulation of Chapter 4 hold with minor modifications, and one could use algorithms derived from the theory of stochastic variational inequalities to compute one such equilibrium [YNS17; RW17].

As second research direction, one could consider receding-horizon implementations of the schemes proposed here. While some of the applications presented in this thesis were of dynamic nature (e.g., the charging coordination for a fleet of electric vehicles), we have been able to model them as single-stage decision problems. This has been possible due to the exact knowledge of the agents' dynamics. As this is hardly the case in a real world scenario, one might consider receding-horizon implementations of the single-stage problems considered here. This research direction follows the same spirit with which model predictive control is used in uncertain dynamic optimization problems [GPM89].

Non monotone games

Most of the results derived in the first part of the thesis were based on the assumptions of Lipschitzianity and monotonicity of the variational inequality operator (or variations thereof such as strong monotonicity, or co-coercivity see Section 3.1). In this regard, a

[1] In a network aggregative game each agent is represented with a node on a graph, while his cost function is influenced by his decision and by a linear combination of the decision variables of his neighbours.

long term research goal is that of weakening the monotonicity assumption. While this direction would have great impact (there are many situations in which the monotonicity property is not satisfied), there seem to be a fundamental roadblock that needs to be resolved or circumvented before embarking on this route. Indeed, as we have seen in Chapter 3, game theory is a generalization of single agent decision making and hence contains the field of optimization as a special case. Thus, the study of non monotone variational inequalities requires a better understanding of non convex optimization first. While there has been a recent surge of interest in other classes of continuous functions that produce tractable optimization problems (e.g., continuous submodular functions), we feel that this direction is currently underdeveloped.

13.2 Part II: programmable machines

In the second part of the thesis we studied a class of combinatorial resource allocation problems arising in various applications connected to multiagent systems and machine learning. More precisely, we considered a setup where a large number of cooperative agents need to select a subset of resources from a common set, with the objective of jointly maximizing a given welfare function. In the considered setup, the welfare function was assumed to be additive over the resources and to be anonymous with respect to the agent identities. An example of problem satisfying these requirements is the well-known and studied weighted maximum coverage.

Since the class of problems investigated is computationally intractable ($\mathcal{NP}$-hard), our goal was to derive distributed algorithms that run in polynomial time and achieve near-optimal performances. We approached the problem from a game-theoretic perspective and aimed at assigning a local utility function to each agent so that their selfish maximization recovers a large portion of the desired system level objective. Towards this goal, we presented a novel framework for the characterization of the equilibrium efficiency (price of anarchy). More precisely, for a given set of utilities, we showed that the problem of computing the worst-case equilibrium efficiency can be posed as a tractable linear problem. This result might be of independent interest to the community concerned with the study of the price of anarchy. We then leveraged the linear programming reformulation to resolve the question previously posed, i.e., to design local utilities that maximize such performance metric. The importance of this results stems from the observation that any algorithm capable of computing a Nash equilibrium would naturally inherit an approximation ratio matching the corresponding equilibrium efficiency. Surprisingly, the optimal price of anarchy (the price of anarchy achieved by optimally designed utility functions) matches or outperforms the guarantees available for many commonly used algorithms. We validate our results with two applications: the vehicle-target assignment problem and a coverage problem arising in distributed caching for mobile networks.

13.2.1 Further research directions

Different equilibrium notion

As discussed in the introduction of Chapter 8, the game design approach for the approximate solution of an optimization problem amounts to the design of three elements: equilibrium concept, agents' utilities and corresponding learning algorithm. While all the efficiency results presented in this thesis are limited to the notion of pure Nash equilibrium, one might be interested in using a different equilibrium concept. As a matter of fact, the choice of pure Nash equilibria originated from the fact that their efficiency is the highest possible. Unfortunately, pure Nash equilibria are intractable to compute in general (see Figure 9.1 for the tradeoff between complexity and efficiency). The way we resolved this issue was by assuming that $\{\mathcal{A}_i\}_{i=1}^N$ are the sets of bases for a matroid, so that the best-response algorithm converges in a polynomial number of steps (Proposition 17). Instead, coarse correlated equilibria are tractable to compute in general. Thus, an interesting research direction is to understand whether the efficiency bounds obtained for pure Nash Equilibria extend to coarse correlated equilibria. Nevertheless, the performance guarantees offered by coarse correlated equilibria are in expected value, and one would have to understand how to derandomize the corresponding solution efficiently (if at all possible).

Non-anonymous agents

The results derived in this thesis are relative to welfare functions of the form (8.1)

$$W(a) = \sum_{r \in \cup a_i} v_r w(|a|_r) \,.$$

We observe that the key ingredient that allowed to reduce the computation of the price of anarchy to a tractable linear program is the *indistinguishability* of the agents (also called anonymity in the following), see the proof of Theorem 8. Formally, the agents are anonymous if any allocation $a = (a_1, \ldots, a_n)$ and any other allocation obtained as a permutation of the former have the same welfare. While it is very much unclear if and how to extend the current results to the case of non-anonymous agents, we remark that this will greatly expand the number of applications that could benefit from this approach.

The tradeoff between anarchy and stability

Throughout Part II of this thesis, we assessed the quality of an algorithm with its worst case performance over a set of instances. This is a common approach to study the performance of an algorithm as it gives a bound that requires no information on the distribution of inputs and holds instance by instance. Nevertheless, an interesting and

underdeveloped question is whether optimizing the worst-case performance comes at the cost of other performance metrics. In relation to the problem studied in this thesis, a different and more optimistic metric to quantify the equilibrium efficiency is known as price of stability. With the same notation previously used, the price of stability can be defined as

$$\mathrm{PoS}(f) \coloneqq \inf_{G \in \mathcal{G}_f} \left(\frac{\max_{a \in \mathrm{ne}(G)} W(a)}{\max_{a \in \mathcal{A}} W(a)} \right).$$

Informally, the price of stability bounds the performance of the best equilibrium over all the possible instances in the set $\mathcal{G}_f$. While preliminary results have shown that there is a fundamental tradeoff between the price of anarchy and the price of stability in specific classes of problems [RPM17; FGL18], this research direction warrants further exploration as it would provide an additional guiding principle in the design of efficient algorithms.

Bibliography

[AJ13] D. Acemoglu and M. K. Jensen. "Aggregate comparative statics". In: *Games and Economic Behavior* 81 (2013), pp. 27–49.

[ARV08] H. Ackermann, H. Röglin, and B. Vöcking. "On the impact of combinatorial structure on congestion games". In: *Journal of the ACM (JACM)* 55.6 (2008), p. 25.

[Aga+12] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang. "Optimization for dynamic ride-sharing: A review". In: *European Journal of Operational Research* 223.2 (2012), pp. 295–303.

[ABS02] E. Altman, T. Basar, and R. Srikant. "Nash equilibria for combined flow control and routing in networks: Asymptotic behavior for a large number of users". In: *IEEE Transactions on automatic control* 47.6 (2002), pp. 917–930.

[Alt+06] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter. "A survey on networking games in telecommunications". In: *Computers & Operations Research* 33.2 (2006), pp. 286–311.

[Alt+11] E. Altman, R. Combes, Z. Altman, and S. Sorin. "Routing games in the many players regime". In: *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). 2011, pp. 525–527.

[AW04] E. Altman and L. Wynter. "Equilibrium, games, and pricing in transportation and telecommunication networks". In: *Networks and Spatial Economics* 4.1 (2004), pp. 7–21.

[And13] J. G. Andrews. "Seven ways that HetNets are a cellular paradigm shift". In: *IEEE Communications Magazine* 51.3 (2013), pp. 136–144.

[AS94] R. Arnott and K. Small. "The economics of traffic congestion". In: *American scientist* 82.5 (1994), pp. 446–455.

[AD54] K. J. Arrow and G. Debreu. "Existence of an equilibrium for a competitive economy". In: *Econometrica: Journal of the Econometric Society* (1954), pp. 265–290.

[AMS07] G. Arslan, J. R. Marden, and J. S. Shamma. "Autonomous vehicle-target assignment: A game-theoretical formulation". In: *Journal of Dynamic Systems, Measurement, and Control* 129.5 (2007), pp. 584–596.

[Bar+15] R. Barbosa, A. Ene, H. Nguyen, and J. Ward. "The power of randomization: Distributed submodular maximization on massive datasets". In: *International Conference on Machine Learning*. 2015, pp. 1236–1244.

[BE13] L. Barenboim and M. Elkin. "Distributed graph coloring: Fundamentals and recent developments". In: *Synthesis Lectures on Distributed Computing Theory* 4.1 (2013), pp. 1–171.

[BC+11] H. H. Bauschke, P. L. Combettes, et al. *Convex analysis and monotone operator theory in Hilbert spaces*. Vol. 408. Springer, 2011.

[BP13] D. Bauso and R. Pesenti. "Mean field linear quadratic games with set up costs". In: *Dynamic Games and Applications* 3.1 (2013), pp. 89–104.

[Ber07] V. Berinde. *Iterative Approximation of Fixed Points*. Springer, 2007.

[BT89] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and distributed computation: numerical methods*. Vol. 23. Prentice hall Englewood Cliffs, NJ, 1989.

[Bil12] V. Bilo. "A unifying tool for bounding the quality of non-cooperative solutions in weighted congestion games". In: *International Workshop on Approximation and Online Algorithms*. Springer. 2012, pp. 215–228.

[Blu93] L. E. Blume. "The statistical mechanics of strategic interaction". In: *Games and economic behavior* 5.3 (1993), pp. 387–424.

[Bom83] I. M. Bomze. "Lotka-Volterra equation and replicator dynamics: a two-dimensional classification". In: *Biological cybernetics* 48.3 (1983), pp. 201–211.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[Bre+99] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker. "Web caching and Zipf-like distributions: Evidence and implications". In: *INFOCOM'99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*. Vol. 1. IEEE. 1999, pp. 126–134.

[Bri02] T. Brinkhoff. "A Framework for Generating Network-Based Moving Objects". In: *GeoInformatica* 6.2 (2002), pp. 153–180.

[Bur+17] G. Burger, D. Paccagnan, B. Gentile, and J. Lygeros. "Guarantees of convergence to a dynamic user equilibrium for a single arc network". In: Elsevier, July 2017, pp. 9674–9679.

[Cal+11] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák. "Maximizing a monotone submodular function subject to a matroid constraint". In: *SIAM Journal on Computing* 40.6 (2011), pp. 1740–1766.

[CL05] C. G. Cassandras and W. Li. "Sensor networks and cooperative control". In: *European Journal of Control* 11.4-5 (2005), pp. 436–463.

[Cha+11] A. C. Chapman, A. Rogers, N. R. Jennings, and D. S. Leslie. "A unifying framework for iterative approximate best-response algorithms for distributed constraint optimization problems 1". In: *The Knowledge Engineering Review* 26.4 (2011), pp. 411–444.

[CK04] C. Chekuri and A. Kumar. "Maximum coverage problem with group budget constraints and applications". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Springer, 2004, pp. 72–83.

[Che+14] H. Chen, Y. Li, R. H. Louie, and B. Vucetic. "Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach". In: *IEEE Transactions on Smart Grid* 5.4 (2014), pp. 1744–1754.

[CD06] X. Chen and X. Deng. "Settling the complexity of two-player Nash equilibrium". In: *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on*. IEEE. 2006, pp. 261–272.

[CKS08] G. Christodoulou, A. Kovács, and M. Schapira. "Bayesian combinatorial auctions". In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2008, pp. 820–832.

[CDT12] J. Cohen, C. Dürr, and N. K. Thang. "Smooth inequalities and equilibrium inefficiency in scheduling games". In: *International Workshop on Internet and Network Economics*. Springer. 2012, pp. 350–363.

[CK08] R. Cohen and L. Katzir. "The generalized maximum coverage problem". In: *Information Processing Letters* 108.1 (2008), pp. 15–22.

[CC84] M. Conforti and G. Cornuéjols. "Submodular set functions, matroids and the greedy algorithm: tight worst-case bounds and some generalizations of the Rado-Edmonds theorem". In: *Discrete applied mathematics* 7.3 (1984), pp. 251–274.

[CS02] V. Conitzer and T. Sandholm. "Complexity results about Nash equilibria". In: *arXiv preprint cs/0205074* (2002).

[Cor94] L. C. Corchon. "Comparative statics for aggregative games the strong concavity case". In: *Mathematical Social Sciences* 28.3 (1994), pp. 151–165.

[CH12] R. Cornes and R. Hartley. “Fully aggregative games”. In: *Economics Letters* 116.3 (2012), pp. 631–633.

[CSS04] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. “Selfish routing in capacitated networks”. In: *Mathematics of Operations Research* 29.4 (2004), pp. 961–976.

[Cor+02] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. “Coverage control for mobile sensing networks”. In: *Robotics and Automation, 2002. Proceedings. ICRA ’02. IEEE International Conference on*. Vol. 2. 2002, pp. 1327–1332.

[Cou38] A.-A. Cournot. *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*. chez L. Hachette, 1838.

[CT14] R. Cressman and Y. Tao. “The replicator equation and other game dynamics”. In: *Proceedings of the National Academy of Sciences* 111.Supplement 3 (2014), pp. 10810–10817.

[DD03] R. D’Andrea and G. E. Dullerud. “Distributed control design for spatially interconnected systems”. In: *IEEE Transactions on Automatic Control* 48.9 (Sept. 2003), pp. 1478–1495.

[Daf80] S. Dafermos. “Traffic equilibrium and variational inequalities”. In: *Transportation science* 14.1 (1980), pp. 42–54.

[DN87] S. Dafermos and A. Nagurney. “Oligopolistic and competitive behavior of spatially separated markets”. In: *Regional Science and Urban Economics* 17.2 (1987), pp. 245–254.

[DGP06] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. “The Complexity of Computing a Nash Equilibrium”. In: *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*. STOC ’06. Seattle, WA, USA: ACM, 2006, pp. 71–78.

[DP05] A. De Palma and N. Picard. “Route choice decision under travel time uncertainty”. In: *Transportation Research Part A: Policy and Practice* 39.4 (2005), pp. 295–324.

[DAS17] A. De Paola, D. Angeli, and G. Strbac. “Convergence and optimality of a new iterative price-based scheme for distributed coordination of flexible loads in the electricity market”. In: *Proceedings of the IEEE Conference on Decision and Control*. 2017.

[De +17] F. De Pellegrini, A. Massaro, L. Goratti, and R. El-Azouzi. “Competitive caching of contents in 5G edge cloud networks”. In: *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017 15th International Symposium on*. IEEE. 2017, pp. 1–8.

[DG08] J. Dean and S. Ghemawat. “MapReduce: simplified data processing on large clusters”. In: *Communications of the ACM* 51.1 (2008), pp. 107–113.

[DMP17] L. Deori, K. Margellos, and M. Prandini. “On the connection between Nash equilibria and social optima in electric vehicle charging control games”. In: *IFAC-PapersOnLine* 50.1 (2017), pp. 14320–14325.

[DJ12] N. R. Devanur and K. Jain. “Online matching with concave returns”. In: *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. ACM. 2012, pp. 137–144.

[DHZ06] P. Dubey, O. Haimanko, and A. Zapechelnyuk. “Strategic complements and substitutes, and potential games”. In: *Games and Economic Behavior* 54.1 (2006), pp. 77–94.

[Ern+04] A. T. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier. “Staff scheduling and rostering: A review of applications, methods and models”. In: *European journal of operational research* 153.1 (2004), pp. 3–27.

[FPT04] A. Fabrikant, C. Papadimitriou, and K. Talwar. “The complexity of pure Nash equilibria”. In: *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*. ACM. 2004, pp. 604–612.

[FFP07] F. Facchinei, A. Fischer, and V. Piccialli. “On generalized Nash games and variational inequalities”. In: *Operations Research Letters* 35.2 (2007), pp. 159–164.

[FK07] F. Facchinei and C. Kanzow. “Generalized Nash equilibrium problems”. In: *4OR* 5.3 (2007), pp. 173–210.

[FP07] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[FH13] P. von Falkenhausen and T. Harks. “Optimal cost sharing for resource selection games”. In: *Mathematics of Operations Research* 38.1 (2013), pp. 184–208.

[Far+12] R. Z. Farahani, N. Asgari, N. Heidari, M. Hosseininia, and M. Goh. “Covering problems in facility location: A review”. In: *Computers & Industrial Engineering* 62.1 (2012), pp. 368–407.

[Fei98] U. Feige. “A threshold of ln n for approximating set cover”. In: *Journal of the ACM (JACM)* 45.4 (1998), pp. 634–652.

[FGL18] A. Filos-Ratsikas, Y. Giannakopoulos, and P. Lazos. “The Pareto Frontier of Inefficiency in Mechanism Design”. In: *arXiv preprint arXiv:1809.03454* (2018).

[FL98] D. Fudenberg and D. K. Levine. *The theory of learning in games*. Vol. 2. MIT press, 1998.

[Gai09] M. Gairing. “Covering games: Approximation through non-cooperation”. In: *International Workshop on Internet and Network Economics*. Springer. 2009, pp. 184–195.

[GTL13] L. Gan, U. Topcu, and S. H. Low. “Optimal decentralized protocol for electric vehicle charging”. In: *IEEE Transactions on Power Systems* 28.2 (2013), pp. 940–951.

[GPM89] C. E. Garcia, D. M. Prett, and M. Morari. “Model predictive control: theory and practice - a survey”. In: *Automatica* 25.3 (1989), pp. 335–348.

[Geb+16] P. Gebraad, F. Teeuwisse, J. Wingerden, P. A. Fleming, S. Ruben, J. Marden, and L. Pao. “Wind plant power optimization through yaw control using a parametric model for wake effects-a CFD simulation study”. In: *Wind Energy* 19.1 (2016), pp. 95–114.

[Gen18] B. Gentile. “Equilibria in Aggregative Games”. PhD thesis. ETH Zurich, 2018.

[Gen+17] B. Gentile, D. Paccagnan, B. Ogunsola, and J. Lygeros. “A novel concept of equilibrium over a network”. In: *2017 IEEE 56th Annual Conference on Decision and Control*. Dec. 2017, pp. 3829–3834.

[Gen+18] B. Gentile, F. Parise, D. Paccagnan, M. Kamgarpour, and J. Lygeros. “A game theoretic approach to decentralized charging of plug-in electric vehicles”. In: *Challenges in Engineering and Management of Cyber-Physical Systems of Systems*. River Publishers, 2018.

[Goe+06] M. X. Goemans, L. Li, V. S. Mirrokni, and M. Thottan. “Market sharing games applied to content distribution in ad hoc networks”. In: *IEEE Journal on Selected areas in Communications* 24.5 (2006), pp. 1020–1033.

[GGL15] M. Gonzales, S. Grammatico, and J. Lygeros. “On the price of being selfish in large populations of plug-in electric vehicles”. In: *Proceedings of the IEEE Conference on Decision and Control*. 2015, pp. 6542–6547.

[GGS05] G. Gottlob, G. Greco, and F. Scarcello. “Pure Nash equilibria: Hard and easy games”. In: *Journal of Artificial Intelligence Research* 24 (2005), pp. 357–406.

[Gra17] S. Grammatico. “Dynamic control of agents playing aggregative games with coupling constraints”. In: *IEEE Transactions on Automatic Control* 62.9 (2017), pp. 4537–4548.

[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 (2016), pp. 3315–3329.

[HS66] P. Hartman and G. Stampacchia. “On some non-linear elliptic differential-functional equations”. In: *Acta mathematica* 115.1 (1966), pp. 271–310.

[HM85] A. Haurie and P. Marcotte. “On the relationship between Nash-Cournot and Wardrop equilibria”. In: *Networks* 15.3 (1985), pp. 295–308.

[Hoc97] D. S. Hochbaum. “Approximation Algorithms for NP-hard Problems”. In: ed. by D. S. Hochbaum. Boston, MA, USA: PWS Publishing Co., 1997. Chap. Approximating Covering and Packing Problems: Set Cover, Vertex Cover, Independent Set, and Related Problems, pp. 94–143.

[HCM07] M. Huang, P. E. Caines, and R. P. Malhamé. “Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria”. In: *IEEE Transactions on Automatic Control* 52.9 (2007), pp. 1560–1571.

[ISA02] K. Ishiyama, M. Sendoh, and K. Arai. “Magnetic micromachines for medical applications”. In: *Journal of Magnetism and Magnetic Materials* 242 (2002), pp. 41–46.

[Jen10] M. K. Jensen. “Aggregative games and best-reply potentials”. In: *Economic Theory* 43.1 (2010), pp. 45–66.

[JM18] E. Jenson and J. R. Marden. “Optimal Utility Design in Convex Distributed Welfare Games”. In: *American Control Conference (ACC), 2018*. IEEE. 2018.

[Jör+14] M. J. Jörgensen, D. Paccagnan, N. K. Poulsen, and M. B. Larsen. “IMU calibration and validation in a factory, remote on land and at sea”. In: *2014 IEEE/ION Position, Location and Navigation Symposium*. 2014, pp. 1384–1391.

[KZ12] T. R. Kaplan and S. Zamir. “Asymmetric first-price auctions with uniform distributions: analytic solutions to the general case”. In: *Economic Theory* 50.2 (2012), pp. 269–302.

[Kar+17] M. Karimi, M. Lucic, H. Hassani, and A. Krause. “Stochastic submodular maximization: The case of coverage functions”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 6853–6863.

[Kit+99] H. Kitano, S. Tadokoro, I. Noda, H. Matsubara, T. Takahashi, A. Shinjou, and S. Shimada. “Robocup rescue: Search and rescue in large-scale disasters as a domain for autonomous agents research”. In: *Systems, Man, and Cybernetics, 1999. IEEE SMC’99 Conference Proceedings. 1999 IEEE International Conference on*. Vol. 6. IEEE. 1999, pp. 739–743.

[KNS16] J. Koshal, A. Nedić, and U. V. Shanbhag. “Distributed algorithms for aggregative games on graphs”. In: *Operations Research* 64.3 (2016), pp. 680–704.

[KP99] E. Koutsoupias and C. Papadimitriou. “Worst-case Equilibria”. In: *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science*. STACS’99. Trier, Germany: Springer-Verlag, 1999, pp. 404–413.

[KSG08] A. Krause, A. Singh, and C. Guestrin. “Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies”. In: *Journal of Machine Learning Research* 9.Feb (2008), pp. 235–284.

[KCM11] T. K. Kristoffersen, K. Capion, and P. Meibom. “Optimal charging of electric drive vehicles in a market environment”. In: *Applied Energy* 88.5 (2011), pp. 1940–1948.

[KM14] J. Kulkarni and V. Mirrokni. “Robust price of anarchy bounds via LP and fenchel duality”. In: *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*. SIAM. 2014, pp. 1030–1049.

[Kun+12] H.-B. Kuntze, C. W. Frey, I. Tchouchenkov, B. Staehle, E. Rome, K. Pfeiffer, A. Wenzel, and J. Wöllenstein. “SENEKA-sensor network with mobile robots for disaster management”. In: *Homeland Security (HST), 2012 IEEE Conference on Technologies for*. IEEE. 2012, pp. 406–410.

[LL07] J.-M. Lasry and P.-L. Lions. “Mean field games”. In: *Japanese Journal of Mathematics* 2.1 (2007), pp. 229–260.

[LLN06] B. Lehmann, D. Lehmann, and N. Nisan. “Combinatorial auctions with decreasing marginal utilities”. In: *Games and Economic Behavior* 55.2 (2006), pp. 270–296.

[LZ11] J. Li and H. M. Zhang. “Fundamental Diagram of Traffic Flow : New identification Scheme and Further Evidence from Empirical Data”. In: *Journal of the Transportation Research Board* (2011).

[LM13] N. Li and J. R. Marden. “Designing games for distributed optimization”. In: *IEEE Journal of Selected Topics in Signal Processing* 7.2 (2013), pp. 230–242.

[LW94] N. Littlestone and M. K. Warmuth. “The weighted majority algorithm”. In: *Information and computation* 108.2 (1994), pp. 212–261.

[MCH13] Z. Ma, D. S. Callaway, and I. Hiskens. “Decentralized charging control of large populations of plug-in electric vehicles”. In: *Control Systems Technology, IEEE Transactions on* 21.1 (2013), pp. 67–78.

[MS07] S. Mannor and J. S. Shamma. “Multi-agent Learning for Engineers”. In: *Artif. Intell.* 171.7 (May 2007), pp. 417–422.

[MW95] P. Marcotte and J. H. Wu. “On the convergence of projection methods: application to the decomposition of affine variational inequalities”. In: *Journal of Optimization Theory and Applications* 85.2 (1995), pp. 347–362.

[MAS09] J. R. Marden, G. Arslan, and J. S. Shamma. "Cooperative Control and Potential Games". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39.6 (Dec. 2009), pp. 1393–1407.

[MP17] J. R. Marden and M. Philips. "Optimizing the price of anarchy in concave cost sharing games". In: *2017 American Control Conference (ACC)*. May 2017, pp. 5237–5242.

[MS18] J. R. Marden and J. S. Shamma. "Game-Theoretic Learning in Distributed Control". In: *Handbook of Dynamic Game Theory*. Ed. by T. Basar and G. Zaccour. Springer International Publishing, 2018, pp. 1–36.

[ME12] J. R. Marden and M. Effros. "The price of selfishness in network coding". In: *IEEE Transactions on Information Theory* 58.4 (2012), pp. 2349–2361.

[MR14] J. R. Marden and T. Roughgarden. "Generalized efficiency bounds in distributed resource allocation". In: *IEEE Transactions on Automatic Control* 59.3 (2014), pp. 571–584.

[MS12] J. R. Marden and J. S. Shamma. "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation". In: *Games and Economic Behavior* 75.2 (2012), pp. 788–808.

[MW08] J. R. Marden and A. Wierman. "Distributed welfare games with applications to sensor coverage". In: *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE. 2008, pp. 1708–1713.

[MW13] J. R. Marden and A. Wierman. "Distributed welfare games". In: *Operations Research* 61.1 (2013), pp. 155–168.

[Mir+16] B. Mirzasoleiman, A. Karbasi, R. Sarkar, and A. Krause. "Distributed submodular maximization". In: *Journal of Machine Learning Research* 17.238 (2016), pp. 1–44.

[MHK18] A. Mokhtari, H. Hassani, and A. Karbasi. "Decentralized Submodular Maximization: Bridging Discrete and Continuous Settings". In: *arXiv preprint arXiv:1802.03825* (2018).

[MS96] D. Monderer and L. S. Shapley. "Potential games". In: *Games and economic behavior* 14.1 (1996), pp. 124–143.

[Mot+16] M. Motalleb, M. Thornton, E. Reihani, and R. Ghorbani. "Providing frequency regulation reserve services using demand response scheduling". In: *Energy Conversion and Management* 124 (2016), pp. 439–452.

[MS01] H. Moulin and S. Shenker. "Strategyproof sharing of submodular costs: budget balance versus efficiency". In: *Economic Theory* 18.3 (2001), pp. 511–533.

[MV78] H. Moulin and J.-P. Vial. "Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon". In: *International Journal of Game Theory* 7.3-4 (1978), pp. 201–221.

[Mur00] R. A. Murphey. "Target-based weapon target assignment problems". In: *Nonlinear Assignment Problems*. Springer, 2000, pp. 39–53.

[MYR17] A. Murthy, C. Yeshwanth, and S. Rao. "Distributed Approximation Algorithms for the Multiple Knapsack Problem". In: *arXiv preprint arXiv:1702.00787* (2017).

[NR10] U. Nadav and T. Roughgarden. "The limits of smoothness: A primal-dual framework for price of anarchy bounds". In: *International Workshop on Internet and Network Economics*. Springer. 2010, pp. 319–326.

[Nag13] A. Nagurney. *Network economics: A variational inequality approach*. Vol. 10. Springer Science & Business Media, 2013.

[Nas50] J. F. Nash. "Equilibrium points in n-person games". In: *Proceedings of the National Academy of Sciences* 36.1 (1950), pp. 48–49.

[NO09] A. Nedic and A. Ozdaglar. "Distributed subgradient methods for multi-agent optimization". In: *IEEE Transactions on Automatic Control* 54.1 (2009), pp. 48–61.

[NWF78] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. "An analysis of approximations for maximizing submodular set functions-I". In: *Mathematical Programming* 14.1 (1978), pp. 265–294.

[NB+10] F. Nemry, M. Brons, et al. *Plug-in hybrid and battery electric vehicles. Market penetration scenarios of electric drive vehicles*. Tech. rep. Institute for Prospective and Technological Studies, 2010.

[Neu28] J. v. Neumann. "Zur theorie der gesellschaftsspiele". In: *Mathematische annalen* 100.1 (1928), pp. 295–320.

[Nis+07] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic game theory*. Vol. 1. Cambridge University Press Cambridge, 2007.

[PCM18] D. Paccagnan, R. Chandan, and J. R. Marden. "Distributed resource allocation through utility design – Part I: optimizing the performance certificates via the price of anarchy". In: *arXiv preprint arXiv: 1807.01333* (2018). *Submitted for journal publication*.

[Pac+16] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros. "Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints". In: *2016 IEEE 55th Conference on Decision and Control*. Dec. 2016, pp. 6123–6128.

[Pac+18] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros. "Nash and Wardrop Equilibria in Aggregative Games with Coupling Constraints". In: *IEEE Transactions on Automatic Control* (2018). *Early access*.

[PKL15] D. Paccagnan, M. Kamgarpour, and J. Lygeros. "On the range of feasible power trajectories for a population of thermostatically controlled loads". In: *2015 54th IEEE Conference on Decision and Control*. Dec. 2015, pp. 5883–5888.

[PKL16] D. Paccagnan, M. Kamgarpour, and J. Lygeros. "On aggregative and mean field games with applications to electricity markets". In: *2016 European Control Conference (ECC)*. June 2016, pp. 196–201.

[PM18a] D. Paccagnan and J. Marden. "The Importance of System-Level Information in Multiagent Systems Design: Cardinality and Covering Problems". In: *IEEE Transactions on Automatic Control* (2018). *Early access*.

[PM17a] D. Paccagnan and J. R. Marden. "The risks and rewards of conditioning noncooperative designs to additional information". In: *2017 55th Annual Allerton Conference on Communication, Control, and Computing*. Oct. 2017, pp. 958–965.

[PM18b] D. Paccagnan and J. R. Marden. "Distributed resource allocation through utility design – Part II: applications to submodular, supermodular and set covering problems". In: *arXiv preprint arXiv: 1807.01343* (2018). *Submitted for journal publication*.

[PPL18] D. Paccagnan, F. Parise, and J. Lygeros. "On the Efficiency of Nash Equilibria in Aggregative Charging Games". In: *IEEE Control Systems Letters* 2.4 (Oct. 2018), pp. 629–634.

[PP09] Y. Pan and L. Pavel. "Games with coupled propagated constraints in optical networks with multi-link topologies". In: *Automatica* 45.4 (2009), pp. 871–880.

[Pan+10] J. Pang, G. Scutari, D. P. Palomar, and F. Facchinei. "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach". In: *IEEE Transactions on Signal Processing* 58.6 (2010), pp. 3251–3271.

[Pav07] L. Pavel. "An extension of duality to a game-theoretic framework". In: *Automatica* 43.2 (2007), pp. 226–237.

[PSM16] M. Philips, Y. Shalaby, and J. R. Marden. "The importance of budget in efficient utility design". In: *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE. 2016, pp. 6117–6122.

[PM17b] M. Phillips and J. R. Marden. "Design Trade-offs in Concave Cost Sharing Games". In: *IEEE Transactions on Automatic Control* (2017).

[Pis95] D. Pisinger. “A minimal algorithm for the multiple-choice knapsack problem”. In: *European Journal of Operational Research* 83.2 (1995). EURO Summer Institute Combinatorial Optimization, pp. 394–410.

[QBL15] G. Qu, D. Brown, and N. Li. “Distributed Greedy Algorithm for Satellite Assignment Problem with Submodular Utility Function”. In: *IFAC-PapersOnLine* 48.22 (2015), pp. 258–263.

[RPM17] V. Ramaswamy, D. Paccagnan, and J. R. Marden. “The Impact of Local Information on the Performance of Multiagent Systems”. In: *arXiv preprint arXiv:1710.01409* (2017). *Submitted for journal publication.*

[Ram27] F. P. Ramsey. “A Contribution to the Theory of Taxation”. In: *The Economic Journal* 37.145 (1927), pp. 47–61.

[RW17] R. T. Rockafellar and R. J. Wets. “Stochastic variational inequalities: single-stage to multistage”. In: *Mathematical Programming* 165.1 (2017), pp. 331–360.

[Ros65] J. B. Rosen. “Existence and uniqueness of equilibrium points for concave N-person games”. In: *Econometrica: Journal of the Econometric Society* (1965), pp. 520–534.

[Ros73] R. W. Rosenthal. “A class of games possessing pure-strategy Nash equilibria”. In: *International Journal of Game Theory* 2.1 (1973), pp. 65–67.

[Rou03] T. Roughgarden. “The price of anarchy is independent of the network topology”. In: *Journal of Computer and System Sciences* 67.2 (2003), pp. 341–364.

[Rou09] T. Roughgarden. “Intrinsic robustness of the price of anarchy”. In: *Proceedings of the forty-first annual ACM symposium on Theory of computing.* ACM. 2009, pp. 513–522.

[Rou16] T. Roughgarden. *Twenty lectures on algorithmic game theory.* Cambridge University Press, 2016.

[RS15] T. Roughgarden and F. Schoppmann. “Local smoothness and the price of anarchy in splittable congestion games”. In: *Journal of Economic Theory* 156 (2015), pp. 317–342.

[RST17] T. Roughgarden, V. Syrgkanis, and E. Tardos. “The price of anarchy in auctions”. In: *Journal of Artificial Intelligence Research* 59 (2017), pp. 59–101.

[SG09] B. Saha and L. Getoor. “On maximum coverage in the streaming model & application to multi-topic blog-watch”. In: *Proceedings of the 2009 siam international conference on data mining.* SIAM. 2009, pp. 697–708.

[SA10] W. H. Sandholm and C. Ansell. *Population games and evolutionary dynamics*. MIT press, 2010.

[San75] A. Sandmo. "Optimal taxation in the presence of externalities". In: *The Swedish Journal of Economics* (1975), pp. 86–98.

[SC17] H. R. Sayarshad and J. Y. Chow. "Non-myopic relocation of idle mobility-on-demand vehicles as a dynamic location-allocation-queueing problem". In: *Transportation Research Part E: Logistics and Transportation Review* 106 (2017), pp. 60–77.

[Sch+96] S. Schaible et al. "Generalized monotone nonsmooth maps". In: *Journal of Convex Analysis* 3 (1996), pp. 195–206.

[SM03] A. S. Schulz and N. Moses. "On the performance of user equilibria in traffic networks". In: *Proc. 14th Symp. Discrete Algorithms*. 2003, p. 86.

[Scu+12] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang. "Monotone games for cognitive radio systems". In: *Distributed Decision Making and Control*. Springer, 2012, pp. 83–112.

[Ser+15] A. Servant, F. Qiu, M. Mazza, K. Kostarelos, and B. J. Nelson. "Controlled in vivo swimming of a swarm of bacteria-like microrobotic flagella". In: *Advanced Materials* 27.19 (2015), pp. 2981–2988.

[Sha07] J. S. Shamma. *Cooperative control of distributed multi-agent systems*. Wiley Online Library, 2007.

[SWL11] Y. Song, S. H. Wong, and K.-W. Lee. "Optimal gateway selection in multi-domain wireless networks: A potential game perspective". In: *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM. 2011, pp. 325–336.

[SSR09] C. Soto, B. Song, and A. K. Roy-Chowdhury. "Distributed multi-target tracking in a self-configuring camera network". In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE. 2009, pp. 1486–1493.

[Spi+14] K. Spieser, K. Treleaven, R. Zhang, E. Frazzoli, D. Morton, and M. Pavone. "Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore". In: *Road vehicle automation*. Springer, 2014, pp. 229–245.

[SK10] P. Stobbe and A. Krause. "Efficient minimization of decomposable submodular functions". In: *Advances in Neural Information Processing Systems*. 2010, pp. 2208–2216.

[SCL16] T. H. Summers, F. L. Cortesi, and J. Lygeros. "On Submodularity and Controllability in Complex Dynamical Networks." In: *IEEE Trans. Control of Network Systems* 3.1 (2016), pp. 91–101.

[SVW17] M. Sviridenko, J. Vondrák, and J. Ward. "Optimal approximation for submodular and supermodular optimization with bounded curvature". In: *Mathematics of Operations Research* 42.4 (2017), pp. 1197–1218.

[ST13] V. Syrgkanis and E. Tardos. "Composable and efficient mechanisms". In: *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. ACM. 2013, pp. 211–220.

[Tha17] N. K. Thang. "Game Efficiency through Linear Programming Duality". In: *arXiv preprint arXiv:1708.06499* (2017).

[Vet02] A. Vetta. "Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions". In: *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*. IEEE. 2002, pp. 416–425.

[VM07] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton university press, 2007.

[Wac13] G. Wachsmuth. "On LICQ and the uniqueness of Lagrange multipliers". In: *Operations Research Letters* 41.1 (2013), pp. 78–80.

[War12] J. Ward. "Oblivious and non-oblivious local search for combinatorial optimization". PhD thesis. University of Toronto (Canada), 2012.

[War52] J. G. Wardrop. "Some theoretical aspects of road traffic research". In: *Inst Civil Engineers Proc London*. 1952.

[WOJ13] E. Wei, A. Ozdaglar, and A. Jadbabaie. "A distributed Newton method for network utility maximization–I: Algorithm". In: *IEEE Transactions on Automatic Control* 58.9 (2013), pp. 2162–2175.

[Wel10] D. J. Welsh. *Matroid theory*. Courier Corporation, 2010.

[Yan+18] G.-Z. Yang, J. Bellingham, P. E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, R. Merrifield, et al. "The grand challenges of Science Robotics". In: *Science Robotics* 3.14 (2018), eaar7650.

[Yas+17] M. Yassin, S. Lahoud, K. Khawam, M. Ibrahim, D. Mezher, and B. Cousin. "Centralized versus decentralized multi-cell resource and power allocation for multiuser OFDMA networks". In: *Computer Communications* 107 (2017), pp. 112–124.

[YP17] P. Yi and L. Pavel. "A distributed primal-dual algorithm for computation of generalized Nash equilibria via operator splitting methods". In: *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE. 2017, pp. 3841–3846.

[YSM11] H. Yin, U. V. Shanbhag, and P. G. Mehta. "Nash equilibrium problems with scaled congestion costs and shared constraints". In: *IEEE Transactions on Automatic Control* 56.7 (2011), pp. 1702–1708.

[YNS17] F. Yousefian, A. Nedić, and U. V. Shanbhag. "On smoothing, regularization, and averaging in stochastic approximation methods for stochastic variational inequality problems". In: *Mathematical Programming* 165.1 (2017), pp. 391–431.